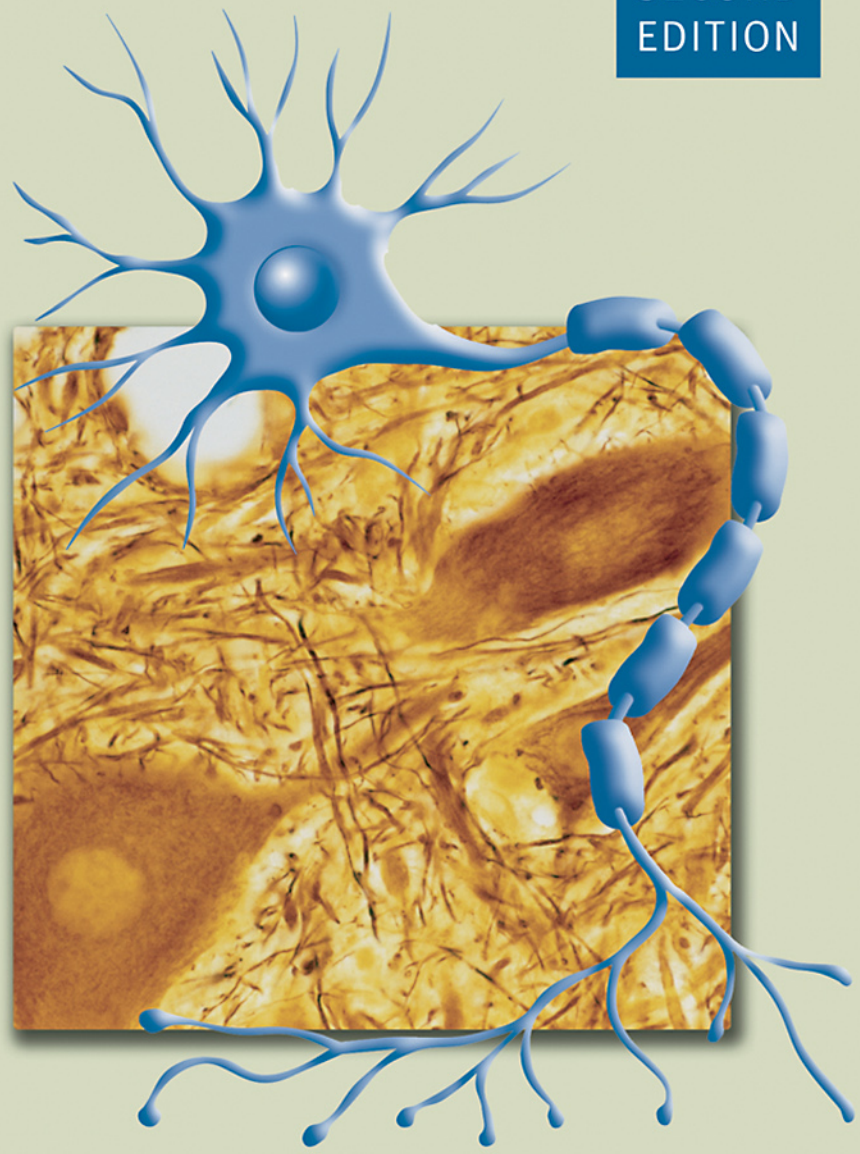
Instant Notes

Neuroscience

SECOND EDITION



A. Longstaff

BIOS Instant Notes

Neuroscience

Second Edition

BIOS INSTANT NOTES

Series Editor: B.D. Hames, School of Biochemistry and Molecular Biology, University of Leeds, Leeds, UK

Biology
Animal Biology, Second Edition
Biochemistry, Third Edition
Bioinformatics
Chemistry for Biologists, Second Edition
Developmental Biology
Ecology, Second Edition
Genetics, Second Edition
Immunology, Second Edition
Mathematics & Statistics for Life Scientists
Medical Microbiology
Microbiology, Second Edition
Molecular Biology, Third Edition
Neuroscience, Second Edition
Plant Biology, Second Edition
Sport & Exercise Biomechanics
Sport & Exercise Physiology

Chemistry
Consulting Editor: Howard Stanbury
Analytical Chemistry
Inorganic Chemistry, Second Edition
Medicinal Chemistry
Organic Chemistry, Second Edition
Physical Chemistry

Psychology
Sub-series Editor: Hugh Wagner, Dept of Psychology, University of Central Lancashire, Preston, UK
Cognitive Psychology
Physiological Psychology
Psychology
Sport & Exercise Psychology

BIOS Instant Notes

Neuroscience

Second Edition

Alan Longstaff

Science writer and neuroscience lecturer

Taylor & Francis
Taylor & Francis Group

Published by:
Taylor & Francis Group
In US: 270 Madison Avenue
New York, NY 10016
In UK: 4 Park Square, Milton Park
Abingdon, OX14 4RN

First published 2000; Second edition published 2005

This edition published in the Taylor & Francis e-Library, 2007.

"To purchase your own copy of this or any of Taylor & Francis or Routledge's collection of thousands of eBooks please go to www.eBookstore.tandf.co.uk."

ISBN 0-203-96985-5 Master e-book ISBN

ISBN: 0-4153-5188-X (Print Edition)

A catalog record for this book is available from the British Library.

Library of Congress Cataloging-in-Publication Data

Longstaff, Alan.
Neuroscience / A. Longstaff. –– 2nd ed.
p. cm. –– (Instant notes)
Includes bibliographical references and index.
1. Neurosciences –– Outlines, syllabi, etc. I. Title. II. Series: Instant notes series
RC343.6.L66 2005
616. 8 –– dc22 2005005818

Editor: Elizabeth Owen
Editorial Assistant: Chris Dixon
Production Editor: Karin Henderson

T&F informa

Taylor & Francis Group
is the Academic Division of T&F Informa plc.

Visit our web site at http://www.garlandscience.com

Contents

Abbreviations

ACh	acetylcholine
AChE	acetylcholinesterase
ACTH	adrenocorticotrophic hormone
AD	Alzheimer's disease
AII	angiotensin II
AGRP	agouti-related peptide
AMPA	α-amino-3-hydroxy-5-methyl-4-isoxazole proprionic acid
ANS	autonomic nervous system
AP	action potential
apoE	apolipoprotein E
APP	amyloid precursor protein
APV	D-2-amino-5-phosphonovalerate
ATN	anterior thalamic nuclei
ATP	adenosine 5′-triphosphate
AVP	arginine vasopressin
βA	β-amyloid
βAR	β adrenoceptors
BAT	brown adipose tissue
BDNF	brain derived neurotrophic factor
bFGF	basic fibroblast growth factor
bl	basal lamina
BMP	bone morphogenetic protein
BOLD	blood oxygen level detection
α-BTX	α-bungarotoxin
CA	cornu ammonis
CaM	calmodulin
CAM	cell adhesion molecule
CaMKII	calcium–calmodulin-dependent protein kinase II
cAMP	cyclic adenosine monophosphate
CART	cocaine- and amphetamine-related transcript
CAT	computer assisted tomography
cbf	cerebral blood flow
CC	cingulate cortex
CCK	cholecystokinin
CF	characteristic frequency
cGMP	3′,5′-cyclic guanosine monophosphate
ChAT	choline acetylesterase
CL	central laminar nucleus (of thalamus)
CNG	cyclic-nucleotide-gated channel
CNS	central nervous system
CNTF	ciliary neurotrophic factor
CoA	coenzyme A
CP	cortical plate
CPG	central pattern generators
CR	conditioned response
CRE	cAMP response element
CREB	cAMP response element binding protein
CREM	cAMP response element modulator
CRH	corticotrophin releasing hormone
CRO	cathode ray oscilloscope
CS	conditioned stimulus
CSF	cerebrospinal fluid
CVA	cerebrovascular accident
CVLM	caudal ventrolateral medulla
CVO	circumventricular organ
DAG	diacylglycerol
DAT	dopamine transporter
DB	diagonal band (of Broca)
DβH	dopamine-β-hydroxylase
DBL	dorsal blastopore lip
DCML	dorsal column–medial lemniscal system
DCN	dorsal column nuclei
2-DG	2-deoxyglucose
DHC	dorsal horn cell
DI	diabetes insipidus
DLPN	dorsolateral pontine nucleus
DOPAC	dihydroxyphenyl acetic acid
DOPEG	3,4-dihydroxy phenylglycol
DRG	dorsal root ganglion
DYN	dynorphin
ECT	electroconvulsive therapy
EEG	electroencephalography
EGF	epidermal growth factor
EGL	external granular layer
EMG	electromyography
ENK	enkephalin
ENS	enteric nervous system
epp	endplate potential
epsp	excitatory postsynaptic potential
ER	endoplasmic reticulum
ERP	event-related potential
F-actin	filamentous actin
FEF	frontal eye field
FF	fast fatiguing
FGF-8	fibroblast growth factor-8
FM	frequency modulation

fMRI	functional magnetic resonance imaging
FR	fatigue resistant
FRA	flexor reflex afferents
FSH	follicle stimulating hormone
G_i	inhibitory G protein
G_q	G protein coupled to phospholipase
G_s	stimulatory G protein
GABA	γ-aminobutyrate
GAD	glutamic acid decarboxylase
GAT	GABA transporter
GC	guanylyl cyclase
GDNF	glial-derived neurotrophic factor
GDP	guanosine 5′-diphosphate
GFAP	glial fibrillary acidic protein
GH	growth hormone
GHRH	growth hormone releasing hormone
GnRH	gonadotrophin releasing hormone
GPe	globus pallidus pars externa
GPi	globus pallidus pars interna
GR	glucocorticoid receptor
GTO	Golgi tendon organs
GTP	guanosine 5′-triphosphate
5-HIAA	5-hydroxyindoleacetic acid
HPA	hypothalamic–pituitary–adrenal (axis)
HPG	hypothalamic–pituitary–gonadal (axis)
HPT	hypothalamic–pituitary–thyroid (axis)
HRP	horseradish peroxidase
5-HT	5-hydroxytryptamine (serotonin)
5-HTP	5-hydroxytryptophan
HVA	high voltage activated
IaIN	Ia inhibitory interneurons
IbIN	Ib inhibitory neurons
IC	inferior colliculus
ICSS	intracranial self-stimulation
Ig	immunoglobulin
IGF-1	insulin-like growth factor 1
IGL	internal granular layer
iGluR	ionotropic glutamate receptor
ILD	interaural level differences
IP_3	inositol 1,4,5-trisphosphate
ipsp	inhibitory postsynaptic potential
IT	inferotemporal cortex
JGA	juxtoglomerular apparatus
L-DOPA	L-3,4-dihydroxyphenylalanine
LC	locus ceruleus
LCN	lateral cervical nucleus
LDCV	large dense-core vesicle
LGN	lateral geniculate nucleus
LH	luteinizing hormone
LSO	lateral superior olivary nucleus
LTD	long-term depression
LTM	long-term memory
LTN	lateral tegmental nucleus
LTP	long-term potentiation
LVA	low voltage activated
M	magnocellular pathway
M/T	mitral/tufted cells
mAChR	muscarinic cholinergic receptor
MAO	monoamine oxidase
MAP	mean arterial (blood) pressure
MB	mammillary bodies
mepp	miniature endplate potential
MFB	medial forebrain bundle
MFS	mossy fiber sprouting
mGluR1	type 1 metabotropic glutamate receptor
MGN	medial geniculate nucleus
MI	primary motor cortex
MII	secondary motor cortex
MLCK	myosin light chain kinase
MLR	mesencephalic locomotor region
MOPEG	3-methoxy,4-hydroxy phenylglycol
MPOA	medial preoptic area
MPP^+	1-methyl-4-phenyl pyridinium
mpsp	miniature postsynaptic potential
MPTP	1-methyl-4-phenyl-1,2,3,6-tetrahydropyridin
MR	mineralocorticoid receptor
MRI	magnetic resonance imaging
MSO	medial superior olivary complex
MST	medial superior temporal cortex
NA	noradrenaline
nAc	nucleus accumbens
nAChR	nicotinic cholinergic receptor
NBM	nucleus basalis of Meynert
NGF	nerve growth factor
NMDA	*N*-methyl-D-aspartate
NMDAR	*N*-methyl-D-aspartate receptor
nmj	neuromuscular junction
NMR	nuclear magnetic resonance
NPY	neuropeptide Y
NREM	nonrapid eye movement sleep
NRM	nucleus raphe magnus
NST	nucleus of the solitary tract
NT3-6	neurotrophins 3-6

OC	olivocochlear
OCD	obsessive–compulsive disorder
OHC	outer hair cells
6-OHDA	6-hydroxydopamine
ORN	olfactory receptor neurons
OVLT	vascular organ of the lamina terminalis
P	parvocellular pathway
PAD	primary afferent depolarization
PAG	periaqueductal gray matter
Pc	Purkinje cells
PD	Parkinson's disease
PDE	phosphodiesterase
PDGF	platelet-derived growth factor
PDS	paroxysmal depolarizing shifts
PET	positron emission tomography
pf	parallel fibers
PFC	prefrontal cortex
PGO	pontine–geniculate–occipital spikes
PHF	paired helical filaments
PIP_2	phosphatidyl inositol-4,5-bisphosphate
PKA	protein kinase A
PM	premotor cortex
PNS	peripheral nervous system
POA	preoptic area
POM	posterior complex (medial nucleus) of thalamus
POMC	pro-opiomelanocortin
PP	posterior parietal cortex
PRL	prolactin
PSNS	parasympathetic nervous system
psp	postsynaptic potential
PVN	paraventricular nucleus
RA	retinoic acid
RAIC	rostral agranular insular cortex
REM	rapid eye movement sleep
RER	rough endoplasmic reticulum
RF	receptive field
RHT	retinohypothalamic tract
RN	red nucleus
RXR	retinoic acid X receptor
S	slow twitch fiber
SCG	superior cervical ganglion
SCN	suprachiasmatic nucleus
Sc	Schaffer collateral
SDN–POA	sexually dimorphic nucleus of the preoptic area
SER	smooth endoplasmic reticulum
SH2	src homology domain 2
SHH	sonic hedgehog protein
SMA	supplementary motor area
SNpc	substantia nigra pars compacta
SNpr	substantia nigra pars reticulata
SNS	sympathetic nervous system
SON	supraoptic nucleus
SP	substance P
SPL	sound pressure level
SR	sarcoplasmic reticulum
SSRI	selective serotonin reuptake inhibitors
SSV	small clear synaptic vesicle
STM	short-term memory
STN	subthalamic nucleus
STT	spinothalamic tract
SVZ	subventricular zone
TB	trapezoid body
TCA	tricyclic antidepressants
TEA	tetraethylammonium
TENS	transcutaneous electrical nerve stimulation
TGF-β	tumor growth factor-β
TH	tyrosine hydroxylase
TM	transmembrane
TRH	thyrotropin releasing hormone
trk	tyrosine kinase receptors
TSH	thyroid releasing hormone
TTX	tetrodotoxin
UR	unconditioned response
US	unconditioned stimulus
VDCC	voltage-dependent calcium channel
VDKC	voltage-dependent potassium channel
VDSC	voltage-dependent sodium channel
VIP	vasoactive intestinal peptide
VLH	ventrolateral hypothalamus
VLPO	ventrolateral preoptic area
VMAT	vesicular monoamine transporter
VMH	ventromedial hypothalamus
VOR	vestibulo-ocular reflexes
VPL	ventroposterolateral nucleus (of thalamus)
VPM	ventroposteromedial nucleus (of thalamus)
VRG	ventral respiratory group
VST	ventral spinocerebellar tract
VZ	ventricular zone
WDR	wide dynamic range

Preface

Neuroscience is one of the most rapidly advancing areas of science and in consequence spawns a literature which is growing dramatically. Moreover, it is multidisciplinary, having contributions from biochemistry and molecular biology, anatomy, physiology, pharmacology, psychology and clinical medicine, to name the most obvious. For these reasons textbooks of neuroscience tend to be large and complex and it can be hard for students to discriminate the conceptual heart of the subject from exemplars and enrichment material. This second edition of *Instant Notes in Neuroscience* is intended as a supplement to lectures which gives rapid and easy access to the core of the subject in an affordable and manageable-sized text.

When coming to a new subject, it is my experience that students commonly express two concerns; firstly how to sort out the important ideas and facts from the wealth of detail and secondly how to get to grips with the unfamiliar terminology. In addition, the best understanding of neuroscience comes by integrating knowledge across the subject. The *Instant Notes* format addresses each of these issues. Each topic is supported by a 'Key Notes' panel, which gives a concise summary of the crucial points. Whenever a term appears for the first time it is in **bold** and immediately followed by a definition or explanation. Extensive cross references are provided between topics so that students can forge the links that are important for integration.

Instant Notes in Neuroscience is a much slimmer volume than most neuroscience texts. A number of features contribute to this. Firstly, I have tried to minimize the amount of detail without compromising the need for students to have a database for subsequent autonomous learning. Secondly, while many of the methods used by neuroscientists are included, *individual* experiments or items of evidence are included only where I thought it essential to illustrate a point or on matters that would need some justification to be convincing. Thirdly, I have restricted examples almost exclusively to mammals, including – particularly for anatomy – humans, even though there is a great deal of remarkable work, for example, on invertebrates. I have always qualified the species, since species differences matter. If not, then rats and cats would behave as humans do, which clearly they do not!

Instant Notes in Neuroscience has 16 sections containing 96 topics. Section A sets the scene by introducing the cells of the nervous system, and looks at how the nervous system is organized by taking a broad view of neuroanatomy and brain-imaging techniques. The next three sections are essentially cellular neuroscience. Section B is concerned mostly with action potentials, section C examines the general properties of synapses, while section D gives an account of the principal neurotransmitters and their receptors. These sections provide an introduction to the electrophysiological techniques used to study nerve cells, and say something about the molecular biology of the ion channels and receptors that govern their behavior. How information is encoded by the firing and connectivity of neurons is considered in section E. All the material thus far might reasonably be found in first year courses. The next seven sections (F–L) form the core of systems neuroscience. Section F reviews the body senses, touch, pain and balance. Sections G and H deal with vision and hearing respectively, while section I looks at the chemical senses, smell and taste. The properties of motor units and the role of the spinal cord and brainstem in movement are the subject of section J, while proprioception and the functions of the cerebral cortex, cerebellum, and the basal ganglia in voluntary movement are covered in section K. Section L tackles neuroendocrinology, how the brain controls metabolism, growth and reproduction by harnessing the power of hormones, and both peripheral and central aspects of the autonomic nervous system. The perspective of the last four sections is rather broader than what has come before. The neuroscience which helps us to understand behavior, such as emotion, motivation and sleep is introduced in section M. Section N is an overview of how the embryonic nervous system develops, ranging from how the basic plan is genetically specified, though the myriad ways in which that plan unfolds. Section O addresses how the nervous system continues to rewire itself on the basis of experience (learning and

memory) and introduces cognitive neuroscience with an account of attention and language. Finally, although quite a number of nervous system disorders are considered at appropriate places throughout the book, section P takes the six most common brain disorders (schizophrenia, depression, stroke, epilepsy, Parkinson's disease and Alzheimer's dementia) and looks in some detail at what has gone amiss and what current and future treatments may do. At the end is a reading list for those who wish to take their studies further.

As a student, how should you use this book? Restrict your reading only to the sections and topics covered by your current course. That said, sections A–E are likely to appear in, or be required knowledge for, just about any neuroscience program; you will probably need to work through these first. Later sections can be dipped into in any order. Read the main sections thoroughly first, making sure that you *understand* the ideas, and use the 'Related topics' to make links. This is the stage to incorporate additional material from lectures, and other textbooks, in the gaps at the end of topics. For areas that particularly interest you, turn to the Further Reading at the end of the book. Studying *Instant Notes* 'little but often' is a good strategy. The information density in the text is high, so many short, concentrated, bursts is much more effective than a few eight-hour stints. The more times you work through a topic, the better your understanding will be, and the more likely you will remember it clearly. When it comes to revision, use the 'Key Notes' as a prompt. In addition, you should aim to be able to write, from memory, a few sentences about each of the terms that appear in bold in the main text. Being able to reproduce the simpler diagrams is also an effective way of getting your point across in an exam.

At one level neuroscience attempts to provide a mechanistic account of the most complex 'device' in the known universe, the human brain. It is an extraordinary endeavor because it aims to reveal what, in essence, it is to be human; how we behave, think and feel as we do. At the moment we are a long way from being able to give a coherent account of any of these faculties; that there is so much still to be done is one reason that this science is so exciting. This book is an account of the remarkable progress made so far, I hope you find that it serves your needs well and, like me, you enjoy discovering neuroscience.

Acknowledgements

A number of people provided thought-provoking comments that helped shape the first edition: Barry Hunt, Vasanta Raman and John Wilkinson, all at the University of Hertfordshire, kindly read through some individual topics; David Hames (University of Leeds, UK), Kevin Alloway (Penn State University, USA), and Patricia Revest (Queen Mary College, London University, UK), each read through the entire text. The second edition was read in entirety by Peter Shortland (Queen Mary College, London University, UK) and Robert J Gayton (University of Liverpool, UK) whose extremely helpful remarks have been a significant contribution. I am very grateful to all of these people for their time and expertise. Finally, for the production of the second edition I thank Sarah Carlson and Elizabeth Owen of BIOS Scientific Publishers (Taylor & Francis Group) for their hard work and encouragement.

A1 Neuron structure

Key Notes

Cell body

The neuron cell body contains all the subcellular organelles found in a typical animal cell but it is specialized to maintain high rates of protein synthesis.

Neurites

Neurites are long projections from the cell body. There are two types of neurite: dendrites and axons. Dendrites are large extensions of the cell body and receive most of the synaptic inputs impinging onto the cell. Neurons may have one or many dendrites. Neurons have a single axon arising from the axon hillock. Axons form the presynaptic components of synapses.

Axon or dendrite?

The two neurites can be distinguished on structural grounds. Dendrites contain many organelles and are capable of protein synthesis. By contrast axons cannot synthesize protein, so axonal proteins are derived from the cell body. Axons and dendrites both have mitochondria.

Related topics

Neuron diversity (A2)
Morphology of chemical synapses (C1)

Cell body

The **cell body** (**soma**, **perikaryon**) of a neuron (see *Fig. 1*) contains the nucleus, Golgi apparatus, ribosomes and other subcellular organelles, and is responsible for most of its routine metabolic 'housekeeping' functions. The neuron cell body resembles other cells, although it is specialized to maintain high levels of biosynthetic activity. The rough endoplasmic reticulum, for example, is so densely packed as to produce distinct structures called **Nissl bodies**, which are extremely rich in ribosomes. This reflects the high rates of protein synthesis of which neurons are capable.

Neurons come in a great variety of shapes and sizes. The smallest cell bodies are 5–8 μm in diameter, the largest 120 μm across.

Neurites

Neurons are distinguished from other cells by **neurites**. These are long cylindrical processes that come in two varieties: dendrites and axons. **Dendrites** are highly branched extensions of the cell body, may be up to 1 mm in length and can be 90% of the surface area of a neuron. Dendrites on some neurons are covered with hundreds of tiny projections termed **dendritic spines** on which synapses are made. Nerve cells with spines are called **spiny neurons**, those lacking them **aspiny neurons**. A neuron may have one or many dendrites, arranged in a pattern which is cell typical and collectively referred to as the **dendritic tree.** The majority of synaptic inputs from other neurons are made on dendrites.

Nerve cells generally have only one **axon.** It usually arises from the cell body but may emerge from a dendrite just where it leaves the cell body. In either case

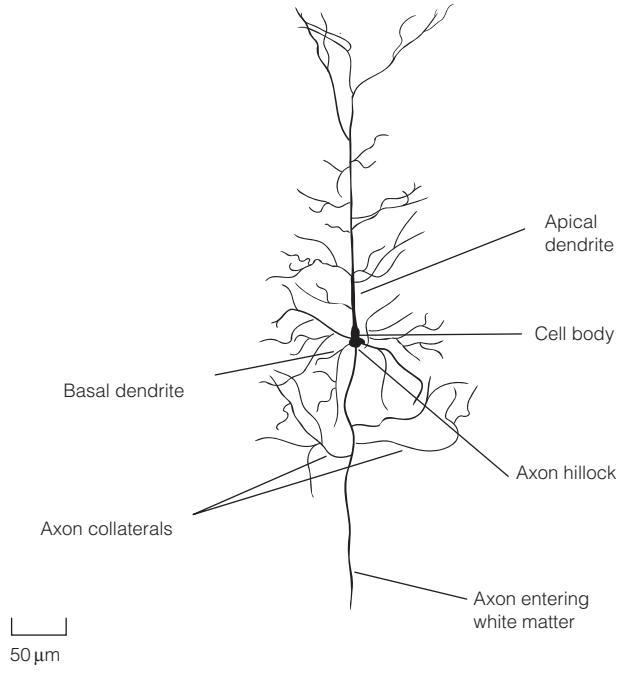


Fig. 1. Key features of a neuron. A drawing of a pyramidal cell showing the distribution of neurites (dendrites and axon).

the site of origin is termed the **axon hillock**. Axons have diameters 0.2–20 µm in humans (though axons of invertebrates can reach 1 mm) and vary in length from a few µm to over a meter. They may be encapsulated in a **myelin sheath**. Axons usually have branches, referred to as **axon collaterals.** The ends of an axon are swollen **terminals** (**boutons**) and usually contain mitochondria and vesicles. Some axons have swellings along their length called **varicosities**. Axon terminals and varicosities are the presynaptic components of chemical synapses.

Axon or dendrite? Axons can be distinguished from dendrites on structural grounds. Axons tend to be long, untapered, less highly branched, are never spiny and may have a myelin sheath, whereas dendrites are shorter, tapered, highly branched and may bear spines. Dendrites are extensions of the cell body in that they contain Golgi apparatus, rough endoplasmic reticulum and ribosomes (organelles not seen in axons). Since axons do not possess protein synthetic machinery, proteins in axons must be made in the cell body and moved along the axon by **axoplasmic transport**. Both axons and dendrites have mitochondria. Axon terminals are rich in mitochondria, indicating their high requirement for metabolic energy.

A2 Neuron diversity

Key Notes

Neuron classification

Neurons may be classified by their morphology, function or by the neurotransmitters they secrete. Cells with one, two or three or more neurites are classed as unipolar, bipolar or multipolar, respectively. The shape of the dendritic tree, whether the dendrites have dendritic spines or not, and the length of the axon, have all proved useful in categorizing neurons. Functional classification distinguishes afferent neurons that provide input and efferent neurons that provide output to a region of the nervous system. Neuron shape is often a good guide to the neurotransmitters it secretes, and so to its function

Neuron numbers

The human nervous system may contain 300–500 billion neurons. Neuron density is quite constant across the cerebral cortex and between cerebral cortices of different mammals. Smaller brains have fewer neurons.

Related topics

Organization of the peripheral nervous system (A4)
Organization of the central nervous system (A5)

Neuron classification

There is no such thing as a 'typical' neuron. Nerve cells come in diverse shapes and sizes, each with their own distinctive patterns of synaptic contacts and chemical transmitters. This allows neurons to be classified according to their morphology, neurotransmitters and function.

Structural ways to classify a nerve cell include the size of its cell body, the number of neurites it has, the pattern of its dendritic tree, axon length and the nature of the connections it makes. A neuron with a single neurite is **unipolar**, one with two neurites is **bipolar**, while a neuron with three or more is said to be **multipolar** (*Fig. 1*). The majority of neurons in the vertebrate nervous system are multipolar but there are important exceptions. For example, **bipolar** neurons in the retina synapse with photoreceptors, and sensory neurons in the dorsal root ganglion are described as **pseudounipolar**; so called because they start life as bipolar cells, having two processes, which subsequently fuse. Invertebrate nervous systems are dominated by unipolar neurons.

Dendrites are used to classify neurons on the basis of whether or not they have dendritic spines and the overall pattern of their dendritic tree. The shape of any dendritic tree helps determine the efficacy of its synaptic connections and so the functioning of the cell. **Pyramidal** cells, so called because of the shape of their cell bodies, comprise some 60% of neurons in the cerebral cortex and have dendrites which extend to fill a pyramidal space. A second population of cortical cells is termed **stellate cells** because of the star-like appearance of their dendritic trees. **Purkinje** cells of the cerebellar cortex have the unique feature that their extensive network of dendrites forms a two-dimensional array.

Neurons can also be classified on the basis of the lengths of their axons. **Projection** (**principal**, **relay** or **Golgi type I**) neurons have long axons, which

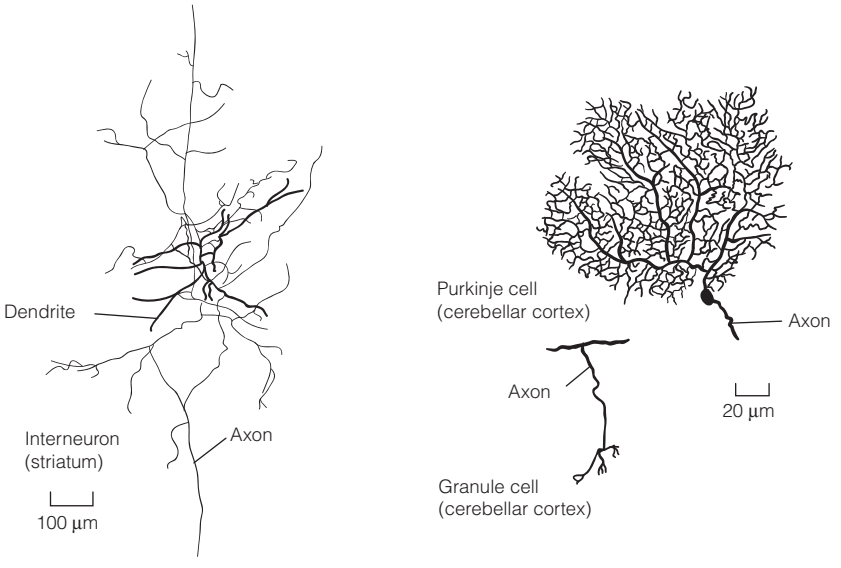


Fig. 1. The morphologies of three common types of neuron. The full length of the axons is not shown. The bifurcating axon of the granule cell extends for several millimeters in each direction. Note how the axon of the interneuron branches extensively.

extend way beyond the region of the nervous system in which their cell body resides. Pyramidal and Purkinje cells fall into this category. In contrast, **interneurons** (**intrinsic or Golgi type II** neurons) have short axons. These local circuit neurons, such as stellate cells, produce direct effects only in their immediate neighborhood.

By examining the connections that a neuron makes it is possible to classify neurons by function. Any given region of the nervous system receives inputs from **afferent** neurons and projects by **efferent** neurons to other regions of the nervous system or an effector organ (such as a muscle or gland). **Afferent neurons** that synapse with sensory receptors or which are themselves capable of responding directly to physiological stimuli are **sensory** neurons. Efferent neurons which synapse with skeletal muscles are termed **motor** neurons. Sometimes the term motor neuron is applied to projection neurons in motor pathways even if they do not directly synapse with a muscle.

Finally, neurons can be classed according to the neurotransmitters which they secrete. Moreover, there is often a clear correlation between neuron morphology and neurotransmitter. In other words, the shape of a neuron allows an intelligent guess to be made about which transmitter it secretes. For example, pyramidal cells release glutamate, whereas stellate and Purkinje cells secrete γ-aminobutyrate. This, in turn, provides very strong circumstantial evidence as to function, because usually glutamate excites, while γ-aminobutyrate inhibits, other neurons.

Neuron numbers

Estimates of the number of neurons in the nervous system can be made by statistical analysis of cell counts in thin-tissue sections viewed under light microscope. This shows that the number of neurons per unit area of the cerebral cortex is remarkably constant from one area of the cortex to another in humans,

and across mammalian species, at around 80 000 mm^{-2}. The exception is the primary visual cortex where the neuron density rises to 200 000 mm^{-2}. Assuming a total surface area for the human cerebral cortex of 2000 mm^2, these figures suggest that there are some 1.6×10^{11} neurons in the cerebral cortex alone. The most populous cells in the mammalian nervous system are small granule cells of the cerebellum; in humans they may number 10^{11}. Hence the human nervous system contains at least 2.5×10^{11} neurons; the total is likely to be between 300 and 500 billion! Smaller mammals have smaller brains because they have fewer neurons, not because their neurons are smaller.

A3 Glial cells and myelination

Key Notes

Classes of glial cells

Glial cells perform a number of functions that support neurons. There are many more glial cells than neurons. Glial cells can be assigned to one of three major populations: astrocytes, oligodendrocytes (including peripheral Schwann cells) and microglia.

Astrocytes

Astrocytes are large, numerous, star-shaped glia which have elongated processes tipped with endfeet. They cover synapses, form contacts with capillary endothelial cells and with the pia mater where they form a limiting glial membrane. The functions of astrocytes include homeostatic regulation of the extracellular K^+ concentration, the synthesis of transmitter glutamate and γ-aminobutyrate, removal of neurotransmitters from the synaptic cleft, storing glycogen, and supplying lactate to neurons.

Oligodendrocytes and Schwann cells

Oligodendrocytes in the central nervous system (CNS) and Schwann cells in the peripheral nervous system are responsible for forming the insulating myelin sheath that surrounds many axons. The sheath is produced by part of the glial cell spiraling around the axon a number of times. The sheath is interrupted at regular intervals by nodes of Ranvier, tiny gaps where the axon membrane is naked.

Microglia

Microglia are small immune cells derived from monocytes. In their macrophage guise they are key players in the inflammatory processes that accompany repair of nervous system injury.

Related topics

Blood–brain barrier (A8)
Neurotransmitter inactivation (C7)
Action potential conduction (B5)
Cell determination (N2)

Classes of glial cells

As well as neurons, the nervous system contains glial cells. These are thought not to be directly involved in information processing but instead perform a variety of supporting functions without which neurons could not operate. Estimates suggest that glial cells outnumber neurons perhaps by as much as ten-fold. It is hardly surprising then that the total cell density in nervous tissue is extremely high and the brain has the lowest extracellular space of any organ in the body. Glial cells are divided into **macroglia** and **microglia**. Several distinct populations of macroglia are recognized: astrocytes, oligodendrocytes and Schwann cells.

Astrocytes

Astrocytes are the largest and most numerous of glial cells. They are irregularly shaped cells and many have long processes, which superficially resemble the

dendrites of neurons. Astrocytes can readily be distinguished from neurons however; they do not have Nissl bodies and can be stained using immunocytochemistry for a specific astrocyte marker, **glial fibrillary acidic protein**. Astrocytes fill most of the space between neurons leaving gaps only about 20 nm across. Astrocyte processes surround synapses and some form **endfeet** which butt onto capillaries or onto the pia mater (the innermost layer of meninges) to produce a layer covering the surface of peripheral nerves and CNS called the **glial membrane.**

Astrocytes have a wide variety of functions:

1. Removing K^+ that accumulates in the extracellular space as a result of neural activity and dumping it, via their endfeet, into capillaries. This maintains appropriate potassium concentrations in the vicinity of neurons.
2. Uptake or synthesis of the precursors for the two major neurotransmitters, glutamate and γ-aminobutyrate.
3. Terminating the actions of small transmitter molecules by removing them from the synaptic cleft by way of specific transporters in their plasma membranes.
4. Providing energy substrates to neurons. Astrocytes take up glucose from blood and either store it as glycogen or convert it to lactate which is exported to neurons. This may be particularly important for highly active neurons that require more energy than can be supplied by glucose crossing the blood–brain barrier.
5. Ammonia detoxification via the ornithine–arginine cycle, and detoxification of free radicals.
6. As radial glial cells they guide neurons to their proper destinations in the developing brain.
7. Regulation of synapse formation in the developing brain and the production of new neurons in the adult brain.
8. Ensuring the integrity of the blood–brain barrier by influencing endothelial cells to form tight junctions.

Oligodendrocytes and Schwann cells

Oligodendrocytes in the CNS and **Schwann cells** in the peripheral nervous system have the common function of providing the **myelin sheath**, an electrically insulating covering around many axons. Those axons with a myelin sheath are said to be **myelinated**, those without are termed **unmyelinated**. The myelin sheath is formed in the peripheral nervous system in the following way. Schwann cells line up along the axon surrounding the axon with a pseudopodium-like structure, the **mesaxon**. For unmyelinated axons the process stops at this point. For myelinated axons the mesaxon spirals around the axon some 8–12 times. During this ensheathing most of the cytoplasm gets left behind (except in the innermost turn) so the majority of layers simply consists of a double thickness of plasma membrane (see *Fig. 1*). Each Schwann cell myelinates between 0.15 and 1.5 mm of axon. The thicker the axon the longer the region myelinated by a single glial cell. Between adjacent ensheathed regions is a tiny (0.5 μm) gap of naked axon called the **node of Ranvier**. Here the axon membrane is directly exposed to the extracellular space. Since a peripheral nerve may be quite long a few hundred Schwann cells might be required to generate the sheath. Myelinated axons vary in total nerve fiber diameter between 3–15 μm but across this range the proportion of the diameter contributed by the myelin sheath is roughly constant.

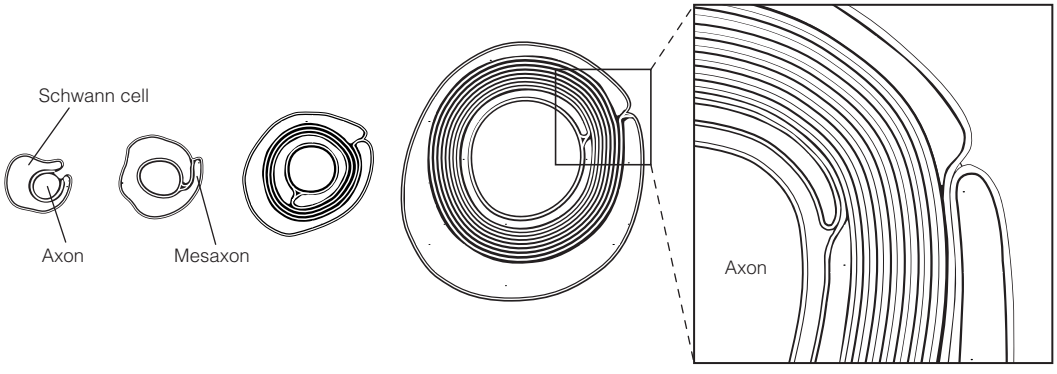


Fig. 1. Myelination of a peripheral axon. The myelin sheath is generated by the growth of the mesaxon which wraps itself around the axon. Redrawn from Gray's Anatomy, *37th edn, 1989, with permission from Harcourt Publishers Ltd.*

Myelination proceeds in a similar way in the CNS except that each oligodendrocyte extends several processes so that it can contribute to the myelination of several adjacent axons. This ensures that fewer glial cells are needed for CNS myelination, which saves space.

Multiple sclerosis is a progressive disorder in which multiple plaques (2–10 mm in size) of demyelination occur in the central nervous system, resulting in defects in the propagation of action potentials. The optic nerves, brainstem and cervical spinal cord are particularly vulnerable. It is thought to be an autoimmune disease in which the immune system mounts an attack on one or more of the proteins of which myelin is made. Immigrants from low- to high-prevalence zones (e.g. tropics to northern Europe) come to have the same risk of developing the disease as the high-prevalence zone natives if they arrive before the age of 10 years. The remissions seen in the disease presumably reflect attempts at repair and re-myelination.

Microglia

The smallest of the glial cells are the **microglia.** These are cells of the immune system. Derived from bone marrow monocytes, they migrate into the nervous system during development where they secrete growth factors, guide axons, and stimulate the differentiation of glial cells and the formation of blood vessels. Acting as macrophages they phagocytose debris generated by the programmed cell death that is normal during development. In the adult nervous system they cease being motile and may have some homeostatic function. Microglia are important in repair of nervous system damage. They proliferate and revert to their macrophage lifestyle, releasing cytokines, in a wide variety of conditions that produce inflammation of the nervous system such as infections, trauma and tumors. Cytokine release by activated microglia also contributes to pathological pain that can occur in a number of diseases, such as cancer. Scar tissue formation in the CNS as a result of the activities of microglia is called **gliosis**.

A4 Organization of the peripheral nervous system

Key Notes

Principal divisions of the nervous system

The brain and spinal cord comprise the central nervous system, while the peripheral nervous system, divided into somatic, autonomic and enteric parts, is everything else.

Somatic nervous system

Thirty-one pairs of spinal nerves originating from the spinal cord and 12 pairs of cranial nerves arising from the brain make up the somatic nervous system. Almost all spinal nerves are mixed nerves, containing both sensory and motor fibers. Of the cranial nerves only four are mixed; some are purely sensory and others purely motor. Every spinal segment gives rise to a pair of spinal nerves, each with a dorsal root containing sensory fibers and a ventral root with motor fibers. The cell bodies of the sensory fibers lie outside the spinal cord in the dorsal root ganglia.

Peripheral nerves

Peripheral nerves consist of nerve fibers (axons surrounded by their associated Schwann cells) organized into bundles (fasciculi) and invested with connective tissue. Peripheral nerve fibers are classified by their diameters and conduction velocities.

Autonomic nervous system (ANS)

This visceral motor system originates with cell bodies in the brainstem and spinal cord that give rise to preganglionic myelinated axons which secrete acetylcholine. They synapse with postganglionic unmyelinated fibers in autonomic ganglia. The ANS has two divisions, sympathetic and parasympathetic. The sympathetic system arises from thoracic and lumbar spinal segments, has its autonomic ganglia close to the cord in the paravertebral chains or subsidary ganglia, and its long postganglionic fibers usually secrete noradrenaline (norepinephrine). The adrenal medulla secretes adrenaline (epinephrine) into the bloodstream under the influence of preganglionic sympathetic innervation. The parasympathetic system originates from the brainstem and sacral spinal cord. Its autonomic ganglia are located on or near the innervated organ. The short postganglionic fibers secrete acetylcholine.

Enteric nervous system (ENS)

The nervous system of the gut is organized into two highly interconnected cylindrical sheets of neurons embedded in the gut wall called the submucosal and myenteric plexuses. The enteric nervous system regulates gut function semi-autonomously although its activity is modified by the ANS.

Related topics

Nerve–muscle synapse (J1)
Autonomic nervous system function (L5)

Principal divisions of the nervous system

The nervous system is comprised of the **central nervous system (CNS)** and **peripheral nervous system (PNS**). These divisions are contiguous both anatomically and functionally. The CNS includes the brain and spinal cord. The peripheral nervous system is everything else; namely nerve trunks going between the CNS and the periphery, and the networks of nerve cells with supporting glia in organs throughout the body. The PNS has three subdivisions, the somatic, autonomic and enteric nervous system.

Somatic nervous system

The somatic nervous system structure reflects both the bilateral symmetry and segmented nature of the vertebrate body plan. In humans it consists of 31 pairs of **spinal nerves**, each pair arising from a single segment of the spinal cord, and 12 pairs of **cranial nerves**, which have their origin in specific regions in the brain. Axons in spinal and cranial nerves that enter the CNS are called **afferent** (**centripetal**) **fibers**, while axons leaving the CNS are **efferent** (**centrifugal**) **fibers**. Afferent fibers carry sensory information from skin, muscles, joints and viscera. The majority of these afferents are wired to **mechanoreceptors** which inform about mechanical forces impinging on the surface or produced within the body. Some act as **nociceptors** which respond to factors associated with tissue damage, and some (restricted to skin) are connected to **thermoreceptors** which are temperature sensitive. Efferent fibers are the axons of motor neurons supplying skeletal muscles. The synapse between a motor neuron and a skeletal muscle fiber is called a **neuromuscular junction**.

All spinal nerves (except C1 which is motor and Cx 1 which is sensory) are **mixed**, that is, they contain both sensory and motor fibers. Of the cranial nerves only four are mixed (see *Table 1*). The olfactory, optic and vestibulocochlear are essentially pure sensory nerves, while the oculomotor, trochlear, abducens, accessory and hypoglossal are purely motor. The optic nerves, unique among the cranial nerves, develop as direct outgrowths of the brain so it is argued that they and the retinae are part of the CNS. All other peripheral nervous system components arise from the neural crest.

Each spinal nerve is formed from a **dorsal root** housing sensory fibers and a **ventral root** carrying motor fibers. The cell bodies of the primary afferent neurons lie within the **dorsal root ganglia (DRG**) just outside the spinal cord. There are a pair of DRG for each spinal segment. Efferent neuron cell bodies lie within the spinal cord (*Fig. 1*).

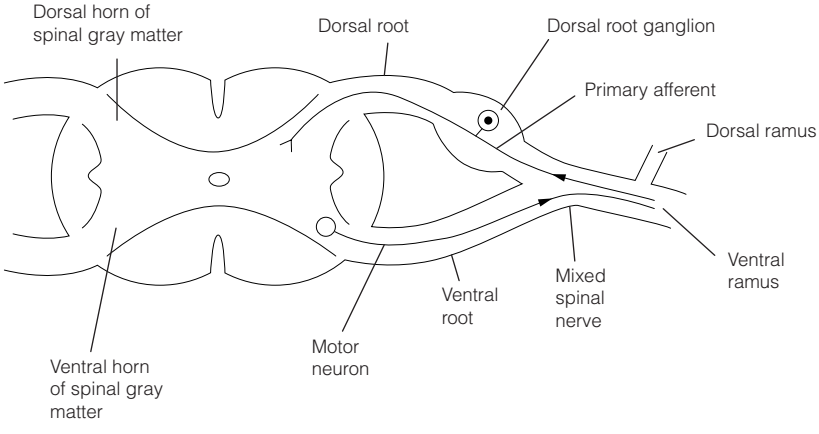


Fig. 1. Origin of a spinal nerve from a spinal cord segment.

Table 1. Peripheral nerves

Nerve		Type	Region of origin or destination in CNS	Function
Cranial nerves				
I	Olfactory	Sensory	Olfactory bulb	Smell
II	Optic	Sensory	Forebrain LGN (thalamus)	Vision
			Midbrain superior colliculus tectum	Visual reflexes
III	Oculomotor	Motor[a]	Midbrain	Motor to extrinsic eye muscles except superior oblique and lateral rectus, autonomic to intrinsic eye muscles
IV	Trochlear	Motor	Midbrain	Motor to superior oblique extrinsic eye muscles
V	Trigeminal	Mixed	Midbrain and hindbrain	Sensory from head and face, motor to jaw muscles
VI	Abducens	Motor	Hindbrain	Motor to lateral rectus extrinsic eye muscles
VII	Facial	Mixed[a]	Ventral lateral thalamus (sensory)	Sensory from tongue (taste) and palate
			Hindbrain (motor)	Motor to face, parasympathetic secretomotor to submandibular, submaxillary salivary glands and lachrymal glands
VIII	Vestibulocochlear	Sensory	MGN (auditory division)	Sensory from inner ear (hearing and balance)
			Hindbrain (vestibular division)	
IX	Glossopharyngeal	Mixed[a]	Hindbrain	Sensory from tongue (taste)
				Motor to pharyngeal muscles
				Parasympathetic secretomotor to parotid salivary glands
X	Vagus	Mixed[b]	Hindbrain	Sensory from viscera
				Somatic motor to pharyngeal and laryngeal muscles
				Parasympathetic to viscera
XI	Accessory	Motor	Medulla, spinal cord C1–C5	Motor to palate and some neck muscles
XII	Hypoglossal	Motor	Medulla	Motor to tongue
Spinal nerves				
C1–8		Mixed		
T1–12		Mixed (including sympathetic autonomic T1–12)		
L1–5		Mixed (including sympathetic autonomic L1, 2)		
S1–5		Mixed (including parasympathetic autonomic S2, 3)		
Cx 1		Mixed		

[a]Including autonomic.

[b]Large autonomic component.

LGN, lateral geniculate nucleus; MGN, medial geniculate nucleus.

Peripheral nerves All peripheral nerves have a common basic structure. A nerve fiber consists of an axon together with accompanying Schwann cells. Several unmyelinated axons are invested by a single glial cell, which comprises the **neurolemma.** In myelinated axons this term is reserved for the outer, nucleated, cytoplasm-rich portion of the Schwann cell. Nerve fibers are collected into bundles, **fasciculi,** surrounded by a connective tissue sheath, the **perineurium**. Within the

fasciculus individual fibers are supported by a connective tissue network, the **endoneurium**, which is continuous with the perineurium. A nerve may be one or several fasciculi, all encapsulated by a connective tissue **epineurium**.

Two systems for the classification of PNS nerve fibers are in common use. They are based on fiber diameter and conduction velocity and are summarized in *Table 2*. The Erlanger and Gasser system is used to classify both afferents and efferents. The Lloyd and Hunt scheme is used exclusively to define afferent fibers.

Table 2. Classification of peripheral nerve fibers

Fiber (type/group)	Mean diameter (μm)	Mean θ ($m\ s^{-1}$)	Functions (example)
Erlanger/Gasser classification (type)			
Aα	15	100	Motor neurons
Aβ	8	50	Skin touch afferents
Aγ	5	20	Motor to muscle spindles
Aδ	4	15	Skin temperature afferents
B	3	7	Unmyelinated pain afferents
C	1	1	Autonomic postganglionic neurons
Lloyd/Hunt classification (group)			
I	13	75	Primary muscle spindle afferents
II	9	55	Skin touch afferents
III	3	11	Muscle pressure afferents
IV	1	1	Unmyelinated pain afferents

Autonomic nervous system (ANS)

The ANS is the visceral motor nervous system. By definition it includes no sensory components. However, the activities of the ANS are modified by sensory input that travels by way of the somatic nervous system, and by the CNS. The target tissues of the ANS are smooth muscle, cardiac muscle, endocrine and exocrine glands, liver, the juxtaglomerular apparatus of the kidney and adipose tissue. The synapses of autonomic neurons with their target cells are called **neuroeffector junctions**.

The **preganglionic** neurons of the ANS have their cell bodies in motor nuclei of cranial nerves in the midbrain or medulla, or the intermediolateral horn of the thoracic and upper lumbar spinal cord. Their axons are myelinated B fibers, which secrete **acetylcholine**. The preganglionic axons synapse with postganglionic neurons in **autonomic ganglia**. The axons of the postganglionic neurons are unmyelinated C fibers. The ANS has two divisions, the **sympathetic** and **parasympathetic**, the main distinguishing features of which are summarized in *Table 3.*

Table 3. Divisions of the autonomic nervous system

Anatomy	Physiology	Postganglionic cell neurotransmitters
Craniosacral Preganglionic axons in cranial nerves III, VII, IX, X and spinal nerves S2, S3	Parasympathetic	Acetylcholine Vasoactive intestinal peptide
Thoracolumbar Preganglionic axons in spinal nerves T1–T12, L1, L2	Sympathetic	Norepinephrine (but acetylcholine at selected neuroeffective junctions) Neuropeptide Y Adenosine 5′-triphosphate

In general the preganglionic axons of the sympathetic division are short and the postganglionic axons are long because the **sympathetic ganglia** lie close to the spinal cord in one of two locations:

- in paired **paravertebral chains** that lie just lateral of the vertebral column, running parallel to it in the neck and down the posterior wall of the thorax and abdomen;
- in **subsidiary ganglia** of autonomic plexuses situated in the midline adjacent to major blood vessels.

The pathway taken by sympathetic axons is illustrated in *Fig. 2.*

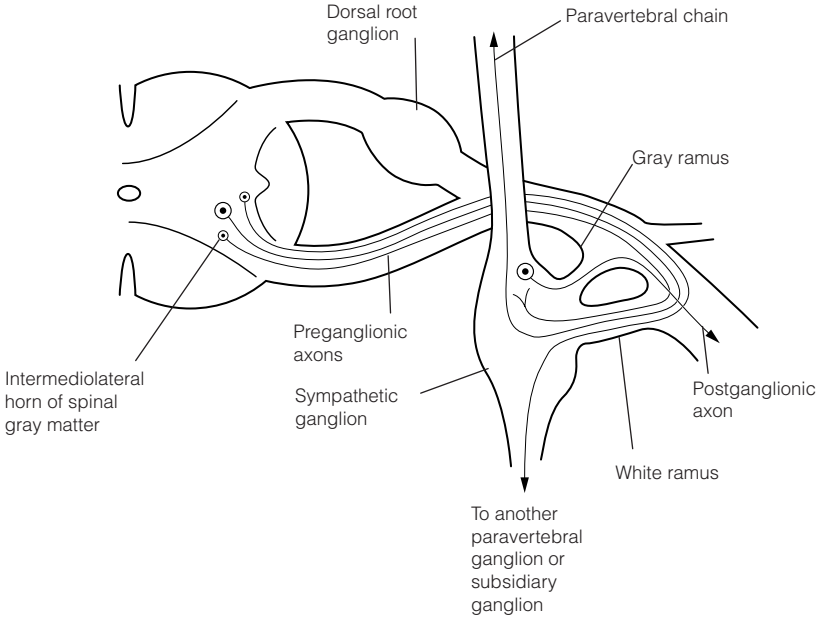


Fig. 2. Sympathetic pathway from the spinal cord.

Preganglionic axons may synapse in the nearest ganglion, traverse the paravertebral chain to synapse in subsidiary ganglia, or ascend or descend the chain to synapse in a ganglion at a different level. Preganglionic sympathetic axons can modify the actions of up to 100 postganglionic cells. This is an example of **divergence**, a phenomenon seen throughout the nervous system, which serves to disseminate neural activity. Most, but not all, **postganglionic sympathetic** axons secrete **noradrenaline** (**norepinephrine**). The **adrenal medulla** is endocrine tissue that secretes **adrenaline** (**epinephrine**) directly into the circulation in response to activity in the preganglionic sympathetic fibers which supply it. The adrenal medulla is therefore regarded as part of the sympathetic system.

Parasympathetic autonomic ganglia are all subsidiary ganglia located close to the target organ. For this reason, in the parasympathetic division the preganglionic fibers are long, and the postganglionic ones are short. All **postganglionic parasympathetic** fibers secrete **acetylcholine**. Although all the major organs, except the liver, have a parasympathetic supply, this division is far less

extensive than the sympathetic. This is partly because only a few specialized blood vessels have a parasympathetic innervation, whereas all blood vessel smooth muscle receives a sympathetic supply.

Enteric nervous system (ENS)

An interconnected network of about 10^8 neurons makes up the nervous system of the gut. It is organized into two thin cylindrical sheets that run along the length of the gut. The **myenteric (Auerbach's)** plexus lies between the longitudinal and circular smooth muscle layers and extends the whole length of the gut. The **submucosal (Meissner's)** plexus lies in the submucosa and extends from the pylorus of the stomach to the anus. There are extensive interconnections between these two plexuses. A number of amines, peptides and nitric oxide (NO) are employed as transmitters by this system. The ENS can act autonomously to coordinate gut motility and secretion. Its activity is modified by input from both divisions of the ANS.

A5 Organization of the central nervous system

Key Notes

Spinal cord

The human spinal cord contains about one hundred million neurons. Peripheral white matter and central gray matter can be seen by the naked eye. The gray matter of the spinal cord contains neuron cell bodies. Sensory neuron fibers enter the dorsal horn of the gray matter in an ordered fashion, larger diameter fibers entering more medially and extending deeper than smaller ones. Motor neuron cell bodies lie in the ventral horn of the gray matter. The spinal gray is divided on morphological grounds into ten columns which, on transverse section, are called Rexed laminae. Each lamina has a distinctive set of inputs and outputs. The white matter contains tracts of axons ascending or descending the cord. Neural tracts or pathways are named for their origin and destination.

The brain

The brain has three structural components. White matter consists of fiber tracts or pathways. Embedded in this are nuclei which are clusters of neuron cell bodies. Two large brain structures, the cerebrum and cerebellum are covered by cortex, a thin rind of gray matter densely packed with neurons. Anatomically the brain has three principal divisions: hindbrain, midbrain and forebrain. The center of the brain is taken up with the cerebrospinal fluid (CSF)-filled ventricular system. The hindbrain consists of medulla, pons and cerebellum, while the midbrain is divided into a ventral tegmentum and a dorsal tectum. Together hindbrain and midbrain are the brainstem from which emerge most of the cranial nerves. With the exception of the cerebellum, which organizes high-level motor functions, the brainstem is mostly concerned with vital functions and functions requiring orchestrated activity of large parts of the brain (e.g. wakefulness). The forebrain consists of diencephalon and cerebrum. The diencephalon contains a dorsal thalamus, serving sensory functions among others, and a ventral hypothalamus, implicated in temperature and endocrine regulation, and appetitive behaviors. The cerebrum has two cerebral hemispheres heavily interconnected across the midline. Its surface is covered by cortex, which has been subdivided on the basis of structural differences into Brodmann areas. The cerebral cortex has motor, perceptual and cognitive functions. The core of the cerebrum is occupied by nuclei, which form two neural systems. The basal ganglia form the extrapyramidal motor system and the limbic system (which includes cortical areas) is concerned with emotion and learning.

Related topics

Early patterning of the nervous system (N1)

Spinal cord

The human spinal cord has about 10^8 neurons. It has 31 segments, each of which gives rise to a pair of spinal nerves. It ends at the level of the first lumbar vertebra. The lumbosacral nerve roots pass down the vertebral canal as the **cauda equina** so that they emerge from the vertebral column at their appropriate levels.

A transverse section through the spinal cord shows a butterfly-shaped central **gray matter** which contains neuron cell bodies, **neuropil** (dendrites and short lengths of axon) and glia. The **white matter** surrounding the gray is largely axons in ascending and descending tracts and gets its color from the high content of myelin. In the middle is the **central canal**, which contains **cerebrospinal fluid** (**CSF**), though in adult humans it is often closed.

Sensory fibers enter the spinal cord via the dorsal roots to synapse largely with cells in the **dorsal horns** of the spinal gray matter. Larger-diameter fibers enter more medially and extend more deeply into the dorsal horn than smaller ones. Motor neuron cell bodies lie in the **ventral horns** of the spinal gray and their axons exit via the ventral roots. The distribution of afferents to dorsal roots and efferents from ventral roots is referred to as the **Bell–Magendie law**. Some **visceral** afferents however enter the spinal cord via the ventral roots.

Ten distinct regions can be distinguished in the spinal gray matter on the basis of cell size. Each region occupies a long column that extends through the cord. On transverse section these columns appear as **Rexed laminae** (*Fig. 1*).

Each lamina has distinctive input–output relations, which reflect a measure of functional specialization. Small-diameter afferents carrying pain and temperature information terminate on **dorsal horn cells** (**DHCs**) in lamina II. Cutaneous mechanoreceptor afferents terminate in deeper layers of the dorsal horn. Lamina VI is present only in spinal segments associated with limbs and receives sensory input from joints and muscles that provides information about the position and movement of the limb in space. Lamina VII includes the cell bodies of the preganglionic autonomic axons. Lamina IX houses both α and γ motor neurons which go to skeletal muscles.

The white matter is organized into columns or **tracts** each specified by its origin and destination. For example, the tract which runs down the cord from the cerebral cortex is termed the corticospinal tract whereas the ascending pathway which terminates in the thalamus is the spinothalamic tract. *Fig. 2* shows the locations of the major tracts.

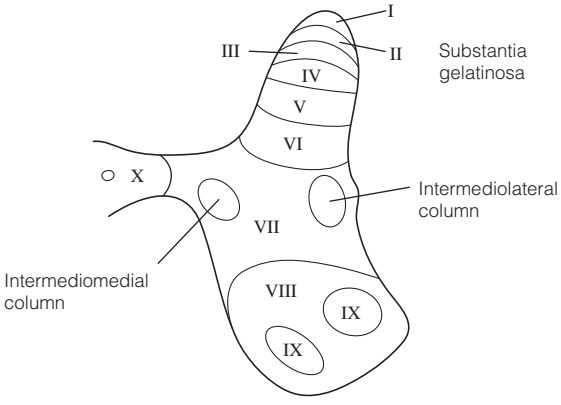


Fig. 1. Rexed laminae. Lamina VI is only present in spinal segments supplying the limbs.

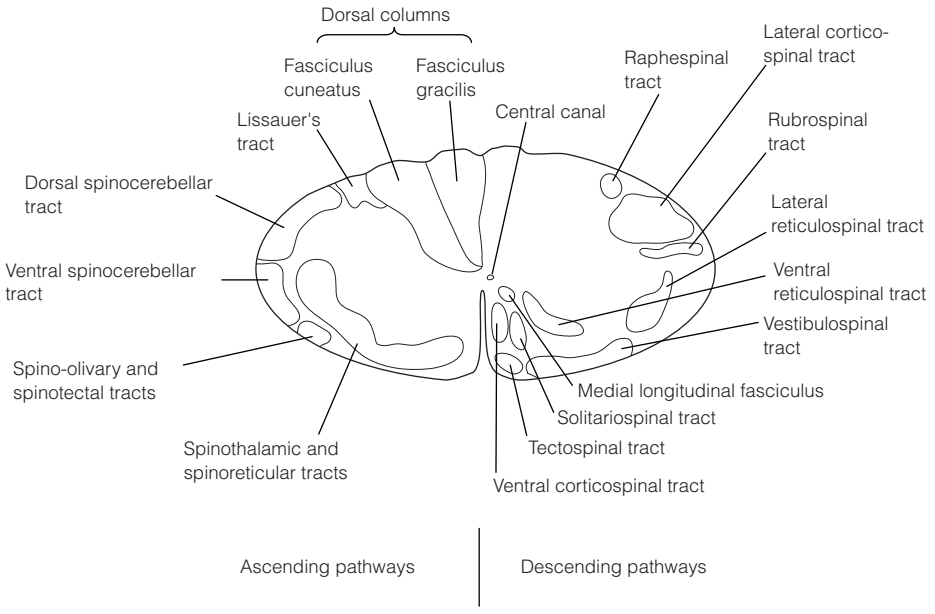


Fig. 2. Pathways in the spinal cord white matter.

The brain

There are three main structural components to the brain:

- **Tracts** or pathways enter the neuraxis at various levels, ascend or descend, and these, together with internal tracts which go from one part of the brain to another, constitute the white matter.
- **Nuclei** embedded in the white matter are clusters of neuron cell bodies and associated neuropil. Some neural structures are composed of groups of nuclei. The thalamus, for example, consists of some 30 nuclei. Interconnected nuclei in turn constitute neural systems.
- Two brain structures, the cerebrum and the cerebellum, are covered by **cortex**, a thin rind with a very high density of neuron cell bodies. In wiring terms cortex appears to be a simple circuit between just a few neuron types, repeated millions of times.

Together, the nuclei and cortex are the gray matter of the brain. **Neural systems** are comprised of interconnected nuclei and cortical regions, which have a common function. The visual system, for example, consists of the retinas, the lateral geniculate nuclei of the thalamus, the visual cortex and the pathways between them.

In the human embryo at the end of the 4th week the CNS is a hollow tube, the **neural tube**, the **caudal** (back) end of which becomes the spinal cord (*Fig. 3*). At its **rostral** (front) end are three swellings, **primary vesicles**, which are the most fundamental anatomical divisions of the brain. These are the **hindbrain**, **midbrain** and **forebrain**. As development unfolds the forebrain differentiates into a caudal **diencephalon** and a rostral **telencephalon** which in turn acquires two lateral swellings, the **cerebral hemispheres**. Down the center of the neural tube is the CSF-filled ventricular system.

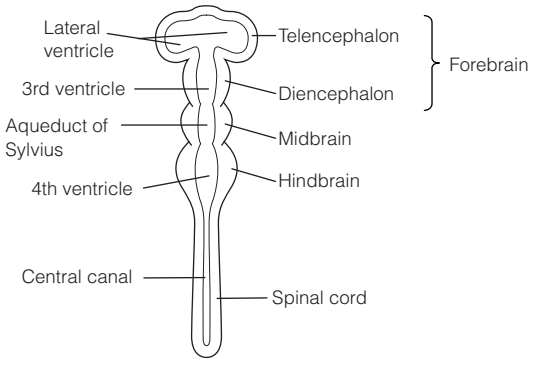


Fig. 3. The human embryo neural tube, at 28 days gestation.

Much of the neural tube is divided into a dorsal **alar plate** along the midline of which runs the **roofplate**, and a ventral **baseplate**, which has along its midline the **floorplate**. In the spinal cord and hindbrain these dorsal and ventral plates organize sensory and motor functions respectively. Such a clear distinction is not so obvious in the midbrain or forebrain.

The hindbrain subsequently develops into a caudal **medulla** and a rostral **pons** and (from about 12 weeks) a dorsal outgrowth, the **cerebellum**. The midbrain, which in the adult is the smallest part, acquires a ventral **tegmentum**, in which are found cell bodies of dopamine-using neurons that are part of a **motivation** system, and a dorsal **tectum**, which organizes visual and auditory reflexes. The hindbrain (minus the cerebellum) and midbrain together are often referred to as the **brainstem**. Much of the brainstem is occupied with vital (life-support) functions; for example, autonomic regulation of the cardiovascular system, and generation of the rhythmic neural output required for breathing. In addition, the brainstem contains the nuclei of most cranial nerves. A core of highly interconnected nuclei extending through the brainstem constitutes the **reticular system**. Many of its neurons (e.g. the midbrain dopamine cells mentioned above) use amine transmitters. The reticular system is involved in orchestrating global brain functions such as motivation, arousal, sleep and wakefulness and connects widely with the forebrain.

The diencephalon in the adult is differentiated into a dorsal thalamus and a ventral hypothalamus. The **thalamus** is a collection of over 30 distinct nuclei organized into five groups. All sensory input enters the cerebral cortex by way of the thalamus (ventral or posterior groups) with the exception of smell. Although clearly involved in sensory processes other nuclear groups are massively interconnected with cortical regions concerned with emotion (anterior group) and memory (medial group).

The **hypothalamus** is concerned with thermoregulation, triggering sleep, regulating endocrine systems, and goal-directed behaviors (eating, drinking and sexual behavior).

The smallest part of the diencephalon, the **pineal gland**, gets visual input and regulates circadian rhythms on the basis of the hours of light and dark.

The dominant part of the telencephalon is the **cerebrum**, two cerebral hemispheres linked across the midline by about 10^6 axons that constitute the **corpus callosum.** The cerebrum is massively developed in humans. Each hemisphere is

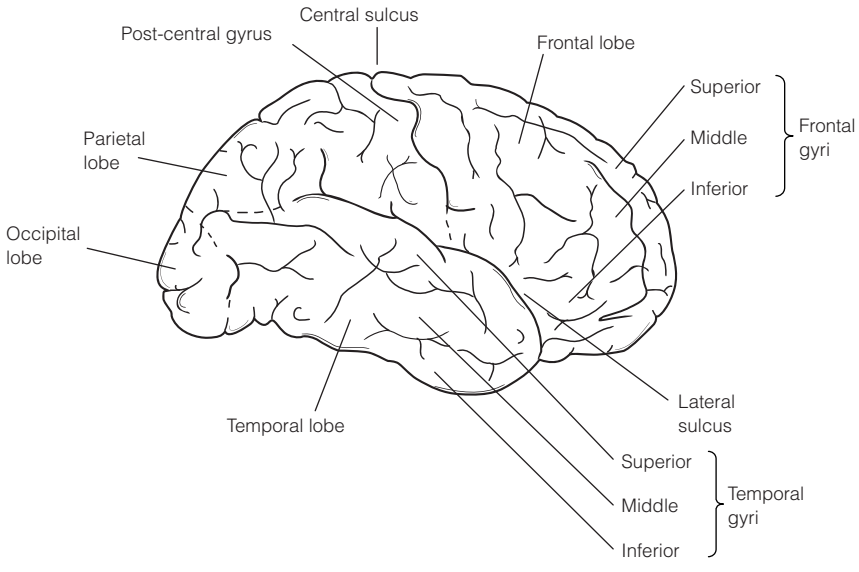


Fig. 4. Lateral surface of the human right cerebral hemisphere.

divided into four lobes named after the bones which overlie them (*Fig. 4*). The surface is covered by cortex and is highly convoluted giving it a high surface area in relation to its volume. The folds are called **gyri** (sing., **gyrus**), and the creases between them **sulci** (sing., **sulcus**). Most of the cerebral cortex is **neocortex** (new cortex), which has six layers. Cortical regions are mapped into **Brodmann** areas on the basis of differences in **cytoarchitecture**, that is, their cellular composition and relative thickness of the layers. The significance of this is that the Brodmann map corresponds quite well to how functions are localized in the cortex, though nowadays its main use is as a numerical guide.

The layers of the cerebral cortex are numbered from I, nearest to the pial surface through to VI which is the deepest (*Fig. 5*). The layers contain different proportions of two types of neurons, pyramidal cells which are output cells, and stellate cells that are interneurons. Layer I consists mostly of axons that run parallel to the cortical surface. Layers II and III have small pyramidal cells that project to other cortical areas. Layer IV is rich in interneurons and the site for the termination of most inputs to the cortex from the thalamus. Layer V has the largest pyramidal cells, which project to subcortical nuclei, brainstem and spinal cord. Layer VI pyramidal cells send their axons back to the same thalamic nucleus that supplied the inputs. The relative size of the layers differs with the function of the cortical region. For example, the sensory cortex has a thick layer IV because of its large number of thalamic inputs, whereas in motor cortex it is layer V which is particularly extensive because these neurons project to the brainstem and spinal cord to mediate motor activity.

The cerebral cortex is implicated in most brain activities, but is most often associated with the planning and execution of intentional movement, sensory perception and cognitive (problem-solving) functions. Those regions, which are not specifically devoted to sensory or motor activities, are called **association cortex**.

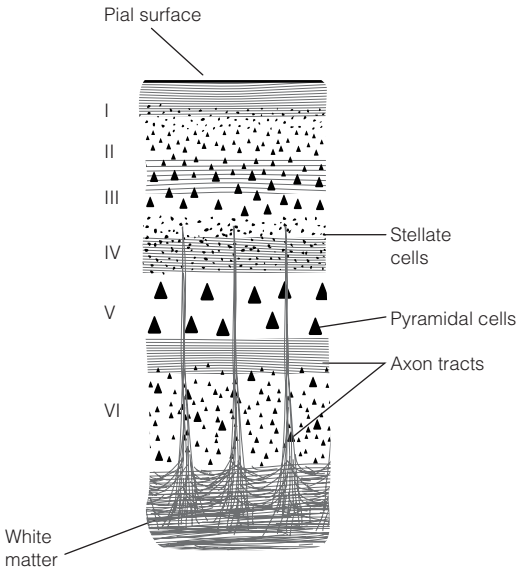


Fig. 5. Representative section through the neocortex (parietal lobe).

Within the core of each hemisphere lie clusters of nuclei that form major components of two neural systems, the basal ganglia and the limbic system (*Fig. 6*). The **basal ganglia**, responsible for organizing stereotyped patterns of movement, consists of the striatum, which lies in the forebrain, and two midbrain nuclei, the **subthalamus** and the **substantia nigra**. The **striatum** is subdivided into the **neostriatum**, itself composed of two nuclei, the **caudate** and **putamen**, and **paleostriatum** or **globus pallidus**. Anatomically the putamen and globus pallidus together form the **lentiform nucleus**.

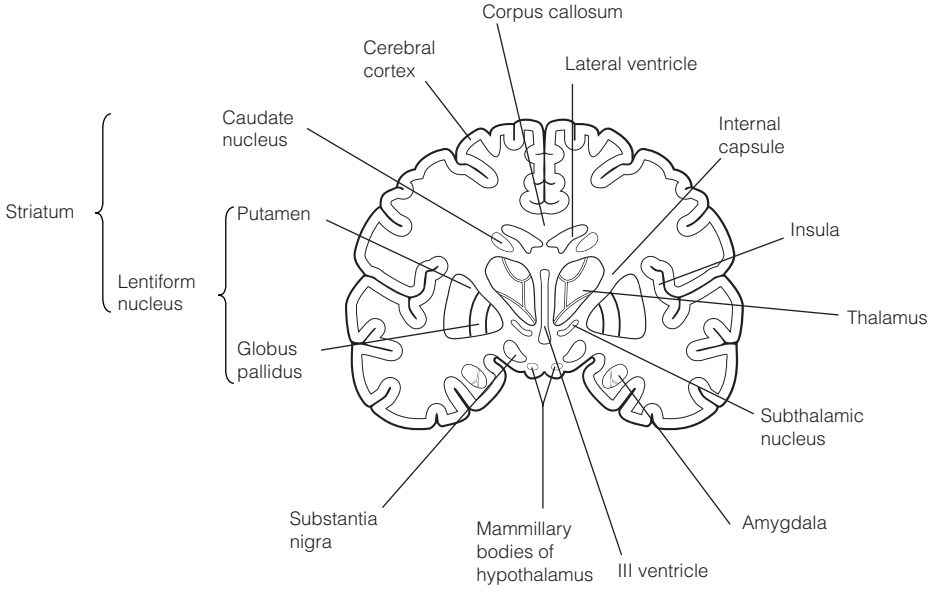


Fig. 6. Coronal section through the human cerebrum at the level of the posterior hypothalamus.

The **limbic system** is made from several heavily interconnected nuclei and several regions of cerebral cortex, which form a ring around the diencephalon (*Fig. 7*). The cortical regions are the **cingulate gyrus**, which lies above the corpus callosum and has contributions from medial, parietal and frontal lobes, and the **parahippocampal gyrus** and **uncus** which are part of the medial surface of the temporal lobe.

The medial and underside of the temporal lobe is occupied by the **hippocampal formation**, most of which is the **hippocampus** and **subiculum**. The hippocampus is **archaecortex** (ancient cortex) and has only three layers. Limbic system nuclei include the **amygdala**, **septal nucleus**, and the **mammillary bodies** (which are part of the hypothalamus). The hippocampus and amygdala are concerned with certain types of learning, and the limbic system in general is implicated in emotion.

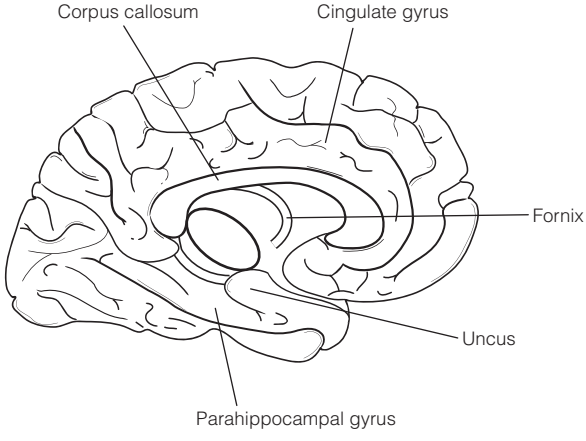


Fig. 7. Medial surface of the human left cerebral hemisphere.

A6 Brain imaging

Key Notes

Computer assisted tomography (CAT)

A CAT scan produces a series of X-ray images, each one a slice, of living brain. The technique can distinguish tissues that differ in their ability to transmit X-rays by as little as 1% and has a spatial resolution of 0.5 mm. Its main use is in the diagnosis of neurological disorders that can be revealed anatomically.

Positron emission tomography (PET)

PET scanning, by revealing the distribution in the brain of a positron-emitting isotope (with a spatial resolution between 4 and 8 mm) can provide functional as well as anatomical information about the living brain. Using uptake of a positron-emitting glucose analog as a marker for neuron activity it can show which regions of the brain are involved in a variety of activities both in health and disease. Positron-emitting neurotransmitters and receptor ligands are used to study neurotransmitter pathways in the living brain.

Magnetic resonance imaging (MRI)

Magnetic resonance imaging takes advantage of the fact that atomic nuclei with odd mass numbers aligned in a strong magnetic field resonate in response to a pulse of radio waves. When the radio pulses are switched off the nuclei relax in a way that depends on their chemical environment. fMRI has a spatial resolution < 1 mm. It can be used to measure changes in brain activity with a time resolution of a few seconds.

Related topics

Parallel processing in the visual system (G7)
Cortical control of voluntary movement (K1)
Language (O5)
Strokes and excitotoxicity (P3)
Parkinson's disease (P5)

Computer assisted tomography (CAT)

The first imaging technique developed to allow visualization of the living brain, the CAT scan, interposes the head between a source which emits a narrow beam of X-rays and an X-ray detector (*Fig. 1*). A series of measurements is made of X-ray transmission. The source and detector are rotated as a pair through a small angle and a further series of measurements taken. This is repeated until the source and detector have rotated through 180°. The radiodensity of each region of the head is computed from the transmission data for all of the beams that have traversed that region, and the results visually displayed. This provides a view through a single slice of brain lying at a known orientation. The key element here is the algorithm – and the computer software to implement it – that calculates the radiodensity for each point in the brain slice; this is **computerized tomography**. By moving the head at right angles to the orientation plane for a short distance another section can be imaged. This is repeated until the whole brain has been scanned.

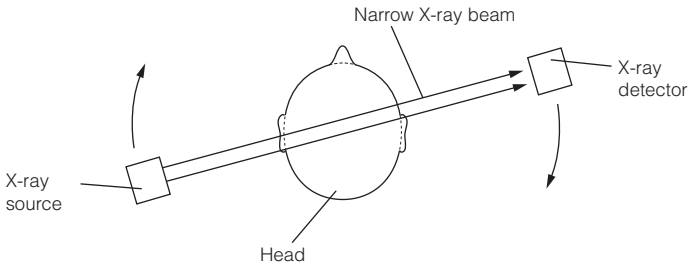


Fig. 1. Computer assisted tomography (CAT). Arrows depict the rotation of the scanner.

The method can distinguish tissues which differ in X-ray opacity by 1% (the lower the density the darker the image) with a spatial resolution of about 0.5 mm. Blood vessels can be seen by injection of radio-opaque dyes. This allows cerebrovascular disease or tumors and abscesses with abnormal vascularities to be revealed.

Positron emission tomography (PET)

A remarkable feature of PET is that it provides insights into the *function* of the living brain as well as its anatomy. It uses the principles of computerized tomography in which a γ-ray detector is rotated around the head, but the source is a positron-emitting compound, either injected or inhaled, which enters the brain (*Fig. 2*). Compounds used include neurotransmitters, receptor ligands, and glucose analogs which are used for studying brain activity. Typically they are radiolabeled with $^{11}_{6}C$, $^{13}_{7}N$, $^{15}_{8}O$ or $^{18}_{9}F$ (which substitutes for hydrogen). These isotopes have short half-lives, decaying to the element with atomic number one less; a proton (p) within the nucleus decays to a neutron (n) emitting a positron (e^+) in the process:

$$^{13}_{7}N \rightarrow {}^{13}_{6}C + e^+$$

$$p^+ \rightarrow n + e^+$$

The positron travels a short distance before colliding with an electron (e^-). The two particles annihilate with the production of two γ-ray photons that shoot off in exactly opposite directions. These are detected simultaneously by a pair of detectors 180° apart. This coincidence detection permits localization of the site of the γ-ray emission, which is between 2 and 8 mm from the positron source, depending on the isotope used.

The spatial resolution of PET is about 4–8 mm, not as good as CAT, but it can be used to follow brain events over time.

The importance of PET in functional studies is illustrated by the use of the non-metabolizable analog of glucose, **2-deoxyglucose** (**2-DG**). This molecule crosses the blood–brain barrier, is transported into neurons and phosphorylated to 2-DG-6-phosphate, so it remains in the cell. However, it cannot be metabolized further. This means it acts as a marker for local glucose uptake and therefore of neuron activity. Imaging the distribution of [$^{18}_{9}$F]2-DG while subjects engage in sensory, motor or cognitive tasks reveals how these functions are localized in the brain. Related studies show that during transient increases in neuronal activity, the rise in local cerebral oxygen consumption (as measured by $^{15}_{8}O$ PET) does not match the increase in glucose utilization (as estimated from 2-DG PET). This implies that brief periods of brain activity can be supported by glycolysis.

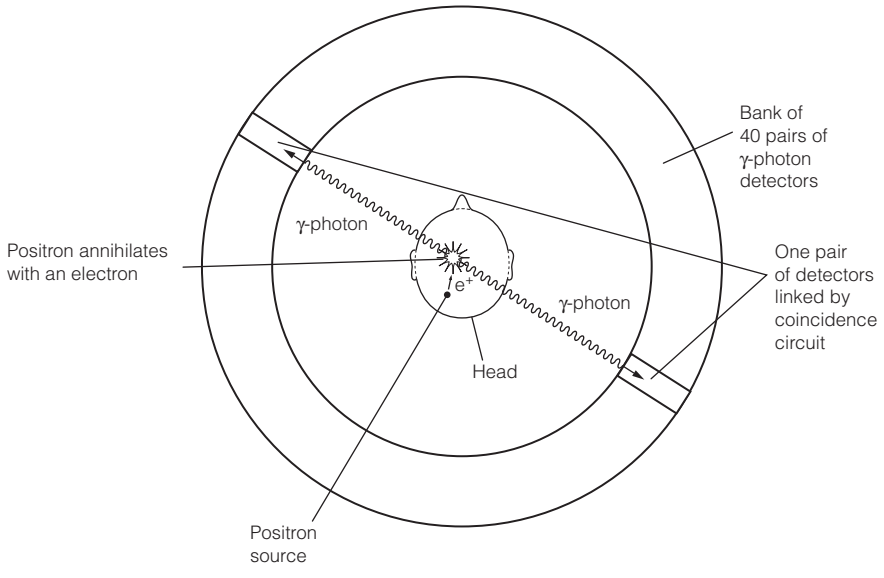


Fig. 2. Positron emission tomography (PET).

Magnetic resonance imaging (MRI)

Like PET, magnetic resonance imaging provides information about brain function as well as anatomy. It combines computerized tomography with **nuclear magnetic resonance** (**NMR**). Nuclei with odd mass number, for example, $^{1}_{1}H$, generate a magnetic field along their spin axis. In the powerful magnetic field of an MRI scanner, hydrogen nuclei can adopt one of two orientations; with their magnetic fields either parallel or antiparallel to the external field. The parallel state has a slightly lower energy and normally a small excess of nuclei will be in this state (*Fig. 3*). This gives rise to a net *longitudinal* magnetic field parallel to the scanner field.

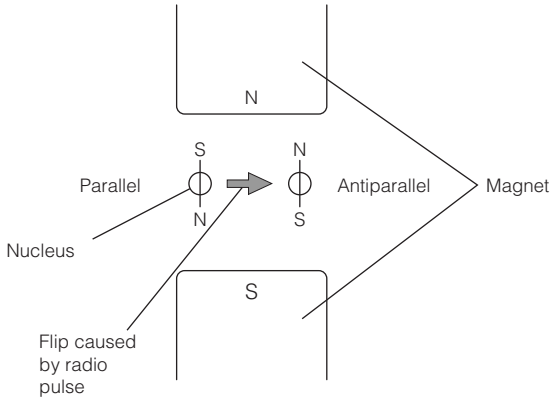


Fig. 3. The principle of nuclear magnetic resonance (NMR). A radiofrequency pulse will excite atomic nuclei, flipping them from the parallel state into the higher energy antiparallel state. Relaxation of the nuclei back into the low energy state generates the magnetic resonance imaging (MRI) signal.

A cylindrical coil placed around the head broadcasts a radio frequency (RF) pulse to a slice of head at right angles to the main scanner field. The RF pulse makes the nuclei wobble around their magnetic axis – rather like a spinning top as it slows down – with the rate of wobbling in resonance with the pulse frequency. The wobble generates an electric field that is received by the coil, producing a *transverse* magnetic field at right angles to the scanner field. When the RF pulse is turned off the nuclei return to their original state, and the longitudinal and transverse fields decay with **relaxation times** that are characteristic for the nucleus and its chemical environment (e.g. lipid or aqueous). Generating an MRI image actually requires a further three coils that produce magnetic field gradients in the x, y and z directions.

MRI has numerous clinical uses, for example: mapping cerebral blood vessels; showing changes in extracellular space that accompany trauma or inflammation; diagnosis and following the progress of a variety of diseases (such as multiple sclerosis); and the precise localization of regions of stroke damage or tumors. It has a better resolution (< 1 mm) than PET.

An MRI method that records changes related to brain function in successive images is termed **functional MRI** (**fMRI**). The most important type of fMRI is **blood oxygen level detection** (**BOLD**) which provides a more sensitive measure of cerebral cortical activity than PET, following changes in activity with a time resolution of a few seconds. It depends on the ratio of oxygenated to deoxygenated hemoglobin and this varies with blood flow, metabolism, and other variables in ways that are not yet completely understood.

A7 Meninges and cerebrospinal fluid

Key Notes

Meninges

The brain and spinal cord are invested by three connective tissue layers, the meninges. Directly covering the brain is the pia mater, above which is the arachnoid mater. Between these layers lies the subarachnoid space, which is filled with cerebrospinal fluid (CSF) and through which run blood vessels, branches of which enter the brain. Passive exchange of water and solutes across the pia mater keeps brain extracellular fluid and CSF in equilibrium. The tough outer layer is the dura mater, which contains venous sinuses. Projections of the arachnoid mater herniate into the venous sinuses. Here, tiny one-way valves allow the bulk flow of CSF from subarachnoid space into the venous circulation. The potential space between the arachnoid mater and the dura mater is the subdural space. Traumatic rupture of the veins passing through this space from brain to venous sinuses causes subdural hemorrhage. Between the dura and the cranial bones is the extradural space through which run major arteries. Traumatic rupture of these results in extradural hemorrhage.

Cerebrospinal fluid circulation

CSF is actively secreted by the choroid plexuses located in the ventricles. The direction of CSF flow is from lateral to 3rd to 4th ventricles, from where it enters the subarachnoid space. Finally it drains into the venous sinuses. Obstruction to the flow of CSF causes hydrocephalus.

Cerebrospinal fluid secretion

About 500 cm^3 of CSF is secreted per day into a volume of 100–150 cm^3. Choroid plexus epithelium contains a variety of active transport mechanisms. This results in secretion into the CSF of sodium, chloride, and bicarbonate but resorption of potassium, glucose, urea, and a number of neurotransmitter metabolites. The protein concentration of CSF is very much lower than that of blood plasma.

CSF and meningeal functions

The CSF acts as a sink for metabolites that eventually are dumped into the blood via arachnoid villi or choroid plexuses. Mechanical functions of CSF and meninges are to reduce the weight of the brain in the skull, to resist changes in intracranial pressure due to altered brain blood flow and to cushion the brain during violent movements of the head.

Related topic Organization of the central nervous system (A5)

Meninges

The brain and spinal cord are surrounded by three connective tissue membranes, the meninges (*Fig. 1*).

The **pia mater** and **arachnoid mater** are together called the leptomeninges. In the **subarachnoid space** that lies between them run superficial cerebral blood vessels. These are invested by a leptomeningeal coat and suspended in the space

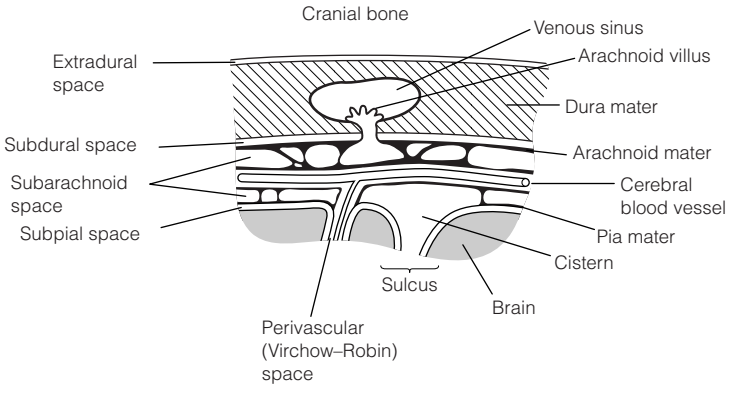


Fig. 1. The meninges.

by trabeculae. The subarachnoid space is filled with cerebrospinal fluid. Branches of the subarachnoid vessels penetrate the brain, becoming surrounded by a cuff of pia mater that extends as far as the capillaries. The **perivascular (Virchow–Robin) space** between the vessel wall and the pia mater is continuous with the subarachnoid space. Here passive exchange of water and solutes across the pia mater keeps the CSF in equilibrium with brain extracellular fluid. At the cerebral capillaries the pia mater is lost and the single layer of capillary endothelial cells, with their basement membrane, are covered by glial cells. Expanded regions of the subarachnoid space are **cisterns**. The lumbar cistern is the target for sampling CSF (**lumbar puncture**) since there is no risk of damage to the cord.

The **dura mater** is a thick, tough, outer layer with **venous sinuses** running through it. Small herniations of the arachnoid mater called **arachnoid villi** (**arachnoid granulations**) protrude through the dura into the venous sinuses. Here bulk flow of CSF into blood occurs via mesothelial tubes in the arachnoid villi that act as valves, closing when the pressure in the venous sinus exceeds that of subarachnoid space to prevent reflux of blood into the CSF.

The **subdural space** is a potential space between the dura mater and the arachnoid mater. It is traversed by cerebral veins entering the venous sinuses in the dura. Traumatic rupture of these vessels as they pass through the space causes **subdural hemorrhage**, which may present clinical problems at any time from the moment of injury to months later. Trauma which shears major vessels going from the dura mater into the cranial bone causes bleeding in the **extradural space** that opens up between meninges and skull. **Extradural hemorrhage** is a life-threatening surgical emergency because it causes brain compression. In the vertebral canal the dura mater forms a loose sheath leaving an **epidural space** between it and the canal wall. Injection of local anesthetics into this space produces **epidural nerve block**.

Cerebrospinal fluid circulation

CSF is actively secreted by choroid plexuses situated in the lateral, third and fourth ventricles (*Fig. 2*). Flow of CSF is from the lateral ventricles through the **foramen of Munro** into the third ventricle, and then through the **aqueduct of Sylvius** into the fourth ventricle. From here it drains via three orifices, a medial **foramen of Magendie** and two lateral **foramina of Lushka**, to enter the subarachnoid space. Here it equilibrates with extracellular fluid in the perivascular spaces. Finally it is dumped into the venous sinuses via the arachnoid villi.

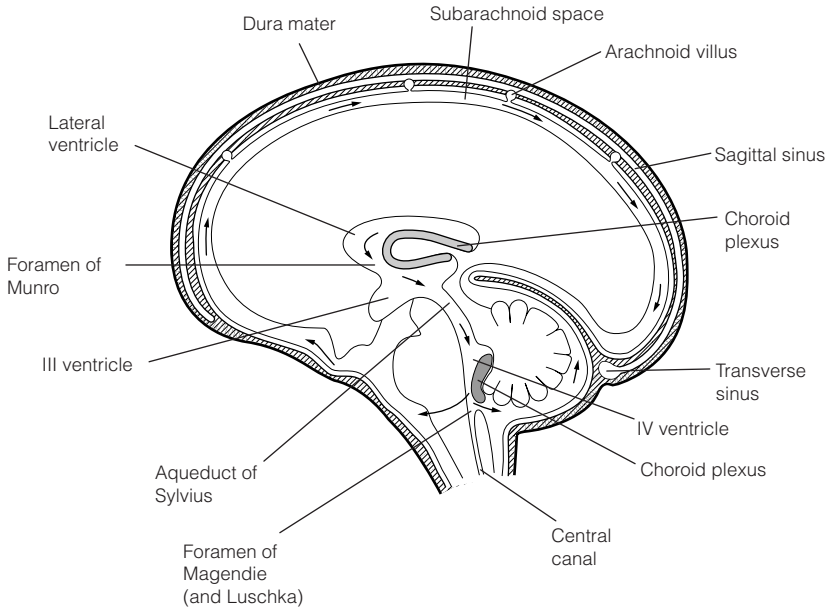


Fig. 2. Cerebrospinal fluid circulation (arrow shows the direction of bulk flow).

Obstruction to the flow of CSF causes **hydrocephalus**, an accumulation of fluid in the cranium. This may increase CSF pressure, distending the ventricles and inflicting damage to the surrounding neural tissue. An obstruction of the ventricular system is **non-communicating hydrocephalus**. It is the result of congenital malformation, scarring or tumors. In **communicating hydrocephalus** there is a failure of CSF flow from the arachnoid villi. This may happen if the concentration of protein in the CSF gets abnormally high, as with some spinal cord tumors, subarachnoid hemorrhage, or meningitis.

CSF secretion

Each **choroid plexus** consists of a cuboidal epithelium derived from the **ependyma (**the lining of the ventricles and spinal cord central canal), covering a core of highly vascular pia mater. In adult humans CSF is secreted at about 500 cm^3 day^{-1} into a steady state volume of 100–150 cm^3. Of this, about 30 cm^3 is in the ventricles and the rest in the subarachnoid space. Cerebrospinal fluid is turned over about every 5–7 hours.

The choroid plexus secretes some substances and absorbs others specifically, most by active transport mechanisms. In this way it acts as a selective interface between blood and CSF, the **blood–CSF barrier**. The result is that by comparison with blood plasma CSF has somewhat higher Na^+, Cl^-, and HCO_3^- concentrations but lower K^+, urea, glucose and amino acid concentrations. Although the protein concentration of CSF is about 1000-fold lower than blood plasma, its higher ionic concentration gives the two fluids the same osmolality.

Some of the mechanisms involved in ion transport across the blood–CSF barrier are shown in *Fig. 3*. Na^+, K^+-ATPase on the apical border of the epithelial cell pumps sodium into the CSF. This generates a sodium gradient that drives two secondary active transport mechanisms bringing Na^+ across the basolateral border; Na^+–H^+ exchange and a Na^+–Cl^- symport. The Cl^- influx in turn drives a Cl^-–HCO_3^- antiport. Bicarbonate brought into the cell in this way is added to that

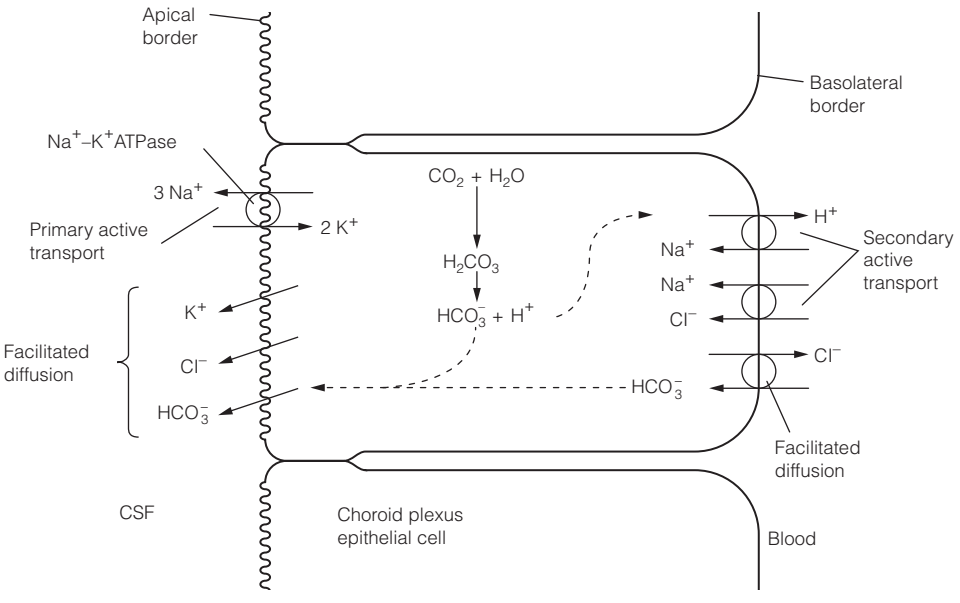


Fig. 3. Aspects of ion transport across the choroid plexus.

formed intracellularly by hydration of CO_2, a reaction greatly accelerated by the high levels of **carbonic anhydrase** present in the choroid plexus. The bicarbonate diffuses via an apical anion transporter into the CSF.

The ability of the choroid plexus to absorb materials from the CSF means that it acts as an excretory organ for the brain. It scavenges choline, dopamine and serotonin metabolites, urea, creatinine and K^+, dumping them into the blood.

CSF and meningeal functions

The functions of the CSF are metabolic and mechanical. By equilibrating with brain extracellular fluid unwanted metabolites are removed to the blood, either via arachnoid villi or choroid plexuses. There are three mechanical effects:

1. Because the subarachnoid space is a fluid-filled compartment in which the brain floats, the effective weight of the brain is reduced from about 1350 g to about 50 g.
2. Adjustments to CSF and meninges prevent changes in intracranial pressure due to alterations in cerebral blood flow. When blood flow increases, CSF is squeezed from ventricles into the subarachnoid space around the spinal cord. Here the dura mater is more elastic and stretches to accommodate the rise in volume. Longer-term increases in intracranial pressure can be offset by a rise in CSF flow into the venous sinuses through the arachnoid villi.
3. The meninges support the brain and the CSF reduces the force with which the brain impacts the inside of the cranium when the head moves.

A8 Blood–brain barrier

Key Notes

Structure of the blood–brain barrier

The blood–brain barrier is formed by capillary endothelial cells, which are coupled by tight junctions. A few regions of the brain, the circumventricular organs, are on the blood side of the blood–brain barrier and are able to secrete substances directly into the blood, or monitor the concentrations of materials in the blood. These regions are isolated from the rest of the brain by specialized ependymal cells, which are coupled together by tight junctions.

Functions of the blood–brain barrier

The blood–brain barrier is a highly selective permeability barrier, which allows the passage of water, some gases, and lipid-soluble molecules by passive diffusion, and contains specific carrier-mediated transporters for the selective transport of molecules crucial to neural function (such as glucose and amino acids). It prevents the entry of circulating neuroactive compounds and is able to exclude lipophilic, potential neurotoxins via P-glycoprotein. Cerebral edema is the accumulation of excess water in the extracellular space of the brain, and results when hypoxia causes the blood–brain barrier to open.

Related topics

Glial cells and myelination (A3)
Organization of the central nervous system (A5)
Posterior pituitary function (L2)

Structure of the blood–brain barrier

The blood–brain barrier governs what is allowed to cross into the brain extracellular fluid from the blood. The physical barrier is provided by brain capillary endothelial cells which are coupled to each other by tight junctions one hundred-fold tighter than is typical for other capillaries. This means that even small ions will not permeate between endothelial cells in brain capillaries. Furthermore brain capillary endothelial cells have a relative lack of two major transport mechanisms possessed by other endothelial cells: **pinocytotic vesicles**, which normally allow the bulk transfer of fluid across the cell; and **receptor-mediated endocytosis**, by which a variety of substrates, for example lipoproteins, are specifically transported. Brain capillaries are entirely covered by the endfeet of astrocytes, which secrete as yet undefined factors which promote the formation of the tight junctions between the endothelial cells (*Fig.1*).

A few regions of the brain, the **circumventricular organs** (CVOs), lie on the blood side of the blood–brain barrier. They are sealed off from the rest of the brain by specialized **ependymal cells** (epithelial cells lining the ventricles) that are coupled together by tight junctions. CVOs, which include the posterior pituitary and choroid plexus, are situated around the ventricles of the brain (see *Fig. 1* of Topic L2). The lack of a blood–brain barrier at the posterior pituitary permits oxytocin and vasopressin to be secreted directly into the systemic circulation. Other CVOs allow the brain to monitor the concentrations of water, ions and selected molecules for homeostatic functions.

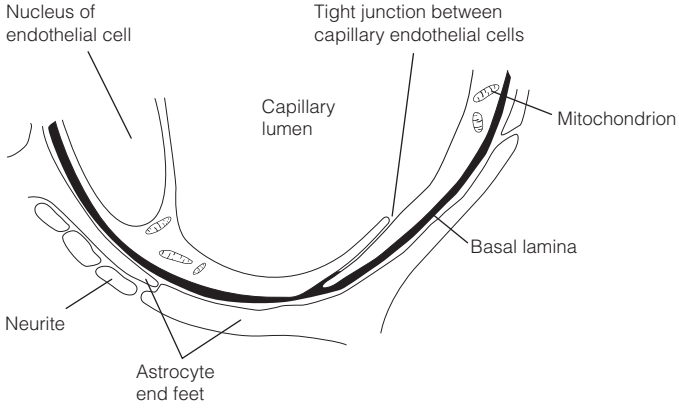


Fig. 1. Structural features of the blood–brain barrier. The barrier is made by the tight junctions between the endothelial cells.

Functions of the blood–brain barrier

The plasma membranes of endothelial cells, like those of any cell, consist of phospholipid bilayers into which are inserted numerous protein species. The lipid component will exclude ions or charged molecules and all but the smallest of polar molecules. Only water, gases which are water- or lipid-soluble (e.g. O_2 or volatile general anesthetics, respectively) and lipophilic molecules (e.g. steroids) will be permeant to any extent. The transport of ions, charged or polar molecules must be by carrier-mediated mechanisms. Many of the proteins in the endothelial cell plasma membrane are transporters or ion channels serving this function.

By means of its selective permeability the blood–brain barrier ensures that crucial molecules, glucose and amino acids, for example, are taken across into the brain. It protects neurons from the actions of neuroactive molecules in the blood, such as circulating catecholamines or glutamate. The blood–brain barrier is also able to actively exclude a wide range of lipophilic compounds that are potentially neurotoxic, many of which are ingested as part of a natural diet. This is achieved by a transport protein, **P-glycoprotein**, expressed in high levels in the plasma membrane of endothelial cells. Lipophilic toxins, which diffuse into the endothelial cell, are rapidly pumped back out into the blood, by P-glycoprotein. Unfortunately, many brain tumors also express P-glycoprotein and so are able to exclude a variety of chemically unrelated chemotherapeutic agents; a phenomenon known as **multi-drug resistance**. It explains why chemotherapy is not very successful in treating brain tumors.

The blood–brain barrier opens in cerebral ischemia causing **cytotoxic cerebral edema**, which is a medical emergency. The lack of oxygen causes a decline in endothelial cell ATP, which compromises the function of the cation pump (see *Instant Notes in Biochemistry*). Consequently, Na^+ accumulates inside the cell, water then enters osmotically and the cell swells. This opens tight junctions, allowing the influx of ions and water into the brain extracellular space.

B1 Resting potentials

Key Notes

Excitable properties

Excitable cells are able to produce action potentials, brief reversals of the electrical potential across their plasma membrane. Excitable cells include neurons and muscle cells.

Intracellular recording

This is a technique for measuring transmembrane potentials. It uses a fine, electrolyte-filled glass microelectrode to impale the cell. The microelectrode output goes to an amplifier and then to a computer for display, storage and analysis.

Resting potentials

The resting potential is the voltage across the plasma membrane of an unstimulated excitable cell. All transmembrane potentials are expressed as inside relative to outside. Resting potentials are inside negative, and range from about –60 mV to –90 mV in neurons. The resting potential is caused largely by the tendency for potassium ions to leak out of the cell, down their concentration gradient, so unmasking a tiny excess of negative charge on the inside of the cell membrane. Other ions (e.g. sodium) make a small contribution to the resting potential. The electrochemical force tending to drive an ion across a membrane is the difference between the resting potential and the equilibrium potential for the ion, at which there is no net flow of ions. Equilibrium potentials can be calculated for individual ions using the Nernst equation.

Related topic Action potentials (B2)

Excitable properties

A modest difference in electrical potential exists across the plasma membrane that surrounds all cells. Cells are said to be **excitable** if, when sufficiently stimulated, they generate **action potentials**. These are rapid, brief, reversals in membrane potential that are actively propagated over the cell surface. Excitable cells include neurons, skeletal, cardiac and smooth muscle cells, and some endocrine cells (for example, the insulin-secreting B cells). The transmembrane potential that exists across an excitable cell when in an unstimulated state is called the **resting potential**.

Intracellular recording

Being able to measure membrane potentials directly is crucial to understanding how excitable cells function. A standard technique for doing this in individual cells is **intracellular recording**.

To record the potential difference across a membrane it is necessary to have two electrodes, one inside the cell, the other outside, both connected to a voltmeter of some description (*Fig. 1*). Because neurons are small the tip of the intracellular electrode impaling the cell needs to be very fine. To achieve this, glass micropipettes are manufactured to have a tip diameter of less than 1 μm. The micropipette is filled with an electrolyte (commonly KCl at a concentration

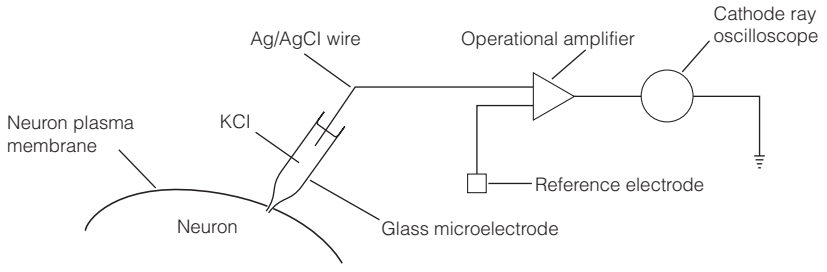


Fig. 1. The circuitry used for intracellular recording.

between 0.15–3 M) to carry the current, so forming the **microelectrode**. Typically transmembrane potentials are less than 0.1 V and so must be amplified with an **operational amplifier**. This has inputs from both the intracellular microelectrode that impales the cell and the **reference** (**bath** or **indifferent**) electrode, which is placed in the solution bathing the cell. If no potential difference exists between the microelectrode and the reference electrode the amplifier output will be zero. If a potential difference exists between the electrodes, however, the amplifier generates a signal, the magnitude of which is proportional to the potential. The output of the amplifier goes to a suitable recording device, traditionally a cathode ray **oscilloscope**, but nowadays it is likely be the analog-to-digital port of a computer running software which emulates an oscilloscope and which allows display, storage and analysis of data.

Resting potentials

Resting potentials (V_m) arise because there is a difference in the concentrations of ions between the inside and outside of the cell and because the cell membrane has different permeabilities for these ions. *Table 1* gives values for concentrations and relative permeabilities for the ions that are most important in relation to the resting potential. The extracellular fluid that bathes cells is essentially a dilute solution of sodium chloride. The intracellular solution, in contrast, has quite a high concentration of potassium ions that are balanced by a variety of anions to which the cell membrane is completely impermeable (although not listed individually in *Table 1* these include organic acids, sulfates, phosphates, some amino acids and some proteins). The cell membrane is permeable to K^+ and because there is a concentration gradient for K^+ across the membrane there is a **diffusional force** acting to drive the K^+ from the inside to the outside of the cell (*Fig. 2*). However, the cell membrane is completely impermeable to the much larger anions, which therefore remain inside the cell. As the potassium diffuses out a potential difference forms across the membrane because some of the intracellular anions are no longer neutralized by K^+. The potential difference now means that an attractive **coulombic (electrostatic) force** is generated which acts to prevent potassium ions diffusing out. At some point the diffusional force

Table 1. Ionic concentrations across mammalian membranes (mmol l^{-1})

Ion	Extracellular fluid	Axoplasm
K^+	2.5	115
Na^+	145	14
Cl^-	90	6

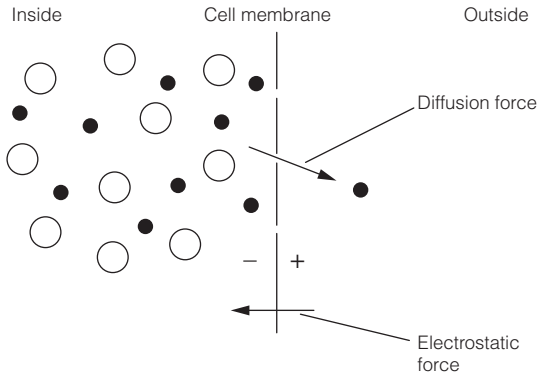


Fig. 2. Illustration of how a potassium equilibrium potential is formed. A small potential exists across the membrane when the diffusional force equals the electrostatic force. Small filled circles represent K^+ ions, large open circles represent anions.

driving K^+ out is exactly balanced by the electrostatic force preventing K^+ leaving.

At this equilibrium a small potential difference exists which is termed an **equilibrium potential**. It is called an equilibrium potential because at this potential there is no *net* flow of K^+ ions across the membrane. In the case where the potential arises as a result of the distribution of diffusible K^+ it is called a potassium equilibrium potential (E_K). Typically nerve cells have potassium equilibrium potentials around –90 mV. Three important points should be noted.

- Transmembrane potentials are always quoted as inside relative to the outside, which is taken to be zero. So, E_K = –90 mV means that the inside of the cell is negative with respect to the outside.
- The number of ions which migrate across the membrane to establish an equilibrium potential is extremely small.
- The potential difference exists only at the plasma membrane, which by storing charge acts as a **capacitor**.

Equilibrium potentials can be calculated using the **Nernst** equation:

$$E = (RT/zF) \ln C_e/C_i$$

where R is the universal gas constant, T is absolute temperature, z is the oxidation state of the ion, F is Faraday's constant and C_e, C_i are extracellular and intracellular concentrations, respectively, of the ion.

The potassium equilibrium potential is close to the resting potential (V_m) for excitable cells. This suggests that V_m arises largely as a result of the distribution of potassium ions across the cell membrane. Neuron resting potentials range between –60 mV and –90 mV. The discrepancy between E_K and V_m arises because ions other than potassium also make a contribution by virtue of their equilibrium potentials. The most important is sodium (E_{Na} = +55 mV) but since the permeability of Na^+ is low compared to K^+ it exerts only a modest influence. The effect of sodium is to drag the resting potential away from E_K towards E_{Na} by an amount that reflects the relative permeabilities of the two ions. The difference between the resting potential and the equilibrium potential for any ion, V_m-E_{ion}, is termed the **ionic driving force** and is a measure of the **electrochemical**

force tending to move the ion across the cell membrane. At rest the driving force for K^+ is quite small but that for Na^+ is high.

In most excitable cells the ionic driving force for chloride ions is close to zero (that is $E_{Cl} = V_m$). This is because Cl^- ions are passively distributed across the membrane according to the resting potential set up by the combined effects of E_K and E_{Na}. The reason for chloride being passively distributed whilst K^+ and Na^+ directly determine the resting potential is because the resting concentration gradients for potassium and sodium are actively maintained by the actions of Na^+/K^+-ATPase (cation pump), whereas there are no active transport mechanisms to maintain a fixed chloride concentration gradient.

B2 Action potentials

Key Notes

Stimulating neurons

Neurons can be stimulated using a stimulator which delivers a current to the cell via a microelectrode. The current usually has a square waveform, the frequency, amplitude and pulse width of which can be varied independently. Inward currents cause neurons to depolarize (i.e. the membrane potential becomes smaller) whereas outward currents cause neurons to hyperpolarize.

Action potentials

An action potential or nerve impulse is a short-lived reversal of the membrane potential. In neurons the spike lasts less than 1 ms and peaks at about +30 mV. The after-hyperpolarization that follows lasts a few milliseconds.

Action potential properties

Action potentials are triggered at the axon hillock and propagated along the axon. They obey the all-or-none rule; a stimulus must be sufficiently large to depolarize a neuron beyond a threshold voltage before it will fire and all action potentials in a given cell are the same size. There is a short delay, the latent period, between the onset of the stimulus and the beginning of the action potential. Neurons become completely inexcitable to further stimulation during the spike and harder to excite during the after-hyperpolarization. These constitute the absolute and relative refractory periods respectively. Refractory periods limit the maximum rate at which neurons can fire, and ensure that action potentials are propagated in only one direction along the axon.

Related topics

Resting potentials (B1)
Voltage-dependent ion channels (B3)
Channel molecular biology (B4)
Action potential conduction (B5)

Stimulating neurons

In vivo, neurons are excited either by the cascade of synaptic inputs onto their dendrites and cell body from other neurons or by receptor potentials generated by sensory organs. Neurophysiologists often stimulate a neuron directly by injecting an electrical current into it via a stimulating microelectrode. The **stimulator** normally delivers a square wave current pulse. Three variables can be altered at will on most stimulators; the **duration** of the pulse, the **amplitude** of the injected current, and the **frequency** of the pulses. The direction of the current (which is defined as the flow of positive charge) determines the response of the neuron. If a small **inward current** is injected into a cell it will become a little more inside positive. This is a decrease in the membrane potential because V_m gets closer to zero and is called a **depolarization**. If, on the other hand, an **outward current** is injected (that is, if current is withdrawn from the cell) then the membrane potential increases; this is called **hyperpolarization**. The sizes and timecourses of depolarizing and hyperpolarizing potentials seen

in nerve cell injected with *small* currents are determined solely by the passive electrical properties of the neuron.

Action potentials

If a sufficiently large inward current is injected into a neuron its membrane potential will depolarize enough to generate an action potential (nerve impulse). This is a brief reversal of the transmembrane potential which is propagated over the surface of the cell. Intracellular recording of a neuronal action potential shows (see *Fig. 1*) that the membrane potential rapidly depolarizes to zero, overshoots to about +30 mV then repolarizes back towards V_m all in less than 1 ms. This constitutes the **spike** of the action potential. Immediately after the spike the neuron membrane hyperpolarizes. This **after-hyperpolarization** decays over a few milliseconds, so the potential returns to its resting value.

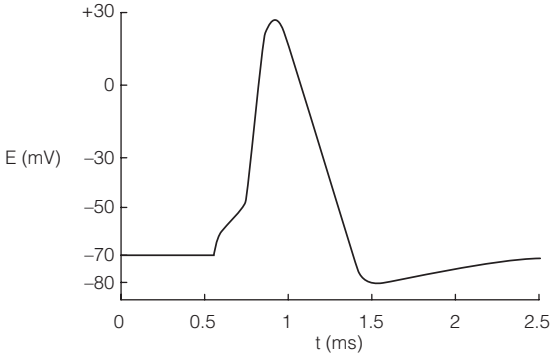


Fig. 1. An intracellular recording of a neuron action potential. The resting membrane potential is –70 mV.

Action potential properties

Action potentials have a number of important properties:

1. Under physiological conditions action potentials are triggered at the axon hillock (initial segment) and are propagated along the axon toward its terminals.
2. They are threshold phenomena. A minimum-size stimulus current is required to produce an action potential. This is the **threshold stimulus**, and is defined as the current that will cause a neuron to fire on 50% of the occasions it is used. Stimuli smaller than this are called **subthreshold**, those which are larger are termed **suprathreshold**. A threshold stimulus causes a neuron to fire an action potential by depolarizing it to a critical **threshold voltage**. The size of the threshold stimulus depends on the size of the neuron. In most neurons the threshold voltage is some 15 mV less than V_m. Action potentials are triggered at the axon hillock because this region of the neuron has the lowest threshold.
3. All action potentials are about the same size (in a given cell) regardless of the amplitude of the stimulus. Fractional action potentials, or superimposing two or more action potentials, are not allowed! The spike of the action potential contains no clue as to the size of the stimulus that produced it. The combined effects of this and property 2 are often paraphrased as the **all-or-none** rule; a neuron either fires completely or not at all.

4. There is a short delay between the onset of the stimulus and the start of the action potential. This is called the **latency** (or **latent period**). The latency gets shorter as the strength of the stimulus increases.
5. During the spike a neuron becomes completely inexcitable. This is the **absolute refractory period**, during which time a nerve cell will not fire again, no matter how large the stimulus. After the spike, while the neuron remains hyperpolarized, the neuron can be excited only by suprathreshold stimuli. This period is the **relative refractory period** and is explained by the fact that when the cell is hyperpolarized bigger stimuli are needed to drive the membrane potential to the threshold voltage. That neurons are temporarily refractory to stimuli during an action potential has three important consequences for nerve cell function. Firstly, it imposes a limit on the maximum rate that a neuron can fire. Secondly, because the region of neuron membrane that has just fired an action potential is inexcitable, the action potential will not reinvade it and so must be propagated forwards. This explains why action potentials are normally propagated in only one direction along an axon (property 1).

All of the above properties can be explained by the behavior of the ion channels responsible for the action potential.

B3 VOLTAGE-DEPENDENT ION CHANNELS

Key Notes

Voltage-dependent ion channels

Voltage-dependent ion channels are transmembrane proteins that are ion selective and voltage sensitive. They are named for the ion species to which they are most permeable. They can exist in at least two interchangeable states, open or closed, depending on the membrane potential across them.

Voltage-dependent sodium channels (VDSCs)

Sodium channels are transmembrane glycoproteins found in most excitable cells. Normally closed, they are opened by depolarization of the membrane potential beyond threshold allowing sodium ions to permeate into the cell. This causes the depolarization phase of the action potential. After about 0.5–1 ms the channels flip into an inactivated state rendering them impermeable to sodium. This, together with the low ionic driving force for sodium at positive membrane potentials, curtails the spike amplitude. Sodium channel inactivation accounts for the absolute refractory period.

Voltage-dependent potassium channels (VDKCs)

The delayed outward rectifying potassium channel is responsible for the downstroke of the action potential spike and the subsequent after-hyperpolarization. These transmembrane channels are activated by depolarization permitting potassium ions to flow out of the cell, carrying the membrane potential to increasingly negative values. The after-hyperpolarization is responsible for the relative refractory period.

Voltage clamping

Voltage clamping is a technique that allows the current that flows across nerve cell membranes to be measured. It does this by holding the membrane potential constant. It provided the crucial evidence that action potentials are caused by an early inward current due to sodium and a late outward current due to potassium. These currents can be blocked independently by tetrodotoxin and tetraethylammonium respectively.

Related topics

Resting potentials (B1)
Action potentials (B2)
Channel molecular biology (B4)

Voltage-dependent ion channels

Neurons are excitable because they have in their plasma membrane a class of transmembrane proteins called voltage-dependent ion channels. They are so called because they have two properties, **ion selectivity** and **voltage sensitivity**. Ion channels will allow only certain small ions to flow through them. Voltage-dependent channels are classified according to the ion (Na^+, K^+ or Ca^{2+}) to which they are preferentially permeable. A large number of different voltage-dependent ion channels have been identified to date. For example, nine distinct voltage-dependent sodium channels have been discovered.

Voltage-dependent ion channels can exist in at least two interchangeable states: **open** (**activated**), when they allow ions to flow through them; or **closed**, when they do not. Whether they are open or closed depends on the voltage across them.

Voltage-dependent sodium channels (VDSCs)

Voltage-dependent sodium channels (VDSCs) are large glycoproteins that span the full thickness of the plasma membrane of most excitable cells. At the resting potential they are closed. If a region of membrane is depolarized by only a modest amount (e.g. 10 mV) they remain closed (*Fig. 1a*). However, if the membrane is depolarized to the threshold voltage or beyond VDSCs change shape so that the channels open, permitting Na^+ ions to flow through. Channel opening, **activation** (*Fig. 1b*), is an extremely fast event. The transition from the closed to the open state takes only about 10 μs. During an action potential in a neuron a given sodium channel will remain open for about 0.5–1 ms. In this time about 6000 Na^+ ions will flow through the channel. The combined effect of sodium influx through a few hundred VDSCs will produce the upstroke of the spike that is the depolarizing phase of the action potential. Only a few voltage-dependent sodium channels need be driven beyond threshold initially to trigger an action potential since the localized influx of Na^+ causes a depolarization which drives other channels to open and so on. It is this **self-regeneration** property of the action potential that causes the explosive rise in Na^+ permeability.

At the top of the spike the increase in sodium permeability is halted for three reasons:

- All of the available VDSCs in the active region of membrane have opened.
- The ionic driving force of sodium gets less as the membrane depolarizes towards the sodium equilibrium potential.
- The voltage-dependent sodium channels flip into an **inactivated** state. During this state the channel does not permit the flow of any ions, but this is *not* the same as the closed state because while in the inactivated state a sodium channel *cannot* be made to open. In addition to curtailing the spike, it is this inactivation of VDSCs that is responsible for the absolute refractory period. The inactivation wears off after a couple of milliseconds or so and the sodium channel goes into the closed state from which it can be reactivated by subsequent depolarization.

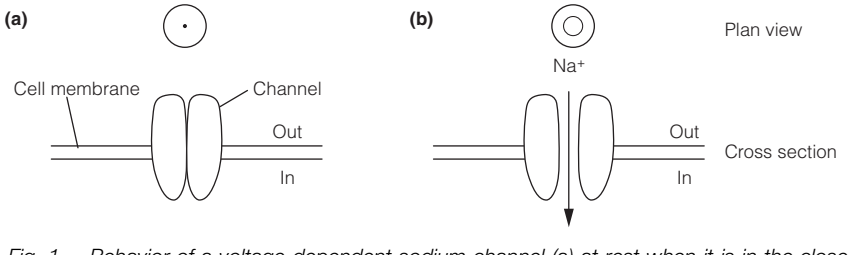


Fig. 1. Behavior of a voltage-dependent sodium channel (a) at rest when it is in the closed state and (b) during the spike of an action potential, when it is activated.

Voltage-dependent potassium channels (VDKCs)

Excitable cells have a great variety of distinct types of potassium channel each with their own particular properties. One type, the **delayed outward rectifier** is involved in the repolarization phase of the action potential. These channels are transmembrane glycoproteins with a molecular structure related to that of voltage-dependent sodium channels. Like them, they are opened by depolarization. This allows potassium ions to flow down their concentration gradient out of the cell. The inside of the neuron becomes less positive, that is, it repolarizes. This accounts for the downstroke of the spike of the action potential. At the bottom of the downstroke most of the VDSCs are inactivated so there is no flow of sodium into the cell. Delayed outward rectifying VDKCs either do not inactivate or inactivate much more slowly than VDSCs (depending on species) so immediately after the spike the neuron remains highly permeable to K^+ whilst being impermeable to Na^+. The consequence is that for a few milliseconds after the spike, potassium ions continue to leave the cell carrying the membrane potential more negative than V_m. This is the **after-hyperpolarization** phase of the action potential. It accounts for the relative refractory period, because during this time, for a stimulus to excite the cell to threshold it must make the neuron depolarize from a more inside-negative state. Finally the membrane potential returns to the resting potential as the potassium channels flip into their closed or inactivated state in a time-dependent fashion. The changes in ion conductance during the action potential are summarized in *Fig. 2.*

A couple of points follow from the above:

- The extent of the hyperpolarization is determined by the potassium equilibrium potential. When the efflux of K^+ is sufficient to bring the membrane potential to E_K the ionic driving force on the potassium ions will be zero. No more K^+ will leave.
- Since depolarization of the neuron to threshold causes the opening of VDSCs (allowing Na^+ in) *and* opening of VDKCs (allowing K^+ out) how does an action potential arise? The reason is that the sodium channels respond to the depolarization earlier than the potassium channels so the increase in sodium permeability precedes the increase in potassium permeability.

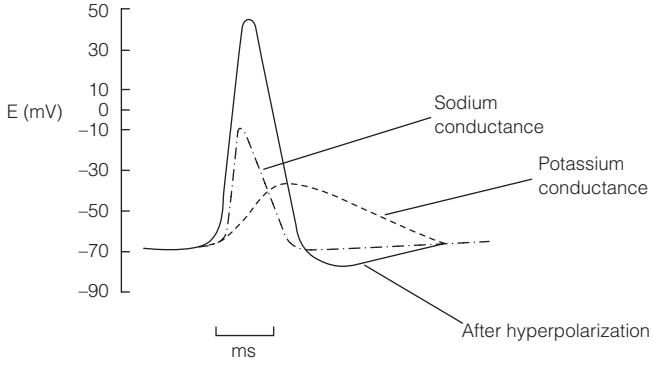


Fig. 2. Changes in ion conductance during the action potential.

Voltage clamping

The ion fluxes that underlie neuron action potentials were discovered in the early 1950s using a technique called **voltage clamping** which remains a key technique in neurophysiology. It measures the currents that flow across an

excitable cell membrane at a fixed potential. Measuring currents is important because it provides information on which ions might be responsible for changes in membrane potential. Current (I) cannot directly be found from potential (V); it requires, in addition, the membrane resistance (R). Only if *both* V and R are known can I be calculated from Ohm's law, $V = IR$. Voltage clamping circumvents this problem by measuring the transmembrane potential and having a feedback amplifier in the circuit which injects into the cell the current which is needed to keep the membrane potential constant; i.e. the voltage is clamped. The current that must be injected by the circuit to keep the voltage fixed has to be the same size as the current flowing through the ion channels which would normally cause the potential to change. The voltage at which the membrane is clamped is called the **command voltage**. By examining the currents that flow across a membrane over a range of command voltages it is possible to determine which ions carry the currents.

The use of voltage clamping is illustrated in *Fig. 3*, which shows an experiment in which the giant axon of a squid is initially clamped at its resting potential of –60 mV and the command voltage is then switched to zero. This first gives rise to a **capacitance current** as the change in voltage alters the amount of charge stored on the nerve cell membrane.

After the capacitance current comes an early inward current followed by a late outward current. These are the currents that normally flow during an action potential. When the squid axon is bathed in a solution that contains no sodium the early inward current is abolished. This shows that the early inward current is carried by Na^+. The same result is seen if an axon immersed in normal seawater is poisoned with the neurotoxin **tetrodotoxin** (TTX). By binding to the external mouth of the VDSC, TTX prevents Na^+ from permeating through the channel. TTX added to any nerve preparation abolishes action potentials. Similarly, **tetraethylammonium** (TEA), which blocks VDKCs, when added to the bathing medium abolishes the late outward current, showing that it is carried by K^+.

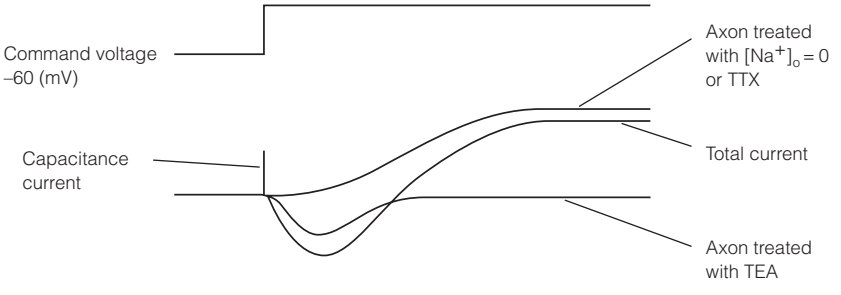


Fig. 3. A voltage clamping experiment to show the currents that flow during an action potential. Outward currents are upward deflections, inward currents downward deflections. Initially the neuron is clamped at –60 mV, and the step to 0 mV generates first a capacitance current, then the ionic currents. The experiment is done with the axon under three different conditions. See text for details. TTX, Tetrodotoxin; TEA, tetraethylammonium.

B4 Channel molecular biology

Key Notes

Patch clamping

Patch clamping measures the currents that flow through single-ion channels. It depends upon electrically isolating the tiny patch of membrane that lies under the tip of a microelectrode. In the cell-attached mode the patch remains on the cell and single channel currents are recorded. Rupturing the patch gives the whole cell mode, which allows recording of the macroscopic currents that flow through the plasma membrane of the cell. Alternatively the patch can be completely removed to give two other single-channel configurations. In outside-out mode the effects of channel ligands applied to the bath can be studied. In inside-out mode agents can be added to the bath to investigate the role of second messengers in modulating channels.

Protein chemistry

Cloning and sequencing the DNA coding for a protein gives its amino acid sequence, from which much about its secondary structure can be inferred. Hydropathicity plots and the location of consensus sequences for glycosylation and phosphorylation sites provide insights into how a protein is oriented in the membrane. Site-directed mutagenesis can reveal the functional importance of specific amino acids. Provided a protein can be crystallized, X-ray crystallography will give its 3-D structure. All these techniques have been used to fathom ion channel function.

Structure of voltage-dependent channels

Ion channels consist of four subunits, each with six α-helical transmembrane segments (S1–S6) arranged around a central pore. Segments S5 and S6 and the H5 loop form the pore. The positively charged S4 segments lie on the outside of the channel and act as voltage sensors. On depolarization they swing towards the extracellular side of the membrane, opening the intracellular mouth of the channel by pulling on S5. Three distinct mechanisms can close potassium channels. In N-inactivation a peptide blocks the internal mouth of the channel.

Related topics

Voltage-dependent ion channels (B3)
Calcium channels (C6)

Patch clamping

Patch clamping is an *in vitro* technique that makes it possible to study the electrophysiology of single-ion channels. It works by forming a very high electrical resistance seal between a glass micropipette and the surface of a cell. Only currents flowing through the patch under the electrode will be recorded. This permits the extremely small currents that flow through single-ion channels (about 1 pA) to be measured. The electronics allows the voltage of the patch to be clamped so voltage-clamping experiments can be performed (*Fig.1a*).

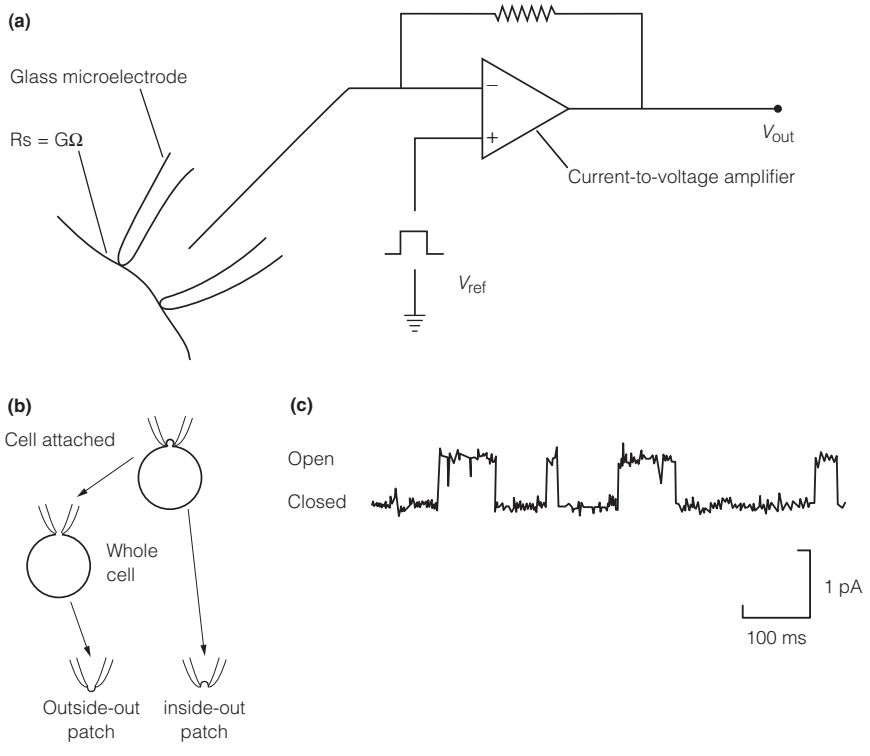


Fig. 1. Patch clamping. (a) The circuitry, V_{ref} is used to clamp the potential of the neuron membrane, the method relies on forming a high resistance seal, R_s, between the pipette tip and the cell membrane. (b) Patch clamp configurations. (c) Single channel currents through a $GABA_A$ receptor.

There are several configurations of patch clamping, each useful for particular types of experiment (*Fig. 1b*).

1. **Cell attached**. This is used for measuring single-channel currents in intact cells. Second messenger-induced modifications of the patched channels can be investigated in response to bathing the cell with specific agents, e.g. neurotransmitters.
2. **Whole cell**. In this configuration the patch under the microelectrode is ruptured. The current flowing through the electrode represents the sum of all the currents flowing through individual ion channels on the cell surface. Hence whole cell patching measures **macroscopic currents**.
3. **Outside-out**. This is one of two patch clamp modes used to study single-channel currents in which the patch is removed from the cell but remains sealed to the pipette tip. This configuration is used to study the effects of ligands such as neurotransmitters, hormones or externally acting drugs on channels. These ligands are added to the bath because bath solutions can be changed much more easily and rapidly than the pipette solution. This has obvious advantages for performing complicated experiments such as investigating dose–response relationships.
4. **Inside-out**. This is the second of the patch-only configurations, usually used for the detailed examination of second messengers since these can be applied directly to the inside face of the membrane via the bath solution.

A representative patch clamp recording is shown in *Fig. 1c*. Each of the square wave events is the opening of a single channel. The height of the wave is the **unitary channel current**, its length the time for which it is open. Statistical analysis of large numbers of opening events provides estimates for parameters such as **mean channel open time**, and allows models of channel kinetics to be tested. Such studies are very useful in fathoming out how neuroactive drugs act at the molecular level.

Protein chemistry

By cloning and sequencing the DNA that codes for a protein it is possible to deduce its primary amino acid sequence. This provides clues as to the secondary structure of the protein such as the presence of α-helical or β-pleated sheet regions. **Hydropathicity plots**, which represent how hydrophobic or hydrophilic particular amino acid stretches are, indicate which regions of the protein might be located in the plasma membrane. Recognizing consensus sequences for glycosylation suggest which parts of a molecule might be extracellular, similarly consensus sequences for phosphorylation imply an intracellular location. Information of this nature provides evidence for how a given protein is orientated in the membrane.

Site-directed mutagenesis, in which the DNA coding for a protein is altered in precise ways using genetic engineering techniques, and the mutated protein subsequently expressed in some convenient way, can be used to investigate the functions of particular amino acids in a channel.

The three-dimensional structure of ion channels in both closed and open states can be deduced by X-ray crystallography from which details of the voltage sensor and exactly how the channel opens can be worked out. A major difficulty in this work is being able to crystallize the proteins. However this has been successfully done with potassium channels from heat-loving bacteria.

Structure of voltage-dependent channels

There are numerous types of voltage-dependent ion channel, structures for many of which have been deduced using the technologies outlined above. Sodium channels and several potassium channels, including the delayed outward rectifier responsible for the downstroke of the action potential, share a common structure. The basic subunit has six highly hydrophobic segments (S1–S6), each about 20 amino acids long, that are thought to be α-helices spanning the membrane (*Fig. 2*). In the case of potassium channels each subunit is a separate molecule and a functioning channel is a tetrameric homo-oligomer; i.e. it consists of four similar subunits arranged around a central pore. In sodium channels the four subunits are linked by cytoplasmic loops to form a single huge molecule, but its tertiary structure is thought to resemble a potassium channel.

Segments 5 and 6 and the H5 loop between them contribute to the pore. This H5 loop is thought to line the channel pore because it contains amino acids that are crucial for conferring ion selectivity. The extracellular throat is lined with oxygen atoms that are crucial for ion permeation.

One of the most striking features of voltage-dependent ion channels is the S4 segment (*Fig. 3*). It is highly conserved between channels, suggesting it has hardly altered during evolution and so serves a critical function. Although mostly hydrophobic, its N-terminal end has several positively charged lysine or arginine residues. Site-directed mutagenesis shows that S4 is required for activation of the channel, so it is part of the voltage sensor of the channel. X-ray crystallography shows the S4 segment lies like a 'paddle' at the periphery of the channel, i.e. at the protein–lipid interface. It is attached to the rest of the channel

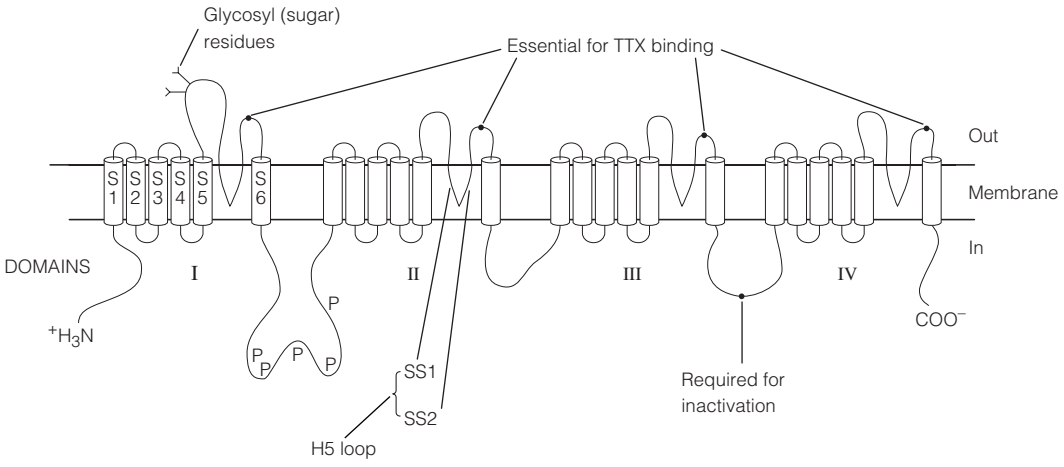


Fig. 2. A cartoon depicting the secondary structure of a VDSC. Segments S1–S6 are labeled only in domain I. P, consensus sequences for phosphorylation.

at each end by flexible regions, an S3 loop and an S4–S5 linker. In response to depolarization the S4 paddle swings about 2 nm through the membrane – towards the extracellular side – as its positive charges are repelled. Acidic amino acids on the adjacent S2 help to stabilize these positive charges as they move. The large displacement of the paddle pulls the S4–S5 linker, opening the intracellular throat of the pore.

Three mechanisms to close potassium channels have been proposed. One, **N-inactivation** occurs by the so called **ball and chain** mechanism. This involves a cluster of amino acids (the 'ball') swinging up to block the internal mouth of the channel by interacting with an amino acid at the internal tip of the H5 loop. In *Drosophila* channels this peptide is the N-terminus of the standard (α) subunit. In mammals the ball is the N-terminus of a separate β-subunit.

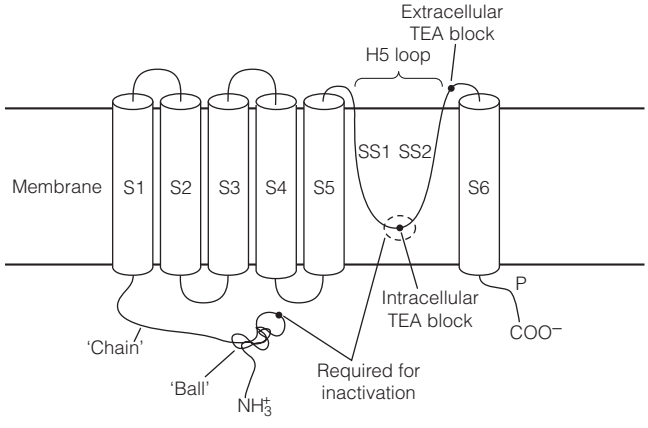


Fig. 3. A cartoon depicting the secondary structure of an outward rectifying potassium channel subunit. Four subunits assemble to form a functioning channel.

B5 Action potential conduction

Key Notes

Propagation of action potentials

Action potentials are generated at the axon hillock (spike initiation zone) and propagate actively, at constant velocity, and without loss of amplitude, down the axon. Because the active zone, the region of the axon at which the action potential sits at a given instant, bears different charges to the axon at rest ahead of it, local circuit currents flow which depolarize the adjacent upstream membrane and so the action potential advances. Local circuit currents also spread backwards but do not allow the action potential to propagate in this direction because the membrane there is refractory.

Conduction velocity in non-myelinated axons

In non-myelinated axons the speed of conduction is between 0.5 and 2 m s^{-1}. The velocity is proportional to the square root of the axon diameter.

Conduction velocity in myelinated axons

Myelination produces dramatic increases in conduction speed for only modest increases in the overall diameter of axons. Myelinated axons conduct faster because local circuit currents flow around the electrically insulating myelin sheath so that only the axon membrane at the node of Ranvier needs to be depolarized to generate an action potential. The action potential appears to jump from one node to the next. The conduction velocity is proportional to axon diameter and varies from 7 to 100 m s^{-1}.

Related topics Glial cells and myelination (A3) Action potentials (B2)

Propagation of action potentials

In a neuron, action potentials are initiated at the axon hillock because this region has the greatest density of VDSCs and so the lowest threshold for excitation. For this reason the axon hillock is sometimes referred to as the **spike initiation zone.** Once generated action potentials are actively propagated (conducted) with constant velocity down the axon without loss of amplitude. Thus, action potentials are undiminished in size even when conducted along peripheral axons that in humans may be up to one meter long. This is one of the features that makes action potentials reliable signals for information transmission. The details of conduction are a little different depending on whether the neuron is myelinated or not.

In non-myelinated neurons conduction works as follows (see *Fig. 1*). The region of an axon invaded by an action potential at a given time is called the **active zone**. It is a few centimeters long. The part of the active zone occupied by the overshoot of the spike will be inside positive. Far away from the active zone, ahead of the oncoming action potential or behind it, the membrane potential will be inside negative. The consequence of this is that a potential difference

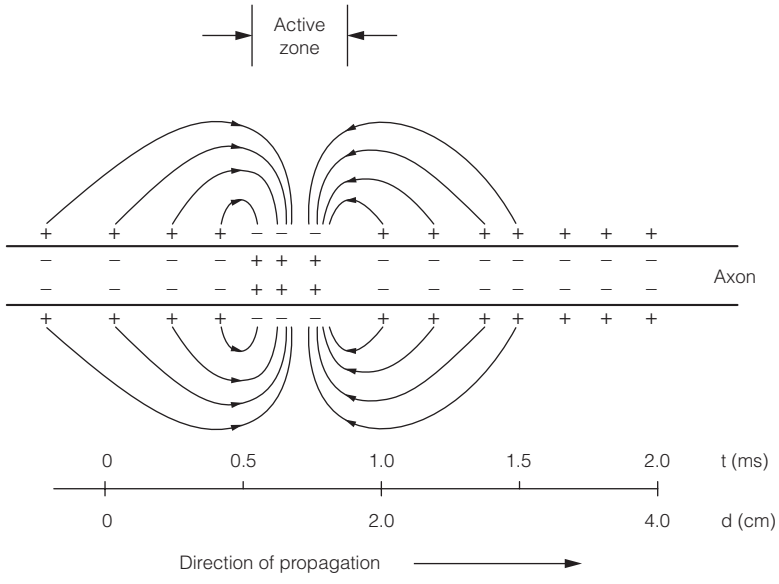


Fig. 1. Local circuit currents involved in the conduction of an action potential. For clarity currents inside the axon are omitted. The action potential is depicted as travelling from left to right along the axon, and the leading edge of the spike (active zone) is 2 cm from the origin (lower scale) after 1 ms (upper scale). t, time; d, distance.

exists between different regions of the external surface of the axon; the outside of the active zone is more negative than its surroundings. A similar situation exists on the inside surface of the membrane except that here the active zone is more positive than its surroundings. These differences in potential cause currents to flow passively over the axon membrane. By convention, current flows in the positive to negative direction and this is depicted by the arrows in *Fig. 1*.

Current flow across the external surface of the axon is from the regions ahead and behind the action potential into the active zone. These currents are called **local circuit currents**. Just ahead of the axon potential the local circuit currents drain positive charge from the external surface of the axon and simultaneously dump positive charge on the inside of the axon membrane. The net effect is to depolarize the axon immediately in front of the action potential. When this depolarization becomes suprathreshold, VDSCs in this region activate and the action potential advances. Of course local circuit currents flow in the same way along the axon behind the action potential but this region is refractory (VDSCs are inactivated and the membrane is hyperpolarized) and so the currents do not excite here. This explains why action potentials propagate physiologically only in one direction.

Conduction velocity in non-myelinated axons

The conduction velocity, θ, the speed with which nerve impulses are propagated, is quite slow in non-myelinated axons. It varies between 0.5 and 2 m s^{-1} depending on the diameter of the axon. Small axons offer a higher resistance to the flow of currents through their cores than large ones, just as thin wires have a higher electrical resistance than thick ones. So, local circuit currents in the

axoplasm of small axons spread less well than in larger axons and this is the reason for the slower speed. Very roughly the relationship is:

$$\theta = ka^{\frac{1}{2}}$$

where a is the axon diameter, k is a constant which depends on the internal resistance of the axon and its membrane capacitance.

Conduction velocity in myelinated axons

A large number of neurons in the vertebrate nervous system, certainly the majority of those in the peripheral nervous system, have myelinated axons. The function of the myelin sheath is to increase the conduction velocity substantially with relatively little increase in total axon diameter. The evolution of myelination has enabled vertebrates to have a large number of rapidly conducting axons without taking up too much cable space.

Because the myelin sheath consists of plasma membrane it has a high content of phospholipid with a high electrical resistance. Local circuit currents are forced to take paths of lesser resistance through the electrolyte solution around the sheath. The effect is that local circuits are established, not between adjacent regions of membrane as they are in non-myelinated axons, but between adjacent nodes of Ranvier, which are relatively far apart. Local circuit currents ahead of an action potential arriving at the next downstream node cause it to depolarize beyond the threshold and trigger an action potential. In this manner action potentials appear to jump from node to node, a mechanism called **saltatory conduction.** The density of VDSCs is about 100-fold greater at nodes than in non-myelinated axon membrane and the node threshold is consequently much lower. This greatly reduces the risk of nodes not firing in response to local circuit currents weakened by the long distances they must spread.

Conduction velocity (θ) is faster in myelinated than non-myelinated axons for two reasons:

1. The presence of a myelin sheath is functionally equivalent to increasing the thickness of the axonal membrane about 100-fold. This greatly reduces the amount of charge stored across the membrane, which means that much less time is taken to depolarize it. More technically put, myelination reduces the membrane capacitance (c_m), because the capacitance of a parallel plate conductor is inversely proportional to the thickness of the insulator, and since $\theta \propto 1/ c_m$, the lower the capacitance the greater the speed of action potential propagation.
2. The time taken up for VDSCs to respond to depolarization is less, because only channels at nodes have to be activated. In a non-myelinated axon each little region of membrane has to be depolarized and respond in succession. For a myelinated axon however, only the node membrane needs to be excited.

Conduction velocities of myelinated axons vary from about 7 to 100 m s^{-1}. As with non-myelinated axons, velocity depends on diameter, a, but the relationship is even simpler:

$$\theta = ka$$

C1 Morphology of chemical synapses

Key Notes

Synapse location

Synapses may be electrical or chemical. Chemical synapses can be classified according to where they are located on the receiving neuron. Axodendritic synapses are made on dendrites, axosomatic on the cell body and axoaxonal on the axon. The majority of synapses are axodendritic.

Synapse structure

An axodendritic synapse is formed between an axon terminal presynaptically and a dendrite postsynaptically. The synaptic cleft between these elements is 30 nm wide. The axon terminal contains mitochondria, spherical synaptic vesicles and dense projections on the presynaptic membrane. The cleft contains proteins which link pre- and postsynaptic membranes. The postsynaptic membrane is thickened to form the postsynaptic density.

Synapse diversity

The majority of cortical synapses fall into just two types. Type I are the axodendritic synapses described above and are usually excitatory. Type II synapses have less-well-developed dense projections, cleft material and postsynaptic density. They have ovoidal vesicles, an axosomatic location and are usually inhibitory. Synapses that secrete catecholamines or peptides have large dense-core vesicles; some of these have little discernible pre- or postsynaptic specialization and a wide cleft. Many synapses contain both small clear vesicles and large dense-core vesicles, evidence that many neurons secrete more than one transmitter.

Related topic

Overview of synaptic function (C2)

Synapse location

Signaling between nerve cells occurs via synapses. There are two types of synapse, **electrical** and **chemical**. Chemical synapses far outnumber electrical ones. A chemical synapse is formed from the close approximation of an axon terminal, which is the presynaptic component, with the postsynaptic cell. The gap between the presynaptic terminal and postsynaptic cell, the **synaptic cleft**, varies between 20 nm and 500 nm depending on the type of synapse. Synapses can be classified on the basis of where on the postsynaptic cell they are located. The majority are on dendrites and so are called **axodendritic** synapses. On spiny dendrites each spine is the target of an axon terminal. Particularly powerful synapses are made between axons and the cell body of a postsynaptic cell. These are called **axosomatic** synapses, named for the alternative term for the nerve cell body, the soma. Synapses between axon terminals and *axons* of postsynaptic neurons are said to be **axoaxonal**.

Synapse structure

Electron microscopy must be used to resolve synapses because of their small size. This has revealed numerous morphologically distinct types of synapse, but all share common features. *Fig. 1* shows a typical axodendritic synapse. The **small clear synaptic vesicles** (SSVs) which store neurotransmitter, are spherical with a diameter of about 50 nm, and are scattered throughout the terminal apparently in close association with microtubules which might be involved in transporting them to the presynaptic membrane. The presynaptic membrane is thickened and may show inwardly directed **dense projections** which are involved in the docking of synaptic vesicles at the **active zone**, the region from which transmitter release occurs.

The synaptic cleft of axodendritic synapses is 30 nm wide and contains protein filaments that stretch transversely from the pre- to the postsynaptic side that may serve to keep the two membranes in close apposition.

The cell membrane of the dendrite in the region of the synapse appears thickened to form the **postsynaptic density**. This is formed by the accumulation of proteins, receptors, enzymes and the like, involved in the postsynaptic cells' response to transmitter.

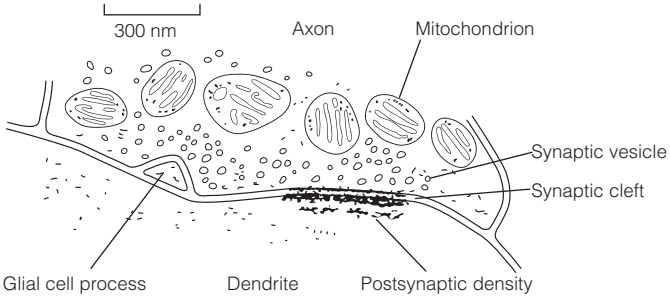


Fig. 1. The structure of a chemical (axodendritic) synapse.

Synapse diversity

Studies of cerebral cortex and cerebellar cortex have revealed that the majority of synapses fall into one of two types. **Type I** are those described above. **Type II** synapses have little or no postsynaptic dense projections, poorly defined material in a cleft only 20 nm across and a thin postsynaptic density. These synapses are frequently axosomatic. Type II synapses contain ovoidal-shaped vesicles. Physiological studies of these two types of synapse show that type I are generally excitatory whereas type II are usually inhibitory.

Although most synapses contain small clear synaptic vesicles, some have spherical vesicles with electron-dense centres called **large dense-core vesicles** (LDCVs) (see *Fig. 2*) which fall into two populations according to size; 40–60 nm vesicles are found in neurons that release catecholamine transmitters, whereas 120–200 nm vesicles are associated with neurosecretory neurons that liberate peptide hormones in the posterior pituitary. Some synapses lack obvious specialized contact zones on both pre- and postsynaptic sides and have extremely wide synaptic clefts; 100–500 nm. These often secrete one or other of the catecholamines, have dense-core vesicles and can be found both in the CNS and in the peripheral nervous system.

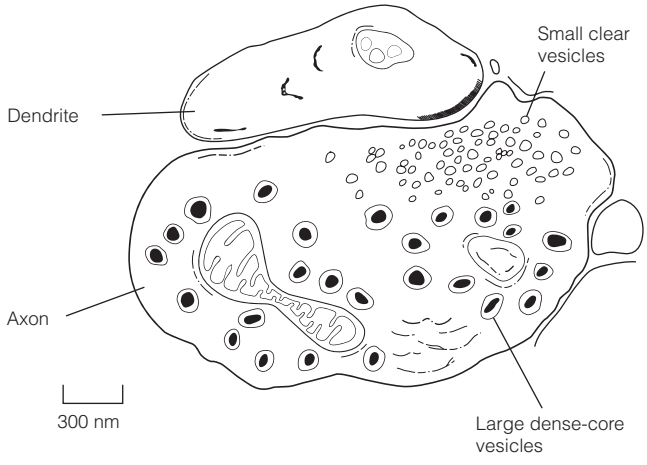


Fig. 2. A type I synapse containing both small clear vesicles and large dense-core vesicles. From Revest, P.A. and Longstaff, A. (1998) Molecular Neuroscience *© BIOS Scientific Publishers Ltd, Oxford.*

Many synapses contain more than one population of vesicles. Commonly, SSVs are found together with LDCVs. This is clear structural evidence that many neurons secrete more than one transmitter.

Synaptic diversity is greater than implied above. In some specialized regions of the nervous system synapses can be found which are quite distinct from the typical, for example, the triad ribbon synapse in the retina. These exceptions will be described later where appropriate.

C2 Overview of synaptic function

Key Notes

Electrical transmission

Electrical synapses are formed by arrays of ion channels, connexons, at gap junctions. Here small ions flow between cells so that they are electrically coupled. Via gap junctions action potentials can spread rapidly between cells without distortion.

Chemical transmission

Neurotransmitter release from the nerve terminal following the arrival of an action potential is triggered by the influx of calcium through voltage-dependent calcium channels. After crossing the cleft, transmitter binds to postsynaptic receptors. These are either ligand-gated ion channels or metabotropic receptors coupled to second messenger systems. Receptor activation either increases or decreases the chance that the postsynaptic cell will fire, responses described as excitatory or inhibitory respectively. Transmission mediated by ionotropic receptors is fast, whereas that mediated by metabotropic receptors is slow. Synapses may secrete more than one transmitter.

Related topics

Morphology of chemical synapses (C1)
Postsynaptic events (C3)
Neurotransmitter release (C5)
Calcium channels (C6)

Electrical transmission

Electrical transmission is mediated by electrical synapses, which are **gap junctions** between adjacent neurons. Gap junctions are arrays of paired hexameric ion channels called **connexons** (*Fig. 1a*). The channel pores are 2–3 nm in diameter, allowing ions and small molecules to permeate between neighboring neurons. By electrically coupling neurons, gap junctions allow any potentials, e.g. action potentials, to spread between cells. Key features of electrical transmission are that it is extremely rapid, it is high fidelity (signals are transmitted with no distortion), and it works in both directions. Gap junctions between cells may close. Each connexon is made up of six subunits called **connexins**. In response to specific chemical signals, such as a rise in intracellular Ca^{2+} concentration, the connexins rotate – rather like the iris diaphragm of a camera – to close the central pore (*Fig. 1b*). Electrical synapses form only a small proportion of all synapses in adults, but are more numerous during development.

Chemical transmission

The vast majority of synapses are chemical. At most central synapses chemical neurotransmission happens in the following way. The arrival of an action potential at the axon terminal may result in the release of neurotransmitter from a single presynaptic vesicle. Neurotransmitter release requires a rise in intracellular Ca^{2+} brought about by calcium entry into the axon terminal via **voltage-dependent calcium channels**. The transmitter diffuses across the synaptic cleft and binds to specific receptors on the postsynaptic membrane. Binding of the

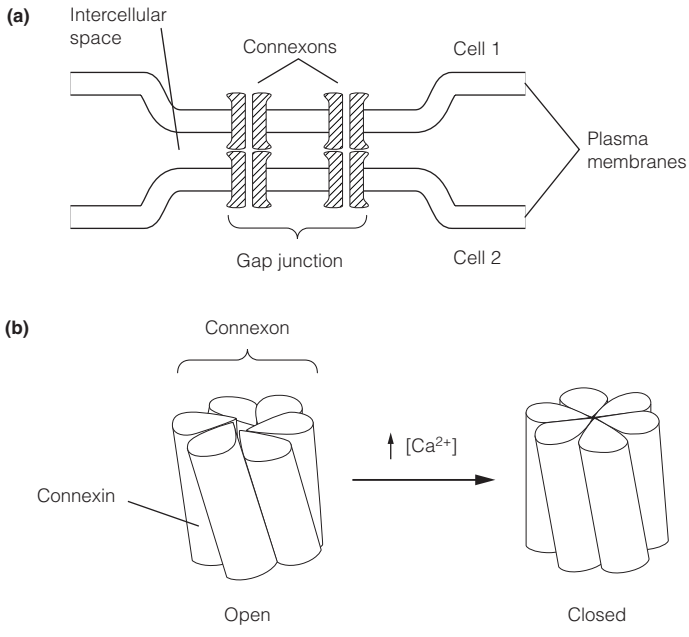


Fig. 1. (a) Gap junction. (b) Change in configuration of the connexons to close a gap junction.

transmitter causes a change in the conformation of the receptor. What happens next depends on the receptor, but the overall result is to change the postsynaptic membrane permeability to specific ions.

Neurotransmitter receptors come in two **superfamilies.** The **ligand-gated ion channel receptors**, or **ionotropic receptors**, have ion-selective channels as part of the receptor. Binding of the transmitter to the receptor opens the channel, directly increasing its permeability. The second superfamily is the **G-protein-linked receptors**, also referred to as **metabotropic receptors.** Binding of transmitter to these receptors activates their associated G-proteins that are capable of diverse and remote effects on membrane permeability, excitability and metabolism. G-proteins can influence permeability either by binding ion channels directly or by modifying the activity of second messenger system enzymes, which phosphorylate ion channels, thereby altering their permeability.

The changes in membrane properties brought about by a transmitter can have essentially one of two effects. It may increase the probability that the postsynaptic neuron fires action potentials, in which case the response is **excitatory**. If the effect is to decrease the probability that the postsynaptic cell might fire, the response is **inhibitory**. Although neuroscientists often describe particular transmitters as excitatory or inhibitory, these actions should more properly be attributed to a *combination* of transmitter and receptor. It is commonly the case that a transmitter is excitatory at one of its receptors, but inhibitory at another. Any given synapse can be described as excitatory or inhibitory.

Some 30 molecules have been unambiguously identified as neurotransmitters and many more are candidates. In general they fall into two groups. The classical transmitters are amino acids or amines. Quantitatively by far the most important are glutamate, which is almost invariably excitatory, and γ-aminobutyrate (GABA), which is usually inhibitory. This group also includes acetylcholine, the

Table 1. Key central nervous system neurotransmitters

Classical	Amino acids	Glutamate
		Aspartate
		γ-aminobutyrate
		Glycine
	(Mono)amines	Acetylcholine
		Dopamine } catecholamines
		Norepinephrine } catecholamines
		Epinephrine } catecholamines
		Serotonin (5-hydroxytryptamine) indolamine
Peptides	Opioids	Dynorphins
		Endorphins
		Enkephalins
	Tachykinins	Substance P
	Hormones	Cholecystokinin
		Somatostatin

catecholamines such as dopamine and noradrenaline (norepinephrine), and the indoleamine, serotonin. The second, larger, group is an eclectic group of peptides which includes the opioids (such as endorphins) and the tachykinins (e.g. substance P). See *Table 1* for a more comprehensive (but far from exhaustive) list.

In general, neurotransmission is thought of as falling into two broad categories based on its time course. **Fast transmission** occurs whenever a neurotransmitter acts via ionotropic receptors, whereas **slow transmission** occurs by transmitters acting through metabotropic receptors. Glutamate, GABA and acetylcholine (ACh) are together responsible for most of the fast transmission. However, each of these molecules also mediates slow transmission by activating their corresponding metabotropic receptors. It is quite common for each of these transmitters to mediate both fast and slow transmission at the same synapse by activating multiple receptor populations. For example, ACh released from preganglionic cells acts on both nicotinic and muscarinic receptors on postganglionic cells in autonomic ganglia, mediating fast and slow effects of ACh respectively. Catecholamine and peptide transmission are invariably slow. It is extremely common for a given synapse to release more than one transmitter. This is termed **cotransmission** and usually involves the release of a classical transmitter, coupled with the co-release of one or more peptides at high stimulus frequencies.

Transmitters are rapidly cleared from the synaptic cleft after release by one of three methods, passive diffusion away from the cleft, reuptake into surrounding neurons or glia, or enzyme degradation.

C3 Postsynaptic events

Key Notes

Fast transmission

Opening of ligand-gated ion channels is responsible for fast transmission. The flow of ions through the open channels causes postsynaptic potentials to be generated on the postsynaptic membrane. These decay with time and distance as they spread passively over the neuron surface. Postsynaptic potentials are either excitatory or inhibitory.

Excitatory postsynaptic potentials

Activation of ionotropic receptors permeable to Na^+, K^+ and Ca^{2+} results in depolarizing excitatory postsynaptic potentials (epsps). When recorded from the cell body they are caused by the activation of several synapses. They are graded in size from about 0.5 to 8 mV depending on the number of afferent inputs stimulated, and decay exponentially after 10–20 ms. Most fast excitatory transmission is mediated by glutamate and acetylcholine.

Inhibitory postsynaptic potentials

Fast inhibitory postsynaptic potentials (ipsps) are caused by activating ionotropic receptors permeable to chloride ions. This is shown by the equality between the reversal potential of the ipsp current and the chloride equilibrium potential, –70 mV. Inhibitory postsynaptic potentials have similar properties to epsps except for their inhibitory nature. An increase in Cl^- permeability will always be inhibitory since it will tend to stabilize the membrane potential at the E_{Cl}. This is true even if the resting potential is greater than E_{Cl} and the ipsp is depolarizing. Most fast inhibitory transmission is mediated by GABA and glycine.

Slow transmission

Metabotropic receptor actions are intrinsically slower and longer lasting than those of ionotropic receptors. Slow transmitters are often referred to as neuromodulators because, apart from producing slow epsps and ipsps, they have other diverse effects on cell behavior, because of the capacity of metabotropic receptor second messenger systems to open or close a large variety of ion channels throughout a neuron. Slow transmitters modulate transmitter release, the action of fast transmitters, neuronal excitability and gene expression.

Related topics

Resting potentials (B1)
Overview of synaptic function (C2)
Neural integration (C4)
Ionotropic receptors (D1)
Metabotropic receptors (D2)

Fast transmission

The action of neurotransmitters on ligand-gated ion channels is rapid. Changes in membrane potential occur within a millisecond or so and return to resting potential in tens of milliseconds. This is **fast neurotransmission**. The binding of transmitter to ligand-gated ion channels increases the permeability of the post-synaptic membrane to whatever ions the channel conducts. This alters the potential of the postsynaptic membrane, generating a **postsynaptic potential**. These potentials are typically just a few millivolts in amplitude. They spread

passively over the plasma membrane of the nerve cell, diminishing in size as they move away from the postsynaptic membrane, and with time since their generation. Postsynaptic potentials come in two flavors, excitatory and inhibitory.

Excitatory postsynaptic potentials

Activation of ionotropic receptors that are non-selective cation conductances (i.e. those permeable to Na^+ plus K^+ and possibly Ca^{2+} also) depolarizes the synaptic membrane. This is because the reversal potential for the current through these receptors is close to zero. Since this brings the membrane potential of the neuron *closer* to the threshold voltage at which action potentials are triggered, it *increases* the probability that the cell might fire. Hence depolarizing synaptic potentials are termed **excitatory postsynaptic potentials** (**epsps**) and the transmitter–receptor combination is described as **excitatory**. It is important to note that an *individual* epsp has no chance of making a central neuron fire, it is far too small to have a significant effect on the postsynaptic cell. Only when many epsps are generated on a neuron within tens of milliseconds of each other is there any prospect of driving the neuron across threshold. The major fast excitatory transmitters are glutamate and acetylcholine. Glutamate transmission was first studied in the spinal cord where sensory nerve axons from muscles synapse directly with motor neurons (see *Fig. 1*). The synapses are axodendritic, located on dendrites within about 600 mm of the cell body. Intracellular recording from the motor neuron shows that the effect of electrically stimulating the sensory nerve axons is to produce an epsp.

There are several important points to note about excitatory postsynaptic potentials generally:

- It is technically hard to record individual epsps at vertebrate synapses, so generally the epsp resulting from the activation of *several* synapses is recorded at the cell body.
- They are small and graded in size, ranging from fractions of a millivolt to about 8 mV, depending on the number of afferent fibers being stimulated. The reason for this is that if more afferent fibers are stimulated, then more synapses are activated.
- There is a short delay of 0.5 to 1 ms between stimulating the afferents and the generation of an epsp. This is called the **synaptic delay**.
- They typically last for about 10–20 ms before decaying exponentially.

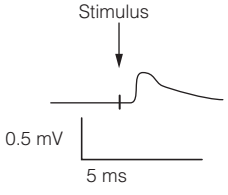


Fig. 1. Excitatory postsynaptic potentials in spinal motor neurons in response to stimulating a single sensory axon.

Inhibitory postsynaptic potentials

Activation of ionotropic receptors that conduct chloride ions channels *reduces* the probability of neuron firing. The most important fast inhibitory transmitters are γ-aminobutyrate (GABA) and glycine. The effect of GABA can be seen in the experiment illustrated in *Fig. 2*. Pyramidal cells in the cerebral cortex have many

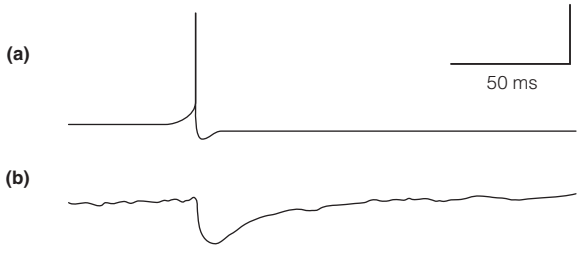


Fig. 2. Inhibitory postsynaptic potential in a pyramidal cell produced by GABA release from an inhibitory (basket) neuron: (a) presynaptic action potential in basket cell, vertical scale bar 25 mV; (b) postsynaptic potential in pyramidal cell, vertical scale bar 0.5 mV.

GABAergic synapses impinging on them from interneurons. Most of these synapses are axosomatic. By intracellular recording from the pyramidal cell it is possible to see that the effect of activating the interneuron is to produce a modest hyperpolarization. This is an **inhibitory postsynaptic potential (ipsp)** because it carries the membrane potential away from the threshold for firing action potentials. Inhibitory postsynaptic potentials have very similar properties to epsps.

Voltage clamping of neurons shows the reversal potential of the current responsible for fast GABA-evoked ipsps to be about –70 mV (*Fig. 3*). This is the equilibrium potential for Cl^-. Hence in response to the release of GABA the permeability of the pyramidal cell to chloride ions increases. When the membrane potential of the neuron is more positive than the reversal potential Cl^- enters the cell, making it more negative inside, i.e. the cell hyperpolarizes.

Fig. 3 shows that when the membrane potential of the neuron is initially more negative than the reversal potential for chloride, Cl^- leaves the cell, the cell becomes less negative inside, i.e., the cell depolarizes. It is important to note that this is usually still inhibitory. In both cases the effect of increasing Cl^- permeability is to force the membrane potential to remain at the chloride equilibrium potential (E_{Cl}) since whenever the membrane potential is not the same as E_{Cl} there will be an ionic driving force causing chloride ions either to leave or enter the cell. The increased Cl^- permeability therefore resists any tendency acting to make the membrane potential move away from –70 mV, and so prevents it from

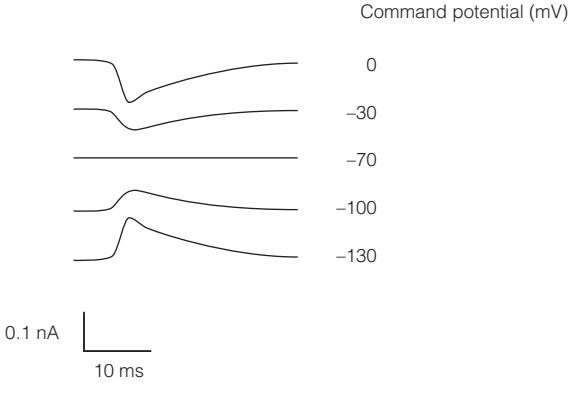


Fig. 3. Reversal potential of fast GABA ipsps, found by voltage clamping. The reversal potential is –70 mV.

being driven towards threshold. Because this inhibition effectively short circuits epsps it is often referred to as **shunting inhibition**.

Slow transmission

Metabotropic receptors affect ion channels only indirectly by way of a cascade of events which take time to switch on and off. Consequently their effects are slower in onset (tens of milliseconds to seconds) and longer lasting (seconds to minutes) than ionotropic receptor action.

Metabotropic receptors can produce slow epsps and ipsps but are also able to produce much more diverse effects on nerve cell behavior than ionotropic receptors. For this reason slow transmitters are sometimes described as **neuromodulators**. The variety of their effects can be attributed to a number of factors.

1. Second messengers are freely diffusible so they spread through dendrites to the cell body and even into the axon, bringing them into contact with channels throughout a nerve cell. Thus metabotropic receptor action is not confined to synapses but determines the global behavior of a neuron.
2. Second messenger systems target a great variety of ion channels including: several types of voltage-dependent K^+ channels responsible for setting the overall excitability of the cell membrane; voltage-dependent channels that generate action potentials; Ca^{2+} channels required for transmitter release; and ligand-gated ion channels. Metabotropic receptor activation can result in channels closing as well as opening.

Key neuromodulatory effects of metabotropic receptors include:

- increases or decreases in transmitter release by acting at presynaptic metabotropic receptors that open or close ion (usually Ca^{2+} or K^+) channels in the presynaptic membrane;
- altering the efficacy of fast transmitters by opening or closing ligand-gated ion channels;
- modifying the excitability of cells by opening or closing voltage-dependent ion channels.

An example of the latter is modulation of the M-type (muscarinic) K^+ channel by acetylcholine acting on muscarinic receptors (*Fig. 4*). The M-channel is a

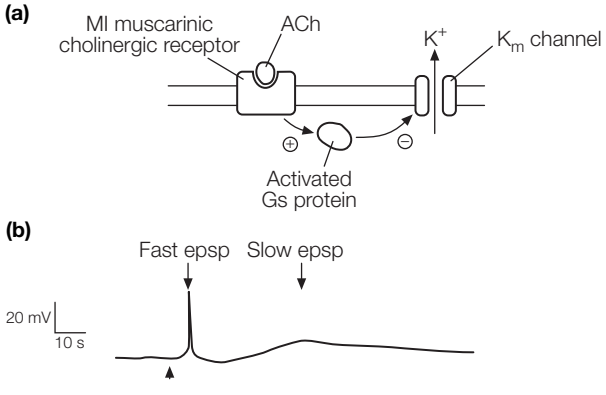


Fig. 4. Modulation of M-type (muscaninic) potassium (K_m) channels. (a) ACh binding to MI muscaninic receptors closes K_m channels. (b) Response of autonomic postganglionic neuron to firing of the preganglionic cell. The slow epsp is due to K_m channel closure by ACh. (The fast epsp is due to activation of nicotinic cholinergic receptors.)

slow-activating, voltage-dependent potassium channel. It is closed by muscarinic receptors. With fewer K^+ ions leaving the cell, it depolarizes producing a slow epsp. Activating the M-channel makes the neuron more sensitive to any excitatory input because, by Ohm's law ($\Delta V = \Delta IR$), closing ion channels – equivalent to increasing membrane resistance (R) – means that less current (ΔI) is needed to depolarize (ΔV) a cell to threshold. M-channels are found in hippocampal pyramidal cells where, controlled by cholinergic inputs from the septum (part of the limbic system), they modify the cell's response to excitatory glutamatergic inputs.

Metabotropic receptors can also have very long-lasting effects on nerve cells, unrelated to their transient influence on cell permeability and electrical excitability, because second messenger systems can act at the nucleus to alter gene transcription. This can alter neuron geometry and bring about functional and structural changes in connectivity between cells. Such mechanisms underlie long-term memory.

C4 Neural integration

Key Notes

Neurons as decision-making devices

Small potentials (e.g. synaptic potentials) decay with time and distance. This behavior underlies how individual nerve cells treat their inputs and hence all information processing in the nervous system. Postsynaptic potentials (psps) generated on a neuron, both excitatory and inhibitory, add together (summate). If the result of this summation is that the axon hillock membrane potential is driven beyond threshold, the neuron will fire. So, whether or not a neuron will fire at any moment depends on how many excitatory and inhibitory synapses are active, and where they are located. It is by integrating synaptic inputs in this way that neurons act as computational devices.

Summation

The summation of psps generated at slightly different times is temporal summation. The summation of potentials arriving on different parts of the neuron is spatial summation. The geometry of a neuron determines the size and time course of synaptic potentials as they spread, and hence the extent of summation that occurs. If summation results in a sufficiently large depolarization of the axon trigger zone a nerve cell will fire.

Related topic Postsynaptic events (C3)

Neurons as decision-making devices

Individual synaptic potentials are too small to activate voltage-dependent sodium channels so they do not trigger action potentials. Instead they are passively conducted over the nerve cell membrane, getting smaller with both time and distance as they spread. This decay of small potentials is determined solely by the physics of the neuron. Generally, the smaller the diameter of a neuron, axon or dendrite along which a potential is spreading, the shorter the distance over which it will decay, the faster this will happen, and the slower the potential is conducted. This is crucial in determining how neurons integrate their inputs and hence how information is processed in the nervous system. In addition it accounts for why action potentials, which do not decay in time and distance, are needed for long-distance transmission. Synaptic potentials decay to zero within a few millimeters in most neurites, so cannot carry information any great distances. However, some short interneurons (e.g. those in the retina) do not fire nerve impulses, but rely on synaptic potentials for transmission along neurites.

Many thousands of synapses are formed on a neuron, both excitatory and inhibitory. At any given time a subset of these will be activated to generate epsps and ipsps. A special property of these graded potentials is that they **summate**, or add together. If a sufficient number of epsps are produced, in summing they will drive the axon hillock membrane potential across the threshold for triggering action potentials and the neuron will fire. The axon hillock is crucial because, being the region of a neuron with the highest density of voltage-dependent sodium channels, it has the lowest threshold. If at any

instant insufficient excitatory synapses are activated, or a high level of excitatory synaptic input is more than offset by the generation of ipsps from inhibitory input, then the axon hillock will not be driven across the threshold and the cell will not fire. So, neurons are decision-making devices. The decision – to fire or not – is actually taken by the axon hillock on the basis of whether the sum total of epsps and ipsps causes its membrane potential to go more positive than the firing threshold. It is this operation that constitutes information processing by individual neurons. In engineering terms, a synapse converts digital signals (action potentials) into analog ones (postsynaptic potentials). The neuron then integrates all its analog signals over a short time and compares the result of that integration with a given threshold to decide whether to fire. When it does fire the output is digital.

Experiments on pyramidal cells show that about 100 excitatory synapses, on average, must be activated at the same time to trigger an action potential. However, the efficacy with which a synapse can influence firing depends on its position. Because postsynaptic potentials decay as they spread passively towards the axon hillock, a synapse far out on a distal dendrite will have less effect than one closer to the cell body. In this context it is noteworthy that on pyramidal cells there are only about 250 inhibitory synapses on the cell body but 10 000 or so excitatory axodendritic synapses. The relative strength of a synapse in contributing to a neuron's output is its **weighting**. This need not be a fixed property but may change with time.

Summation

If an afferent neuron fires a series of action potentials in quick succession (a **volley**), then the earliest psps generated in the postsynaptic cell will not have time to decay before the next psps arrive. Hence successive psps summate over time. This is referred to as **temporal summation**. If sufficient, temporal summation will cause the postsynaptic cell to reach firing threshold.

The summing of postsynaptic potentials generated at separate points on the neuron surface is called **spatial summation** (*Fig. 1*). If a sufficient number of excitatory synapses are activated in relation to inhibitory ones the cell will fire.

Although temporal and spatial summation are described as separate processes, both occur together as a neuron is stimulated and it is their combined effect which dictates whether it will fire. The precise details of how summation works depend on a neuron's geometry because this determines exactly how all the synaptic potentials set up on the cell decay as they spread towards the axon hillock. The frequency with which a cell fires, and how long it fires, is determined by the amplitude and duration of the depolarization of the axon hillock membrane.

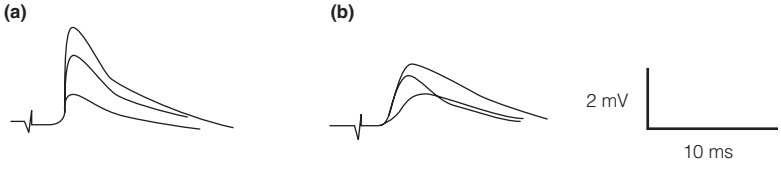


Fig. 1. Spatial summation. In each case the upper trace is the summed response of the two lower epsps generated at synapses: (a) a long way apart, (b) close together.

C5 Neurotransmitter release

Key Notes

Vesicular release

Neurotransmitter release occurs most commonly by calcium-dependent exocytosis from vesicles, in response to excitation of the axon terminal by action potentials. Non-vesicular calcium-independent release of glutamate and GABA via transporters can occur under some circumstances.

Release is quantal

Transmitter is released in discrete packets, quanta, that correspond to exocytosis from a single vesicle. The spontaneous, random release of a single quantum causes miniature endplate potentials (at the neuromuscular junction) or miniature postsynaptic potentials (at CNS synapses). Postsynaptic potentials arise from the release of several quanta simultaneously. At central synapses action potentials trigger neurotransmitter release in only a proportion of occasions.

The role of calcium

Calcium imaging shows how Ca^{2+} moves in space and real time through cells. This reveals that following excitation of the nerve terminal calcium influx is restricted to a small region, but the local concentration reaches 200 mM, sufficient to trigger the exocytosis mechanism for small synaptic vesicles very rapidly.

Exocytosis from large dense-core vesicles

Amines and peptides are released by high-frequency stimulation, only after an appreciable delay, because large dense-core vesicles are situated some distance from the active zone.

Biochemistry of exocytosis

Several linked steps are involved in exocytosis. Recruitment shifts vesicles from a reserve pool into a releasable pool. Binding of vesicle-associated proteins and plasma membrane proteins permits the vesicles to be docked at the active zone in close proximity to voltage-dependent calcium channels. Partial fusion of the vesicle is achieved by priming, mediated by the assembly of a macromolecular complex, and involving the hydrolysis of ATP. The final rapid stage of exocytosis occurs when excitation triggers Ca^{2+} influx. Binding of calcium to synaptotagmin permits fusion to go to completion.

Endocytosis

In the Heuser–Reese cycle vesicles are recycled. Vesicle membrane is coated with clathrin so that it invaginates. Fission of coated vesicle is then triggered by hydrolysis of GTP bound to dynamin. Once in the cytoplasm the vesicle loses its clathrin coat.

Kiss-and-run cycle

In central synapses, in addition to the Heuser–Reese cycle, a much faster kiss-and-run cycle, with much simpler endocytosis, allows high levels of release to be maintained by a small pool of vesicles.

Refilling Classical transmitters are imported into vesicles driven by the efflux of H^+ via specific transporters. The proton gradient is generated by a vesicular proton ATPase. Peptides are packaged in the Golgi apparatus from which vesicles bud to be transported to the axon terminal, itself incapable of protein synthesis.

Autoreceptors Autoreceptors respond to the transmitter released by the neuron in which they are located. They occur at the presynaptic terminal, the soma and dendrites. They regulate neurotransmitter release, synthesis, and neuron firing rate, usually homeostatically.

Related topics Overview of synaptic function (C2) Nerve–muscle synapse (J1) Calcium channels (C6)

Vesicular release

Most neurotransmitter release occurs by transmitter-loaded synaptic vesicles fusing with the presynaptic membrane so that the contents of the vesicle are discharged into the synaptic cleft. This is an example of **exocytosis.** It is triggered by the arrival at the nerve terminal of an action potential which causes a transient and highly localized influx of Ca^{2+}. After release the vesicle membrane is recycled from the presynaptic membrane to form new vesicles by **endocytosis.** The vesicles are subsequently loaded with transmitter via active transporters localized in the vesicle membrane.

Under some circumstances non-vesicular, Ca^{2+}-independent release of transmitters, particularly GABA and glutamate, can be seen. This is thought to occur by the reversal of transport mechanisms that normally serve to reuptake transmitter from the synaptic cleft back into the nerve terminal.

Release is quantal

In vesicular release, neurotransmitter is secreted in discrete packets or **quanta**. Each quantum represents the release of the contents of a single vesicle, about 4000 molecules of transmitter. At the **neuromuscular junction** (nmj), ACh released from a single vesicle diffuses across the cleft in about 2 μs, reaching a peak concentration of around 1 mM, activating 1000–2000 nAChR, to give a depolarization of the muscle fiber membrane locally of approximately 0.5 mV. Such events occur randomly and spontaneously under resting conditions and are called **miniature endplate potentials** (mepps). The **endplate potential** produced by a single action potential arriving at the motor neuron terminal results from the summation of about 300 quanta being liberated simultaneously from around 1000 **active zones**, presynaptic membrane regions specialized for transmitter release.

At CNS synapses **miniature postsynaptic potentials** (mpsps) are seen. They are the equivalent of mepps at the nmj. Miniature endplate potentials are excitatory or inhibitory, depending on the transmitter, and are due to the release of transmitter from a single vesicle acting on only 30–100 receptors that lie under the active zone of central synapses. Excitatory and inhibitory postsynaptic potentials represent the summation of multiple mpsps, generated by an action potential invading several active zones simultaneously. This happens either because axons branch to form several discrete terminals or because some terminals have more than one active zone.

The active zone of many CNS synapses appears to have only one release site.

This is known as the **one vesicle** or **one quantum** hypothesis. However transmitter release does not happen every time an action potential arrives at the presynaptic terminal. Individual active zones behave in an all-or-none fashion because an action potential will either trigger the release of the single quantum or not. The proportion of successes will reflect the probability of release. At central synapses the probability of release varies between different sites and at least at some synapses it also depends on the recent history of the synapse.

The role of calcium

The arrival of an action potential at a nerve terminal causes an influx of Ca^{2+} through voltage-dependent calcium channels. Direct evidence for the role of calcium is provided by **calcium imaging**, a technique which makes visible how Ca^{2+} signals spread in time and space through cells. Fluorescent dyes are used which on binding Ca^{2+} absorb UV light at a different wavelength than they do in the unbound state. Neurons are preloaded with the dye and the emission of UV from the dye is observed in response to its excitation by the two distinct absorption wavelengths. This gives a quantitative measure of how the concentration of Ca^{2+} changes in the neuron in real time.

This technique shows that it takes about 300 μs for calcium channels at the active zone to open in response to an action potential. The driving force for calcium entry is extremely high because of the large concentration gradient. The free Ca^{2+} concentration at rest in a terminal is 100 nM whilst the external concentration is about 1 mM. Despite this huge concentration gradient, the presence of diffusion barriers and calcium buffers in the terminal restrict the rise in calcium concentration to within 50 nm of the channel mouth. This region is called a **calcium microdomain**. The $[Ca^{2+}]$ within 10 nm of the channel mouth rises to 100–200 μM, which matches the half-maximal concentration of Ca^{2+} for glutamate release. Several overlapping microdomains cooperate to trigger the release of a vesicle in close proximity.

Exocytosis from large dense-core vesicles

In contrast to small clear synaptic vesicles (SSVs), the mechanism for release from large dense-core vesicles (LDCVs) vesicles takes longer and has a higher affinity for calcium (half-maximal release occurs at about 0.4 μM), because only a small amount of Ca^{2+} manages to diffuse to the LDCVs, which are some distance from the active zone. Hence exocytosis of amines and peptides occurs with a delay of about 50 ms and only in response to high-frequency stimulation of the neuron which causes high levels of calcium influx.

Biochemistry of exocytosis

Exocytosis from SSVs involves several linked steps, most of which need calcium. Nerve terminals contain two pools of SSVs. The **releasable pool** is located at the active zone and can take part in repeated cycles of exocytosis and endocytosis at low neuron firing frequencies. The **reserve pool** consists of vesicles tethered to cytoskeletal proteins, and can be mobilized by repetitive stimulation to join the releasable pool. This is called **recruitment**. Liberation of a vesicle from the cytoskeleton requires Ca^{2+}-dependent phosphorylation of **synapsin I**, a protein which anchors vesicles to actin filaments in the terminal.

Vesicles are aligned at specific sites in the active zone by a process termed **docking,** which involves **SNARE** proteins (*Fig. 1*). A vesicle-associated protein, **synaptobrevin** (v-SNARE, VAMP) binds with high affinity to a presynaptic membrane protein, **syntaxin** (t-SNARE). Syntaxin is closely associated with voltage-dependent calcium channels ensuring that the release machinery is optimally placed to receive the Ca^{2+} signal. Synaptobrevin and syntaxin, together

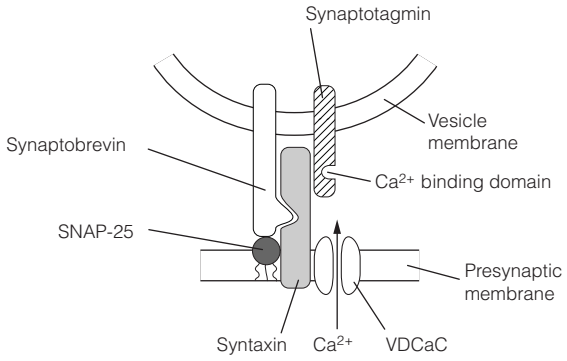


Fig. 1. Proteins involved in the docking of neurotransmitter vesicles. VDCaC, voltage-dependent calcium channel.

with a third protein crucial for docking, **SNAP-25**, are targets for **botulinum** and **tetanus toxins** that are powerful inhibitors of neurotransmitter secretion.

After docking comes another calcium-dependent step, **priming**, in which a number of soluble cytoplasmic proteins form a transient complex with the SNAREs, resulting in partial fusion of vesicle and presynaptic membranes. This step involves the hydrolysis of ATP.

Primed vesicles are poised for exocytosis, requiring only a large pulse of Ca^{2+} to permit complete fusion of the vesicle and presynaptic membranes and opening of the **fusion pore** through which exocytosis occurs. A calcium-binding protein located in the vesicle membrane, **synaptotagmin**, is the Ca^{2+} sensor in the exocytotic machinery. In the absence of calcium it prevents complete fusion but when it binds Ca^{2+} it undergoes a conformational change which allows fusion to proceed. This final stage is fast. It must occur within 200 μs.

Endocytosis

Following exocytosis, synaptic vesicles are recycled within 30–60 s by endocytosis. Firstly, the vesicle membrane acquires a clathrin coat, distorting it so that it invaginates into the terminal. Next a GTP-binding protein, **dynamin**, forms a collar around the neck of the invagination. Dynamin has an intrinsic GTPase activity which hydrolyzes the bound GTP so triggering the fission of the coated vesicle from the presynaptic membrane. The GTP bound form of dynamin requires calcium, so the same rise in nerve terminal Ca^{2+} concentration responsible for exocytosis also enables endocytosis. Once free in the terminal the vesicle loses its clathrin coat (*Fig. 2*).

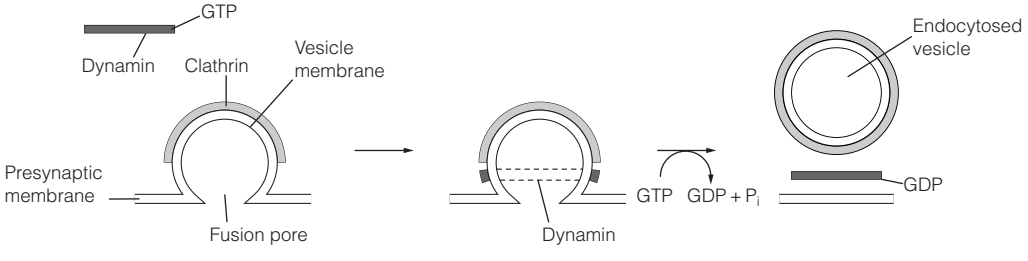


Fig. 2. Vesicle endocytosis. Reprinted from Revest, P.A. and Longstaff, A. (1998) Molecular Neuroscience*. © BIOS Scientific Publishers Ltd, Oxford.*

Kiss-and-run cycle The above account describes the **Heuser–Reese cycle**, based on studies of the neuromuscular junction. A recently documented **kiss-and-run cycle** also operates at central synapses. Here the vesicle membrane fuses with the presynaptic membrane to open a pore through which transmitter discharges, after which the pore closes and the vesicle disengages. There is no complicated endocytosis. Because this mechanism has a cycle time of only one second it is able to support extended periods of high release with only 35–40 vesicles in the recycling pool.

Refilling Vesicles are reloaded with neurotransmitter in the nerve terminals. The vesicles are acidified by the action of a proton ATPase. The transport of transmitter into vesicles is then driven by secondary active transport with H^+ efflux providing the energy (*Fig. 3*). Vesicle transporters have been identified for a number of transmitters including glutamate, ACh and catecholamines, but not yet for GABA. They are large glycoproteins with 12 transmembrane segments. Surprisingly, they seem unrelated to neurotransmitter transporters located in plasma membranes of neurons or glia. Peptide transmitters, after synthesis on ribosomes in the cell body, are secreted into the lumen of the rough endoplasmic reticulum (RER), and packaged for export by the Golgi apparatus, from which the loaded vesicles are budded. These are then moved to the terminal by fast axoplasmic transport. This is necessary because nerve terminals are devoid of ribosomes and incapable of protein synthesis.

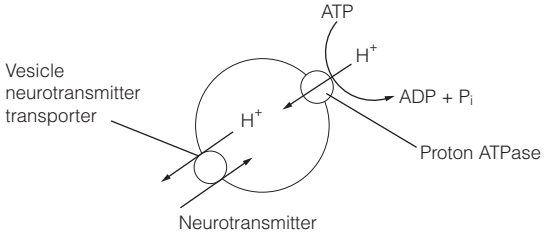


Fig. 3. Vesicle refilling. Reprinted from Revest, P.A. and Longstaff, A. (1998) Molecular Neuroscience*. © BIOS Scientific Publishers Ltd, Oxford.*

Autoreceptors Neurotransmitter receptors are not confined to the postsynaptic membrane but also exist in the presynaptic membrane, where they are termed **presynaptic receptors,** and over the cell body and dendrites. If these are receptors for the transmitter released by the neuron in which they are located, they are **autoreceptors**. Autoreceptors, which are invariably metabotropic receptors, have several functions that are normally homeostatic. Those on the presynaptic membrane are involved in regulating neurotransmitter release. Most (but not all) presynaptic autoreceptors *decrease* the release of neurotransmitter by reducing calcium influx into the presynaptic terminal. This is a negative feedback mechanism either to avoid excessive excitation, or to curtail postsynaptic receptor desensitization, which would reduce the sensitivity of the synapse.

Autoreceptors can also decrease the synthesis of transmitter (e.g. of catecholamines and serotonin by their respective neurons) and some dopaminergic neurons have dopamine receptors on their dendrites and cell body which regulate neuron firing rate.

Some presynaptic receptors are receptors for transmitters *not* secreted by the neuron in which they are situated. These **heteroceptors** also regulate transmitter release. For example, $GABA_B$ receptors exist presynaptically at glutamatergic synapses where they reduce glutamate release. It is assumed that they are activated by GABA that has diffused from neighboring synapses.

C6 Calcium channels

Key Notes

Channel characterization

Calcium channels are responsible for excitation–secretion coupling in neurons, dendritic action potentials and excitation–contraction coupling in muscles. There are several types of calcium channel that can be characterized by their electrophysiology (activation voltage, conductance, time course of inactivation), by their susceptibility to blockade by specific drugs or toxins, and by their distribution.

Channel types

Most calcium channel types are activated by quite large depolarization. L-type channels are responsible for excitation–contraction coupling in muscle. They are blocked selectively by calcium channel antagonists. N-, P- and Q-type channels are all implicated in neurotransmitter release. They can each be blocked selectively by toxins. These channels may coexist at some synapses, each contributing to the calcium influx required for exocytosis. T-type channels are activated by small depolarizations. This property underlies burst firing of thalamic neurons.

Molecular biology of voltage-dependent calcium channels

A functional calcium channel consists of an α1 subunit which closely resembles a voltage-dependent sodium channel, together with auxiliary proteins which modify channel properties. That there are many different isoforms of α1 subunits is the cause of the diversity of channel types.

Related topics

Channel molecular biology (B4)
Neurotransmitter release (C5)
Nerve–muscle synapse (J1)
Sleep (M5)
Epilepsy (P4)

Channel characterization

Voltage-dependent calcium channels control the influx of calcium, which couples excitation to secretion of transmitter. They are also responsible for calcium action potentials in dendrites, and for excitation–contraction coupling in skeletal, cardiac and smooth muscle. There are several distinct types of calcium channel. They are all Ca^{2+} selective and activated by depolarization but can be differentiated by their electrophysiological properties, their sensitivity to drugs or toxins, and their distributions and functions in the nervous system. The electrophysiological criteria used to distinguish the channels include:

- The size of the depolarization needed to activate them. High-voltage activated (HVA) need a large depolarization, and low-voltage activated (LVA) require only a small depolarization.
- The conductance of the channel.
- The time course of inactivation.

Channel types

These are summarized in *Table 1*. **L-type** channels are HVA and require depolarization to –20 mV for activation. L-type channels are located in proximal

Table 1. Calcium channel types

Type	Named for	Electrophysiology	Location
L	Long-lasting	HVA (–20 mV) Slowly inactivating	Pyramidal cells Skeletal, cardiac and smooth muscle Endocrine cells
T	Transient	LVA (–65 mV) Rapidly inactivating	Cardiac muscle Neurons (e.g. thalamic) Endocrine cells
N	Neuronal	HVA (–20 mV) Moderate inactivation	Neurons
P	Purkinje cell	HVA (–50 mV) Non-inactivating	Cerebellar Purkinje cells Mammalian neuromuscular junction
Q	Q after P	HVA	Cerebellar granule cells
R	Remaining	HVA and LVA	

HVA, high voltage activated; LVA, low voltage activated.

dendrites of pyramidal neurons and contribute to their excitability but are not the presynaptic calcium channels involved in neurotransmitter release. They are the major calcium channels of excitation–contraction coupling. They are the only channel type so far which can be targeted by therapeutically useful agents, the calcium channel antagonists. These are mostly used in cardiovascular medicine. Unfortunately they have not proved effective for the treatment of strokes in clinical trials.

Three types of HVA calcium channel are implicated in transmitter release by the ability of selective toxins to block release. **N-type** channels, found on a variety of neurons, can be blocked by ω-conotoxin from the cone snail *Conus geographus* and play a more prominent part in GABA than glutamate release**. P-type** channels can be blocked by toxins of the funnel web spider *Agenelopsis aperta*, and are responsible for GABA release from cerebellar Purkinje cells (for which these channels are named), and also for a substantial fraction of glutamate release from pyramidal cells and ACh release at the mammalian neuromuscular junction. Glutamate release from cerebellar granule cells and a fraction of glutamate release from pyramidal cells is controlled by **Q-type** channels blocked by ω-conotoxin from *Conus magus*. Different types of calcium channel may coexist in the same terminal where each makes a contribution to transmitter release.

T-type channels are LVA channels, activated by depolarizations beyond –65 mV and relatively rapidly inactivating. These properties mean that they can generate burst firing of excitable cells. They are important in the thalamus.

Molecular biology of voltage-dependent calcium channels

Calcium channels are macromolecular complexes composed of five different subunits. One, the α1 subunit is the functional channel (the other subunits are auxiliary proteins which can modify channel properties) and resembles the voltage-dependent sodium channel. There are many distinct α1 subunits because there are six genes which code for different versions and each is subjected to alternative splicing. This accounts for the diversity of channels.

C7 Neurotransmitter inactivation

Key Notes

The reason for inactivation

Inactivation of transmitter action occurs so that synapses can be modulated on a fast timescale, and to curtail receptor desensitization. Inactivation can occur by enzyme degradation, by transport out of the cleft back into neurons or glia, or by diffusion away from the synapse.

Enzyme degradation

Acetylcholine is hydrolyzed by acetylcholinesterase. The liberated choline is taken up into the nerve terminal via a high affinity Na^+-dependent cotransport system. ATP is also inactivated enzymically.

Transport

Reuptake from the synaptic cleft into neurons (and glia in the case of the amino acids) is a major mechanism for the inactivation of the classical transmitters. Two major families of transporters are involved. These molecules are unrelated to the vesicular transporters for transmitters. Cocaine exerts its effects by blocking the dopamine transporter. Serotonin transporters are thought to be the crucial targets for antidepressant drugs.

Diffusion

Diffusion away from the synaptic cleft is the major mode of inactivation of the peptides and is also probably important for glutamate and GABA. The large size of the peptides makes their diffusion slow and accounts for their protracted action.

Related topics

Overview of synaptic function (C2)
Acetylcholine (D7)
Nerve–muscle synapse (J1)
Depression (P2)

The reason for inactivation

Inactivation is necessary to ensure that synapses respond to rapid changes in presynaptic neuron firing frequency. Without it the postsynaptic cell could not be updated on recent changes in incoming signals. In addition, by clearing transmitter from the synaptic cleft, inactivation limits receptor desensitization. There are three ways in which transmitter can be inactivated and they are not mutually exclusive: enzymic degradation, transport out of the synaptic cleft back into neurons or glia, or by passive diffusion away from the synapse.

Enzyme degradation

Many enzymes are involved in the catabolism of both classical and peptide transmitters and pharmacological manipulation (e.g. inhibition) of many of these can have consequences for synaptic transmission. However, only in a couple of cases is enzyme-catalyzed degradation important in the *physiological* inactivation of transmitter. Acetylcholine is hydrolyzed by **acetylcholinesterase** (AChE), which cleaves the transmitter molecule into choline and acetate.

Choline is taken back into the presynaptic nerve terminal by a Na^+-dependent transporter. AChE has an extremely high catalytic activity and at the nmj can reduce the concentration of ACh from about 1 mM immediately after release to virtually zero in about 1 ms.

ATP, a cotransmitter at some synapses, is the only other molecule in which enzyme degradation is important for its synaptic inactivation.

Transport

Many classical transmitters are inactivated by their removal from the cleft via high-affinity, saturable, secondary active transport. The amino acid transmitters may be transported into both neurons or glia whereas amines are transported only into neurons. Two families of transporter have been identified which serve this function:

1. Na^+/K^+-cotransporter family constitutes the glutamate (and aspartate) transporters. Three have been discovered so far, two present in glia (astrocytes) and one localized in neurons. Glutamate transport is electrogenic, that is, it results in a modest potential difference being set up across the membrane, inside positive (see *Fig. 1*). A consequence of this is that excessive depolarization of the membrane can reverse the direction of the transport causing glutamate *efflux* into the cleft. This can result in excitotoxicity. Glutamate transporters have been cloned and sequenced.

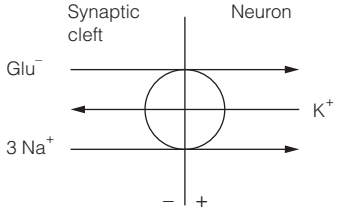


Fig. 1. Glutamate transport by a Na^+/K^+ cotransporter.

2. Na^+/Cl^--cotransporter family (*Fig. 2*). This is a large family and includes three GABA transporters (GATs), the transporters for noradrenaline/adrenaline (norepinephrine/epinephrine), dopamine, serotonin, glycine, and the high-affinity transporter for choline. The three GABA transporters are all expressed in both neurons and glia yet pharmacological experiments can distinguish glial from neuronal GABA uptake. It is likely then that more GABA transporters remain to be discovered. The noradrenaline (norepinephrine) and serotonin transporters are the targets for the tricyclic antidepressant

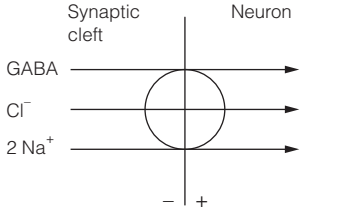


Fig. 2. GABA transport by a Na^+/Cl^- cotransporter.

class of drugs. The recent development of drugs acting at the serotonin transporter, selective serotonin reuptake inhibitors (SSRIs), such as fluoxetine (Prozac), is proving useful in the treatment of depression, although recently questions have been raised about their safety and efficacy. The dopamine transporter is the target for cocaine. By inhibiting dopamine reuptake, cocaine deranges dopamine transmission in reward pathways in the brain, accounting partly for its addictive properties. Many members of this family have been sequenced. They are large glycoproteins with 12 transmembrane segments, but have no homology with vesicle dopamine transporters.

Diffusion

Despite the existence of transporters, diffusion out of the synapse may be important for the inactivation of glutamate and GABA at synapses in the cerebral cortex. There is no high-affinity reuptake for peptides and although peptides may be internalized by neurons via receptor-mediated endocytosis and then degraded by non-specific peptidases, this is probably not an important mechanism for their inactivation. Hence the major route for terminating the synaptic action of peptides is by diffusion. However peptides are very much larger than the small classical transmitter molecules and there are significant barriers for free diffusion out of the cleft. This means that peptides clear only slowly from a synapse, which helps to explain why their actions can be so prolonged.

D1 Ionotropic receptors

Key Notes

Ligand-gated ion channel receptors

These receptors fall into two classes, the nicotinic receptor family and the ionotropic glutamate receptors. All consist of several subunits arranged around a central pore.

Nicotinic receptor family

The nicotinic receptor superfamily members are pentamers assembled from several distinct subunits. Each subunit has four transmembrane segments that are either α-helices or β-pleated sheets and extracellular N and C terminals. Both nicotinic cholinergic receptors (nAChR) and $GABA_A$ receptors show allostery (positive cooperativity) in ligand binding. The nAChR receptor ion channel is permeable to both Na^+ and K^+ while the $GABA_A$ receptor is a chloride permeable channel.

Pharmacology of $GABA_A$ receptors

The $GABA_A$ receptor has allosteric sites for binding benzodiazepines, barbiturates and steroid anesthetics. These drugs all enhance inhibition by increasing Cl^- conductance. Benzodiazepines have antianxiety and anticonvulsant profiles. Inverse agonists *decrease* Cl^- conductance and have an anxiety-producing, pro-convulsant profile. Barbiturates are sedative and used for anesthetic induction. Both benzodiazepines and barbiturates have roles in epilepsy treatment.

Ionotropic glutamate receptor family

Most fast transmission by glutamate is via AMPA/kainate receptors, which have a tetrameric structure. Their pore region resembles a potassium channel, but inside-out in membrane orientation. Most native AMPA receptors are permeable to Na^+ and K^+ but some are also Ca^{2+} permeable. NMDA receptors are important because of their involvement in learning and memory, and neuropathologies, including epilepsy and strokes. They share only weak homologies with other ligand-gated ion channels. They are Ca^{2+} permeable and are blocked by Mg^{2+}except at depolarizing potentials. They have a requirement for glycine as a co-agonist and binding sites for Zn^{2+}, polyamines and several drugs.

Related topics

Overview of synaptic function (C2)
Postsynaptic events (C3)
Nerve–muscle synapse (J1)
Cell physiology of learning (O3)
Strokes and excitotoxicity (P3)
Epilepsy (P4)

Ligand-gated ion channel receptors

Ligand-gated ion channel receptors can be divided into two families (see *Table 1*), the conventional ligand-gated ion channels typified by the nicotinic cholinergic receptor, and the ionotropic glutamate receptor family. All ligand-gated ion channels are thought to consist of several subunits clustered around a central pore. Usually several distinct subunits aggregate to form a receptor, with different combinations conferring different functions.

Table 1. Ligand-gated ion channels

Nicotinic receptor family	nAChR	
	$GABA_A$	
	$GABA_C$	
	glycine	
	$5\text{-}HT_3$	
Glutamate receptor family	GluR1–GluR4	(AMPA receptors)
	GluR6	(kainate receptors)
	NMDAR	

Nicotinic receptor family

The nicotinic receptor family are pentamers, that is they consist of five subunits. The nicotinic receptor (nAChR) is responsible for fast acetylcholine transmission. It has the subunit configuration 2α, β, γ, δ. The subunits share modest amino acid homologies both with each other and with corresponding subunits across widely differing species. Each subunit has extracellular N and C terminals and four transmembrane segments (M1–M4) of uncertain secondary structure. Although an α-helical structure has usually been accepted, there are powerful theoretical arguments in favor of a β-pleated sheet configuration for the transmembrane segments (*Fig. 1*).

Each of the α-subunits has a binding site for acetylcholine, thus each receptor binds two molecules of ACh. Binding of one molecule of ACh to a receptor makes binding of the second easier. This is an example of **allostery** (see *Instant Notes in Biochemistry*). The channel is a non-selective cation conductance allowing Na^+ influx and K^+ efflux. The net current is inward (i.e. depolarizing) so ACh is excitatory at nicotinic receptors.

$GABA_A$ receptors are responsible for all *fast* GABA transmission. Their activation opens a channel selective for Cl^-. They are aggregates of various combinations of subunits designated α, β, γ, δ and ρ (which should not be confused with nAChR subunits with the same designation), but because each subunit comes in a number of versions, **isoforms**, a large number of distinct $GABA_A$ receptors may exist. $GABA_A$ subunits have some homology with nAChR subunits. Indeed the binding site for GABA-binding site is structurally similar to the acetylcholine-binding site, with two molecules of GABA acting allosterically to open the channel.

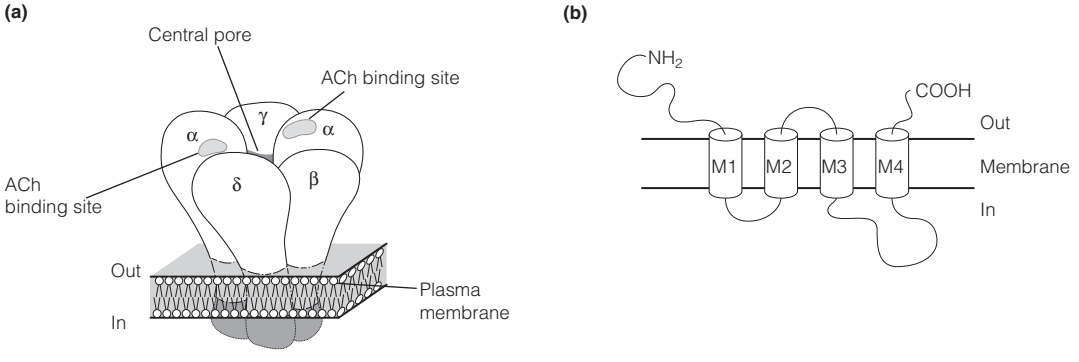


Fig. 1. The nicotinic receptor family: (a) pentameric arrangement of subunits; (b) cartoon of subunit secondary structure.

Pharmacology of $GABA_A$ receptors

The $GABA_A$ receptor contains binding sites for several major classes of drugs. Typical **benzodiazepines** (e.g. diazepam) are allosteric agonists in that they bind to a **benzodiazepine-binding** (**BZ**) **site** distinct from the GABA-binding site, increasing the affinity of the receptor for GABA, which causes the chloride channel to open more frequently. The overall effect is to potentiate the inhibitory effect of GABA without prolonging it, and this presumably accounts for their anti-anxiety and anticonvulsant actions.

Benzodiazepines have a place in the *short-term* treatment of anxiety (dependence is seen with long-term use) and (intravenously) as sedatives or basal anesthetics in minor surgery. Most benzodiazepines are too sedative to be used for maintenance therapy in epilepsy, but intravenously they can be life-saving in severe seizures.

Inverse agonists bind the BZ site and *decrease* channel opening. Not surprisingly they have the unpleasant pharmacological profile of being anxiogenic (anxiety-causing) and pro-convulsant. One benzodiazepine, **flumazenil**, is an antagonist at the BZ site. It reverses the effect of intoxication by benzodiazepine agonists and ethanol. For some years there has been speculation that there may be an endogenous ligand for the BZ site, presumed to play a role in fear responses. To date this is unproven.

The **barbiturates** and **steroid anesthetics** also have allosteric actions at the $GABA_A$ receptor – each at their own distinct binding sites – but the effect of these drugs is to prolong the length of time that the chloride channel is open so that GABA inhibition lasts longer. Highly lipophillic barbiturates, such as **thiopentone**, cross the blood–brain barrier easily, and are administered intravenously for the rapid induction of anesthesia. Although all barbiturates are sedative, phenobarbitone remains an option in the treatment of some types of epilepsy.

Ionotropic glutamate receptor family

The **ionotropic glutamate receptors** (**iGluRs**), which have a tetrameric quaternary structure, share only weak homologies with the nicotinic receptor superfamily. There are three populations of iGluRs, defined by selective agonists; **AMPA receptors**, **kainate receptors** and **NMDA receptors**. All are relatively non-selective cation channels, though NMDA receptors and some AMPA receptors favor calcium permeation.

The most likely secondary structure for the AMPA and kainate receptor subunits is shown in *Fig. 2*. Hydropathicity profiles suggest the presence of three transmembrane (TM) segments and a loop (TMII), which inserts into the membrane and probably contributes to the pore. The region spanning TMI–TMIII bears a striking resemblance to the S5–H5–S6 region of potassium channels, albeit oriented in the membrane in the opposite sense. These receptors may thus have evolved from some ancient potassium channel.

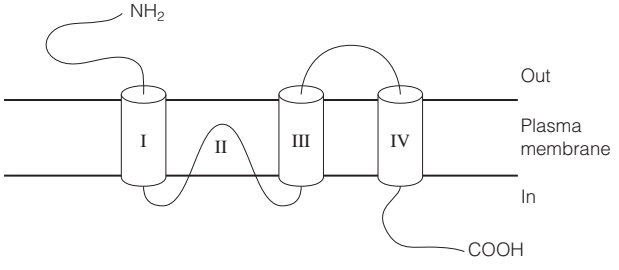


Fig. 2. Ionotropic glutamate receptor subunit structure.

Most AMPA receptors are formed by GluR2 subunits together with either GluR1 or GluR3 subunits. GluR2 subunits have several rather special roles in AMPA receptors:

1. They control the voltage gating of the receptor.
2. They come in two forms that differ in their TMII region by just one amino acid. Receptors with one version are Na^+ and K^+ channels, whereas those with the other version are Ca^{2+} channels.
3. The GluR2 subunit controls the transport of AMPA receptors to the post-synaptic membrane.
4. By binding an as yet unidentified protein, GluR2 subunits of AMPA receptors stimulate the growth of dendritic spines on cortical pyramidal cells.

The NMDA receptor is named after the selective agonist ***N*-methyl-D-aspartate.** The receptor is important because it is implicated in key aspects of brain function such as development, learning and memory, and in pathologies, e.g. strokes and epilepsy. NMDA receptors are the target for **dissociative anesthetics** (e.g. **ketamine**). They also harbor the sigma opioid binding site, so called because it is responsible for the psychotomimetic – schizophrenia-mimicking – effects of some opioids and **phencyclidine**. NMDA receptors have some unusual properties that are summarized below:

1. They are both ligand- and voltage-gated. At resting membrane potentials glutamate will bind to the receptor but the ion channel is blocked by Mg^{2+} ions. This blockade is lifted only by a large depolarization. In other words, the ion channel is only opened if glutamate binds and the receptor experiences depolarization at the same time.
2. The ion channel is permeable to Ca^{2+} as well as Na^+ and K^+. Under some circumstances calcium entry through NMDA receptors can be a significant factor in raising intracellular Ca^{2+} concentrations.
3. Glycine, normally an inhibitory transmitter that acts via receptors that are very similar to $GABA_A$ receptors, acts as a **co-agonist** at NMDA receptors. Glycine acts allosterically to dramatically potentiate the effects of glutamate. The concentration of glycine available to the receptor *in vivo* is about that needed to give maximum potentiation.
4. NMDA receptors can be modulated in complicated fashion by Zn^{2+} ions and by polyamines such as spermine. Generally, *in vivo*, zinc inhibits while spermine potentiates the action of glutamate on NMDA receptors.

The structural complexity of NMDA receptors matches their functional complexity. The subunits from which they are assembled are broadly similar to the AMPA and kainate receptor subunits. Five genes code for NMDA receptor subunits. The *nmdar1* gene codes for the NR1 subunit. These can form homomeric channels with all the properties of native NMDA receptors. The remaining four genes code for NR2 subunits A–D. In addition numerous subunit isoforms are made by alternative splicing (see *Instant Notes in Molecular Biology*) which increases the number of possible distinct NMDA receptors. On their own NR2 subunits cannot form channels. However, they will form functional channels together with NR1 subunits.

The NR1 subunit can simultaneously bind the proteins Eph B2 and ephrin B, located in the post- and presynaptic membranes respectively. This increases the Ca^{2+} entry through the NMDA receptor in response to glutamate, necessary for forming dendritic spines during brain development.

D2 Metabotropic receptors

Key Notes

G protein-linked receptors

Many neurotransmitter or hormone receptors and sensory transduction molecules are G protein-linked receptors. They have seven transmembrane segments (TMI–TMVII). The third cytoplasmic loop between segments V and VI interacts with the G protein. Ligand binding of peptides occurs to external parts of the receptor, amines bind to sites embedded in the membrane. The metabotropic glutamate receptors form a separate family distinguished by their large N terminal end which contains the glutamate-binding site.

G proteins

Metabotropic receptors are coupled to ion channels or second messenger enzymes via trimeric GTP-binding proteins, G proteins. There are several different G protein families with their own particular targets. Binding of ligand to receptor causes liberation of a GTP bound form of G protein which activates its targets. The intrinsic GTPase activity of the G protein rapidly hydrolyzes the GTP, hence curtailing its own activity.

Activation of adenylyl cyclase

G_S proteins activate adenylyl cyclase, which converts ATP to cyclic adenosine monophosphate (cAMP). This second messenger molecule activates protein kinase A, which phosphorylates its target proteins. The cAMP is subsequently degraded by phosphodiesterases. Depletion of cAMP, the action of phosphatases and receptor desensitization all act to curtail the effects of the cAMP second messenger system.

Inhibition of adenylyl cyclase

G_i proteins inhibit adenylyl cyclase. The activity of this enzyme and so the concentration of cAMP in a cell at any time therefore depends on activation of receptors coupled to G_S relative to those coupled to G_i.

Phosphoinositide second messenger system

G_q proteins activate phospholipase C, which cleaves a membrane phospholipid to generate two second messenger molecules. Diacylglycerol (DAG) activates protein kinase C. Inositol trisphosphate (IP_3) mobilizes calcium from internal stores to raise cytoplasmic Ca^{2+} concentration, which activates calcium-dependent protein kinases.

Related topics

Overview of synaptic function (C2)
Retina (G3)
Olfactory receptor neurons (I1)
Neurotrophic factors (N6)
Cell physiology of learning (O3)

G protein-linked receptors

G protein-linked receptors form a huge superfamily. Its members include receptors for slow neurotransmitters, many hormones, and sensory transduction molecules important in vision, smell and taste. Responses produced by these receptors may last for seconds or minutes. Detailed X-ray diffraction studies of one member of the superfamily allowed its structure to be deduced and all other

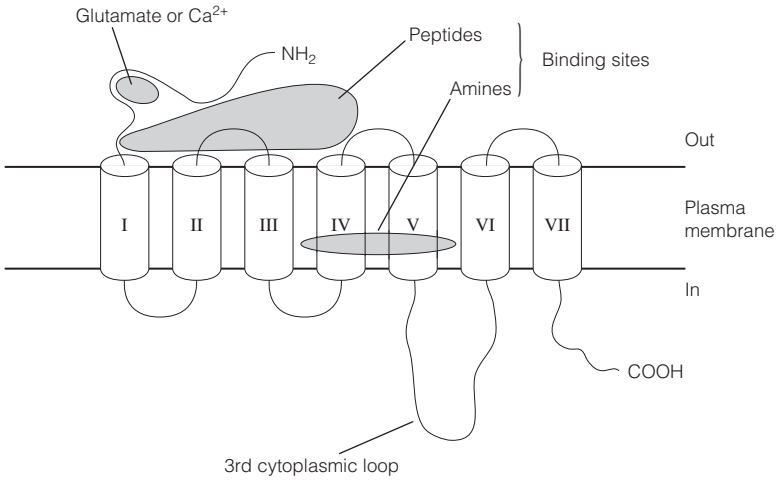


Fig. 1. G protein-linked receptors: cartoon showing transmembrane segments and ligand binding sites; each of the Roman numerals designates a transmembrane segment.

members are assumed to be the same, based on the homologies they share in their primary sequences. This structure is illustrated in *Fig. 1*.

The key features include seven membrane-spanning segments (I–VII) and the third cytoplasmic loop that couples to the G protein. The peptide-binding receptors have their binding domains associated with several extracellular regions. In those receptors that bind small amines the ligand binds to membrane-spanning regions embedded quite deeply in the membrane. Although metabotropic glutamate receptors are G protein-linked receptors they have little homology with the others, and have a large N terminal end which binds glutamate.

G proteins

G proteins, trimers consisting of α-, β-, and γ-subunits, are so called because the α-subunit binds GTP. Binding of neurotransmitter to metabotropic receptors activates their associated G proteins, which may do one or both of the following:

- They may interact directly with ion channels causing them to open or close.
- They may interact with enzymes, e.g. adenylyl cyclase and phospholipase C, to switch on or off second messenger cascades that regulate ion channels and other proteins by phosphorylation.

The cycle of events by which a G protein couples metabotropic receptor activation to second messenger modulation is shown in *Fig. 2*.

Binding of the transmitter allows the receptor and G protein to couple. GDP leaves the α-subunit in exchange for GTP. In its GTP-bound form the G protein dissociates into separate α and β/γ-subunits. The α-subunit activates the enzyme. The α-subunit has an intrinsic GTPase activity that cleaves the terminal phosphodiester bond in the GTP converting it to GDP. In its GDP-bound form the α-subunit uncouples from the enzyme, which reverts to its basal activity. One purpose of this cycle is to act as an amplifier. A single transmitter-binding event results in several cycles of G protein shuttling between receptor and enzyme. Furthermore the enzyme will have time to catalyze the synthesis of hundreds of second messenger molecules before it is switched off by the hydrolysis of the G protein-bound GTP.

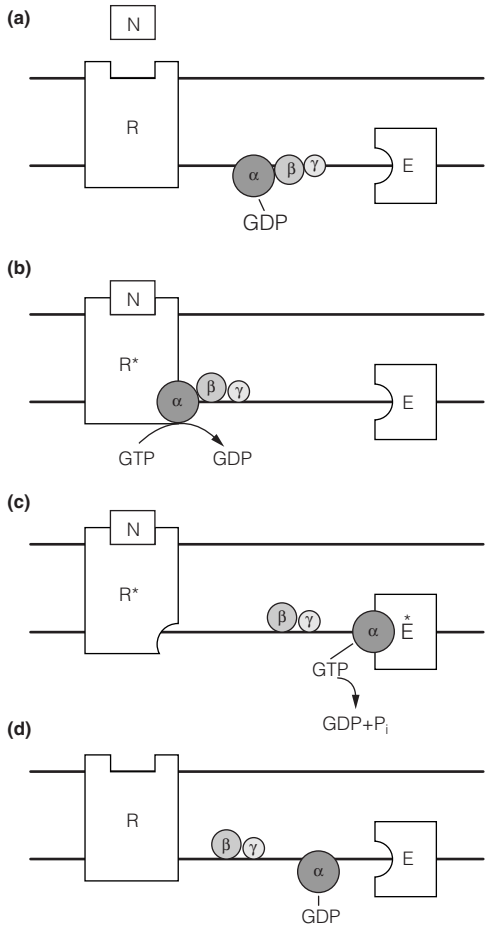


Fig. 2. Coupling of metabotropic receptors to second messenger systems by G proteins. N, Neurotransmitter; R, receptor; E, enzyme.

There are several distinct G proteins, differing largely in their α-subunits. G_S and G_I interact with adenylyl cyclase, G_q with phospholipase C. Despite this multiplicity, G proteins serve as a point for convergence of signals impinging on a neuron because many receptors talk to just a few second messenger systems. *Table 1* lists some of the major G-protein-linked receptors for selected transmitters, together with the second messenger systems they are coupled to.

Activation of adenylyl cyclase

Adenylyl cyclase is activated by a specific family of G proteins, the **G_S proteins**, so called because their action on adenylyl cyclase is stimulatory. The enzyme catalyzes the conversion of ATP to **cyclic adenosine-3′,5′-monophosphate** (**cAMP**). This second messenger molecule diffuses freely through the cytoplasm and binds to a kinase enzyme, **protein kinase A** (**PKA**), which is thereby switched on (*Fig. 3*). The kinase then phosphorylates target proteins that have the appropriate amino acid sequence to recognize the kinase. Targets include ion channels (the phosphorylation state of a channel often determines whether it

Table 1. Second messenger coupling to selected neurotransmitter receptors

G protein	Second messenger	Receptor
G_s	Increased cAMP	β1, β2, β3 adrenoceptors D1, D5 (dopamine) H2 (histamine)
G_i	Decreased cAMP and/or opening of K^+ channels closing of Ca^{2+} channels	α2 adrenoceptors D2, D4 (dopamine) $GABA_B$ 5-HT^1 (serotonin) mGlu, types II and III (glutamate) M2, M4 (muscarinic) μ, δ and κ opioid
G_q	Increased phosphoinositide metabolism	α1 adrenoceptors CCK (cholecystokinin) mGlu, type I (glutamate) 5-HT^2 (serotonin) M1, M3, M5 (muscarinic) H1 (histamine) NK (tachykinin)

is open or closed; phosphorylation opens many but closes some), and transcription factors – allowing the cAMP second messenger system to modify gene expression.

Clearly a single activated PKA molecule is able to phosphorylate many target proteins, adding to the amplification achieved by the system. Second messenger cascades must rapidly turn off if their signals are to be modulated over a time course of tens or hundreds of milliseconds. For the cAMP system this occurs in three ways.

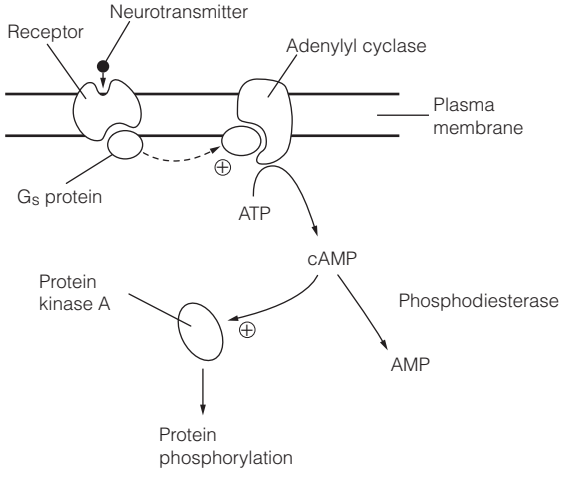


Fig. 3. The adenylyl cyclase-cAMP second messenger system. The activated G_s protein uncouples from the receptor to switch on adenylyl cyclase.

- Cyclic AMP is hydrolyzed to AMP by the action of a specific **phosphodiesterase** in the cytoplasm.
- There are specific **phosphatases** responsible for dephosphorylating the target proteins. Hence the phosphorylation state of a protein at a given time will depend on the balance of the activities of kinases and phosphatases.
- Prolonged occupation of the receptor by the transmitter causes it to **desensitize**. This involves phosphorylation by a specific **kinase** that recognizes the agonist bound form of the receptor followed by the binding of an **arrestin** protein. The resulting complex is unable to recognize the G protein.

Inhibition of adenylyl cyclase

Some neurotransmitter receptors are negatively coupled to adenylyl cyclase. These receptors associate with **G_I proteins** that inhibit the activity of the enzyme. Both their α-subunit and β/γ-subunits independently block the isoform of adenylyl cyclase common in neurons. The outcome is that the activity of adenylyl cyclase, and so the amount of cAMP in the cell at any given instant, will reflect the balance of activation of receptors coupled to G_S and those coupled to G_I.

Phosphoinositide second messenger system

Many receptors are coupled via the **G_q protein** to activation of phospholipase C (*Fig. 4*). This enzyme cleaves a minor phospholipid in the inner leaflet of the plasma membrane, **phosphatidyl inositol-4,5-bisphosphate** (**PIP_2**), to give **diacylglycerol** (**DAG**) and **inositol-1,4,5-trisphosphate** (**IP_3**), both of which are second messengers. DAG, a hydrophobic molecule, diffuses within the lipid where it activates **protein kinase C** (**PKC**). In turn this kinase phosphorylates its protein targets, affecting metabolic, receptor and ion channel functions.

IP_3 is water soluble and freely diffusable in the cytosol. Its target is the **IP_3 receptor**, a large IP_3-gated calcium channel located in the membrane of **smooth endoplasmic reticulum** (**SER**). The SER in neurons (and its equivalent, the **sarcoplasmic reticulum** in muscle cells) acts as an intracellular Ca^{2+} store. The binding of IP_3 to its receptors causes the calcium channels to open and Ca^{2+} flows out of the SER into the cytosol. A rise in intracellular calcium concentration has diverse and widespread effects that are cell typical. An obvious example is that by binding the protein **troponin** in striated muscle, calcium triggers the cascade of biochemical events that leads to muscle contraction. Neurons

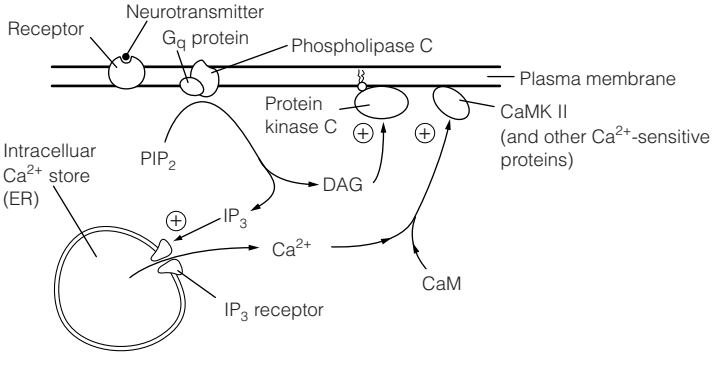


Fig. 4. The phosphoinositide second messenger system. CaM, calmodulin; CaMKII, calcium–calmodulin-dependent protein kinase II; DAG, diacylglycerol; ER, endoplasmic reticulum; IP_3, inositol trisphosphate; PIP_2, phosphatidyl inositol bisphosphate.

contain a calcium-binding protein called **calmodulin** (**CaM**), which shares considerable homology with troponin. On binding Ca^{2+}, calmodulin activates a number of enzymes including **calcium–calmodulin-dependent protein kinase II** (**CaMKII**). CaMKII, and the many other calcium-sensitive proteins, mediate the effects of raised intracellular calcium, such as changes in membrane permeability and gene expression.

D3 Amino acid transmitters

Key Notes

Excitatory amino acids

Glutamate is the major mammalian CNS excitatory transmitter. Microiontophoresis is used to apply low doses of putative transmitters onto CNS neurons so that their responses can be studied. In the glutamate–glutamine cycle, astrocytes remove glutamate from the cleft and convert it to glutamine for re-export to neurons which use it to synthesize transmitter glutamate.

Inhibitory amino acids

GABA is the major inhibitory transmitter in the mammalian CNS. Glycine is important in the spinal cord. GABA is synthesized from glutamate by an enzyme found only in GABAergic neurons. After reuptake from the cleft it is broken down by GABA transaminase. This enzyme is a target for the anticonvulsant vigabatrin.

Related topics

Neurotransmitter inactivation (C7)
Epilepsy (P4)
Strokes and excitotoxicity (P3)

Excitatory amino acids

Glutamate and aspartate are the major CNS excitatory transmitters with glutamate by far the most predominant. It is estimated that 35–40% of synapses use glutamate as a transmitter. Over 90% of nerve cells in the cat spinal cord will respond to the application of low doses of glutamate by **microiontophoresis**. This technique allows delivery of precise amounts of a charged molecule onto the surface of a neuron through a micropipette. In the case of glutamate, which carries a net negative charge at physiological pH, initially a current of a few nA is injected, making the inside of the pipette positively charged so that the glutamate is retained. For delivery of transmitter the current is reversed for a brief period. Most of the major sensory pathways and some motor pathways are glutamatergic (*Table 1*). All pyramidal cells in the cerebral cortex and granule cells in the cerebellar cortex (the most abundant neuron in the mammalian brain) release glutamate.

Neurotransmitter glutamate is synthesized in neurons from glutamine, a reaction catalyzed by **glutaminase**. Glutamate is then pumped into vesicles. After release glutamate is removed from the synaptic cleft by glutamate transporters in neurons and glia. In neurons the glutamate is probably metabolized, although some may be recycled as a transmitter. In glia the glutamate is converted by **glutamine synthetase** to glutamine, which is then liberated into the extracellular space for uptake by neurons. This closes the **glutamate–glutamine cycle** (*Fig. 1*). It allows glial cells (predominantly astrocytes) to export transmitter glutamate to neurons in a form – glutamine – that cannot spuriously activate glutamate receptors.

Inhibitory amino acids

Gamma-aminobutyrate (**GABA**) and glycine are the predominant inhibitory amino acids in the CNS, with glycine confined to the spinal cord. Estimates

Table 1. Major glutamatergic and GABAergic neurons/pathways

Glutamate	GABA
Primary afferents of cranial and spinal nerves	Interneurons of cerebral cortex
Visual system; photoreceptors, bipolar cells, ganglion cells	Interneurons of cerebellar cortex
Dorsal column-medial lemniscus	Cerebellar Purkinje cells
Thalamocortical neurons	Efferents of caudate nucleus, putamen (dorsal striatum)
2nd order neurons of proprioceptor pathways	Efferents of nucleus accumbens (ventral striatum)
Cerebral cortical pyramidal cells; corticopontinecerebellar tract, corticospinal tract	Efferents of globus pallidus and substantia nigra pars reticulata
Hippocampal pyramidal cells	Interneurons of hippocampus
Granule cells of dentate gyrus	Inhibitory dorsal horn cells
Rubrospinal tract	
Lower motor neurons	
Propriospinal neurons	
Cerebellar granule cells	

suggest that 17–30% of the synapses in the mammalian brain use γ-aminobutyrate as a transmitter, making it by far the most important inhibitory neurotransmitter in the CNS. Many pathways involved in motor control are GABAergic, as are most of the interneurons in both cerebral and cerebellar cortices, and the Purkinje cells that provide the entire output of the cerebellar cortex (see *Table 1*).

GABA is synthesized from glutamate by **glutamic acid decarboxylase** (**GAD**), an enzyme virtually exclusive to GABAergic neurons. After release it is taken up by transporters into both neurons and glia. It is catabolized to succinic semi-aldehyde by the mitochondrial enzyme **GABA transaminase** (*Fig. 2*). The GABA analog **vigabatrin**, used in the treatment of epilepsy, is an irreversible

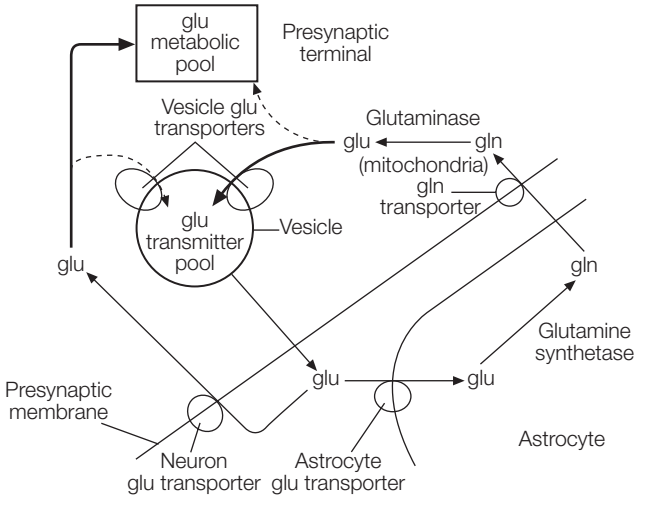


Fig. 1. The glutamate (glu)–glutamine (gln) cycle.

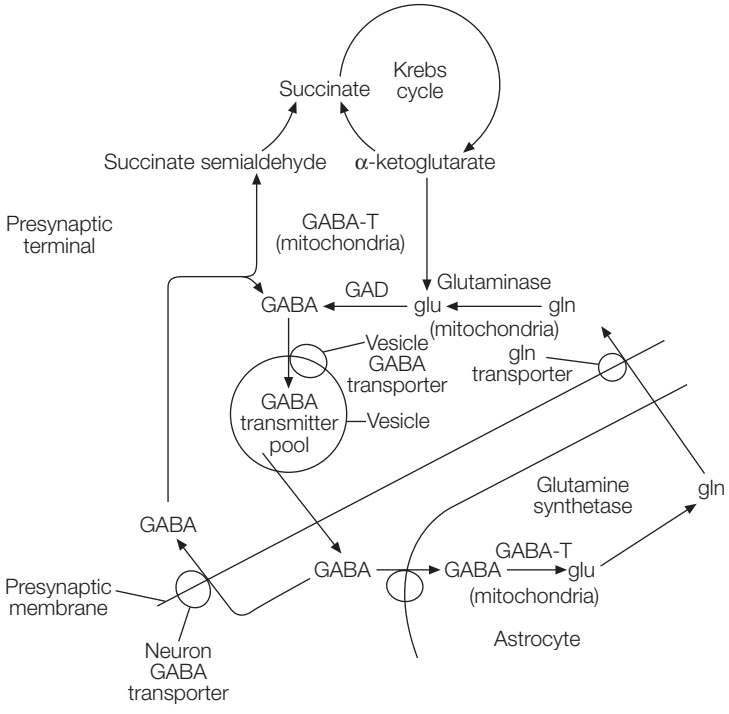


Fig. 2. GABA shunt. GAD, glutamic acid decarboxylase; GABA-T, GABA-transaminase; glu, glutamate; gln, glutamine.

inhibitor of this action and is presumed to be anticonvulsant by increasing the neurotransmitter pool of GABA.

Neurotransmitter **glycine** is synthesized from serine by mitochondrial **serine transhydroxymethylase**. Glycine transporters remove it from the synapse. In the spinal cord **Renshaw cells** express nAChR and are excited by collaterals of motor neurons. Renshaw cells use glycine as a transmitter and inhibit the motor neurons that excite them, among others. This is an example of recurrent inhibition. It serves to dampen the output of motor neurons. The **glycine receptor** resembles the $GABA_A$ receptors and is a Cl^- channel. It is blocked by **strychnine**. Moreover **tetanus toxin** blocks glycine release. Both of these agents are convulsants because they remove Renshaw cell inhibition.

D4 Dopamine

Key Notes

Dopaminergic pathways

The major dopaminergic pathways arise from the midbrain and go to the forebrain. The nigrostriatal tract from the substantia nigra to the striatum contains most of the brain dopamine neurons and is involved in movement. Dopaminergic neurons in the ventral tegmentum project to limbic structures via the mesolimbic pathway and to the cortex by way of the mesocortical pathway. These form a motivation system. Dopamine cells in the hypothalamus control pituitary hormone secretion.

Dopamine synthesis

The catecholamines (dopamine, noradrenaline (norepinephrine), adrenaline (epinephrine)) are synthesized from tyrosine. The first, rate-limiting step which generates L-DOPA is catalyzed by tyrosine hydroxylase. This enzyme is inhibited by catecholamines. This endpoint inhibition is one method by which the synthesis of catecholamines is controlled. L-DOPA is decarboxylated to give dopamine.

Inactivation of dopamine

Synaptic dopamine is taken back into nerve terminals by a high-affinity dopamine transporter. This process is inhibited by amphetamines and cocaine. Dopamine which escapes reuptake is catabolized to homovanillic acid by catechol-O-methyltransferase then monoamine oxidase (MAO). Dopamine free in the cytoplasm is converted to dihydroxyphenyl acetic acid by mitochondrial MAO.

Dopamine receptors

The five metabotropic receptors for dopamine fall into two families. The D1 receptor family (D1 and D5) increase cAMP concentrations, whereas the D2 receptor family (D2, D3 and D4) decrease cAMP concentrations. Generally D1 receptors are postsynaptic, D2 receptors are localized both pre- and postsynaptically.

Related topics

Neurotransmitter release (C5)
Anatomy of the basal ganglia (K7)
Basal ganglia function (K8)
Motivation and addiction (M2)
Schizophrenia (P1)
Parkinson's disease (P5)

Dopaminergic pathways

Dopamine neurons are widely distributed in the nervous system, being found in the retina (amacrine cells), olfactory bulb, adjacent to the ventricles of the brain and in autonomic ganglia. Most dopaminergic neurons, however, are confined to a few nuclei in the brainstem, sending their axons to many regions of the forebrain including the cerebral cortex. These major pathways are illustrated in *Fig. 1*.

About 80% of dopamine neurons are in the pars compacta of the **substantia nigra** (SNpc), which constitutes the A9 group of catecholaminergic cells. (These groups range from A1–A16, the higher the number the more rostrally they are located.) SNpc neurons project to the striatum as the **nigrostriatal pathway**. These cells are involved in basal ganglia regulation of movement and their loss

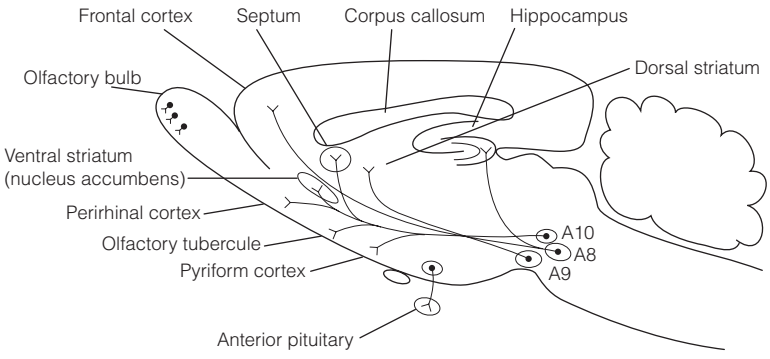


Fig. 1. Major dopamine pathways in a sagittal section of the rat brain. The A8 and A10 group of dopamine neurons give rise to the mesolimbic and mesocortical tracts. The nigrostriatal tract originates in the substantia nigra (A9). A12 neuron axons run in the tuberoinfundibular pathway.

results in Parkinson's disease. Dopamine cell clusters (groups A8 and A10) in the ventral tegmentum of the midbrain project to limbic structures (e.g. nucleus accumbens) or to associated cortical areas (e.g. medial prefrontal and cingulate cortex), giving rise to the **mesolimbic** and **mesocortical** systems respectively. These are implicated in motivation, drug addiction and in schizophrenia. Several small groups of dopaminergic cells in the hypothalamus project axons in the **tuberoinfundibular pathway** to the pituitary to inhibit the secretion of prolactin or growth hormone.

Dopaminergic neurons are small with a thin unmyelinated axon (which arises from one of the dendrites), which bears numerous varicosities along its length. Action potentials in dopamine neurons are long lasting (2–5 ms) and propagated very slowly (0.5 m s^{-1}).

Dopamine synthesis

The precursor for all catecholamine transmitters (dopamine, noradrenaline (norepinephrine), adrenaline (epinephrine)) is the amino acid, L-tyrosine. This is hydroxylated by **tyrosine hydroxylase** to give 3,4-dihydroxyphenylalanine (L-DOPA) which is rapidly decarboxylated by the unspecific enzyme **L-aromatic amino acid decarboxylase** to give dopamine (see *Fig.* 2).

Tyrosine is actively transported into the brain, and the brain concentration of tyrosine is normally enough to saturate tyrosine hydroxylase (TH), so administration of tyrosine cannot alter the rate of dopamine synthesis. The hydroxylation of tyrosine is the rate-limiting step for catecholamine synthesis under basal conditions and TH is subject to regulation:

- increased expression of TH genes, leading to *de novo* synthesis of the enzyme;
- phosphorylation by protein kinases which increases its activity;
- inhibition by catecholamines. This is an example of **endpoint inhibition**.

Dopamine is taken into vesicles by a vesicular monoamine transporter (VMAT), which actively transports catecholamines and serotonin using the efflux of protons from the vesicle to provide the energy. VMATs are blocked by the drug **reserpine** which, by preventing vesicular storage, drastically impairs monoamine neurotransmission. Reserpine has been a useful research tool for investigating the contribution of monoamines to behavior and psychiatric disease.

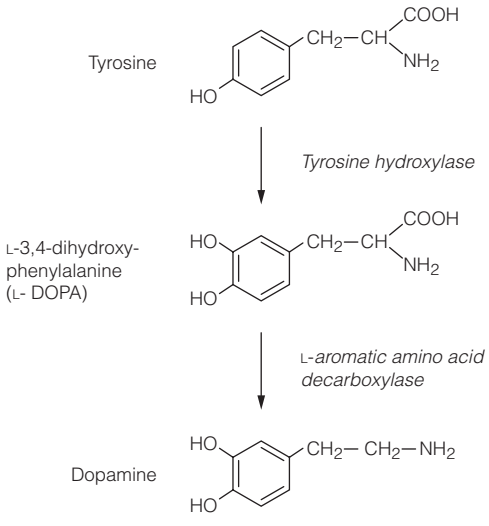


Fig. 2. Synthesis of dopamine from the amino acid tyrosine.

Inactivation of dopamine

Inactivation is by diffusion and reuptake into the nerve cell by a high-affinity Na^+/Cl^--dependent **dopamine transporter** (**DAT**). Neurons that release their dopamine into the hypothalamic–pituitary portal system lack the dopamine transporter. The dopamine transporter is competitively inhibited by amphetamines and by cocaine, which thus potentiate the effects of dopamine at the synapse. This mechanism may underlie the powerful reinforcing properties of these drugs which make them addictive.

Two major enzymes are involved in catecholamine catabolism, although catabolism is not important in inactivating dopamine at the synapse. The primary dopamine metabolites in the CNS are **homovanillic acid** (**HVA**) and **dihydroxyphenyl acetic acid**. In primates, the major catabolic route is via HVA and this is the fate of dopamine released into the cleft which escapes reuptake. It requires the sequential action of **catechol-O-methyl transferase** (**COMT**) and **monoamine oxidase** (**MAO**), both of which are present in neuronal membranes (*Fig. 3*). Cytoplasmic dopamine that is not transported into vesicles and hence remains free in the axon is catabolized by MAO located on the outer membrane of mitochondria, then by **aldehyde dehydrogenase**, a soluble cytosolic enzyme, to dihydroxyphenyl acetic acid.

Dopamine receptors

Five dopamine receptors have been identified and sequenced by genetic engineering techniques. All are metabotropic G-protein-linked receptors and fall into two groups. The D1 family is coupled to G_s and activates adenylyl cyclase to increase cAMP synthesis. It consists of two members, D1 and D5.

The D2 family consists of D2, D3 and D4 receptors. The D2 superfamily is coupled to G_i and inhibits adenylyl cyclase to reduce cAMP synthesis. Both D1 and D2 receptors are located postsynaptically (e.g. in the striatum). In addition, D2 receptors are autoreceptors on dopamine neurons in the substantia nigra and ventral tegmentum where they help to regulate dopamine synthesis. When occupied by dopamine they lower cAMP concentrations and so phosphorylation of tyrosine hydroxylase by protein kinase A falls. This reduces the synthesis of

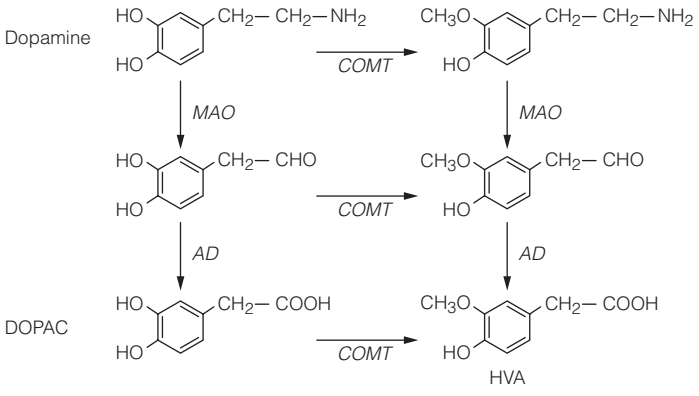


Fig. 3. Metabolism of dopamine. DOPAC, Dihydroxyphenyl acetic acid; HVA, homovanillic acid; COMT, catechol-O-methyl transferase; MAO, monoamine oxidase; AD, alcohol dehydrogenase.

dopamine. D3 receptors are presynaptic autoreceptors. By closing presynaptic Ca^{2+} channels they reduce dopamine release.

The mesocortical pathway differs from the nigrostriatal pathway in terms of its complement of dopamine receptors. Firstly, the mesocortical neurons have no autoreceptors which means they lack the normal regulation of synthesis and release of dopamine. Secondly, the cortex, but not the striatum, expresses D4 receptors. These differences are significant for the treatment of schizophrenia.

D5 Noradrenaline (norepinephrine)

Key Notes

Noradrenergic pathways

Noradrenergic neurons are located in the pons and medulla. The largest group is the locus ceruleus. Noradrenergic axons project via the medial forebrain bundle to most forebrain structures including the cortex, forming wide synapses that allow considerable diffusion of the transmitter. Noradrenergic pathways are probably an arousal system.

Noradrenaline (norepinephrine) and adrenaline (epinephrine) synthesis

Dopamine-β-hydroxylase catalyzes the synthesis of noradrenaline (NA; norepinephrine) from dopamine. In adrenergic neurons in the brain and chromaffin cells of the adrenal medulla, NA is further metabolized to adrenaline (epinephrine).

Noradrenaline (norepinephrine) inactivation

A high-affinity transporter is responsible for reuptake of NA from the synaptic cleft. The transporter is inhibited by tricyclic antidepressant drugs. Compounds structurally related to NA (e.g. tyramine) are taken up by the transporter and they enhance NA release (indirect sympathomimetics) or are subsequently metabolized to weak adrenergic agonists and then released (false transmitters). The enzymes MAO and COMT are responsible for NA catabolism producing 3-methyl-4-hydroxyphenyl glycol, which is then excreted.

Adrenergic receptors

Adrenoceptors are metabotropic receptors activated by NA and adrenaline (epinephrine). α1 receptors are typically postsynaptic and coupled to the IP_3/DAG second messenger system. β2 receptors are presynaptic and reduce cAMP. All β adrenoceptors are coupled to G_s proteins and raise cAMP levels.

Related topics

Pain (F3)
Autonomic nervous system function (L5)
Sleep (M5)
Arousal and attention (O4)
Depression (P2)

Noradrenergic pathways

Cell bodies of noradrenergic neurons are located in the pons and medulla (cell groups A1–A6, except A3). The most caudal groups A1 and A2 send their axons into the spinal cord where they form synapses with the terminals of primary afferents. The others project in two bundles, a dorsal and a ventral bundle, which unite to form the **medial forebrain bundle** (**MFB**) that ascends to supply the hypothalamus, amygdala, thalamus, limbic structures, hippocampus and neocortex. The major noradrenergic cell group is the **locus ceruleus** (**LC**, group A6), which contributes most of the axons of the dorsal noradrenergic bundle and projects to the cerebellum. In the rat, the LC contains some 200 000 neurons (see *Fig. 1*). Noradrenergic neurons are small with fine, highly branched axons that ramify

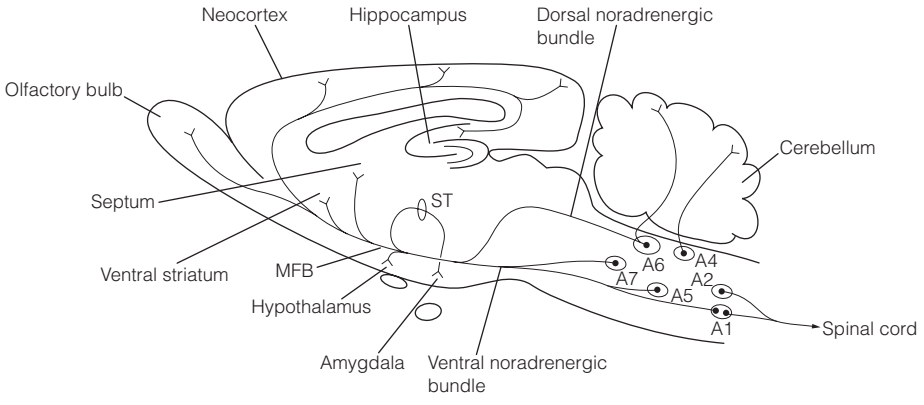


Fig. 1. Major noradrenergic pathways in a sagittal section of the rat brain. A6 is the locus ceruleus. MFB, medial forebrain bundle; ST, stria terminalis.

widely. The axons bear varicosities along their length, but they do not form close synaptic contacts, so noradrenaline (norepinephrine) is released some distance from its targets. This and the wide distribution of its terminals have led to the description of noradrenergic transmission as being a 'neural aerosol'. Firing of noradrenergic cells is low in sleeping animals and increases with arousal level. Hence noradrenaline (norepinephrine) is implicated in controlling sleep–waking cycles and in maintaining arousal. It increases the signal-to-noise ratio of cortical processing.

Noradrenaline (norepinephrine) and adrenaline (epinephrine) synthesis

The first steps in noradrenaline (NA; norepinephrine) synthesis require the synthesis of dopamine from tyrosine. **Dopamine-β-hydroxylase** (**DβH**), an enzyme present in the synaptic vesicle membrane then catalyzes the synthesis of noradrenaline (norepinephrine) (see *Fig. 2*). NA is actively taken into synaptic vesicles by the vesicular monoamine transporter where it is stored together with ATP.

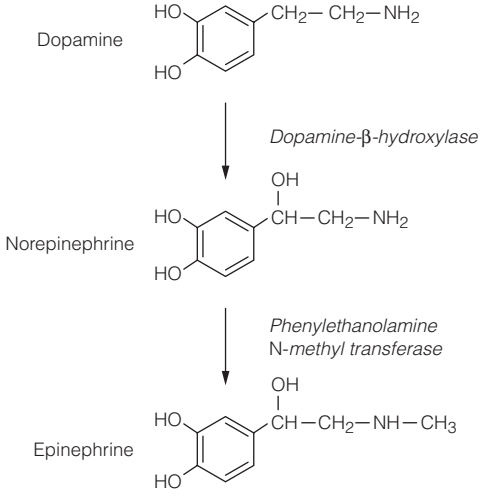


Fig. 2. Synthesis of norepinephrine and epinephrine. These catecholamines, like dopamine, are derived from tyrosine. Early synthetic steps are shown in Topic D4, Fig. 2.

For noradrenergic neurons the reactions stop at this point. However, for the relatively few neurons in the hindbrain that are adrenergic (and for the chromaffin cells of the adrenal medulla), the enzyme **phenyletholamine N-methyltransferase** (**PNMT**) catalyzes the N-methylation of noradrenaline (norepinephrine) to adrenaline (epinephrine).

High activity by locus ceruleus neurons results in increased expression of the tyrosine hydroxylase genes and *de novo* synthesis of the enzyme so that the demand for NA synthesis can be met. The effect of this is that DβH becomes the rate-limiting enzyme rather than tyrosine hydroxylase; thus dopamine and its metabolites may be co-released with NA.

Noradrenaline (norepinephrine) inactivation

Diffusion and reuptake are the key mechanisms removing NA from the synapse. The noradrenaline (norepinephrine) transporter is a saturable Na^+/Cl^--dependent transporter expressed in noradrenergic neurons. It shares homology with the dopamine transporter. The noradrenaline (norepinephrine) transporter is inhibited by the tricyclic antidepressant group of drugs.

The NA transporter does not show a high degree of substrate specificity. Amphetamines, tyramine and other compounds structurally related to NA are taken up by the NA transporter. This inhibits reuptake of NA itself, so its effect at the synapse is prolonged. Furthermore, these compounds displace stored NA from vesicles into the cytoplasm where some of it is degraded by mitochondrial MAO and the rest is released into the cleft via the NA transporter operating in reverse. Because they enhance NA transmission they are referred to as **indirectly acting sympathomimetics**. As repeated doses of amphetamine produce ever greater depletion of stored NA, the quantity required to produce a given effect gets progressively larger. This is an example of tolerance. Some substrates for the NA transporter are metabolized in the terminal and the metabolites stored in the vesicles. When released these **false transmitters** exert weak effects on adrenoceptors.

The metabolic degradation of NA is not important for its inactivation and occurs via different routes in the periphery and CNS. In the CNS, monoamine oxidase (MAO) catalyzes the formation of 3,4-dihydroxy phenylglycoaldehyde which is then reduced to the corresponding alcohol, **3,4-dihydroxy phenylglycol** (**DOPEG**). Finally, this is methylated by COMT to give **3-methoxy,4-hydroxy phenylglycol** (**MOPEG**) which is excreted in the urine. MOPEG excretion has been used as a measure of NA turnover in the CNS.

Adrenergic receptors

Adrenoceptors are metabotropic receptors that are activated by both NA and adrenaline (epinephrine). *Table 1* summarizes the G proteins and second

Table 1. Adrenergic receptors

Receptor	G protein	Second messenger/effector
α1	Gq	IP_3/DAG
	Go	↓ gK
α2	Gi	↓ cAMP
		↑ gK ↓ gCa
β1	Gs	↑ cAMP
		↑ gCa
β2	Gs	↑ cAMP
β3	Gs	↑ cAMP

IP_3, inositol trisphosphate; DAG, diacylglycerol; cAMP, cyclic adenosine monophosphate.

messenger systems that are linked to the different receptors. In the CNS, α2 receptors are presynaptic autoreceptors, where they reduce NA release by reducing cAMP-mediated phosphorylation of N-type Ca^{2+} channels, thus suppressing Ca^{2+} entry. Presynaptic β receptors are also found on noradrenergic terminals in the brain. These facilitate NA release by increasing cAMP-mediated phosphorylation and opening of Ca^{2+} channels. Both excitatory and inhibitory effects of NA release are seen postsynaptically in CNS neurons.

D6 SEROTONIN

Key Notes

Serotonergic pathways

Serotonin neurons are located in raphe nuclei, which lie close to the midline throughout the brainstem. Some axons descend into the spinal cord and inhibit nociceptor input into the spinothalamic tract. Axons run in the medial forebrain bundle to most forebrain structures including the choroid plexus and cerebral blood vessels. Serotonin pathways are implicated in regulating nociceptor input, sleep, anxiety, mood and the control of CSF secretion and cerebral blood flow. Depleted serotonin transmission is associated with depression.

Synthesis of serotonin

Serotonin (5-HT) is synthesized from tryptophan, the plasma concentration of which can affect serotonin levels in the brain. The rate-limiting step in 5-HT synthesis is the hydroxylation of tryptophan catalyzed by tryptophan hydroxylase. The activity of this enzyme increases with neuron firing rate so that transmitter synthesis keeps pace with neural activity.

Inactivation of serotonin

Reuptake of serotonin by a transporter terminates its transmitter action. The transporter is inhibited by tricyclic antidepressants and selective serotonin reuptake inhibitors. 5-HT is catabolized by MAO to 5-hydroxyindoleacetic acid.

Serotonin receptors

Of the many subtypes of 5-HT receptor all but 5-HT_3 receptors are metabotropic. 5-HT_3 receptors are ligand-gated non-specific cation channels. Most 5-HT receptors are postsynaptic but the 5-HT_{1A} subtype is a presynaptic autoreceptor that inhibits serotonin release. Some serotonin receptors are targets for antianxiety agents.

Related topics

Pain (F3)
Brain biological clocks (M4)
Sleep (M5)
Depression (P2)

Serotonergic pathways

Clusters of serotonin neurons (designated B1–B9) are scattered throughout the brainstem, mostly towards the midline in the **raphe nuclei**. Projections into the spinal cord that terminate in the dorsal horn are important in pain sensation by reducing nociceptor input into the spinothalamic tract. Forward projections run into the medial forebrain bundle to go to the hypothalamus, amygdala, striatum, thalamus, hippocampus and neocortex (*Fig. 1*). Some of these serotonergic neurons may be involved in the expression of anxiety. They are inhibited by GABAergic neurons, enhancing the activity of which reduces anxiety. Moreover, chemical destruction of serotonergic neurons in animals reduces behaviors associated with anxiety. Several potent anxiolytic drugs are 5-HT receptor ligands. Hence serotonergic cells comprise a brain anxiogenic (anxiety-generating) system that is presumably important in learning about aversive situations. Serotonin transmission is an important determinant of mood. Deficits are associated with depression and increased risk of suicide. Serotonin is also

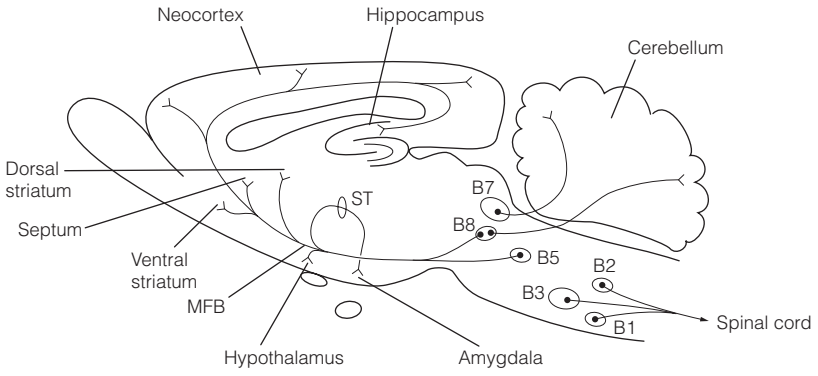


Fig. 1. Major serotonin (5-HT) pathways in a sagittal section of the rat brain. The cell groups B1–B8 correspond to the 5-HT containing raphe nuclei (except B4 and B6). MFB, Medial forebrain bundle; ST, stria terminalis.

involved in sleep. Most brain structures have a serotonergic innervation, including the choroid plexus and cerebral blood vessels, where it regulates CSF secretion and cerebral blood flow respectively.

Synthesis of serotonin

The precursor for serotonin is the amino acid tryptophan. The plasma concentration of tryptophan, which varies according to dietary intake, can alter brain serotonin levels. Serotonin is hydroxylated by **tryptophan hydroxylase** to give 5-hydroxytryptophan (5-HTP) and this reaction is the rate-limiting step in serotonin synthesis. Decarboxylation of 5-HTP by L-aromatic amino acid decarboxy-

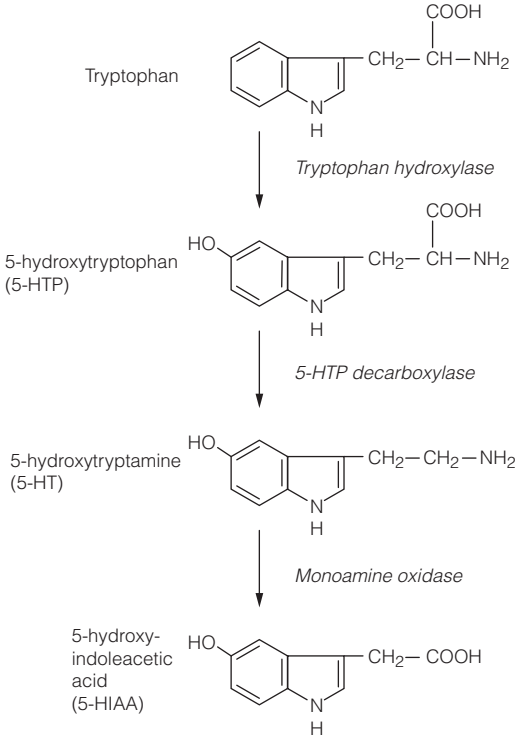


Fig. 2. Synthesis of serotonin from the amino acid tryptophan.

lase (the same enzyme found in catecholaminergic neurons) gives **serotonin**, also referred to as **5-hydroxytryptamine** (**5-HT**), which is an indoleamine (*Fig. 2*).

Serotonin synthesis is matched to the firing frequency of the neuron. Higher firing rates allow increased Ca^{2+}-dependent phosphorylation of tryptophan hydroxylase, the activity of which goes up.

Inactivation of serotonin

Diffusion and reuptake via a saturable Na^+/Cl^--dependent transporter are the prime means by which the action of serotonin in the synapse is terminated. The transporter is inhibited by tricyclic antidepressants and the **selective serotonin reuptake inhibitors** (**SSRIs**). The psychostimulant ecstasy (3,4-methylenedioxymethamphetamine) competes with serotonin for this reuptake system, which may partly explain the euphoria it produces. Animal studies suggest ecstasy may kill serotonin neurons, and negative mood changes have been reported in heavy users. This is worrying given the link between serotonin depletion and depression and the widespread abuse of ecstasy. Oxidative deamination of serotonin by MAO yields its principal metabolite, **5-hydroxyindoleacetic acid** (**5-HIAA**).

Serotonin receptors

There are numerous subtypes of serotonin receptor, all except the 5-HT_3 receptor (which is a ligand-gated ion channel) are metabotropic. *Table 1* summarizes the G protein and second messenger coupling of 5-HT receptors. 5-HT_{1A} autoreceptors inhibit the release of serotonin. Most 5-HT_1, 5-HT_2 and 5-HT_3 receptor subtypes are located postsynaptically. Agents which reduce serotonin transmission (5-HT_2 and 5-HT_3 antagonists, and 5-HT_{1A} agonists such as **sumatriptan**) have proved to be potent anti-anxiety agents.

Table 1. Serotonin (5-HT) receptors

Receptor	Second messenger/ effector	Agonists	Antagonists	Actions
1A	↓cAMP	5-CT[a], 8-OH-DPAT[b], Buspirone	Spiperone, methiothepin	Autoreceptors (agonists are anxiolytic)
1B	↓cAMP	5-CT, Ergotamine	Methiothepin	Presynaptic autoreceptors
1D	↓cAMP	5-CT, Sumatriptan	Methiothepin	Cerebral vasoconstriction (agonists are effective in migraine)
2A	↑IP_3/DAG	LSD[c]	Ketanserin	Excitatory in CNS and PNS
2B	↑IP_3/DAG	α-methyl-5HT		
2C	↑IP_3/DAG	α-methyl-5HT	Methysergide	CSF secretion
3	↑IP_3/DAG Ionotropic cation channel	2-methyl-5HT	Ondansetron	Fast excitatory transmission, located on nociceptors and in area postrema (antagonists are anti-emetic)
4	↑cAMP			Excitatory in CNS and enteric nervous system
7[d]	↑cAMP	5-CT	No selective antagonists	Excitatory in CNS

a 5-carboxamidotryptamine
b 8-hydroxy-2-(di-n-propylamino)tetraline
c Lysergic acid diethylamide
d 5-HT_5 and 5-HT_6 subtypes have been identified but almost nothing is known about them.

D7 Acetylcholine

Key Notes

Cholinergic pathways

Somatic and autonomic preganglionic motor neurons that project from the brainstem and spinal cord are cholinergic. Central cholinergic projections come from three principal sources. The pontine reticular formation sends axons to the spinal cord or forward to forebrain structures. These help regulate sleep and waking. Basal forebrain nuclei make massive connections with the cortex and the septum projects to the hippocampus. Basal forebrain neurons in primates fire in response to presentation of reinforcing stimuli, and their effect is to cause a long-term facilitation of cortical neurons. They produce cortical arousal in response to salient stimuli, thereby facilitating learning and memory.

Acetylcholine synthesis

Acetylcholine (ACh) is produced from acetyl CoA and choline by acetylcholine transferase, a marker enzyme for cholinergic neurons.

Acetylcholine inactivation

ACh is hydrolyzed to choline and acetate in the synaptic cleft by acetylcholinesterase which terminates its transmitter action. Choline is taken back into the nerve terminal by a Na^+-dependent choline transporter.

Acetylcholine receptors

Nicotinic cholinergic receptors (nAChR) are ligand-gated ion channels and muscarinic cholinergic receptors (mAChR) are metabotropic receptors. In the brain nAChR are found in sensory cortex, the hippocampus, and in the ventral tegmentum, where they are presumed to mediate nicotine addiction. Presynaptic nAChR enhance transmitter release. In the spinal cord nAChR are found on the Renshaw cells that mediate recurrent inhibition of motor neurons. Central mAChR are widely distributed with M1 receptors being postsynaptic and M2 receptors being presynaptic where they regulate transmitter release. In the periphery nAChR allow fast transmission in autonomic ganglia and at the neuromuscular junction of skeletal muscle. Muscarinic receptors are present in smooth muscle, cardiac muscle and glands and respond to ACh released from the ANS.

Related topics

- Nerve–muscle synapse (J1)
- Autonomic nervous system function (L5)
- Sleep (M5)
- Physiological psychology of memory (O2)
- Alzheimer's disease (P6)

Cholinergic pathways

Motor neurons in motor nuclei of the cranial nerves and ventral horn of the spinal cord are cholinergic, as are preganglionic autonomic neurons, and the axons of all these cells project into the peripheral nervous system. Cholinergic interneurons are present in the striatum and the nucleus accumbens.

Three regions within the brain contain cholinergic neurons that project centrally. Most caudal are neurons of the **laterodorsal tegmental** and **interpeduncular nuclei** (part of the **pontine reticular formation**) that send axons into the spinal cord or forward to the amygdala, thalamus and basal forebrain. These are important in regulating sleep and wakefulness.

A second region, the basal forebrain, contains the **magnocellular forebrain nuclei**, including the **nucleus basalis of Meynert** (**NBM**) and the **diagonal band of Broca** (**DB**), which project extensively to the cerebral cortex. The third region, the **medial septum**, gives rise to the **septohippocampal pathway** (*Fig. 1*). In primates, cholinergic neurons of the NBM show brief changes in firing rate during behavioral tasks, particularly when reinforcing stimuli are presented. Lesions of the NBM and DB produce impairment in recall for tasks learnt before the surgery and in acquisition of new learning. ACh produces long-term facilitation of neurons in the neocortex and hippocampus by acting at muscarinic receptors to close potassium (K_M) channels, making the cells more likely to fire in response to excitatory inputs. Hence the forebrain cholinergic system seems to be a *selective* arousal system, activated by rewarding or salient events, which facilitates learning. Loss of the cholinergic cells in the basal forebrain is an inevitable finding in Alzheimer's dementia.

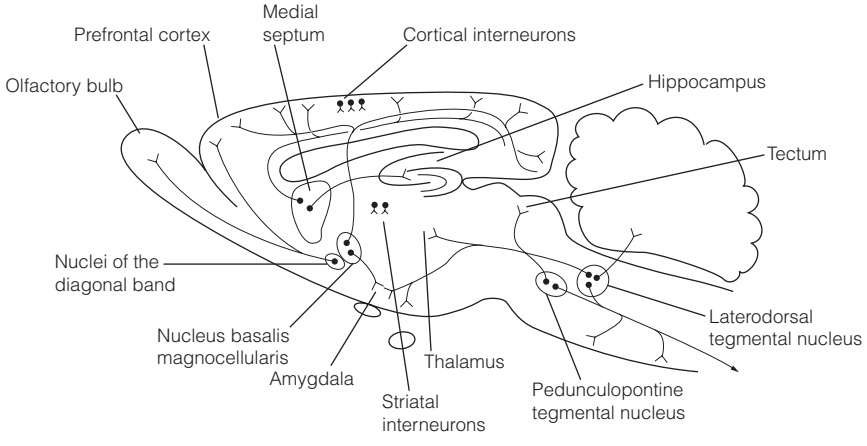


Fig. 1. Major cholinergic pathways in a sagittal section of the rat brain. The nucleus basalis magnocellularis of the rat is known as the nucleus basalis of Meynert in primates.

Acetylcholine synthesis

Synthesis of acetylcholine (ACh) from choline and acetyl CoA is catalyzed by **choline acetyl transferase** (**ChAT**), a cytoplasmic enzyme. Acetyl CoA is derived from glycolysis and must be transported out of the mitochondria of cholinergic neurons. This supply of acetyl CoA, rather than ChAT activity, is thought to be rate limiting for ACh synthesis. Cholinergic neurons express a Na^+-dependent choline transporter, which is saturated at plasma choline concentrations and is responsible for the uptake of choline into neurons. ACh is loaded into vesicles by a transporter that is related to the vesicular monoamine transporters. Interestingly, the ACh vesicle transporter is coded by the first intron of the ChAT gene. Hence synthesis of transporter and enzyme are co-regulated.

Acetylcholine inactivation

The synaptic action of ACh is terminated by its hydrolysis in the cleft to choline and acetate by **acetylcholinesterase** (**AChE**). The liberated choline is recovered by the Na^+-dependent choline transporter. AChE can be secreted in a Ca^{2+}-dependent manner and may act as a neuromodulator in the substantia nigra and cerebellum, where in addition to its catalytic activity AChE enhances the responses of cerebellar neurons to glutamate.

Acetylcholine receptors

Receptors for ACh are either ligand-gated ion channels, **nicotinic receptors** (**nAChR**) or metabotropic, G-protein linked **muscarinic receptors** (**mAChR**) (*Table 1*).

Table 1. Muscarinic receptors

Receptor	G protein	Second messenger/effector
M1	Gq	IP_3/DAG
M2	Gi	↓ cAMP
	Go	↑ gK
M3	Gq	IP_3/DAG
M4	Gi	↓ cAMP

IP_3, inositol trisphosphate; DAG, diacylglycerol; cAMP, cyclic adenosine monophosphate.

Nicotinic receptors are found throughout the CNS. They are important in critical periods of development of the sensory cortex. They mediate fast acetylcholine transmission from the septum to GABAergic inhibitory interneurons in the hippocampus. This is thought to help synchronize the rhythmic firing of hippocampal pyramidal cells. They also mediate fast ACh transmission from brainstem nuclei to the ventral tegmental area, stimulating the dopamine reward pathways. This may be the route by which nicotine is addictive. Presynaptic nicotinic receptors enhance transmitter release at several sites. Interestingly, they potentiate glutamate transmission via NMDA but not AMPA receptors in the prefrontal cortex, suggesting a role in memory.

Slow cholinergic transmission by acetylcholine in the CNS is mediated by muscarinic receptors. Postsynaptic mAChR are commonly M1 while presynaptic autoreceptors are M2 receptors and inhibit the release of ACh. M1 receptors are implicated in the facilitation of cortical neuron responses to excitatory input (by closing K_M potassium channels) and in learning and memory.

In the peripheral nervous system, both nicotinic and muscarinic receptors are involved in cholinergic transmission in autonomic ganglia, whereas muscarinic receptors only are found at the neuroeffector junctions of the ANS; that is, on smooth muscle, cardiac muscle and glands. Nicotinic receptors mediate transmission at the neuromuscular junction between somatic motor neurons and skeletal muscle.

D8 Purines and peptides

Key Notes

Purines

The purine transmitters are ATP and adenosine. ATP acts on ionotropic receptors and is excitatory on smooth muscle and neurons. It is implicated in transmission in the hippocampus, by autonomic neurons, sensory neurons and is implicated in pain signaling. It is inactivated enzymically. Adenosine is not stored in vesicles or released in a calcium-dependent way. It is generated from ATP and ADP and acts on metabotropic receptors. It is probably the molecule responsible for physiological termination of seizure activity. It is inactivated by re-uptake.

Peptides

There are over 50 peptide transmitters. They are grouped by amino acid homology and by derivation from a common precursor. Peptides are packaged into vesicles, subject to posttranslational processing and moved by axoplasmic transport to nerve terminals for secretion.

Tachykinins

Tachykinins are a group of excitatory peptides. They act via receptors coupled to phospholipase C second messenger systems. Substance P is the transmitter of small-diameter primary afferents and implicated in pain transmission and neurogenic inflammation.

Opioids

Opioid peptides are implicated in natural analgesic pathways, sexual and aggressive/submissive behaviors. Falling into three families, the enkephalins, endorphin and dynorphins, they act on metabotropic receptors that produce inhibitory responses by decreasing cAMP.

Related topics

Postsynaptic events (C3)
Pain (F3)
Motivation and addiction (M2)

Purines

Both ATP and its catabolite adenosine are purine transmitters. ATP is stored in synaptic vesicles and co-released with classical transmitters from postganglionic autonomic fibers and central synapses.

Some purinergic receptors for ATP (*Table 1*) constitute a family of ligand-gated ion channels that is distinct from either the nicotinic receptor family or the ionotropic glutamate receptors. They are permeable to Na^+, K^+, and Ca^{2+} and the current has a reversal potential close to zero, which explains why their activation is excitatory. Others are G protein-coupled receptors. ATP is synaptically inactivated by an ecto-5′-nucleotidase.

Examples of ATP transmission include:

- The fast phase of smooth muscle contraction in response to sympathetic stimulation.
- Excitation of dorsal horn cells and motor neurons in the spinal cord by ATP release from primary afferents.

Table 1. Purine receptors

Receptor	Second messenger/ effector	Endogenous ligand	Actions
A_1	↓ cAMP	Adenosine	Presynaptic, inhibition of transmitter release. Postsynaptic inhibition (terminate seizure activity, anxiolytic, hypnogenic)
A_2	↑ cAMP	Adenosine	Heart nociceptors (mediate ischemic pain)
P_{2X}	Ionotropic cation channel	ATP/ADP	Fast transmission in CNS and sympathetic terminals. Located on polymodal nociceptors
P_{2Y}[a]	↑ cAMP, ↑ IP_3/DAG	ATP/ADP	Excitatory

a There are multiple subtypes of P_{2Y} receptors that couple to different G proteins

- In the CA3 region of the hippocampus;
- Nociceptor signaling at a number of sites (e.g. urinary bladder).

Adenosine is an atypical transmitter in that it is not stored in vesicles or released in a Ca^{2+}-dependent way. It is generated locally by enzyme-catalyzed breakdown of released ATP and ADP. Adenosine receptors are G protein-coupled receptors which modulate cAMP. Synaptic actions of adenosine are inactivated by a nucleoside transporter. Adenosine transmission has roles in:

- increasing cardiac muscle blood flow by coronary vasodilation;
- physiological termination of epileptic seizures;
- protecting neurons against the deleterious effects of oxidative stress.

Peptides

Over 50 small peptides are thought to be neurotransmitters. Some are also hormones or neuroendocrines. They can be grouped into families on the basis of:

- similarities in their amino acid sequence;
- being derived by cleavage of a common large precursor polypeptide encoded by a single mRNA molecule. Often the peptides generated from a common polypeptide have related functions. Different cells may process the same precursor or its mRNA in different ways. For example, different peptides are made from proopiomelanocortin by hypothalamic neurons and endocrine cells in the anterior pituitary.

As the mRNA encoding a peptide neurotransmitter is translated on ribosomes, the newborn **prepropeptide** is transported across the membrane of the rough endoplasmic reticulum into the lumen with the aid of a hydrophobic signal sequence at the N-terminal end. The secreted prepropeptide is then cleaved to remove the signal sequence, to give the propeptide. Further proteolysis generates the functional peptide. Peptide-containing vesicles are moved by motor proteins along microtubules, a component of axoplasmic transport, to nerve terminals.

Tachykinins

The first peptide transmitter to be discovered was **substance P (SP)**. It is an excitatory transmitter in several brain regions including the cerebral cortex, striatum and substantia nigra. It is released by both central and peripheral terminals of C

fiber primary afferents. The central terminals synapse with dorsal horn cells to convey information about pain and temperature. Release from the peripheral terminals results in **neurogenic inflammation**. SP-containing terminals are found adjacent to cerebral blood vessels and abnormal release of SP may play a role in migraine and other headaches.

The gene which codes for SP also encodes other transmitters of the tachykinin family, such as **substance K** and **neurokinins A** and **B**.

The three receptors for tachykinins (NK1, NK2 and NK3, where NK is neurokinin) are G protein-coupled receptors linked to phospholipase C. The tachykinins have different affinities for the receptors, with SP the preferred ligand of NK1. Synthetic peptides that are NK1 antagonists, although good analgesics in rodents, are ineffective in humans.

Opioids

The opioids are a group of neurotransmitters which act on opioid receptors, the targets for opiate drugs such as morphine. They are generally co-released with classical transmitters, typically GABA and serotonin, and are usually inhibitory. Opioid transmission is thought to be important in analgesia pathways in the CNS and is also implicated in sexual and aggressive/submissive behaviors. Opioids are encoded by three precursor genes.

- The enkephalin precursor encodes **met-enkephalin** and **leu-enkephalin** (so called because they differ in just one amino acid) and is expressed mainly in short interneurons throughout the brain.
- Proopiomelanocortin encodes **β-endorphin** and is expressed in neurons of the hypothalamus which project to the thalamus or brainstem.
- The dynorphin precursor codes for leu-enkephalin and **dynorphins**.

There are three populations of opioid receptors, the properties of which are summarized in *Table 2.* They are G protein-coupled receptors that allow direct coupling of G proteins to ion channels. By opening K^+ channels and closing Ca^{2+} channels they hyperpolarize neurons.

Remarkably, endogenous morphine has been found in mammalian brain and there is some evidence that it may be a *bona fide* neurotransmitter.

Table 2. Opioid receptors[a]

		μ	δ	κ
Location:	supraspinal	+++[b]	–	–
	spinal	++	++	+
	peripheral	++	–	++
Endogenous ligands:	β-endophin	+++	+++	+++
	enkephalins	+	+++	–
	dynorphin	++	+	+++
Agonists:	morphine, meperidine, fentanyl	+++	+	+/–
Weak agonists:	methadone	+++	–	–
Mixed partial agonists-antagonists:	buprenorphine	+++	–	××
	nalorphine	××	–	++
	pentazocine	×	+	++
Antagonists:	naloxone, naltrexone	×××	×	×××

a All opoid receptors ↓ cAMP and are inhibitory by opening K^+ channels, closing Ca^{2+} channels and inhibiting presynaptic transmitter release.

b +, agonist activity; ×, antagonist activity; –, inactive.

E1 Information representation by neurons

Key Notes

Information coding

The frequency with which a sensory neuron fires conveys information about the timing and intensity of a stimulus. How a sensory neuron is connected encodes the location of a stimulus and its qualitative nature (modality). Motor functions are similarly encoded. Often information is represented by the concerted activity of a number of cells; this is population coding.

Extracellular recording

A technique for recording from single cells or groups of cells in a variety of situations both *in vitro* and *in vivo*, extracellular recording works by amplifying the potentials that arise between a focal electrode close to the neuron(s) and a distant, indifferent electrode.

Related topics

Frequency coding (E2)
Location coding (E3)
Modality (E4)
Cell physiology of learning (O3)

Information coding

Neurons encode information by virtue of two properties. Firstly, the frequency with which a sensory neuron fires conveys information about the **duration** of a stimulus, its **intensity** and how the intensity changes over time. In the same way motor neuron firing rate encodes the timing and force of contraction of a discrete population of muscle fibers. Secondly, the **address** of an afferent neuron, that is, how it is connected via its inputs and outputs, encodes the **spatial location** of a stimulus, and the qualitative nature of the stimulus or **modality.** The address of a motor neuron contributes to the type of movement executed and its direction. In both sensory and motor systems the accurate encoding of a given feature (e.g. skin temperature or the direction of a limb movement) often depends on activity in an array of cells. This is referred to as **population coding**.

Extracellular recording

The firing patterns of either single neurons or clusters of neurons in living animals in response to physiological stimuli are obtained by extracellular recording. This technique uses two fine electrodes usually of tungsten or stainless steel. One, the **exploring (focal) electrode** is placed as close as possible to the surface of the neuron of interest but does not impale it. The second, **indifferent electrode** is placed at a convenient distance. Neuron activity will cause currents to flow between the two electrodes. These currents are amplified and fed to a cathode ray oscilloscope or to the analog to digital port of a computer running software to capture, store and analyze such data. By convention, if the exploring electrode is positive with respect to the indifferent electrode an upward

deflection is recorded. The polarity, shape, amplitude and timing of the recorded waveform generated by neural activity will depend on the position of the electrodes. The closer the exploring electrode is to the neuron the larger the measured signal. Changing the distance between the two electrodes or altering their relative positions will modify all the above parameters. All of this can make extracellular recordings hard to interpret. The technique can be used in brain slices or other *in vitro* preparations, in anesthetized animals, or via **chronically implanted electrodes** (electrodes can be very precisely inserted into the brain under anesthetic, ahead of time, and attached to a connector cemented into the skull), recordings can be made in conscious animals while they are behaving over long periods.

E2 Frequency coding

Key Notes

Static and dynamic coding

Stimulus intensity is encoded by the firing frequency of a neuron. The firing pattern of a sensory neuron is determined by the nature of its sensory receptor and the spatio-temporal characteristics of the stimulus. Slowly adapting receptors cause their afferents to fire at a rate that reflects the size of a constant stimulus so are said to show static responses. Rapidly adapting receptors result in their afferents showing reduced firing to application of a constant stimulus. These afferents respond to the rate of change of stimulus intensity and therefore show dynamic responses.

Stimulus intensity

The relationship between the intensity of a stimulus and the firing frequency of a sensory neuron may be linear or more complicated. For many mechanoreceptors and photoreceptors, the firing frequency is proportional to the logarithm of the stimulus intensity.

Related topics

Sensory receptors (F1)
Touch (F2)
Retina (G3)
Anatomy and physiology of the ear (H2)

Static and dynamic coding

Stimulus intensity is encoded by the mean frequency with which a sensory neuron fires. This is called **frequency modulated (FM) coding**. Broadly speaking, the behavior of afferents falls into two categories depending on the nature of their sensory receptor. **Slowly adapting receptors** respond to a protracted stimulus for as long as the stimulus lasts, causing its sensory neuron to fire repetitively with a frequency that relates to the magnitude of the stimulus. These neurons exhibit **static** (**tonic**) responses to a constant stimulus. In contrast, **rapidly adapting receptors** respond only briefly to a constant stimulus because they soon become insensitive, or adapt, to it. These receptors respond best to *changes* in stimulus intensity. Their afferents show **dynamic** (**phasic**) responses. Many afferents display a mixture of dynamic and static responses. Examples of static and dynamic responses are shown in *Fig. 1* which compares three classes of afferent in the skin that are wired to different types of mechanoreceptor. The **Ruffini organ** is slowly adapting so its afferent has a frequency of firing that is directly proportional to the extent to which overlying skin is indented by a mechanical force. This receptor codes skin *position*. The **Meissner's corpuscle** is rapidly adapting and its afferent fires only when skin displacement is changing with time. It codes the *velocity* with which the skin is displaced. Finally, the **Pacinian corpuscle** adapts so rapidly that its afferents respond to skin *acceleration*. Hence the three afferents between them encode a wealth of dynamical information about the stimulus.

The beginning and end of a stimulus will be signaled by changes in the rates of firing of slowly adapting afferents, and by transient bursts of firing from rapidly adapting afferents. In this way stimulus duration is encoded.

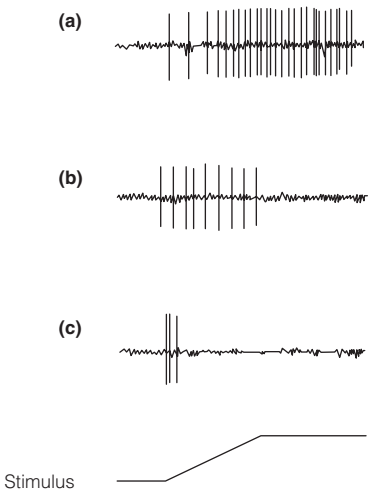


Fig. 1. Static and dynamic afferent neuron responses to skin displacement: (a) static response of Ruffini organ afferent; (b) dynamic response of Meissner's corpuscle afferent to velocity of displacement; (c) dynamic response of Pacinian corpuscle afferent to acceleration.

Stimulus intensity

The relationship between stimulus intensity and response for static cells may be a simple linear one, as for example in skin thermoreceptor afferents. Often however the relationship is more complicated. Commonly, for example, the firing rate is linear with respect to the *logarithm* (log_{10}) of the stimulus intensity. Many skin mechanoreceptors and photoreceptors fall into this category. This type of relationship allows a very wide range of stimulus intensities to be encoded by quite small differences in firing frequency. It has the disadvantage, however, that as intensity gets higher the ability to discriminate differences in intensity falls off.

E3 Location coding

Key Notes

Receptive fields

The region of a sensory surface which when stimulated causes a neuron to respond is the cell's receptive field. In sensory systems, proximal neurons have larger receptive fields than distal neurons because of convergence onto proximal cells of inputs from several distal neurons, and more complex receptive fields because proximal cells can receive inputs from many sources. Many receptive fields show lateral inhibition, in which the cell is excited if this stimulus is directed at the center of the field, but inhibited when directed onto the surround of the field, or vice versa. Lateral inhibition enhances contrast at sensory boundaries.

Topographic mapping

Sensory pathways are organized anatomically so that information about the location of a stimulus in sensory space is preserved. In consequence many structures in the brain contain ordered maps of sensory space. Three broad categories of map exist. Discrete maps are anatomically accurate representations of a sensory surface, though area is usually distorted, and reflect the presence of largely local interactions. Patchy maps have discontinuities which distort anatomical relations and represent interactions between distant parts of the body. Diffuse maps are not ordered by any property of the sensation.

Related topics

Touch (F2)
Retinal processing (G5)
Early visual processing (G6)
Central auditory processing (H4)
Olfactory pathways (I2)

Receptive fields

The spatial location of a stimulus on a sensory surface (skin, retina, etc.) is given by which particular subset of neurons responds. The **receptive field (RF**) of a neuron is the region of a sensory surface which when stimulated causes a change in the firing rate of the neuron. Primary afferents generally have small RFs, the size of which is governed by the distribution of the cluster of sensory receptors which supply the afferent. Receptive fields of neighboring neurons responding to the same type of stimulus tend to overlap.

More proximal neurons in a sensory pathway have receptive fields that are composites of the RFs of more distal neurons. This gives rise to two features: firstly, that proximal neurons have larger receptive fields than distal ones; and secondly, that proximal neurons have receptive fields that are more complicated than those of distal cells.

In general, proximal neurons have larger RFs because several afferents may synapse on a single more proximal (i.e., downstream) neuron. The more proximal a neuron the more complex its receptive field. This is because proximal neurons get inputs from a wider range of sources than distal neurons. This reflects the fact that extensive information processing occurs in sensory systems. Greater complexity of receptive fields also arises as a consequence of an extremely common characteristic of sensory pathways**, lateral (surround)**

inhibition. At its simplest, this is where the RF of a neuron has two zones, a central area and a surround, from which opposite and antagonistic effects are produced in the cell when stimulated. It is seen in somatosensory, visual and auditory pathways. *Fig. 1* shows the receptive field of a somatosensory cell. Stimulation of the center causes an increase in firing so the RF is said to have an excitatory center. Stimulation of the surround reduces firing and is brought about by inhibitory interactions. A cell behaving in this way is described as an **on-center cell**. **Off-center cells** are also common. For the on-center cell, maximum firing rate would be seen with a stimulus that just managed to stimulate the entire center. A larger stimulus that encroached upon the surround would be less effective, by causing some inhibition. In this way lateral inhibition sharpens spatial resolution and enhances contrast at boundaries between stimuli. In skin mechanoreceptor afferents this improves **two-point discrimination**.

By similar means, light–dark contrast at edges is enhanced in the retina, and tone discrimination sharpened by central auditory neurons. In general, lateral inhibition happens between neurons coding the same type of sensation. However, color vision depends on lateral inhibition between cells that respond to different wavelengths.

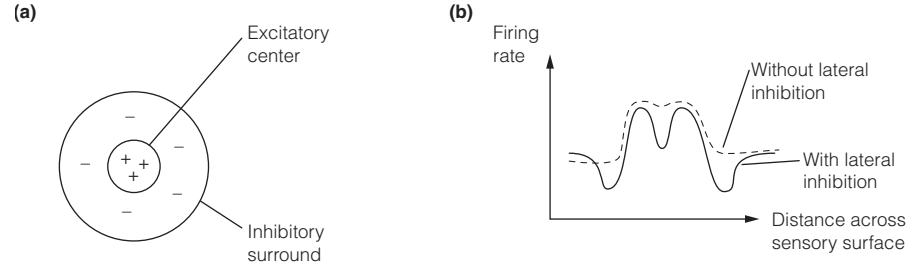


Fig. 1. Lateral inhibition. (a) Receptive field of an on-center sensory neuron showing lateral inhibition; an off-center neuron would have an inhibitory center and an excitatory surround. (b) Contrast enhancement in the presence and absence of lateral inhibition.

Topographic mapping

In most sensory pathways primary afferents are wired to specific subsets of more central neurons in a strictly ordered fashion so that nearest neighborhood relations are conserved. This means that information about stimulus location is not lost in more proximal parts of a pathway. This arrangement is called **topographic mapping**. Receptive fields are aligned to produce an ordered map across brain structures such as the thalamus or the cerebral cortex. These maps are neural representations of a sensory surface or some feature of a sensation. Key examples are: somatotopic maps which represent skin surface; retinotopic maps that reflect the visual fields; and tonotopic maps which represent the pitch of a sound. In addition there are numerous motor maps, particularly in the cerebral and cerebellar cortices in which movements are represented in a systematic way. The motor mapping is preserved in descending pathways so that connections with motor neurons are precisely those needed to execute the mapped movement.

Three broad types of map are recognized, thought to be determined by the extent of the connections between the neurons involved in the mapping:

1. **Discrete maps**, such as somatotopic or retinotopic maps, are anatomically accurate and complete representations of a sensory surface, though they are usually distorted, in that the area of the surface is not faithfully proportioned. Fingers and lips get far more than their fair share of the cortex in somatotopic maps, because they have the highest density of receptors. Discrete maps arise because neurons are connected mostly to their neighbors, allowing **local interactions** between cells. In other words, most of the comparisons the CNS needs to make of, say an image, are between adjacent pixels of retina.
2. **Patchy maps** consist of several domains within each of which the body is accurately represented. However, adjacent domains map regions that are not anatomically close or which are disoriented. Cerebellar motor maps are of this kind and said to exhibit **fractured somatotopy**. Patchy maps arise because while some groups of neurons are locally connected, others are wired to distant neurons allowing **global interactions** to take place. For example, serving a tennis ball requires the coordination of movements in distant parts of the body.
3. **Diffuse maps** are those which have no underlying topography. Distinct smells are mapped to particular sites in the olfactory bulb but not in any orderly fashion. Smells are not arranged within the brain in any systematic way by property.

E4 MODALITY

Key Notes

Sensory receptors

Sensory receptors can be classified in several ways, for example, by their location in the body or by the nature of the stimulus which excites them.

Labeled line

A neural pathway that receives its input from a single class of sensory receptor is a labeled line. It mediates just one type of sensation. Individual receptors within a single class vary in terms of threshold and dynamic range. Stimulus intensity is partly conferred by the number of neurons activated. Compound sensations are the result of simultaneous activation of several labeled lines.

Modality

Modality is the set of all qualitatively similar sensations produced by a single sensory organ. Stimulus quality is probably conferred by the sense organ.

Related topics

Information representation by neurons (E1)

Sensory receptors (F1)

Sensory receptors

Sensory receptors can be classified in a variety of ways. These are summarized in *Tables 1, 2* and *3*.

Labeled line

The neural pathway which is connected to a single class of sensory receptor and which when stimulated gives rise to a readily identifiable type of sensation is called a **labeled line**. The correspondence between receptor class and the nature of the sensation occurs because a sensory receptor responds only to a specific type of stimulus (e.g. photons, pressure). For example, a labeled line exists for

Table 1. Sensory receptors classified by location

Location	Organ/receptor	Sense
Exteroreceptors		
Special	Retina	Vision
	Cochlea	Hearing
	Olfactory epithelium	Smell
	Gustatory epithelium	Taste
	Vestibular inner ear	Balance
Superficial[a]	Cutaneous mechano-, thermo- and nociceptors	Touch, temperature and pain
Proprioceptors		
Deep[a]	Muscle/joint mechanoreceptors	Body position and movement
Interoceptors		
Visceral	Visceral mechanoreceptors	Visceral senses

[a] These are classified as somatosensory receptors.

Table 2. Sensory receptors classified by nature of the stimulus

Receptor	Stimulus	Sense
Photoreceptors	Light	Vision
Mechanoreceptors	Mechanical forces	Hearing, balance, touch Proprioception, visceral stretch
Thermoreceptors	Heat	Temperature
Chemoreceptors	A diverse variety of molecules	Olfaction, taste

Note: Nociceptors are mixed and may be mechano-, thermo or chemoreceptors or **polymodal**, that is responding to several stimuli.

Table 3. A classification of sensation quality

Organ	Modality	Submodality
Retina	Vision	Gray scale brightness, color
Cochlea	Hearing	Tone
Olfactory epithelium	Smell	No agreed primary qualities
Tongue epithelium	Taste	Salt, sweet, sour, bitter
Vestibular inner ear	Balance	Direction of gravitational field, angular acceleration of the head
Muscle and joint mechanoreceptors	Proprioception	
Visceral mechanoreceptors	Visceral stretch	
Skin mechanoreceptors	Touch	Light touch, pressure, vibration/flutter
Skin warm thermoreceptors	Warmth	
Skin cold thermoreceptors	Cold	
Skin/visceral nociceptors	Pain	
Skin itch receptors	Itch	

skin warming. This is because warm thermoreceptors respond optimally to increases in skin temperature. Receptors even in a single class may differ in their properties from each other. Individual receptors may vary in the strength of the stimulus that will make their afferent fire on 50% of the occasions it is delivered, the **threshold stimulus**. Individual warm thermoreceptors respond over different ranges of temperature; that is, they differ in the **dynamic range** over which they operate. The sensation mediated by a labeled line involves activating a population of afferents. This population coding contributes to the intensity of a sensory experience. Many topographic maps are of single or a few closely related labeled lines. Some topographic maps are really a series of embedded submaps, each representing a labeled line.

Many perceived sensations do not correspond to what can be produced by activating a single labeled line. These **compound sensations** arise from the activation of several receptor types by a single stimulus. By this means a rich variety of higher-order sensory experience is made possible (e.g. texture, wetness).

Modality

The concept of **modality** is ill-defined in the neuroscience literature. One common definition is that modality is the group of qualitatively similar sensations detected by a particular sensory organ and recognizes **submodalities** based on perception. In this scheme, some submodalities correspond to a

receptor class, others encompass a diversity of receptors. A version of this approach is given in *Table 3*.

Experiment suggests that stimulus quality is determined by the sense organ. Surgically re-routing visual pathways to auditory cortex resulted in animals which behaved as if they interpreted input into the redirected pathway as light, not sound. This also suggests that sensory cortex may be a rather general-purpose machine.

E5 Elementary neural circuits

Key Notes

Neural networks The operation of groups of interconnected neurons is not well understood. This is a problem because much of the nervous system consists of such neural networks.

Divergence Neural pathways in which few neurons upstream connect to many neurons downstream exhibit divergence. It allows signals to be spread to many targets.

Convergence Neural pathways in which many upstream neurons connect to few downstream ones exhibit convergence. It allows data compression and integration of weak signals.

Feedforward In feedforward signals flow downstream from lower-order to higher-order neurons; i.e. in the input-to-output direction. Feedforward inhibition is responsible for the surround inhibition in sensory pathways and the reciprocal inhibition seen in motor reflexes.

Feedback In feedback circuits signals flow upstream. They can be excitatory but are more usually inhibitory when their effect is negative feedback. Neurons that feedback onto themselves do so via recurrent axon collaterals.

Central pattern generators Neural circuits that generate cyclical patterns of neural activity, such as respiration or locomotion are called central pattern generators.

Related topics

- Location coding (E3)
- Spinal motor function (J4)
- Brainstem postural reflexes (J5)
- Cerebellar function (K6)
- Control of autonomic function (L6)

Neural networks

Neurons connect together to form networks. The operation of neural networks, even those containing just a few distinct types of nerve cell, is poorly understood in general. This is arguably the most serious problem for contemporary neuroscience because large regions of the nervous system (e.g. the cerebral cortex) apparently consist of the same circuit repeated millions of times. How the same circuit serves functions as diverse as those of the cerebral cortex is currently a mystery. However some patterns of neural organization are both common and comprehensible.

Divergence

Few cells connecting with many downstream is termed **divergence** (*Fig. 1a*). It serves to disseminate information to a wide variety of targets. Examples include:

- Primary afferents which relay with many interneurons so that other inputs and motor output in the cord and brainstem can be modified.
- Small numbers of preganglionic autonomic neurons supply up to 100-fold greater numbers of postganglionic neurons.

Convergence

The funneling of connections from many cells to few is called **convergence** (*Fig. 1a*). It is the means by which target cells are able to integrate information from several sources. Convergence must involve data compression. Examples include:

- The retina, which has 100 million photoreceptors but only one million output neurons.
- The spinal cord, where motor neurons are outnumbered by primary afferents about 10-fold.

Low convergence is seen where high **spatial resolution** (the ability to sense stimuli that are close together as independent) is important, such as between cones and bipolar cells in the retina. In contrast, high convergence is required where it is necessary to integrate weak signals from a number of receptors to achieve high sensitivity. This is the case between rods and bipolar cells in the retina, where it permits vision in dim light.

Feedforward

In feedforward circuits (*Fig. 1b*), input neurons establish connections (either excitatory or inhibitory) with cells that are closer to the output (i.e. higher order) neurons than themselves. In **feedforward inhibition** lower-order cells excite

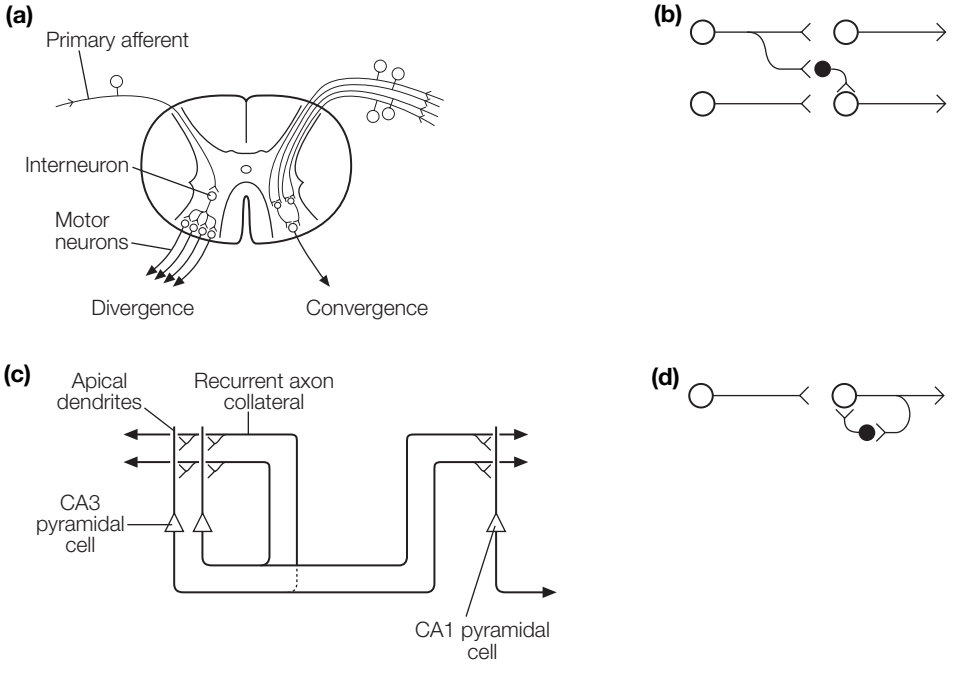


Fig. 1. Basic neural circuits. (a) Divergence and convergence in the spinal cord. (b) Feedforward inhibition. (c) Recurrent excitation in the CA3 region of the hippocampus. (d) Feedback (recurrent) inhibition. Excitatory neurons are open circles, inhibitory neurons are filled circles.

inhibitory interneurons which project forward to dampen the activity of neighboring higher-order cells. This results in only the strongest signals being propagated further. Feedforward inhibition, in the form of GABAergic interneurons, is responsible for generating the surround inhibition seen in sensory pathways, and may also contribute to **selective attention**, the facility to attend to one stimulus in preference to others.

A special case of feedforward inhibition is the enhancing of a response by attenuating an opposing action, a mechanism known as **reciprocal inhibition**. This operates in spinal cord reflexes that time the activities of limb flexors and extensors. Here, inhibitory interneurons reduce the activity of extensor muscles during flexion and vice versa.

Feedback

In feedback circuits, higher-order cells establish connections to lower-order cells. The connections may be excitatory but are more usually made via inhibitory interneurons to cause feedback inhibition, the neural equivalent of negative feedback. Feedback inhibition allows motor systems to correct errors during the execution of a movement. A neuron may feedback on itself by making recurrent connections. **Recurrent excitation** (*Fig. 1c*) by axon collaterals is important in the hippocampus, while **recurrent inhibition** (*Fig. 1d*) of motor neurons in the spinal cord by Renshaw cells is crucial.

Central pattern generators

Neural networks that produce cyclical patterns of activity autonomously are called **central pattern generators** (**CPGs**). They mediate, for example:

- the inspiratory–expiratory cycle of ventilation;
- limb movements during locomotion which involves alternate activation of flexors and extensors.

The basic operation of CPGs is modified or overridden by extraneous pathways.

F1 Sensory receptors

Key Notes

Receptor potentials

Sensory receptors respond to a stimulus with a change in membrane potential, a receptor potential. In vertebrates this is depolarizing at all receptors, except photoreceptors which hyperpolarize, and inner ear hair cells which make both responses. A cutaneous sensory receptor is part of a primary afferent. In other sensory systems the sensory receptor is a separate receptor cell. Receptor potentials are small amplitude, graded, passively conducted potentials that decay with time and distance (c.f. synaptic potentials). Receptors adapt to a constant stimulus in that the response declines with time. Receptor potentials that are sufficiently large will trigger action potentials in sensory pathways; if they do so directly they are called generator potentials.

Mechanoreceptors

Skin mechanoreceptors respond to mechanical forces. They are classified as slowly or rapidly adapting and within each of these they fall into two types. Type I have small receptive fields (RFs) with clear boundaries and are concerned with shape and texture sensation. Type II have large RFs with fuzzy edges. The density of receptors is variable, being highest in fingertips and lips. Skin regions with a high density of receptors have a bigger representation on somatotopic maps than regions with low density.

Thermoreceptors

Thermoreceptors are slowly adapting. Warm receptors increase firing in response to a rise in skin temperature, whereas cold receptors respond to a decrease in temperature. Individual thermoreceptors are most sensitive to rapid changes in temperature but are poor at signaling absolute temperature. Temperature perception relies on comparing the responses of populations of warm and cold receptors.

Nociceptors

Mechanical nociceptor afferents are Aδ fibers responsible for the sensation of sharp pricking pain. Thermal nociceptors respond to high or low skin temperature activating Aδ afferents. Polymodal nociceptor afferents are C fibers and respond to intense mechanical forces, heat and a number of chemicals released during tissue damage. They cause the sensation of poorly localized burning pain. Several inflammatory mediators lower the threshold of polymodal nociceptors, resulting in primary hyperalgesia. Itch receptors have C fiber afferents and respond to histamine.

Clinical pharmacology of nociceptors

Glutamate and peptides (e.g. substance P) are released from the central and peripheral terminals of nociceptor afferents. Peripheral release is responsible for changes in blood flow and histamine secretion that causes the redness, heat and swelling characteristic of neurogenic inflammation. Blocking the actions of mediators that excite or sensitize nociceptors can produce clinically useful analgesia. This is the case for bradykinin receptor antagonists, non-steroidal anti-inflammatory analgesics (e.g. aspirin), which inhibit the synthesis of prostaglandins, and vallinoid receptor ligands which act on the ion channels that transduce noxious heat.

Related topics Information representation by neurons (E1) Modality (E4) Pain (F3)

Receptor potentials

Sensory receptors convert the stimulus to which they are sensitive to a change in membrane voltage by making the membrane more permeable to one or more ions. This process is called **transduction** and is different in different receptors. For somatosensory systems the sensory receptor is the modified ending of the primary afferent neuron, and is depolarized directly by the stimulus. In all other sensory systems the sensory receptor is a specialized cell type, which forms synaptic connections with the first afferent neuron. For these, alterations in membrane potential translate into changes in sensory cell neurotransmitter release, with corresponding effects on the primary afferent. In vertebrates, all sensory receptors, except photoreceptors, depolarize in response to stimulation. Photoreceptors hyperpolarize when exposed to light. The hair cells of the inner ear responsible for balance and hearing depolarize or hyperpolarize depending on the stimulus.

The stimulus-evoked change in membrane potential is called a **receptor potential**. In some sensory systems, for example somatosensory systems, the effect is to trigger actions potentials if the stimulus is sufficiently strong. Receptor potentials which directly generate action potentials are often referred to as **generator potentials**. Receptor potentials share many of the properties of synaptic potentials (*Fig. 1*).

They are small amplitude, graded in size depending on the stimulus strength, passively conducted over the receptor cell surface or along neurites, decay with time and distance and can be summated. A generator potential will trigger action potentials for as long as it remains beyond the firing threshold, the frequency of firing will be higher the greater its amplitude. Receptors demonstrate **adaptation**, a decline in response over time to a constant stimulus, and are classified as **slowly** or **rapidly adapting**. Cutaneous receptors are classified as mechanoreceptors, thermoreceptors and nociceptors. Their properties are summarized in *Table 1*.

Table 1. Cutaneous receptors

Receptor	Adaptation		Fiber type	Sensation
Mechanoreceptors				
Meissner's corpuscle	RA1	velocity	Aβ	Touch, flutter, stretch
Pacinian corpuscle	RA2	acceleration	Aβ	Vibration
Merkel's disc	SA1	velocity and displacement	Aβ	Touch, pressure
Ruffini corpuscle	SA2	displacement	Aβ	Stretch
Lanceolate ending[a]	RA1	velocity	Aα	Hair movement
Pilo-Ruffini ending[a]	SA2	displacement.	Aβ	Hair movement
Hair follicle receptor[a]	RA1	displacement.	Aβ	Hair movement
Thermoreceptors				
Warm, bare nerve ending	SA		C	↑Skin temperature
Cold, bare nerve ending	SA		Aδ	↓Skin temperature
Nociceptors				
Mechano- bare nerve ending	Nonadapting		Aδ	Sharp pain
Polymodal bare nerve ending	Nonadapting		C	Burning pain

[a] Hairy skin only.

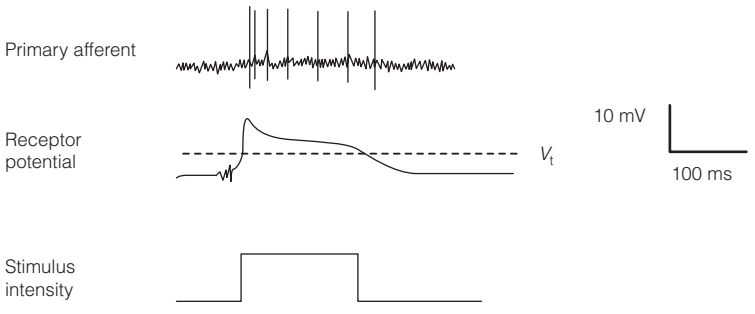


Fig. 1. Receptor (generator) potential (middle trace) and discharge (upper trace) of a slowly adapting cutaneous mechanoreceptor afferent in response to 150 ms indention of skin (lower trace). V_t*, threshold voltage.*

Mechanoreceptors Skin mechanoreceptors are classified as slowly or rapidly adapting and, separately, as being of two types, type I and II, distinguished by their location and receptive fields (RFs). Type I are superficial, lying at the boundary of epidermis and dermis and have small RFs with well-defined boundaries. These include **Meissner's corpuscles** (*Fig. 2a*) and **Merkel's disks** (*Fig. 2b*). Type II are deep in the dermis and have large RFs with poorly defined edges, and include **Ruffini corpuscles** (*Fig. 2c*) and **Pacinian corpuscles** (*Fig. 2d*).

Type I receptors are more directly concerned with form and texture perception than type II. The density of type I receptors varies across the body surface being highest in the fingertips, lips and tongue and lowest in the trunk. Areas with higher density have proportionally greater representations in somatotopic maps. Receptor convergence varies with receptor; whereas each Merkel's disk afferent receives input from 2–7 receptors, a one-to-one ratio is the case for Pacinian corpuscles and their afferents.

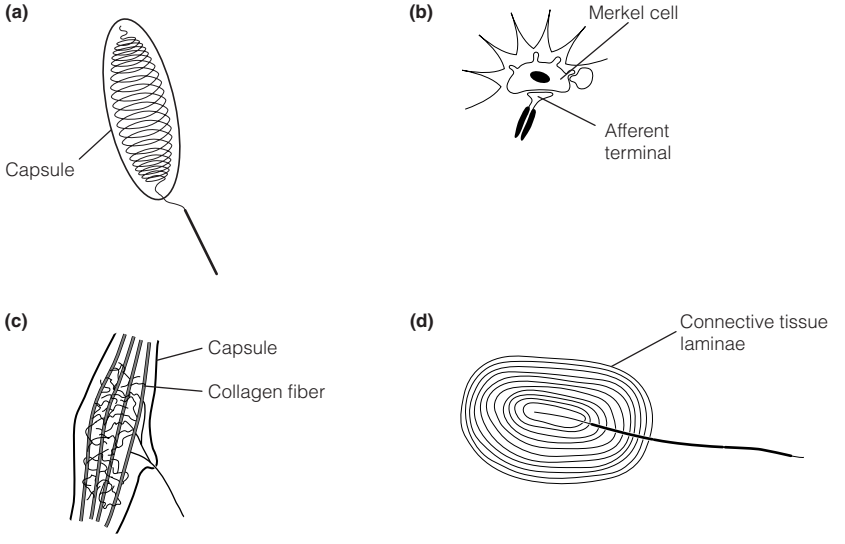


Fig. 2. Morphology of glabrous (non-hairy) skin mechanoreceptors: (a) Meissner's corpuscle; (b) Merkel's disc; (c) Ruffini corpuscle; (d) Pacinian corpuscle.

Meissner's corpuscle signals are important in adjusting grip force since they are very sensitive to small movements of an object over the skin of a grasping hand. Human skin is sensitive to vibration over a wide range of frequencies (5–500 Hz). For frequencies < 40 Hz the term flutter is used and this sensation is largely attributable to Meissner's corpuscles. Higher-frequency vibration is detected by Pacinian corpuscles which follow a sinusoidal vibration stimulus by triggering a single action potential per period. Optimal sensitivity is about 200 Hz. Frequencies in this range can be perceived even at skin indentations less than 1 mm.

Transduction has been most extensively studied for Pacinian corpuscles. Skin indentation force is transmitted through the corpuscles to deform the neurite within. This causes the opening of stretch-sensitive Na^+ channels in the membrane (not to be confused with voltage-dependent Na^+ channels) resulting in a brief depolarization. Membrane potential returns to resting extremely fast because the receptor adapts. Adaptation is brought about by the corpuscle, which consists of concentric layers of connective tissue. An applied force is rapidly dissipated as shear force as the layers slide over each other, cutting off deformation of the neurite within.

Human skin is either **hairy** or **glabrous** (non-hairy). Innervation of hairy skin differs in having a lower density of Merkel's disks and in possessing two additional types of mechanoreceptor closely associated with hairs (*see Table 1*).

Thermoreceptors

Skin thermoreceptors are the naked terminals of small-diameter afferents. They are slowly adapting, tonically active, more sensitive to rapid than slow temperature changes, and poor indicators of absolute temperature. Thermoreceptor afferents get input from 3–4 receptors, and have very small RFs (1 mm diameter in glabrous skin), yet surprisingly infrared radiation is very poorly localized. There are two populations of thermoreceptors (*Fig. 3*).

- Warm receptors fire in the range 29–45°C, with peak frequency at 45°C. Noxious heat is not sensed by warm receptors.
- Cold receptors are sensitive to skin temperatures between 5–40°C and have maximal activity at 25°C.

Perceived skin temperature is given by comparing the relative activities of the warm and cold receptors; an example of population coding. Thermoreceptors also signal the *direction* in which temperature changes. If the skin is cooled, warm receptors are briefly silenced and cold receptor firing rates rise rapidly and then drop slightly as they adapt to the tonic frequency that codes the new

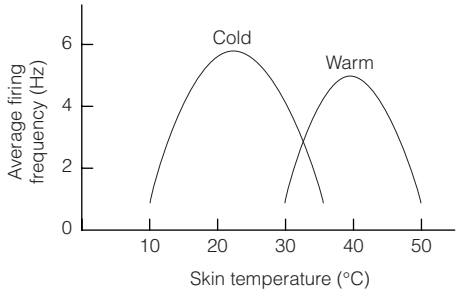


Fig. 3. Frequency response of populations of cutaneous cold and warm thermoreceptors.

temperature. In like manner, skin warming silences the cold receptors and boosts warm receptor firing.

Nociceptors

The bare endings of small-diameter high-threshold afferents are *nociceptors*, receptors for noxious, pain-producing stimuli. Unlike thermoreceptors, normally they have no background firing. They are classified by what excites them:

1. **Mechanical nociceptors** are stimulated by intense mechanical forces and those in the skin give rise to sharp, pricking pain. Each nociceptor is the ending of one of 5–20 branches of an Aβ afferent with a conduction velocity of 5–30 m s^{-1}. Similar receptors are also found in the visceral peritoneum that invests the gut, where they respond to excessive distension.
2. **Thermal nociceptors** fall into two groups, one excited either by temperatures greater than 45°C, the other by temperatures less than 5°C. They also respond to intense mechanical stimuli. Their afferents are Aδ or C fibers.
3. **Polymodal nociceptors** respond to puncture, temperatures in excess of 48°C, and to a wide variety of molecules liberated as a result of tissue damage including: K^+, H^+, bradykinin, prostaglandins, serotonin and histamine. Stimulation of these nociceptors causes burning or aching pain that is poorly tolerated. Their afferents are C fibers which conduct at less than 1.0 m s^{-1}. Because C fiber conduction is so slow, the pain they produce arrives last after a mechanical injury. They are also responsible for visceral pain and toothache. Release of bradykinin and prostaglandin E_2 from damaged tissue reduces the threshold of nociceptors to mechanical and thermal stimuli making the site of an injury more painful, a phenomenon called primary hyperalgesia. Even non-noxious stimuli may elicit pain. It is postulated that some visceral C fiber nociceptors are silent even in the face of noxious mechanical or thermal stimuli until sensitized by a chemical mediator.
4. **Itch receptors** belong to a separate class of C fiber that responds to histamine released from mast cells.

Clinical pharmacology of nociceptors

Primary nociceptor afferents co-release glutamate (which acts as a fast transmitter) and peptides, most notably **substance P**. The peptides enhance and prolong the effects of the excitatory amino acid because they diffuse away from their site of release to increase the excitability of neighboring dorsal horn cells. Glutamate and peptides are released from the peripheral nerve terminals as well as the central endings of the nociceptor afferent axon. Action potentials triggered by exciting polymodal nociceptor terminals are not only conducted centrally but, in what is termed an **axon reflex**, can also travel **antidromically** along neighboring branches to stimulate secretion of substance P from their peripheral terminals. This contributes to the classic signs of inflammation at an injury site in what is termed **neurogenic inflammation**. Substance P acts on post-capillary venules to produce vasodilation (heat and redness) and increased permeability (swelling), and can cause itching by liberating histamine from mast cells which excites itch C fibers. Surprisingly, antagonists of the neurokinin (NK1) receptor at which substance P is most efficacious have not proved to be good analgesics in humans.

Bradykinin is a potent excitant of nociceptors. Competitive antagonists of bradykinin (B_2) receptors are analgesic and anti-inflammatory in the laboratory and have potential as novel analgesics.

Cell damage in the periphery causes the production of arachidonate. Cyclooxygenases act on this phospholipid, producing cyclic endoperoxides

from which **prostaglandins** are synthesized by prostaglandin synthases. Prostaglandins **sensitize** (lower the threshold of) nociceptors to other inflammatory mediators such as serotonin and bradykinin. **Aspirin** and other **non-steroidal anti-inflammatory dugs** (**NSAIDs**) work by inhibiting cyclooxygenases, particularly cyclooxygenase 2 (COX 2). Aspirin produces an irreversible inhibition by acetylating the enzymes (i.e. a covalent modification), whereas many NSAIDs, e.g. ibuprofen, are reversible competitive inhibitors. These drugs are particularly effective in pain associated with inflammation, e.g. toothache.

Capsaicin, the active compound responsible for the hot taste of chilli peppers, acts on vanilloid receptors in thermal and polymodal nociceptors. **Vanilloid receptors** are ligand-gated ion channels with a high conductance to Ca^{2+} and are the molecular receptors which transduce noxious heat stimuli. Capsaicin causes pain by releasing substance P from nociceptors, but repeated application causes depletion of the transmitter and hence a reduced sensitivity to nociceptor stimuli. Recovery takes days to weeks. Vanilloid receptor ligands are being explored as potential novel analgesics.

F2 Touch

Key Notes

Dorsal column–medial lemniscal pathway

Each dorsal root receives input from a skin dermatome. Mechanoreceptor and proprioceptor afferents enter the dorsal roots to synapse with interneurons involved in spinal reflexes in the dorsal horn. A branch from each afferent ascends the spinal cord in the dorsal columns to synapse with neurons in the dorsal column nuclei (DCN) in the medulla. Lateral inhibition occurs in the DCN. Axons of DCN neurons cross the midline and ascend in the medial lemniscus to terminate in the ventroposterolateral thalamus. From here neurons project to the primary somatosensory cortex (SI). Somatotopic mapping at each stage preserves stimulus location (the somatosensory cortex has several somatotopic maps over its surface, each representing a different class of receptor), and the dynamic features of stimuli are transmitted faithfully to the cortex. The primary somatosensory cortex (SI) is organized into radially arranged columns, each of which gets input from a single type of receptor in a particular place on the skin. Adjacent regions of skin have adjacent columns. SI is concerned with tactile discrimination and stereognosis, the ability to perceive shape by touch. The secondary somatosensory area (SII) gets input from both sides of the body and is involved in guiding movement in the light of somatosensory input.

Descending connections

Reciprocal connections between the somatosensory cortex and DCML system nuclei are formed which have the same somatotopic mapping as the ascending pathway. These descending connections probably act to filter somatosensory inputs.

Related topics

Location coding (E3)
Modality (E4)
Sensory receptors (F1)
Pain (F3)

Dorsal column–medial lemniscal pathway

The region of skin innervated by a dorsal root is a **dermatome**. These are numbered for the spinal cord segment served by the dorsal root (e.g. T1–T12). Cutaneous low-threshold mechanoreceptor primary afferent axons relaying skin mechanoreceptor and proprioceptor signals, enter the spinal cord via the dorsal roots to synapse with interneurons, **dorsal horn cells (DHCs**), in Rexed laminae III–VI. These neurons mediate or modify spinal reflexes. Each afferent sends a collateral up the dorsal columns to synapse with neurons in the **dorsal column nuclei (DCN)** in the medulla. The **cuneate nucleus** receives input from C1–8 and T1–6, whereas the **gracile nucleus** gets its inputs from T7–12, lumbar and sacral spinal segments. Dorsal column nuclei are a site for lateral inhibition (*Fig. 1*).

Axons of the DCN neurons cross the midline to ascend on the opposite side of the spinal cord as the **medial lemniscus**, terminating in the **ventroposterolateral (VPL)** division of the ventrobasal thalamus (*Fig. 2*). VPL neurons give rise

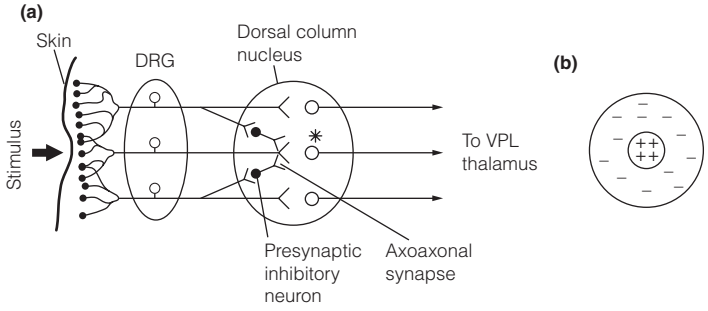


Fig. 1. Lateral inhibition in the dorsal column nucleus: (a) circuitry, the dorsal column neuron marked by an asterisk has a receptive field with inhibitory surround because transmitter release from its primary afferents is reduced by presynaptic inhibitory neurons driven by afferents in the surrounding skin; (b) the receptive field of the indicated cell.

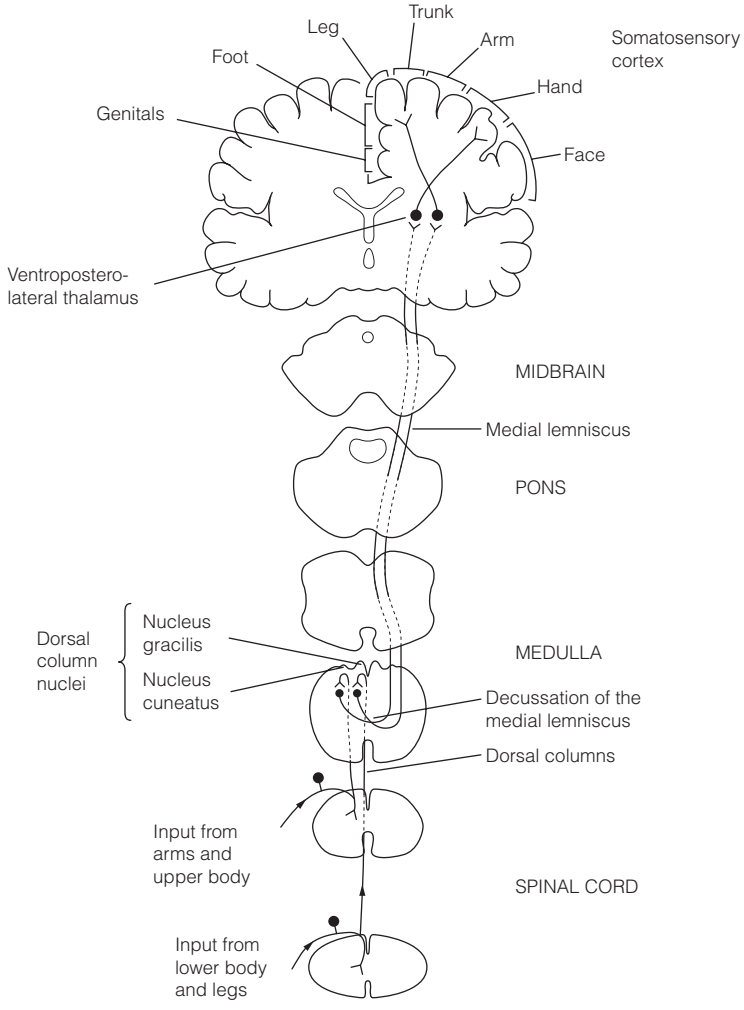


Fig. 2. The dorsal column–medial lemniscal system. All neurons shown are excitatory.

to thalamo-cortical axons, which project to the **primary somatosensory cortex** SI (Brodmann's areas 1, 2, 3a and 3b), situated over the postcentral gyrus. SI neurons in turn project to SII (*Fig. 3a*).

The general properties of the **dorsal column medial lemniscal** (**DCML**) **system** are:

- The great strength of its synaptic connections.
- The properties of its neurons are matched to those of the sensory receptors supplying them, so the dynamic features of stimuli are transmitted with high fidelity through the whole system.
- Somatotopic mapping preserves localization at every stage. Body maps are found in dorsal column nuclei, VPL and somatosensory cortex. Each of the four regions in SI has a distinct map. Cutaneous input maps to the core of the VPL and then to areas 1 and 3b whereas proprioceptor input maps more peripherally in the VPL thalamus and then to areas 2 and 3a (*Fig. 3b*).

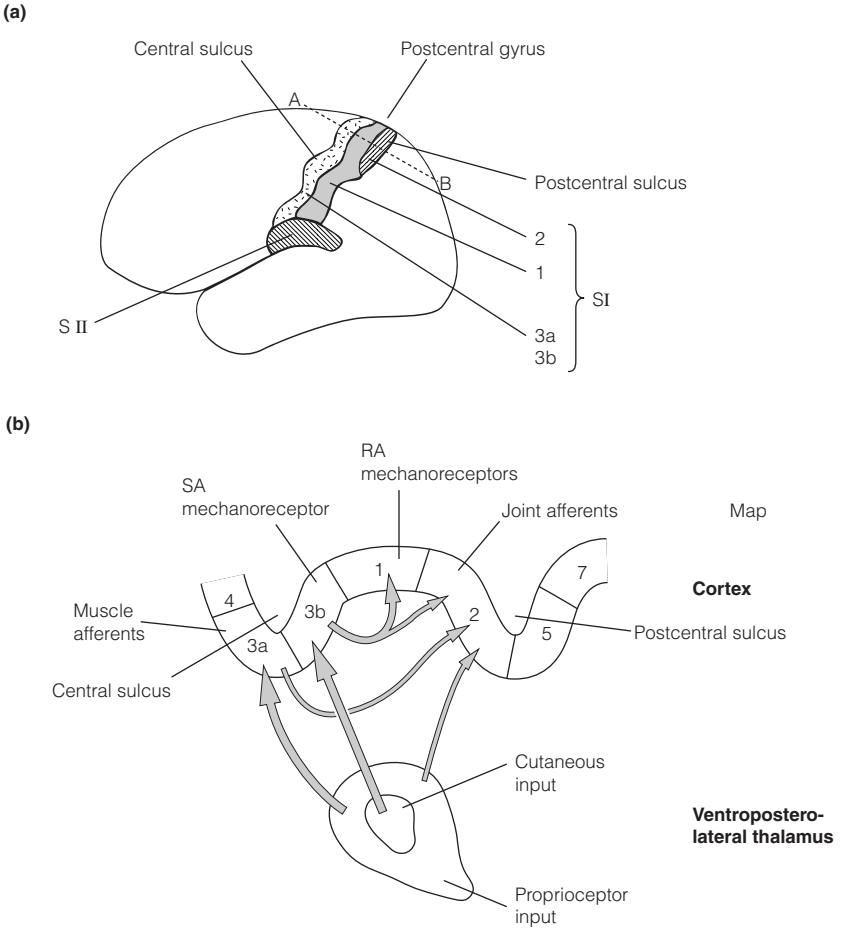


Fig. 3. The somatosensory cortex: (a) location of the primary (SI) and secondary (SII) somatosensory cortex of the left cerebral hemisphere, lateral aspect. Numbers refer to Brodmann areas; (b) interconnections of the thalamus and somatosensory cortex viewed across line A–B in (a).

Neurons in SI are organized into radially arranged **columns**. Each column gets input from just a single type of receptor, and from a specific location. Adjacent locations are represented in adjacent columns in a somatotopic manner. Extensive neural connections exist within a column; connections between columns are sparse.

The first cortical stage in somatosensory processing is in area 3b to which most VPL neurons project. Cells in 3b then project to layer IV of areas 1 or 2. The RFs of neurons in area 3b are relatively simple and those in 1 and 2 more complex. Lesion studies show that area 3b is important for all tactile discrimination, area 1 is concerned with analysis of texture and area 2 with **stereognosis**, the ability to perceive the three-dimensional shape of an object by touch. In addition to cutaneous input, area 2 gets proprioceptor input (directly and from 3a) and has reciprocal connections with the motor cortex. These are not involved in modifying ongoing movements but may inform the motor system of the sensory consequences of moving.

The secondary somatosensory cortex (SII) gets inputs directly from the thalamus and from SI. Many neurons in SII have bilateral RFs, that is, stimuli in the corresponding region on both sides of the body will evoke a response. Inputs from the contralateral body surface arise as a direct consequence of the **decussation** (crossing over) of the medial lemniscus. Inputs from the ipsilateral body surface enter SII from the contralateral side via the corpus callosum. By integrating information from both sides of the body, SII is the first stage in the formation of a unified percept of the whole body. It enables tactile discriminations learned using one hand to be easily performed with the other.

SII is important in controlling movement in the light of somatosensory input via its connections with the motor cortex. In addition, SII has inputs to the limbic cortex and so to the hippocampus and amygdala. By this route tactile learning is brought about.

Descending connections

The somatosensory cortex has reciprocal connections with all of the subcortical structures which relay sensory input to it. The descending pathway is made by the corticospinal (pyramidal) tract either directly or via its connections with the brainstem reticular nuclei. These back projections have a somatotopic mapping precisely in register with the ascending DCML pathway. They are probably the vehicle by which somatosensory input can be selectively filtered as an attention mechanism. For example, much of the time we are not aware of the sensations caused by clothes.

F3 Pain

Key Notes

Definition and classification of pain

Pain is the unpleasant sensory and emotional experience that accompanies stimuli that cause tissue damage. Normally it has a protective role. Clinical pain is associated with disease (e.g. arthritis) and has no obvious physiological role. Neuropathic pain results from the damage to pain pathways.

Anterolateral pathways

Small-diameter afferents carrying temperature and nociceptor input enter the dorsal horn, terminating on dorsal horn cells which mediate temperature, pain and poorly localized (crude) touch sensation. Most axons of these neurons cross the midline within one or two spinal segments to ascend in the anterolateral columns. Some nociceptor input is by way of pathways outside the anterolateral columns.

Physiology of spinothalamic neurons

Lamina I neurons are nociceptor specific and give rise to the lateral spinothalamic tract. Fast pain signals are relayed in the neospinothalamic pathway to the posterior nucleus of the thalamus, whereas slow pain is sent to via the spinoreticular pathway to reticular thalamic nuclei. Lamina V cells get inputs from both nociceptors and mechanoreceptors and project in the anterior spinothalamic tract which terminates in the ventral posterior thalamic nuclei. Neurons in the spinal nucleus of the trigeminal nerve equivalent to those in laminae I and V of the spinal cord are responsible for crude touch and pain sensation in face and head.

Crude touch sensation

Activation of lamina V cells by mechanoreceptor afferents only generates a well-localized but poorly discriminated sense of touch.

Discriminative pain perception

Noxious stimuli excite both lamina I cells, which carry the pain signal, and lamina V cells, which allow the pain to be localized. Nociceptor input is widely disseminated in the brain to regions collectively termed the pain matrix (which includes the primary somatosensory cortex), but the site for discriminative aspects of pain is uncertain.

Enhanced pain responses

An injury site may become more sensitive to nociceptor input (hyperalgesia), innocuous input may cause pain (allodynia), or pain can occur in the absence of any stimulus. This ensures that the injured region is protected. Primary hyperalgesia is due to sensitization of polymodal nociceptors by substances produced at the injury site, while allodynia is due to increased excitability of mechanical nociceptors. Secondary hyperalgesia results from central sensitization, in which spinal cord neurons become hyperexcitable to low-threshold primary afferent input and acquire larger receptive fields. One component of central sensitization is wind-up, a progressive increase in the firing rate of spinal cord neurons produced by low-frequency electrical stimulation of C fibers sufficient to produce temporal summation. It is attributed to NMDA receptor activation. Central sensitization is a common feature of chronic inflammatory pain.

Autonomic, arousal and emotional aspects of pain

Connections of the spinoreticular pathway to several brainstem nuclei, hypothalamus and thalamus are responsible for autonomic and arousal responses to pain. The cingulate cortex, part of the limbic system, is concerned with the emotional response to pain.

Related topics

Sensory receptors (F1)
Touch (F2)
Pain modulation (F4)

Definition and classification of pain

Pain is the unpleasant sensory and emotional experience associated with **noxious stimuli**, those which may cause tissue damage. The protective role of **nociceptive pain**, felt only in the presence of acute injury or inflammation, is illustrated by the serious injury that individuals with congenital insensitivity to noxious stimuli unwittingly inflict on themselves. **Clinical** (**pathological**) **pain**, such as that associated with chronic inflammatory states (e.g. arthritis), or which persists long after an injury has apparently healed, has no obvious physiological role. Clinical pain resulting from damage to pain pathways is termed **neuropathic pain**.

Anterolateral pathways

The primary afferents of the anterolateral pathways are small-diameter dorsal root ganglion (DRG) cells driven by thermoreceptors or nociceptors and DRG cells with large 'receptive fields' stimulated by mechanoreceptor input. The anterolateral system thus mediates temperature and pain sensations and poorly localized (crude) touch sensation (*Fig. 1*).

Because conduction by Aδ mechanical nociceptor afferents is so much faster than that of C fiber polymodal nociceptors, a painful blow produces a sharp **fast** (**first**) **pain** initially, followed by an aching or burning **slow** (**second**) **pain**.

The small-diameter fiber afferents, situated laterally in the dorsal roots, enter the **dorsolateral tract of Lissauer**, where they divide into ascending and descending branches giving off collaterals that enter the dorsal horn usually within 1–2 segments. Nociceptive Aδ afferents synapse with projection neurons in lamina I that convey fast pain, and with the distal dendrites of cells in lamina V that extend dorsally into lamina I responsible for crude touch sensation. Nociceptive C fibers synapse both monosynaptically and via interneurons with lamina I projection neurons conveying slow pain. The axons of the projection neurons cross the midline within one or two segments to ascend in the anterolateral columns.

The **anterolateral columns** are the conduits for the spinothalamic pathways, which are the axons of second-order sensory neurons. Those conveying temperature and pain sensations are found in the **lateral spinothalamic tract** while those involved in crude touch sensation, in which localization is preserved but modality is poorly represented, run in the **anterior spinothalamic tract**.

Some nociceptor input is transmitted outside the anterolateral columns, for example by the **spinocervical** tract. This explains why **anterolateral cordotomy** (a surgical procedure to cut the anterolateral columns at a specific spinal level to treat intractable pain) is often followed by the recovery of some pain sensation. Moreover, visceral pain fibers travel in the dorsal column–medial lemniscal system.

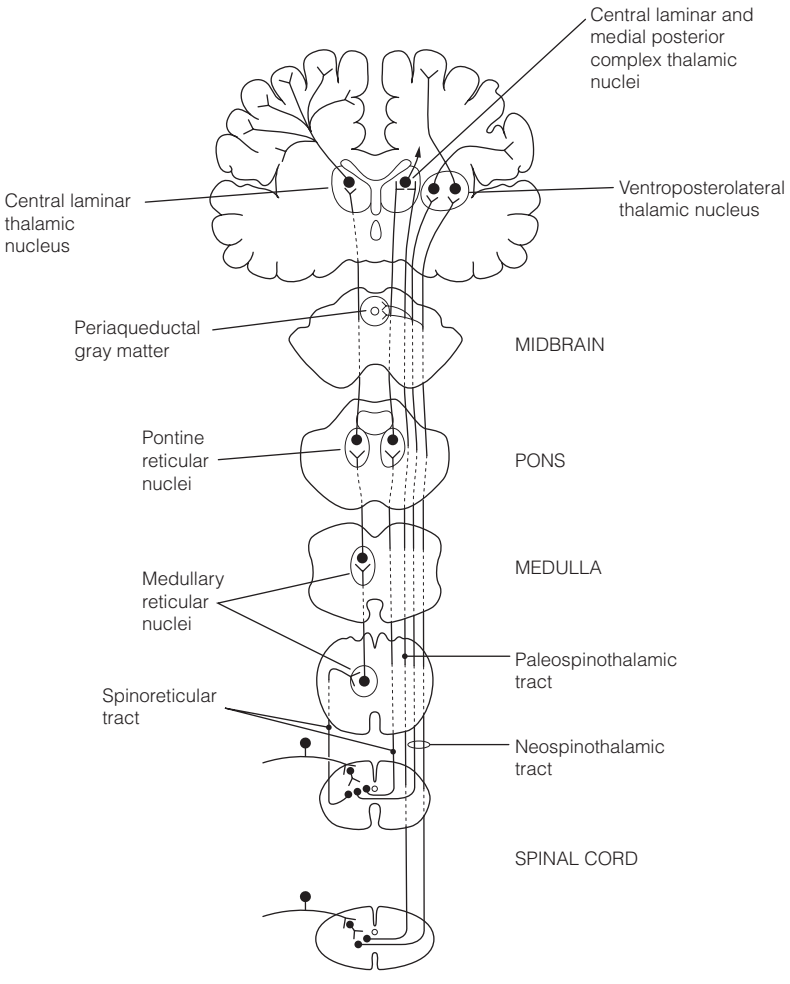


Fig. 1. Anterolateral pathways. See text for details of the individual tracts.

Physiology of spinothalamic neurons

The projection neurons in laminae I are nociceptor specific – that is they respond only to high-threshold afferent activity – and convey fast pain, slow pain or itch; these are separate submodalities and are processed in parallel but differently.

Lamina I projection neurons enter the lateral spinothalamic tract. Those conveying fast (Aδ fiber) pain decussate to ascend on the contralateral side in the **neospinothalamic pathway**, terminating in the **posterior nucleus** of the thalamus. Projection neurons transmitting slow (C fiber) pain send their axons (together with axons from cells in laminae VII and VIII) into the **spinoreticular** (**paleospinothalamic**) **pathway**, a part of the spinothalamic tract that makes extensive connections in the reticular system of the brainstem (e.g. the parabrachial nucleus) and the **reticular nuclei** of the thalamus. Many spinoreticular axons do not cross the midline but ascend ipsilaterally.

The largest number of spinothalamic tract cells are in lamina V. The dendrites of some of these neurons extend into lamina I and receive inputs from

nociceptor afferents (both Aδ and C fiber). But lamina V neurons also receive connections from large-diameter (Aβ) mechanoreceptor afferents so are said to be **wide dynamic range** (**WDR**) cells, because they get inputs from low-threshold *and* high-threshold afferents (*Fig. 2*). Hence most spinothalamic tract neurons can be excited by *both* innocuous and noxious stimuli. Their axons cross the midline to enter the neospinothalamic pathway and terminate in the ventral posterior and posterior nuclei of the thalamus.

Neurons equivalent to the lamina I and lamina V cells in the spinal cord are also found in the **spinal nucleus of the trigeminal nerve**, which get their inputs from the primary afferents of the face and head. They project to the ventral posterior medial and posterior thalamic nuclei.

Note that although the target thalamic nuclei of lamina V and trigeminal nucleus cells are the same as those of the dorsal column–medial lemniscus pathway, the thalamic neurons on which these two systems synapse are different.

Neurons in lamina IV also project to the contralateral anterior spinothalamic tract. These cells receive input from large-diameter (Aβ) mechanoreceptor afferents but not from nociceptor afferents. They too mediate crude touch.

Crude touch sensation

Lamina V (and lamina IV) cells when activated only by large-diameter low-threshold afferents convey crude touch sensation. These cells have small receptive fields and their projections are organized topographically and so are able to localize stimuli precisely. However, the ability of the neospinothalamic pathway to discriminate between the different mechanoreceptor submodalities is poor.

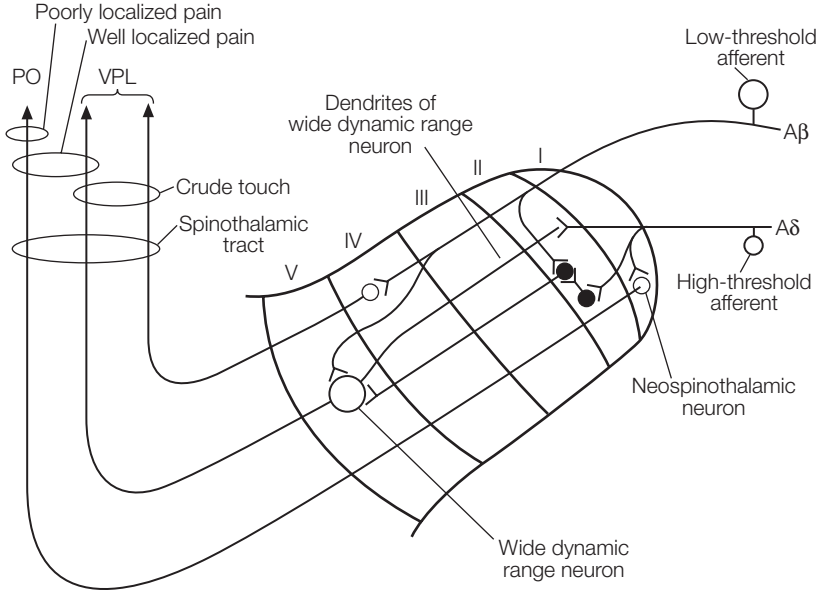


Fig. 2. Circuitry to account for fast pain and crude touch transmission by dorsal horn. C-fiber input for slow pain sensation is not shown. PO, posterior nucleus of thalamus; VPL, ventral posterior lateral nucleus of thalamus; ○–≺, excitatory neurons; ●–≺, inhibitory neurons.

Discriminative pain perception

Any noxious stimulus will activate both lamina I and lamina V cells (stimuli exciting only lamina V cells are perceived as innocuous). Lamina I cells signal pain, but because they have large receptive fields, and relay to posterior lateral and reticular thalamic nuclei which do not have topographic projections to the cortex, they cannot localize it well. The localization of painful stimuli depends on the simultaneous firing of lamina V cells, which project in somatotopic fashion to the ventral posterior lateral nucleus (and the homologous ventral posterior medial nucleus for the trigeminal system) of the thalamus. These thalamic nuclei project to the primary somatosensory cortex (SI).

Brain-imaging studies reveal that activity increases in a number of brain areas when human volunteers experience pain. These regions are collectively termed the **pain matrix** and include the periaqueductal gray, secondary somatosensory cortex (SII), insula, anterior cingulate cortex, movement-related areas of the cerebellum, supplementary motor areas, and much less robustly SI and the thalamus. Clinical ablation of large parts of SI does not alter pain perception, hence discriminative aspects of pain must be conferred elsewhere, perhaps by SII. Other cortical (limbic) areas are concerned with emotional responses to pain.

Enhanced pain responses

At the site of an injury abnormal pain responses can occur:

- noxious stimulation becomes more painful than usual (**hyperalgesia**);
- stimuli that are normally innocuous can become painful (**allodynia**);
- pain can occur in the absence of any stimulus (**spontaneous pain**).

These responses, when transient, have a protective role in that they encourage guarding of the site of injury. However, when they outlast their usefulness they become pathological pain.

Hyperalgesia arising from peripheral mechanisms is termed primary hyperalgesia, while that arising from central mechanisms is secondary hyperalgesia. **Primary hyperalgesia** is due to sensitization of *polymodal* nociceptors by substances produced at the injury site. For example, prostaglandins enhance nociceptor excitability by activating protein kinase A and phosphorylation of sodium channels. Glutamate released from nociceptors, macrophages, and leaky blood vessels following injury is also a key player. Allodynia is due to glutamate increasing the excitability of *mechanical* nociceptors. The effect of glutamate on nerve terminals is mediated by metabotropic glutamate receptors coupled to phospholipase C.

Secondary hyperalgesia arises partly from changes in the behavior of wide dynamic range (WDR) spinal cord cells, termed **central sensitization**, which has the following characteristics:

- it is induced by tissue or nerve injury, so it occurs secondary to activity in primary nociceptor afferents;
- increased spontaneous (background) firing;
- hyperexcitability to low-threshold (Aβ fiber) primary afferent input;
- increased size of receptive fields.

Under normal circumstances stimuli applied near the edge of the receptive field of a WDR cell in the spinal cord usually elicit subthreshold responses (epsps). However when inflammation arises in the periphery as a result of injury, persistent firing of primary nociceptor afferents causes the continual release of glutamate and a number of neuropeptides (e.g. substance P, neurokinin A and calcitonin gene-related peptide). These transmitters depolarize the WDR cells,

thus increasing their probability of firing. Now the cell is more sensitive to excitatory input, and previously subthreshold stimuli at the edge of its receptive field become *suprathreshold*; in effect the receptive field is enlarged.

In animals, electrical stimulation of C fibers at a frequency above about 0.3 Hz causes temporal summation and a progressive increase in the firing rate of the lamina I cells, a phenomenon called **wind-up**. The enhanced responsiveness lasts for several minutes after the stimulus train has stopped. Wind-up is attributed to activation of NMDA receptors and blocked by NMDA receptor antagonists, such as the anesthetic ketamine. However, it is also blocked by substance P receptor antagonists, and by opioid receptor (μ and δ) antagonists given preemptively, so it is clearly a complex phenomenon. Wind-up occurs at frequencies as low as 0.5 Hz that are typical of background firing rates of sensitized WDR cells. Hence it is thought to be a component of central sensitization. In patients with secondary hyperalgesia arising from a variety of causes, temporal summation of pain can be elicited and this can be blocked by ketamine, suggesting that wind-up is indeed important. However, it is not the whole story. While temporal summation is a correlate of wind-up, central sensitization can occur without temporal summation. In addition, high-frequency NMDA receptor activation can trigger *persistent* increases in the excitability of the lamina I cells, and enlargement of their receptive fields, by formation of new connections, in a manner akin to long-term potentiation.

Central sensitization is a common feature of chronic inflammatory pain. Recent work is beginning to reveal underlying mechanisms and suggest treatment *(Fig. 3)*. Peripheral nerve injury somehow activates microglia in the local spinal cord. One consequence of this is to up-regulate $P2X_4$ receptors, one of a family of ligand-gated ion channels which respond to ATP, probably released by primary afferent terminals, dorsal horn cells or astrocytes. Stimulation of the $P2X_4$ receptors causes the microglia to release cytokines which increase the sensitivity of lamina I dorsal horn cells.

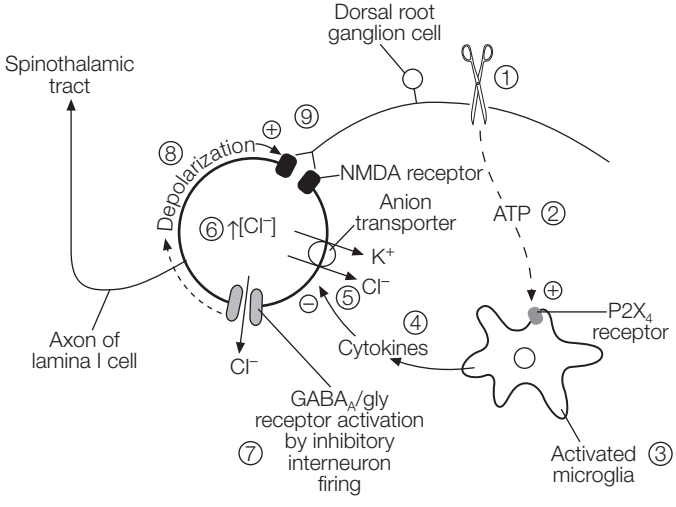


Fig. 3. A model for central sensitization in response to peripheral nerve injury. Numbers indicate the sequence of events. Glutamatergic input from remaining intact fibres cause prolonged firing of the lamina I cell because the voltage-dependent blockade on its NMDA receptors has been removed by the abnormal depolarization (8).

The identity of the offending cytokines is not known, but a likely sequence in the lamina I cells is as follows. Peripheral nerve injury causes the down-regulation of an anion exporter and this results in a large rise in intracellular Cl^- concentration. Lamina I cells now respond to activation of their $GABA_A$ and glycine receptors by a large depolarization, caused by Cl^- *leaving* the cell down its concentration gradient. The depolarization allows NMDA receptors to be opened by any incoming excitation. Analgesic cover after surgery is optimal if given immediately and topped-up frequently. This curtails central sensitization, dramatically reducing post-operative pain. For established secondary hyperalgesia, antagonists of $P2X_4$ receptors and NMDA receptors could prove clinically important. Unfortunately, NMDA antagonists currently available, e.g. ketamine, have side effects which curtail their usefulness.

Autonomic, arousal and emotional aspects of pain

The connectivity of the spinoreticular pathway brings about autonomic and arousal responses to painful stimuli.

1. Brain stem reticular system connections generate autonomic responses to pain (e.g. increases in heart rate and blood pressure) and changes in ventilation.
2. Direct connections with the hypothalamus are responsible for activation of the hypothalamic–pituitary–adrenal axis stress response.
3. Activation of noradrenergic pathways in the brainstem reticular system contributes to the increased arousal experienced during pain.
4. Connections run to midbrain reticular structures including the periaqueductal gray (PAG) and the parabrachial nucleus in the **spinomesencephalic tract**, which lies laterally in the anterolateral columns. The PAG is part of a system for modulating nociceptor input, and the parabrachial nucleus projects to the amygdala and so is presumably involved in learning to fear situations associated with pain.
5. The thalamic targets of spinoreticular neurons establish widespread connections with the basal ganglia, influencing motor activity, and with the cerebral cortex.
6. All of the thalamic nuclei that receive nociceptor input send direct connections to the insular cortex, lesions of which reduce the emotional responses to pain without altering discriminative aspects. Imaging studies show that the anterior cingulate cortex is involved in emotional responses to pain. Interestingly, recent brain-imaging studies reveal that selected parts of the pain matrix (anterior insula and rostral anterior cingulate) are activated by watching pain inflicted on a loved one.

F4 Pain modulation

Key Notes

Spinal control of nociceptor input

Control of nociceptor input into the CNS is exerted at two levels. At the spinal cord the input into the spinothalamic tract from small-diameter primary afferents is inhibited by concurrent activity in large-diameter mechanoreceptor afferents, via enkephalinergic interneurons in the substantia gelatinosa.

Supraspinal pain modulation

An anti-nociception system descends from the periaqueductal gray matter to activate serotonin and noradrenergic neurons in the brain stem that inhibit nociceptor input. Two systems in insular (limbic) cortex, modulated by tonic GABAergic inhibition, modify pain sensation, presumably in response to cognitive and emotional factors. One influences the anti-nociception system and produces analgesia, the second projects to the amygdala and is a pain-producing system. Natural analgesic systems appear to be engaged during injury in highly arousing contexts (e.g. sport, war) and after stressful encounters.

Opioid involvement in pain

Some neurons in pain-modifying pathways use opioid neurotransmitters such as met-enkephalin, endorphin and morphine. It is for this reason that opioids are such potent analgesics. Opioid transmission has been implicated in acupuncture and the placebo effect.

Pathological pain

Pain arising from internal organs is often referred to the body surface because nociceptor input from viscera and skin often converges in the spinal cord. Unmasking of central connections may be responsible for phantom pain that follows amputation of limbs. Pain can be caused by damage to central pain pathways, e.g. the thalamus.

Related topics

Purines and peptides (D8)
Sensory receptors (F1)
Touch (F2)
Pain (F3)

Spinal control of nociceptor input

Mechanisms to reduce nociceptor input operate at both the spinal and supraspinal level. At the level of the spinal cord, whether a stimulus is perceived as painful or innocuous depends on the relative activity of large- and small-diameter fibers. An explanation for this, the **gate control theory** (*Fig. 1*) is based on three experimentally verifiable postulates. Firstly, that the wide dynamic range neurons in lamina V get convergent excitatory input from large (Aβ) mechanoreceptor afferents and small (Aδ and C) nociceptor afferents. Secondly, that the large-diameter Aβ fibers inhibit lamina V neurons by activating inhibitory interneurons in lamina II. Thirdly, that nociceptor afferents inhibit the lamina II interneurons.

How does this circuitry operate? Nociceptor afferents directly excite the distal dendrites of lamina V cells that lie in lamina I, and indirectly excite the lamina V cells by inhibiting the lamina II interneurons (inhibiting an inhibition is excitation). This opens the pain gate. Large-diameter afferent activity offsets the

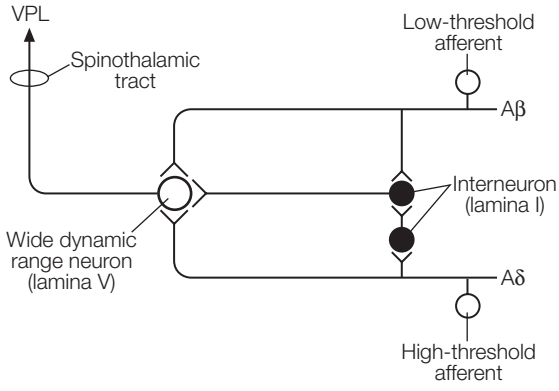


Fig. 1. Circuitry required by the gate control theory. VPL, ventral posterior lateral nucleus of thalamus.

effect of nociceptor input, reducing the firing of lamina V cells by increasing the inhibitory drive from the lamina II interneurons. Hence co-activation of low- and high-threshold afferents closes the pain gate. Similar circuitry probably also operates to modulate nociceptor input via the projection neurons in lamina I.

The gate control theory accounts for **counter-stimulation analgesia** in which pain is reduced by stimulating low-threshold afferents; we instinctively rub the site of a painful blow, stimulating mechanoreceptors. **Transcutaneous electrical nerve stimulation** (**TENS**) delivers high-frequency, low-intensity currents, sufficient to stimulate Aβ and Aδ fibers. This is used particularly to produce analgesia during labor and physiotherapy. The analgesia produced by TENS can be blocked by opioid receptor antagonists.

Supraspinal pain modulation

Brainstem nuclei that receive connections from the spinoreticular pain pathways give rise to descending supraspinal analgesia systems which *reduce* nociceptor input (*Fig. 2*).

1. Axons of the **nucleus raphe magnus** (**NRM**) terminate in the dorsal horn, releasing serotonin which reduces transmission from nociceptors into the lamina I dorsal horn cells and hence up spinothalamic tracts. Serotonin exerts its effects by:
   - acting at serotonin receptors on the nociceptor terminals to reduce transmitter release (i.e. presynaptic inhibition);
   - acting at postsynaptic serotonin receptors that open K^+ channels, thereby hyperpolarizing the lamina I cells;
   - stimulating enkephalinergic interneurons which hyperpolarize the lamina I cells by acting on μ-opioid receptors that open K^+ channels.
2. Noradrenergic neurons of the locus ceruleus project to the superficial dorsal horn. Their activation reduces the response of lamina I cells to nociceptor input by similar mechanisms to serotonin.
3. The **periaqueductal gray** (**PAG**) matter surrounds the cerebral aqueduct in the midbrain. Stimulation of the PAG in rats causes a powerful suppression of pain responses, **stimulus-induced analgesia**. This has been exploited by neurosurgeons who implant chronic electrodes into the PAG of patients with intractable pain. The PAG exerts its anti-nociception effects by activating the brainstem serotoninergic and noradrenergic neurons.

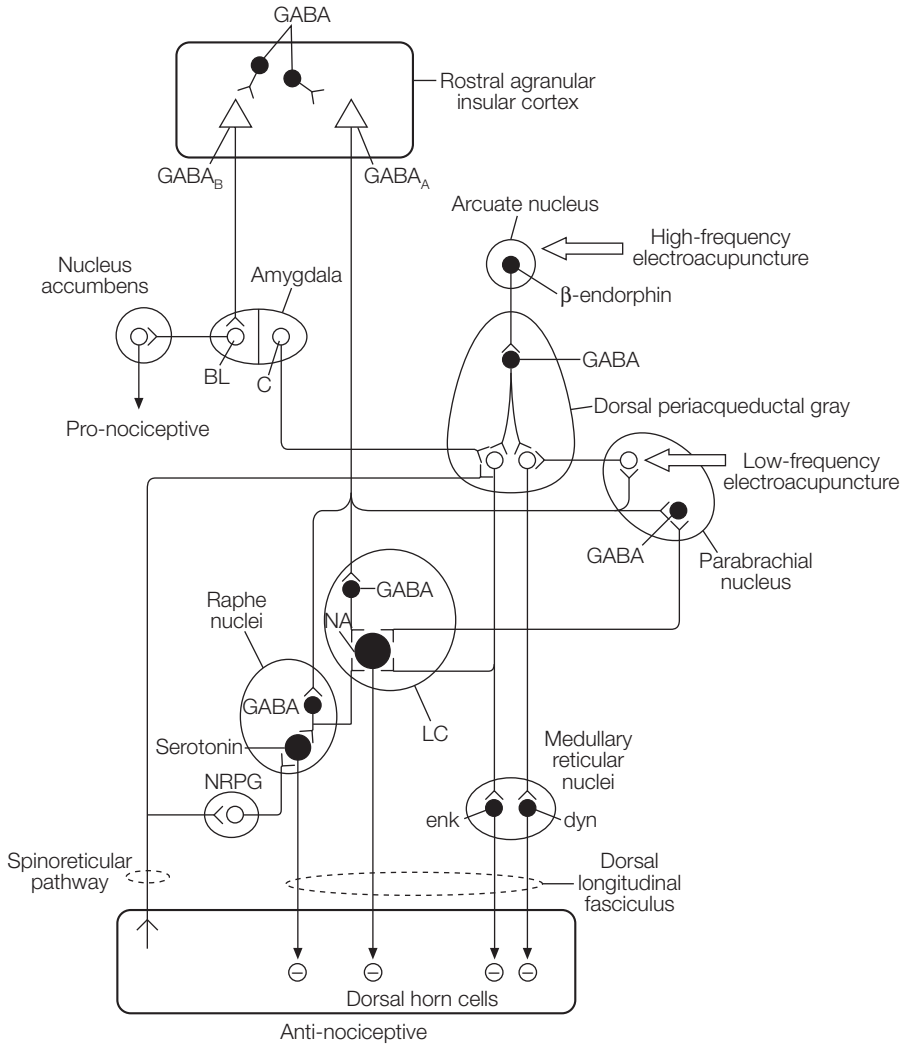


Fig. 2. Descending pathways modulating nociceptor input. BL, basolateral nucleus of amygdala; C, central nucleus of amygdala; LC, locus ceruleus; NA, noadrenaline (norepinephrine); NRPG, nucleus reticularis pars gigantocellularis; enk, enkephalin; dyn, dynorphin; ○≺, excitatory neurons; ●≺, inhibitory neurons.

The anti-nociception pathway is probably activated by intense nociceptor input into the spinothalamic tract by way of a medullary reticular nucleus (NRPG, see *Fig.* 2) and by stress. But, in addition, the limbic cortex can exert anti-nociceptive or pro-nociceptive influences via two separate pathways. Both are comprised of pyramidal neurons in the **rostral agranular insular cortex** (**RAIC**) that are under tonic inhibitory control from cortical GABAergic interneurons.

1. Activation of the anti-nociceptive pathway in rats by injecting vigabatrin (which raises GABA concentration) into the RAIC causes analgesia. It runs to brainstem GABAergic interneurons (e.g. in the raphe nuclei) that inhibit

noradrenergic neurons of the locus ceruleus. The tonic cortical inhibition dampens RAIC pyramidal cell excitation of the brainstem GABAergic cells, allowing the noradrenergic cells to silence lamina I cells.
2. The pro-nociceptive pathway goes to the amygdala and then to the nucleus accumbens. Activating this pathway results in hyperalgesia.

It is presumably by these cortical pathways that pain sensation is modified by cognitive and emotional factors.

The descending pain-modulating pathways appear to be brain analgesia systems. In the event of injury they produce the natural counterpart of stimulus-induced analgesia. **Emergency analgesia**, seen in individuals who have sustained severe injuries when they are in threatening or highly arousing situations (e.g. during warfare or sport), is rapid in onset and wears off after a few hours. The analgesia is localized to the injured area, autonomic reflexes and emotional responses reflect a typical defense reaction. The selective advantage of emergency analgesia is that it allows an individual to continue to function so they can remove themselves from danger. **Stress analgesia** occurs in individuals recovering from a stressful encounter. Here the analgesia is generalized, the individual is relaxed and has low levels of sympathetic activity. These two behaviors may be organized by different parts of the periaqueductal gray matter.

Opioid involvement in pain

Some neurons in pain-modulating pathways use opioid neurotransmitters:

- Met-enkephalin is used by interneurons in lamina II responsible for inhibiting the lamina I nociceptor-specific spinothalamic neurons.
- Neurons using β-endorphin or met-enkephalin are found in the PAG where they inhibit GABAergic interneurons that normally tonically suppress the anti-nociception neurons.
- Endogenous **morphine** has been identified in terminals forming synapses with neurons expressing μ-opioid receptors in pain-modulating pathways, where it appears to be a *bona fide* neurotransmitter.

Opioid peptides and opiate drugs such as **morphine** and **pethidine** are full or partial agonists at one or more of the opioid receptors, and produce analgesia by mimicking the actions of the endogenous opioids in the pain-modulating pathways. The effects of opioid analgesics can be rapidly reversed by the competitive antagonists such as **naloxone**. Receptors in the medulla are responsible for two of the unwanted effects of opioid analgesics, respiratory depression (which occurs at therapeutic doses) and vomiting. **Electro-acupuncture**, which delivers electrical stimulation through needles inserted through the skin, relies on stimulation of small-diameter (group II and III) afferents from skeletal muscle and joints. Low-frequency stimulation activates a descending endorphinergic pathway from the arcuate nucleus of the hypothalamus to the ventrolateral PAG, activating enkephalinergic neurons that run from the medulla to inhibit dorsal horn cells. High-frequency stimulation activates a separate PAG pathway that uses dynorphin-secreting neurons in the medulla. Alternating between low- and high-frequency electro-acupuncture produces stronger analgesia in patients than either alone. Interestingly the ventrolateral PAG is responsible for stress analgesia. Acupuncture produces a state rather like stress analgesia, and both can be reversed by naloxone.

The **placebo effect** refers to the therapeutic efficacy of agents or procedures that are without any physiological or pharmacological action. In general, the

more elaborate the placebo treatment the better its effectiveness. There is evidence that the placebo effect, at least when harnessed to reduce postoperative pain, can be blocked by naloxone and so depends upon endogenous opioid neurotransmission.

Functional MRI imaging shows that placebo analgesia is associated with decreased activity in the pain matrix but *increased* activity in the dorsal lateral prefrontal cortex, an area thought to be involved in cognition. A working hypothesis is that the expectation of pain relief activates a top–down recruitment of supraspinal opioid pathways.

Pathological pain

Pathological pain arises in many situations, including chronic inflammation or following damage to pain pathways.

Convergence of visceral and somatic nociceptor afferents in a given spinal root onto common spinothalamic neurons means that stimulation of nociceptors in internal organs is often perceived as pain in distant superficial regions. This is termed **referred pain** and the areas to which it is referred are called **Head's zones** (*Fig. 3*). These are useful in diagnosis.

After amputation of a limb almost all patients experience the sensation that the limb is still present and for some it is painful (**phantom pain**). Phantom sensations can also follow loss of other body parts (e.g. mastectomy). Children born without limbs have powerful, non-painful, phantom limb sensations. This implies that complete somatotopic representations can exist in the absence of peripheral inputs, and this is the situation in amputees. Sensations in phantom limbs are poorly localized and not abolished by cutting the dorsal roots or by surgical lesions to the deafferented regions of the thalamus or to SI. It is attributed to spurious firing of dorsal horn cells resulting from the lack of proprioceptor feedback. However, considerable rewiring of the somatosensory cortex occurs after amputation so that neurons that have lost their original inputs acquire functional connections from previously silent synapses established by neighboring cells. Phantom limb sensations can now be elicited by touching other parts of the body, commonly the face. With continuing cortical reorganization phantom sensations may fade over time.

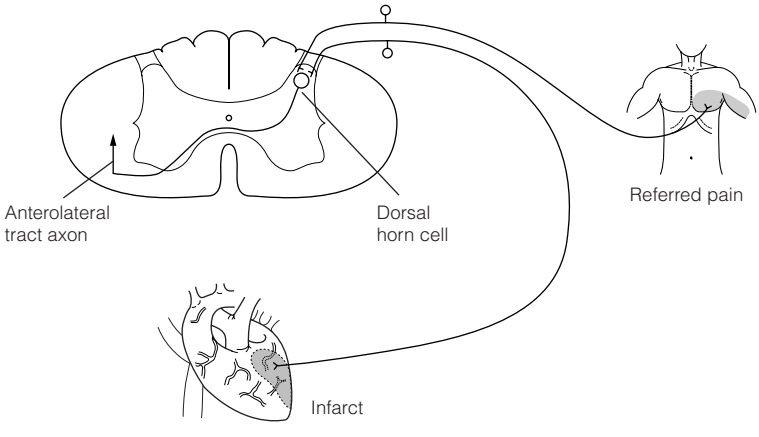


Fig. 3. Referred pain. Typically the pain of ischemic heart disease is referred to the chest and left arm.

Strokes which cause infarcts of the ventrobasal thalamus can cause the **thalamic syndrome**. This is an example of neuropathic pain – one arising from damage to pain pathways – which is perceived as spontaneous burning pain. It is thought to arise as a result of re-wiring of neural pathways as a consequence of the injury.

Trigeminal neuralgia is a disorder in which innocuous stimulation of a particular area on the face triggers brief episodes of excruciating pain, despite lack of an obvious pathology. It results from compression of the trigeminal nerve root by rogue blood vessels. As with other severe pain syndromes, treatment with anticonvulsant drugs which work by blocking voltage-dependent sodium channels can be effective. Surgical treatment is microvascular decompression.

F5 BALANCE

Key Notes

Vestibular functions

Receptors in the inner ear detect the position and motion of the head in space and this information is used to maintain body posture in the face of forces tending to disrupt it. In addition it allows gaze to be controlled independently of head movement via vestibulo-ocular reflexes.

Vestibular labyrinth

The bony labyrinth houses the inner ear. Within it lies the epithelial membranous labyrinth, the vestibular part of which consists of the utricle, saccule (the otolith organs) and three mutually perpendicular semicircular ducts. The otolith organs detect linear acceleration; the utricle responding to tilt away from the horizontal and the saccule to acceleration due to gravity. The semicircular ducts detect angular acceleration of the head.

Vestibular fluids

The vestibular labyrinth is filled with endolymph, while the space between the bony and membranous labyrinths contains perilymph. Active transport of potassium ions into the endolymph creates a potential difference between endolymph and perilymph of +80 mV. This is needed for sensory transduction by hair cells.

Transduction in otolith organs

The macula is the sensory structure of the otolith organs. It consists of a sheet of epithelium with sensory hair cells, overlain by a gelatinous matrix, the otolith membrane, containing crystals of calcium carbonate. Hair cells possess a single kinocilium and a number of stereocilia, the tips of which are embedded in the otolith membrane. Any force which displaces the otolith membrane with respect to the hair cells causes their kinocilia and stereocilia to bend. This either opens or closes K^+ channels in the hair cells, depending on the direction of the force, causing a depolarizing or hyperpolarizing response respectively. Alterations in the membrane potential of a hair cell alter its release of transmitter, so modifying the firing rate of its postsynaptic primary afferent.

Transduction in semicircular ducts

At one end of a semicircular duct is an ampullary crest, the hair cells of which have their stereocilia embedded in a gelatinous sheet, the cupula. Rotation of the head in the plane of the duct distorts the cupula because inertia causes the rotation of the endolymph to lag behind. Distortion of the cupula stimulates the hair cells, the transduction mechanism of which is identical to that of hair cells in the otolith organs.

Meniere's syndrome

Abnormal increase in the volume of endolymph is believed to be responsible for Meniere's syndrome of progressive hearing loss, tinnitus and vertigo.

Central vestibular connections

Vestibular primary afferents live in the vestibular ganglion and send their centrally directed axons through the 8th cranial nerve to form connections with neurons in four vestibular nuclei. The inferior nucleus projects to the contralateral ventral posterior thalamus and so to a region of parietal cortex close to SI. This pathway mediates the conscious perception of balance.

Related topics

Oculomotor control (G8) Brainstem postural reflexes (J5)
Anatomy and physiology of the ear (H2)

Vestibular functions

The sense of balance is conferred by receptors which detect the position and motion of the head in space. The receptors are located in organs which are part of the **inner ear** (**labyrinth**) located in a hollow **vestibule** and three **semicircular canals** within the petrous portion of the temporal bone. Conscious perception of balance is normally overshadowed, except when head acceleration is high, by visual and proprioceptive cues to head position and motion. Vestibular input is used to maintain body posture in the face of forces which shift the center of mass to cause pitch (rocking backwards and forwards), yaw (rocking from side to side) or roll (rotation around the long axis of the body), by adjusting output to antigravity muscles.

Vestibular input is also used to execute eye movements which are independent of head movement. These **vestibulo-ocular reflexes** constitute one of several mechanisms for maintaining a fixed gaze.

Vestibular labyrinth

Within the **bony labyrinth** which contains all the inner ear structures lies the **membranous labyrinth**, a sensory epithelium subserving hearing and balance (*Fig. 1*). The vestibular labyrinth, concerned with balance, consists of two **otolith organs**, the **utricle** and **saccule** and three **semicircular ducts**. The sensory structure of the otolith organs, the **macula**, which detects linear acceleration, is

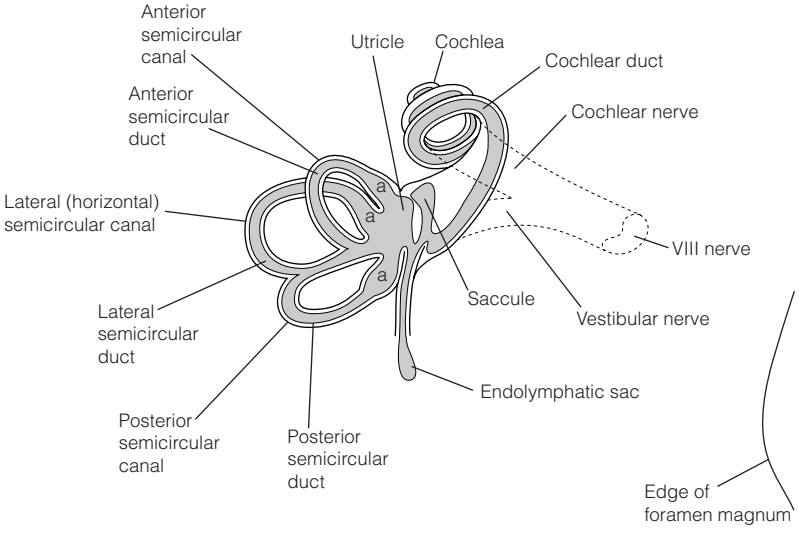


Fig. 1. Left vestibular labyrinth viewed from above. The membranous labyrinth is shaded. a, ampullae of the semicircular ducts.

horizontal in the utricle and vertical in the saccule for a person standing upright. The effect of this is that the utricle is sensitive to tilting of the head (pitch and yaw) whilst the saccule is sensitive to vertically acting forces such as the acceleration due to gravity. The three semicircular ducts are approximately mutually orthogonal. Each contains a sensory structure, the **ampullary crest**, which detects angular acceleration in the plane in which the duct lies. Using the signals coming from all six semicircular ducts the brain computes the magnitude and direction of the angular acceleration of the head.

Vestibular fluids

The vestibular labyrinth is filled with **endolymph** which has a potassium concentration of about 160 mM and a sodium concentration of about 2 mM, and has a composition similar to intracellular fluid. It is secreted by a specialized epithelium, the **stria vascularis**, lining the outer wall of the **cochlear duct**, and drains into a venous sinus of the dura via the **endolymphatic sac**. The space between the bony and membranous labyrinths is filled with a CSF-like fluid, **perilymph**, secreted by arterioles of the periosteum (the connective tissue layer covering the bone), which drains into the subarachnoid space via the **perilymphatic duct**. The high potassium concentration in the endolymph arises because the marginal cells of the stria vascularis have Na, K-ATPase on their basolateral border allowing them to concentrate potassium ions for secretion into the endolymph (*Fig. 2*). The K^+ transport results in the endolymph having a potential difference of +80 mV. Since the resting potential of the hair cell is about –60 mV, the effective potential across its apical border is 140 mV. Hence there are sizeable electrical and chemical gradients favoring facilitated passive diffusion of K^+ across the hair cell. This makes hair cells extremely sensitive.

Transduction in otolith organs

The macula consist of an epithelial sheet of **supporting cells**, embedded in which are an array of **hair cells**, which are sensory epithelial cells. Each hair cell is innervated at its base by a single vestibular afferent and an efferent fiber. The apical border of a hair cell has a single motile **kinocilium**, resembling a cilium, and 40–100 **stereocilia**, microvilli which are progressively shorter the further they are from the kinocilium (*Fig. 3*). This defines an **axis of polarity** for a hair

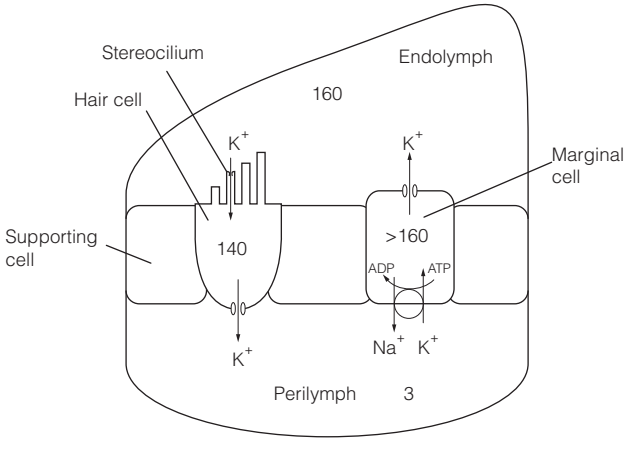


Fig. 2. Simplified model of K^+ transport in the inner ear. Figures are approximate concentrations of K^+ (mM).

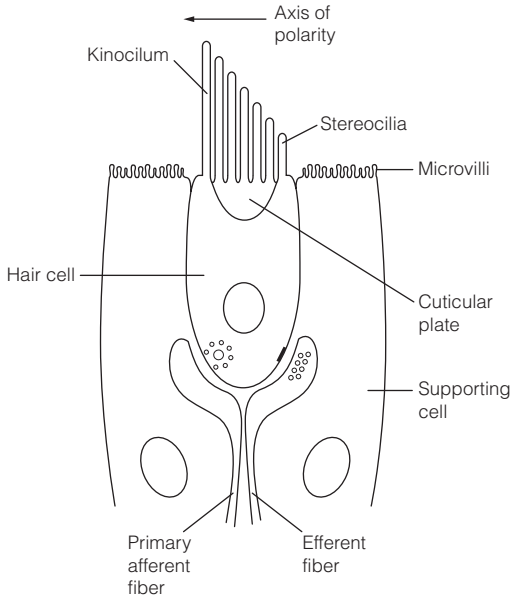


Fig. 3. The hair cell of an otolith organ surrounded by supporting cells of the sensory epithelium.

cell, with a direction going from the smallest stereocilium to the kinocilium. Stereocilia lying along this axis are connected at their tips, those perpendicular to the axis are free.

The kinocilium and stereocilia are embedded in a gelatinous matrix, the **otolith membrane**, containing tiny crystals of calcium carbonate, called **otoliths**. The hair cell is at rest if no force acts on the otolith membrane to cause the stereocilia to pivot. In this state the tension in the links (*Fig. 4*) connecting adjacent stereocilia is slight so only about 10% of the potassium ion channels gated by these links are open, causing a small depolarization.

This is sufficient to sustain tonic release of an excitatory transmitter, probably glutamate, which maintains baseline firing of the primary afferent. Head tilt in the

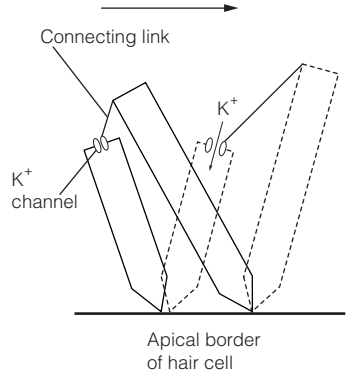


Fig. 4. Transduction by otolith hair cells. Two stereocilia are shown with their connecting link. When pivoted to the right (dotted lines) the connecting link is stretched, opening K^+ channels to cause an influx of K^+.

direction of the axis of polarity causes the otolith membrane to pull on the stereocilia, making them pivot, so increasing the tension in the connecting links. This opens the K^+ channels in the tips of the stereocilia, allowing K^+ influx to depolarize the hair cell, increasing its transmitter release and raising the afferent firing rate. Tilt in exactly the opposite direction, by reducing tension in the links, causes hair cell hyperpolarization which reduces primary afferent firing. Tilt which is perpendicular to the axis of polarity of a hair cell has no effect because stereocilia are not linked in this direction. Tilts in intermediate directions cause graded receptor potentials. The responses of individual otolith afferents are proportional to tilt angle and adapt only with prolonged stimulation. The axes of hair cells are orientated in an orderly pattern towards (utricle) or away from (saccule) the **striola**, a curved boundary running across the middle of the macula surface. This means that a given stimulus will depolarize some hair cells and hyperpolarize others.

Transduction in semicircular ducts

Velocity is a vector quantity, it has a magnitude (speed) and direction. For circular motion, such as head rotation, even if the angular speed is kept constant, the direction in which the velocity vector is acting is continuously changing. Hence head rotation is an angular acceleration.

Both ends of each semicircular canal insert into the utricle. Within the canal is the endolymph-filled semicircular duct. At one end of each duct is a dilation, the **ampulla**; in which sits the ampullary crest. Vestibular hair cells in the crest have their stereocilia embedded in a gelatinous sheet, the **cupula**, which stretches from the crest to the foot of the ampulla. Rotation of the head maximally stimulates hair cells in the canals lying in the same plane as the rotation. Rotation of the endolymph lags behind head rotation because of its inertia, so the endolymph exerts a pressure distorting the cupula, bending the stereocilia. The transduction mechanism is identical to that of hair cells in otolith organs. Because the cupula is not an ideal pressure transducer the signals transmitted by the duct afferents measure angular acceleration for slow and fast rotations, but encode velocity for mid-range rotation speeds. Semicircular ducts on each side lying in the same plane operate in pairs. Head rotation that causes depolarization of hair cells in the horizontal duct of the left ear will hyperpolarize hair cells in the horizontal duct of the right ear. This works because the axis of the hair cells points towards the utricle on both sides. Since the anterior duct on one side lies on approximately the same plane as the posterior canal on the other side these make up the two further pairs.

Meniere's syndrome

Typically this disorder involves progressive hearing loss, **tinnitus** (high- or low-pitched sounds generated within the head) in one ear, and episodes of **vertigo** (loss of balance sensation often accompanied by nausea and vomiting). It is associated with an increase in the volume of the endolymph causing herniations and ruptures of the membranous labyrinth. It may be caused by a viral infection of the ear which impairs reabsorption of endolymph. Cutting the vestibular nerve or destroying the labyrinth on the affected side can abolish the vertigo but has no effect on the other symptoms.

Central vestibular connections

The vestibular primary afferents (about 20 000 on each side) are pseudobipolar cells with their cell bodies in the **vestibular (Scarpa's) ganglion**. Their axons run in the vestibulocochlear nerve to enter the vestibular nuclei, which lie laterally in the medulla and pons. There are four vestibular nuclei. Their roles in postural reflexes and vestibulo-ocular reflexes are dealt with in J5 and G8 respectively.

The pathway for conscious perception of balance are axons of the inferior vestibular nucleus which cross to the contralateral side, ascend close to the medial lemniscus to terminate in the ventral posterior thalamus. Cortical representation is in the lateral parietal cortex, adjacent to the SI map of the head. Other representations may exist in the superior temporal cortex adjacent to the auditory area.

G1 ATTRIBUTES OF VISION

Key Notes

Visual perception

From two-dimensional retinal images the brain constructs a three-dimensional percept by which it can identify what is present in the world and where it is. Different aspects of the visual image (color, form, movement and depth perception) are processed in parallel by distinct pathways. Visual perception arises because the brain has internal representations which it compares with retinal images to make hypotheses about what they are. This allows objects to be recognized from any viewpoint, even if the image is data poor. Object recognition is facilitated by perceptual constancy in which parameters like size and color are preserved over big differences in viewing conditions. Some internal representations are developmentally programmed but most are learned.

Sensitivity

The human eye responds to light between 400 and 700 nm over a 10^{11}-fold intensity range, though discrimination declines for higher light levels.

Acuity

The ability to see fine detail is highest at the central portion of the retina, the fovea. In ideal conditions, two points are separable if they subtend an angle of one arc minute at the retina. Acuity falls with decreasing light levels.

Depth perception

The perception of distance comes from monocular cues for distant objects and binocular cues for nearer objects. Monocular cues include parallax, perspective and shadows. Binocular vision, stereopsis, arises because each eye gets a slightly different view of the world so the image of an object may lie on different points in each retina. For small degrees of retinal disparity the brain reconstructs a single percept and computes the object distance from the disparity.

Color vision

Color vision allows boundaries to be discerned simply on the basis of differences in the wavelength composition of reflected light. It requires a minimum of two types of receptor that respond over different wavelengths so that two values for brightness can be assigned for each bit of an image. This dichromatic vision is the case for most mammals. Many primates, including humans, have trichromatic vision mediated by three types of receptor, which allows three brightness values to be ascribed to an object. The brain compares these values to give the perception of color.

Related topics

Retina (G3)
Early visual processing (G6)
Parallel processing in the visual system (G7)

Visual perception

Vision has been defined as the process of discovering from images what is present in the world and where it is. This requires that the brain is able to use a two-dimensional shifting pattern of light intensity values on the two retinas (the

light-sensitive layers of the eyes) to form a representation of the form of an object, its color, movement and position in three-dimensional space. There is good evidence that each of the visual subroutines (color, form, movement and depth perception) are handled simultaneously by distinct (but interdependent) pathways. This is called **parallel processing**. It contrasts with serial processing in which a task is segmented into several subroutines which must be executed, in sequence, one after the other. Parallel processing has the advantage of speed. The final visual representation is a unified percept in which all the independently processed subroutines must somehow be combined together. How this might be achieved by the brain is termed the binding problem, which applies not only to vision but other sensory modalities as well.

Visual perception involves processing of the retinal image so that its key features can be abstracted. The visual system is more concerned with regions of the visual world that are changing in time (movement) and space (contrast) than regions that are static. It is widely accepted that perception requires the existence of internal representations of the visual world which allow the brain to make hypotheses about what the retinal image is. Internal representations account for the fact that the visual system shows **pattern completion**, in that it can generate a complete percept even when the raw sensory data are incomplete or corrupted by noise, and **generalization**, the ability to recognize objects from a wide variety of vantage points. Internal representations are probably encoded in the firing patterns of assemblies of neurons. Some internal representations are specified during development and are immutable, but most probably depend on early learning. Mental images of objects are thought to be manifestations of the internal representations of the objects and can be manipulated by most people in predictable ways. Whenever an unresolvable mismatch occurs between the sensory input and the internal representation the result is a **visual illusion**.

Perceptual constancy is a key property of vision. Visual perception can be invariant over wide differences in the properties of the retinal image. For example, with **size constancy**, familiar objects do not diminish in size in proportion to the reduction of the retinal image, but appear larger than they should. **Color constancy** preserves the colors of objects in the face of alterations in the wavelength composition of the light source. Perceptual constancy permits successful object recognition under a wide variety of ambient conditions.

Sensitivity

The human eye is sensitive to the electromagnetic spectrum between the wavelengths 400 nm (violet) and 700 nm (red). The range of light intensity to which we are exposed is huge, about 10^{11}-fold. Although the human eye can respond to a single photon of light, 5–8 photons arriving within a short time are required to give the experience of a flash of light in the dark-adapted state. Because intensity is encoded by the visual system logarithmically, it is difficult to distinguish differences in intensity at high light levels.

Acuity

Visual acuity is the ability to see fine detail. It is highest at the central region of the retina, the **fovea**, and depends on the ambient illumination. Under ideal conditions two points of light can be separately resolved if they subtend an angle at the retina of one arc minute. For ruled gratings acuity is much better; lines can be resolved if separated by only a few arc seconds. Visual acuity, V, is defined by the Snellen fraction:

$$V = d/D$$

where d is the distance at which a symbol (e.g. a letter) can just be discriminated and D is the distance at which the symbol subtends an angle of one arc minute at the retina. An individual with good normal vision can just discriminate a symbol, the size of which subtends one arc minute at 20 feet (about 6 meters), from a distance of 20 feet; such a person has a visual acuity of 20/20. Loss of acuity in dim light occurs because insufficient photons fall on the retina to build up a complete image within the time span over which the light-sensitive cells can integrate the energy.

Depth perception

The retinal image is two dimensional but from it the visual system can infer the three-dimensional structure of the world. There are both monocular and binocular clues to depth perception. Monocular clues are most important for distant objects, where binocular cues cannot be used, and include:

- **Parallax**. Movement of the head causes an apparent movement of near objects with respect to distant ones. The closer the object the bigger this apparent movement.
- **Perspective**. Parallel lines appear to converge with distance. Artists from the early 1400s onwards used perspective as a major depth cue in painting.
- The relative sizes of objects of known dimensions.
- **Occultation**, in which a more distant object is partly hidden by a nearer one.
- **Shadows**.

The binocular cue to depth perception is called **stereopsis**. It is possible only for the field of view for which the two monocular visual fields overlap. Because the eyes are about 6.3 cm apart each has a slightly different view of the world and the image of a nearby object falls onto different horizontal positions on the left and right retinas, a phenomenon called **retinal** (**binocular**) **disparity**. This can be visualized by viewing a scene first through one eye, then the other, when nearby objects appear to jump sideways; **binocular parallax**. When the eyes are made to converge so as to fixate on a nearby point the images are formed at the fovea on each retina, and the two images are perceived as fused into a single point. All other points that are fused appear so because their images lie at corresponding positions on left and right retinas. Points in space that lie closer to or beyond those that form images at corresponding positions will generate binocular disparity (*Fig. 1*). Images of these points will also fuse. The brain is able to compute depth from disparity simply by comparing where the same pattern lies on left and right retinas. Stereopsis does not require form, movement or color. For binocular disparity that is too great (binocular disparity > 0.6 mm or 2° arc) fusion cannot occur and two images are seen; **double vision** (**diplopia**).

Color vision

Color vision permits boundaries to be seen between regions that have equal brightness, provided the spectrum of wavelengths they reflect is different. The spectrum of light reflected from an object depends on the wavelength composition of the illuminating light and the reflectance of the surface, but color vision is not just a matter of measuring all the wavelengths in the reflected light.

Color vision requires a minimum of two types of receptor that respond over different wavelength ranges. This is dichromatic color vision and is the case for all mammals except Old World monkeys, apes and humans. With two receptors the visual system can assign two brightness values for each pixel of the visual field. By comparing these values, colors may be perceived. For example, if a pixel reflects more short-wavelength light it will appear brighter to a short-wavelength

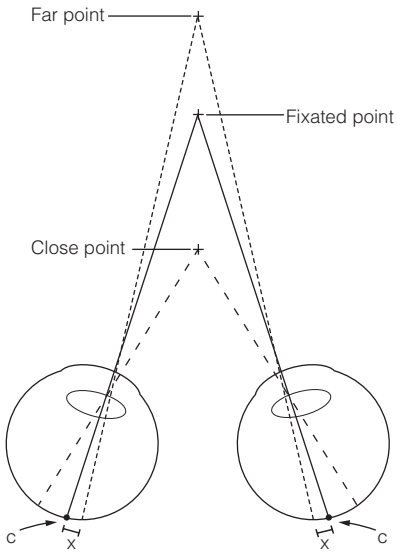


Fig. 1. Stereopsis. The fixated point produces images on corresponding positions, c, on the retinae so that the images are fused. The images of the far point are displaced from the corresponding points by a distance x giving a binocular disparity = 2x. A similar argument applies to the close point images.

receptor than a long-wavelength receptor, and will be seen as blue. If a pixel reflects more long-wavelength light it will be seen as red. In the case that a pixel reflects equal amounts of short- and long-wavelength light it will appear monochrome, either white or shades of gray depending on the intensity of the light.

Human color vision is trichromatic because the eye has three populations of receptors (cones) that can function in daylight, each sensitive to a different (but wide and overlapping) range of wavelengths. The three types of cones have maximum absorptions corresponding approximately to violet, green and yellow light. The wavelength of the light does not affect the character of the response of the cone. A given cone simply has a higher probability of absorbing a photon that is close to its peak wavelength. This means that the visual system has no way of detecting the absolute wavelength composition of any light. The trichromatic visual system abstracts three brightness values for an object and comparisons of these values determines the color.

Color vision has several remarkable properties. It shows **color constancy**. An object can be viewed under a variety of light sources with different spectral compositions, e.g. neon lighting, sunlight or tungsten light, and appear to be the same color even though the wavelengths of light it reflects in each case will be quite different. While some colors in the same pixel of visual space perceptually mix to produce other color categories (e.g. blue and green mix to give cyan) complementary colors (e.g. red and green) do not perceptually mix; reddish-green colors are never seen. This is **perceptual cancellation**.

Simultaneous color contrast is the perceptual facilitation of complementary colors that occurs across boundaries. For example a gray disk within a red background looks slightly green, while a gray disk in a green background appears slightly red. Each of these features can be accounted for in terms of visual system physiology.

G2 Eye and visual pathways

Key Notes

Structure of the eye

The eye has three layers. The tough outer sclera maintains the shape of the eye and serves as an attachment for extraocular muscles. The choroid is pigmented to prevent light being reflected within the eye. The inner layer is the light-sensitive retina. At the front the sclera becomes the transparent cornea responsible for most of the refraction of light rays entering the eye. The front of the choroid forms the ciliary body and iris. The biconvex lens is attached to the ciliary body by a suspensory ligament. The iris is a diaphragm surrounding the pupil and contains smooth muscles, which act as pupillary sphincter or dilator. In front of the lens is the aqueous humor, which determines the pressure of the eyeball. Behind the lens is the vitreous humor, a gel-like refractile medium.

Anatomy of visual pathways

The optic nerves meet at the optic chiasm where nerve fibers from the nasal half of each retina cross to the other side. Beyond this point nerve fibers enter the optic tract. A small number go to the pretectum which controls pupil and accommodation reflexes, others go to the superior colliculus, which mediates many visual reflexes, but most go to the lateral geniculate nucleus of the thalamus. From here the optic radiation goes to the primary visual cortex located in the occipital cortex. This last pathway is responsible for visual perception.

Visual reflexes

The amount of light entering the pupil can be altered 30-fold by changing the size of the pupil. The pupil light reflex causes pupil constriction in bright light. Light shone in one eye causes a light reflex in both eyes. The reflex pathway goes by way of optic nerve axons to the pretectum, which sends output to preganglionic parasympathetic neurons in the accessory oculomotor nucleus. The postganglionic cells lie in the ciliary ganglion and their axons supply the pupillary sphincter. Light stimulation elicits contraction of the muscle. For close objects greater refraction is needed to focus the image. This is achieved by the accommodation reflex; contraction of the ciliary body eases tension in the suspensory ligament, allowing the lens to become more spherical. The accommodation reflex is brought about by parasympathetic action on ciliary muscles. When observing a close object the eyes converge so that both eyes can fixate on it (the vergence reflex) and the pupils constrict, which increases depth of field and acuity.

Related topics

Attributes of vision (G1)
Early visual processing (G6)
Oculomotor control (G8)
Autonomic nervous system function (L5)

Structure of the eye

The eye consists of three layers enclosing its contents, the sclera, the choroid and the retina (*Fig. 1*). The **sclera** is a thick, stiff outer layer of connective tissue. At its anterior end it is continuous with the cornea. At its posterior end it becomes the dura mater covering the optic nerve. The function of the sclera is to maintain the shape of the eyeball and to serve as an attachment for the extraocular muscles. The **cornea** is a curved circular transparent layer at the front of the eye. Its outer layer is continuous with the **conjunctiva,** an epithelial sheet covering the front of the eyeball. Most of the focusing power of the eye is due to refraction of light by the cornea. The **choroid** is a thin, highly vascular layer, dark brown in color because of the presence of **choroidal pigment cells**. By absorbing light it prevents total internal reflection within the eye.

Anteriorly the choroid becomes the **ciliary body** and the **iris**. The ciliary body gives rise to numerous, thin **zonular fibers**, which attach to the capsule of the lens and constitute the **suspensory ligament**. Inside the ciliary body lies the ciliary muscle composed of smooth muscle fibers arranged in both radial and circular directions.

The iris is a diaphragm surrounding a central hole, the **pupil**. The iris contains two intraocular muscles which act in concert to control the size of the pupil. Innermost is a flat ring of circularly arranged smooth muscle fibers, the **pupillary sphincter**. Surrounding the sphincter is a thin layer of radially organized myoepithelial cells which form the **pupillary dilator**.

The innermost layer of the eye is the light-sensitive **retina**.

Aqueous humor is actively secreted by the epithelium of the ciliary body into the **posterior chamber**. It percolates through the pupil to the **anterior chamber** from where it drains into the venous system via the **canals of Schlemm** located in the **irideocorneal angle**. The pressure of the aqueous humor determines the pressure of the eyeball. This is normally less than 3 kPa. Obstruction to the proper drainage of the aqueous humor causes raised intraocular pressure, a condition called **glaucoma** which can result in blindness due to the reduced

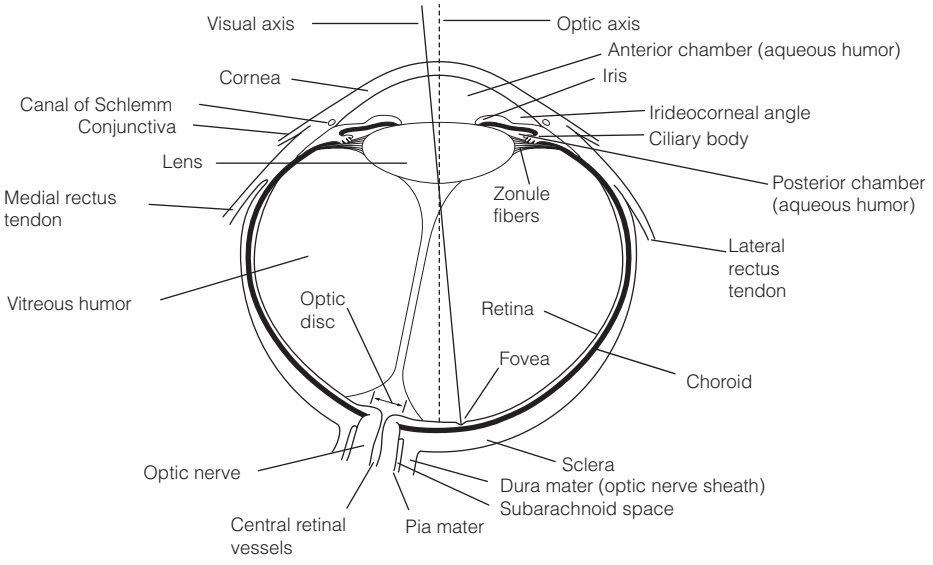


Fig. 1. Horizontal section through the right human eye.

perfusion of blood through the retina. Aqueous humor transports metabolic requirements (e.g. glucose, amino acids and ascorbate) for the lens and cornea, which have no blood supply. **Vitreous humor** is a gel of extracellular fluid which refracts light rays appropriately so that they come to focus on the retina.

The biconvex **lens** of the human eye has a diameter of 9 mm. It is encapsulated within an elastic connective tissue membrane, which is attached to the suspensory ligament.

Anatomy of visual pathways

The optic nerves meet in the midline at the **optic chiasm** (*Fig. 2*). Here, 53% of optic nerve fibers, those from the nasal halves of the retina, cross to the contralateral side in the **optic decussation**. Axons from the temporal halves of the retina remain on the ipsilateral side. Retinal axons leave the optic chiasm to enter the **optic tracts** from where they go to three targets. A small proportion go

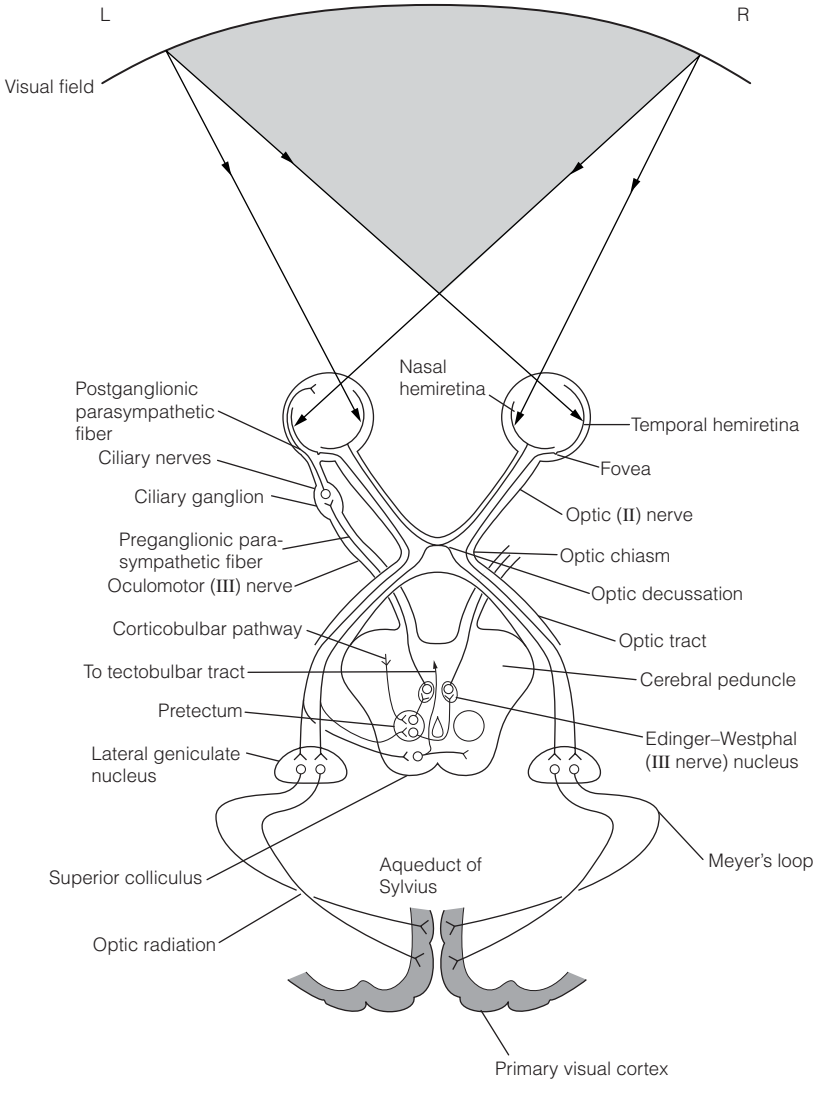


Fig. 2. Visual pathways. Reflex pathways are shown complete only on the left side. The direction of light rays from left and right halves of the visual field into the eyes is depicted. Note that light from the left visual field falls on the right halves of each retina (nasal hemiretina of the left eye, temporal hemiretina of the right eye) and light from the right visual field goes to the left hemiretinae. Binocular vision is possible only in the shaded region.

to the **pretectum** of the midbrain which controls pupil and accommodation reflexes. Other axons go to the **superior colliculus** in the tectum of the midbrain which organizes several visual reflexes, and a visual pathway runs to the hypothalamus to entrain circadian rhythms. The great majority of axons go to the **lateral geniculate nucleus** (LGN), part of the thalamus. From here, the **optic radiation** sweeps to the medial aspect of the pole of the occipital cortex, most axons terminating in layer IV of Brodmann area 17, the **striate** or **primary visual cortex** (**V1**). The retina–LGN–visual cortex pathway is responsible for visual perception. Visual defects characterized clinically can provide clues to the site of the lesion within the visual system (*Fig. 3*).

Visual reflexes

The **pupil light reflex** controls the amount of light entering the eye by altering pupil size. This ranges between 1.5 and 8 mm in diameter, being maximal in complete darkness. Although this allows only a 30-fold change in light entry (which is small compared with the range of light intensity the visual system can

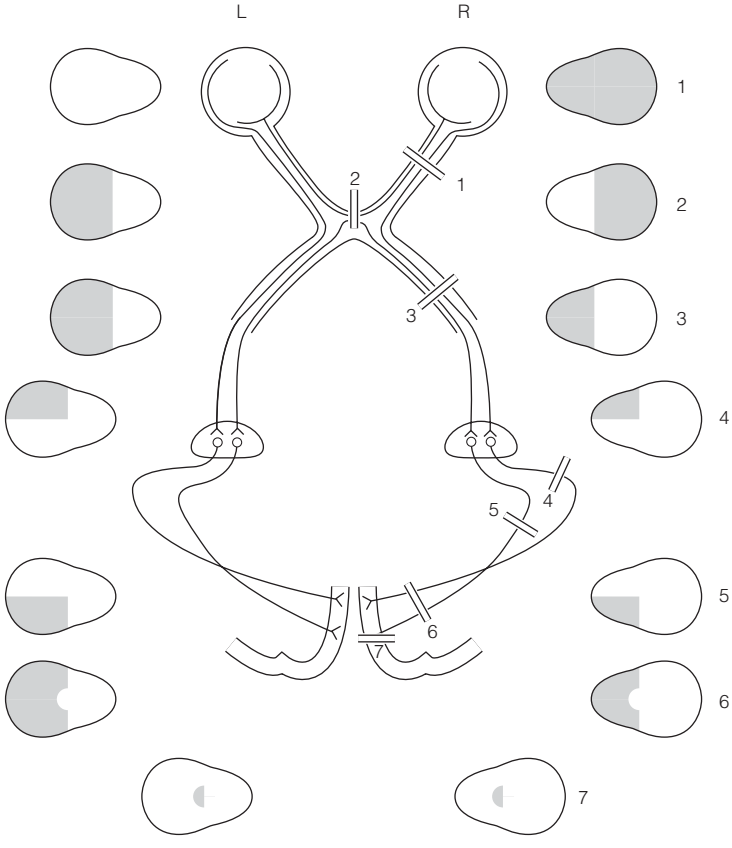


Fig. 3. Visual defects arising from damage to visual pathways. Lesion 2 commonly arises from compression of the central part of the optic chiasm by a pituitary tumor. Optic tract lesions (3) are rare. Optic radiation damage is usually due to an infarct or a tumor in the temporal lobe (4) or parietal lobe (5). Lesion 6 is usually caused by occlusion of the posterior cerebral artery. The fovea area is spared because it is supplied by the middle cerebral artery. Damage to one occipital lobe pole is usually caused by trauma and, since the fovea has by far the largest representation, selective loss of foveal vision is typical (lesion 7).

experience) the reflex is useful because it operates over the light levels typically encountered during daylight. Light shone in one eye produces pupil constriction of the same eye (the **direct** reflex) *and* of the contralateral eye (the **consensual** reflex) because of reciprocal crossed connections in the midbrain. The reflex pathway is shown in *Fig. 4*. Optic nerve axons synapse in the pretectum, which sends output to preganglionic parasympathetic fibers in the **Edinger–Westphal (accessory) oculomotor** nucleus. These autonomic fibers travel in the oculomotor nerve to the ciliary ganglion which lies in the orbit. Postganglionic fibers from there go to the pupillary sphincter. Light stimulation of optic nerve fibers excites the parasympathetic terminals to release acetylcholine, which contracts the sphincter. The latency of the reflex is about 200 ms. Lesions of the optic and oculomotor nerves, or of the midbrain can be diagnosed by examining defects in the pupil light reflex (*Fig. 4*).

For close objects, light rays are diverging as they enter the eye and so greater refraction is needed to bring them to focus at the fovea. This is achieved by the **accommodation reflex**. Contraction of the ciliary muscles pulls the ciliary body forwards and inwards, easing the tension in the suspensory ligament and lens capsule, allowing the lens to become more spherical and reducing its focal length. The stimulus for the accommodation reflex is blurring of the retinal image. This is monitored by the visual cortex which projects to the pretectum via the corticobulbar pathway. Via connections between the pretectum and the Edinger–Westphal nucleus, parasympathetic fibers are activated which contract the ciliary muscles. Accommodation occurs in both eyes equally and takes nearly one second to execute.

Observing a close object also causes convergence of the visual axes of both eyes, the **vergence reflex**. This enables both eyes to fix their gaze on an object. In addition, the degree of convergence provides a cue for stereopsis, since the closer an object is, the greater the convergence must be. Vergence can be triggered by a blurred retinal image or by consciously altering gaze to a point at a different distance. The circuitry is from the visual cortex to a region of the frontal cortex concerned with the planning and execution of eye movements.

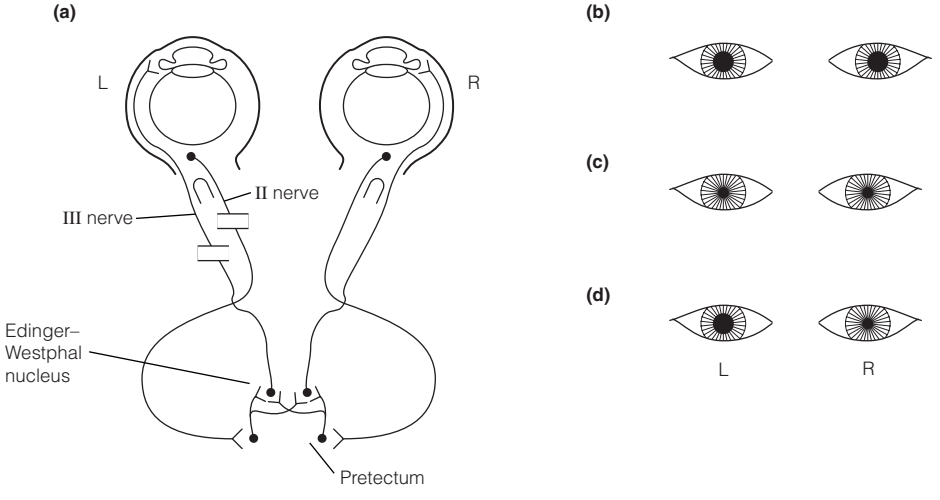


Fig. 4. Altered pupil reflexes following damage to either optic (II) or oculomotor (III) nerves on the left side: (a) reflex pathway; (b) optic nerve damage, left eye stimulated; (c) optic nerve damage, right eye stimulated; (d) oculomotor nerve damage, left or right eye stimulated.

Accommodation and convergence are both accompanied by pupil constriction, which has two effects. Firstly, it increases the depth of field, the distance range in which objects are in focus. Secondly, it reduces spherical aberration (a defect of lenses in which parallel rays of light are not brought to focus at the same point); this improves acuity when looking at near objects. Pupil constriction in this instance is mediated by a pathway from the primary visual cortex to the pretectum. The **Argyll–Robertson pupil** fails to constrict in response to the light reflex, but will constrict during accommodation and convergence. It results from damage to the light reflex pathway in the region of the tectum or aqueduct of Sylvius.

G3 RETINA

Key Notes

Structure of the retina

The light-sensitive retina has five basic neuron types arranged in several layers. The outer nuclear layer consists of the cell bodies of the photoreceptors and the inner nuclear layer contains the cell bodies of the retinal interneurons, bipolar cells, horizontal cells and amacrine cells. The innermost ganglion cell layer harbors the output cells of the retina which send their axons into the optic nerve. The number of ganglion cell axons is 100-fold less than the number of photoreceptors, which indicates that a great deal of visual processing must go on in the retina. Only ganglion cells are excitable, all other retinal neurons signal via passively conducted synaptic potentials. The central 1.5 mm of retina is the fovea and has the highest visual acuity. The optic disk, the site where optic nerve and blood vessels pierce the retina, accounts for the blind spot.

Rod and cone cells

Rod photoreceptors, situated throughout the retina except the fovea and optic disk, are extremely sensitive to light and are used in dim light vision. In daylight rod cells saturate and are unresponsive. Rod cell vision has a low acuity because many rod cell signals converge, which, while maximizing light sensitivity, causes loss of information about location. Cone cells are found at the fovea. They are 1000-fold less sensitive to light than the rods, fail in dim light, but do not saturate except in very bright light so operate in daylight vision. Daylight vision has high acuity because there is little convergence of cone signals. There are three populations of cones, distinguished by the range of wavelengths to which they are sensitive. Short-wavelength (blue) cones constitute only a few percent of all cones and are absent from the center of the fovea. Medium- (green) and long- (red) wavelength cones are randomly but patchily distributed, so color vision cannot be used to see fine detail. The maximum sensitivity of the human eye is to yellow light but this shifts towards green in low light conditions when rod cells become active. Going from bright to dim light causes a massive increase in sensitivity of the retina to occur, dark adaptation, which takes about 30 minutes.

Color blindness

Virtually all color blindness is genetic and caused by loss or abnormality of cones. Trichromats retain all their cones, but have a defect in one type. Dichromats lack one population of cone cell while monochromats, missing two or all three types of cone, have no color vision. Since the genes for the medium- and long-wavelength cone pigments lie on the X chromosome a defect in either, resulting in red–green color blindness, is an X-linked recessive trait afflicting predominantly males.

Related topics Attributes of vision (G1) Retinal processing (G5)

Structure of the retina

The retina is the light-sensitive innermost layer of the eye. It contains five distinct neuron types, interconnected in circuits that are repeated millions of times. Under light microscopy, the retina is seen to be composed of several layers (*Fig. 1*).

Closest to the choroid is a single layer of pigmented epithelial cells. These contain melanin and absorb light not absorbed by the retina so that it is not reflected back to degrade the image. The outer nuclear layer contains the cell bodies of the photoreceptors. The inner nuclear layer consists of the cell bodies of retinal interneurons; bipolar cells, horizontal cells and amacrine cells. The ganglion cell layer contains ganglion cell bodies, the axons of which provide the output of the retina via the optic nerve. Ganglion cell axons only become myelinated at the optic disk (see below). The two plexiform layers are the locations for the connections between the retinal cells. Note that although light must pass through the full thickness of the retina before striking the photoreceptors the retina is highly transparent. Each human retina has roughly 10^8 photoreceptors, but an output via only about 10^6 optic nerve axons. This is a massive convergence and shows that considerable processing of visual input is done by the retina to achieve this level of data compression.

Although all the cells in the retina (except pigment cells) are neurons, only the ganglion cells are able to fire action potentials. Photoreceptors and retinal interneurons signal by way of passively conducted synaptic potentials.

The gaze of the eye is usually adjusted so that the images are brought to focus at the **fovea**. This region of the retina with a diameter of 1.5 mm has the greatest visual acuity. The high visual acuity at the fovea is in part due to:

- the great density of photoreceptors;
- the displacement of overlying layers of the retina to the side so that light hits the photoreceptors directly;
- the lack of blood vessels;
- the fovea lying on the optical axis of the eye so image distortion by the optics (e.g. spherical or chromatic aberration) is minimal.

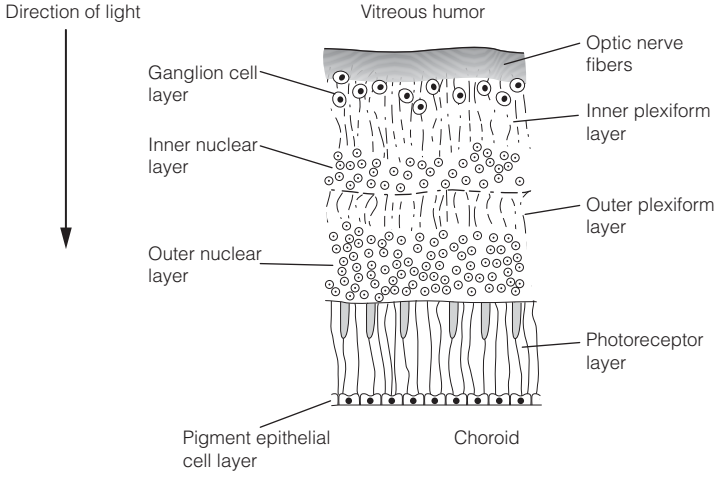


Fig. 1. Section through the human retina viewed by light microscopy (× 1500).

About 4 mm from the fovea towards the nose lies the **optic disk**, where optic nerve fibers and retinal blood vessels pierce the retina. This region lacks photoreceptors and accounts for the **blind spot** that occurs in the visual fields.

Retinal detachment can occur as a result of head trauma, the separation occurring at the interface between the pigment epithelial cells and the photoreceptors.

Rod and cone cells

There are two populations of photoreceptors, rods and cones. Only about 10% of light entering the eye succeeds in exciting photoreceptors, the rest is scattered or absorbed.

Rod cells are 20-fold more numerous than cones and are distributed across the entire retina except for the fovea and the optic disk (*Fig. 2*). They are about 1000 times more sensitive to light than cone cells and are used for **scotopic** (dim light) vision. Indeed, rod cells saturate in daylight so that they become unresponsive. The high sensitivity of rod cells comes about partly because they integrate responses to incoming photons over a long period (~100 ms). The disadvantage of this is that rods are unable to discern flickering light if the flicker rate is faster than about 12 Hz. A second feature contributing to rod cell sensitivity is that they are able to greatly amplify the effect of the photons impinging upon them. Scotopic vision has a low **acuity** for two reasons. Firstly, the image formed on the peripheral retina is quite distorted. Secondly, many rods converge onto a single bipolar cell. Although this maximizes the chance of a rod bipolar cell responding to a dim light signal, because it can capture light signals from a large area of retina, by the same token information about localization is less precise.

Cone photoreceptors are most dense at the fovea and their numbers fall off sharply beyond 5 degrees of it. They have a low sensitivity to light and do not saturate except in very intense light, so are used for **photopic** (daylight) vision. Photopic vision has high acuity because there is little or no convergence between cone cells and bipolar cells. Cone cells integrate photon responses over a short time and so are able to resolve a flicker frequency of less than about 55 Hz.

At low light levels (e.g. dusk) rod cells operate alongside cones, boosting the cone cell signal by means of electrical synapses to preserve color perception. This is termed **mesopic** vision.

There are three populations of cone cells which differ in their spectral sensitivity (*Fig. 3*).

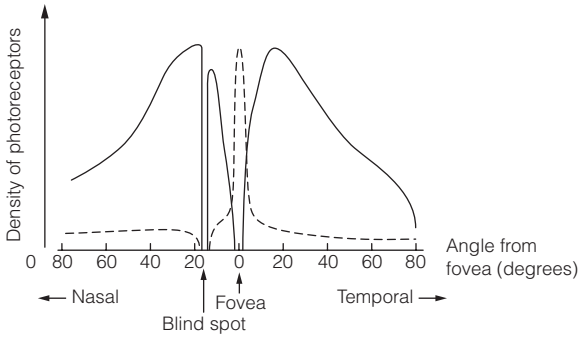


Fig. 2. Distribution of cone (----) and rod (——) photoreceptors in the human retina.

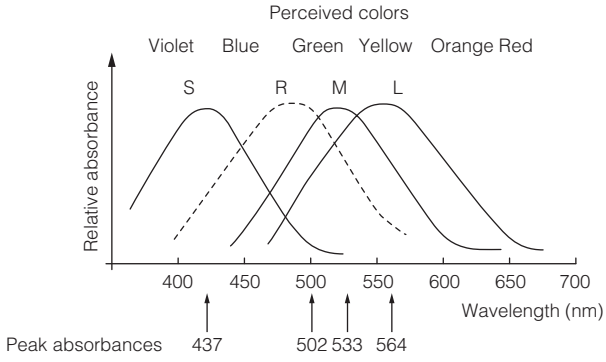


Fig. 3. Spectral sensitivities of photoreceptors: S, short wavelength cones; R, rods; M, medium wavelength cones; L, long wavelength cones.

Although more properly called short- (S), medium- (M) and long- (L) wavelength cones they are often referred to as blue, green and red cones, respectively, although their peak sensitivities are not best described by these colors. The S cones are sensitive to wavelengths down to 315 nm. However, the normal eye does not see wavelengths shorter than 400 nm because they are absorbed by the lens. Absorption of ultraviolet light is an important cause of **cataract**, in which the lens becomes progressively more opaque. This is especially prevalent in countries at lower latitudes where sunlight is most intense.

Color vision requires comparisons of the relative strengths in the outputs of the S, M and L cones. S cones constitute only about 5–10% of the total number of cones and are absent from the center of the fovea. This is because the eye suffers from **chromatic aberration**, in which short-wavelength light is not brought to focus at the same point as longer wavelengths and so causes slight blurring of the image. This would compromise high-acuity vision. Consequently, color vision at the central fovea is dichromatic and furthermore, M and L cones are distributed randomly leaving patches in which there is only one population of cone. These features mean that color vision is coarse grained and cannot resolve fine detail. Scotopic vision is achromatic because all rod cells have the same spectral sensitivity curve. They are unable to distinguish between wavelengths on the rising and falling limbs of the spectrum that excite the cell to the same extent.

Under scotopic vision the wavelength sensitivity of the eye is determined by the rod cells and peaks at around 500 nm. Under photopic conditions the wavelength sensitivity is governed by cones and is maximal at 555 nm. This shift in wavelength sensitivity between scotopic and photopic vision is called the **Purkinje shift**. It accounts for the fact that as dusk falls the last color sensations to be lost are blues and greens.

When moving from bright to very dim light the sensitivity of the retina to light increases a million-fold over a period of 30 minutes or longer. This is called dark adaptation and is a property of photoreceptors. **Dark adaptation** has two phases (*Fig. 4*). The first is due to cone cells which increase sensitivity about 100-fold, the second longer phase is due to rod cells.

Light adaptation occurs when going from dim to brightly lit conditions. It is very much faster than dark adaptation.

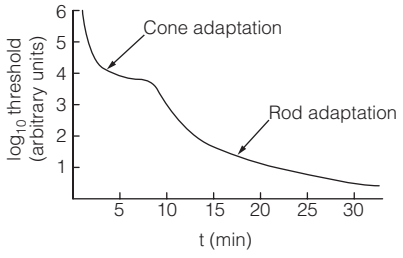


Fig. 4. The timecourse of dark adaptation.

Color blindness

Almost all **color blindness** is genetic in origin and results from loss or abnormality of cone cells. **Trichromats** retain all three cone populations but have abnormalities in the opsins (visual pigments) of S, M or L cones, most commonly the M opsin. **Dichromats** have only two cone populations and their visual defect is more severe than that of the trichromats. Dichromats are classified as **protanopes** (without L cones), **deuteranopes** (lacking M cones) and **tritanopes** (who lack S cones). Most affected, but rare, are **monochromats** who, lacking two or all three types of cone cell, have no color vision. Those without any cones have no photopic vision and are effectively blind in daylight.

Abnormality of S opsin or loss of S cones is quite rare and afflicted individuals may be unable to distinguish colors having a short-wavelength component (violet) from those without (yellow); both appear gray. The gene for S opsin is on chromosome 7 and defects in S cone vision are inherited as an autosomal dominant trait. In defects of M or L cone vision it is not possible to discriminate red from green or either from gray. M and L opsin genes lie on the X chromosome so red–green color blindness is an X-linked recessive trait. Not surprisingly it is much more common in males (4–8% in Europe, depending on ethnicity) than females (about 0.4%).

Loss of rod cell function has also been described. Afflicted individuals have only narrow, central visual fields and, lacking scotopic vision, are blind whenever light levels fall below those needed to excite cones.

G4 PHOTOTRANSDUCTION

Key Notes

Photoreceptor structure

Photoreceptors have an inner segment (containing the nucleus), which has a synaptic terminal and an outer segment, the plasma membrane of which is invaginated into deep folds to form disks. The visual pigments, rhodopsin of rods or cone opsins are situated in the disk membrane. Photoreceptors cannot divide but their outer segments are continuously turned over.

Photoreceptor transduction

Photoreceptors in the dark have quite a low (depolarized) membrane potential because of the influx of sodium and calcium ions through cyclic nucleotide-gated ion channels in the outer segment plasma membrane. Light causes a hyperpolarizing receptor potential by closing the channels. Transduction in rod cells starts when photons are captured by the prosthetic group in rhodopsin, retinal, which undergoes photo-isomerization. This activates the rhodopsin which in consequence couples with a G protein, transducin. Transducin stimulates a phosphodiesterase that hydrolyzes cyclic guanosine monophosphate, reducing its concentration, so closing the cyclic nucleotide-gated channels. Subsequently the photo-isomerized retinal dissociates from the rhodopsin leaving the pigment bleached. Dark adaptation is the regeneration of rhodopsin.

Related topics Metabotropic receptors (D2) Retina (G3)

Photoreceptor structure

Rod and cone photoreceptors have similar structures (*Fig. 1*). Photoreceptors have diameters of 1–4 μm, being smaller at the fovea, a factor contributing to the higher visual acuity achieved by this area. The inner segment contains the nucleus, is rich in mitochondria and has an axon-like process connected to a synaptic terminal. The outer segment in the cone cell has its plasma membrane invaginated into numerous closely packed parallel folds, forming disks. In rod cells the disks are pinched off the plasma membrane to become completely intracellular. The disk membrane is densely packed with **visual pigment**. In rod cells this is **rhodopsin**. Each population of cone cells has its characteristic **cone opsin**. The outer segment is continually regenerated from the base, whilst its apical tip is phagocytosed by pigment epithelial cells at the rate of 3–4 disks per hour. Failure of this phagocytotic mechanism may underlie some forms of the X-linked disorder **retinitis pigmentosa**. Photoreceptors are neuroepithelial cells and incapable of mitotic division.

Photoreceptor transduction

The resting potential of the photoreceptor plasma membrane in the dark is quite low, about –40 mV. Light produces a hyperpolarizing receptor potential, the amplitude of which is related to the light intensity (*Fig. 2*).

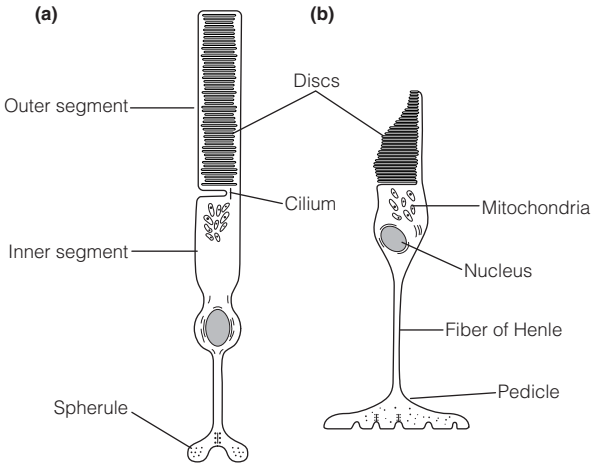


Fig. 1. Photoreceptors: (a) a rod cell; (b) a cone cell.

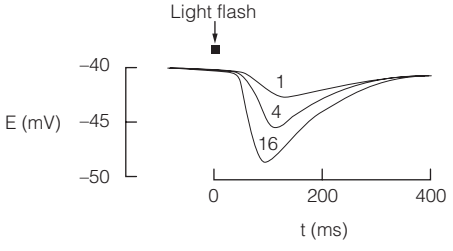


Fig. 2. Cone receptor potentials in response to light flashes of increasing relative intensity (1, 4 and 16).

The hyperpolarization is produced by the light-evoked closure of cyclic nucleotide-gated cation channels that are open in the dark. The normal, relatively depolarized, state of the photoreceptor is caused by the flow of a **dark current**, as shown in *Fig. 3*.

The cyclic nucleotide-gated cation channel allows Na^+ and Ca^{2+} ions to flow into the outer segment in darkness. Na^+ ions are actively extruded by the Na^+–K^+ ATPase in the inner segment. Ca^{2+} leaves the photoreceptor via a Na^+–K^+–Ca^{2+} transporter. When light photons strike the outer segments a cascade of biochemical events is initiated which results in the closure of the cation channels, reducing the dark current, and hyperpolarizing the photoreceptor. The transduction process in rod cells is well understood. Rhodopsin consists of a G protein receptor, **opsin**, and a prosthetic group, **retinal**, synthesized by retinol dehydrogenase from **retinol** (**vitamin A**). Retinol cannot be synthesized *de novo* in mammals and hence must be supplied in the diet. Lack of vitamin A in the diet causes night blindness and if prolonged results in irreversible damage to rod cells.

In the dark, retinal is present as the 11-*cis*-isomer. Light causes photo-isomerization to the all-*trans* isomer. The isomerization occurs within a few picoseconds of the photon being absorbed and triggers a series of conformational changes in the rhodopsin to form photoexcited rhodopsin (R*). Photoexcited

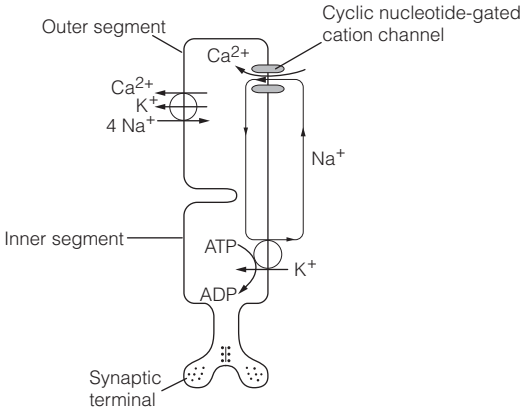


Fig. 3. The ionic basis of the rod photoreceptor dark current.

rhodopsin couples with a G protein, **transducin** (G_t) and exchange of GDP for GTP occurs. The GTP-bound form of the transducin alpha subunit activates a phosphodiesterase (PDE) which catalyzes the hydrolysis of 3′,5′-cyclic guanosine monophosphate (cGMP) to 5′-GMP. This reduces the concentration of cGMP in the photoreceptor and so the cation channels, normally kept open by the cyclic nucleotide, close. The opening of each channel requires three molecules of cGMP with high positive cooperativity. This means that small changes in cGMP concentration effect big changes in the number of open channels. This sequence of events is summarized below (*Fig. 4*).

This second messenger cascade has a large amplification. A single photon activates about 500 transducin molecules, closes hundreds of cation channels, blocking the influx of 10^6 Na^+ ions to cause a hyperpolarization of about 1 mV.

Several mechanisms act sequentially to terminate the cascade:

- Like other G proteins, transducin has an intrinsic GTPase which hydrolyzes the bound GTP to GDP, stopping the activation of PDE.
- Photoexcited rhodopsin is phosphorylated by **rhodopsin kinase**, and then binds **arrestin** which blocks the binding of transducin.
- Within a few seconds the bond between retinal and opsin in photoexcited rhodopsin spontaneously hydrolyzes and the all-*trans* retinal diffuses away

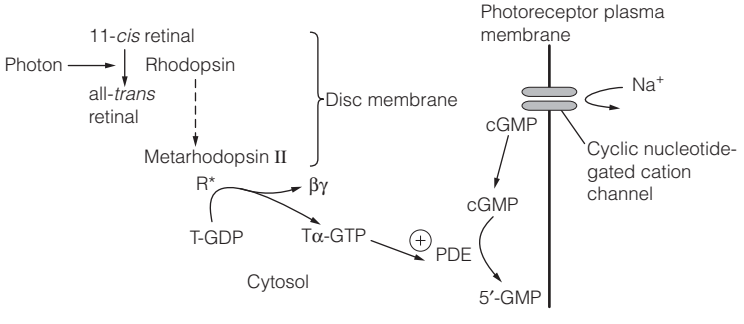


Fig. 4. The role of transducin in photoreceptor transduction. T, transducin; PDE, phosphodiesterase.

from the opsin. In high light levels most of the rhodopsin exists in this dissociated state in which it is described as being **bleached** and the rod said to be **saturated**. Regeneration of rhodopsin occurs in the dark: retinal isomerase catalyzes the isomerization of the all-*trans* isomer to the 11-*cis* isomer which then reassociates with the opsin. This process underlies **dark adaptation**.

Restoration of the dark state, in addition, requires the synthesis of cGMP. This is catalyzed by guanylate cyclase.

Light adaptation, in which photoreceptors become less sensitive during light exposure, allows them to respond to levels of illumination that vary by as much as four orders of magnitude (*Fig. 5*). Light-evoked closure of the cation channels reduces Ca^{2+} influx, so the Ca^{2+} concentration in the rod outer segment falls. Since Ca^{2+} normally inhibits the guanylyl cyclase needed for cGMP synthesis, this drop in Ca^{2+} concentration increases the production of cGMP, offsetting its destruction by the light.

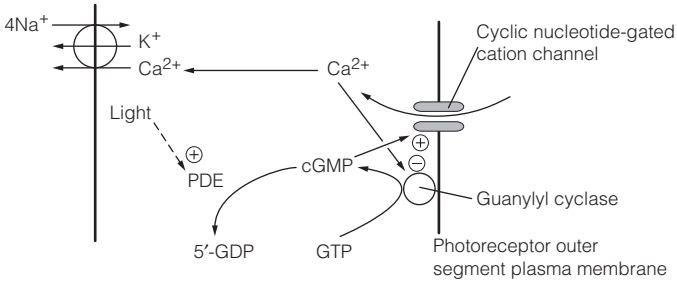


Fig. 5. The role of Ca^{2+} in photoreceptor light adaptation.

G5 Retinal processing

Key Notes

Bipolar cells and on and off channels

Photoreceptors synapse with either invaginating bipolar cells which depolarize in response to light, or flat bipolar cells which hyperpolarize when light stimulated. Midget bipolar cells, which get their input from cones, are of either type and synapse directly with ganglion cells, which respond in the same way to light as their bipolar cells; light increases the firing of on ganglion cells, and silences off ganglion cells. Hence the retina has on channels, formed from cone-depolarizing bipolar cell (on ganglion cell), and off channels, routed through cone-hyperpolarizing bipolar cell (off ganglion cell). The transmitter of all photoreceptors is glutamate and the opposite responses of the two types of bipolar cell are because they have different glutamate receptors. On channels signal the presence of local bright patches and off channels local dark regions of an image.

Horizontal cells and lateral inhibition

Bipolar and ganglion cell receptive fields are circular and divided into center and surround. Light stimulation has the opposite effect on the cell depending on whether it falls on the center or surround. This is lateral inhibition and its effect is to enhance contrast at boundaries. It is mediated by GABAergic horizontal cells that form reciprocal connections with adjacent photoreceptors. Because horizontal cells are heavily interconnected, the surround signal they generate represents light intensity averaged over an area of retina.

Ganglion cells

There are two populations of ganglion cell. P ganglion cells are small, slowly conducting, get their input from single cone types, are wavelength selective and show sustained responses. They are implicated in form and color vision. P ganglion cells exhibit color opponency, a sort of lateral inhibition in which they are excited by one type of cone but inhibited by one or both of the others. M ganglion cells in contrast are large, rapidly conducting, get input from middle- and long-wavelength cones together and show transient responses. M ganglion cells detect brightness (but not color) contrast and movement.

Rod signaling

In daylight only the cone channels are functional. At dusk rod cells become sensitive but transmit their signals via gap junctions to cone cells, boosting their function to maintain high acuity and color vision. When it is very dark cone cells fail and the rods cells signal exclusively through their own pathway via depolarizing (rod) bipolar cells and amacrine cells.

Amacrine cells

These interneurons, with neurites that have properties of both axons and dendrites, are extremely diverse. Different types are involved in rod cell pathways, surround inhibition and in signaling the direction of movement of a stimulus.

Related topics

Location coding (E3)
Attributes of vision (G1)
Retina (G3)

Bipolar cells and on and off channels

Photoreceptors synapse with bipolar cells. Two types of bipolar cell can be distinguished on the basis of both morphology and physiological responses. Those with processes that form **triad ribbon** synapses (*Fig. 1*) deep in the photoreceptor terminal are **invaginating bipolar** cells and they depolarize in response to light striking the photoreceptor. Triad ribbon synapses are so called because they have three postsynaptic components, the bipolar cell dendrite and dendrites of two horizontal cells. By contrast, **flat bipolar** cells form superficial **basal** synapses with photoreceptors and hyperpolarize in response to light shone on the photoreceptor.

Cone cells form synapses with **midget bipolar** cells (so called because of their size) of either one or the other type. Midget bipolar cells synapse directly with **ganglion cells** which respond to light in the same sense as their bipolar cells. This arrangement gives rise to two labeled lines: **on channels** are formed by cone-depolarizing bipolar cell (on ganglion cell), whereas **off channels** are cone-hyperpolarizing bipolar cell (off ganglion cell). On ganglion cells are depolarized and increase their firing rate as a function of light intensity. Off ganglion cells are silenced by hyperpolarization (*Fig. 2*).

All photoreceptors use glutamate as a transmitter. The opposite responses of invaginating and flat bipolar cells come about because they have different glutamate receptors. For invaginating bipolar cells, tonic release of glutamate from

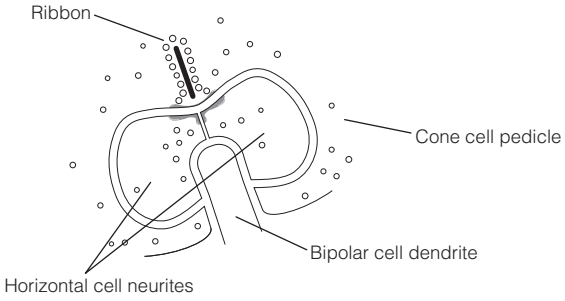


Fig. 1. A triad ribbon (invaginating) synapse.

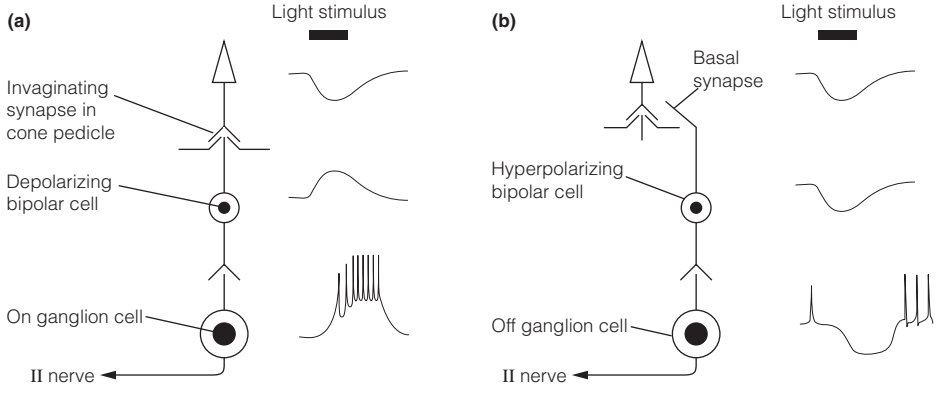


Fig. 2. (a) On channel and (b) off channel in the retina. In each case electrophysiological responses of the cells to light stimulation recorded intracellularly is shown on the right. All cells depicted use glutamate as a neurotransmitter.

photoreceptors in the dark is inhibitory. When light hyperpolarizes the photoreceptor glutamate release is suppressed, and inhibition lifted, so the bipolar cell depolarizes. For flat bipolar cells the response to tonic glutamate release is excitatory and light, by reducing that excitation, causes the bipolar cell to hyperpolarize.

On channels respond with increased firing to light levels that are greater than the local average. Off channels show increased firing in response to dark regions, i.e. where light levels are lower than the local average. In this way the existence of separate on and off channels is a mechanism to enhance the boundaries between regions that reflect different amounts of light. It is one of several processes by which the visual system is adapted to respond preferentially to stimulus change as opposed to steady-state stimulation.

Horizontal cells and lateral inhibition

An important mechanism for enhancing contrast in the retina is lateral inhibition brought about by **horizontal cells**. Lateral inhibition can be seen in the receptive fields (RFs) of both bipolar and ganglion cells. These are circular and divided into an inner center and an outer surround. Stimulation of these two regions separately produces opposite effects on the cells. In the case of an on ganglion cell, for example (see *Fig. 3*), background firing rate is dramatically increased by illumination of the center but silenced by light on the surround. When light fills the whole of the receptive field there is little change in the background firing rate. Off ganglion cell RFs have the converse response, with center illumination producing inhibition and surround illumination excitation.

Lateral inhibition arises because horizontal cells form reciprocal connections at triad synapses between neighboring photoreceptors. The details are a bit complicated. In the dark, horizontal cells are excited by glutamate release from photoreceptors but themselves release GABA which tends to inhibit the photoreceptors. Light which hyperpolarizes surrounding photoreceptors causes them to secrete less glutamate so reducing horizontal cell excitation. This means that GABA release from the horizontal cells is lowered, which in turn allows the central cone to depolarize somewhat, so that it releases more glutamate. The

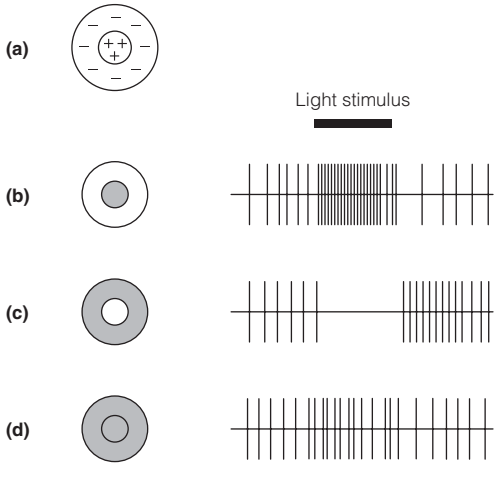


Fig. 3. Extracellular recording from on ganglion cells: (a) receptive field; (b) central illumination; (c) surround illumination; (d) overall illumination.

final step in the sequence depends precisely on the type of bipolar cell the central cone synapses with. If it is a depolarizing (on) bipolar cell the increased glutamate will cause it to hyperpolarize, as glutamate is inhibitory at invaginating synapses. If it is a hyperpolarizing (off) bipolar cell the increased glutamate will make the bipolar cell depolarize. Note that in each case the bipolar cell response (and hence that of the ganglion cell with which it synapses) is the opposite for surround compared with central illumination (*Fig. 4*).

Horizontal cells are extensively interconnected via gap junctions forming a network which spans an area of retina termed the **S space**. The S space horizontal cells provide the signal for surround inhibition and it is thought that this signal is a measure of the mean luminance over quite a wide area of retina.

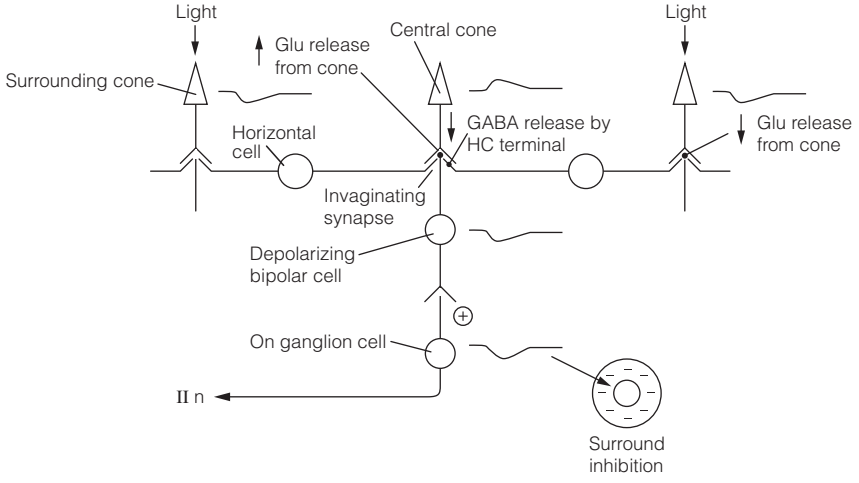


Fig. 4. The mechanism of lateral inhibition by horizontal cells (HC).

Ganglion cells

Ganglion cells are the output cells of the retina. Their axons become myelinated at the optic disk and form the optic (second cranial) nerve. Ganglion cells are the only retinal cells capable of firing action potentials. In the macaque Old World monkey (which has vision similar to that of humans) there are two major populations of ganglion cells. **Parvocellular** (**P**) **ganglion** cells are small and by far the most numerous; about a million in each retina. **Magnocellular** (**M**) **ganglion** cells are large and number about 100 000 per retina.

These two types differ in several important respects:

- P cells have smaller RFs than M cells.
- Because of their smaller size, P cells have a slower conduction velocity than M cells.
- Whereas P cells often show sustained responses, M cells respond transiently to a prolonged visual stimulus.
- P cells are usually wavelength selective whereas M cells are not.
- M cells are much more sensitive than P cells to low-contrast stimuli.

From these differences it can be inferred that P cells must get their input from single cones, or from several cones with the same wavelength sensitivity (S, M or L). By contrast, M cells get input from M and L cones together (but not S

cones), and from rods. Hence P, but not M, cells must be involved in color vision. The rapid, transient responses of M cells make them adapted for motion detection. The small RFs of P cells, and their sustained responses are suitable for fine-form discrimination. The distinct functional properties of P and M ganglion cells are the starting point for parallel processing in the visual system.

P cells are **color single opponent cells**. They have RFs that are excited by one type of cone cell but inhibited by another. Two types of P cell can be distinguished by the nature of their RFs (*Fig. 5*). Most common are the **red–green cells** in which the RF compares input from M and L cones. **Blue–yellow cells** do not have a center surround pattern but are either excited by S cones and inhibited by a combined M plus L cone signal, or vice versa. They are called blue–yellow cells because combining the inputs of M and L cones gives the sensation of yellow.

The red–green cells respond differently to small or large spots of light. For example, a green-on/red-off cell will be *equally* stimulated by a small green or white spot that covers the center of the RF. This is because white light contains the wavelengths that excite the green cones. However, the same cell will be excited by a large green spot but not by a large white spot. This is because the white light contains the wavelengths that stimulate the red-off surround. A large red spot will silence the cell. So, in general these cells are more wavelength selective for large stimuli than small ones. For small spots, they cannot distinguish red or green from white light, but they can signal brightness.

M cells get *combined* input from two types of cones so they are called **broad band** cells and their RFs measure only brightness contrast; M cells are color blind.

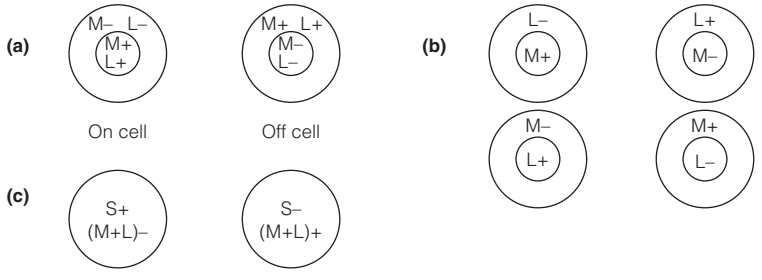


Fig. 5. Receptive fields of retinal ganglion cells: (a) M ganglion cells; (b) concentric single opponent (red–green) P cells; (c) coextensive single opponent (blue–yellow) P cells.

Rod signaling

Signaling by rod cells depends on the light intensity. At high light levels rods are saturated, and only cones operate. In the partially dark-adapted eye (e.g. at dusk) rod cells come on stream but signal through gap junctions to neighboring cones. This effectively augments cone cell function so maintaining acuity and color vision (mesopic vision). However, when it is very dark (e.g. moonless night sky) cone cells fail even with signal boosting from rods. Rod cells in the dark will have a greater influx of Ca^{2+} via the dark current. One effect of this increase in Ca^{2+} is to close the gap junctions between the rod and cone cells. Rod signaling is now relayed via depolarizing (rod) bipolar cells which synapse with a population of amacrine cells. The effect of this is to increase contrast sensitivity.

Amacrine cells

Amacrine cells have no axons but their extensive neurites share properties of both axons and dendrites. They are a very diverse group morphologically, and most of the neurotransmitters identified in the nervous system are used by one or other of the 30 or so types of amacrine cell.

Amacrine cells are implicated in rod signaling, surround inhibition and detecting the direction of motion of an object across the visual field. Dopaminergic amacrine cells are only about 1% of all amacrine cells but their long dendrites interconnect, possibly via gap junctions, to form a network. These cells get input from cone bipolar cells so the network is able to signal average illumination which is used to produce surround inhibition. In the dark-adapted eye, ganglion cells become much more sensitive to light by virtue of the fact that this dopamine surround inhibition is turned off. Some ganglion cells are sensitive to the direction of motion of a stimulus. Direction sensitivity is conferred by amacrine cell circuits.

G6 Early visual processing

Key Notes

Lateral geniculate nucleus

The sorting of fibers in the optic chiasm means that the left LGN maps the right side of the visual field. The primate LGN has six layers, two magnocellular layers get input from M ganglion cells and contain movement-sensitive cells, four parvocellular layers are innervated by P ganglion cells, and cells in these layers are wavelength selective. Each layer gets a precise retinotopic input from just one eye. The properties of LGN cells are similar to those of the ganglion cells which supply them and have circular receptive fields with surround inhibition. The LGN projects to the primary visual cortex and receives extensive connections from it which may be involved in visual attention.

Primary visual cortex

The striate cortex in the occipital lobe is the primary visual cortex (V1) and gets a retinotopic projection from the LGN, with the fovea occupying a disproportionately big area. The M and P LGN cells project to different sublayers in layer 4C of the cortex, giving rise to distinct streams of information flow through the cortex. Most cells in V1 have elongated receptive fields and respond to linear features rather than spots. Simple cells are position sensitive. Complex cells are less sensitive to position than simple cells, many having a preference for linear stimuli moving at right angles to the long axis of their receptive fields.

Orientation columns

All cells lying within radial columns extending through V1 respond to linear features having approximately the same orientation. All orientations are represented for each point on the retina. These orientation columns are ordered so that a smooth gradient for orientation exists; columns with the same orientation are aligned in stripes across the cortex.

Binocular cells

Many cells in V1 get input from both eyes, but most show ocular dominance in that they are preferentially driven by one eye. These cells occur in discrete ocular dominance columns that are aligned in stripes across the cortex in which ipsilateral and contralateral eye dominance alternates. Binocular cells get input from corresponding positions on the two retinas and measure retinal disparity, from which the visual system computes the depth of an object in three-dimensional space.

Hypercolumns

The volume of cortex in which every orientation is mapped for corresponding positions on both retinas is called a hypercolumn. It consists of a complete set of orientation columns and ocular dominance columns for a single pixel of the visual field.

Related topics

Organization of the central nervous system (A5)
Attributes of vision (G1)
Eye and visual pathways (G2)
Parallel processing in the visual system (G7)

Lateral geniculate nucleus

Pathways for visual perception start with the retinogeniculate fibers, axons of ganglion cells that end in the lateral geniculate nucleus (LGN). Because of the manner in which fibers are sorted in the optic chiasm, the left optic tract and left LGN carry axons from the left side of both retinas. Thus the left LGN represents the right side of the visual field (see *Fig. 2*, Topic G2).

The primate LGN has six layers (*Fig. 1*). The two most ventral are the **magnocellular** (large-cell) **layers** which receive input from M ganglion cells. Dorsal to these are the four **parvocellular** (small-cell) **layers** that are innervated by P ganglion cells. Interleaved between these major layers are **koniocellular layers** containing very small cells. These are thought to receive input from small slowly conducting retinal ganglion cells with large dendritic trees and large receptive fields which signal average illumination.

LGN cells have circular RFs with surround antagonism. They show little or no response to diffuse light covering the whole receptive field. Each layer in the LGN gets input from only one eye and no cells show binocular responses (responses to both eyes). It is not until the visual cortex is reached that input from both eyes is integrated. The responses of the LGN cells match those of the ganglion cells which supply them, so on and off channels remain independent and P cells display precisely the same color opponency properties as retinal ganglion cells. There is a very precise topographic (retinotopic) mapping from the retina onto each of the layers of the LGN, with the representation of the fovea taking up about half of the nucleus. The maps in each layer are in precise register with each other so that any given vertical axis through the LGN passes through cells with RFs representing the same place in visual space.

The LGN contains two populations of neurons. Those that project to the primary visual cortex are **geniculostriate** neurons. In addition, there is a substantial population of smaller interneurons the exact function of which is not known. Furthermore only about 20% of the synaptic connections on geniculostriate neurons are from retinal ganglion cells. Other synapses are made by back projections from the visual cortex and by the reticular formation. They may play a role in visual attention, modifying geniculostriate neuron responses so that only a selected proportion of retinal input is transmitted through to the visual cortex.

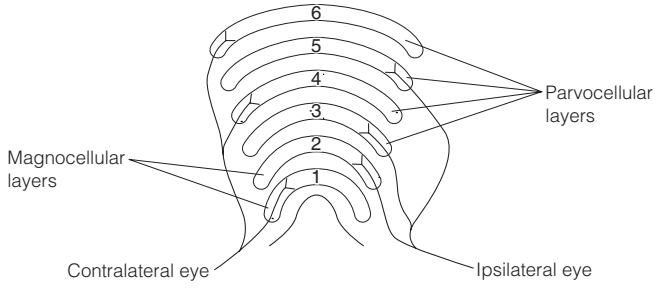


Fig. 1. The structure and inputs of the lateral geniculate nucleus.

Primary visual cortex

The fibers of the optic tract terminate in the **striate** cortex (Brodmann's area 17) on the medial surface of the tip of the occipital lobe. This region is the **primary visual cortex** (**V1**). Precise retinotopic mapping is maintained up to V1 with the fovea having a disproportionate representation.

There are at least three parallel streams of information into the primary visual cortex. The movement-sensitive M LGN cells input into layer 4Cα, the P LGN cells go to layer 4Cβ, whereas the koniocellular layers of the LGN project to layers 2 and 3. These streams remain quasi-independent throughout the visual system. The connections of the primary visual cortex are illustrated for the primate in *Fig. 2*.

The great majority of cells in V1 have elongated receptive fields (RFs) with both inhibitory and excitatory regions, and respond to bars, slits, edges and corners rather than spots of light. Most fall into two categories based on their RF properties, simple or complex cells. Both are orientation selective, in that they respond to linear features in only a narrow range of orientations.

1. **Simple cells** are pyramidal cells found mostly in layers 4 and 6. They are highly sensitive to the position of a stimulus on the retina. They have small oval RFs with center–surround antagonism (*Fig. 3*). A simple cell gets its input from a linear array of LGN cells having the same RF properties, so the RF of the simple cell emerges as a consequence of the RFs of the LGN inputs.
2 **Complex cells** are most abundant in layers 2, 3 and 5. They have larger RFs than simple cells and, lacking distinct inhibitory or excitatory regions, a stimulus of the appropriate orientation anywhere in the RF evokes a response. Hence, complex cells are much less fussy about position than simple cells. Many complex cells show a preference for movement at right angles to the long axis of the stimulus. Some complex cells receive their inputs from simple cells but others get their input directly from the LGN.

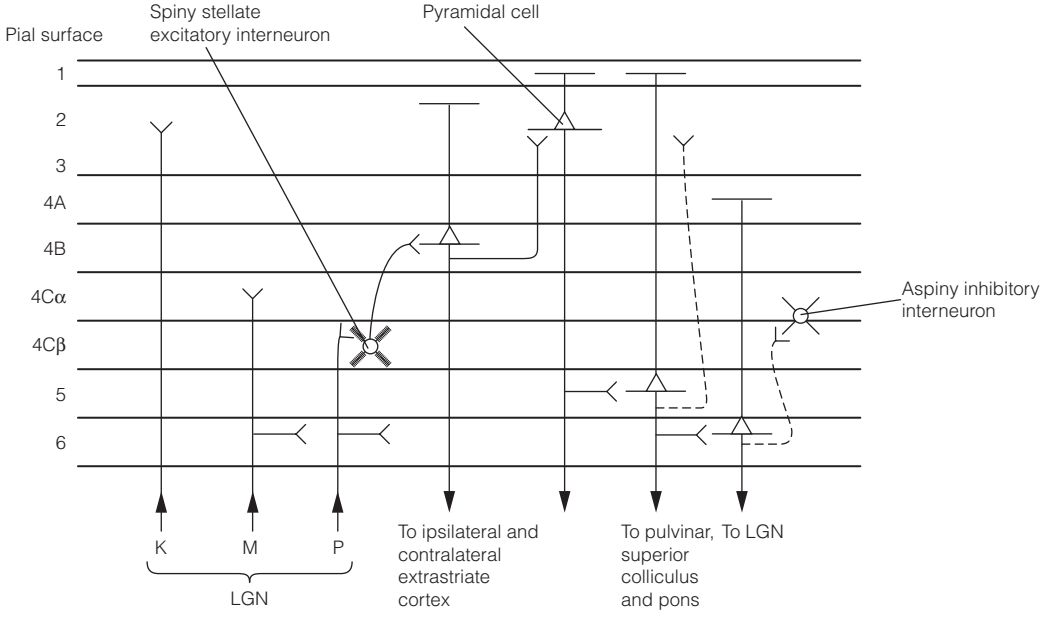


Fig. 2. Canonical circuitry in the primary visual cortex illustrated for parvocellular (P) LGN input. Magnocellular (M) circuitry (not illustrated) has its input to 4Cα. Spiny stellate cells here send axons to pyramidal cells in 4B and these send collaterals to pyramidal cells in layers 5 and 6 directly rather than via more superficial layers. Koniocellular input is directly to pyramidal cells in blobs of 2 + 3. Feedback collaterals are dotted. Cortical layers are designated by Arabic numbers.

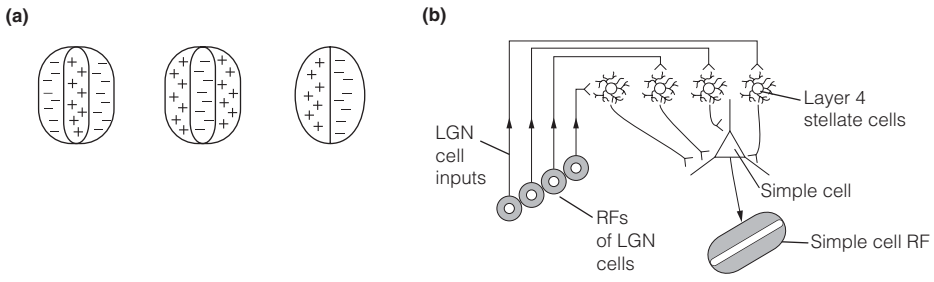


Fig. 3. (a) Receptive fields of three simple cells; (b) a diagram depicting how lateral geniculate nucleus (LGN) cells contribute to the simple cell receptive field (RF), four on-center LGN cell RFs generate an on-centre simple cell RF.

Orientation columns

In common with other sensory cortex, the primary visual cortex is divided into radial columns 30–100 μm across. In each of these all cells respond preferentially to linear features with a given orientation so they are called **orientation columns**. The cortex is organized so that adjacent columns have an orientation preference that differs by only about 15°; in other words, orientation is represented in a systematic way across the cortex. Columns which have the same orientation are arranged in stripes across the cortex. The obvious inference, that orientation selectivity is how the visual system represents straight-line segments which can be built up to give the form of an image, need not be true. Computer modeling shows that orientation selectivity is a property of neural networks that learn the curvature of curved surfaces from their shading. Hence orientation selectivity might, counter-intuitively, be concerned with representations of curves rather than linear features in the visual world.

Binocular cells

V1 is the first region in which input from both eyes is combined. Many cells, particularly in layers 4B, 2 and 3 show binocular responses in that they can be driven by either eye. This is a necessary condition for stereopsis. Most **binocular cells** show a preference for one eye, a phenomenon referred to as **ocular dominance**. Cells which have the same ocular dominance (e.g. those that are driven preferentially by the ipsilateral eye) occupy **ocular dominance columns** that are situated in long stripes about 500 μm across. Columns representing ipsilateral and contralateral input alternate regularly over the cortex which when visualized at the level of layer 4C look like the pattern of stripes on a zebra.

The receptive fields of binocularly driven cells resemble those of simple or complex cells, lie in corresponding positions in the two retinas, have identical orientation properties and have similar arrangements of excitatory and inhibitory regions. Similar input from both eyes into arrays of binocular cells is needed for perception of a fused image. To the extent that inputs into these cells are unequal they measure retinal disparity and so the depth of an object in three-dimensional visual space. Cells that respond to visual disparity have been discovered in the primate visual cortex (including V1). These are responsible for stereopsis.

Hypercolumns

A higher-order modularity exists in the primary visual cortex. Called a **hypercolumn** (*Fig. 4*), it represents a given corresponding position for both retinas, and maps every orientation for that position. It consists of a full-thickness slab

of cortex with an area of about 1 mm^2 containing a complete set of orientation columns for both ipsilateral and contralateral ocular dominance. The retinotopic map in V1 occurs because adjacent pixels of the retina map to adjacent columns in an orderly fashion.

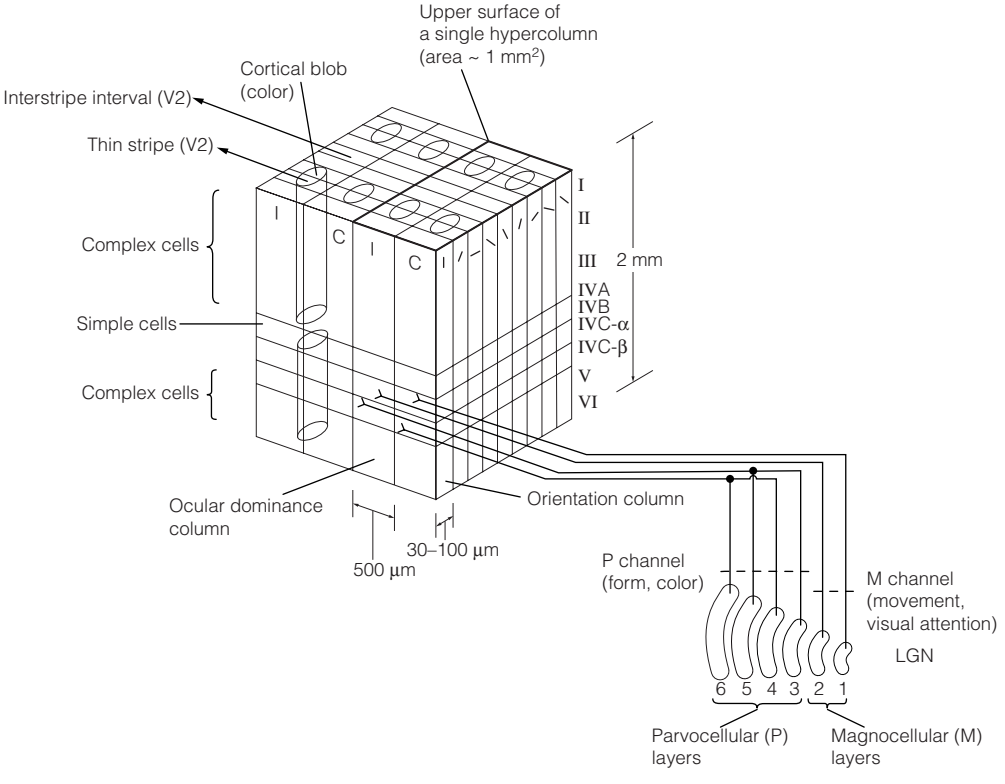


Fig. 4. Modular structure of the primary visual cortex. Cortical layers are designated by Roman numerals. I, ipsilateral; C, contralateral (blobs are described in Topic G7).

G7 Parallel processing in the visual system

Key Notes

Parallel processing in V1

There are three relatively independent pathways for processing information in the visual system. The magnocellular (M) pathway gets its input from M LGN cells which synapse with neurons in 4Cα. The M pathway is color blind and involved in analysis of moving stimuli, control of gaze and stereopsis. The parvocellular (P) systems arise from P LGN cells which synapse with cells in 4Cβ. There are two P systems. The parvocellular–interblob pathway is concerned with form, its cells are orientation selective and binocular. The parvocellular–blob pathway mediates color vision; wavelength-selective cells in this pathway have double opponent receptive fields in which the center is excited by some cones and inhibited by others, whilst the converse situation occurs in the surround. Three features of color perception – color constancy, perceptual cancellation and simultaneous color contrast – can be explained by the manner in which double opponency is organized.

Extrastriate visual cortex

All cortical areas involved in vision other than V1 are together referred to as extrastriate visual cortex. It includes much of the occipital cortex and parts of the parietal and temporal cortex. The secondary visual cortex (V2) receives input from V1 which then projects to other extrastriate cortex. The three streams of visual information remain segregated in V2, as revealed by stripes in cytochrome oxidase staining, and throughout the extrastriate visual cortex. The M pathway goes via thick stripes in V2 to V3 and then V5. Destruction of human V5 causes a loss of ability to see objects in motion. The parvocellular–interblob pathway goes via V2 interstripes to V3 and V4 whilst the parvocellular–blob pathway goes from V2 thin stripes to V4, cells of which show color constancy. Destruction of human V4 causes loss of color vision.

Where and what streams

Beyond V5 and V4 information is divided into two streams. From V5 a dorsal stream to the medial superior temporal cortex and the posterior parietal cortex is concerned with object location. The ventral stream from V4 to the inferotemporal cortex is concerned with recognition of objects. The two streams are called 'where' and 'what' streams respectively.

Related topics

Attributes of vision (G1)
Early visual processing (G6)
Arousal and attention (O4)

Parallel processing in V1

Three relatively independent pathways, each of which processes different aspects of vision in parallel, in a quasi-autonomous fashion, can be delineated in the primary visual cortex.

The **magnocellular pathway** from M ganglion cells to M LGN cells has its input to spiny stellate cells in layer 4Cα (see *Fig. 2, Topic G6*). These excitatory

interneurons synapse with pyramidal cells in layer 4B which show orientation and direction selectivity. These cells send axon collaterals to pyramidal cells in layers 5 and 6. Layer 5 cells project to subcortical regions, the pulvinar (a thalamic nucleus involved in visual attention), the superior colliculus and pons. Layer 6 pyramidal cells go to the extrastriate cortex. The M pathway is specialized for analysis of motion. Its outputs via layer 5 are important in visual attention and gaze reflexes. Some cells in the M pathway are binocular so it contributes to stereopsis. Because it originates with ganglion cells which combine input from two classes of cone cell it is not wavelength selective, the M system is color blind.

There are two **parvocellular pathways**. They arise from P ganglion cells via P LGN cells which synapse with spiny stellate cells in 4Cβ. Like the M pathway the interneurons connect with pyramidal cells in 4B. However, in the parvocellular paths, 4B cells (which are orientation-selective simple cells) synapse with pyramidal cells in layers 2 and 3 which then relay with deep pyramidal cells in layer 5. Segregation of the two parvocellular pathways occurs in layers 2 and 3. When stained for the mitochondrial enzyme, **cytochrome oxidase**, layers 2 and 3 show pillars of high activity, **blobs**. Each blob is centered on an ocular dominance column. Between the blobs lies the **interblob** region. Cells in the interblob region are orientation-selective, binocularly driven, complex cells. They are not wavelength selective or motion sensitive. They are part of the **parvocellular–interblob** (PI) pathway which processes high-resolution analysis of form in the visual world. By contrast cells in the blobs are wavelength selective, show poor orientation selectivity and are monocular. The **parvocellular–blob** (PB) pathway mediates color vision. Blob pyramidal cells get direct input from the koniocellular LGN layers but the function of this input is not yet understood.

Wavelength selective blob cells are **double opponent cells** with receptive field (RF) properties derived from their inputs, the single opponent parvocellular LGN cells. Double opponent cells have center–surround antagonist RF configuration, they signal color contrast and come in four classes categorized by their preferred stimuli. The top left cell in *Fig. 1* is excited by L cones in the center and inhibited by L cones in the surround. In addition, it is inhibited by M cones centrally but excited by M cones in the surround. The preferred stimulus for this cell is a red spot in a green background. However the cell gives **off responses** if exposed to a green spot in a red background (*Fig. 2*).

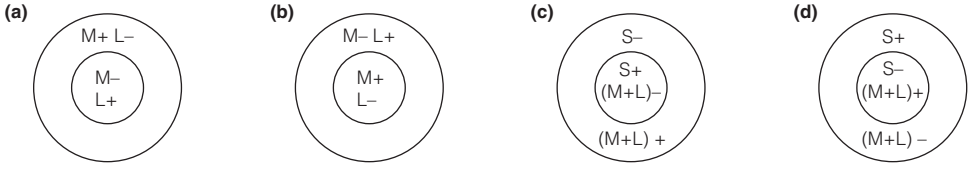


Fig. 1. Double opponent cells in V1 blobs. Preferred stimuli: (a) red spot, green surround; (b) green spot, red surround; (c) blue spot, yellow surround; (d) yellow spot, blue surround.

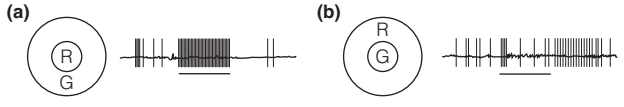


Fig. 2. Responses of the double opponent cell in Fig. 1(a): (a) preferred stimulus; (b) off response which might account for successive color contrast (see text for details).

Unlike single opponent cells which are excited by small spots of white light, double opponent cells are unaffected by white light stimuli of any size, so they are more selective detectors of color contrast. The organization of double opponent cell RFs explains some of the properties of color vision described in Topic G1.

1. **Color constancy**. The way in which the brain computes color constancy is not understood in detail but is partly accounted for by the behavior of double opponent cells. A shift in the wavelength composition of light will produce equal but opposite effects on the responses of the center and surround of double opponent cells. There will be little effect overall on the RF of the cell which will continue to signal the same color. On the scale of the entire visual field, color constancy is thought to involve comparing red–green brightness, blue–yellow brightness (from color single opponent cells) and total brightness (added outputs of S, M and L cones) over large areas of retina.
2. **Perceptual cancellation** is explained by the way in which color opponency happens to be organized as red (R) versus green (G) and yellow (R+G) versus blue channels. Since mutual antagonism occurs between red and green or between yellow and blue only one color in each pair can be seen at a single pixel of the retina at any time.
3. **Simultaneous color contrast** can also be accounted for by the properties of double opponent cells. For example, the cell in *Figure 2* cannot discriminate between a green stimulus to its surround or a red stimulus to the center; the response is the same for both. So a gray disk viewed in a green background is interpreted as red. A similar mechanism explains the complementary afterimages that appear after staring at a uniform patch of color (*Fig. 2b*).

Extrastriate visual cortex

The segregation of visual information for motion, form and color in V1 is maintained in the **extrastriate visual cortex**, which is a term applied to all of the visual cortex except V1. The extrastriate cortex of primates contains about 30 regions that can be differentiated on the grounds of cytoarchitecture, connections and physiological properties. Most have a retinotopic map of some aspect of the visual world. It includes not only occipital cortex areas 18 and 19 but also areas of parietal and temporal cortex. In humans it is estimated that almost one half of the cerebral cortex is implicated in vision, more than is devoted to any other single function. This implies that vision is the most complex task that the brain performs. The terminology usually adopted for extrastriate cortex is based on studies of the macaque monkey. It is thought that most regions in this primate have counterparts in humans. The location and connections of the major visual cortical areas are depicted in *Figure 3a* and *3b* respectively.

Most outputs from V1 go to **V2**, the **secondary visual cortex**, which occupies part of area 18. V2 shows a characteristic cytochrome oxidase staining pattern, alternating thick and thin stripes running at right angles to the V1/V2 border. Pathway-tracing techniques and electrophysiological studies reveal how the magnocellular and parvocellular pathways continue into V2 and beyond.

Cells in the V2 thick stripes are motion sensitive and binocular, being driven by a preferential retinal disparity. The thick stripe receives inputs from layer 4 of the interblob region of V1 and sends much of its output, via **V3,** to the **medial temporal (MT)** visual cortex, **V5**. Lesions of human V5 result in loss of the ability to perceive motion (**akinetopsia**). Hence, the V2 thick stripe–V3–V5 (MT) connection is the extension of the magnocellular pathway (*Fig. 3b*) and concerned with motion and depth perception.

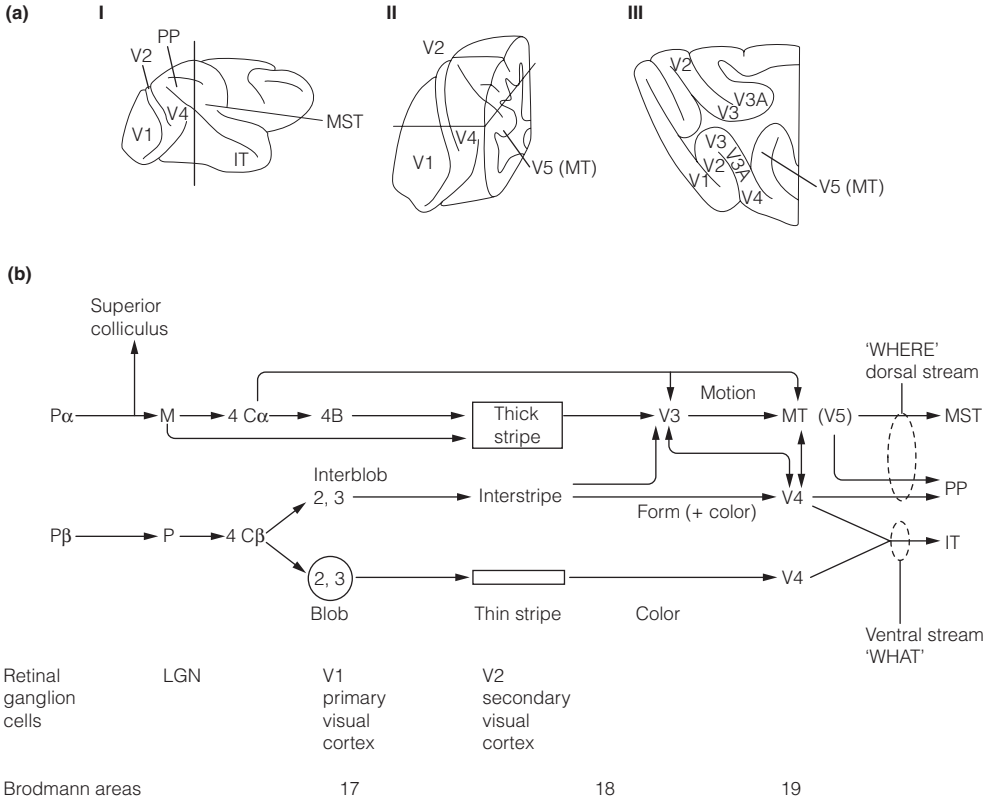


Fig. 3. Parallel processing in the visual system. (a) Anatomy of the visual areas in the macaque: (i) left cerebral hemisphere; (ii) coronal section through the posterior third of the hemisphere; (iii) horizontal section. Modified from Kandel, Schwartz, Jessell (eds) (1991) Principles of Neural Science, *3rd edn. (b) Flow diagram of M and P channels in the primate visual system. IT, Inferotemporal cortex; MST, medial superior temporal cortex; MT, medial temporal cortex; PP, posterior parietal cortex.*

The interstripe region of V2 gets its inputs from the V1 interblob regions (layers 2 and 3) and sends outputs to V3 and then to **V4**. Many cells in V3 and a proportion of those in V4 are orientation selective and these represent a continuation of the PI parvocellular pathway. It is primarily concerned with form perception.

The blobs of V1 project to the thin stripe of V2 which in turn sends outputs to visual area V4. Thin stripe V2 cells, and some V4, cells are both wavelength selective and show color constancy, so the blob–V2 thin stripe–V4 route is the extension of the PB parvocellular pathway for color vision. This is supported by the loss of color vision (**achromatopsia**) that occurs in patients with damage to human V4.

Although the M, PI and PB pathways operate in parallel they are not completely independent. Reciprocal pathways exist between V3–V4 and V5–V4, which presumably allow interactions between M and PI systems, both of which contribute to stereopsis. Interaction of motion and form analysis is probably required for the identification of moving objects. There seems to be no cross talk however between M and PB pathways. The M system is color blind and for

equiluminant stimuli (those varying in color but not in brightness), which can only be perceived by the PB system, the perception of motion vanishes. The PI system must receive input from wavelength-selective cells in V4 since it is able to use color contrast to localize borders as part of its role in analyzing form. However, form information seems not to be available to the PB pathway. When viewing equiluminant blocks of color they appear to 'jump around' because the PB system is unable to localize boundaries.

Saccades are rapid stereotyped movements of the eyes which serve to bring an object at the periphery of the visual field to the fovea. During saccades black and white gratings detectable only to the M system vanish while equiluminant stimuli detectable by the parvocellular pathway remain visible; thus the M, but not the P, system is shut down during saccades. This means that the M (motion) system is not confused by the rapid eye movement. The response times of the cells in the P system are sufficiently slow that they are unaffected by the shifting image.

Where and what streams

Parallel processing beyond V5 and V4 results from the segregation of information into two streams. The dorsal stream, largely from MT, goes to the **medial superior temporal** (**MST**) and **posterior parietal** (**PP**) cortex. Cells in the PP cortex have large RFs, show selectivity for size and orientation of objects and fire as a monkey makes hand movements to grasp an object. Many cells show gaze-dependent responses, i.e. their firing depends on where an animal is looking. Lesions to MST and PP in primates results in **optic ataxia**, in which visuospatial tasks are profoundly affected, but are without any effect on the ability of animals to recognize objects.

By contrast, the ventral stream from V4 to **the inferotemporal cortex** (**IT**) is crucial for object recognition. Cells of the IT cortex have extremely large RFs, usually bilateral, are sensitive to form and color but are relatively unfussy about object size, retinal position or orientation. Many of these cells respond selectively to specific objects such as hands or faces. Unusually for visual cortex, the IT area has no retinotopic map. Lesions of the IT cortex cause **visual agnosia** in which animals fail to perform or learn tasks that require the recognition of objects. Visuospatial tasks are unaffected.

The very different functions of the dorsal and ventral streams are epitomized by their being referred to as **where** and **what** streams respectively. Clinical data suggest that a similar dichotomy exists in humans. Optic ataxia occurs with posterior parietal damage. These patients have no difficulty in recognizing objects but cannot seem to reach for, or grasp them. By contrast, patients with damage to the occipito-temporal cortex fail to recognize common objects, including once familiar faces, a disorder termed **prosopagnosia**. These people experience no difficulty in understanding where objects are located in space or how to reach for, or avoid them.

Bilateral loss of V1 causes total loss of visual perception. However, there are examples of primates and humans with this damage who are able to avoid obstacles whilst moving through space much better than chance. This phenomenon is called **blindsight**. Humans possessing it report that they are completely unaware of the visual world and do not understand how they are able to navigate through space. Blindsight is mediated by a pathway that goes directly from the magnocellular LGN to the thick stripe of V2. This provides input to the where system. The implication of this condition is that V1 is required for conscious visual perception.

G8 Oculomotor control

Key Notes

Eye movements

Eye movements either keep the gaze fixed on an object when the head is turning, or shift the gaze to follow a moving object. Gaze fixation is brought about by the vestibulo-ocular reflex, which relies on signals from the semicircular ducts, and the optokinetic reflex which depends on visual input. Gaze shift can be produced either by saccades (fast), smooth pursuit (slow) or vergence movements. Vergence allows an object to be tracked as it approaches or recedes and requires the eyes to move in opposite directions.

Extraocular eye muscle control

The actions of three pairs of muscles allow the eyes to be rotated about three principal axes. During conjugate eye movements in which both eyes move in the same direction, eye muscle activity in one eye is complementary to eye muscle activity in the other. The extraocular muscles are controlled by motor neurons in the nuclei of the oculomotor, trochlear and abducens nerves, which in turn are driven by brainstem reticular and medial vestibular nuclei. Firing of eye motor neurons encodes the velocity of the movement and the change in eye position.

Vestibulo-ocular reflexes (VORs)

Rotation of the head, detected by the semicircular ducts, causes a well-matched opposite rotation of the eyes to keep the retinal image stationary. For large horizontal head rotations, once the eyes have rotated as far as possible they are rapidly reset to frontwards gaze. This causes nystagmus, flickering eye movements with slow phases in which the gaze is fixed and fast phases when gaze is reset. Vestibulo-ocular reflexes adapt in response to alterations in visual input. This is an example of motor learning by the cerebellum.

Optokinetic reflexes (OKRs)

Slow head rotation causes images to move across the retina. This triggers eye movement in the opposite direction. Nystagmus occurs for large rotations.

Saccades

Fast movements taking the fovea to a new point in visual space are saccades, produced reflexly by visual, auditory or somatosensory stimuli. Burst cells in brainstem reticular nuclei are directly responsible for the signals to the eye motor neurons, but saccades are produced in response to activity in the superior colliculus and frontal cortex. The superior colliculus generates reflex saccades. Because it has sensory and motor maps, each point in the superior colliculus represents a location in sensory space and specifies the saccades necessary to point the gaze towards it. The size and direction of saccades is actually determined by the average firing of many collicular neurons. The frontal eye fields, located in the frontal cortex, trigger saccades via connections with the superior colliculus and brainstem, and basal ganglia. The frontal cortex is responsible for intentional saccades.

Smooth pursuit movements

These are used to voluntarily track an object that is moving in the visual field. The velocity of the object is signaled by the cortex of the visual 'where' system to neurons in the pons. These cells convert the velocity signals to smooth pursuit motor commands.

Vergence

Signals for vergence include blurring of the retinal image or the degree of accommodation and require the visual cortex. Fast vergence movements are made during saccades.

Related topics

Balance (F5)
Eye and visual pathways (G2)
Cortical control of voluntary movement (K1)
Cerebellar function (K6)
Basal ganglia function (K8)

Eye movements

The purpose of eye movements is either **gaze stabilization**, in which the eyes remain fixated on an object during rotation of the head, or **gaze shifting** which allows the central part of the retina, the fovea, to be brought to bear on an object, or track a moving object. Five types of eye movements, each controlled by a distinct neural system, bring about these aims.

Gaze stabilization is controlled by the vestibulo-ocular and optokinetic systems. Rapid head rotation, detected by the semi-circular ducts provides input for vestibulo-ocular reflexes (VOR), whereas optokinetic reflexes depend on visual input to monitor slow head rotations. For both systems their output causes conjugate eye movements in the opposite direction to the head rotation, so that retinal images do not shift.

Three systems organize gaze shift. The saccadic system generates extremely rapid eye movements, saccades, which move the gaze from one point in the visual field to another, bringing new targets onto the fovea. The smooth pursuit system permits gaze to follow a moving target, so that its image remains on the fovea. Finally, for animals with binocular vision, the vergence system allows the eyes to move in opposite directions (disjunctive movements); either both converge or both diverge, so that both eyes can remain directed towards an object as it gets closer or recedes.

The output of all five eye motor systems is via oculomotor neurons in the brainstem, the axons of which run in three pairs of cranial nerves to the skeletal muscles that move the eyes.

Extraocular eye muscle control

Each eye is moved by three pairs of **extraocular eye muscles**. Two pairs of rectus muscles (superior and inferior, medial and lateral) originate from a common annular tendon attached at the back of the orbit. These muscles insert into the sclera in front of the equator of the eyeball. The third pair is the oblique muscles (superior and inferior) which insert into the sclera behind the equator of the eyeball (*Fig. 1*).

Working in concert these muscles act to rotate the eye about three principal axes (*Fig. 2*). The actions of the medial and lateral rectus muscles are simple. They cause the eye to rotate about the vertical axis so that the gaze moves horizontally. The medial rectus brings about rotation towards the midline (adduction) while the lateral rectus causes lateral rotation (abduction). The other two pairs of muscles produce rotations that have components along two of the

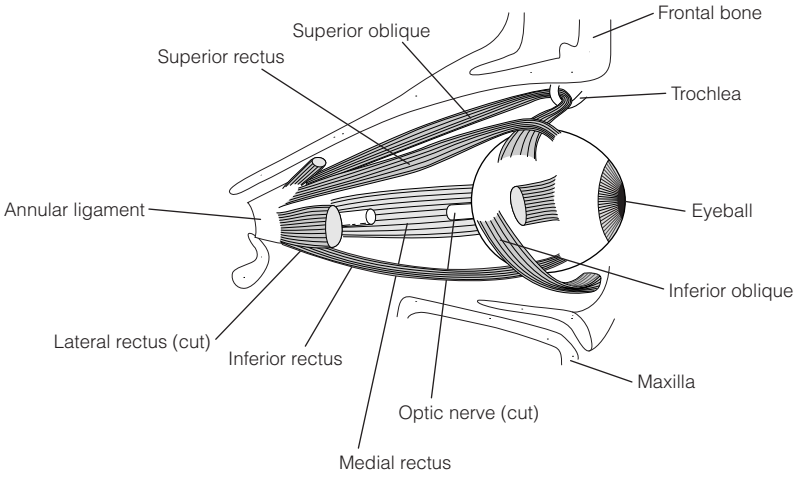


Fig. 1. The right orbit showing the extraocular muscles.

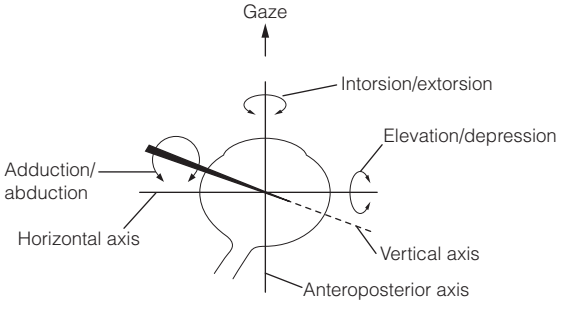


Fig. 2. Principal axis for rotation of the eye, shown for the right eye. In health, torsional movements (rotation about the anteroposterior axis) are small.

principal axes, and the components change depending on the horizontal position of the eye. These actions are summarized in *Table 1*.

Eye muscles act in complementary fashion in the two eyes during conjugate movements in which the two visual axes move in parallel. Thus, contraction of the lateral rectus in one eye is coupled with contraction of the medial rectus in the other eye for a conjugate horizontal shift in gaze (see *Table 1*).

The extraocular muscles are innervated by motor neurons in the nuclei of the oculomotor (III), trochlear (IV) or abducens (VI) cranial nerves. These neurons are the final common path for the output of all five eye-movement systems and are driven by brainstem reticular and medial vestibular nuclei axons that run in the **medial longitudinal fasciculus**. Eye motor neurons fire both statically, in a manner relating to eye position, and dynamically, reflecting eye velocity. To hold the eye steady in a given position requires tonic discharge by a particular set of motor neurons. The set will be different for different positions. Each motor neuron fires with a frequency needed to maintain eye position, so its firing rate is linearly related to position. Activity by a given neuron is not needed for all positions (e.g. sustained leftward gaze requires high firing rates of motor units in the left lateral rectus, but the left medial rectus is an antagonist of this movement so its motor neurons remain silent).

Table 1. Actions of extraocular eye muscles. For the muscles moving the eye vertically the action is different depending on whether the eye is also abducted or adducted. For example, the superior rectus elevates the eye if the lateral rectus is active at the same time, but causes intorsion if the eye is adducted by the medial rectus.

Muscle	Innervation	Movement	Contralateral eye complementary muscle
Lateral rectus	Abducens (VI)	Abduction	Medial rectus
Medial rectus	Oculomotor (III)	Adduction	Lateral rectus
Superior rectus	Oculomotor (III)	Elevation and intorsion	Inferior oblique
Inferior rectus	Oculomotor (III)	Depression and extorsion	Superior oblique
Inferior oblique	Oculomotor (III)	Extorsion and elevation	Superior rectus
Superior oblique	Trochlear (IV)	Intorsion and depression	Inferior rectus

Eye movements are brought about by high-frequency pulses of action potentials in oculomotor neurons. The discharge rate during a pulse is directly proportional to the velocity of the movement. The eye movement will bring the eye to a new position, which needs the generation of a new position signal. This is thought to be achieved by integrating the velocity signal, an operation probably done by the vestibulocerebellum and **prepositus nucleus** of the brainstem reticular system.

Vestibulo-ocular reflexes (VORs)

Head rotation detected by the semicircular ducts triggers equal and opposite rotation of both eyes. For large head rotations the eyes cannot continue to rotate but must be reset to a central position by rapidly moving in the *same* direction as the head. This gives rise to **nystagmus,** eye movements characterized by slow phases that stabilize the retinal image, and quick phases that reset the eyes. By convention the direction of the nystagmus is the direction of the quick phase (*Fig. 3*).

The horizontal semicircular ducts are effectively wired to the medial and lateral rectus muscles to produce the eye movements that counter the head rotation (*Fig. 4*).

The gain of the VOR (the magnitude of the eye rotation divided by the magnitude of the head rotation) is quite close to one for fast head rotations. This means there is a good match between eye and head movements, which makes for a stable retinal image. The VOR can be modified by visual experience. When human subjects wear magnifying lenses, which means that they should make bigger eye movements to match head rotations, the gain of their vestibulo-ocular reflexes increases appropriately over the next few days. The cerebellum is required for this adaptation to occur, but not for it to be maintained once established. The instability of the retinal image acts as an error signal which is

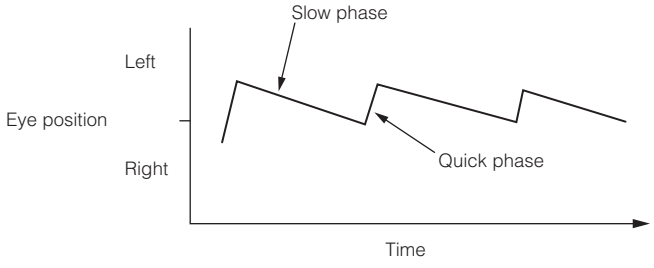


Fig. 3. Leftward nystagmus during head rotation.

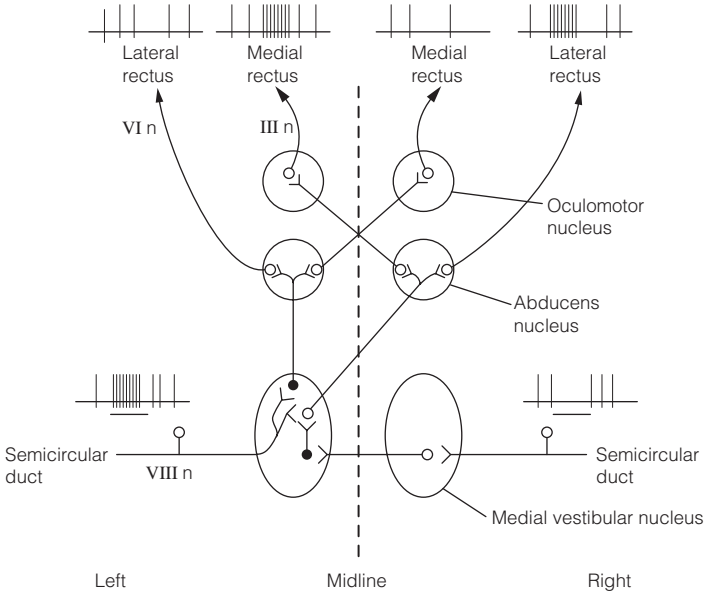


Fig. 4. Circuitry for the vestibulo-ocular reflex. Stimulation of the horizontal semi-circular ducts by the leftward rotation of the head excites the ipsilateral medial rectus and the contralateral lateral rectus and inhibits their antagonists. Excitatory neurons, open circles; inhibitory neurons, filled circles. Firing patterns are shown for cranial nerve neurons.

relayed to the cerebellum from the inferior olivary nucleus. The cerebellum learns to minimize the error and alters its drive to the extraocular muscles. This is an example of motor learning. Lesions of the vestibulocerebellum impair the ability to maintain steady gaze, causing inappropriate nystagmus.

Optokinetic reflexes (OKRs)

Slow rotation of the head causes an apparent movement of the visual world in the opposite direction called **retinal slip**, This is detected by large, movement-sensitive retinal ganglion cells and is used to produce eye movements which are equal in speed but of opposite direction to the retinal slip. As in the VOR, nystagmus occurs for large head rotations.

Saccades

Saccades are very fast conjugate eye movements that move the fovea to target a different point in visual space. The saccade system uses visual, auditory and somatosensory input to determine the eye rotation required to realign the gaze. Horizontal saccades are controlled by the **paramedian pontine reticular formation** (**PPRF**) which lies at the midline adjacent to the nuclei of the oculomotor, trochlear and abducens cranial nerves. Vertical saccades are organized by the **rostral interstitial nucleus** of the medial longitudinal fasciculus, situated in the midbrain rostral to the oculomotor (III) nerve nucleus.

Both of these structures contain burst cells which code for the size and direction of the eye movement and produce saccades by exciting the oculomotor neurons. The signals that trigger saccades come from two sources, the superior colliculus and the frontal eye fields in the cortex of the frontal lobes. Both these structures can generate saccades independently of the other. Destruction of both renders primates incapable of making saccades.

The **superior colliculus** lies in the tectum of the midbrain and is divided into superficial, intermediate and deep layers. The superficial layers receive visual information from the retina and visual cortex that gives rise to a map of the contralateral visual field. The deep layers get auditory and somatosensory input and so have two maps; an auditory map depicting the location of sounds in space, and a somatotopic map in which the body parts closest to the eyes get the greatest representation. The intermediate layers are the site of a motor map. Neurons here are called **collicular saccade-related burst neurons** because they fire a high-frequency burst of action potentials about 20 ms before a saccade. Each one has a **movement field** (the equivalent of a receptive field) that covers the sizes and directions of the saccades it participates in. These movement fields are large, in that the cells are active for many similar saccades, but they fire maximally for a preferred saccade. The broad tuning of these cells means that the direction of any given saccade is encoded by a population of neurons whose firing precisely determines the direction of the required saccade. This is exactly the way in which the primary motor cortex uses population coding to determine the direction of a movement.

A key function of the superior colliculus is to turn sensory coordinates into motor coordinates. All of its four maps are in register so each point on the superior colliculus represents a specific location in sensory space and the saccades necessary to direct gaze towards it. Visual input to the superficial layers need not lead to firing of collicular saccade-related burst cells in the intermediate layer. This is because superficial layer cells do not synapse directly with intermediate layer cells but instead connect with them indirectly via a relay through

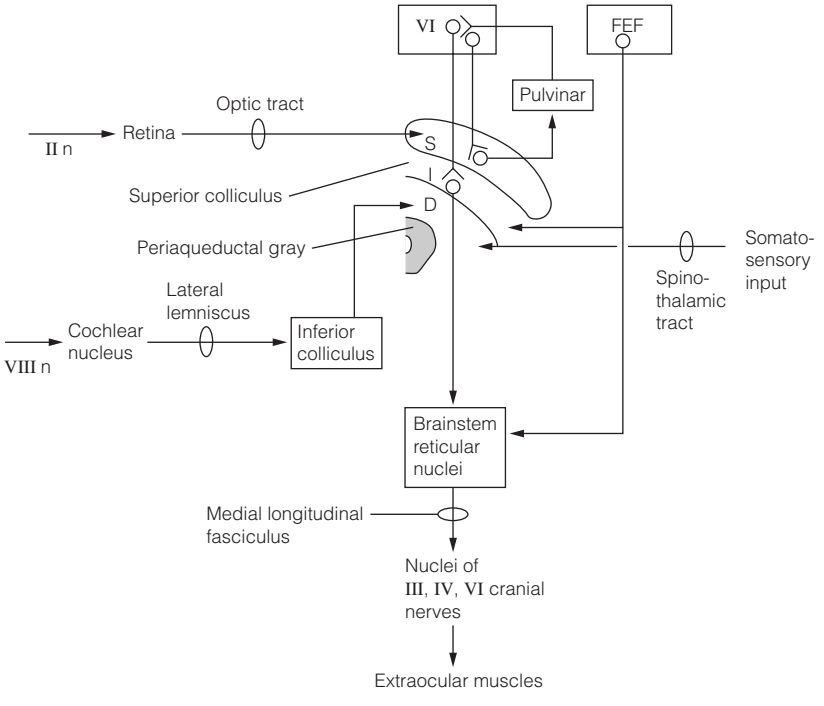


Fig. 5. Circuitry for saccades. The superior colliculus has three layers: S, superficial; I, intermediate; D, deep. FEF, frontal eye fields; VI, primary visual cortex.

the **pulvinar** of the thalamus and the visual cortex. The purpose of this relay may be to decide the importance (**salience**) of a particular visual stimulus and so allow saccades only to stimuli with high salience. *Fig. 5* is a diagram of the circuitry involved in saccades.

In addition to generating saccades the superior colliculus causes head rotation, by way of the **tectospinal tract**, to neck muscle motor neurons. This allows orientation towards a stimulus, so called **orienting responses**.

The **frontal eye field** (**FEF**) of the frontal cortex triggers saccades via the intermediate layers of the superior colliculus, and the pontine and midbrain reticular nuclei. The FEF directly stimulates the collicular saccade-related burst neurons in the intermediate layers of the superior colliculus. Additional control of saccades by the FEF is achieved by an oculomotor circuit in the basal ganglia.

Lesions of the superior colliculus cause temporary impairment in triggering saccades, but recovery occurs because the FEF can trigger saccades by its direct connections with the pons and midbrain. Damage to the FEF causes transient paralysis of gaze towards the opposite side, but reflex saccades soon return, produced by the superior colliculus. However, loss of the FEF prevents intentional or anticipatory saccades.

Smooth pursuit movements

Intentionally tracking a moving object so that its image remains on the fovea is done by the smooth pursuit system. Smooth pursuit movements differ from optokinetic reflexes in being voluntary, and in attending to movement over a small part of visual space. By contrast, optokinetic reflexes are involuntary and are responses to movement of the entire visual world.

Signals relating to the velocity (i.e. speed *and* direction) of the target are generated by the medial temporal cortex of the visual 'where' system which analyzes motion. Lesions of this cortex impair pursuit movements. These signals are transmitted to the **dorsolateral pontine nucleus** (**DLPN**), which translates the target velocity into the motor commands for the pursuit movement. The DLPN projects to the vestibulocerebellum, cells of which fire in a precisely correlated way with smooth pursuit movements. Its output to the medial vestibular nuclei drives the smooth pursuit movements.

Vergence

Vergence is the only disjunctive eye movement. For example, shifting gaze to a closer target requires adduction of both eyes, which is achieved by contracting both medial rectus muscles. Signals to produce vergence include blurring of the retinal image by a large degree of retinal disparity, the extent of accommodation, or monocular cues to distance. These all require the visual cortex. Fast vergence movements occur during saccades.

H1 Acoustics and audition

Key Notes

Sound waves

Sound is longitudinal pressure waves in a medium. The frequency of the wave is the pitch of the sound and given in Hertz (cycles per second).

Sound pressure amplitude

The change in pressure produced by a sound wave is the sound pressure amplitude (P). By comparing P with a reference amplitude at the threshold of human hearing a sound pressure level (SPL) can be calculated. The unit of SPL is the decibel. Differences in SPL are perceived as differences in the loudness of a sound. Loudness varies with frequency in a manner that is determined by the sensitivity of the ear.

Frequency response of human hearing

In the young the detectable frequency range is between 20 Hz and 20 kHz. Peak sensitivity (the ability to hear quiet sounds) and auditory acuity (the ability to discriminate tones) both lie between 1000–4000 Hz. Two pure tones are better discriminated if they are heard in succession than at the same time.

Related topics

Anatomy and physiology of the ear (H2)

Peripheral auditory processing (H3)

Central auditory processing (H4)

Sound waves

Sound is the oscillation of molecules or atoms in a compressible medium. The energy of the oscillations is transmitted as a longitudinal wave in which the medium is alternately compressed and rarefied, causing periodic variations in the pressure of the medium (*Fig. 1*).

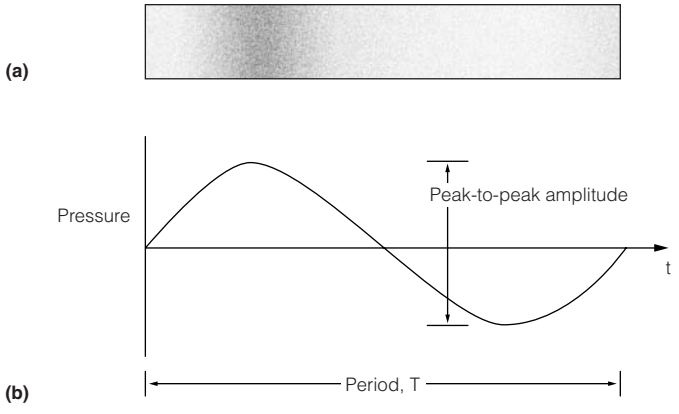


Fig. 1. Sound waves: (a) density of air molecules during propagation of a longitudinal pressure wave; (b) sine wave representation of a pressure wave.

For a sine wave the **period**, T, is the time taken for one complete cycle. The **frequency** of the wave, the perceived **pitch** of the sound, is the reciprocal of the period (i.e. f = 1/T). The unit of frequency is the **Hertz** (**Hz**); one cycle per second.

Sound pressure amplitude

The amplitude of a sound wave is the total change in pressure that occurs during a single cycle. Because of the huge range in sound wave amplitudes, P, it is expressed in a logarithmic scale as a ratio of a reference pressure, P_{ref}.

$$\text{Sound pressure level (SPL)} = 20 \log_{10} P/P_{ref}$$

P_{ref} is 2×10^{-5} Pa, a sound pressure which is at the threshold of human hearing. The unit of SPL is the **decibel** (dB). Each ten-fold increase in SPL is equivalent to 20 dB. Sound pressure levels in excess of 100 dB can result in damage to hearing, and at 120 dB auditory pain results.

Differences in sound pressure level are perceived as differences in **loudness**. Loudness varies with frequency so that for **tones** (single-frequency sounds) that have the same SPL; it is greatest at 4000 Hz and falls off dramatically above this value and below 250 Hz. This reflects the sensitivity of the ear to sounds of different frequencies.

Frequency response of human hearing

The **frequency response** of the human ear is from 20 Hz to 20 kHz optimally, but rapidly narrows with age, with most of the loss occurring at the higher frequencies. By 50 years the upper limit averages about 12 kHz. The highest **sensitivity** (the ability to detect quiet sounds) occurs over 1000–4000 Hz. This closely matches the frequency range of human speech, 250 Hz–4000 Hz. Greatest **auditory acuity** (the ability to discriminate between tones) is also found between 1000 and 4000 Hz. Over this bandwidth, tones differing in frequency by only 2–6 Hz can be differentiated if they are heard in succession. For comparison, the smallest interval in conventional Western music, the semitone, corresponds to frequency differences of 62–234 Hz over this bandwidth.

The same resolution does not, however, apply to pure tones played *simultaneously*, which must differ by about one third of an octave before they can be told apart. However, virtually no naturally occurring sounds or musical notes are pure tones, but have harmonics that span many octaves and permit them to be much better resolved even when they occur at the same time.

H2 Anatomy and physiology of the ear

Key Notes

The middle ear

The middle ear converts pressure waves in the air into vibrations of perilymph in the inner ear. Sound waves striking the ear drum cause it to vibrate. This vibration is transmitted by three articulated middle ear bones, the malleus, incus and stapes, to the oval window and so to the perilymph. Because perilymph is incompressible it is set in motion *en mass*, with the pressure being transmitted through to the round window. The surface area of the oval window is 20 times less than that of the ear drum. This gives a four-fold amplification of the sound across the middle ear. Two middle ear muscles, upon contraction, act on the middle ear bones to reduce sound transmission. They are activated by tympanic reflexes which may afford some protection against loud sounds.

Inner ear anatomy

The auditory inner ear is the cochlea, a bony canal arranged in a coil. Within it lies the cochlea duct, a part of the membranous labyrinth, which divides the cochlea cross-section into three compartments. Within the cochlea duct is the scala media, which contains endolymph. On either side lie the perilymph-containing scala vestibuli and scala tympani. These are continuous with each other at the apex of the cochlea so that vibrations of the oval window are propagated through the scala vestibuli to the scala tympani and then to the round window. Pressure waves moving through the perilymph cause the basilar membrane on the floor of the scala media to oscillate. The organ of Corti, a sheet of epithelium running the length of the cochlea duct rests on the basilar membrane. Hair cells in the organ of Corti have their stereocilia embedded in a gelatinous matrix, the tectorial membrane.

Cochlea function

Vibrations of the basilar membrane cause it to shear with respect to the tectorial membrane and bend the stereocilia to and fro. This results in periodic hair cell depolarization and hyperpolarization by the same transduction mechanism that operates in vestibular hair cells. The periodic changes in the release of transmitter from the hair cells alter the firing of the auditory primary afferents with which they synapse. Because the basilar membrane differs in width, mass and stiffness systematically along its length, different frequencies of sound make the membrane vibrate maximally at different distances along it. This is the basis of pitch discrimination.

Related topics Balance (F5) Acoustics and audition (H1)

The middle ear

The function of the middle ear is to convert pressure waves in the air into vibrations of the perilymph in the inner ear. Sound waves pass along the **external auditory meatus** striking the **tympanic membrane** (ear drum), which resonates faithfully in response. The ear drum is critically damped in that it stops vibrating the instant the sound ceases. A sound at hearing threshold causes the ear drum to vibrate with an amplitude of about 0.01 nm – one tenth the diameter of a hydrogen atom! The movement of the ear drum is transferred with an overall efficiency of about 30% to the fluid in the inner ear by a lever system, composed of three **ear ossicles**, lying in the **tympanic cavity** (middle ear) (*Fig. 1*).

The **malleus** (hammer) is fixed at its thin end (the handle) to the tympanic membrane. Its thick end (the head) articulates with the head of the **incus** (anvil) via a saddle-shaped joint. The long process of the incus makes a ball and socket joint with the head of the **stapes** (stirrup). The base of the stapes is attached by an annular ligament to the **oval window** (fenestra vestibuli). The malleus vibrates with the tympanic membrane. Inward movement locks the joint between the malleus and the incus driving the long process of the incus inward, pushing the stapes in the same direction to exert a pressure on the perilymph beyond the oval window. This pressure wave is transmitted through the perilymph to cause a compensatory bulge of the **round window** (fenestra cochleae). Outward movement reverses these motions. Since the area of the oval window is 20 times smaller than the tympanic membrane, the pressure (force per unit area) at the oval window is proportionally greater. This is important because perilymph is incompressible and so must be driven to vibrate *en masse*. This needs more force than it takes to transmit sound waves through air. In addition it results in an amplification of the sound by about 20 dB, corresponding to a four-fold increase in loudness, by the middle ear.

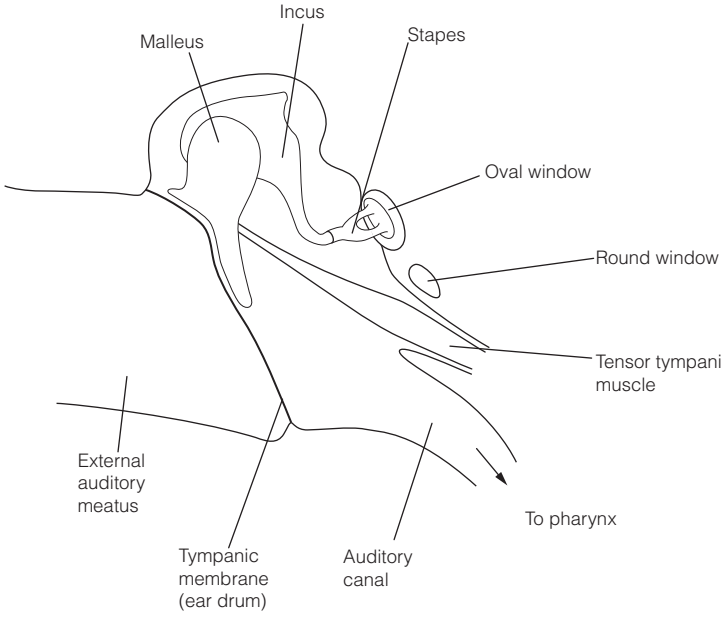


Fig. 1. The anatomy of the middle ear.

There are two middle ear muscles, the **tensor tympani** and **stapedius**. When they contract together the handle of the malleus and the tympanic membrane are pulled inwards and the base of the stapes is pulled away from the oval window. This reduces sound transmission by 20 dB, especially for low frequencies. Reflex contraction of these muscles in response to loud noise may prevent damage to the inner ear but since the reaction time is 40–60 ms this **tympanic reflex** affords no protection against *brief* loud sounds. The **auditory canal** connects the middle ear to the pharynx which allows the air pressure to be equalized to ambient pressure. This is important when going to high altitude. Impairment to middle ear function causes **conduction deafness**. The most common cause is **otosclerosis**, a bone disease in which the stapes becomes fused to the oval window. It is amenable to surgical correction.

Inner ear anatomy

The auditory part of the inner ear is the **cochlea**, a bony canal 3.5 cm long, which spirals two and three-quarter turns around a central pillar, the **modiolus**. Within the cochlea lies a tubular extension of the membranous labyrinth, the **cochlear duct**, attached to the modiolus and the outer wall of the cochlea. This divides the cochlea into three compartments, the **scala media**, which contains endolymph, the **scala vestibuli** and the **scala tympani**.

The two latter compartments contain perilymph and are continuous with each other via a small gap known as the **helicotrema** situated at the apex end of the cochlea, where the cochlear duct ends blindly (*Fig. 2*). Pressure waves generated at the oval window are propagated through the scala vestibuli into the scala tympani and so to the round window where the energy dissipates. During their passage the pressure waves cause oscillations of the **basilar membrane**, the floor of the scala media on which rests the sensory apparatus, the **spiral organ of Corti** (*Fig. 3*).

The spiral organ is a narrow sheet of columnar epithelium running the length of the cochlear duct. The epithelium consists of supporting pillar cells and Hensen's cells, and sensory hair cells resembling those in the vestibular apparatus. A single row of 3500 **inner hair cells** form **ribbon synapses** with myelinated axons of large bipolar cells (type I) in the **spiral ganglion** of the cochlear nerve. Each inner hair cell is innervated by about tan such axons, a large degree of divergence. There are about 12 000 **outer hair cells** arranged in three rows. These are innervated by an unmyelinated axon from small bipolar cells (type II)

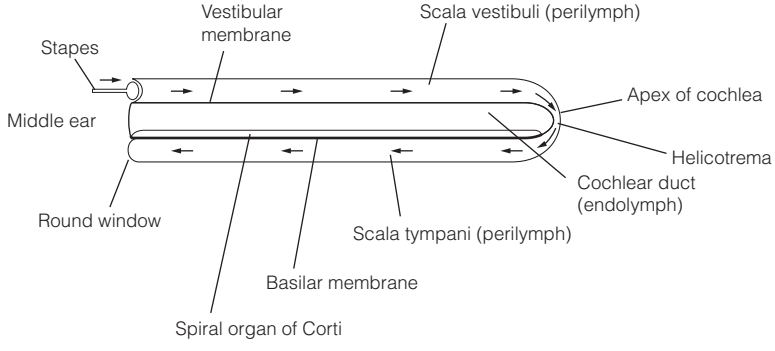


Fig. 2. The cochlea, depicted unfurled. Arrows show the direction of propagation of sound waves through the perilymph.

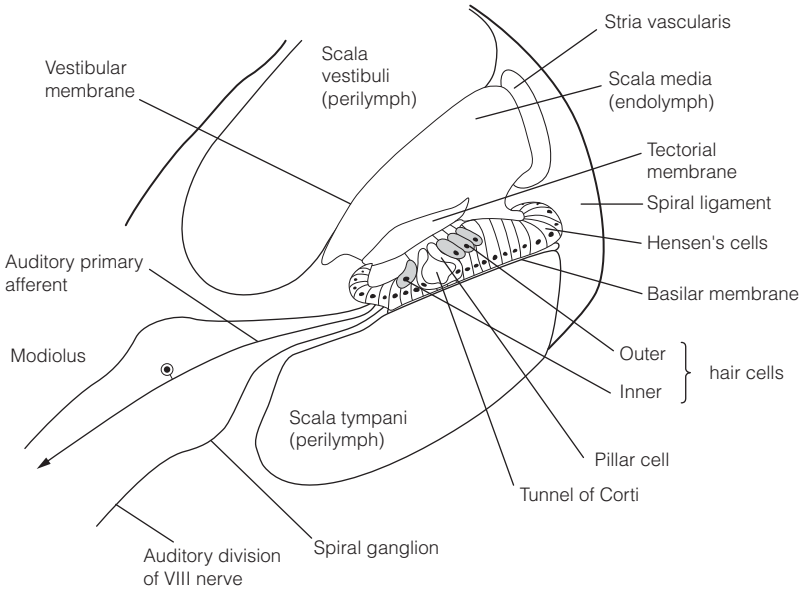


Fig. 3. Transverse section through the cochlea showing the organ of Corti.

in the spiral ganglion, each of which synapses with ten hair cells, representing considerable convergence.

Cochlea function

Cochlea hair cells lose their kinocilia during development and the tips of their tallest stereocilia are embedded in the overlying **tectorial membrane**, a matrix of mucopolysaccharides and proteins. Oscillations of the basilar membrane in response to a sound stimulus cause it to shear with respect to the tectorial membrane, bending the stereocilia first one way and then the other. This results in periodic depolarization and hyperpolarization of the hair cells, producing cyclical alterations in the tonic secretion of glutamate. The transduction mechanism for hair cells is like that of vestibular hair cells.

A sound stimulus causes a **traveling wave** (like that generated by twitching the free end of a rope fixed at its other end) to spread from the base to the apex of the basilar membrane. High frequencies cause vibration at the basal end whereas low frequencies cause vibration towards the apex. This frequency sorting is a result of the continuous variation in the width, mass and **stiffness** of the basilar membrane along its length. The basilar membrane is narrow (50 μm) and stiff at the base, wider (500 μm) and less stiff at the apex. The relationship between frequency and length is logarithmic. At a given frequency, as the SPL increases so do the amplitude of the displacement and the length of the basilar membrane vibrating.

Outer hair cells (OHCs) contract in a voltage-dependent manner. Depolarization causes them to shorten. The speed with which they change length is so fast that they are able to follow the high-frequency voltage changes produced by sound stimuli. By this means OHCs augment the vibrations of the basilar membrane, a process called **cochlear amplification**. It probably contributes to the high sensitivity and fine tuning to frequency exhibited by the basilar membrane, since these features are lost when OHCs are selectively

damaged by **aminoglycoside antibiotics** such as streptomycin. Cochlear amplification causes vibrations of perilymph that are transmitted to the oval window across the middle ear in the 'wrong' direction to the tympanic membrane which now acts as a loudspeaker producing **otoacoustic emissions**. They may occur spontaneously or be evoked by sound. They are usually pure tones, inaudible and are *not* the cause of **tinnitus** (ringing in the ears), the origin of which is unknown. Otoacoustic emissions are not necessary for normal audition but provide insights into ear function and are used clinically to test the hearing of babies.

H3 Peripheral auditory processing

Key Notes

Primary auditory afferents

Primary auditory afferents have their cell bodies in the spiral ganglion and send their central axons to the pons via the auditory (VIII cranial) nerve. Auditory afferents fire tonically and increase firing in response to a tone. Most are sharply tuned in that, for low-sound pressures, they show highest sensitivity for a narrow range of frequencies.

Coding of sound frequency

Frequency is coded in two ways. For frequencies above 3000 Hz the frequency response of an afferent depends on where along the basilar membrane it is from. This is a place coding; specifically the mapping of frequency to position is referred to as tonotopic mapping. For lower frequencies, afferents fire during a particular phase of a sound wave. This is called phase locking. A population of afferents encodes the entire waveform.

Coding of sound level

An auditory afferent responds only to a limited range of sound pressure levels. The full range is encoded by afferents with different dynamic ranges. Afferents that display the highest rate of spontaneous firing are the most sensitive. Efferents from the superior olivary complex make inhibitory synapses with outer hair cells, reducing the sensitivity of auditory afferents with increasing sound pressure levels.

Related topics Acoustics and audition (H1) Central auditory processing (H4)

Primary auditory afferents

The primary afferents have their cell bodies in the **spiral ganglion** located in the modiolus. Their centrally directed axons project through the vestibulocochlear (VIII) nerve to synapse in the cochlear nuclei of the lower pons. In humans, about 30 000 type I afferents from inner hair cells provide the bulk of the output from the cochlea. Three-quarters of the hair cells (the OHCs) send their output to only about 3000 type II afferents. The nature of type II cell signaling is unknown. Auditory afferents fire tonically.

In response to a tone, type I afferents show an increase in firing which adapts. When the sound stops, firing ceases for a brief period. Hence they exhibit both dynamic and static responses (*Fig. 1*). Responses of type I afferents plotted as **tuning curves** (*Fig. 2*) show that they are sharply tuned at low sound pressure

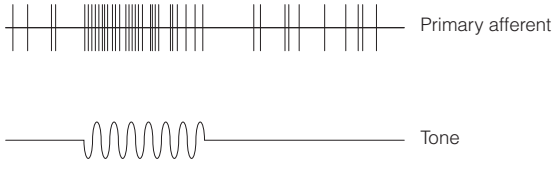


Fig. 1. Firing of auditory primary afferent (upper trace) in response to a tone (lower trace).

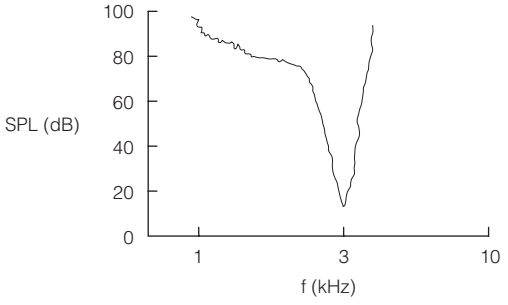


Fig. 2. Tuning curve for a type I cochlear afferent. The plot shows the minimum SPL required to evoke a response over a range of frequencies.

levels. The frequency to which the unit is most sensitive is the **characteristic frequency** (CF). At high SPLs the primary afferents respond to a much wider frequency range.

Coding of sound frequency

The nervous system has two ways of encoding the frequencies of a sound. **Place coding** is possible because the CF of an afferent is determined by the position of its hair cells along the cochlea. Fibers with successively lower CFs are found closer to the apex of the cochlea. This mapping of frequency to position is known as **tonotopic mapping** and is retained in the central auditory pathway. Place coding is most important for frequencies above 1–3 kHz. For lower frequencies coding uses the property that afferents fire with greatest probability during a particular phase of a sound wave, **phase-locking**. It is only necessary that an individual afferent fires during some cycles if a group of cells is involved. Moreover, if different groups phase-lock onto different parts of the cycle then a whole population of cells acting in concert can encode frequency.

Coding of sound level

Auditory afferents have a **dynamic range** of about 30 dB, beyond which further increase in SPL has no additional effect. The full range of SPL (0–100 dB) is signaled by afferents with different sensitivities. Cells with the same CF may differ in threshold SPL by 70 dB. Afferent sensitivity correlates with its spontaneous rate (SR) of firing. High SR cells have the greatest sensitivity. Low SR cells are least sensitive and are concerned with encoding the frequencies of loud sounds.

The sensitivities of afferents can be modified by efferents which have their cell bodies in the **superior olivary complex** (**SOC**) and form inhibitory synapses on OHCs. These neurons alter the gain of the cochlear amplifier, reducing the sensitivity of type I afferents as sound pressure levels increase.

H4 Central auditory processing

Key Notes

Central auditory pathways

Primary auditory afferents terminate in the cochlear nuclei in the pons. Ventral cochlear axons go to the superior olivary nucleus on both sides. This structure projects to the nuclei of the lateral lemniscus, and is primarily concerned with localizing the direction of a sound source. The dorsal cochlear nucleus projects directly to the contralateral nucleus of the lateral lemniscus. The nuclei of the lateral lemniscus sends axons to the inferior colliculus, which in turn projects to the medial geniculate nucleus (MGN). The auditory radiation, which originates in the MGN, goes to the primary auditory cortex, and is responsible for conscious sound perception. Although the largest auditory pathway is contralateral, extensive connections across the midline ensure interactions between sides.

Cochlear nuclei

Distinct cell types within the cochlear nuclei are able to process different features of a sound stimulus. Bushy cells signal exact timing information to the superior olivary nucleus which, by comparing input from both ears, is able to localize sound. Stellate cells signal sound level. Many cells are finely tuned to particular frequencies and show lateral inhibition which sharpens this tuning.

Tonotopic mapping

Maps in which frequency is represented in a systematic way occur in all auditory structures. In humans all frequencies have about equal neural representation. In the primary auditory cortex isofrequency columns are found perpendicular to the cortical surface. These are arranged in bands which form an ordered tonotopic map.

Sound level

Cells responding to differences in sound level are found throughout the auditory system. Some are finely tuned to a characteristic sound level. In humans there are no maps of sound level.

Localization of sound

The source of a sound can be localized in the vertical plane (elevation) and in the horizontal plane (azimuth). Elevation is signaled by the delay caused by sound waves being reflected from the pinna (external ear). The azimuth of a sound source is measured in two ways. For higher frequencies, the difference in sound level between the ear nearest and that furthest from the sound source is computed by neurons in the lateral superior olivary nucleus. This nucleus projects to the tectum, which controls eye and head reflexes in response to sound. For lower frequencies, cells in the medial superior olivary nucleus compute the phase difference that occurs because sound entering the ear furthest from the source is slightly delayed. These time differences are mapped topographically in the medial superior olivary nucleus. In the auditory cortex most cells respond preferentially to input in the contralateral ear and are either excited or inhibited by ipsilateral input.

Related topics

Location coding (E3)
Acoustics and audition (H1)
Peripheral auditory processing (H3)

Central auditory pathways

Primary auditory afferents bifurcate to terminate in both the ventral and dorsal **cochlear nuclei**. From the ventral cochlear nucleus axons run to the **superior olivary nucleus** (**SON**) on both sides and the contralateral inferior colliculus. The auditory fibers that cross the pons constitute the **trapezoid body** (TB). The SON compares input from the two ears to compute the whereabouts of a sound source. It projects to the nuclei of the lateral lemniscus. The dorsal cochlear nucleus sends axons directly to the contralateral nucleus of the lateral lemniscus (*Fig. 1*).

The nucleus of the **lateral lemniscus** projects to the **inferior colliculus** (**IC**). The IC relays with the **medial geniculate nucleus** (**MGN**) of the thalamus which sends its output via the **auditory radiation** to the **primary auditory cortex**, A1 (Brodmann's areas 41 and 42) located in the superior temporal gyrus. This is the pathway for conscious auditory perception. The biggest auditory pathway is contralateral. However, reciprocal connections between the nuclei of the lateral lemniscus (via Probst's commissure) and between the IC, ensure extensive interactions between the input from both ears.

The ascending sensory pathways are matched by descending projections. The auditory cortex sends axons back to both the medial geniculate nucleus and the inferior colliculus and the latter in turn projects to the SON and the cochlear nuclei. The SON sends efferents to the spiral organ to modify the sensitivity of cochlear afferents.

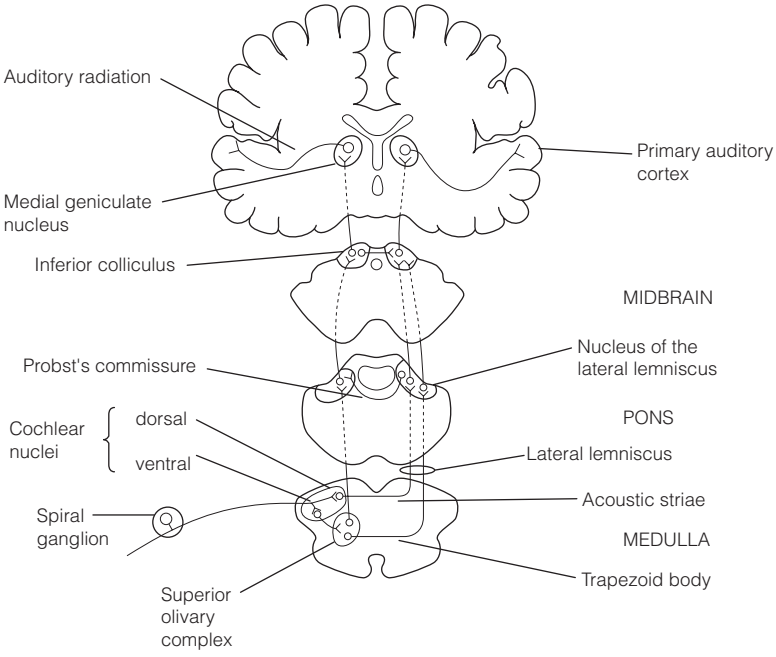


Fig. 1. Central auditory pathways.

Cochlear nuclei

Central auditory pathways process three features of sound input in parallel, tone, loudness and timing. From the last two the brain calculates the location of the sound in space. Parallel processing begins in the cochlear nuclei.

Several neuron types are present in the cochlear nuclei that can be distinguished both by their shape and responses. Numerous **bushy cells** reproduce

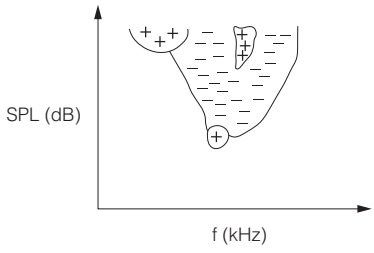


Fig. 2. Response map of a type IV cell in the dorsal cochlear nucleus. Excitatory region, +; inhibitory region, –.

the firing pattern of the primary afferents faithfully, including phase-locking. Their output, which precisely signals the timing of sound, goes to the superior olivary nucleus which compares the input from both ears to compute the location of a sound source. By contrast **stellate cells** have a much greater dynamic range than bushy cells and signal sound level. Hence, timing and sound level are processed in parallel.

Receptive fields of auditory neurons are called **response maps** and are plotted in the same way as the primary afferent tuning curve. Five classes of cell can be distinguished in the cochlear nuclei on the basis of their RFs. Type I cells have a purely excitable RF that precisely matches primary afferent tuning curves, but all the other types have inhibitory responses which arise by lateral inhibition and which fine tune their frequency response (*Fig. 2*). Type IV cell axons are the main output of the dorsal cochlear nucleus.

Tonotopic mapping

Tonotopic maps are found in the cochlear nuclei, superior olivary nucleus, inferior colliculus and auditory cortex. Some structures have multiple maps. The cochlear nucleus is divided into isofrequency strips, each containing cells with similar characteristic frequencies (CFs). Strips representing increasingly higher

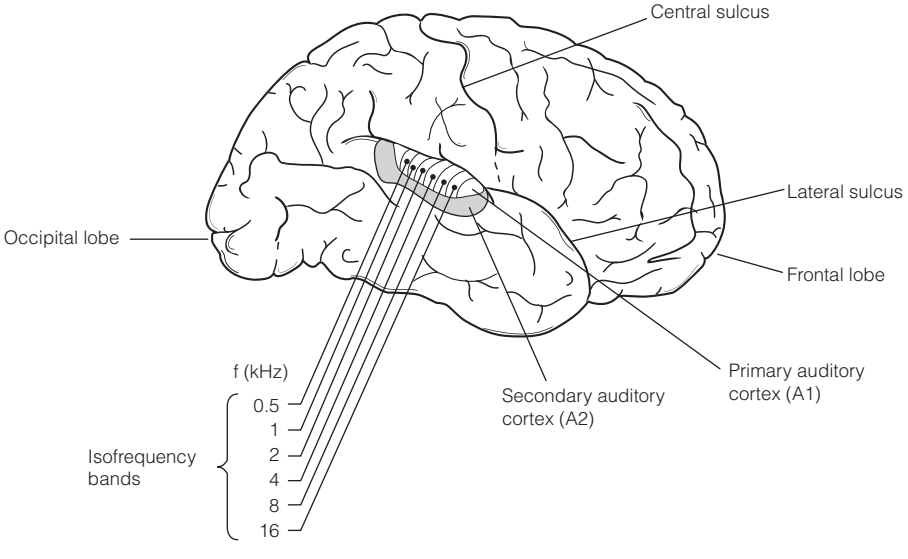


Fig. 3. Tonotopic map of the human primary auditory cortex.

frequencies are found progressively more posteriorly. In the primary auditory cortex, **isofrequency columns** running through the entire thickness of the cortex are arranged in **isofrequency strips** running mediolaterally, with low frequencies represented rostrally and high frequencies caudally (*Fig. 3*). There are at least three other tonotopic maps in the auditory cortex. Adjacent maps are always mirror images of each other. In humans there is no great over-representation of particular frequencies.

Some regions of auditory cortex, e.g. **secondary auditory cortex** (**AII**), are less well tonotopically organized and contain cells that respond to a wider range of frequencies.

Sound level

Cells throughout the auditory system respond to differences in sound level and fall into two broad classes. Monotonic cells have sigmoid plots of sound level against firing rates. Non-monotonic cells are more finely tuned with a maximum firing rate at a characteristic sound level. Maps of sound level are found in bats which echolocate but have not been found in other species, including humans.

Localization of sound

The ability to localize the source of a sound in space is very important in avoiding danger. The co-ordinates of a sound source in vertical and horizontal planes are **elevation** and **azimuth** respectively. Different mechanisms are involved in determining these two co-ordinates.

For finding elevation the **pinna** of the outer ear is crucial. Sound waves entering the ear do so by two routes, one direct, the other reflected from the pinna, will arrive at slightly different times at the ear drum. Sound coming from different directions in the vertical plane will be reflected differently, because of the peculiarities of the shape of the pinna, and so have different delay times (*Fig. 4*). The auditory system uses the delay times to compute the sounds' position in the vertical plane. Although the human pinna is small and immobile it is important in this respect.

The superior olivary complex uses two methods to localize sound in the horizontal plane. Both compare input into the two ears (**binaural sound localization**) and allow azimuth to be pinpointed with a precision of about one degree of arc.

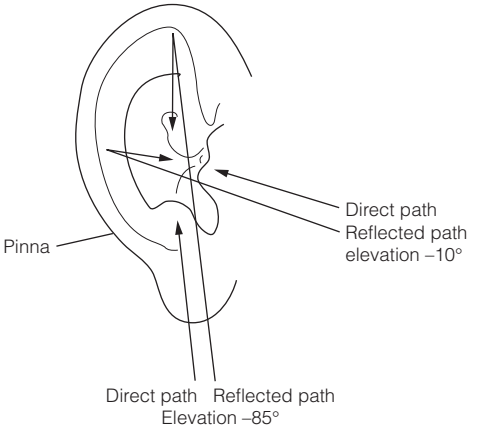


Fig. 4. Role of the pinna in localizing the sound direction in elevation.

1. **Interaural level differences** (ILDs). If the head is orientated so that one ear is closer to the sound source, then the head forms a shadow which reduces the sound level entering the other ear. Using ILD to find azimuth is most accurate for high frequencies. The brain can use ILDs as low as 1 dB to compute azimuth. Neurons of the **lateral superior olivary nucleus** (LSO) have a tonotopic map restricted largely to high-frequency input. These cells receive inputs from both ipsilateral and contralateral cochlear nuclei. However, the contralateral route is by way of a glycinergic inhibitory neuron. Equal sound level in both ears causes overall inhibition of the LSO neuron and increasing the sound level in the contralateral only serves to augment the inhibition. However, increased sound level to the ipsilateral ear causes LSO firing. Maximum firing rate is seen when ILD is 2 dB or more. Corresponding cells in the opposite LSO will show reverse responses to the same sound. The LSO projects to the ventromedial part of the IC central nucleus. The IC connects extensively with the deep layers of the **superior colliculus** to form an **auditory space map** in register with the retinotopic map. Hence the superior colliculus is implicated in the auditory reflexes organizing gaze and head rotation towards the sound source.
2. **Interaural time differences** (ITDs). A sound wave enters the closer ear slightly earlier than the further one. For low frequencies (< 3 kHz) this results in a phase difference, in which the time delay is less than one period, which can be analyzed by neurons capable of phase-locking. At higher frequencies input into the furthest ear is delayed by more than a single period, and this makes phase-locking unreliable, so ITDs cannot provide an unambiguous cue to location. ITDs as short as 20 μs can be detected.

The neural system for measuring ITDs depends on cells in the **medial superior olivary nucleus** (MSO) acting as coincidence detectors. The MSO has inputs from bushy cells in both cochlear nuclei that phase-lock in response to low-frequency stimuli. If a phase difference exists between the two ears then the bushy cells corresponding to the furthest ear fire slightly later. The MSO circuitry for transforming this timing difference to azimuth has been worked out for the owl, but the mammalian circuit is obviously different and not yet understood.

Most cells in A1 are binaural (respond to input from either ear) and fall into two groups of cortical columns. Cells within a column show similar binaural responses. Those in **summation columns** show bigger responses to input from both ears than to one. By contrast, cells in **suppression columns** have a preference for input from one ear (*Fig. 5*), with weaker responses to input from the

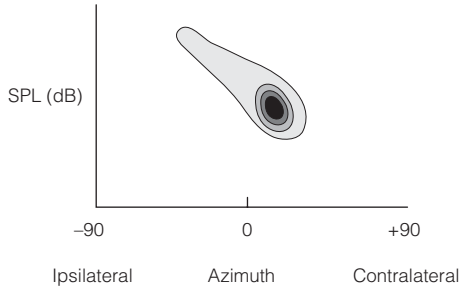


Fig. 5. Location tuned neuron in a suppression column of the auditory cortex. Higher density of shading corresponds to a greater firing rate.

other ear or both ears. Summation and suppression columns are arrayed alternately and at right angles to isofrequency strips. There are clear similarities between the organization of the primary auditory and visual cortices, and it is likely that summation and suppression columns form the basis for a spatial map of sound location. However, despite the fact that lesions of A1 (in cats) do impair localization of sound on the contralateral side, no orderly map of sound location has been discovered in the auditory cortex.

I1 Olfactory receptor neurons

Key Notes

Olfactory epithelium

Bipolar olfactory receptor neurons lie in the olfactory epithelium. Their dendrites extend to the surface of the epithelium to form a swelling with a cluster of olfactory cilia. Their axons form the olfactory (cranial I) nerve and synapse with neurons in the olfactory bulb.

Olfactory transduction

Odor molecules are detected by odorant receptors in the olfactory cilia. The receptors are a family of about 1000 G protein-coupled receptors in which the associated G protein (G_{olf}) is related to G_s. Most odorant receptors are positively coupled to the cAMP second messenger system. The rise in cAMP on binding of an odorant molecule opens a cation channel, resulting in a depolarizing generator potential that is graded according to the concentration of odorant. Each olfactory sensory neuron expresses just a single species of odorant receptor and each odorant receptor binds a range of related molecules with differing affinities.

Related topics Metabotropic receptors (D2) Taste (I3)

Olfactory epithelium

A sense of smell in humans is important in feeding and probably also in sexual behavior. The olfactory epithelium lies in the dorsal nasal cavity. It consists of bipolar **olfactory receptor neurons** (ORNs) and supporting cells. A dendrite emerges from one pole of the ORN and extends to the surface of the epithelium, where it forms a swelling that gives rise to a cluster of 6–12 immobile **olfactory cilia**. These project into a layer of mucus secreted by the supporting cells. The centrally directed unmyelinated axon passes, in the olfactory (cranial I) nerve, through the **cribiform plate** of the **ethmoid** bone to synapse with cells in the **olfactory bulb**. Human olfactory epithelium contains about 10^8 ORNs. The mucus provides a medium to absorb airborne odor molecules, which then reach the high surface area presented by the densely packed layer of olfactory cilia.

Olfactory transduction

Odor molecules are usually small (M_r < 200 Da), lipid soluble and volatile. Initially they bind to **odor-binding proteins** in the mucus which concentrate the odor molecules in the vicinity of the cilia. Odor molecules are recognized by **odorant receptors** in the cilia plasma membrane. These are G-protein-coupled receptors, and around 1000 have been identified in mammals. Each odorant receptor, unlike the G-protein-coupled receptors for neurotransmitters, binds a range of related odor molecules with various affinities. Each odor molecule can interact with 2–6 odorant receptor subtypes. Each ORN expresses just a single subtype of odorant receptor. Because odorant receptors are relatively non-specific, individual ORNs respond to a number of odors. The mammalian nervous system is able to discriminate some 10 000 distinct odors on the basis of

precisely which array of odorant receptors (and so which sensory neurons) are stimulated, and with what relative intensities.

Odorant receptors are coupled to G proteins which are closely related to the G_s proteins that stimulate adenylyl cyclase, and are termed $\mathbf{G_{olf}}$. Most odorant receptors are linked to the cAMP second messenger system. Binding of an odor molecule causes a rise in cAMP within about 50 ms. This activates a cyclic-nucleotide-gated (CNG) channel, a non-specific cation conductance allowing the flow of Na^+, K^+ and Ca^{2+} ions (*Fig. 1*). The resulting depolarization of the dendritic knob of the ORN spreads passively across the cell membrane to trigger action potentials at the axon hillock. The generator potential is graded with an amplitude that signals the concentration of the odor molecule. However, a maximal response is produced by the opening of only a small fraction (3–4%) of the CNG channels available. This means that the concentration range that can be signaled by firing of an ORN is narrow, about a 10-fold difference.

High odor concentration or prolonged exposure allows a high Ca^{2+} influx through the CNG channels. This ion has a number of modulatory effects in olfactory receptor neurons. Ca^{2+} activates heme oxygenase 2, an enzyme that synthesizes carbon monoxide (CO), which can activate **guanylyl cyclase** (GC) as shown in *Fig. 1*. Because Ca^{2+} also inhibits GC, there is no overall activation of the cyclase in the target ORN. However, CO is freely diffusible so it can activate GC in adjacent *unstimulated* ORNs, producing cyclic guanosine monophosphate (cGMP) which binds to and opens the CNG channels. In this way odorant excitation spreads to a cluster of ORNs. Since neighboring ORNs respond to the same odors this does not produce a loss of specificity.

ORNs show **adaptation** to protracted stimulation. Ca^{2+} binds to calmodulin (CaM) which can then bind to CNG channels, reducing the efficacy with which the cyclic nucleotides can open them. Hence Ca^{2+} attenuates the size of the generator potential.

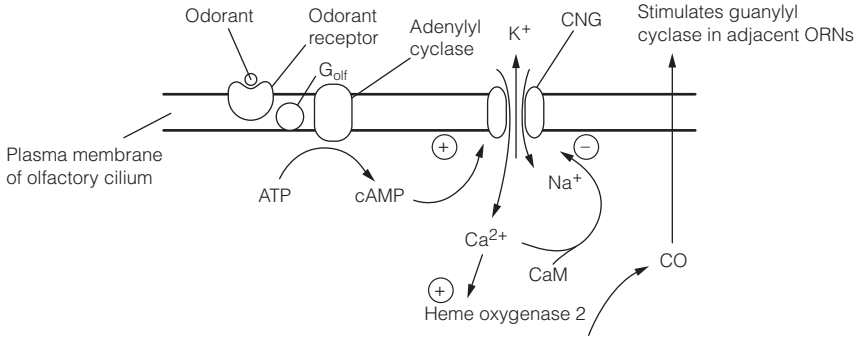


Fig. 1. Olfactory transduction in olfactory receptor neurons mediated by receptors coupled to the cAMP second messenger system. CaM, calmodulin; CNG, cyclic-nucleotide-gated channel.

I2 Olfactory pathways

Key Notes

Olfactory bulb

Olfactory receptor neurons synapse with mitral cells or tufted cells (M/T) and periglomerular cells within glomeruli in the olfactory bulb. Glomeruli are odor specific; each one receives terminals from receptor neurons that respond to the same set of odors. Odor discrimination is maximized by lateral inhibition which suppresses glomeruli responding to slightly different odorants. A given odor excites a group of glomeruli.

Central olfactory connections

M/T neurons project to the olfactory cortex via the olfactory tract. Olfactory cortex is three-layered paleocortex (old cortex); the only cortex to receive sensory input directly rather than by way of the thalamus. Olfactory cortex projections to hypothalamus and amygdala are important for emotional and motivational aspects of odors, a pathway to the hippocampus is concerned with smell memory, and output via the thalamus to orbitofrontal cortex mediates the conscious perception of smell. Brainstem aminergic pathways alter the responsiveness of M/T cells in the context of the behavioral state (e.g. hungry or sated).

Related topic Olfactory receptor neurons (I1) Cell determination (N2)

Olfactory bulb

The axons of the ORNs run in the olfactory nerve to make excitatory synapses on the dendrites of **mitral cells** or **tufted cells** (M/T) and short axon inhibitory **periglomerular cells** in the olfactory bulb. M/T cells send their axons into the olfactory tract. Synapses between ORNs and M/T and periglomerular cells are found in **olfactory glomeruli**, spherical zones some 150 μm across. The olfactory bulb contains about 2000 glomeruli. Each glomerulus receives the terminals of 25 000 ORNs which respond to the same odors. Hence, glomeruli are odor-specific functional units (*Fig. 1*). Low concentrations of a given odor molecule activate cells in the single glomerulus which gets input from the ORNs bearing odorant receptors with the highest affinity for the molecule. At higher concentrations, cells in other glomeruli are activated as their ORN odorant receptors' low affinity binding sites for the molecule are occupied. Each glomerulus has dendrites from about 75 M/T cells. The M/T cells integrate weak inputs from a large number of ORNs within a glomerulus to generate a strong signal.

Lateral inhibition dampens responses from glomeruli with slightly different odor specificities so as to heighten odor discrimination. This is brought about by reciprocal dendrodendritic synapses between M/T cells and inhibitory interneurons termed **granule cells**. Via these synapses, M/T cells excite granule cells, which then inhibit the same, and adjacent M/T cells.

There is a topographical organization to the fibers of the olfactory nerve and their projections to the olfactory bulb. Thin strips of olfactory epithelium running in an anteroposterior direction go to neighboring glomeruli. A given odor excites a particular array of glomeruli across the olfactory bulbs, an **odor**

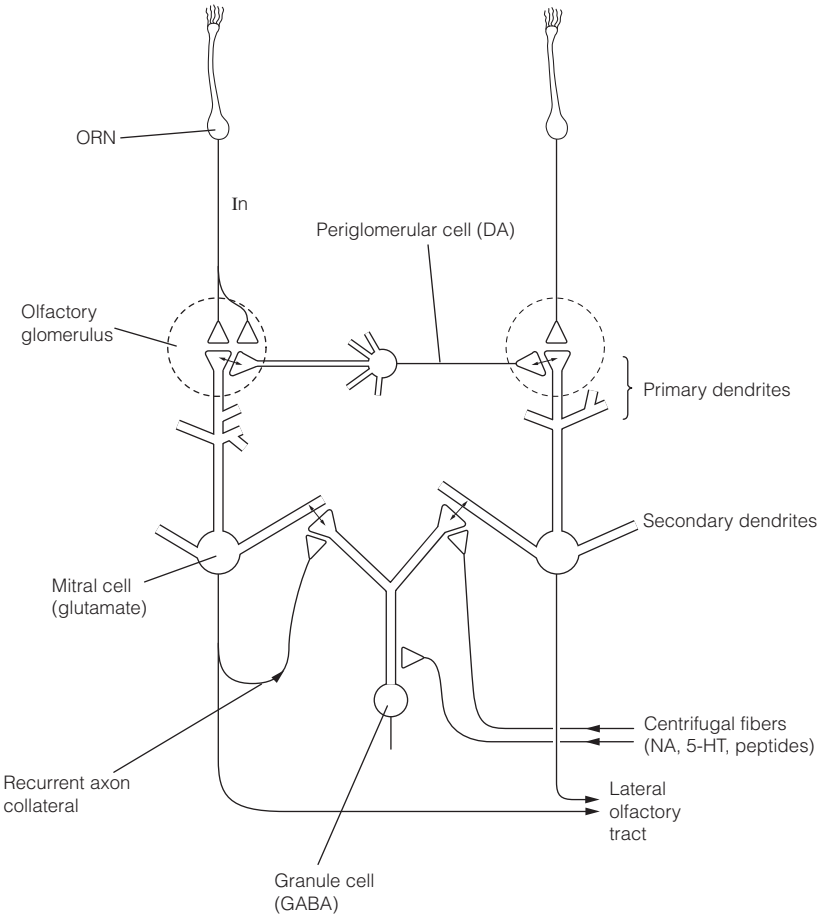


Fig. 1. Circuitry of the olfactory bulb. Reciprocal synapses are denoted by ↔. Neurotransmitters used by specific cell types are shown in parentheses: DA, dopamine; GABA, γ-aminobutyric acid; 5-HT, serotonin; NA, norepinephrine.

image. The higher the concentration of the odor molecule the bigger the area activated.

Central olfactory connections

M/T cell axons project via the **olfactory tract (OT)** to the **olfactory cortex**. This cortex is unusual is two respects. Firstly, it is **paleocortex** (old cortex), structurally resembling the forebrain cortex of non-mammalian vertebrates, in only having three layers. Secondly, it is the only cortex to receive sensory input directly, rather than via the thalamus.

There are five regions of the olfactory cortex with distinct connections and functions but all receive input from the olfactory tract. The **anterior olfactory nucleus** gives rise to axons that cross the midline in the **anterior commissure** to go to the contralateral olfactory bulb (*Fig*. 2). The **anterior perforated substance** (called the **olfactory tubercle** in non-primates) sends output to the posterior hypothalamus. This pathway, together with that to the **corticomedial amygdala**, which then projects to the medial hypothalamus, are concerned with the

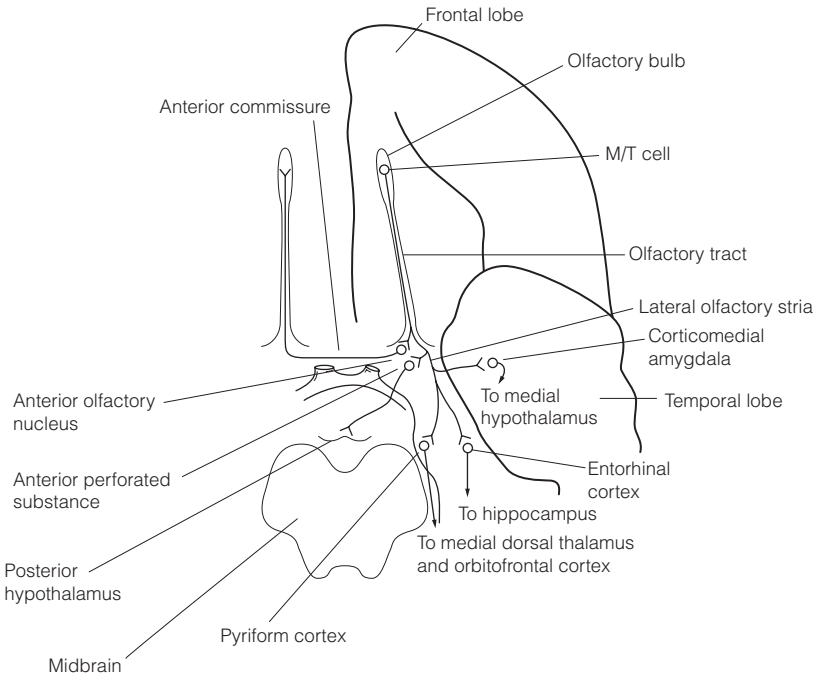


Fig. 2. Connections of the left olfactory cortex viewed from below. M/T, mitral and tufted cells.

affective and motivational aspects of odors and directly influence feeding and mating. A pathway from the olfactory tract to the entorhinal cortex, which sends its entire output to the hippocampus, presumably encodes olfactory components of episodic memories.

A large part of the olfactory cortex is the **pyriform cortex**. This is concerned with olfactory discrimination. It send axons which terminate in the **medial dorsal thalamus**, which in turn projects to the **orbitofrontal cortex**. This cortex mediates the conscious perception of smell.

Olfactory processing is subject to considerable modulation. The olfactory bulb receives inputs from noradrenergic and serotonergic neurons in the brainstem and cholinergic neurons in the forebrain. In addition the anterior perforated substance receives a projection from the brainstem dopaminergic system. These various inputs are implicated in modifying olfaction on the basis of the behavioral state and in olfactory learning. This is likely to be important in feeding and mating. In rats, for example, mitral cell responses to food odors depend on whether the animal is hungry or sated.

I3 TASTE

Key Notes

Gustation

The sense of taste provides a way in which harmful foods may be avoided and nutritious foods selected. There are at least five tastes: salty, sweet, sour, bitter and umami (glutamate). It is one of several senses (others include smell) involved in the oral sensory experience of eating. Taste and smell regulate autonomic responses to feeding.

Taste buds

Taste buds are clusters of neuron-like epithelial cells, the gustatory receptor cells. Microvilli on the apical border of gustatory receptor cells are in contact with the contents of the mouth via taste pores. Receptor cells make synaptic connections with gustatory primary afferents, the axons of which travel through the VII, IX or X cranial nerves. Taste buds are found not only in the tongue but also in the pharynx and upper esophagus.

Taste transduction

Salt taste is mediated by receptor cell depolarization brought about by the opening of amiloride-sensitive sodium channels. Hydrogen ions (sour taste) produce depolarization by blocking voltage-dependent K^+ channels. The sensation of sweetness involves the taste molecule binding to a G protein-linked receptor coupled (usually) to the cAMP second messenger system. This causes depolarization by closing a potassium channel. There are multiple transduction pathways for bitter molecules but all result in receptor cell depolarization. Umami is mediated by a metabotropic glutamate receptor.

Related topics

Metabotropic receptors (D2)

Gustation

The sense of taste, **gustation**, provides a means of avoiding potentially noxious foodstuffs or selecting for foods which have a high-energy content. Five tastes are well defined – salty, sour, sweet, bitter and unami (due to monosodium glutamate) – on the basis that no cross adaptation occurs between them. Plant alkaloids, some of which are toxic in high concentrations, are extremely bitter. A sour taste may signify a food degraded by microbiological action. By contrast a sweet food has a high content of sugars and so a readily available supply of metabolic energy. The sensory experience produced by having food in the mouth is called **flavor perception**, and relies on several sensory modalities. Apart from smell and taste, information about food texture is provided by mechanoreceptors and proprioceptors in the mouth and jaw innervated by trigeminal afferents. Flavor perception is important in triggering or modifying autonomic responses to feeding, e.g. salivation, gastric secretion and changes to gastrointestinal motility.

Taste buds

Gustatory receptor cells are epithelial cells but have many neuron-like features. They are organized into small clusters of 50–100 which, with supporting cells,

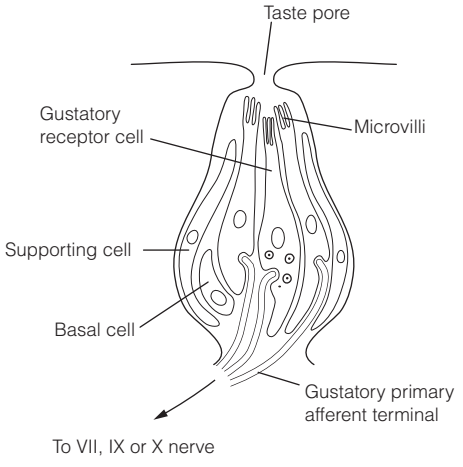


Fig. 1. A taste bud.

form **taste buds**. As with other epithelial cells, gustatory receptor cells are continually replaced from basal cells about every ten days. Microvilli on the apical border of each receptor cell project through **taste pores** in the gustatory epithelium, bringing them into contact with the contents of the mouth. The microvilli carry out taste transduction (*Fig. 1*).

Receptor cells form synaptic connections with gustatory primary afferent neurons. Each afferent branches to synapse with receptor cells in more than one taste bud. The axons of the primary afferents run through the facial (VII), glossopharyngeal (IX) and vagus (X) cranial nerves.

Taste buds are located in the epithelium of the tongue, palate, pharynx, epiglottis and the upper part of the esophagus. In the tongue they are present in small projections, **papillae**, supplied by the VII and IX nerves. The few taste buds in the epiglottis and esophagus are innervated by the X nerve.

Taste transduction

Many of the ions or molecules responsible for taste sensation are hydrophilic and freely diffusible. Those which are hydrophobic include plant alkaloids which may bind to proteins in the saliva, equivalent to odorant binding proteins, for presentation to gustatory receptor cells. Transduction involves changes in membrane conductance which causes a depolarizing generator potential, triggering action potentials, calcium influx and so neurotransmitter release. Gustatory receptor cells have voltage-dependent Na^+, K^+ and Ca^{2+} and are excitable.

Salt taste is caused by Na^+ ions. Salt transduction (*Fig. 2a*) occurs by the influx of Na^+ through an amiloride-sensitive Na^+ channel which depolarizes the receptor cell (i.e. producing a generator potential) so causing it to fire.

H^+ ions responsible for sour (acid) sensation cause a generator potential by blocking voltage-dependent K^+ channels in the apical membrane which at rest carry an outward, hyperpolarizing current. Blockade of other channels by protons may also contribute. The amplitude of the generator potential is proportional to the H^+ concentration.

Sugars, some amino acids and some proteins produce sweet sensations by interacting with metabotropic G protein-linked receptors coupled to second messengers. Sugars activate adenylyl cyclase and the consequent rise in cAMP

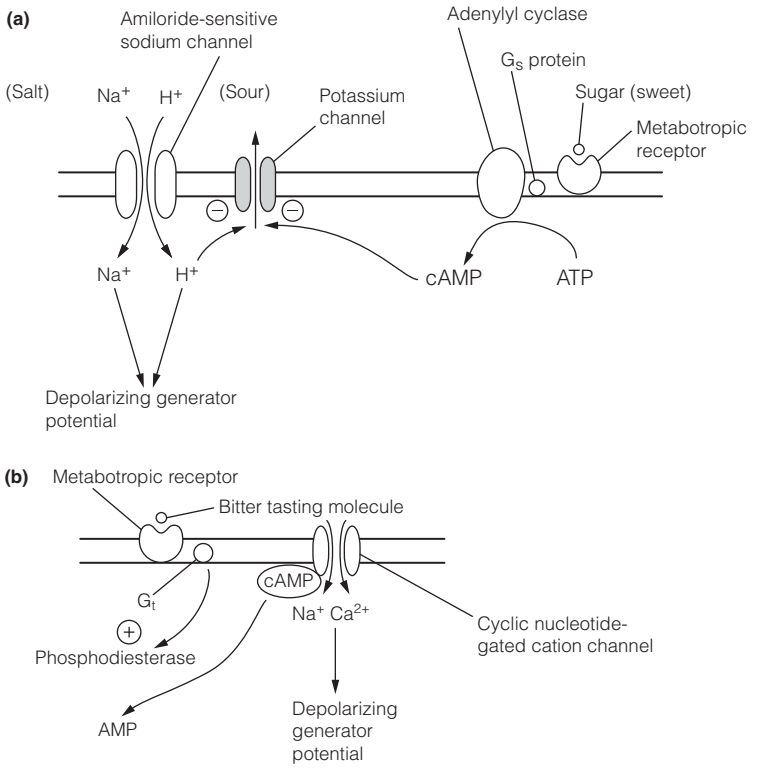


Fig. 2. Taste transduction: (a) salt, sour and sweet transduction mechanisms, note that an amiloride-sensitive Na^+ channel is implicated in both salt and sour taste; (b) one of several mechanisms involved in bitter transduction.

produces depolarization by closing a K^+ channel. Some compounds responsible for sweetness (e.g. artificial sweeteners) increase inositol trisphosphate (IP_3) levels and mobilize Ca^{2+} within receptor cells. How this triggers action potentials in these cells is not known.

Multiple pathways mediate bitter taste transduction (*Fig. 2b*). This reflects the wide diversity of molecules that are bitter flavored; divalent salts, alkaloids, some amino acids and some proteins. Divalent salts and quinine block K^+ channels and so produce depolarization by reducing an outward potassium current. In a mechanism with striking parallels to phototransduction, some bitter-tasting agents bind metabotropic receptors coupled to transducin (G_t) which activates a phosphodiesterase. This enzyme breaks down cAMP, lowering its concentration within the cytoplasm of the gustatory receptor cell. The final consequence of this is dissociation of cAMP from a cyclic-nucleotide-gated (CNG) cation conductance, allowing influx of Na^+ and Ca^{2+} and so depolarization. In addition bitter stimuli can activate phospholipase C, producing synthesis of IP_3 and mobilization of Ca^{2+} from internal stores.

The umami taste sensation produced by L-glutamate seems to involve metabotropic glutamate receptors of the mGluR4 subtype which are coupled via G_i proteins to the inhibition of adenylyl cyclase.

I4 Taste pathways

Key Notes

Anatomy of taste pathways

Gustatory primary afferent cell bodies lie in the ganglia of either the VII, IX, or X cranial nerves. Their axons terminate in the nucleus of the solitary tract (NST) in the medulla. While some NST cells project to the lateral hypothalamus for regulating autonomic responses to feeding, others project to the ipsilateral ventral posterior medial thalamic nucleus. This nucleus sends axons to the ipsilateral cortex taste area I, located adjacent to the somatosensory area for the tongue in the postcentral gyrus, which mediates conscious taste perception. Taste area II in the insula is thought to be involved in emotional responses to taste.

Gustatory coding

Gustatory afferents are broadly tuned. Those in the facial nerve respond best to salt or sweet stimuli, those in the glossopharyngeal respond preferentially to sour and bitter stimuli, while vagal afferents measure the extent to which the ionic concentration differs from extracellular fluid. The classic taste sensations do not correspond to separate labeled lines, neither are they topographically represented in the brain.

Related topic Retina (G3) Taste (I3)

Anatomy of taste pathways

The gustatory primary afferents of cranial nerves, VII, IX and X have their cell bodies in the **geniculate**, **petrosal** and **nodose** ganglia respectively. Their centrally directed axons end in the rostral portion of the **nucleus** of the **solitary tract** (NST) which lies in the dorsal medulla (*Fig. 1*). Taste primary afferents secrete glutamate and substance P.

Some NST cells project to the lateral hypothalamus which organizes autonomic responses to feeding. Gustatory neurons in the NST project via the central tegmental tract to the ipsilateral **ventroposterior medial nucleus** (VPM) of the thalamus, terminating on a population of small cells distinct from those receiving somatosensory input from the tongue or pharynx. These cells send their axons to the ipsilateral cortex. So, unlike most sensory pathways, that for taste is uncrossed. Taste area I (Brodmann's area 43) is located on the dorsal wall of the lateral sulcus at the junction with the insula and adjacent to the somatotopic mapping of the tongue. It is thought to be concerned with the conscious perception of taste. Taste area II is in the insula, a region of cortex buried deeply in the lateral sulcus, which may be concerned with the affective aspects of taste.

Gustatory coding

Afferents in the VII nerve commonly exhibit preferences for either salty or sweet stimuli, whereas most of those in the IX nerve, supplied by the posterior tongue, are tuned to acids (sour) or bitter stimuli. Many vagal (X) afferents respond to distilled water. These neurons have their lowest firing rate in 154 mM NaCl, and increase firing as salt concentration increases or decreases from this value. Vagal afferents thus appear to measure to what extent the pharyngeal contents differ in ionic concentration from extracellular fluid.

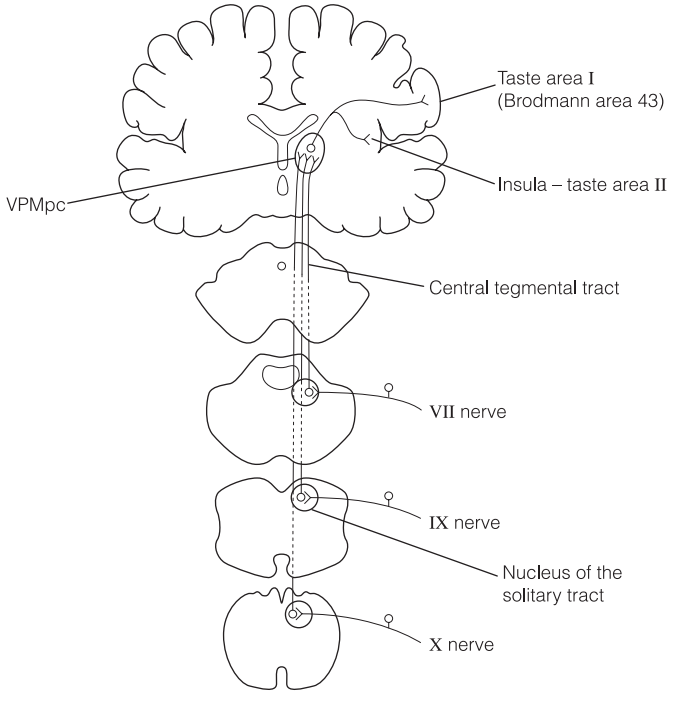


Fig. 1. Central gustatory pathways. VPMpc, parvocellular part of the ventral posterior medial nucleus of the thalamus.

The fact that gustatory neurons are generally quite non-specific argues against the existence of labeled lines corresponding to the classical taste sensations. Furthermore there is no topographical organization apparent in gustatory pathways. Hence distinctive taste sensations arise from neurons with opponent receptive fields that compare the outputs of differently tuned populations of afferents. This is analogous to how color vision arises from opponent processing that compares output from just three populations of cone photoreceptors.

J1 Nerve–muscle synapse

Key Notes

Neuromuscular junction

The neuromuscular junction (nmj) is the synapse between a motor neuron and a muscle fiber. Acetylcholine (ACh) released from the nerve terminal activates nicotinic cholinergic receptors (nAChR) in the postjunctional membrane, the endplate, causing it to depolarize. Secretion of ACh from a single vesicle causes a miniature endplate potential (mepp) of 0.4 mV. Release of ACh from many vesicles in response to the arrival of an action potential in the terminal causes summation of many mepps to give an endplate potential, a large depolarization sufficient to trigger muscle fiber action potentials and so contraction.

Neuromuscular blocking agents

Muscle relaxants, which block neuromuscular transmission, are used to paralyze skeletal muscles during surgery. They fall into two categories. Non-depolarizing drugs are competitive antagonists of nAChR, and their effects can be reversed by acetylcholinesterase inhibitors which increase ACh concentrations in the cleft. Depolarizing drugs are nAChR agonists that produce blockade by inactivating muscle calcium channels and nAChR desensitization.

Related topics

Ionotropic receptors (D1)
Elementary motor reflexes (J3)
Synaptogenesis and developmental plasticity (N5)

Neuromuscular junction

The axon of a motor neuron divides into a number of branches at the surface of the muscle fiber. Each branch ends in a bouton which forms a synapse with the muscle fiber, called a **neuromuscular junction** (**nmj**). The cleft of the nmj (*Fig. 1*) is about 50 nm across. The postjunctional membrane, the **endplate**, which is thrown into folds, has an extraordinarily high density of nicotinic acetylcholine receptors concentrated under the active zones where acetylcholine (ACh) is released. Overlying the endplate is a collagenous basement membrane (**basal lamina**) to which is bound acetylcholinesterase (AChE). Soluble forms of the same enzyme are also secreted into the cleft.

Nicotinic acetylcholine receptors (nAChR) are members of the ligand-gated ion channel superfamily of receptors and mediate fast transmission by ACh. The binding of ACh causing the opening of a cation channel, allowing Na^+ influx and K^+ efflux. The reversal potential for this current is close to 0 mV, so activating nAChR causes depolarization.

Spontaneous release of a single quantum of ACh at the nmj causes a 0.4 mV depolarization at the endplate called a **miniature endplate potential** (**mepp**). The arrival of an action potential at the motor nerve terminal triggers the release of 200–300 quanta which results in a massive depolarization, the summed effect of all the individual mepps, an **endplate potential** (**epp**) to about –20 mV. This greatly exceeds the threshold for activating voltage-dependent sodium channels in the muscle membrane, so the effect of the epp is to set up an action potential

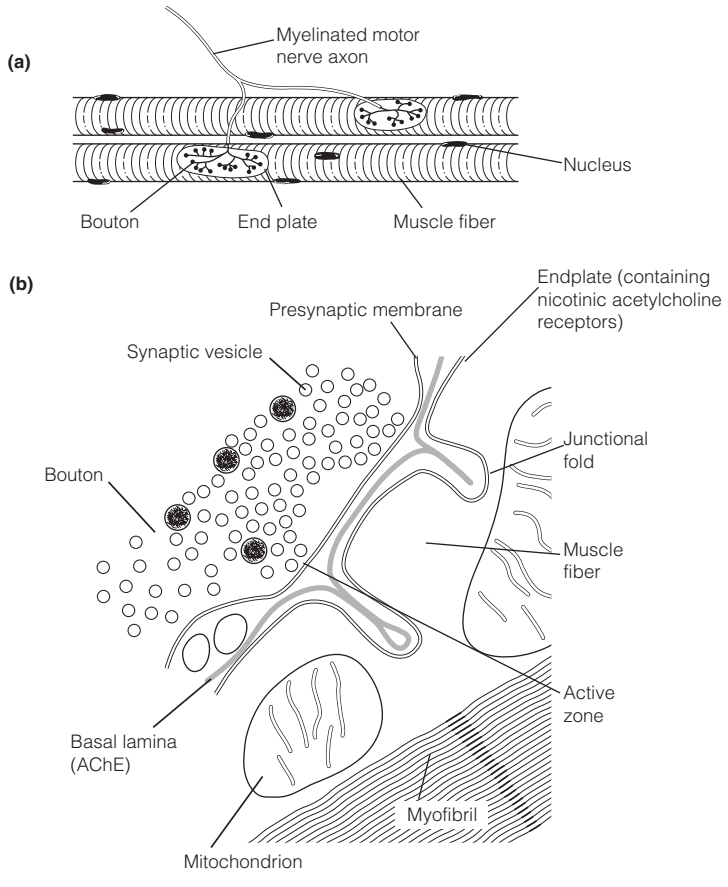


Fig. 1. The neuromuscular junction: (a) a motor neuron forming synapses on two muscle fibers (× 150); (b) a drawing of an electron micrograph of a neuromuscular junction.

which is propagated over the muscle fiber membrane. The neuromuscular junction is unique among vertebrate synapses in that firing of the motor neuron almost invariably results in the triggering of muscle action potentials.

The concentration of ACh reaches 1 mM in the nmj within about 200 μs of the arrival of an action potential at the motor nerve terminal but within a millisecond or so the ACh concentration has fallen back to baseline levels because of the high activity of AChE in the cleft. The enzyme hydrolyzes ACh to choline and acetate. Choline is taken back into the nerve terminal via a Na^+-dependent transporter.

Neuromuscular blocking agents

Blockade of transmission at the nmj is used during surgery to produce relaxation of skeletal muscle. They are effective within one minute of injection. Muscle relaxant drugs fall into two categories depending on their mode of action and all have a structural resemblance to ACh.

Non-depolarizing drugs, such as **tubocurarine**, are competitive antagonists of nAChR. The duration of action of these drugs ranges from 15–60 minutes. Their action can be rapidly reversed by AChE inhibitors, such as **neostigmine**

that cause a rise in the concentration of ACh which can then compete with the drug for the nicotinic receptor.

Depolarizing drugs, of which **succinylcholine** is the only agent of clinical importance, are nAChR agonists. Initially, agonist binding opens the nicotinic receptor channel causing persistent depolarization of the endplate. This first causes generalized disorganized contractions of muscles called **fasciculations**, and is followed by **flaccid paralysis** as muscle Ca^{2+} channels inactivate and the contraction mechanism fails as a result. This early stage in the action of depolarizing drugs (called **phase I** block) arises as a result of an ACh-like depolarization and so is augmented rather than reversed by AChE inhibitors. With continuing exposure **phase II** block occurs in which the nAChR either desensitizes, or suffers open channel blockade by the drug. Phase II block can be reversed by AChE inhibitors. Succinylcholine is rapidly hydrolyzed by circulating esterases, so its duration of action is only about 5 minutes.

J2 Motor units and motor pools

Key Notes

Motor units

A motor unit consists of a motor neuron together with all the muscle fibers it innervates, which ranges from six to a few thousand. In mammals, each muscle fiber gets input from just one motor neuron. An action potential in a motor neuron causes a twitch, a single contraction, in all the fibers it supplies. At high firing rates individual twitches summate to produce tetanus, a prolonged maximal contraction. There are three types of motor unit. Slow twitch (S) units drive type 1 muscle fibers that are adapted for aerobic metabolism and capable of sustaining low forces for very long periods. These dominate in postural muscles. Fast twitch fibers (divided into fatigue resistant (FR) and fast fatigue (FF)) innervate type 2 muscle fibers and can produce large forces rapidly, but only for short periods.

Motor pools

A motor pool is the set of motor neurons that innervates a single muscle. The force of contraction of a muscle is determined by the firing frequencies of individual motor neurons, and by the number of motor neurons in the pool that are firing. Larger forces are generated by recruiting an increasing number of motor units. Recruitment generally (but not always) follows the size principle in which smaller motor neurons come on line before larger ones. This gives the order S–FR–FF. Motor neurons are made hyperexcitable to glutamatergic input by monoaminergic neurons in the reticular system. This allows the generation of higher forces from the same input when arousal is high and during locomotion.

Motor unit disorders

Myasthenic diseases are those in which transmission at the neuromuscular junction is compromised. Most common is myasthenia gravis in which autoantibodies are made against nicotinic cholinergic receptors. The result is that the endplate becomes less sensitive to acetylcholine. Muscular dystrophies are disorders in which muscle fibers die and are replaced at abnormally high rates. The X-linked recessive Duchenne dystrophy is the commonest. Trauma that severs α motor neuron axons causes flaccid paralysis of the disconnected muscle.

Related topics

Nerve–muscle synapse (J1)
Elementary motor reflexes (J3)
Axon pathfinding (N4)
Synaptogenesis and developmental plasticity (N5)

Motor units

The final functional component of motor pathways is the **motor unit**. It consists of a motor neuron and the muscle fibers it innervates. In mammals each muscle fiber is supplied by only one motor neuron. However, each motor neuron synapses with anything from six to a few thousand muscle fibers within a single muscle. The size of a motor unit is related to the precision of motor control

required of a given muscle. Finely regulated muscles (e.g. extraocular eye muscles) consist of small motor units, less finely regulated muscles have larger ones. The fibers of a single unit are scattered widely throughout a muscle so no part of a muscle is controlled by just one motor unit.

A single action potential in the motor neuron causes a **twitch**, a single contraction, in all of the muscle fibers to which it is attached (*Fig. 1a*). The contraction and relaxation of muscle fibers is very much longer than the muscle action potential of about 3 ms. If a volley of action potentials is fired and there is insufficient time for the muscle to relax between successive impulses the twitches summate to increase the force which oscillates about a plateau value. This is called **unfused tetanus** (*Fig. 1b*). As the firing frequency increases the oscillations smooth out and the plateau reaches maximum force. This is **fused tetanus** (*Fig. 1c*).

Three types of motor unit can be distinguished by the firing behavior of their motor neurons and the properties of their muscle fibers.

The most numerous are the **slow twitch** (**S**) motor units which take about 50 ms to develop peak force and show little decline in force after even an hour of repetitive stimulation. The motor neurons of S units are small, have a low conduction velocity and quite long refractory periods because they contain a high density of Ca^{2+}-activated K^+ channels which cause a long after-hyperpolarization. This limits maximum firing frequencies to quite low rates, but fused tetanus is achieved at low frequencies (15–20 Hz). The **type 1** muscle fibers of S motor units are rich in mitochondria, have high activities of Krebs cycle enzymes (which permits them to be selectively stained histologically), which fits them for high rates of aerobic metabolism. Slow twitch motor units are capable of exerting low force for very long times. They form the bulk of the antigravity or postural muscles of the trunk and legs. These are **red** muscles because of their high myoglobin content.

By contrast, fast twitch units contract maximally in 5–10 ms but cannot sustain the contraction for very long. With repetitive stimuli, **fatigue resistant** (**FR**) units can sustain moderate force for 5 minutes or so before a steady decline sets in that takes many minutes. **Fast fatigue** (**FF**) motor units can achieve the greatest force of the three types, but with repetitive stimuli the force falls precipitously after 30 seconds or so. The motor neurons of both FR and FF units are large with high conduction velocities. For brief periods they fire at high rates but action potential volleys are of short duration, particularly for FF units. Fast twitch units contain **type 2** muscle fibers which require firing frequencies of 40–60 Hz to produce fused tetanus. Type 2 fibers come in two varieties that differ in their metabolism. The **type 2b** fibers of FF motor units are anerobic, which explains why these units fatigue so quickly. **Type 2a** fibers, found in FR units, are intermediate between types 1 and 2b in terms of metabolism. Both FR

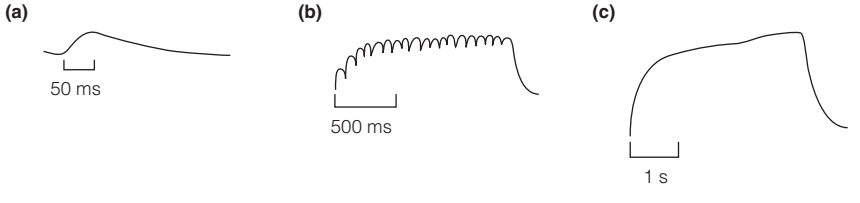


Fig. 1. Muscle fiber contraction: (a) single twitch; (b) unfused tetanus (firing frequency 12 Hz); (c) fused tetanus (30 Hz). Note the increase in force of contraction in going from (a) to (c).

and FF are adapted for producing rapid, large forces and so are found particularly in muscles involved in executing fast movements. Muscle in which fast twitch units predominate is **white muscle** because of its low myoglobin content.

In motor units the properties of the muscle fibers and motor neurons are matched for optimal performance. This is brought about because muscle fiber properties are determined by the motor neurons which innervate them. If type 1 muscle fibers are denervated and the axon of an FF unit sprouts to establish new connections with the denervated fibers, they acquire the characteristics of type 2b muscle fibers.

Motor pools

Motor neurons that innervate the same muscle form a common **motor pool**. Motor pools are topographically localized in motor nuclei of the brainstem and spinal cord. Spinal motor nuclei extend over several spinal segments. Axons of motor neurons leave the ventral horn of the spinal cord to run in the spinal nerve of the same spinal segment. Sorting of fibers destined for the same muscle but originating from different spinal segments occurs in the **nerve plexuses**. Axon collaterals of motor neurons ascend and descend a few segments to influence the behavior of other motor neurons in the same pool.

The force of contraction of a muscle is determined by the motor pool in two ways; the rate at which individual motor neurons fire and the number of motor neurons in the pool that are firing. Small increases in force are met mostly by increased firing rate, but larger contractions involve increasing the number of active motor units, a process called **recruitment**. This is done in an orderly manner. In general, the earliest units to be recruited are S, followed by FR and finally FF, an order determined by the **size principle**. Two effects are at work here. Firstly, the *size* of the motor neurons, secondly, how synapses onto them are organized.

How does size of a motor neuron determine the size principle? Small cells offer a bigger resistance to the flow of current than large ones. Ohm's law says that the relationship between the membrane voltage, V, and the current, I, flowing into a cell is given by:

$$V = IR$$

This means that a given current will produce a greater change in membrane voltage in a small cell (with large R) than it will in a big cell (with small R). Neurons in a motor pool are excited by common inputs. For a given sized synaptic current input into cells in the pool, the small cell body of an S motor neuron will have a bigger excitatory postsynaptic potential than the larger cell body of a fast twitch unit, because the S cell has the greater resistance (*Fig. 2*). This means that the weakest inputs recruit the S units, because they have the lowest threshold for synaptic activation. As the inputs to the pool get progressively stronger the other motor neurons are excited in turn.

The second effect determining the size principle is that the synaptic inputs to the three classes of motor unit are weighted in such a way that as input strength increases so motor units are recruited in the sequence S–FR–FF. However, recruitment does not always obey the size principle. In some instances synapses are arranged so that large motor neurons get more excitation than small ones. For example, in humans, cutaneous afferents preferentially excite fast twitch motor units.

Motor neurons are subject to tonic modulation by monoaminergic neurons (NA, 5-HT) that project in the reticulospinal tracts. This modulation greatly

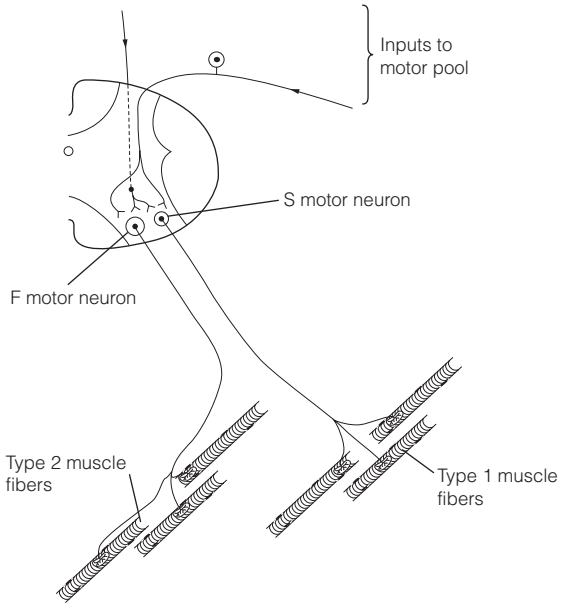


Fig. 2. The size principle in recruitment. The smaller slow twitch (S) motor neurons are recruited before fast twitch (F) motor neurons because they have a bigger excitatory postsynaptic potential in response to a given input.

enhances the response of the motor neurons to excitatory (glutamatergic) input. For example, normally S motor units fire tonically at 20–30 Hz, but with elevated monoamine input Ia afferents will produce firing rates in excess of 50 Hz. The firing rate of monoaminergic neurons increases with arousal so the modulation by NA allows the generation of higher forces from the same input when arousal is high, such as 'fight or flight' situations. Serotonin neurons fire during locomotion and their firing rate increases with the speed of locomotion, driving motor neurons to fire at a higher frequency for a given excitatory input than they otherwise would. Monoamines bring about motor neuron hyperexcitability largely by the enhancement of a depolarizing current through L-type calcium channels in their dendrites.

Motor unit disorders

Normally the endplate potential (epp) generated by motor neuron firing considerably exceeds the threshold for firing muscle fiber action potentials. The difference between the epp amplitude and the muscle firing threshold is the **safety margin** for nmj transmission. This is compromised in **myasthenic** diseases, which arise either presynaptically as a result of reduced ACh release or postsynaptically due to defects in nAChRs or AChE. In **Lambert–Eaton myasthenic syndrome** autoantibodies are produced against voltage-dependent Ca^{2+} channels in motor nerve terminals and this reduces ACh release. **Myasthenia gravis** is an example of a postsynaptic disorder in which autoantibodies are directed against the nAChR. This causes enhanced internalization and degradation of the receptor so the muscle fiber becomes less responsive to ACh.

Muscular dystrophies are a group of diseases characterized by an increased turnover of muscle fibers. The most common is **Duchenne dystrophy**, an X-linked recessive disorder in which a large cytoskeletal protein, **dystrophin**, is

abnormal or absent. Afflicted muscle fibers are weak and fatally damaged by normal mechanical forces. High rates of formation of new fibers from myocytes occur but they resemble fetal muscle fibers. They are small and do not propagate action potentials effectively. The number of motor units and their recruitment is normal in muscular dystrophies.

Peripheral nerve injury, which severs motor neuron axons, causes permanent **flaccid paralysis** and loss of stretch reflexes in the affected muscle, which consequently suffers disuse atrophy. The denervated muscle fibers synthesize large numbers of nicotinic cholinergic receptors which become inserted throughout the plasma membrane rather than being restricted to the endplate. These **extrajunctional receptors** cause muscle fibers to become exquisitely sensitive to acetylcholine (**denervation supersensitivity**) and they exhibit minute contractions, **fibrillations**, to the low concentration of circulating acetylcholine. Surgical nerve repair may effect recovery, if done with sufficient precision, since peripheral axons will regrow at a few millimeters per day to reinnervate muscle fibers.

J3 Elementary motor reflexes

Key Notes

Properties of reflexes

Reflexes are stereotyped responses to sensory input mediated by reflex arcs, which generally include a sensory neuron, a motor neuron and one or more interneurons. Reflexes may be monosynaptic, disynaptic or polysynaptic depending on whether their circuits have one, two or more than two central synapses. The elapsed time between a stimulus and a response, the reflex latency, is determined by the time for conduction in the circuit and the number of synapses. Increasing the stimulus strength or changing its location alters the character of a reflex. Reflexes are modified by experience in a variety of ways such as habituation, sensitization and conditioning.

Muscle spindle reflexes

Muscle spindle reflexes are monosynaptic reflexes which cause a muscle to contract when it is stretched. They control muscle length by negative feedback. The sensory component of the reflex is the muscle spindle, which contains small intrafusal fibers that lie in parallel with the ordinary extrafusal fibers. Afferents from the intrafusal fibers synapse with motor neurons supplying the same and synergistic muscles. Intrafusal fibers receive input from γ fusimotor neurons. Contraction of the intrafusal fiber by γ fusimotor activity keeps it taut over all muscle lengths, so that it is kept sensitive to stretch whatever the muscle length. During a movement muscles shorten. To allow this, muscle spindle reflexes must be overridden. This is achieved by the simultaneous firing of both the α motor neurons and γ fusimotor neurons so that both extrafusal and intrafusal fibers shorten together.

Inverse myotatic reflexes

Inverse myotatic reflexes control muscle tension by negative feedback. Golgi tendon organs located in tendons measure muscle tension. Their Ib sensory neurons synapse, via dedicated Ib inhibitory neurons, with α motor neurons which go to the same and synergistic muscles. An increase in muscle tension causes inhibition of the α motor neurons, reducing the force of contraction of the muscle, and hence the tension.

Control of muscle stiffness

In many normal situations it is not possible to maintain a constant muscle length and constant muscle tension at the same time. Hence rather than controlling muscle length and tension independently, CNS motor systems probably control muscle stiffness.

Related topics

Frequency coding (E2)
Nerve–muscle synapse (J1)
Motor units and motor pools (J2)
Spinal motor function (J4)

Properties of reflexes

The simplest operation the nervous system can execute is the **reflex**, which couples sensory input to motor output. A reflex is a stereotyped response to a particular stimulus. When it involves the autonomic nervous system it is an **autonomic reflex** and the effector is typically cardiac or smooth muscle or a gland. When it occurs in the somatic nervous system it is a **motor reflex** and the effector is skeletal muscle. They are mediated by specified neural circuits sometimes called **reflex arcs** (*Fig. 1*), which consist of a sensory neuron, a motor neuron and usually interneurons interposed between the two, which may be excitatory or inhibitory.

In humans, all but one reflex arc includes interneurons so they have several central synapses (three in *Fig. 1*) and are said to be **polysynaptic**. In the case that the reflex arc has only a single interneuron the reflex is **disynaptic**. The only example of a **monosynaptic** reflex in which there is no interneuron is the stretch reflex.

A sensory neuron will form synapses with several interneurons (or motor neurons in the case of a monosynaptic reflex). Usually the effect of afferent firing is to produce quite large epsps on a few neurons and more modest epsps in a bigger group, depending on the number of synapses. The connections made by several sensory neurons on interneurons overlap and this provides for the possibility of integration. This integration is often non-linear; that is the excitation of the motor neuron can be bigger than the sum of the individual inputs, **facilitation**, or dominated by one input and hence little affected by additional ones, **occlusion**.

The time between the stimulus and response is called the **reflex latency** or **reflex time**. It results chiefly from the conduction time along afferent and efferent fibers, but also includes the time taken for sensory transduction and for activation of the effector (muscle or gland). A small interval is taken up by the **synaptic delay**, usually between 0.5 and 1 ms. When conduction time is taken into account reflex latencies reflect the number of central synapses.

Increasing the stimulus intensity will change the amplitude of the reflex (e.g. the amount by which a limb moves) but may also alter the form of the response

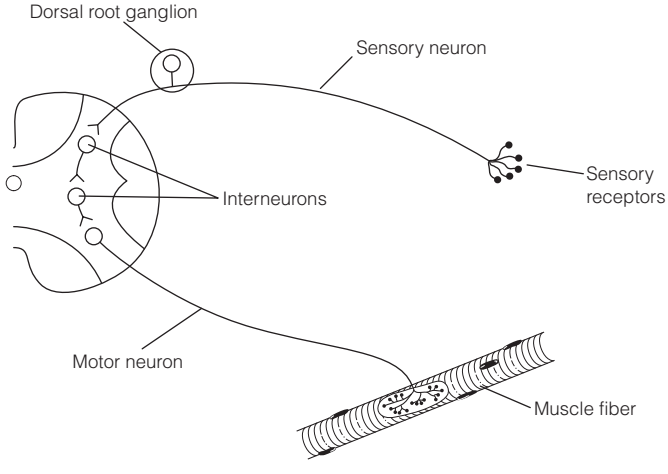


Fig. 1. A polysynaptic spinal motor reflex arc.

by recruiting additional muscles, this is called **irradiation**. The exact form of the reflex response depends on precisely where the stimulus is applied and so which afferents are excited. This is called the **local sign**.

Probably all reflexes can be modified by experience. The attenuation of a reflex by the repeated application of a constant innocuous stimulus is **habituation**. It is caused by synaptic depression. Any change to the stimulus (e.g. in its intensity) causes **dishabituation** in which the reflex returns to its baseline state. By contrast, repeated application of a noxious stimulus can enhance a reflex, by a decrease in latency, increased amplitude or irradiation. This is known as **sensitization** and results from increased transmitter release. Both habituation and sensitization are examples of non-associative learning because only one stimulus is involved. Some reflexes are capable of the more complicated associative learning, in which a response occurs if two stimuli are paired in time. These are **conditioned reflexes**.

Muscle spindle reflexes

The most elementary modulation of motor unit output is made by sensory input from the **muscle spindles** which measure the length and rate of change of length (velocity) of the muscle. Any attempt to stretch the muscle rapidly, for example by suddenly loading it, is met by contraction. This is the **muscle spindle reflex** (**stretch reflex**, **myotatic reflex**) and is a negative feedback mechanism which defends a constant muscle length in the face of external forces which act to perturb it. A stretch reflex can be elicited from any skeletal muscle by sharply tapping its tendon. The resultant stretch causes the muscle to contract. The stretch reflex is most easily demonstrated by tapping the patellar ligament between its insertion into the tibia and the kneecap, causing the contraction of the quadriceps femoris, the powerful group of extensor muscles on the front of the thigh (*Fig. 2*).

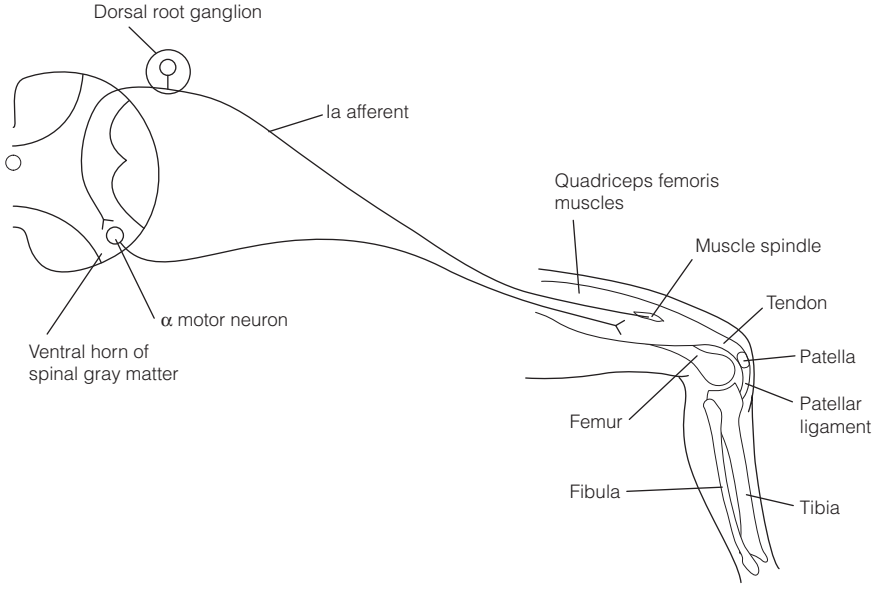


Fig. 2. Basic circuit of a stretch reflex. Striking the patellar ligament excites a few hundred Ia afferents.

The sensory side of the stretch reflex consists of the **muscle spindle** and its afferents. Muscle spindles lie in parallel with the standard **extrafusal fibers** so any force acting on the whole muscle acts in the same way on the spindle. Each muscle spindle is a fluid-filled capsule of connective tissue, 4–10 mm long and 100 μm in diameter, containing about seven modified muscle fibers called **intrafusal fibers** (*Fig. 3*). Intrafusal fibers have contractile ends but their central regions are non-contractile. There are two types of intrafusal fiber, nuclear bag and nuclear chain.

Nuclear bag fibers (**b**) are swollen at their center, where their nuclei are clustered, and are innervated by large-diameter myelinated (Ia) primary afferents,

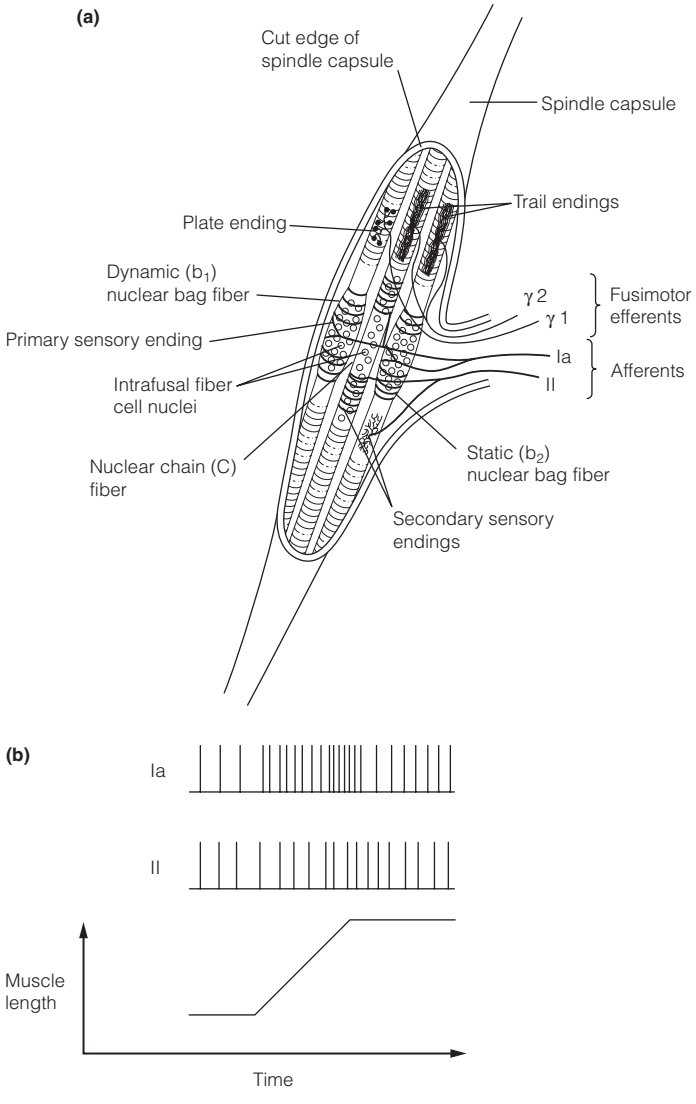


Fig. 3. Muscle spindles: (a) a spindle opened to show intrafusal fibers and their innervation. A spindle normally contains one b_1, one b_2 and several c fibers; (b) responses of Ia and II afferents to muscle stretch.

the ends of which spiral around the central region of the fiber. There are two sorts of nuclear bag fibers which can be recognized by whether, in addition to primary afferent innervation they also receive secondary, group II, myelinated afferents. Those that do not are **dynamic** ($\mathbf{b}_1$), those that do are **static** ($\mathbf{b}_2$).

Primary afferents show dynamic responses, responding to the rate of change of length (velocity). This is because of the properties of the dynamic nuclear bag fiber. When stretched the central region elongates causing the Ia afferent to fire a volley of action potentials. Subsequently however the poles of the fiber elongate slowly permitting the central region to creep back to a shorter length so that the firing rate of the Ia afferent drops off. Primary afferents also show static responses, signaling muscle length, by virtue of their innervation of static (b_2) nuclear bag fibers which are stiffer than the dynamic fibers and hence elongate in proportion to muscle stretch.

Nuclear chain fibers (**c**) are of uniform diameter, are about half the size of b fibers and their central region contains a line (chain) of nuclei. They are innervated by primary and secondary afferents and being stiff (like the b_2 fibers) these afferents respond to muscle length. Typically a spindle will contain one b_1, one b_2 and 3–5 c intrafusal fibers.

The majority of Ia spinal afferents form synapses on **homonymous** motor neurons, i.e. motor neurons going to the same muscle. However about 40% make synapses with motor neurons which go to synergistic muscles. For example, the quadriceps femoris consists of four muscles that act synergistically (they are all leg extensors). Afferents from spindles in any one of them will establish connections with its own motor pool and the pools of the other three muscles.

A stretch reflex has two components. The **phasic** component is that seen by tapping the tendon of a muscle. It occurs rapidly, is brief and occurs because of the dynamic activity of the Ia afferents. The **tonic** component is the much more sustained contraction brought about by the static activity of the Ia afferents and the secondary, group II afferents. This component is particularly important in maintaining posture. Standing in a moving vehicle, for example, the muscles in the legs and trunk that are stretched by the swaying will be contracted, so keeping the body upright. A sudden jolt will, of course, also trigger the phasic component.

Cell size is bimodally distributed in motor pools. The neurons which drive the extrafusal fibers to produce muscle contraction are Aα class with a cell body diameter averaging 80 μm, usually referred to as **a motor neurons**. In addition there is a population of smaller cells belonging to the Aγ class, called **γ motor neurons**, axons of which (**fusimotor fibers**) go to the muscle spindles. All intrafusal fibers have their contractile ends innervated by these γ motor neurons. Contraction of the ends of the intrafusal fibers keeps the central region taut so that it can respond to muscle stretch. So, one purpose of γ efferent discharge is to maintain the sensitivity of the muscle spindle to changes in length over a wide range of lengths. Without it, muscle contraction would cause the intrafusal fibers to slacken and fail to respond to stretch.

There are two categories of γ motor neuron that can be activated independently by the CNS; γ_1 (dynamic) innervate b_1 while γ_2 (static) innervate b_2 and c fibers. Stimulation of γ_1 fibers increases the sensitivity of the b_1 fibers, so that the primary afferent firing rate in response to rapid stretch is higher. Stimulation of γ_2 fibers enhances firing of secondary afferents in response to constant stretch. In both cases the γ efferents are increasing the **gain** (sensitivity) of the spindle.

Firing rates of γ-efferents are raised when performing movements that are particularly complex.

Stretch reflexes must be overridden to allow the execution of a movement since the muscle must contract isotonically and shorten. This is achieved by descending motor pathways exciting both α and γ motor neurons at the same time. This is called **coactivation**. It makes the extrafusal and intrafusal fibers shorten together in such a way that the intrafusal fibers are always sufficiently taut to respond to stretch.

Clinically, stretch reflexes are rather misleadingly termed **tendon reflexes**. A neurological examination includes eliciting stretch reflexes from several muscle groups throughout the body since abnormal or absent reflexes can reveal the level of any damage to the nervous system. An absent reflex may signify a lesion anywhere in the reflex arc: sensory or motor neuron or CNS. The stretch reflex can be studied by electrically stimulating the nerve which supplies a muscle and recording the electrical activity of the muscle with electrodes placed on the skin overlying the muscle or needle electrodes inserted into the muscle. Recording muscle activity in this way is **electromyography** (**EMG**). This procedure is most easily done by stimulating the **tibial nerve** at the back of the knee and recording from the **gastrocnemius** and **soleus** (calf) muscles. Ia fibers have the lowest threshold of any nerve fibers, so low-intensity stimulation of the nerve elicits the stretch reflex which can be seen as the H (Hoffman) wave on the EMG and occurs about 30 ms after the stimulation. This is the reflex latency. Progressively increasing the stimulus strength eventually results in exciting the α motor neurons, as well as the Ia afferent, with the appearance of the M (motor neuron) wave with a latency of only 5–10 ms. In the clinic this procedure can show whether an absent reflex is due to loss of sensory or motor function.

Inverse myotatic reflexes

Located in the tendons, in series with muscle fibers are **Golgi tendon organs** (**GTO**) which measure muscle tension. Increases in muscle tension activate a negative feedback reflex, the **inverse myotatic** (**Golgi tendon**) **reflex**, which opposes the increases in tension. It is brought about by GTO input activating inhibitory interneurons that synapse with α motor neurons supplying the muscle (*Fig. 4*).

The GTO is composed of the collagen fibers which join muscle fibers to tendons, interwoven through which are axon branches of a group Ib afferent neuron. The increased tension that occurs with muscle contraction stretches the collagen fibers, distorting the terminals of the Ib afferent which fires. Individual Ib afferents respond statically, reflecting the level of tension, to the activation of a single motor unit. GTOs do not measure the average tension of the muscle, firstly because only the tension developed by the muscle fibers inserted into its particular region of the tendon is measured, secondly because less than 1% of fibers in motor units are coupled to GTOs. The Ib afferents enter the spinal cord to synapse in the intermediate zone (Rexed laminae VI–VIII) on inhibitory neurons, which then synapse with motor neurons of the homonymous and synergistic muscles. These inhibitory neurons are all specific to this disynaptic reflex pathway and so are designated **Ib inhibitory neurons** (IbINs). Inhibition of homonymous and synergistic muscle by GTOs is augmented by inputs from Ia spindle afferents, joint afferents and cutaneous mechanoreceptor afferents onto IbINs. The functional importance of these connections is not clear. Descending motor systems may either excite or inhibit IbINs.

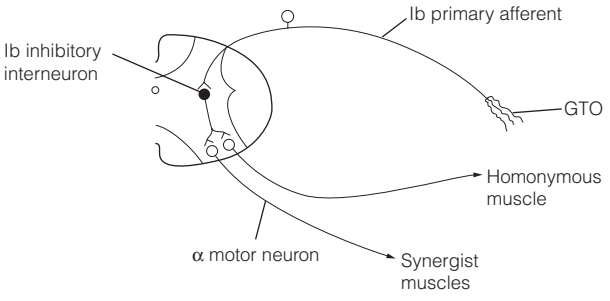


Fig. 4. Circuitry of the inverse myotatic reflex. An increase in muscle tension causes the Golgi tendon organ (GTO) Ib afferent to fire at a greater rate.

Control of muscle stiffness

The muscle spindle reflex for maintaining muscle length and the GTO reflex for keeping constant tension often work in opposition. When the loading on a muscle is altered either the muscle is stretched, in which case its *passive* tension increases (just as it does in a stretched rubber band), or it must contract isometrically to maintain a constant length, in which case the *active* tension rises due to contraction. It is impossible for both length and tension to be held constant at the same time. This suggests that overall, motor systems do not control length or tension independently, but **muscle stiffness**, k, a constant that describes how much the length of a muscle is altered, ΔL, by a change in load, ΔF.

$$k = \Delta F / \Delta L$$

It appears that the muscle spindle and GTO reflexes between them compensate for the complicated way in which muscle stiffness changes over the normal working range of muscle. This allows the organization of supraspinal motor systems to be simpler.

J4 Spinal motor function

Key Notes

Elements of spinal cord motor function

Most neurons in the spinal cord are interneurons, which implies that considerable processing of sensory input takes place before it influences motor output. Much of the circuitry of the spinal cord is either organized for the execution of motor reflexes, or into central pattern generators, which generate the cycles of muscle activity involved in locomotion. Three types of inhibition are important for spinal cord motor function.

Reciprocal inhibition

Reciprocal inhibition occurs between mutually antagonistic muscle motor neurons. Muscle spindle Ia sensory neurons synapses with Ia inhibitory interneurons which supply the motor neurons of antagonist muscles. This circuit allows for relaxation of antagonist muscle while the agonist is contracting. Reciprocal inhibition is important in allowing movement that needs alternating activity in agonist and antagonist muscles.

Presynaptic inhibition

Primary afferent inhibits the terminals of other primary afferent via inhibitory GABAergic interneurons that make axoaxonal synapses. The effect of GABA is to depolarize the terminal inhibiting its release of transmitter in response to an invading action potential. Presynaptic inhibition is organized so that flexor afferents inhibit extensor afferents and *vice versa*. It is modified by descending motor systems.

Recurrent inhibition

Glycinergic Renshaw cells in the ventral horn activated by motor neurons produce inhibition of neighboring synergistic α motor neurons. This recurrent inhibition is the motor equivalent of lateral inhibition in sensory systems and sculpts economic movements from a broader range of motor neuron activity.

Flexor reflexes

Flexor reflex afferents (FRAs) elicit flexor reflexes in the ipsilateral limb, and other reflexes in the contralateral limb. The neural circuitry for these reflexes is recruited during normal limb movements, which are continuously modified by peripheral sensory input via FRAs. Flexor reflexes generated by nociceptor input enable a limb to be withdrawn from a noxious stimulus.

Central pattern generators

Locomotion, movement from place to place, involves alternate flexion and extension of limbs, which are phased in specific ways depending on the manner of locomotion (e.g. walking or running). The basic rhythms of locomotion are produced by central pattern generators (CPGs), networks of spinal interneurons. Each CPG acts as an oscillator, driving a limb to flex then extend alternately. CPG activity is regulated by a midbrain locomotor region which projects to the spinal cord via the reticulospinal tract.

Related topics

Information representation by neurons (E1)
Elementary motor reflexes (J3)
Brainstem postural reflexes (J5)

Elements of spinal cord motor function

About twice as many sensory fibers (half of them unmyelinated C fibers) enter the dorsal roots of the spinal cord, as there are motor neurons in the ventral horn. However, motor neurons make up less than 2% of the total number of neurons in the spinal cord, most of which are interneurons. This shows that massive processing of sensory input occurs before it influences motor neurons.

Organization of movement by the spinal cord depends on reflex circuits and central pattern generators (CPGs). Motor reflexes should not be thought of as independent behaviors but as elements which allow descending motor control of muscles to be continually modified on the basis of proprioceptor input from muscles and joints, and input from the skin, enabling the smooth execution of the right movement for the current situation. With the exception of a few protective reflexes such as the flexion withdrawal reflex, which serves to remove a limb from a noxious stimulus, motor reflexes are not seen in isolation under physiological conditions.

Locomotion involves cycles of activity in which muscles groups are made to contract in a precisely timed sequence. This requires neural networks that can generate the required rhythmic output. These networks, thought to be autonomous, though modifiable by reflexes and activated by supraspinal influences, are called central pattern generators (CPGs). The presence of CPGs can be inferred by experiment in a wide variety of vertebrates, including primates and humans, and putative circuits have been modeled using computer simulations.

Three types of inhibition contribute to spinal cord function, reciprocal, presynaptic and recurrent inhibition.

Reciprocal inhibition

Axon collaterals of Ia afferents from muscle spindles synapse with **Ia inhibitory interneurons** in lamina VII (IaINs) that use glycine as a transmitter and project to motor neurons of antagonist muscles. This disynaptic circuit allows antagonist muscles to be relaxed during agonist contraction (*Fig. 1*). The inhibition of mutually antagonistic muscle motor neurons is called **reciprocal inhibition**.

IaIN activity mediating reciprocal inhibition is modified by descending motor pathways (corticospinal, rubrospinal and vestibulospinal) and locomotor networks in the cord. Two reasons for this are:

- To facilitate rapid movements. Because muscle contractions are so long lasting, muscles follow faithfully only slowly changing neural input. Rapid changes in motor neuron firing cannot be translated into corresponding alterations in muscle tension. Hence to produce fast fluctuations in tension the motor system alternates contraction in agonist and antagonist muscles. This is helped by reciprocal inhibition (*Fig. 2*).

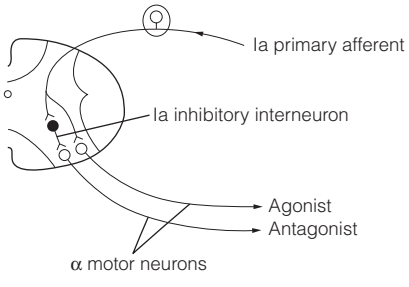


Fig. 1. A disynaptic reflex for reciprocal inhibition of antagonist muscle.

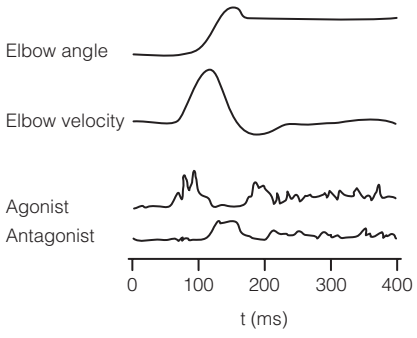


Fig. 2. Pattern of activation of agonist (biceps) and antagonist (triceps) muscles to facilitate rapid movement (elbow flexion).

- To adjust muscle stiffness so that it is appropriate for the load. The stiffness of a joint is increased by co-contraction of muscles with opposing actions at the joint. For example, co-contraction of the biceps and triceps increases stiffness of the elbow. Provided one muscle contracts move powerfully than the other the joint moves. Co-contraction stabilizes a joint, it provides better control when loads change unexpectedly because a given difference between expected and actual load will have a smaller effect on limb trajectory if a joint is stiffer. Co-contraction requires suppression of reciprocal inhibition.

Presynaptic inhibition

Reflexes dependent on sensory input via Ia, Ib and II afferents, can be modified by presynaptic inhibition from inhibitory GABAergic interneurons in the spinal cord. These interneurons make axoaxonal synapses on the afferent terminals. GABA secreted at these synapses acts on $GABA_A$ receptors to bring about *depolarization* because in these sensory neurons the membrane potential is more negative than the reversal potential of the chloride current through the $GABA_A$ receptor channels. The effect of this **primary afferent depolarization (PAD)** is inhibitory in that when an action potential sweeps into the terminal its amplitude gets smaller, less Ca^{2+} influx occurs, so transmitter release from the terminal is reduced.

Presynaptic inhibition of Ia terminals is organized on a reciprocal basis in which flexor afferents inhibit extensor afferents and *vice versa*. In humans this presynaptic inhibition causes a more prolonged reciprocal inhibition between antagonist muscles. It can be modified by descending motor pathways which project to one or other of the interneurons in the circuit. Normally at the start of a movement presynaptic inhibition to Ia terminals going to agonist motor neurons is reduced whereas inhibition of antagonist Ia terminals is increased. The effect of this is that spindle activity in the agonist is reinforced whereas the antagonist muscle myotatic reflex is dampened. This makes sense; agonist action lengthens the antagonist muscle and this would provoke it to contract were its stretch reflex not suppressed.

Similar but separate circuitry is involved in presynaptic inhibition of Ib and II afferents, allowing control of tension and tonic length signals independently of phasic length signals.

Recurrent inhibition

A population of interneurons called **Renshaw cells** in the ventral horn is activated by axon collaterals of α motor neurons and projects to neighboring,

homonymous and synergistic a motor neurons. These cells fire high-frequency bursts of action potentials that produce fast large ipsps. The effect of the inhibition is to silence motor neurons excited weakly and firing at low rates, and to dampen the firing frequency of strongly excited motor neurons. This is very like the lateral inhibition seen in sensory systems. It enhances contrast, and so enables economical movements. Renshaw cells are glycinergic and blockade of glycine receptors with **strychnine** results in convulsions due to the failure of recurrent inhibition. **Tetanus**, caused by infection with the organism *Clostridium tetani*, produces the same effect because its toxin blocks glycine release from Renshaw cell terminals.

Flexor reflexes

A variety of afferents, including group II and III muscle afferents, joint afferent and skin mechanoreceptor and nociceptor afferents elicit flexor reflexes in the ipsilateral limb and so are known as **flexor reflex afferents (FRAs)**. These same afferents trigger an extension of the contralateral limb, the **crossed extensor reflex** or alternative reflex pathways. Different types of FRAs are connected to specific subsets of interneurons so the reflexes they excite differ in form and timing. Many normal limb movements consist of either flexion or alternate flexion and extension and are brought about by the actions of supraspinal motor systems on the interneurons targeted by the flexor reflex afferents. In one model to account for this these interneurons are organized into sets called **half centers**, between which reciprocal inhibitory connections exist. A flexor half center gets inputs from ipsilateral FRAs and excites flexor motor neurons, while an extensor half center is activated by contralateral FRAs driving extensor motor neurons (*Fig. 3*). These half centers are also components of central pattern generators (see below).

For the execution of a particular movement, descending motor axons are activated which project to the specific set of FRA interneurons that bring about the movement. The reciprocal connections ensure that the alternative set is inhibited. As the movement proceeds FRA input from muscles, joints and skin reinforces the movement and allows it to be fine tuned.

Although flexor reflexes are usually recruited as elements of normal movement, the withdrawal flexor reflex, triggered by Aδ or C fiber (group IV) nociceptors is quite distinct in character. It overrides ongoing movement, involves flexor muscles throughout the limb so the response is dramatic, and it is long lasting. It is protective.

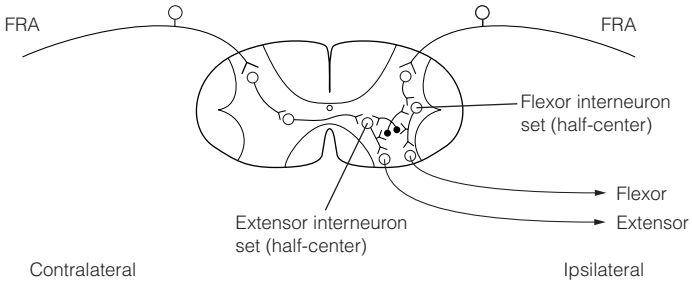


Fig. 3. Organization of flexor reflex afferent (FRA) interneurons. Excitatory neuron cell bodies are open circles, inhibitory interneuron cell bodies are filled circles.

Central pattern generators

Locomotion is movement from place to place. There are numerous modes of locomotion (e.g. stepping, swimming, flying) but all depend on cycles of muscle activity, alternate flexion and extension of each limb, though the **gait** or manner of locomotion (e.g. walking or running) depends on the speed. The gait adopted is the one which minimizes the energy expenditure for that speed.

In human stepping each leg has a **stance phase**, when extensors are most active, and a **swing phase**, when the flexors are the most active. The same sequence in the opposite leg is out of phase, though during walking there is a brief overlap of the stance phase in both legs. As the speed increases the stance phase shortens until there is no overlap and the switch to running occurs.

The basic rhythms of locomotor activity are produced by **central pattern generators** (**CPGs**), networks of interneurons in the spinal cord that generate the precisely timed sequences of α motor neuron activation without the need for sensory input. Each limb has an array of CPGs. Each CPG is an oscillator with two half-centers, one driving flexors, the other driving extensors, with reciprocal connections between them. Each half-center produces rhythmic bursts of action potentials that are terminated in a time and manner determined by the intrinsic excitable properties of its constituent neurons. The cessation of firing of one half-center releases its opposite number from reciprocal inhibition, allowing it to fire a burst. In this way burst firing alternates between the two half-centers.

CPGs have been studied in the lamprey, a primitive vertebrate and it is likely that mammalian CPGs use similar principles (*Fig. 4*). Depolarization of excitatory (E) cells in one half-center activates their NMDA receptors. The resulting calcium influx prolongs the depolarization so that the E cell fires a burst of action potentials. The burst firing of the E cells stimulates motor neurons. Two features of the circuit allow it to flip-flop between bursting of first one half cell then the other. Firstly, the E cells stimulate L cells which inhibit the I neurons responsible for the reciprocal inhibition. This disinhibits the opposite half-center. Secondly, the E cell burst ends because the calcium activates K_{Ca} channels, allowing K^+ efflux and hyperpolarization.

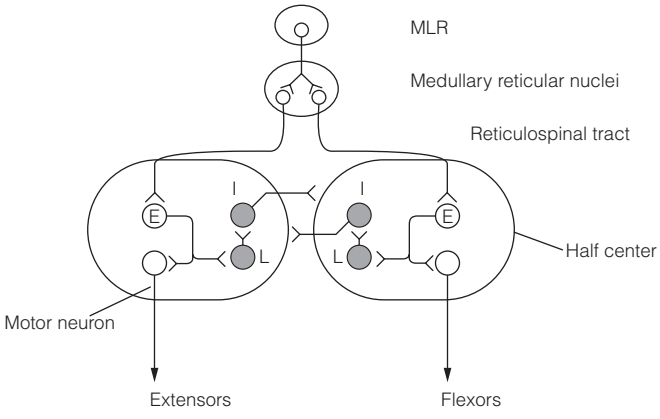


Fig. 4. A simplified hypothetical model of a central pattern generator (CPG) based on studies of the lamprey. Each of the neuron symbols represents several cells. The excitatory (E) cells use glutamate and show burst firing in response to supraspinal input. The inhibitory (filled symbol) neurons are glycinergic; contralaterally projecting interneuron (I) inhibits the opposite half center; L, lateral interneuron. MLR, mesencephalic locomotor region.

Locomotor activity is initiated by activity in the **mesencephalic locomotor region** (MLR) which projects to reticular nuclei in the medulla (*Fig. 4*). Axons from here run in the **reticulospinal** tracts to the spinal cord. These reticular nuclei are excitatory, releasing glutamate to produce a large depolarization of the CPG neurons, which then produce oscillating output for as long as the MLR input continues. CPGs are interconnected so that timing of events in all limbs is coordinated. The basic locomotor rhythms of the CPGs are extensively modified by the supraspinal motor systems.

J5 Brainstem postural reflexes

Key Notes

Postural mechanisms

Postural mechanisms stabilize the body against forces which shift the center of mass, including limb movements. Posture is maintained largely by the action of antigravity muscles which include extensors of the back and legs (and in humans, arm flexors). Feedforward motor commands from motor cortex and cerebellum allow postural adjustments needed for an anticipated movement. Postural reflexes, organized by the brainstem in response to vestibular, proprioceptor and visual input use negative feedback to correct for unanticipated perturbations.

Vestibular (labyrinthine) reflexes

Vestibular input is used to maintain the orientation of the head in space in the face of tilting or rotation of the head and body as a unit. These vestibular reflexes achieve this by acting on neck muscles (vestibulocollic reflexes) or limb muscles (vestibulospinal reflexes) giving ipsilateral extension and contralateral flexion.

Neck reflexes

When the head moves with respect to the trunk, neck muscle stretch produces reflex contraction of neck muscles (cervicollic reflexes) to return the head to its previous position, and of limb muscles (cervicospinal reflexes) to produce ipsilateral flexion and contralateral extension. Cervicocollic and vestibulocollic reflexes act synergistically, cervicospinal and vestibulospinal reflexes are antagonistic.

Postural reflex pathways

The vestibular and reticular nuclei integrate vestibular and proprioceptor (e.g. muscle spindle) input and execute postural reflexes by way of the vestibulospinal and reticulospinal tracts of the medial motor system. Inputs from the cerebral cortex to reticular nuclei allow postural adjustments to be made during voluntary movements.

Related topics

Balance (F5)
Cortical control of voluntary movement (K1)
Spinal motor function (J4)

Postural mechanisms

Postural mechanisms prevent the body from being destabilized by forces, including gravity, which tend to shift the center of mass. They also maintain the center of mass during limb movements. Muscles either oppose or assist gravity when contracting; those that oppose are described as **antigravity muscles**. Many antigravity muscles, such as the leg extensors and the short deep extensor muscles of the back (axial muscles), are involved in maintaining posture. In humans, the flexor muscles of the arms are also antigravity muscles. Since antigravity muscles are generally more powerful than muscles assisted by gravity, in human limbs the strongest muscles are the leg extensors and arm flexors.

Motor commands from the forebrain and cerebellum include those for the postural adjustments necessary during the execution of a movement. For example, if a standing dancer abducts her leg she must shift her center of mass above the other leg to avoid falling over. This needs *feedforward* postural adjustments to trunk and arms which anticipate the unbalancing forces that act during the movement. The adjustments (**postural set**) depend precisely on the initial position and nature of the intended movement and they must be learnt.

However, postural adjustments often have to be made to compensate for *unpredicted* disturbances in body position and movement. This cannot be done by feedforward since, by definition, the disturbance is not known ahead of time. So, these adjustments are made by **postural reflexes**, organized by the brainstem, which are *negative feedback* mechanisms (*Fig. 1*). The sensory input to the reflex circuitry is from three sources:

- vestibular, from the otolith organs;
- proprioceptor, from muscle spindles, Golgi tendon organs and joint receptors;
- visual, from the superior colliculus.

These inputs are highly integrated to recruit the required sequence of corrective muscle activity.

The exact nature of postural reflex adjustments made by humans depends on context, i.e. the initial position of the body and the size and direction of the destabilizing force. Swaying, produced by sudden displacement of the surface on which a person is standing, will activate quite different sets of muscles depending on the direction of sway, but in general distal muscles are excited before proximal ones, with most movement occurring at the ankle joint. Rotation or tilt of the surface, however, results in bending at the hips. *In extremis* postural reflexes try to maintain the center of gravity to prevent falling, or to put the limbs in a position to brace against falling.

Several distinct postural reflexes can be seen in animals by surgically transecting the brainstem (**decerebration**) and in humans who have suffered severe brain damage. These reflexes cannot *easily* be elicited in isolation in healthy behaving humans because motor functions are normally so highly integrated. They may be seen in newborn infants, in whom motor systems are immature.

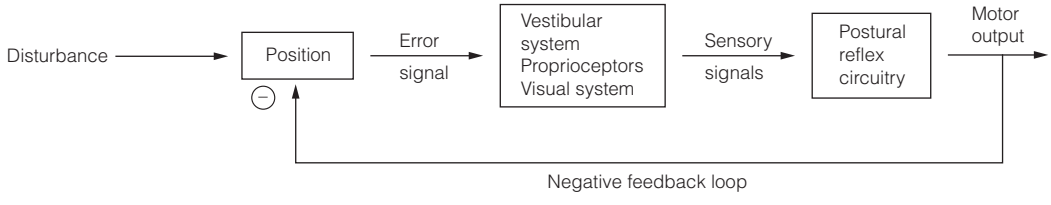


Fig. 1. Postural reflex negative feedback. The circuitry produces a motor output which reduces the mismatch between the desired position of the body and its actual position. The error is detected by the sensory systems which feed information into the reflex circuitry.

Vestibular (labyrinthine) reflexes

Sensory input from otolith organs and semicircular ducts is used to stabilize the orientation of the head in space. Any tilting or rotation of the head and body as a unit activates motor neurons to muscles that maintain the head vertical with respect to gravity. These mainly tonic reflexes have a latency of about 40–200 ms. Those which activate motor neurons to neck muscles are called vestibulocollic reflexes, those to limb muscle motor neurons are vestibulospinal reflexes.

Vestibulocollic reflexes act on the neck muscles to keep the head upright. If the body sways forward the neck extensors contract bringing the head up. If the body sways backwards, neck flexors are activated. **Vestibulospinal reflexes** act on limb muscles. They trigger contraction of arm extensor muscles and leg flexor muscles when falling, to reduce the impact of landing. Swaying sideways triggers extension of the ipsilateral limbs to brace against further tilt in that direction, and contralateral flexion.

Neck reflexes

Turning the head *relative to* the body excites spindles in neck muscles and afferents from the cervical vertebral joints. This evokes reflex contractions of neck muscles (cervicocollic reflexes) and limb muscles (cervicospinal reflexes) with both phasic and tonic components. **Cervicocollic reflexes** contract neck muscles that are stretched and so act to reorientate the head on the body. Cervicocollic and vestibulocollic reflexes are synergistic. **Cervicospinal reflexes** cause contraction of limb muscles in response to *rapid* head movement. In standing humans a force which throws the head backwards on the trunk activates all the limb extensors, whereas a force throwing the head forwards activates all limb flexors. Cervicospinal and vestibulospinal reflexes are antagonistic. If the head and trunk are tilted *as one* to the left, the vestibulospinal reflex causes left arm extension. But if the trunk *alone* is passively tilted to the left (with the head remaining fixed in relation to space) the cervicospinal reflex will flex the left arm. In the more usual situation that the head is tilted while the trunk remains stationary, these opposing reflexes cancel out.

Postural reflex pathways

Supraspinal descending control of movement uses two sets of pathways. Those mediating postural reflexes, which go from brainstem to spinal cord, are collec-

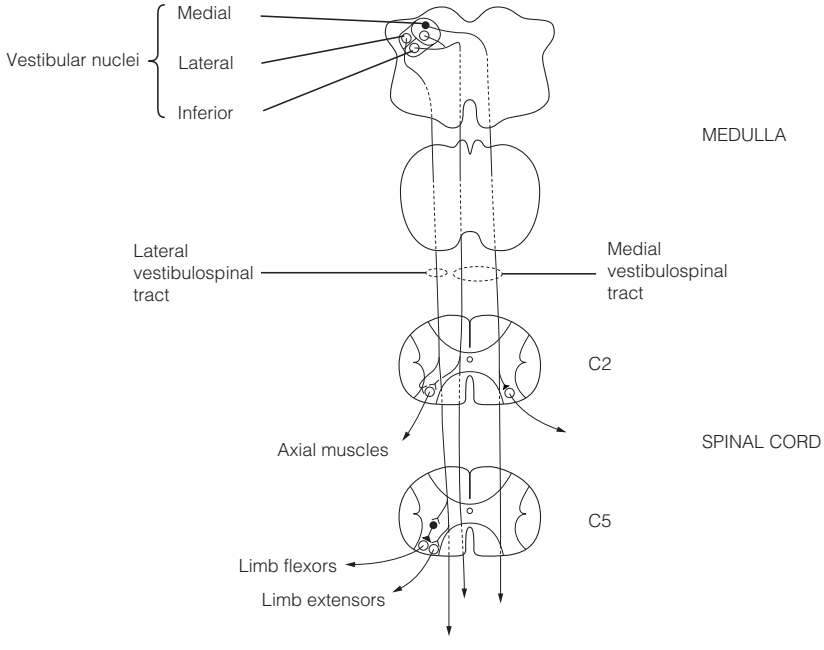


Fig. 2. The vestibulospinal tracts. Filled circles and triangles represent the cell bodies and axon terminals, respectively, of inhibitory neurons.

tively called the **medial motor pathways** to distinguish them from the **lateral motor pathways** for making voluntary movements. The medial motor system comprises the vestibulospinal (*Fig. 2*) and reticulospinal tracts (*Fig. 3*) which terminate in the intermediate and ventral gray matter of the spinal cord. Input from the vestibular labyrinth goes to **medial** and **inferior vestibular nuclei**, which give rise to the bilateral **medial vestibulospinal tract**. This makes connections (many monosynaptic) with neck muscle motor neurons. In general, ipsilateral motor neurons are excited whilst contralateral ones are inhibited. The pathway mediates some vestibulocollic reflexes.

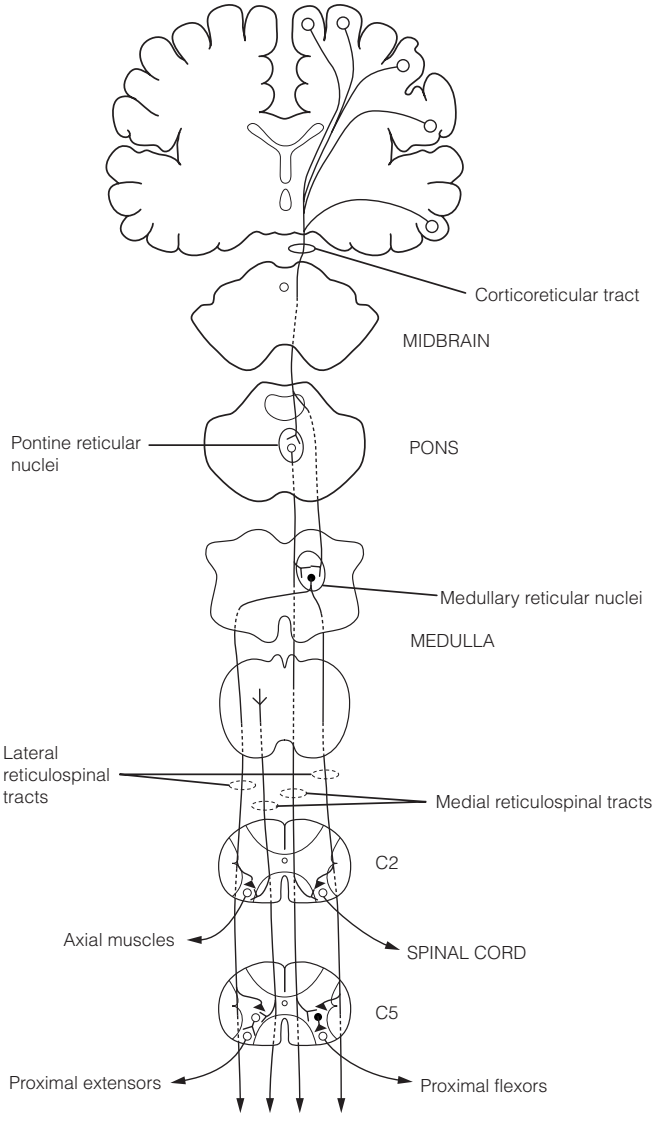


Fig. 3. The reticulospinal tracts. The full extent of the medial reticulospinal tracts is shown on one side only. The action of the tracts on extensors and flexors is shown on opposite sides for clarity. Filled circles and triangles are the cell bodies and axon terminals, respectively, of inhibitory neurons.

Vestibular afferents to the **lateral vestibular (Deiter's) nucleus** are implicated in the control of limb muscles. The lateral vestibular nucleus projects via the uncrossed **lateral vestibulospinal tract** to all segments of the spinal cord. Neurons in this pathway facilitate extensor motor neurons and inhibit flexor motor neurons. This pathway is responsible for some vestibulospinal reflexes.

There are two reticulospinal tracts. The **medial reticulospinal tract** originates from the **pontine reticular nuclei** and is ipsilateral. Its neurons excite axial and limb extensor motor neurons, but are inhibitory via polysynaptic pathways to limb flexors. The **medullary reticular nuclei** give rise to the **lateral reticulospinal tract** which is bilateral, and produces monosynaptic inhibition of neck and axial motor neurons and polysynaptic inhibition (excitation) of proximal limb extensors (flexors). The medullary reticular nuclei get input from the mesencephalic locomotor region and project to the central pattern generators. *Table 1* summarizes the effects of the descending motor pathways on motor neurons.

Table 1. A summary of the major features of the descending motor pathways[a]

Motor system	Tract	Distribution	Principal effects on motor neurons	
			Excitatory to:	*Inhibitory to:*
Medial	Lateral vestibulospinal	Ipsilateral	Axial and proximal limb extensors	Axial and proximal limb flexors
	Medial vestibulospinal	Bilateral	Axial ipsilateral	Axial contralateral
	Pontine (medial) reticulospinal	Ipsilateral	Axial and proximal limb extensors	Proximal limb flexors
	Medullary (lateral) reticulospinal	Bilateral	Proximal limb flexors	Axial and proximal limb extensors
Lateral	Corticospinal	Largely contralateral	Distal limb flexors	Distal limb extensors
	Rubrospinal	Bilateral	Distal limb flexors	Distal limb extensors

[a] The lateral motor system is described in Topic K1.

The reticular nuclei receive proprioceptor input from muscle spindles and vertebral joint receptors, and vestibular input, so reticulospinal tracts service both neck and vestibular reflexes. In fact, otolith signals for head-forward and back-pitching components of the vestibulocollic reflexes are not transmitted through vestibulospinal tracts but probably involve the reticulospinal tracts. The reticular nuclei receive input from the premotor cortex, which allows postural reflexes to be modified by feedforward during intentional movement. After lesions of the medullary reticular nuclei in cats, anticipatory adjustments are compromised and the animals lose balance temporarily when attempting to move a forelimb.

K1 Cortical control of voluntary movement

Key Notes

Intentional movement

Voluntary movements are those made intentionally. They are planned, and then executed by output of motor commands which specify the correct sequence of muscle activation. Sensory feedback can be used to optimize performance.

Lateral motor pathways

The motor cortex gives rise to two lateral pathways for the execution of intentional movements. The corticospinal tract is made up of axons of pyramidal cells in layer V of the cortex which descend through the internal capsule into the brainstem where most cross the midline giving rise either to corticonuclear fibers that go to the motor nuclei of cranial nerves or to the lateral corticospinal tract. Corticospinal tract neurons are excitatory, make monosynaptic connections with α motor neurons to distal limb muscles and polysynaptic connections to α motor neurons to proximal and axial muscles. Fusimotor (γ) motor neurons are coactivated polysynaptically. The corticospinal tract is commonly excitatory to flexors and inhibitory (via Ia inhibitory interneurons) to extensors. The rubrospinal tract from the red nucleus is adjacent to the corticospinal tract, but is much less prominent in humans than in other mammals.

Motor cortex

The motor cortex is divided into three reciprocally interconnected areas, the primary motor cortex (M1), and the supplementary motor area (SMA) and premotor area (PM). The SMA and PM together constitute the secondary motor cortex. The SMA forms a motor loop with the basal ganglia while M1 forms a motor loop with the cerebellum. The motor cortex has several somatotopic maps. The map in M1 has more cortical space devoted to regions such as the hands and face where the variety and complexity of movements is greatest. An M1 cell does not have exclusive control over individual muscles or movements but its activities correlate with a wide variety of movement parameters (e.g. force). Direction of limb movement is not encoded by the behavior of a single cell, but by the average firing of a population of neurons. The secondary motor cortex is involved in planning movements and its neurons may fire hundreds of milliseconds before a movement begins. The SMA has a bilateral somatotopic map and is crucial for performing complicated tasks involving both sides of the body, usually previously learnt. The premotor area is particularly concerned with planning movements that require sensory cues.

Red nucleus

The red nucleus has a motor map, firing of rubrospinal tract neurons correlates with motor parameters in a similar manner to corticospinal tract neurons, and they are distributed to motor neuron pools in a comparable fashion. This implies that the rubrospinal and corticospinal tracts are very similar. However the rubrospinal tract may operate when learnt, automated movements are being executed, while the corticospinal tract is active when new motor tasks are being acquired.

Related topics

- Organization of the central nervous system (A5)
- Touch (F2)
- Elementary motor reflexes (J3)
- Cerebellar function (K6)
- Anatomy of the basal ganglia (K7)
- Basal ganglia function (K8)

Intentional movement

Voluntary movements are those made intentionally to achieve a particular aim or reach for a target. Many neurons in the motor system fire hundreds of milliseconds before any muscle contraction occurs showing that movements are planned. This is necessary because the same motor tasks can be performed in various ways depending on the context, a property called **motor equivalence**. For example, driving a large truck needs a different motor strategy than driving a small car. The movement is executed by the output of **motor commands** which specify the correct temporal sequence of muscle activation. Sensory feedback during the movement, particularly from proprioceptors such as muscle spindles and Golgi tendon organs, and perhaps from the visual system is used to fine tune the execution to ensure that the performance matches the desired goal. The planning of voluntary movements and the elaboration of motor commands for their execution is done by the motor cortex, which has its outputs via the lateral motor pathways.

Lateral motor pathways

There are two lateral pathways for descending control of voluntary movement. Both originate in the motor cortex, which lies on the frontal lobe just anterior to the central sulcus. The **corticospinal tract** consists of the axons of about one million pyramidal cells in layer V of the cortex. Over half come from the **primary motor cortex** (**M1**), Brodmann area 4, and the **supplementary motor area** (**SMA**) or **premotor area** (**PM**) in Brodmann area 6. These project to the ventral horns of the spinal cord. About 40% of corticospinal tract axons come from the somatosensory cortex (Brodmann areas 1, 2 and 3) or other regions of parietal cortex (Brodmann areas 5 and 7). These axons from the parietal lobe terminate in the dorsal horns of the spinal cord and regulate sensory input. Most of the corticospinal tract consists of fine myelinated and unmyelinated axons with conduction velocities between 1 and 25 m s^{-1}. However, there are about 30 000 extremely large (20–80 μm diameter) pyramidal cells in area 4, called **Betz cells**, with big myelinated axons that conduct with velocities of 60–120 m s^{-1}. Axons of the corticospinal tract pack tightly to pass through the internal capsule which lies between the thalamus and the lentiform nucleus (*Fig. 1*) and descend into the brainstem. Here the most medial fibers peel off and cross the midline to go to nuclei (trigeminal (V), facial (VII), hypoglossal (XII) and accessory (XI)) of the cranial nerves. These are **corticonuclear** (**cortico-bulbar**) fibers and are motor to the face, tongue, pharynx, larynx and to the sternomastoid and trapezius muscles. The remaining axons descend through the medulla causing a swelling on its ventral surface, the **pyramid**, and for this reason the corticospinal tract is often referred to as the **pyramidal tract** (the term is unrelated to its derivation from pyramidal cells). At the caudal medulla 85% of fibers cross the midline as the **pyramidal decussation**, giving rise to the **lateral corticospinal tract**. The remaining ipsilateral axons form the **anterior corticospinal tract**, which crosses over at spinal cord level.

Corticospinal tract neurons use glutamate as a transmitter and are excitatory.

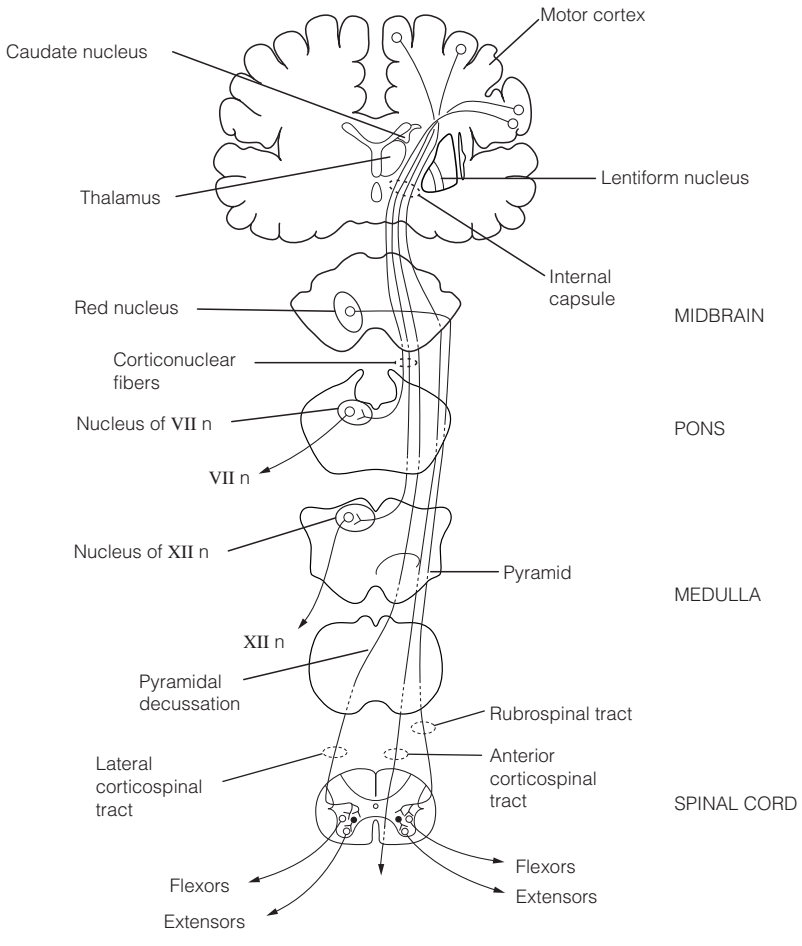


Fig. 1. The lateral motor pathways. Only corticonuclear fibers in the facial (VII n) and hypoglossal (XII n) nerves are shown. Synapses of the anterior (ipsilateral) corticospinal tract with spinal cord interneurons are not shown.

They either synapse directly with α motor neurons supplying distal limb muscles in Rexed lamina IX, or synapse with interneurons in laminae VII and VIII, which make polysynaptic connections with α motor neurons of proximal limb muscles and axial muscles. Fusimotor (γ-efferent) neurons that must be coactivated with α motor neurons to override the stretch reflex during voluntary movement are excited polysynaptically. Stimulation of the corticospinal tract is predominantly excitatory to flexors but inhibitory to extensors. The corticospinal tract inhibits motor neurons disynaptically via Ia inhibitory interneurons.

The corticospinal tract axons arising from the somatosensory cortex project to cranial nerve sensory nuclei and dorsal horns and produce presynaptic inhibition on primary afferent terminals *except* for Ia spindle afferents.

Some pyramidal cells in layer V of the motor cortex send their axons in the **corticorubral tract** to the **red nucleus** in the midbrain, which also receives collaterals from the corticospinal tract. The red nucleus gives rise to the **rubrospinal**

tract, the second of the lateral motor pathways. Some of its axons go to cranial nerve nuclei in the pons and medulla. The tract descends as far as the lumbar cord in the macaque monkey and terminates in the dorsolateral gray matter. In humans the rubrospinal axons descend as part of the corticospinal tract.

Motor cortex

The motor cortex is subdivided into three areas on the basis of cytoarchitecture, connectivity and function. The primary motor cortex, M1, is **agranular cortex**, in that layer IV which receives inputs from the thalamus is very sparse, whereas layer V is well developed and contains numerous large pyramidal cells. Brodmann area 6 in which layer IV is better developed contains the SMA and PM areas. The motor areas are reciprocally connected with each other (*Fig. 2*). They are also connected with subcortical structures which send inputs back to the motor cortex via the thalamus forming closed motor loops. Reciprocal back projections from the motor cortex to the thalamus also exist. These are not shown on *Fig. 2*.

The supplementary motor area is part of a motor loop with the basal ganglia. It sends output to the striatum which projects back to the SMA via its connections with the globus pallidus and ventrolateral (VL_o) thalamus. Many corticospinal tract axons from M1 either terminate in the pons or give off collaterals there. These make synapses with pontine neurons which project to the cerebellum. Outputs from the cerebellum to the thalamus, which in turn project

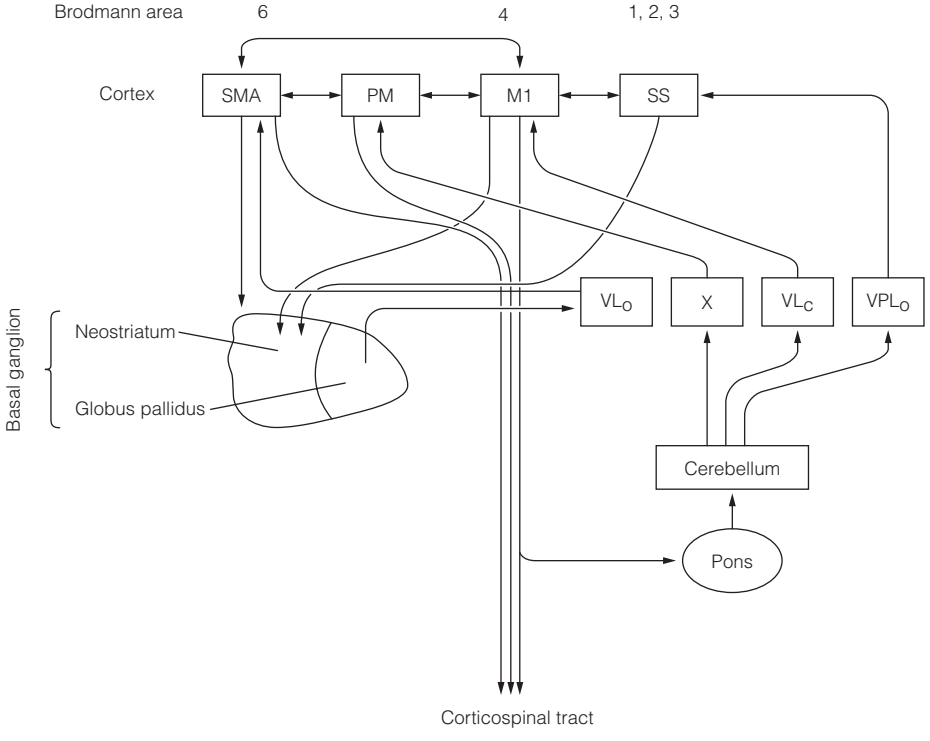


Fig. 2. The connections of the motor cortex establish motor loops with the basal ganglia and the cerebellum in which the thalamus provides the input that closes the loop. SS, somatosensory cortex. Thalamic nuclei: VL_c, ventroposterior nucleus pars caudalis; VL_o, ventroposterior nucleus pars oralis; VPL_o, ventroposterolateral nucleus pars oralis; X, nucleus X.

back to M1 and PM motor areas, form other motor loops. These motor loops with basal ganglia and cerebellum are required for the initiation of specific motor patterns and their coordination.

Several somatotopic maps exist in the motor cortex (*Fig. 3*). Their topography is preserved in the orderly arrangement of the fibers in the corticospinal tract; axons synapsing with leg or arm lower motor neurons being lateral and medial respectively. Like somatosensory maps, these motor maps are grossly distorted. More cortical space is devoted to the face, tongue and hands than other regions because the variety and precision of movements executed by them is so much greater. M1 receives a substantial input from the somatosensory cortex and many M1 neurons have sensory receptive fields. M1 neuron RFs are located in the position where activity of the neuron is likely to cause movement. That is, M1 neurons are wired to be responsive to the sensory consequences of their actions. Those in caudal M1 respond mostly to cutaneous mechanoreceptor input, those more rostral respond to proprioceptor input, particularly from muscle spindles. This contributes to a cortical loop (**long loop**) of the stretch reflex circuitry. This long loop conveys muscle spindle input to M1 from where corticospinal tract axons go to influence the lower motor neurons. By informing the motor cortex about the state of muscles the long loop circuit allows the motor cortex to rapidly modify the stretch reflex in the event of unexpected changes in load. The long loop modifies muscle contraction on a timescale slower that that of the myotatic reflex, but faster than voluntary movements.

What the somatotopic mapping in the motor cortex represents is uncertain. It is not a one-to-one mapping to individual muscles or movements. This is shown by:

- output from single cortical neurons diverging to several motor neuron pools;
- converging outputs from quite a wide area of M1 onto the motor neuron pools for muscles moving a specific body part;
- a given muscle being subserved by a region in the motor cortex which overlaps with regions controlling neighboring muscles.

During the execution of a movement involving a muscle the set of neurons within the region that are activated depends on the nature of the movement; for example, its direction and force.

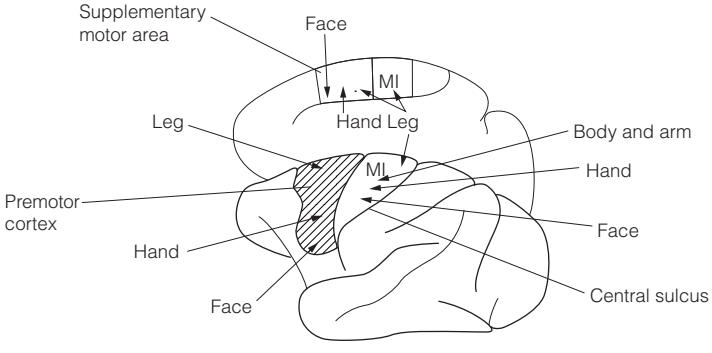


Fig. 3. Approximate somatotopic mapping in the motor cortex of the macaque. M1, primary motor cortex.

Recording from single cells in the motor cortex of conscious monkeys engaged in intentional limb movements shows that M1 cell firing can correlate with force, rate of change of force, velocity, acceleration, direction of movement or joint position. None of these parameters is mapped in an orderly way in the cortex. Firing of an M1 cell during a task is usually related to two or three of these variables so M1 cells do not exclusively encode a single movement parameter. Many M1 cells are rather broadly tuned for movement direction so a single M1 cell cannot encode direction of movement very well. However, direction of movement is very precisely coded by the average firing of a few hundred cells (*Fig. 4*). This is an example of population coding. The population of cells encoding the direction of a given movement are not localized to a discrete site, but are quite widely distributed across the cortex.

The supplementary motor area (SMA) and premotor area (together called the **secondary motor cortex**) contain neurons that fire in a way that correlates with the direction and force of a movement. The SMA controls proximal limb muscles directly via its output to the corticospinal tract, whereas the premotor area neurons synapse with brainstem reticular neurons that go to axial and proximal limb muscles. Both control distal limb muscles via M1.

The SMA has bilateral representations of the body and is crucial for movements involving both sides of the body, particularly those that have been learnt.

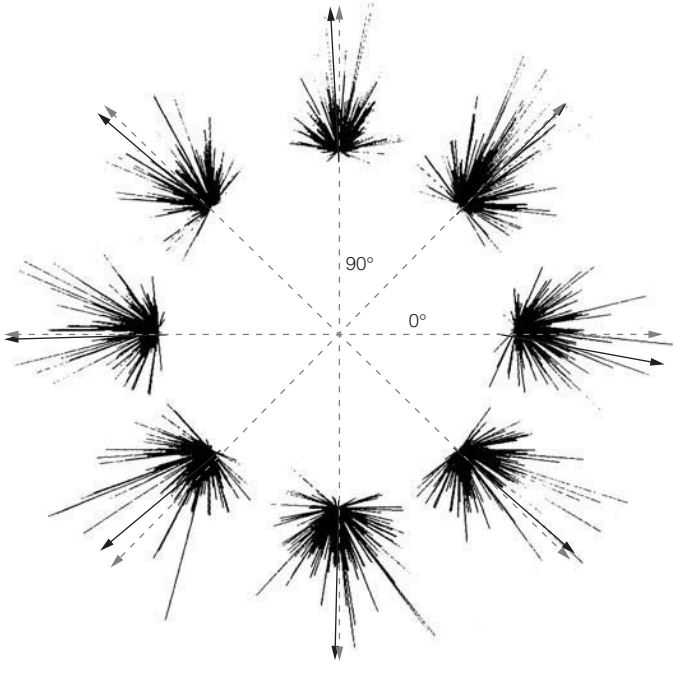


Fig. 4. Population coding of movement direction. A monkey was trained to move its arm in eight different directions. Each cluster of lines represents the activity of a population of neurons encoding that direction. The direction of each line represents the preferred direction of the cell, and the length is proportional to its firing rate. The thick arrow is the vector average for the population. In most cases it points closely in the desired direction. Reproduced from Georgopoulos, A.P. et al. *(1982)* J. Neurosci. ***2****, pp. 1527–1537, with permission. Copyright 1982, The Society for Neuroscience.*

For this, the motor loop involving the basal ganglia is important. The premotor area consists of a number of discrete motor regions and so has several maps. It is implicated in movements in response to sensory (mostly visual) cues. The premotor area gets a large input from the **posterior parietal cortex** (Brodmann areas 5 and 7) association cortex, which receives visual, somatosensory and vestibular sensory input. The posterior parietal cortex thus provides sensory input for targeted movements. Some posterior parietal neurons are context specific, firing only during goal-directed behavior (e.g. reaching for food) but remaining silent if the limb moves in the same way in the absence of the goal.

A key role of the secondary motor cortex is in planning movements. This is shown, firstly, by the fact that neurons here fire a long time (possibly up to 800 ms) before a voluntary movement begins. Secondly, measurements of cerebral blood flow (cbf) in humans doing motor tasks show that a simple movement involves increased cbf in MI only, a more complex task is accompanied by increased cbf in secondary motor cortex as well as MI, but remarkably when subjects were required to mentally rehearse a complex task (but not execute it) an increase in cbf was seen restricted to the secondary motor cortex.

Red nucleus

The red nucleus has a somatotopic map. Its activity precedes intentional movements and correlates with parameters such as force, velocity and direction much like corticospinal tract neurons. Furthermore, in many primates rubrospinal axons have the same distribution to proximal and distal limb motor neurons as the corticospinal tract and their activity moves individual digits. However, in humans the distinction between the two lateral pathways seems not to be as important as in many mammals, and the rubrospinal tract is involved in gross limb movements not fine ones.

Although the two pathways appear strikingly similar, studies in sub-human primates suggest that they operate in different contexts. While the rubrospinal tract is active when previously learnt automated movements are executed, the corticospinal tract is required when novel movements are being learnt. Another pathway acts to switch activity between the two lateral motor systems. As a new movement is successfully learnt its execution is switched from the corticospinal tract to rubrospinal tract control. The switch operates in the opposite direction if an automatic movement needs to be adapted. It is because of the switch that each lateral motor system can compensate for the loss of the other. Corticospinal tract lesions have a more severe and protracted effect than rubrospinal tract lesions because new movements cannot be executed and only the old rubrospinal tract repertoire can be called upon. The switch pathway involves the inferior olive, one function of which is to detect and correct errors in motor performance.

K2 Motor lesions

Key Notes

Brown–Sequard syndrome

Severing the spinal cord on one side causes paralysis and loss of touch and proprioceptor sensation below the lesion on the same side and loss of pain and temperature sensation on the contralateral side. This is the Brown–Sequard syndrome.

Decerebrate rigidity

Lesions which sever the brainstem between the red nucleus and the vestibular nuclei cause an increase in extensor tone known as decerebrate rigidity. It is caused by the loss of facilitation of flexor motor neurons by the rubrospinal tract.

Lesions of spinal motor pathways

Cutting the corticospinal tract causes an ipsilateral lesion if the transection is below the pyramidal decussation, and a contralateral lesion if above it. A pure corticospinal tract lesion in non-human primates causes a loss of fine movements by distal muscles. Cutting vestibulospinal and reticulospinal tracts causes deficits in posture and locomotion.

Cerebrovascular accidents

In clinical terminology, lower motor neurons are those that innervate skeletal muscles while upper motor neurons are pyramidal tract neurons. However the symptoms of an upper motor neuron lesion cannot be accounted for by lesions of the corticospinal tract alone, but involve the loss of corticoreticular neurons. The major cause of an upper motor neuron lesion is a cerebrovascular accident (stroke), the commonest of which is an infarction of the internal capsule due to blockage of the artery which supplies it. The long-term symptoms are muscle weakness and spasticity on the side opposite the lesion. Spasticity is an increase in muscle tone, particularly in extensor muscles and is caused by hyperexcitability of stretch reflexes resulting from the loss of reticulospinal tract-driven presynaptic inhibition on Ia terminals.

Related topics

- Touch (F2)
- Pain (F3)
- Brainstem postural reflexes (J5)
- Cortical control of voluntary movement (K1)
- Strokes and excitotoxicity (P3)

Brown–Sequard syndrome

A classic pattern of sensory and motor deficits, the **Brown–Sequard syndrome**, is seen when the spinal cord is severed on one side. On the side of the lesion there is a motor paralysis and loss of all sensation transmitted through the dorsal columns (touch and proprioception). On the contralateral side there is loss of nociceptor and thermoreceptor sensation from a few spinal segments below the lesion. This is due to the interruption of the anterolateral columns which contain spinothalamic axons that have crossed from the opposite side (*Fig. 1*).

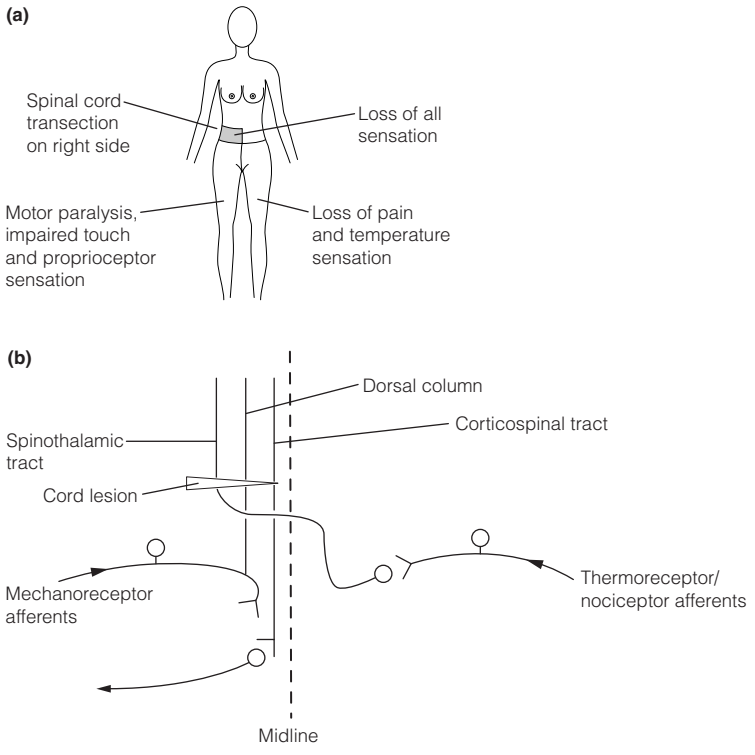


Fig. 1. Brown–Sequard syndrome results from the hemisection of the spinal cord which interrupts dorsal column input and motor output on the side of the lesion and spinothalamic input from the contralateral side. (a) Signs and symptoms; (b) lesion.

Decerebrate rigidity

Patients in whom brain trauma or a tumor produces functional disconnection of the brainstem from the rest of the brain at the level between the red nucleus and the vestibular nuclei show an increase in extensor tone called **decerebrate rigidity**. It is caused by tonic activity in the vestibulospinal and reticulospinal neurons that is no longer opposed by the powerful facilitation of flexor motor neurons by the rubrospinal tract (*Fig. 2*). The overall effect of the vestibulospinal activity is the activation of extensors. Reticulospinal inhibition and excitation of extensor motor neurons tend to cancel, but reticulospinal inhibition of interneurons mediating flexor reflexes enhances, net extensor activity. The high firing rates of both α and γ motor neurons facilitated by vestibulospinal and reticulospinal inputs requires tonic muscle spindle afferent input, since cutting dorsal roots in decerebrate animals abolishes decerebrate rigidity. Ablation of the cerebellum in animal studies, by removing inhibitory input to the lateral vestibular nucleus intensifies decerebrate rigidity.

Lesions of spinal motor pathways

Experimental transection of the corticospinal tract below the pyramidal decussation in primates causes an ipsilateral motor deficit below the level of the section (see Brown–Sequard syndrome above). Corticospinal tract lesions above the pyramidal decussation cause a contralateral deficit. The deficit seen with a *pure* corticospinal lesion is a loss of the ability to make fine movements with distal muscles, for example an inability to make independent finger movements

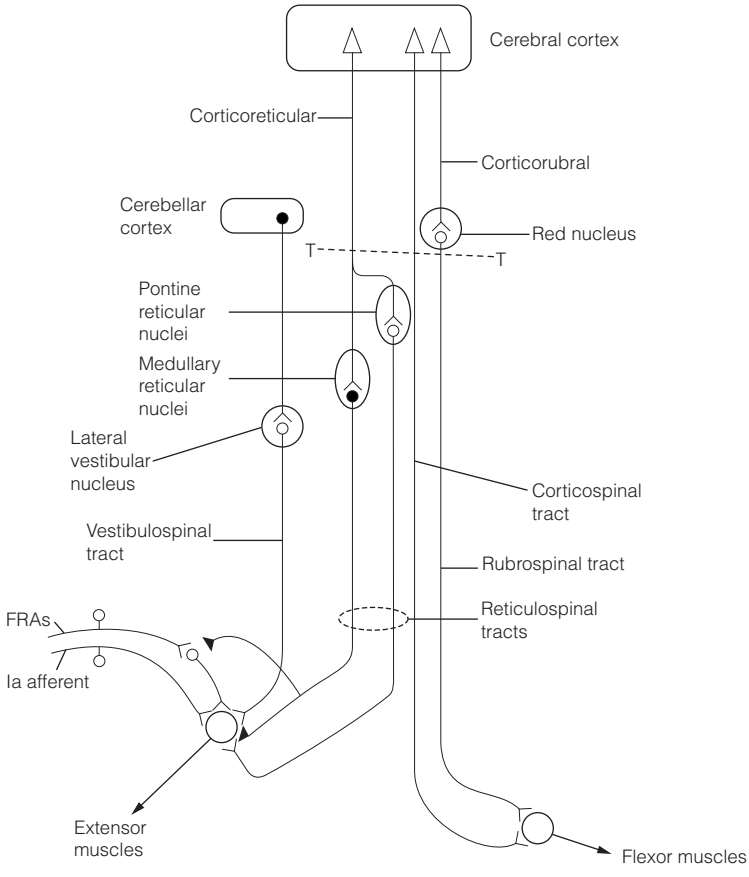


Fig. 2. Illustration of the principal descending inputs to motor neuron pools. Fusimotor fibers (not shown) are generally affected in the same manner as α motor neurons. A transection of the brainstem at the level of T–T causes decerebrate rigidity. FRAs, flexor reflex afferents.

needed to retrieve an object from a narrow hole. Almost complete recovery is eventually seen. A similar though less severe and transient deficit is seen if the rubrospinal tract alone is cut. However, a lesion of both lateral motor pathways results in the deficit being permanent.

In contrast, experimental transection of the vestibulospinal and reticulospinal tracts which control output to proximal limb and axial muscles produces much more extensive deficits in posture, walking and climbing but leaves fine control of distal muscles intact.

Cerebrovascular accidents

Clinicians distinguish deficits due to lesions of **lower motor neurons** (brainstem and spinal cord neurons innervating skeletal muscle) from those due to lesions of **upper motor neurons**, which refers not only to the corticospinal and corticobulbar neurons of the pyramidal tract, but also to cortical cells that drive reticulospinal (medial pathway) neurons. The symptoms attributed to upper motor neuron (pyramidal) lesions *cannot* be explained by damage to corticospinal or corticobulbar neurons alone. This is exemplified in strokes, the major cause of upper motor neuron lesions.

The commonest **cerebrovascular accident** (**CVA, stroke**) is caused by a thromboembolism affecting the branch of the middle cerebral artery that supplies the internal capsule. Infarction of the internal capsule produces a syndrome which *does not* resemble experimental lesions of the lateral motor pathways because the internal capsule also contains corticoreticular axons which drive lateral and medial reticulospinal tracts. After an initial period of flaccid paralysis and lack of reflexes on the side opposite of the lesion two major deficits are seen.

1. **Hemiparesis**. Muscle weakness on one side that is greatest in arm extensors and leg flexors, because arm flexors are stronger than extensors and in the legs the reverse is true. If the corticobulbar fibers are affected, voluntary facial movements are compromised. When the weakness is so severe that paralysis results the term **hemiplegia** is used. Weakness occurs because the loss of descending excitation means fewer motor units are recruited.
2. **Spasticity**. An increase in **muscle tone** is seen in the stronger (antigravity) limb muscles. It is caused by enhanced excitability of the stretch reflex, particularly the phasic component, since attempts at rapid muscle stretch are met with much greater resistance than slow stretch. Forceful attempts to stretch a muscle are met with great resistance (caused by the stretch reflex), which fails suddenly (the **clasp knife response**) due to firing of high-threshold muscle (non-spindle) afferents.

Spasticity in part results from the loss of presynaptic inhibition on Ia terminals. Normally, presynaptic inhibition is brought about by the action of the reticulospinal tracts on GABAergic Ia presynaptic inhibitory interneurons. GABA released from these interneurons acts on $GABA_B$ and $GABA_A$ receptors on the Ia terminals. $GABA_B$ receptors are metabotropic receptors and when stimulated act via G_i proteins to increase the K^+ conductance. The resulting hyperpolarization reduces Ca^{2+} influx into the primary afferent terminal, so curtailing the release of glutamate onto the motor neurons. In spasticity, this descending reticulospinal input is lost. This leads to failure of presynaptic inhibition and so hyperexcitability of the stretch reflex.

Baclofen is an agonist at $GABA_B$ receptors and is used orally and intrathecally in the treatment of spasticity. Because benzodiazepines (e.g. diazepam) are agonists at the $GABA_A$ receptors involved in presynaptic inhibition, they too can be used in spasticity, but have the disadvantage of producing sedation at the doses needed to reduce muscle tone.

K3 Anatomy of the cerebellum

Key Notes

Gross anatomy of the cerebellum

The cerebellum is divided into three lobes: anterior, posterior and flocculonodular, each of which is subdivided into lobules. Longitudinally it has a central vermis and two lateral hemispheres. The cerebellum is covered with a cortex and embedded within its core of white matter are deep nuclei which, with the vestibular nuclei, provide the cerebellar output.

Cerebellar connections

Input to the cerebellum is from the spinal cord and brainstem sensory systems, cerebral cortex and inferior olive.

Proprioceptor pathways

Sensory input from proprioceptors is used by the cerebellum to provide feedback on motor performance. Proprioceptor input to the cerebellum from the upper part of the body comes from collaterals of axons ascending in the dorsal columns which enter the accessory cuneate nucleus. This nucleus projects to the cerebellum by way of the cuneocerebellar tract. Proprioceptor afferents from the lower body terminate in Clark's column of the dorsal horn which gives rise to the dorsal spinocerebellar tract. A ventral spinocerebellar tract comes from the ventral horn and conveys information about the state of spinal circuits controlling locomotion.

Cerebellar zones

The human cerebellum is parceled into three sagittal zones each with distinct deep nuclei outputs and connections with the rest of the CNS. The medial zone (vermis) sends its output to the fastigial nucleus which relays to the vestibular nuclei. The intermediate zone output is via the interposed nucleus while the lateral zone output goes by way of the dentate nucleus. The interposed and dentate nuclei project to the red nucleus and ventrolateral thalamus.

Related topics

Touch (F2)
Subdivisions of the cerebellum (K4)
Spinal motor function (J4)
Cerebellar cortical circuitry (K5)

Gross anatomy of the cerebellum

The human cerebellum is about one quarter of the mass of the brain and contains in excess of 10^{11} neurons. It is divided into three lobes, the **anterior lobe** and **posterior lobe**, separated by the primary fissure, and the **flocculonodular lobe**, separated from the posterior lobe by the posterolateral fissure (*Fig. 1*). The lobes are further subdivided into **lobules** which are differently named in humans and other mammals. Longitudinally the cerebellum has a central **vermis** and two **lateral hemispheres**.

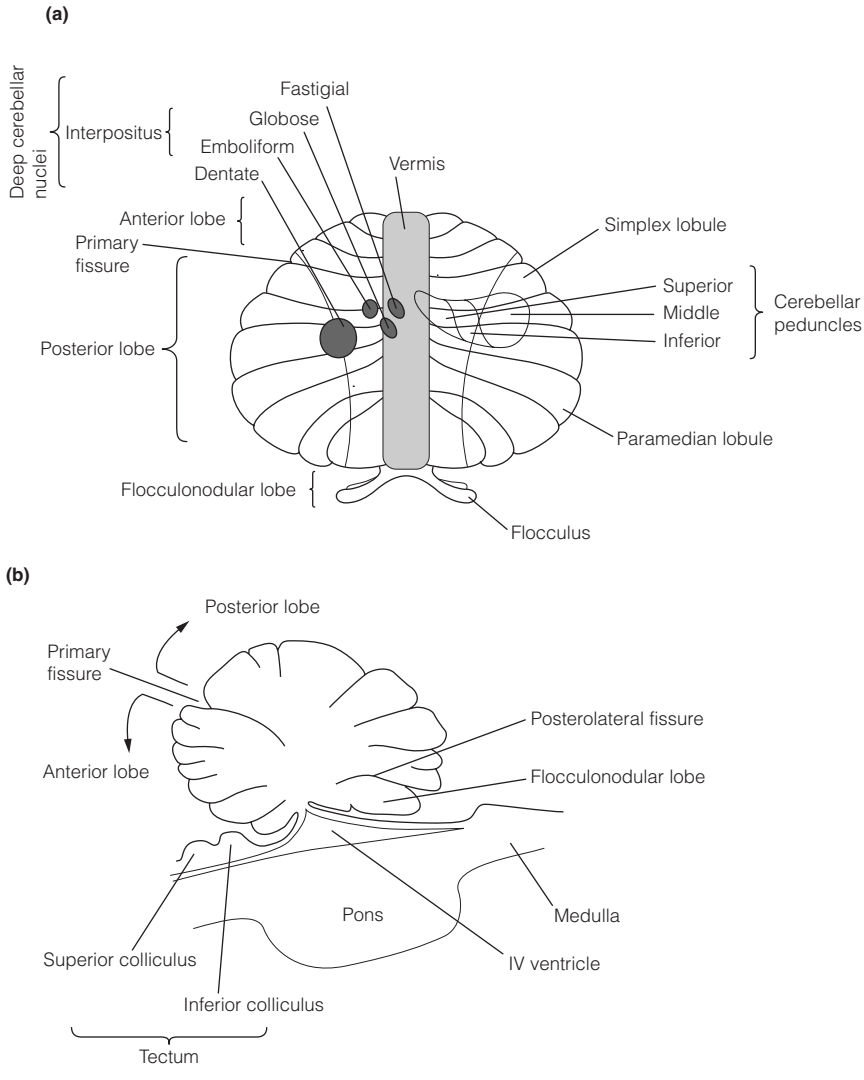


Fig. 1. Cerebellar anatomy. (a) A diagram of the cerebellum, unfurled and viewed from above. Locations of the deep cerebellar nuclei are shown on the left and of the cerebellar peduncles on the right. The medial zone is pale gray, the intermediate zone stippled and the lateral zone clear. (b) Sagittal section through the brainstem and cerebellum.

Over the surface of the cerebellum lies the cerebellar cortex, which is folded into coronal strips called **folia** (singular **folium**). The cerebellum has an internal core of white matter containing **deep** (**intracerebellar**) **nuclei**. These, together with the vestibular nuclei, provide the output of the cerebellum. Afferent and efferent connections of the cerebellum go by way of three pairs of **cerebellar peduncles**.

Cerebellar connections

Input to the cerebellum comes from three major sources:

- the spinal cord and brainstem, conveying sensory information from several modalities;

Table 1. Principle inputs to the cerebellum

Tract	Origin	Peduncle	Distribution	Modality
Vestibulocerebellar	Vestibular nuclei	ICP	Crossed and uncrossed to flocculonodular lobe cortex and fastigeal nucleus	Vestibular
Trigeminocerebellar	Secondary afferents in trigeminal nerve (V nerve) nuclei	ICP	Crossed and uncrossed	Proprioceptive and cutaneous somatosensory from jaw and face
Cuneocerebellar	Accessory cuneate nucleus	ICP	Uncrossed	Proprioceptive from arm and neck
Dorsal spinocerebellar	Clark's column	ICP	Uncrossed	Proprioceptive and cutaneous somatosensory from trunk and leg
Ventral spinocerebellar	Ventral horn	SCP	Crossed and uncrossed	Proprioceptive and cutaneous somatosensory from all parts of the body
Tectocerebellar	Superior colliculi, inferior colliculi	SCP	Crossed	Visual and auditory
Pontocerebellar	Pontine nuclei	MCP	Crossed	Cognitive, motor, somatosensory and visual from cerebral cortex
Olivocerebellar[a]	Inferior olivary nucleus	ICP	Crossed to all deep cerebellar nuclei	Motor error signals

[a]The olivocerebellar input is via climbing fibers, all other afferents are mossy fibers. ICP, Inferior cerebellar peduncle; SCP, superior cerebellar peduncle; MCP, middle cerebellar peduncle.

- the cerebral cortex, by way of pontine nuclei, in the massive corticopontinecerebellar tract relaying motor and sensory input;
- the inferior olivary nucleus by way of the olivocerebellar tract.

Details of these inputs are given in *Table 1*. Cerebellar output goes to two principal destinations:

- the ventrobasal thalamus, which projects to the motor cortex, to modify corticospinal tract outflow;
- the red nucleus, to modify the rubrospinal tract output;
- vestibular and reticular nuclei to modulate the activity of medial motor pathways.

These general input–output relations are depicted in *Fig. 2*, although there are variations on this arrangement for different parts of the cerebellum.

Proprioceptor pathways

Proprioceptor information coming from muscle spindles, Golgi tendon organs and receptors in joints provides for conscious awareness of body position and movement. It is also used by the cerebellum to guide motor performance. Proprioceptor input from the neck, arms and upper trunk is relayed in the

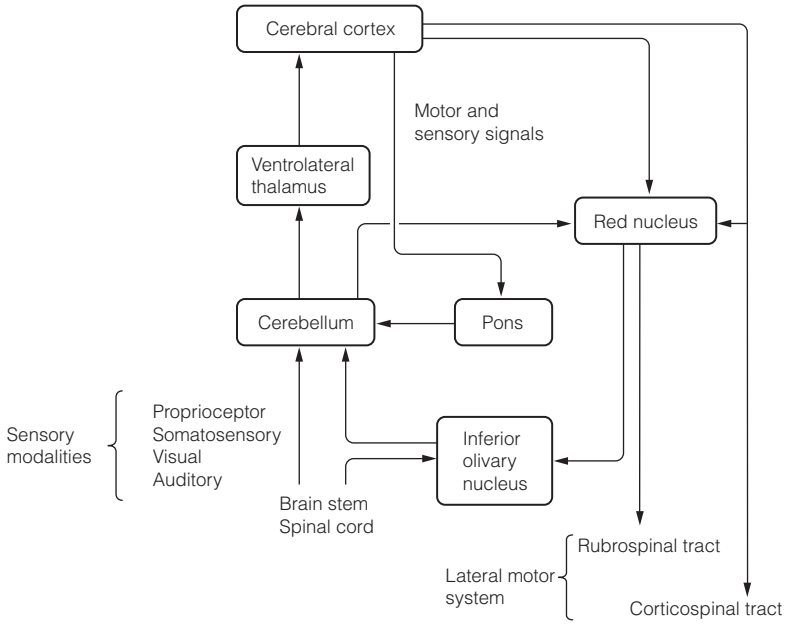


Fig. 2. Major connections of the cerebellum. Cerebellar output also goes to the medial motor system (not shown).

dorsal columns to the cuneate nucleus, from where it follows exactly the same path as touch sensation from the same areas. This is the route for conscious upper body proprioception. Cerebellar input comes by way of the **cuneocerebellar tract** (**CCT**), which arises from the **accessory cuneate** (**external arcuate**) **nucleus**. This receives the collaterals of axons ascending to the cuneate nucleus on the same side.

The proprioceptor pathway from the lower trunk and legs is distinct from that serving the upper body. Proprioceptor afferent terminals from the lower body enter the **nucleus dorsalis** (**Clark's column**), which is located in lamina VII of the medial dorsal horn and extends along the spinal cord between segments C8–L3. The nucleus dorsalis houses second-order neurons, axons from which ascend on the same side to form the **dorsal spinocerebellar tract** (**DST**). Collaterals of the DST axons enter **nucleus Z**, which is located just above the gracile nucleus in the medulla. Nucleus Z neurons send their axons into the medial lemniscus and thereby provide input for conscious lower-body proprioception.

The **ventral spinocerebellar tract** (**VST**) arises from the ventral horn and transmits signals reflecting the current state of spinal cord central pattern generators involved in locomotion.

Cerebellar zones

The output of the cerebellar cortex to deep cerebellar nuclei, and climbing fiber input (see *Topic K5*), is organized into parallel sagittal zones that extend the complete rostrocaudal length of the cerebellum. The exact number and arrangement of the zones depends on the species. In humans there are three. The medial zone occupies the vermis and sends its output to the **fastigial nucleus**.

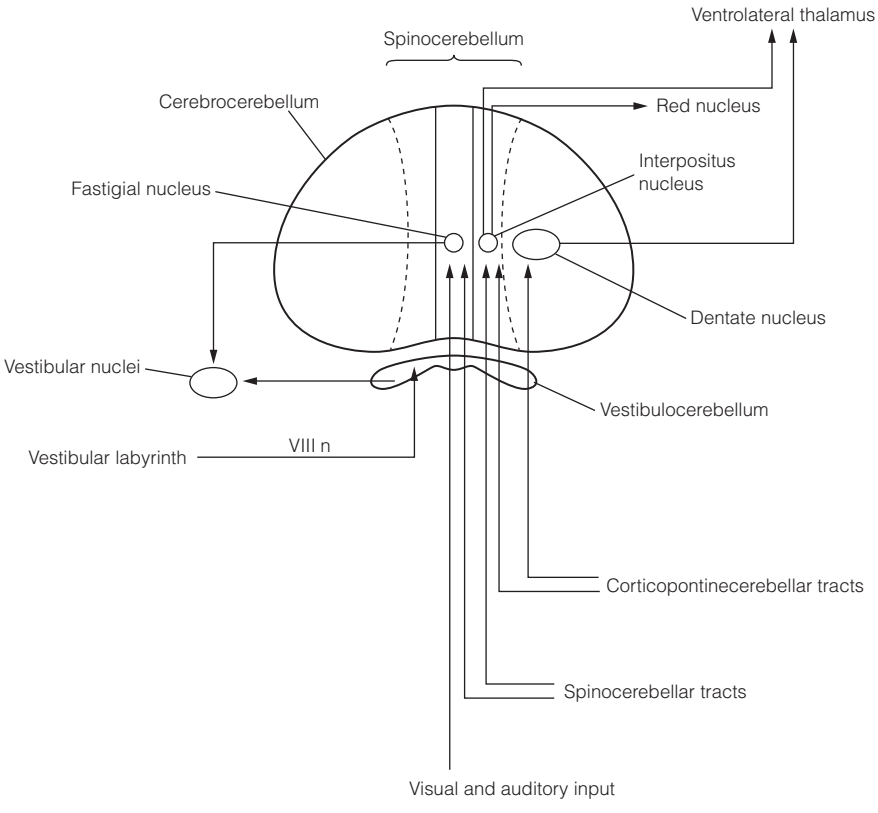


Fig. 3. The major connections and functional subdivisions of the cerebellum.

The intermediate zone projects to the **interpositus nucleus** (which in humans consists of two separate **emboliform** and **globose** nuclei), while the lateral zone output is to the **dentate nucleus**. The zones, and the flocculonodular lobe, each have fairly well segregated and distinct connections with the rest of the CNS that are summarized in *Fig. 3.*

K4 Subdivisions of the cerebellum

Key Notes

Somatotopic mapping

Inputs to the cerebellum are organized topographically and this gives rise to somatotopic maps in the cerebellar cortex, the deep cerebellar nuclei, and in their outputs to the red nucleus and thalamus. Each map represents not only sensory input but is also a motor output map.

Vestibulocerebellum

The vestibulocerebellum is the flocculonodular lobe and gets input from the vestibular system. Lesions cause defects in balance.

Spinocerebellum

The spinocerebellum is divided into a medial zone and an intermediate zone. The medial zone division controls postural adjustments by way of the medial motor system, using inputs from several sensory modalities. When lesioned, animals cannot stand or walk. The intermediate zone part receives proprioceptor input and sensorimotor input from the cerebral cortex via the corticopontinecerebellar pathway. It exerts control over limb movements via the lateral motor system and damage disrupts the ability to make accurate limb movements.

Cerebrocerebellum

The lateral zones comprise the cerebrocerebellum, which gets sensory, motor and cognitive input from the cerebral cortex. Lesions have only small effects on movements involving single joints but severe effects on more complex multijoint movements, and on language.

Related topics

Oculomotor control (G8)
Cortical control of voluntary movement (K1)
Anatomy of the cerebellum (K3)
Cerebellar function (K6)
Language (O5)

Somatotopic mapping

Mossy fiber input and climbing fiber input are organized topographically, giving somatotopic maps in the cerebellar cortex that are retained in the deep cerebellar nuclei, and in their output to the thalamus and red nucleus. The cerebellar cortex has several maps which exhibit fractured somatotopy (see *Fig. 1*) and encode visual and auditory input from the tectum as well as somatosensory input. Each of these representations is actually three maps in register. One is formed by mossy fiber input, a second is corticopontine input and the third is an output map that preserves a somatotopic projection of movements.

Vestibulocerebellum

The **vestibulocerebellum** corresponds to the flocculonodular lobe. It gets input from the ipsilateral vestibular labyrinth via the vestibulocochlear (VIII) nerve, and projects directly to the vestibular nuclei (*Fig. 2*). Lesions of the vestibulocerebellum in primates cause swaying and **truncal ataxia** (staggering gait). If the lesion is unilateral the head is tilted to the side of the injury and **nystagmus** is

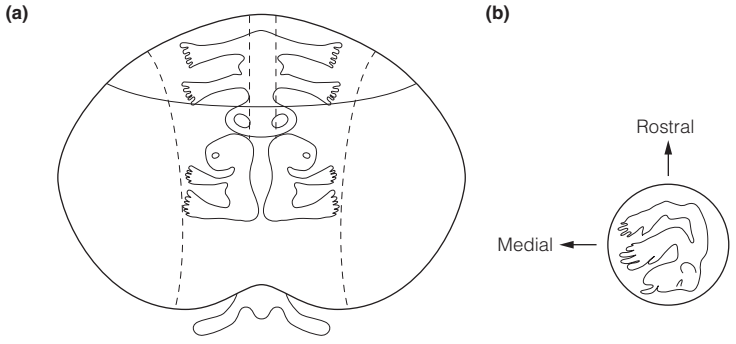


Fig. 1. Somatotopic maps in: (a) cerebellar cortex based on inputs and clinical lesions, the fractured nature of these inputs revealed by detailed studies is not shown; (b) deep cerebellar nucleus.

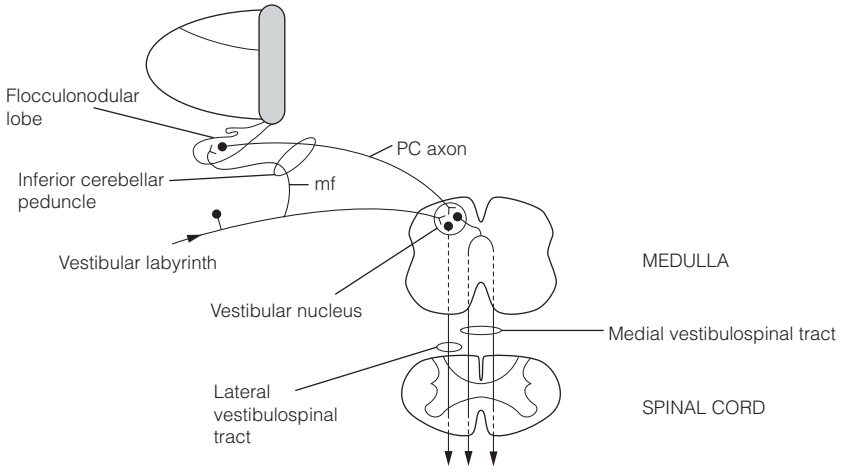


Fig. 2. Connections of the vestibulocerebellum. All cell bodies are shown as filled circles, as in other figures in this topic.

seen. The most common cause of damage to the vestibulocerebellum in humans is medulloblastoma of the fourth ventricle.

Spinocerebellum

The **spinocerebellum** consists of the anterior lobe and part of the posterior lobe, i.e. vermis, simplex and paramedian lobules. The medial zone of the spinocerebellum receives sensory input from several modalities: vestibular, proprioceptor and cutaneous somatosensory input from the trunk, together with visual and auditory input. The output from this part of the spinocerebellum goes by way of the fastigial nucleus to the vestibular nuclei (*Fig. 3*). It controls postural adjustments in response to sensory input by signaling to axial muscles via the medial motor systems. In primates, inactivation of the fastigial nucleus prevents standing and walking; the animals fall towards the side of the lesion.

In the intermediate zone regions of the spinocerebellum, input is from the cuneocerebellar and dorsal spinocerebellar tracts conveying proprioceptor and cutaneous somatosensory data and the ventral spinocerebellar tract which

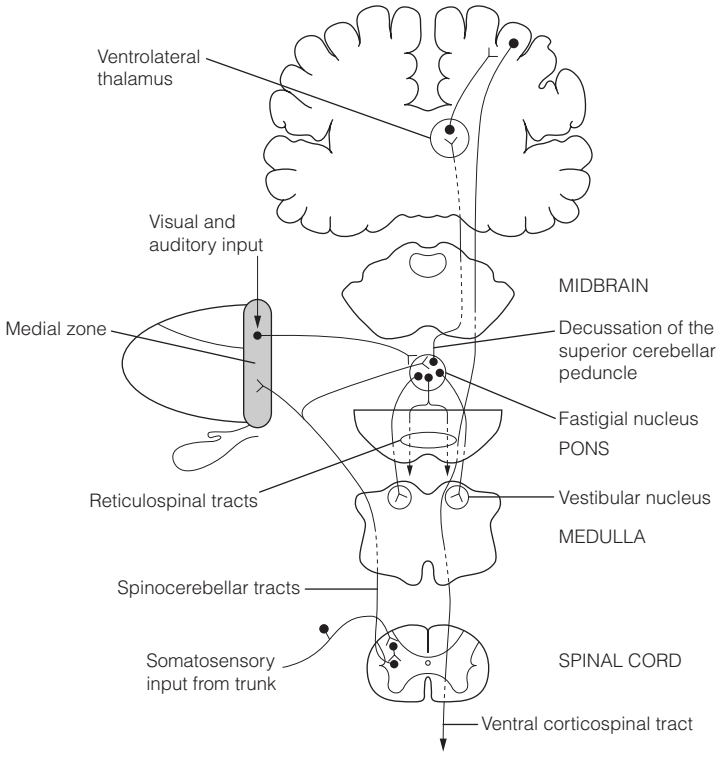


Fig. 3. Connections of the medial zone of the spinocerebellum. The output of the vestibular nuclei are as shown in Fig. 2. The inferior olivary nucleus is not shown.

imparts information about the activity of spinal motor circuits (*Fig. 4*). In addition, it receives inputs from the somatosensory and motor cortex via axon collaterals of the corticospinal tract that synapse with nuclei in the pons (**nuclei pontis**). These pontine nuclei give rise to fibers that cross the midline to enter the contralateral cerebellar cortex by way of the middle cerebellar peduncle. This **corticopontinecerebellar pathway**, containing 20 million axons is one of the largest tracts in the CNS, and also goes to the cerebrocerebellum.

The output of the intermediate zone region of the spinocerebellum is via the interpositus nucleus which projects to the ventrolateral thalamus and the red nucleus. By this route the spinocerebellum controls the lateral motor pathways to the limbs. Inactivation of the interpositus nuclei has little effect on standing or walking but results in a large-amplitude tremor of the limbs with a frequency of 3–5 Hz when an animal attempts to reach for an object. This is known as an **action** (**intention**) **tremor** and is seen often in humans with cerebellar damage.

Cerebrocerebellum The **cerebrocerebellum** corresponds roughly with the lateral zones of the posterior lobe. It receives inputs from frontal, parietal and occipital cerebral cortex relaying sensory, motor and visual information by way of the corticopontinecerebellar tract. However, it also gets input from prefrontal cortex concerned with cognitive *not* motor functions. The cerebrocerebellar outflow via the dentate nucleus goes to the ventrolateral thalamus which in turn projects to

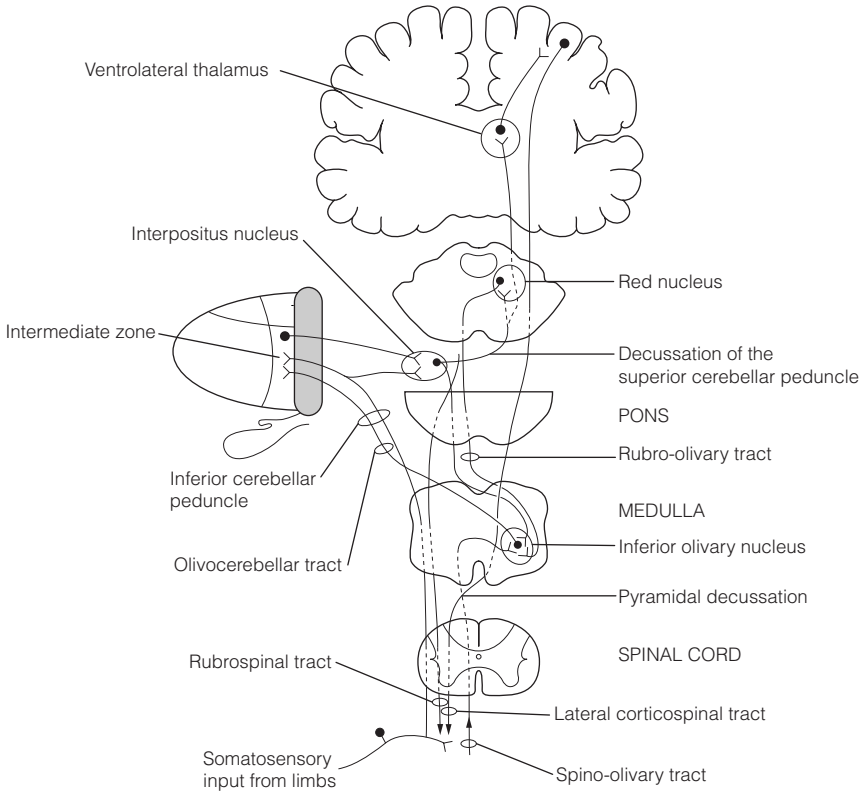


Fig. 4. Connections of the intermediate zone of the spinocerebellum. The corticopontine cerebellar tract is not shown.

frontal cortex motor and prefrontal areas (*Fig. 5*). Additionally the dentate nucleus has reciprocal connections with the red nucleus. Lesions of the cerebrocerebellar cortex or dentate nucleus cause slight delays and modest overshooting of movements involving single joints (**dysmetria**). However, for multijointed movements the deficits are much more severe, so animals and patients have difficulty using their fingers (**dysdiachokinesia**). Cognitive, including language, deficits are also seen in patients with cerebrocerebellar damage.

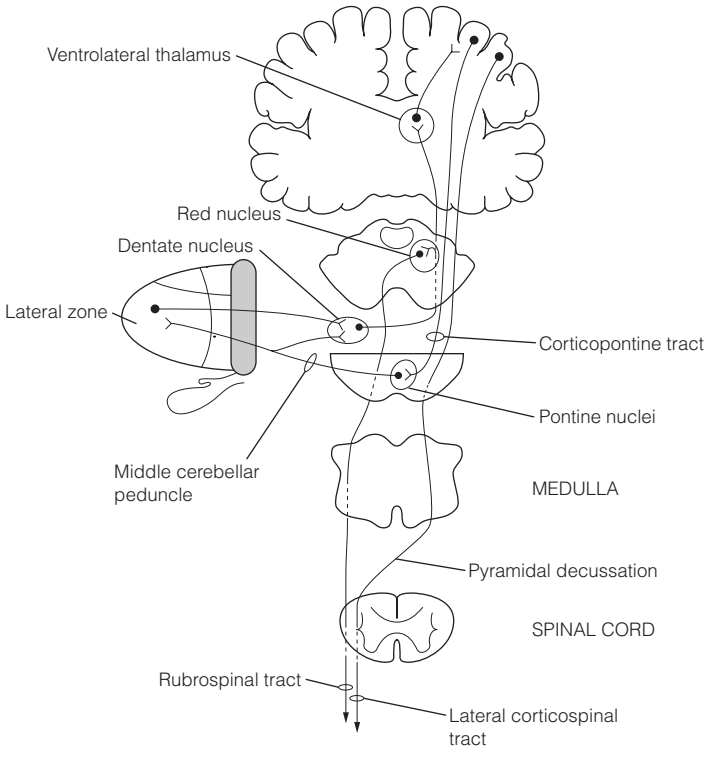


Fig. 5. Connections of the lateral zone (cerebrocerebellum). The connections of the inferior olivary nucleus are not shown.

K5 Cerebellar cortex circuitry

Key Notes

Inputs to the cerebellar cortex

Axons of spinal cord and brainstem neurons conveying sensory and motor information enter the cerebellum as mossy fibers to synapse with granule cells in multisynaptic complexes called glomeruli. Granule cell axons bifurcate into parallel fibers which form synapses with thousands of Purkinje cells aligned in a row. Each parallel fiber makes just one synapse with each Purkinje cell but every Purkinje cell is contacted by 200 000 parallel fibers. Parallel fibers excite Purkinje cells to fire simple spikes. Climbing fibers from the inferior olive wrap themselves around ten Purkinje cells making very powerful excitatory synapses with each one. Climbing fiber stimulation causes Purkinje cells to fire complex spikes.

Cerebellar cortex output

The output of the cerebellar cortex is provided exclusively by large GABAergic inhibitory Purkinje cells and goes to the deep cerebellar nuclei.

Cerebellar cortical interneurons

There are three types of GABAergic inhibitory neuron in the cerebellar cortex. Basket and stellate cells produce lateral inhibition by inhibiting those Purkinje cells that are the immediate neighbors of those activated by parallel fibers. Golgi II cells terminate the effect of parallel fibers on Purkinje cells.

Related topics

Cerebellar function (K6)
Cell physiology of learning (O3)
Cortical development (N3)

Inputs to the cerebellar cortex

The cerebellar cortex has three layers and contains five cell types that are organized into a simple circuit repeated millions of times (*Fig. 1*).

The major input to the cerebellum is **mossy fibers**, axons of second-order neurons from the spinal cord and brainstem conveying proprioceptor input, or the pontine–cerebellar relay from the cerebral cortex conveying sensory and motor signals. Each mossy fiber terminates in a discrete patch confined within a single lobule. Mossy fibers are glutamatergic and excitatory, and after giving off an axon collateral which goes to the appropriate deep cerebellar (or vestibular) nucleus, they synapse with granule cells in synaptic complexes called **glomeruli** (*Fig. 2*). Each glomerulus consists of the swollen terminal of a single mossy fiber which forms 15–20 synapses with the surrounding dendrites of 4–5 granule cells. Mossy fibers branch, so each one can excite about 30 granule cells. Every granule cell is contacted by 5–8 mossy fibers. Glomeruli also contain axodendritic synapses between Golgi cells (see below) and granule cells.

Granule cells in the **granule cell layer** are small (5–8 μm diameter) and one of the most numerous neuron types in the brain. It is estimated that there are

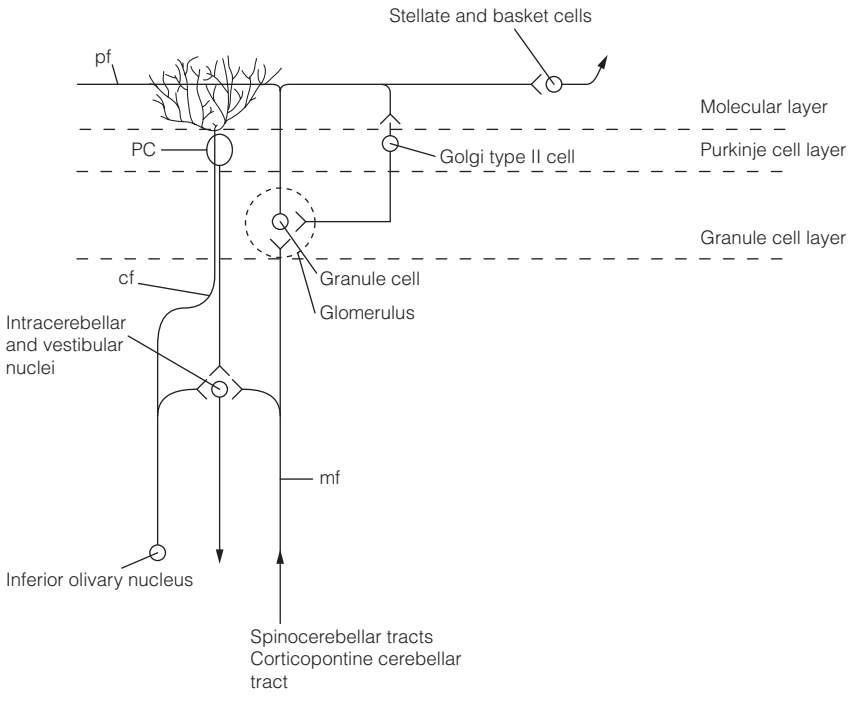


Fig. 1. The basic circuitry of the cerebellar cortex. Mossy fibers (mf) and climbing fibers (cf) are excitatory as are granule cells and the intracerebellar nuclei cells. All other cell types are inhibitory. Stellate and basket cells inhibit adjacent Purkinje cells (PCs).

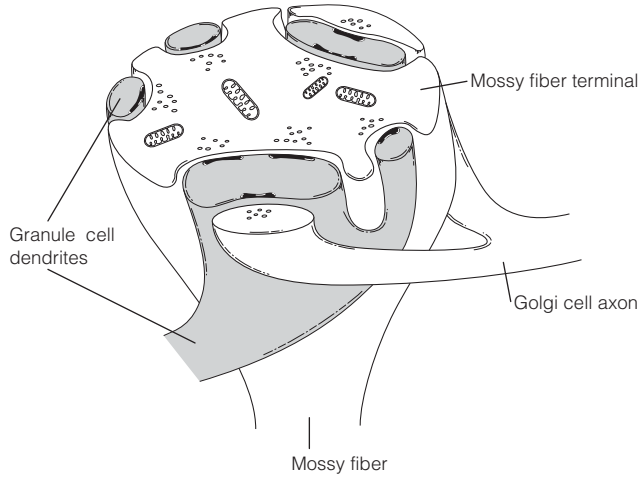


Fig. 2. The structure of a cerebellar glomerulus. All the synapses are axodendritic.

10^{11} granule cells in the human cerebellum. Granule cell axons ascend to the most superficial layer of the cerebellar cortex, **the molecular layer**, where they bifurcate into **parallel fibers** that in primates extend for 6 mm or so in each direction along the long axis of a folium. Parallel fibers intersect with the perpendicularly oriented planar dendritic trees of **Purkinje neurons**. Because of

this arrangement every parallel fiber excites a longitudinal beam of 2000–3000 Purkinje cells, making just a single synapse with each one. Every Purkinje cell is contacted by about 200 000 parallel fibers. Mossy fibers, and the granule cells driven by them have high background firing rates (50–100 Hz) that is changed by sensory input and during movements. The effect of parallel fiber activity is to cause the Purkinje cell to fire **simple spikes** repetitively (*Fig. 3a*). Purkinje cells fire at background rates of 20–50 Hz and even weak mossy fiber input produces increased Purkinje cell firing. Moreover, since Purkinje cells can fire in excess of 400 Hz they can follow a wide range in firing of mossy fiber inputs.

The second input to the cerebellum is the **climbing fibers** that come exclusively from the **inferior olivary nucleus** (**ION**) via the **olivocerebellar tract**. Each climbing fiber, of which there are about 15 million in humans, establishes contact with around ten Purkinje cells; each Purkinje cell getting input from just one climbing fiber that winds its way round the soma and dendrites making about 300 powerful glutamatergic, excitatory synapses. Climbing fibers fire with a frequency of 1–10 Hz, each time causing the Purkinje cell to discharge a **complex spike** (*Fig. 3b*).

A third diffuse set of inputs into the cerebellum comes from monoaminergic cells in the brainstem and includes noradrenergic neurons in the locus ceruleus, serotinergic cells in the raphe nuclei and cholinergic neurons in the reticular system. They establish sparse connections with deep cerebellar nuclei and cortex to produce modulating effects.

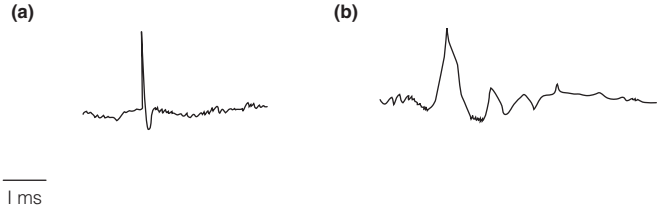


Fig. 3. Simple (a) and complex (b) spikes of Purkinje cells produced by mossy fiber or climbing fiber activation respectively.

Cerebellar cortex output

The sole output of the cerebellar cortex is via the Purkinje cells which have their large cell bodies (50 μm diameter) in the **Purkinje cell layer** of the cortex. Their extensive dendrites are aligned into a flat plane and are all oriented in the same direction, at right angles to the long axis of the folium in which they are located. The axons of the Purkinje cells go to the deep cerebellar nuclei. Purkinje cells use GABA as a neurotransmitter so the entire output of the cerebellar cortex is inhibitory.

Cerebellar cortical interneurons

The cerebellar cortex has three types of GABAergic inhibitory interneuron. **Basket cells** and **stellate cells** in the molecular layer receive input from parallel fibers and send axons at right angles to them to synapse with proximal or distal dendrites, respectively, of neighboring Purkinje cells. Activation of a mossy fiber excites a cluster of granule cells and hence stimulates linear arrays of **on-beam** Purkinje cells via the parallel fibers. However, the basket and stellate cells inhibit surrounding **off-beam** Purkinje cells. This is a surround antagonism mechanism which produces spatial focusing of cerebellar cortex output. It is akin to the lateral inhibition seen in sensory systems.

Golgi cells get input from parallel fibers and synapse with granule cells to produce feedback inhibition. In this way the Golgi cell brings about temporal focusing so that the net effect of mossy fiber input is *brief* firing of Purkinje cells.

In summary, the inhibitory interneurons constrain Purkinje cell output both in space and time.

K6 Cerebellar function

Key Notes

General principles

The cerebellum coordinates movements initiated from the motor cortex, can initiate movements itself, and learn new motor tasks. Parallel fibers excite arrays of Purkinje cells that control muscles spanning several joints. This supports the idea that the cerebellum is important for complicated multijoint movements. The cerebellum can operate in feedback or feedforward mode.

Feedback operation

Here the cerebellum compares motor intentions with motor performance. A discrepancy between these generates an error signal which is used to make the mismatch smaller. Sensory signals caused by movement errors activate mossy fibers, which excite Purkinje cells. These inhibit the deep cerebellar nuclei which drive the red nucleus and thalamus and so the erroneous movement is prevented.

Feedforward operation

For movements that are too fast for feedback to operate, the cerebellum runs pre-programmed sequences that have predictable effects on motor function. This feedforward works well provided nothing unexpected happens.

Motor learning

In humans most voluntary movements must be learnt. In motor learning the cerebellum acquires a program which specifies the motor commands needed for a given movement by trial and error. Sensory errors are translated into motor errors by the inferior olive and sent to the cerebellum via climbing fibers. These motor error signals cause the Purkinje cells to become less responsive to the mossy fiber input occurring at the same time. Whenever the same input recurs the Purkinje cells are excited less than before the learning.

Related topics

Spinal motor function (J4)
Cortical control of voluntary movement (K1)
Subdivisions of the cerebellum (K4)
Cerebellar cortical circuitry (K5)
Cell physiology of learning (O3)

General principles

Despite the functional subdivisions, the same circuit is repeated across the entire cerebellum, so it is likely that the same computations are performed by all parts of the cerebellum. Although the cerebellum is usually regarded as coordinating the execution of movements initiated by the motor cortex, the cerebrocerebellum can *initiate* movements, particularly in response to visual and auditory stimulation. In movements triggered in this way the order of activation is: dentate nucleus – motor cortex – interpositus nucleus – muscle. Furthermore, central to the operation of the cerebellum is motor learning, the acquisition of new motor skills.

In primates, parallel fibers average 6 mm in length and so affect a comparable length array of Purkinje cells (PCs) that lie across the cerebellum. This is

sufficiently long to span an entire deep cerebellar nucleus or bridge adjacent nuclei. PC arrays coupling both fastigial nuclei, for example, would ensure coordination of postural muscles across the midline, which is important in gait. The PC arrays influenced by a given set of parallel fibers span muscles over several joints. This anatomical configuration supports the results of recording and lesion studies showing that the cerebellum is much more concerned with control of movements involving many joints rather than single joints.

The cerebellum is thought to operate in one of two modes, feedback or feedforward, depending on the circumstances.

Feedback operation

During the execution of well-rehearsed movements that are not too fast the cerebellum acts as a feedback device to compare motor intentions with motor performance, and works to reduce any mismatch between them. For the spinocerebellum, the motor intentions are the signals relayed by the corticopontinecerebellar tract. Motor performance is monitored by proprioceptor (and other sensory) input, and by the ventral spinocerebellar signals reporting on the activity of spinal cord and brainstem motor circuits. Similarly, the cerebrocerebellum compares inputs from the supplementary motor cortex and the primary motor to produce error signals that reflect a discrepancy between motor planning and motor commands. In each case the error signals are used to correct the mismatch.

The intermediate spinocerebellum seems to be involved in correcting errors in limb movements, such as when a limb is perturbed by an unexpected force. The order in which various neural elements fire is: muscle afferents – interpositus nucleus – motor cortex – dentate nucleus.

Feedback error correction probably works as follows. An error means that the actual position of a limb is not the intended one. This produces unpredicted muscle stretch. Precisely the same thing happens if an unexpected force is applied to a limb. In either case the muscle stretch will excite Ia and Ib afferents in the loaded muscles. These proprioceptor signals are relayed to the cerebellum by mossy fibers. The stimulated mossy fibers deliver tonic excitation to the intracerebellar (interpositus) nucleus via axon collaterals and stimulate a group of granule cells. The granule cell parallel fibers activate several arrays of on-beam Purkinje cells (*Fig. 1*), which strongly inhibit their target neurons in the interpositus nucleus. Normally, when mossy fibers are firing at background rates their tonic facilitation of the interpositus dominates the inhibition by

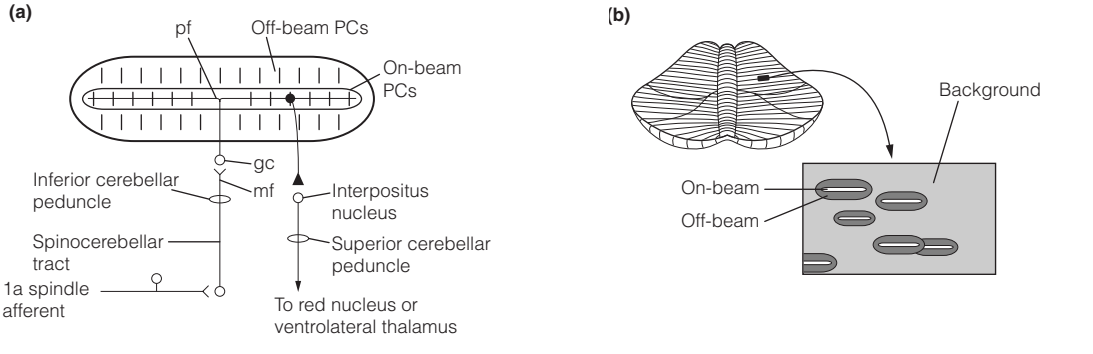


Fig. 1. (a) Mossy fiber (mf) input activates an array on on-beam Purkinje cells (PCs). Each short vertical line represents the planar dendritic field of a PC, viewed from above. The output of only one PC is shown. (b) Pattern of activation produced by single mossy fiber input. gc, Granule cell; pf, parallel fiber.

Purkinje cells. Consequently interpositus neurons maintain excitation of the red nucleus and the ventrolateral thalamus. When the mossy fibers are activated during a movement however, Purkinje cell inhibition **disfacilitates** the interpositus neurons and this inhibition is transmitted downstream to the red nucleus and thalamus. In contrast, neighboring off-beam PC arrays, inhibited by the GABAergic interneurons in the cortex allow *their* interpositus cells to fire at higher than background rates. Thus, the pattern of activation of the deep cerebellar nucleus is a negative image of the input activation.

The overall effect, mediated by the rubrospinal and corticospinal tracts, is to correct the movement error by activating spinal reflexes that defend the correct limb position and dampen those that do not. There is some evidence that this might involve altering the firing of fusimotor efferents to muscle spindles.

The intermediate spinocerebellum seems to control the precise timing of the contraction of agonist and antagonist muscles during a movement. During reciprocal activation of agonist and antagonist muscles, Purkinje cells responsible for controlling these muscles fire alternately, driving interpositus neurons to do the same. During co-contraction, however, the Purkinje cells are silent. A role for the cerebellum in organizing these patterns of muscle activity is supported by the fact that the action tremor resulting from lesions of the interpositus nucleus or intermediate cerebellar cortex appears to be due to derangement in the timing of agonist contraction. In normal humans, a rapid wrist movement involves an initial burst of activity in the agonist, followed by a burst in the antagonist to produce braking, and finally a second agonist burst to stabilize the joint at the desired end point (refer to *Fig. 2* of Topic J4 for an illustration of this pattern of activity). In cerebellar tremor the start of the movement is normal but the second agonist burst is late. Consequently the antagonist burst moves the wrist beyond the end point. This sets up the tremor.

Feedforward operation

For the execution of well-practiced, very rapid movements (e.g. playing fast passages on a musical instrument or a tennis serve), there is insufficient time for feedback correction of errors. For these movements the cerebellum operates in a feedforward mode in which it runs a program that predicts the motor consequences of its own action. *Unexpected* perturbations that occur when the cerebellum is in this mode cannot be corrected for in time and so performance will be degraded.

Motor learning

The predictions inherent to feedforward operation must be learnt during numerous trials attempting to perform the task. This is motor learning and is probably important in acquiring skill in all voluntary motor tasks, including (in humans) learning to walk.

In one model of motor learning (the **cerebellar feedback-error-learning model**) the cerebellum acquires a program called an **inverse model** of the motor task, which is the transformation from the desired trajectory to the motor commands needed to bring about the movement (*Fig. 2*). Errors in a movement

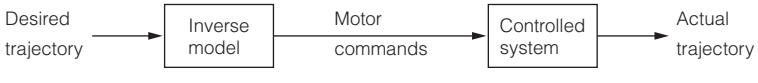


Fig. 2. The operation of the inverse model on the controlled system (motor neurons and muscles) turns a desired into an actual trajectory. The better the inverse model the closer the actual is to the desired trajectory.

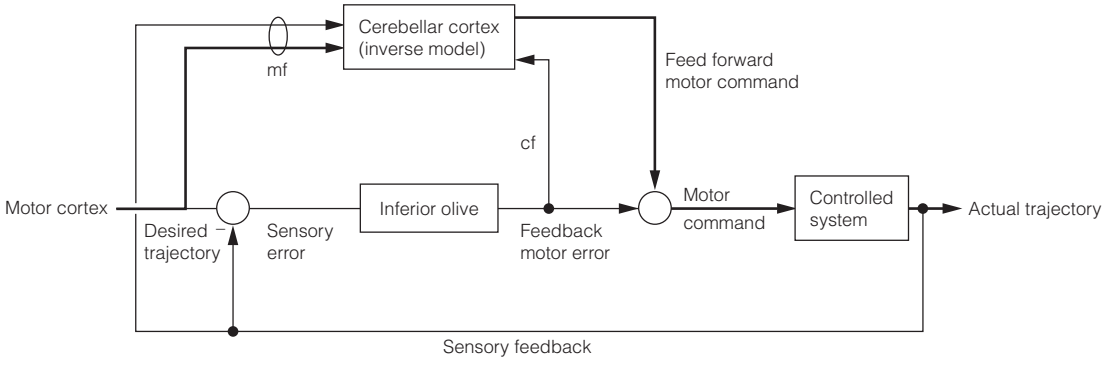


Fig. 3. A model for motor learning by the cerebellum. When the cerebellum is operating in feedforward mode the bold pathway is activated. cf, climbing fibers; mf, mossy fibers.

are initially represented as sensory errors. For example, during a tennis match an incorrect arm movement will be visually obvious from the direction of the ball, and an error in playing a musical instrument will sound wrong. Sensory errors need to be converted into a pattern of nerve impulses that specify errors in motor performance, an operation of the **inferior olivary nucleus** (**inferior olive**). The inferior olive sends these motor error signals to the cerebellum via the olivocerebellar tract climbing fibers (*Fig. 3*). Hence, the cerebellum gets mossy fiber input that represents the desired trajectory (from motor cortex), sensory input (e.g. from visual cortex or proprioceptor pathways) and feedback from the inferior olive that represent errors in motor performance.

The error signals carried by the climbing fibers cause concurrent mossy fiber input to become less effective in activating Purkinje cells. Whenever the same pattern of mossy fiber input occurs subsequently, the Purkinje cells fire fewer simple spikes and cause less inhibition on downstream motor pathways. This corrects the motor output. The cellular events that alter the responsiveness of the Purkinje cells are called long-term depression.

K7 Anatomy of the basal ganglia

Key Notes

Overview

Several interconnected structures make up the basal ganglia: striatum, globus pallidus, substantia nigra and subthalamus. Cerebral cortical input to the basal ganglia goes to the striatum. The basal ganglia output goes from the globus pallidus and substantia nigra to the cortex via the thalamus. The basal ganglia are responsible for producing motor sequences during voluntary movement.

Striatum

The caudate nucleus and putamen together make up the striatum. Glutamatergic axons from the cerebral cortex and dopaminergic axons from the substantia nigra (SNpc) terminate on the medium spiny neurons which are GABAergic. There are two populations of medium spiny neurons, one is excited by dopamine, the other inhibited. The inhibitory output of the striatum goes to the globus pallidus and substantia nigra.

Output structures of the basal ganglia

Parts of the globus pallidus pars interna (GPi) and substantia nigra (SNpr) send axons to specific thalamic nuclei which in turn project to particular areas of the cerebral cortex. This circuitry provides basal ganglia control of limb, facial and eye movements. Part of the globus pallidus pars externa (GPe) projects to the subthalamic nucleus.

Subthalamic nucleus

Excitatory neurons of this nucleus are excited by the motor cortex, inhibited by the GPe and send their axons to the globus pallidus and substantia nigra.

Parallel processing in the basal ganglia

Five circuits form loops between specific regions of the cortex and basal ganglia and each circuit seems to have a distinct functional role. Two of the circuits are concerned with motor functions, the others with aspects of memory, cognition and emotion.

Related topics Dopamine (D4) Cortical control of voluntary movement (K1)

Overview

The basal ganglia consist of several extensively interconnected structures, the striatum (the caudate and putamen), the globus pallidus (pars interna and pars externa), the substantia nigra (pars compacta and pars reticulata) and the subthalamic nucleus. Most inputs to the basal ganglia are from the cerebral cortex and enter the striatum. The output of the basal ganglia emerges from the pars interna (internal segment) of the globus pallidus, and the substantia nigra pars reticulata, to go to the thalamus. The thalamus projects back to the cortex thus closing a loop. The thalamocortical axons return to the same region of cortex which gave rise to the striatal inputs (*Fig. 1*).

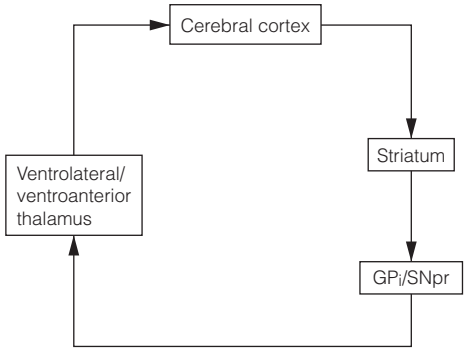


Fig. 1. Block diagram of the interface between basal ganglia and cerebral cortex. GPi, globus pallidus pars interna; SNpr, substantia nigra pars reticulata.

Basal ganglia circuitry is responsible for the execution of appropriate preprogrammed motor sequences during voluntary movements. Firing of many neurons in the basal ganglia correlates with movement but occurs quite late, so the basal ganglia are not implicated in the initiation of movement. Classically the basal ganglia constitute part of the **extrapyramidal system**, on the basis that lesions of the basal ganglia produce quite different symptoms from lesions of the pyramidal (corticospinal) tract. Recently it has been proposed that the basal ganglia also has cognitive and affective functions.

Striatum

The caudate nucleus and putamen are functionally a single unit, the **dorsal striatum** (neostriatum), but split anatomically by the internal capsule. Although distinct from the ventral striatum, which is part of the limbic system, the two have similar circuitry.

The striatum receives excitatory input from the cortex via the glutamatergic corticostriate pathway (*Fig. 2*). The input is organized topographically so that somatotopy is preserved in the projections from the somatosensory cortex and motor cortex. Corticostriate axons terminate on the major neuron type in the striatum, the **medium spiny neuron**. These make up 95% of striatal neurons, use GABA as their transmitter, and provide the inhibitory output of the striatum. There are two populations of medium spiny neuron which although morphologically indistinguishable have different connections and neurochemistry. One type has **substance P** (**SP**) and **dynorphin** (**DYN**) as cotransmitters, express dopamine D1 receptors, and projects to the globus pallidus pars interna (GPi) and **substantia nigra pars reticulata** (**SNpr**). The second uses **enkephalin** (**ENK**) as a cotransmitter, express D2 dopamine receptors, and projects to the globus pallidus pars externa (GPe).

The medium spiny neurons receive projections via the nigrostriatal pathway from the **substantia nigra pars compacta** (**SNpc**). The SNpc uses dopamine as a transmitter. Because the two types of medium spiny neuron express different dopamine receptors they are differently modulated by this input. At the GABA/SP/DYN cells, dopamine acting on D1 receptors enhances the effect of excitatory cortical input. In contrast, the action of dopamine on D2 receptors on GABA/ENK cells is to reduce the effect of cortical excitation. Medium spiny neurons have a third input from large **aspiny interneurons** that constitute about

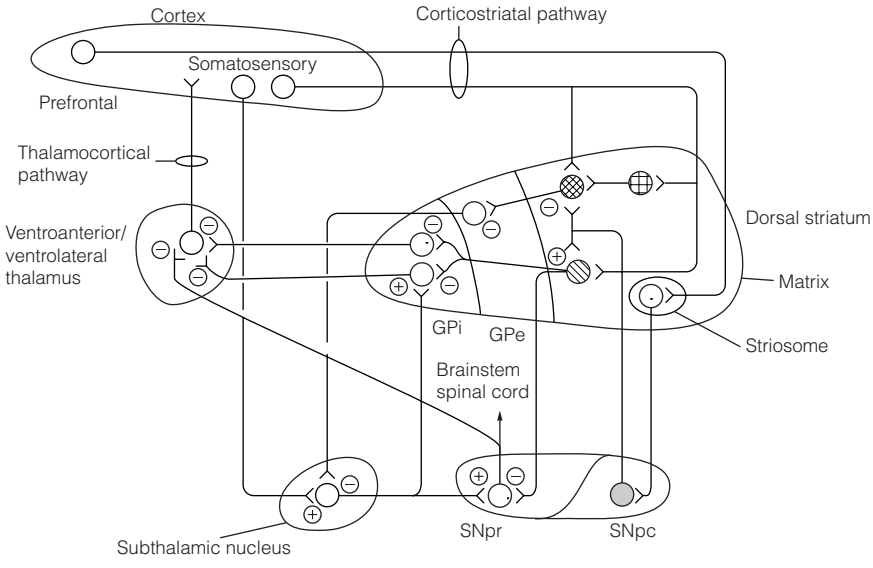


Fig. 2. Connections of the basal ganglia. GABAergic neurons in GPe also project to GPi and SNpr (not shown). GPe, globus pallidus pars externa; GPi, globus pallidus pars interna; SNpc, substantia nigra pars compacta; SNpr, substantia nigra pars reticulata. ⊕, excitatory synapse; ⊖, inhibitory synapse. The neurons are coded by their neurotransmitter: ◯, glutamate; ⊕, acetylcholine; ●, dopamine; ⊖, GABA; ▨, GABA/substance P/dynorphin; ▩, GABA/enkephalin.

2% of striatal neurons. These cells use ACh as a transmitter, are excitatory, and driven by cortical inputs.

Staining of the striatum for acetylcholinesterase shows it to be compartmented into a heavily stained **matrix** and a lightly stained three-dimensional labyrinth, the **striosomes**, that are about 10–20% of the striatal bulk. The connectivity of cells in these two compartments is different. The matrix gets inputs from throughout the cerebral cortex and sends outputs to the entire globus pallidus and SNpr, whereas the striosomes get restricted input from the prefrontal cortex and project to the SNpc. The matrix is concerned with sensorimotor function, while striosomes are associated with the limbic system, and may control the dopaminergic pathway from SNpc to striatum.

Output structures of the basal ganglia

Both the globus pallidus and substantia nigra are divided into two parts. The **globus pallidus pars interna** (**GPi**) of primates (equivalent to the **entopeduncular nucleus** in rodents) and the substantia nigra pars reticulata (SNpr) have very similar structures and are functionally equivalent. Both get inhibitory connections from the GABA/SP/DYN population of striatal neurons and excitatory inputs from the subthalamic nucleus, and both send GABAergic inhibitory outputs to the thalamus. The thalamus in turn projects to specific locations in the cerebral cortex. The GPi and SNpr make connections with several thalamic nuclei that project to motor cortex, providing basal ganglia output for limb and facial movements. Part of the SNpr is concerned with eye movements.

The **globus pallidus pars externa** (**GPe**) receives its striatal connections from

the GABA/ENK medium spiny neurons. The GPe neurons are GABAergic and project mostly to the **subthalamic nucleus**.

Subthalamic nucleus

The subthalamic nucleus (**STN**) lies at the junction between midbrain and diencephalon, and is particularly well developed in primates. It gets excitatory input from the motor cortex and GABAergic input from the GPe. The STN neurons use glutamate as a transmitter, are excitatory, and send their axons principally to the GPi and SNpr.

A good somatotopic representation is present in all structures of the basal ganglia, except the pars compacta of the substantia nigra. The output of the basal ganglia to the thalamus establishes somatotopic maps there, but they are distinct from the somatotopic maps in the thalamus produced by the cerebellum. In other words inputs of the basal ganglia and cerebellum to the thalamus are kept separate.

Parallel processing in the basal ganglia

Five parallel circuits like that illustrated in *Fig. 1* are thought to exist in the basal ganglia, each one getting corticostriatal inputs from several functionally related areas of cortex and projecting back to a restricted locus of the same area by way of specific thalamic nuclei. They are referred to as **basal ganglia–thalamocortical circuits** (*Table 1*). A massive convergence occurs between the striatum and GPi/SNpr. This funneling of input from a large to a restricted region in the cortex by the basal ganglia serves as a way of integrating information.

Table 1. Parallel basal ganglia-thalamocortical circuits

Component	Circuit				
	Motor	Oculomotor	Executive	Limbic	Orbitofrontal
Cortical input	Pre-motor, motor, somatosensory	Prefrontal, posterior parietal	Prefrontal, pre-motor, posterior parietal	Medial orbitofrontal, hippocampus, temporal	Temporal, anterior cingulate
Striatal target	Putamen	Caudate	Caudate	Nucleus accumbens	Caudate
Striatal output	GPi/SNpr, GPe[a]	GPi/SNpr, GPe	GPi/SNpr, GPe	Ventral pallidum	GPi/SNpr, GPe
Thalamic nuclei	Ventral lateral	Ventral anterior, medial dorsal	Ventral lateral, medial dorsal	Medial dorsal	Ventral anterior, medial dorsal
Cortical output	Supplementary motor area	Frontal eye fields	Dorsolateral prefrontal	Anterior cingulate	Lateral orbitofrontal
Functions	Motor	Eye movements	Motor planning, executive, cognitive	Motivation, motor expression of emotions	Motivation, social responses
Lesions	Hypo- and hyper-kinetic disorders	Oculomotor hypokinesia	Deficits in executive functions	Akinetic mutism	Defects in social behavior, obsessive-compulsive disorder

[a] GPi, globus pallidus pars interna; GPe, globus pallidus pars externa; SNpr, substantia nigra pars reticulata.

Of the five, only the motor and oculomotor circuits serve the functions classically attributed to the basal ganglia, the others, dorsolateral prefrontal, orbitofrontal and anterior cingulate, are associated with memory, cognition and emotion, respectively. The anterior cingulate circuit, unlike the others, goes by way of the ventral striatum (nucleus accumbens) which is part of the dopamine motivation system, rather than the caudate or putamen. Since the circuits are all wired in much the same way it is likely that they all perform the same computations. The different outcomes of the operation of each of the circuits depends on the cortical regions they are connected to, and the contexts in which they are activated. Interestingly, basal ganglia disorders can produce deficits in thought processes not unlike those that affect movement.

K8 Basal ganglia function

Key Notes

Direct and indirect pathways

The basal ganglia do not initiate movement but seem to be important in movement associated with a reward. There are two pathways in basal ganglia circuits with opposing effects on firing of thalamic and cortical neurons. The direct pathway activates thalamic neurons and this allows movement sequences to occur. The indirect pathway inhibits thalamic neurons and suppresses unwanted movement. Both of these pathways are activated when the motor cortex initiates a specific movement.

Dopamine modulation

These two pathways are modulated by dopaminergic axons running from the substantia nigra to the striatum. Activity in this nigrostriatal tract enhances the direct pathway but suppresses the indirect pathway. Hence dopamine neurotransmission enables movements to occur.

Basal ganglia operation

The group of cells that represents a given movement sequence are normally inhibited by tonic discharge from the GPe and SNpr. Executing a given movement needs activation of the direct pathway which inhibits the GPe and SNpr.

Basal ganglia disorders

Motor disorders due to malfunction of the basal ganglia are of two types. Hyperkinesias feature motor overactivity and include Huntington's disease, a genetic disorder in which the GABAergic medium spiny neurons of the striatum controlling the indirect pathway die. Hypokinesias are conditions in which motor activity is reduced. The most common, characterized by rigidity, slowness of movement and tremor is Parkinson's disease. Obsessive–compulsive disorder is characterized by an inability to stop repeating the same actions or thoughts, even when they are recognized to be unnecessary. Its cause may lie with reduced activity in the orbitofrontal basal ganglia circuit.

Related topics Anatomy of the basal ganglia (K7) Parkinson's disease (P5)

Direct and indirect pathways

An important feature for understanding the role of the basal ganglia in movement is the presence of two routes through the basal ganglia circuitry with opposite effects on firing of thalamic, and hence cortical neurons (*Fig. 1*). The **direct pathway** uses the GABA/SP/DYN medium spiny striatal neurons which inhibits GABAergic outflow of the GPi and SNpr to the thalamus. Cortical activation of this pathway *increases* the firing of thalamic neurons (since inhibiting an inhibition is excitation).

The **indirect pathway** starts with the GABA/ENK medium spiny neuron output to the GPe, inhibitory neurons from which go to the STN. The STN excites inhibitory neurons in the GPi and SNpr that go to the thalamus. Corticostriate activation of the indirect pathway results in *decreased* firing of

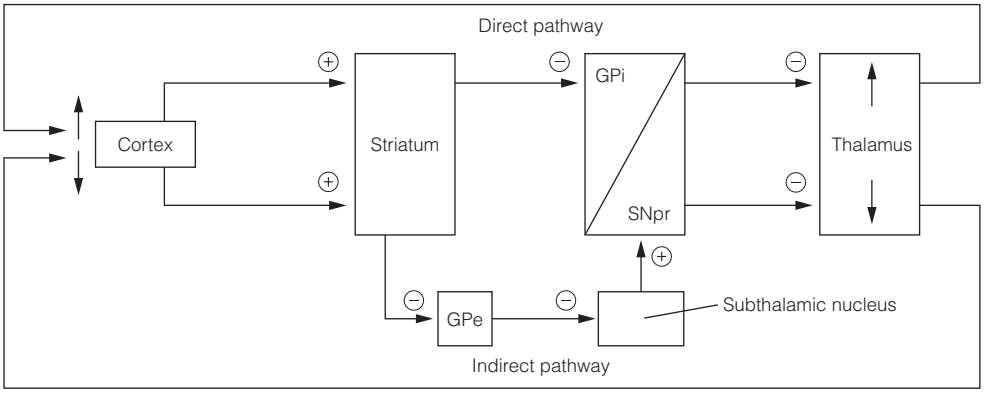


Fig. 1. Direct and indirect pathway activation causes increased (↑) or decreased (↓) firing of thalamic (and cortical) neurons respectively. GPe, globus pallidus pars externa; GPi, globus pallidus pars interna; SNpr, substantia nigra pars reticulata.

thalamic neurons. This dual circuitry allows the possibility that given movement sequences may be triggered or suppressed by differential activation of direct or indirect pathways respectively.

Dopamine modulation

The dopaminergic neurons of the substantia nigra pars compacta (SNpc) alter their firing pattern in response to stimuli that reward a movement. They modulate the response of medium spiny neurons in the striatum to corticostriate inputs but have opposite effects on the two populations. While the GABA/SP/DYN neurons are made more excitable, the GABA/ENK cells become less excitable in the face of SNpc inputs. In summary, the direct pathway is enhanced by and the indirect pathway suppressed by the nigrostriatal tract from the SNpc.

Basal ganglia operation

One function of the basal ganglia is to enable the execution of motor sequences. Each sequence is represented by an array of cells, a micro-loop, within the basal ganglia–thalamocortical motor or oculomotor circuit and can be either activated or (if unwanted) inhibited. While some sequences are stereotyped movements, the circuitry for which is genetically specified, many sequences are learnt; that is, many micro-loops are entrained by experience.

At rest, most medium striatal neurons fire at low frequencies (0.1–1 Hz), while GPi and SNpr neurons have high background firing rates (about 100 Hz). The current model is that movements are initiated by activity in the motor cortex, which is relayed to the striatum. During a movement, striatal neurons increase firing, as a result of elevated activity of the corticostriatal neurons that drive them. Tonic inhibitory output of the GPi and SNpr at rest, which is increased about 50 ms before a movement by excitatory drive from the subthalamus, is due to the operation of the indirect pathway and results in widespread, and complete, suppression of unwanted movement sequences. Making a particular movement requires that the direct striatopallidal pathway to the GPi/SNpr cells that enable the movement (those belonging to the correct micro-loop) become activated. These GPi/SNpr cells reduce their firing, releasing their corresponding thalamocortical cells from inhibition. The nigrostriatal dopamine system acts to raise the likelihood that a movement sequence is actually made.

The similarity of basal ganglia–thalamocortical circuits implies that they all perform the same computation. If this is so, it seems likely that the cognitive circuits may operate to select particular behavioral sequences in a manner that is appropriate to the context. In the light of this it is interesting that lesions of the orbitofrontal cortex result in perseveration in which a behavior is continued long after it is needed or appropriate.

Basal ganglia disorders

Motor disorders arising from dysfunction of the basal ganglia, whether caused by disease, or by lesions in animal studies, fall into two distinct categories, hyperkinesias and hypokinesias.

1. **Hyperkinesias**. These are disorders in which motor overactivity occurs. Characteristically these disorders consist of frequent, random, twitch-like or writhing movements, resembling fragments of normal movements, termed **choreoathetosis**. It is the principal symptom of Huntington's disease, and of **tardive dyskinesia**, an unwanted effect of the treatment of Parkinson's disease with L-DOPA, or infarcts of the subthalamic nucleus. **Huntington's disease** (**HD**) is a progressive neurodegenerative disorder in which symptoms (cognitive as well as motor) begin between 40 and 50 years of age. It is an autosomal dominant disease caused by an abnormality in a gene on chromosome 4 which codes for a widely distributed protein, **huntingtin**, the function of which is not known. The abnormality is an excessive number of trinucleotide (CAG) repeats which code for a string of glutamine residues near the N terminus of the protein. This causes huntingtin molecules to form aggregates in the nuclei of specific neurons. Particularly afflicted are the GABA/ENK medium spiny neurons of the striatum. Their death causes abnormal inhibition of the subthalamic nucleus, so increased and inappropriate firing of thalamocortical neurons. In summary, the chorea is a failure of the indirect pathway to block unwanted movement sequences. Infarcts of the subthalamic nucleus (STN) are rare and cause a greatly exaggerated chorea characterized by large flinging movements of limbs contralateral to the lesion, termed **hemiballismus.** Loss of the tonic excitatory drive from the STN is thought to cause GABAergic neurons in the GPi and SNpr to go into a burst firing mode. Why this happens is not known, neither is it clear what adaptation allows the recovery from hemiballismus that is seen after a few weeks both in humans and in STN-lesioned animals. **Tics** are hyperkinesias in which highly stereotyped and sometimes quite complex movements of face and hands appear. Sometimes tics are associated with behavioral disorders, such as in the rare **Gilles de la Tourette** syndrome where they are accompanied by involuntary uttering of sexual obscenities.
2. **Hypokinesias**. These are disorders in which motor activity is reduced. In animals lesions of the globus pallidus result in abnormal co-contraction of agonist and antagonist muscles in the contralateral limbs. The effect of this is to raise the stiffness of joints causing **rigidity** and slowness of movement, **bradykinesia**. Bilateral lesions result in animals adopting abnormal flexed postures that they seem unable to move out of. This resembles **dystonias** seen in humans in a variety of conditions including end-stage Huntington's disease, strokes, and an unwanted effect of treatment with dopamine D2 receptors such as metoclopramide. The prototypical hypokinetic disorder is **Parkinson's disease**, characterized by rigidity, bradykinesia and tremor. Its pathology and treatment are dealt with in Topic P5.

Obsessive–compulsive disorder (**OCD**) is a chronic psychiatric disorder in which a person is unable to prevent themselves endlessly repeating the same actions or thoughts. The afflicted individual may spend many hours each day acting out pointless rituals, such as hand washing because of an obsessional fear of contamination, or having to check that the front door is locked a set number of times before going out. As with other neuroses, it represents an exaggeration of normal behavior, and in OCDs this appears as excessive perseveration. Brain-imaging studies show a reduction in cerebral blood flow in the orbitofrontal cortex that correlates with the severity of the disorder in patients. Lesions of the orbitofrontal cortex in primates causes perseveration. This suggests that the cause of OCD may lie with dysfunction of the orbitofrontal basal ganglia circuit.

L1 Anatomy and connections of the hypothalamus

Key Notes

Hypothalamic anatomy

The hypothalamus is a collection of nuclei implicated in sleep, appetitive behaviors, and the control of autonomic and endocrine functions which are largely the responsibility of the paraventricular nuclei (PVN). The PVN contain large neurons that project into the posterior pituitary, via the pituitary stalk, and small neurons which have their terminals in the median eminence which lies immediately above the pituitary stalk. The hypothalamus is divided into three zones in the mediolateral direction and four subdivisions along the rostro-caudal axis.

Connections of the hypothalamus

The hypothalamus is part of the limbic system. It gets input from the hippocampus via the fornix, fibers of which go largely to the mammillary bodies. The output of the mammillary bodies goes to the anterior thalamic nuclei. The anterior thalamus projects to the cingulate cortex which, in turn, sends output to the hippocampus, thus closing a loop, the Papez circuit. Input from the amygdala arrives by two routes, the stria terminalis and the amygdalofugal pathway. The medial forebrain bundle, which consists of monoaminergic axons, passes through the hypothalamus making many connections *en route*.

Pituitary gland

The pituitary gland consists of a posterior lobe, which is a direct outgrowth of the hypothalamus, and an anterior lobe. It is connected to the base of the brain by the pituitary stalk. Large neurosecretory neurons in the hypothalamus send axons into the posterior lobe and release hormones there. Hormones secreted into the median eminence by small hypothalamic cells are, in contrast, carried in a vascular network, the hypothalamic–pituitary portal circulation, to the anterior lobe.

Related topics

- Organization of the central nervous system (A5)
- Posterior pituitary function (L2)
- Neuroendocrine control of metabolism and growth (L3)
- Neuroendocrine control of reproduction (L4)
- Control of feeding (M3)
- Sleep (M5)

Hypothalamic anatomy

The hypothalamus is involved in the control of a variety of functions: sleep–wakefulness, thermoregulation, feeding and the control of metabolic energy expenditure, drinking and fluid homeostasis, growth and reproduction.

Some of these functions are brought about by the hypothalamus acting to control the autonomic nervous system or the output of hormones from the pituitary gland.

The hypothalamus consists of many nuclei clustered around the third ventricle (*Fig. 1*). At its rostral end lies the optic chiasm. The most caudal part of the hypothalamus are the paired mammillary bodies. The floor of the third ventricle is a sheet of gray matter extending from the optic chiasm to the mammillary bodies called the **tuber cinereum**. At its anterior end this thickens to become the **median eminence**, which projects as the **infundibular stalk** (part of the pituitary stalk) to the posterior pituitary gland. The hypothalamus is divided into three longitudinal zones, the **periventricular zone**, which surrounds the lining of the third ventricle, the **medial (intermediate) zone**, and the **lateral zone**. It also has four subdivisions along its rostro-caudal axis,

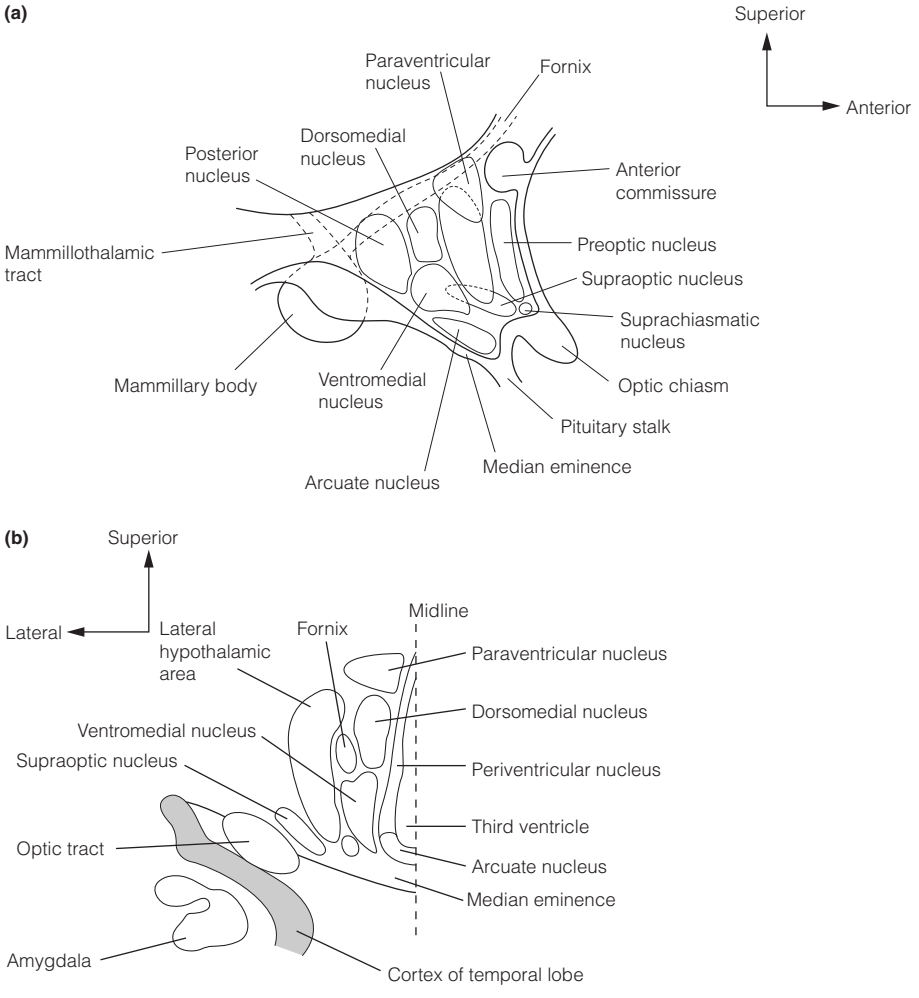


Fig. 1. Human hypothalamus of the left cerebral hemisphere shown diagrammatically: (a) midsagittal section; (b) coronal section.

preoptic, **anterior**, **tuberal** and **mammillary**. The locations of major nuclei are specified in *Table 1.* Many of the endocrine and autonomic functions of the hypothalamus involve the paraventricular nuclei. They contain several populations of neuroendocrine cells, each secreting specific peptides, but they fall into two groups. Magnocellular (large) cells which project to the posterior pituitary and parvocellular (small) cells that project to the median eminence.

Table 1. Location of some hypothalamic nuclei

	Zone		
Subdivision	Periventricular	Medial	Lateral
Preoptic		Medial preoptic nucleus	Lateral preoptic nucleus
Anterior	Suprachiasmatic nucleus Paraventricular nucleus Anterior periventricular nucleus	Anterior nucleus	Supraoptic nucleus
Tuberal	Arcuate nucleus	Ventromedial nucleus Dorsomedial nucleus	Lateral hypothalamic area
Mammillary	Posterior hypothalamic nucleus	Medial mammillary nuclei[a]	Lateral hypothalamic area Lateral mammillary nuclei[a]

[a] Mammillary bodies.

Connections of the hypothalamus

The hypothalamus is connected with limbic structures that are concerned with emotion and its expression (*Fig. 2*). It gets input from the hippocampus:

- by way of the **subiculum** (a region of transitional cortex), via the **postcommisural fornix**, which projects mainly to the mammillary bodies;
- by way of the **septum**, via the **precommisural fornix**, which makes connections with all three zones of the hypothalamus.

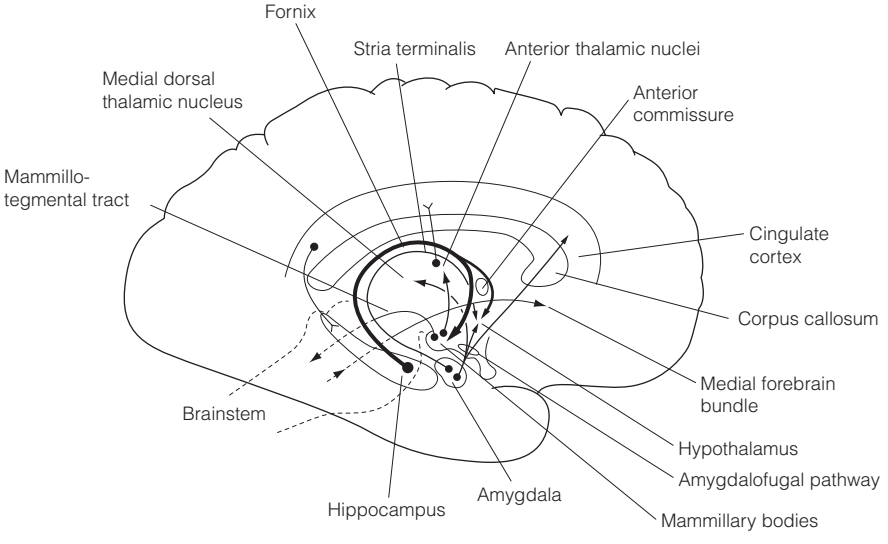


Fig. 2. Major connections of the human hypothalamus.

Input from the amygdala arrives at the hypothalamus via the **stria terminalis**, a loop that follows a similar course to the fornix, and the **amygdalofugal pathway**.

Hypothalamic output from mammillary bodies (MB) via the **mammillothalamic tract** goes to the **anterior thalamic nuclei** (**ATN**) that are connected to the **cingulate cortex** (**CC**). The cingulate cortex projects to the hippocampus, so closing a loop (MB–ATN–CC–hippocampus–hypothalamus) termed the **Papez circuit**, originally proposed for conscious awareness of emotions and cognitive effects on emotion. The mammillary bodies also project via the **mammillotegmental tract** to the midbrain. In addition the hypothalamus projects to the prefrontal cortex.

The **medial forebrain bundle** passes through the lateral hypothalamic zone. It consists mainly of monoaminergic axons ascending from brainstem nuclei. Many noradrenergic and serotonergic axons synapse with hypothalamic neurons. Dopaminergic axons from the substantia nigra and the ventral tegmentum however simply traverse the hypothalamus without establishing connections.

The paraventricular hypothalamus and the lateral hypothalamic area receive visceral sensory input from the nucleus of the solitary tract (NST), which is important for hypothalamic control of the ANS.

Pituitary gland

The **pituitary gland** (**hypophysis**) is divided into the **neurohypophysis** and the **adenohypophysis**. The neurohypophysis, which is a direct outgrowth of the hypothalamus, consists of the **posterior lobe**, the infundibular stalk and the median eminence. The adenohypophysis consists of the **anterior lobe**, an intermediate lobe which is poorly developed in humans, and the **pars tuberalis**, an extension surrounding the infundibular stalk. The pars tuberalis and the infundibular stalk are together termed the **pituitary stalk** (**infundibulum**) (*Fig. 3*).

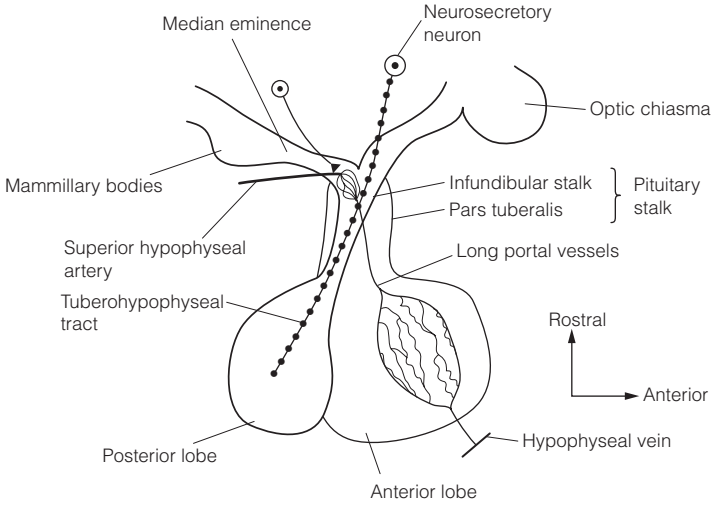


Fig. 3. The pituitary showing its neural and vascular connections with the hypothalamus. In humans the tuberohypophyseal tract contains about 100 000 axons.

There are two routes by which the hypothalamus controls hormone output from the pituitary. The functional connection between the hypothalamus and the posterior lobe is *neural*. Large **neurosecretory neurons** with their cell bodies in the hypothalamus send their axons through the median eminence down the infundibular stalk into the posterior lobe as the **tuberohypophyseal tract**. Hormones released from the posterior lobe are actually made in the cell bodies of the neurosecretory cells and secreted from their axons. By contrast, the functional connection between the hypothalamus and the anterior lobe is *vascular*. The superior hypophyseal artery forms a plexus of capillaries in the medial eminence which drains into **long portal vessels** that descend to the anterior lobe. The portal vessels open into venous sinusoids which supply the cells of the anterior pituitary with blood. The function of this **hypothalamic–pituitary portal circulation** is to deliver hormones, secreted by small hypothalamic neurons into the median eminence, to the anterior lobe.

L2 Posterior pituitary function

Key Notes

Posterior lobe hormones

Neurons in the supraoptic (SON) and paraventricular (PVN) nuclei send their axons into the posterior lobe and release two peptide hormones, arginine vasopressin (AVP) and oxytocin.

Arginine vasopressin (AVP)

AVP is secreted from the posterior pituitary in response to a rise in the osmolality of the extracellular fluid or a fall in blood volume. It acts to restore these by increasing the reabsorption of water by the nephrons of the kidney. Changes in osmolality are detected by neurons in a circumventricular organ which synapses on the PVN and SON. Alterations in blood volume are detected in two ways. Firstly, as changes in mean arterial blood pressure that are signaled by baroreceptors, afferents which go to the nucleus of the solitary tract (NST). The NST relays the baroreceptor signals to the PVN and SON. Reduced blood pressure causes a rise in AVP secretion. Secondly, a fall in blood volume is detected by the nephron, which responds by secreting renin. This enzyme triggers a cascade that generates angiotensin II (AII). AII stimulates both AVP secretion and drinking.

Oxytocin

Oxytocin, released from cells in the PVN and SON different from those that secrete AVP, stimulates contraction of smooth muscle. Suckling stimulates reflex milk ejection by release of oxytocin, and uterine contractions during childbirth occur in response to oxytocin release triggered reflexly by the pressure of the fetus on the neck of the uterus.

Oxytocin and behavior

Oxytocin neurons in the hypothalamus project to limbic structures. These are implicated in reproductive behavior. Pair bonding is viewed as addiction to a specific individual organized by oxytocin release, initially during sexual intercourse, subsequently conditioned to the presence of the individual. Maternal behavior is triggered by oxytocin release during suckling acting on oxytocin receptors in limbic structures. The oxytocin receptors are up-regulated by the higher maternal concentration of estradiol, relative to progesterone, after birth.

Related topics

Blood–brain barrier (A8)
Autonomic nervous system function (L5)
Emotion (M1)

Posterior lobe hormones

Magnocellular (large) neurosecretory neurons in the **supraoptic** (**SON**) and **paraventricular** (**PVN**) nuclei send their axons into the posterior lobe. These neurons secrete the nonapeptides **arginine vasopressin** (**AVP**), also known as **antidiuretic hormone** (**ADH**), and **oxytocin**. Both are synthesized as pro-

hormones in the cell bodies of the neurons and packed into large (120 nm) neurosecretory vesicles which are delivered to the terminals by axoplasmic transport. The prohormones are cleaved in the vesicles to give both the hormone and a cleavage product, a **neurophysin**.

Arginine vasopressin (AVP)

AVP is secreted from the posterior lobe into the systemic circulation in response to an increase in extracellular fluid osmolality or reduced blood volume. It increases the water permeability of the collecting ducts of the nephron, thereby promoting water reabsorption. This has the effect of reducing extracellular fluid osmolality and urine output (an antidiuretic effect) and restoring blood volume. Thus, AVP acts as a negative feedback regulator, defending set points in body fluid osmolality and blood volume.

The stores of AVP in the posterior lobe are large, sufficient to maintain maximum antidiuresis during several days of dehydration. The osmoreceptors which respond to changes in osmolality are in the **vascular organ of the lamina terminalis** (**OVLT**). The OVLT is one of the **circumventricular organs** of the brain which lie on the blood side of the blood–brain barrier situated at the anterior end of the hypothalamus. Osmolality-sensitive neurons in the OVLT synapse with the PVN and SON cells (*Fig. 1*), increasing their tonic firing rate as osmolality rises. There is a linear relationship between plasma osmolality and AVP secretion.

A reduction in blood volume greater than about 10% stimulates AVP secretion. This is seen in dehydration from any cause (e.g. water deprivation, vomiting or diarrhea) or with hemorrhage. Two mechanisms operate to trigger AVP release in **hypovolemia** (low blood volume).

1. Hypovolemia lowers mean arterial blood pressure. This is detected by stretch receptors (**baroreceptors**) in the walls of the carotid sinus and aorta. The afferents of these pressure sensors run in the glossopharyngeal (IX) and vagus (X) cranial nerves to the **nucleus of the solitary tract** (**NST**) in the medulla. The NST activates noradrenergic neurons in the ventrolateral medulla which project to the PVN and SON to bring about AVP release. A reduced blood pressure causes *decreased* firing of the baroreceptor afferents and hence disinhibition of the circuitry triggering AVP secretion. This is shown in *Fig. 2*.

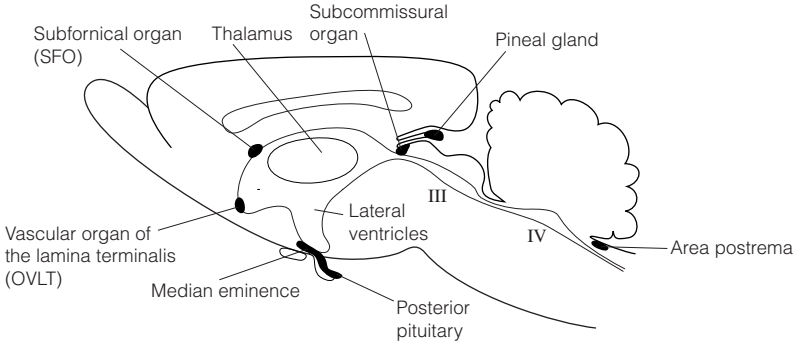


Fig. 1. Location of the circumventricular organs (CVOs, shaded) in the rat brain (midsaggital section); the ventricles are stippled.

2. Activation of the renin–angiotensin cascade (*Fig. 3*). Renin is secreted by the **juxtaglomerular apparatus** (**JGA**) of the kidney in response to several factors contingent on a fall in blood volume.

Renin is a proteolytic enzyme which cleaves a plasma protein, **angiotensinogen**, to yield a decapeptide, **angiotensin I**. This is further cleaved by angiotensin-converting enzyme (ACE), expressed on pulmonary endothelial cells, to the octapeptide, **angiotensin II** (**AII**). Angiotensin II stimulates the **subfornical organ** (a circumventricular organ), neurons of which stimulate AVP

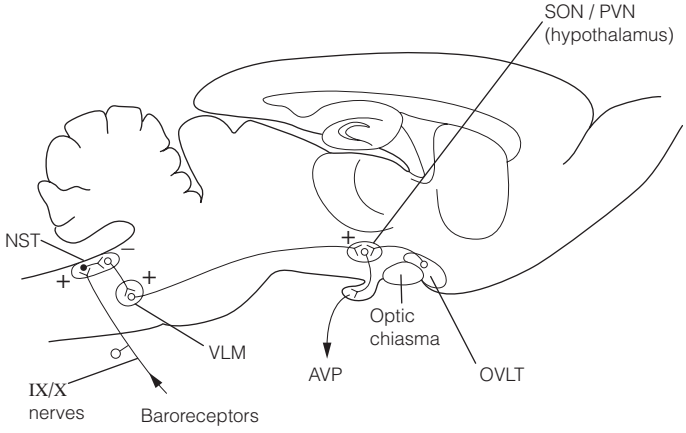


Fig. 2. A model for neural control of arginine vasopressin (AVP) secretion. Increased osmolality detected by the vascular organ of the lamina terminalis (OVLT) stimulates supraoptic and paraventricular nuclei (SON and PVN) cells to secrete AVP. Reduced arterial blood pressure is signaled via the nucleus of the solitary tract (NST) and the ventrolateral medulla (VLM) to the SON/PVN.

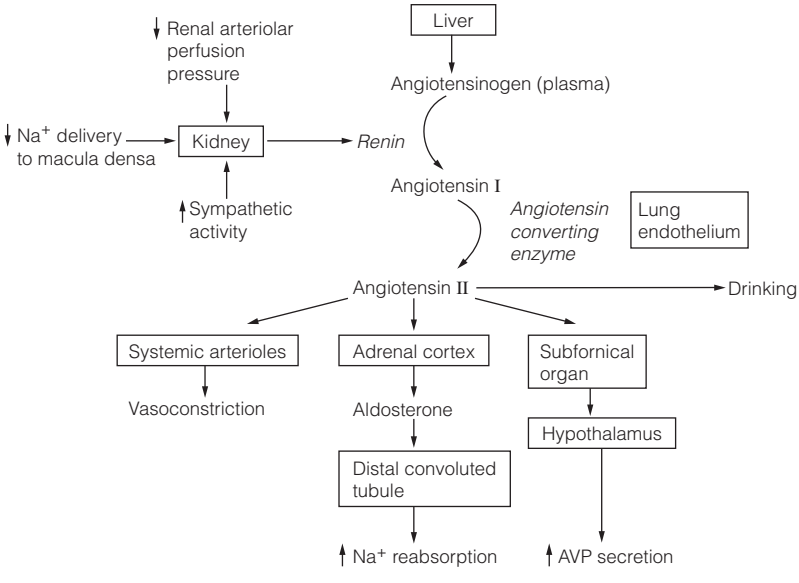


Fig. 3. The renin–angiotensin cascade helps to maintain body fluid osmolality and blood volume.

secretion. In addition, AII stimulates drinking, vasoconstriction, and the secretion of the aldosterone; all actions which help restore blood volume and pressure.

In dehydration due to water deprivation, about 70% of the fluid restoration is brought about by osmoreceptor-driven AVP secretion and the remainder is due to responses to reduced blood volume.

Failure or impairment of AVP secretion or inability to respond to the hormone causes **diabetes insipidus** (**DI**), characterized by very high urine output (10–20 liters per day), and excessive drinking. DI is most commonly caused by destruction of the magnocellular cells of the PVN and SON by tumors or autonomic disease. Long half-life analogs of AVP can be used to treat DI.

Oxytocin

Oxytocin is implicated in several aspects of reproductive function. Its stimulation of smooth muscle contraction underlies the **milk ejection reflex** in lactating females, and maintenance of uterine contractions during **parturition** (birth).

Suckling is the most potent stimulus for milk ejection. Primary afferents from the areolar and nipple skin relay with spinothalamic tract neurons in the dorsal horn of the spinal cord. Spinothalamic input causes oxytocin secretion via an undefined neural pathway from midbrain to the PVN and SON. Neurons that secrete oxytocin are distinct from those that secrete AVP.

Oxytocin is not the trigger for parturition. However, at term, a rise in maternal estradiol/progesterone ratio up-regulates oxytocin receptors in uterine smooth muscle which consequently becomes very sensitive to oxytocin. Once parturition is established, pressure of the fetal head on the cervix evokes the secretion of oxytocin from the posterior pituitary via a reflex pathway similar to that for milk ejection. The oxytocin stimulates contractions of the uterine smooth muscle, further increasing the pressure of the fetus on the cervix. This positive feedback mechanism is the **Ferguson reflex**. However, since parturition can be normal in spinally transected women, or those with diabetes insipidus who lack oxytocin, additional mechanisms must also be important.

Oxytocin and behavior

Oxytocin is implicated in reproductive behaviors such as pair bonding and parenting. These are mediated by oxytocin neurons in the hypothalamus that project to limbic structures involved in emotion and its expression. **Pair bonding**, the tendency for two individuals in a sexual relationship to stay together, is facilitated by oxytocin in monogamous voles and this may also happen in humans. Oxytocin neurons fire during sexual intercourse in women and men, and oxytocin release produces feelings of pleasure and increased libido. It is postulated that oxytocin release becomes conditioned by repeatedly having intercourse with the same partner, so eventually their presence is enough to trigger it. Brain-imaging studies show that brain regions (e.g. the anterior cingulate cortex, part of the limbic system) that are active when a woman is shown photographs of her lover (but not platonic friends) are those that respond to oxytocin. It is currently argued that sexual and romantic love can be thought of as addictive behavior towards a specific individual organized by oxytocin.

In rats estradiol triggers **maternal behavior** (pup retrieval, licking), probably by up-regulating oxytocin receptors in the limbic system. Several studies show that increasing brain oxytocin enhances maternal behavior, while oxytocin antagonists and lesions of the paraventricular nucleus reduce it. Under normal circumstances the rise in maternal blood estradiol/progesterone ratio at birth

triggers the increase in oxytocin receptor number and episodic bursts of oxytocin release occur at suckling. Over time classical conditioning takes place, in which the smell of the pups becomes the conditioned stimulus for oxytocin release. Imaging studies reveal that when women are shown pictures of their children (but not those of friends) they engage brain areas that overlap extensively with those implicated in romantic love and oxytocin neurotransmission.

L3 Neuroendocrine control of metabolism and growth

Key Notes

Hypothalamic–anterior pituitary axes

The hypothalamus and anterior pituitary, acting in concert, control five endocrine axes that regulate aspects of metabolism, reproduction, development and growth. Hypothalamic neurons secrete hormones that either stimulate or inhibit the anterior pituitary secretion of trophic hormones. Trophic hormones released into the circulation in turn stimulate target tissues (e.g. adrenals, thyroid and gonads) to secrete their hormones. The secretion of hypothalamic hormones and hence of the trophic hormones is pulsatile. The size and period of the pulses varies cyclically over 24 hours, and sometimes over longer times also. Secretion from the endocrine axes is under negative feedback regulation which maintains set point concentrations of hormones. By varying the set point the endocrine axes can alter their hormone output.

Hypothalamic–pituitary–adrenal (HPA) axis

The HPA axis controls the secretion of glucocorticoids by the adrenal cortex. Paraventricular neurons secrete corticotrophin-releasing hormone (CRH) which causes a population of anterior pituitary cells to release adrenocorticotrophic hormone (ACTH) into the circulation. ACTH is the trophic hormone that stimulates the adrenals to release glucocorticoids. Glucocorticoids act at two types of receptor that belong to the superfamily of intracellular steroid receptors. On ligand binding they translocate from cytoplasm to nucleus where they alter gene transcription.

Stress

Stress can be defined as a state in which there is a protracted elevation in concentrations of ACTH and glucocorticoids. Glucocorticoids are adaptive in stress by promoting the synthesis of glucose and glycogen from non-carbohydrate precursors. The HPA is activated in stress by catecholaminergic neurons concerned with arousal, or hunger and thirst sensations, by brainstem cholinergic neurons conveying sensory input associated with the stressor, and from other hypothalamic nuclei relaying limbic system information.

Hypothalamic–pituitary–thyroid (HPT) axis

Paraventricular neurons release thyrotrophin-releasing hormone (TRH), which causes anterior pituitary cells to secrete thyroid stimulating hormone (TSH), a trophic hormone which stimulates growth of the thyroid gland and release of the thyroid hormones (T3 and T4). Thyroid hormone receptors are members of the intracellular steroid receptor superfamily, but unlike steroid receptors, they are bound to nuclear DNA in the absence of hormone. On binding T3 the receptor activates gene transcription. The output of thyroid hormone is regulated by negative feedback modification of both TRH and TSH secretion. Cold exposure excites neurons in the preoptic hypothalamus that activates the HPT axis. The increased secretion of thyroid hormones raises the metabolic rate

helping to maintain core temperature. Thyroid hormones are required for fetal brain development, and maternal thyroid hormone deficiency, due to lack of dietary iodide, can cause neurological cretinism in infants.

Growth hormone (GH)

GH released from the anterior pituitary stimulates cell division and growth of many tissues, and mobilizes fatty acids as energy substrates. It is secreted in increased amounts during exercise, stress and fasting. GH secretion is stimulated by growth-hormone-releasing hormone (GHRH), and inhibited by somatostatin. At both hypothalamus and pituitary, GH stimulates the production of insulin-like growth factor, either in the brain or by peripheral tissues. Insulin-like growth factor exerts negative feedback control on the release of growth hormone. Most GH release occurs at night. Several neurotransmitter systems and hormones modify GH output; sex steroids particularly stimulate the high GH output responsible for the growth spurt of puberty.

Related topics

Anatomy and connections of the hypothalamus (L1)
Neuroendocrine control of reproduction (L4)
Emotion (M1)
Brain biological clocks (M4)
Cell physiology of learning (O3)

Hypothalamic–anterior pituitary axes

Acting through the anterior lobe of the pituitary gland the hypothalamus controls five endocrine **axes**. Between them these neuroendocrine axes regulate key aspects of metabolism, reproduction, development and growth. The five axes share many common features. Neurons, located in several hypothalamic nuclei, send their axons to the external zone of the median eminence and the **tuberoinfundibular tract**. These axons secrete **hypophysiotropic hormones** into the hypothalamic–pituitary portal circulation which carries them into the anterior lobe. Each hypophysiotropic hormone acts on a particular population of cells in the anterior lobe, either exciting or inhibiting their secretion of a specific **stimulating** (**trophic**) hormone. Hypophysiotropic hormones that excite secretion are termed **releasing** hormones, those that inhibit are called **release-inhibiting** hormones. Trophic hormones of the anterior pituitary are secreted into the systemic circulation and have endocrine effects on target tissues, particularly endocrine glands (*Table 1*).

Secretion of hypothalamic hormones is pulsatile with a period of 60–180 minutes. This drives pulsatile release of anterior pituitary hormones. The amplitude and period of the pulses is varied on a circadian basis and in some cases on longer timescales. Secretion from the neuroendocrine axes is modulated by feedback acting at several levels which tends to defend a set point in the concentration of the end product (*Fig. 1*).

Negative feedback is an extremely common homeostatic principle in biology. It acts to hold some variable at a constant level, the **set point**. In *Fig. 1*, if the concentration of the end product hormone exceeds the set point, more receptors are activated in the hypothalamus and anterior pituitary which consequently *reduce* their output of hormones. The effect is that, after some delay, the concentration of end product hormone falls. If it falls below the set point the hypothalamus and pituitary secrete more hormones, provoking an increase in the synthesis of the end product. **Autofeedback inhibition** is a special case of

Table 1. Five hypothalamic–anterior pituitary neuroendocrine axes

Hypophysiotropic hormone		Anterior pituitary stimulating (trophic) hormone [anterior pituitary cell type]	Target tissue for trophic hormone	Secreted hormone
Releasing hormone	Release-inhibiting hormone			
Corticotrophin releasing hormone	–	Adrenocorticotrophic hormone [corticotroph]	Adrenal cortex	Glucocorticoids
Thyrotrophin releasing hormone	–	Thyroid stimulating hormone [thyrotroph]	Thyroid	Triiodothyronine (T3) and thyroxine (T4)
Gonadotrophin releasing hormone	–	Follicle stimulating hormone Luteinizing hormone [gonadotroph]	Gonads	Sex steroids: estrogens, progestogens and androgens
Growth hormone releasing hormone	Somatostatin	Growth hormone (GH, somatotrophin) [somatotroph]	Liver, fibroblasts, myoblasts, chondrocytes, osteoblasts and others	Somatomedins (insulin-like growth factors)
Prolactin releasing factor[a]	Dopamine (acting at D2 receptors)	Prolactin [lactotroph]	Mammary glands	

[a] The molecule responsible for stimulating prolactin release has not been unambiguously identified.

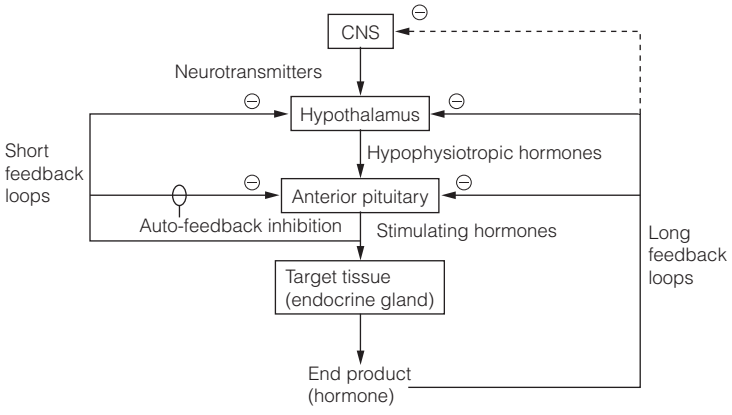


Fig. 1. Negative feedback loops that control neuroendocrine secretion.

negative feedback in which a substance directly inhibits its own synthesis. Several mechanisms exist to alter the set points of physiological systems so that hormone concentrations can be varied as circumstances change. For example, in many endocrine systems hormone concentrations fluctuate in a cyclical manner during the day because set points are adjusted by biological clocks in the brain.

Hypothalamic–pituitary–adrenal (HPA) axis

The HPA axis regulates the synthesis and secretion of **glucocorticoids**, a group of steroid hormones which help control the metabolism of energy substrates. The most important glucocorticoid in humans is **cortisol.** Cells in the paraventricular nucleus of the hypothalamus secrete **corticotrophin-releasing hormone** (**CRH**), a 41 amino acid residue peptide which acts synergistically with arginine vasopressin to stimulate corticotrophs to release **adrenocorticotrophic hormone** (**ACTH**). This is cleaved from a large precursor, **pro-opiomelanocortin**. In response to ACTH, cells in the adrenal cortex synthesize and secrete glucocorticoids. Negative feedback by glucocorticoids at the hippocampus, hypothalamus and pituitary regulates secretion of the steroids (*Fig. 2*).

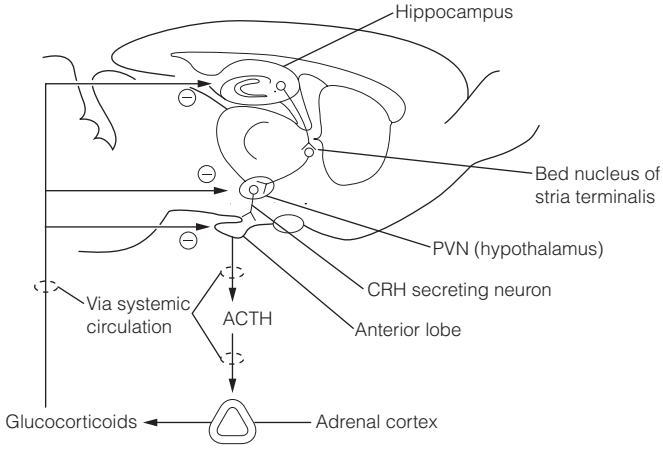


Fig. 2. Feedback in the hypothalamic–pituitary–adrenal axis. Stippled regions harbor glucocorticoid receptors. ACTH, adrenocorticotrophic hormone; CRH, corticotrophin releasing hormone; PVN, paraventricular nucleus.

A circadian rhythm in glucocorticoid output is driven by the suprachiasmatic nucleus acting on CRH-secreting cells. In humans ACTH pulses are greatest early in the morning and decline through the day to reach a low point around midnight. Glucocorticoid secretion follows a similar pattern with a delay of about 30 minutes. This daily rhythm is influenced by light and dark, sleep and meals.

The effects of glucocorticoids are mediated via two distinct receptors, coded for by separate genes: the high-affinity **mineralocorticoid receptor** (**MR**) and the low-affinity **glucocorticoid receptor** (**GR**). Mineralocorticoid receptors are in greatest numbers in limbic structures. Glucocorticoid receptors are more widespread, and expressed in glia as well as neurons. Both receptors colocalize in the same cells in the hippocampus. At basal levels of cortisol secretion MRs are largely occupied, while occupation of significant numbers of GRs occurs only when the cortisol concentration is high such as during the early morning circadian peak.

MRs and GRs are members of a nuclear receptor superfamily that includes receptors for other steroids (estrogen receptors, progesterone receptors, androgen receptors), thyroid hormone receptors, and receptors for vitamin D3 and retinoic acid. Glucocorticoids are lipophilic and so diffuse readily across plasma membranes of cells. MRs and GRs are present in the cytoplasm, complexed with **heat shock proteins** (**HSPs**) that act as molecular chaperones, stabilizing the receptors into their functional configuration. Binding of the ligand to the receptor causes it to translocate into the nucleus where it binds to specific sequences of DNA, **hormone responsive elements** (**HREs**), which results in increased or decreased transcription of specific genes (*Fig. 3*).

The corticotrophs of the anterior pituitary, CRH-secreting neurons of the hypothalamus, and hippocampal neurons express GRs. When the concentration of glucocorticoids is high these GRs are activated and they inhibit transcription of the genes for CRH and arginine vasopressin. This is one mechanism for negative feedback control of glucocorticoid concentrations.

Stress

The hypothalamic–pituitary–adrenal axis is activated by **stress**. No *comprehensive* definition of stress exists. Physiological stressors such as hunger, thirst or trauma threaten homeostasis, and the physiological responses they evoke are adaptive because they tend to restore/maintain homeostasis. Psychological

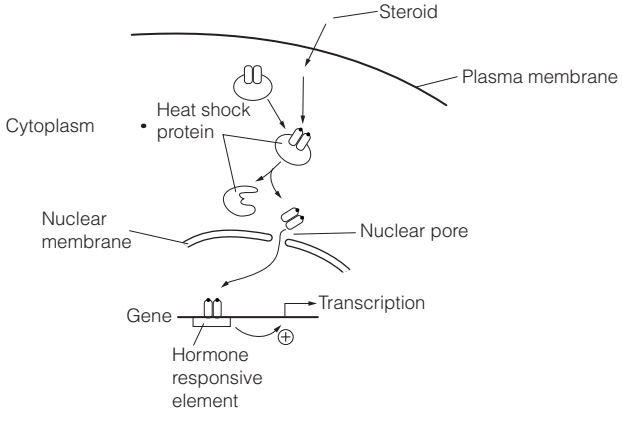


Fig. 3. Model of steroid hormone receptor action to modulate gene transcription.

stressors do not directly derange homeostasis, do not affect individuals equally, and stress responses to them are often learnt. Psychological stress often arises from the uncertainty inherent in social interactions, or in situations over which there is little control. Psychological stressors produce emotional (**affective**) states – anxiety, fear, anger, frustration, depression, etc. – the nature and intensity of which depends on an individual's evaluation of the situation in the light of prior experience. A frequently adopted *operational* definition of stress is any state in which there is a prolonged rise in ACTH and glucocorticoid concentrations.

The increased secretion of glucocorticoids is useful in stress because their overall effect is to harness long-term energy substrates and convert them to readily available substrates, glycogen and glucose. The early morning peak in glucocorticoids is timed to correspond with what is generally the longest interval each day without food. In addition, glucocorticoids potentiate the effects of catecholamines.

Activation of the HPA in stress occurs through inputs from a variety of sources converging on the CRH-secreting cells of the paraventricular nucleus (PVN).

- Stress-evoked arousal activates noradrenergic neurons in the locus ceruleus which project to the PVN.
- Visceral sensations associated with thirst and hunger are transmitted via the glossopharyngeal (IX) and vagus (X) nerves to the nucleus of the solitary tract and adjacent regions of the medulla. These structures project catecholaminergic axons to activate the PVN.
- Inputs from the circumventricular organs, which monitor osmolality or release arginine vasopressin, go to the PVN to activate the HPA during dehydration.
- Neurons in the midbrain and pons, many cholinergic, project to the PVN and are thought to transmit visual, auditory and somatosensory input associated with stressful situations.
- Most hypothalamic nuclei project to the PVN and these connections funnel information about stressful situations from prefrontal cortex and limbic structures such as the amygdala or hippocampus.

The high concentrations of glucocorticoids seen in stress produce substantial occupancy of the low-affinity glucocorticoid receptors and this terminates the stress response via negative feedback. The HPA is more sensitive to stress activation and negative feedback inhibition at times when blood glucocorticoid concentrations are at their lowest.

Protracted activation of the HPA by chronic stress has deleterious effects. High concentrations of corticosteroids acting via glucocorticoid receptors enhance excitatory amino acid transmission, increasing the calcium influx through voltage-dependent Ca^{2+} channels into hippocampal cells, which can lead to excitotoxic cell death. This may account for the finding that old rats have fewer hippocampal pyramidal cells. The cell loss reduces the number of glucocorticoid receptors, blunting the effectiveness of glucocorticoid negative feedback. In both old rats and humans corticosteroid concentrations take longer to return to basal levels after stress compared with younger individuals.

Hypothalamic–pituitary–thyroid (HPT) axis

Thyroid hormones, amongst other functions, regulate basal metabolic rate by increasing metabolic heat production. Thyroid hormone secretion is regulated by the hypothalamus and pituitary, and is influenced by several factors, such as ambient temperature.

Thyrotrophin-releasing hormone (**TRH**) is a tripeptide synthesized, as part of a precursor, in small neurons of the paraventricular nucleus. The axons of these cells run in the tuberoinfundibular tract to the median eminence. TRH secreted here reaches the anterior pituitary by way of the hypothalamic–pituitary portal system and stimulates thyrotrophs to secrete **thyroid-stimulating hormone** (**TSH**). TSH is a glycoprotein consisting of two chains, α and β. It is liberated into the systemic circulation and stimulates division and growth of cells in the thyroid gland, and the synthesis and secretion of the thyroid hormones. There are two thyroid hormones, **thyroxine** (**T4**) and **triidothyronine** (**T3**), named for the number of iodine atoms they contain.

Thyroid hormone receptors (**TRs**) are members of the steroid receptor superfamily. They form heterodimers together with retinoid X receptors (RXRs), and differ from glucocorticoid receptors in that the heterodimer binds to hormone-responsive elements in the DNA in the *absence* of ligand. These receptors have a higher affinity for T3 than T4. T4, which forms the bulk of the secreted hormones, is a prohormone which is converted to T3 by the neuronal cytosolic enzyme, 5′-deiodinase II. On binding of T3 the thyroid hormone receptor activates gene transcription.

Thyroid hormone output is controlled by negative feedback acting at several levels of the HPT axis. A drop in thyroid hormone concentration causes the increased secretion of TSH by thyrotrophs of the anterior pituitary. TRH secretion from the hypothalamus is also subject to feedback inhibition by both T4 and T3.

Pulses of TRH secretion drive pulsatile TSH output. The frequency and amplitude of the pulses is entrained into a circadian rhythm by the suprachiasmatic nucleus, rising throughout the night, falling during the morning and remaining low throughout the afternoon. This circadian rhythm is sensitive to light and dark, but unaffected by sleep patterns.

Thyroid hormone secretion is increased in cold exposure. Temperature-sensitive neurons in the preoptic hypothalamus which get input from skin thermoreceptors project to brainstem noradrenergic neurons. These, in turn, synapse with the TRH-secreting cells of the PVN. Cold exposure activates the noradrenergic neurons, provoking a rise in TRH secretion. The resulting increase in the concentrations of thyroid hormones enhances metabolic rate, helping to maintain core temperature. The effect is rapid, elevation of thyroid hormone output is seen within 30 minutes of cold exposure.

Thyroid hormones are crucially required for human brain development from very early pregnancy, long before the fetal thyroid begins to function at about 17 weeks gestation. Before 17 weeks, brain development is driven by maternal T4 which crosses the placenta and is converted to T3 in the fetal brain. After 17 weeks, T4 entering the placenta is deiodinated by a placental 5-deiodinase to L-3,3′,5′-triidothyronine (reverse T3, rT3) which is inactive. However, the liberated iodide is used by the *fetal* thyroid to synthesize its own hormones. Maternal **hypothyroxinemia**, usually caused by a lack of dietary iodide, can result in **neurological cretinism** in infants, in whom reduced thyroid hormone compromises synaptogenesis, myelination and axoplasmic transport, particularly in the cerebral and cerebellar cortices. About half a billion women live in iodine-deficient regions (e.g. parts of India). Iodized table salt is a cheap and effective remedy for hypothyroxinemia.

Growth hormone (GH)

Growth hormone (**somatotrophin**) stimulates cell division and growth of many tissues, particularly during the perinatal period and the adolescent growth spurt, enhancing protein synthesis by increasing transcription and translation.

GH mobilizes fatty acids as energy substrates. This is adaptive during exercise, stress and fasting, three major physiological variables which increase GH secretion.

GH secretion from somatotrophs of the anterior pituitary is regulated by two hormones, **growth-hormone-releasing hormone** (**GHRH**) and **somatostatin**. GHRH is a 44 amino acid residue peptide synthesized from a precursor by neurons in the **arcuate nucleus** of the hypothalamus, the axons of which terminate in the median eminence. Secreted GHRH is transported to the anterior pituitary via the hypothalamic pituitary portal system. Somatostatin is an important transmitter throughout the CNS, but the somatostatin-containing cells responsible for inhibiting GH secretion are restricted to the **periventricular nucleus**, and their axons project to the median eminence. Somatostatin is a 14 amino acid residue peptide synthesized from a precursor. GHRH and somatostatin exert their opposing effects on GH secretion via metabotropic receptors coupled to the cAMP second messenger system. GHRH receptors enhance cAMP concentrations via Gs proteins while somatostatin receptors reduce cAMP by coupling to Gi proteins. The cAMP second messenger system modulates GH secretion by altering Ca^{2+} influx in the somatotrophs.

The secretion of GH is circadian and pulsatile, driven mostly by pulses of GHRH from the hypothalamus. The pulses are much bigger at night and triggered by deep (stages 3 and 4) slow-wave sleep. This nocturnal GH secretion is greatest in children and declines with age. It is brought about by a serotonergic pathway from the brainstem to the hypothalamus. GHRH secretion is also stimulated by dopaminergic, noradrenergic and enkephalinergic pathways in the brain. Thyroid hormones are required for normal levels of GH synthesis and secretion, and basal concentrations of glucocorticoids enhance, while high concentrations inhibit, GH synthesis.

Negative feedback control of GH secretion occurs at the pituitary by suppression of the synthesis and secretion of GH, and at the hypothalamus by reduction of GHRH secretion. GH also stimulates the secretion of somatostatin. These negative feedback effects are exerted by **insulin-like growth factor-1** (**IGF-1**), one of a group of peptides called **somatomedins** which mediate the effects of GH. IGF-1 is produced either in the brain, or peripherally, in response to GH.

A rapid rise in the rate of growth due to high levels of GH secretion, the **growth spurt**, occurs during puberty. During this time gonadal secretion of androgens and estrogens rises, stimulating GH secretion.

L4 Neuroendocrine control of reproduction

Key Notes

The hypothalamic–pituitary–gonadal (HPG) axis

Gonadotrophin-releasing hormone is synthesized by neurons in several hypothalamic nuclei. It stimulates the anterior pituitary to release two gonadotrophins, follicle-stimulating hormone (FSH) and luteinizing hormone (LH). These trophic hormones stimulate the gonads to produce sex steroids. Gonadotrophin secretion is pulsatile. In males the period of the pulses is constant, but in females the period depends on the phase of her reproductive cycle.

Feedback in males

Gonadotrophins stimulate the testis to produce testosterone and inhibin, which both produce negative feedback suppression of the HPG axis. Testosterone acts at both the hypothalamus and the anterior pituitary but the effects of inhibin are confined to suppression of anterior lobe secretion of FSH.

Feedback in female reproductive cycles

In women, gonadotrophins stimulate the ovarian follicles to grow, producing estradiol and inhibin during the first half of the menstrual cycles (the follicular phase). After ovulation, the follicle becomes a corpus luteum which secretes progesterone through the second half of the cycles (the luteal phase) in response to FSH and LH. During most of the menstrual cycles sex steroids exert a negative feedback suppression of gonadotrophin output. In the follicular phase it is mediated by estradiol (and inhibin) whereas in the luteal phase it results from both estradiol and progesterone. However, just before ovulation the high concentrations of estradiol produced by the mature follicle cause the HPG to switch briefly into a positive feedback mode. Now, estradiol generates a midcycle surge of gonadotrophin which triggers ovulation.

How steroid feedback works

Negative feedback by estradiol is due to down-regulation of GnRH receptors so the anterior pituitary becomes less sensitive to GnRH. High estradiol concentration near mid-cycle causes GnRH release to occur in high-frequency, low-amplitude pulses, allowing up-regulation of GnRH receptors in the anterior pituitary so it becomes very sensitive to GnRH. This results in a rapid rise in LH and FSH secretion (positive feedback). The high estradiol concentration also induces the expression of progesterone receptors in the hypothalamus. Rising progesterone concentrations after ovulation act on these to switch the secretion of GnRH into a low-frequency, high-amplitude pattern of release which down-regulates GnRH receptors, re-establishing negative feedback. Studies of women with congenital deficiency in GnRH secretion show that normal menstrual cycles can be supported even with constant-size GnRH pulses.

Puberty and menopause

Puberty is due to the activation of the previously quiescent HPG axis. What initiates puberty is not known but it is thought to be a metabolic signal of growth or body mass. The menopause results from ovarian failure.

Prolactin (PRL)

Secreted by the anterior pituitary, prolactin stimulates breast tissue development during pregnancy and is responsible for reflex synthesis and secretion of milk by suckling in lactating women. PRL secretion, like that of GH, is subject to dual regulation by the hypothalamus. Dopamine acts on the anterior pituitary to inhibit PRL release. Several peptides can stimulate PRL release, but which one does so *in vivo* is uncertain. High levels of PRL secretion, arising as a result of lactation or pathology causes infertility by suppressing LH secretion.

Related topics

Dopamine (D4)
Anatomy and connections of the hypothalamus (L1)
Posterior pituitary function (L2)
Neuroendocrine control of metabolism and growth (L3)

The hypothalamic–pituitary–gonadal (HPG) axis

The hypothalamic–pituitary–gonadal axis controls reproduction. In primates, neurons scattered widely throughout the hypothalamus synthesize a decapeptide, **gonadotrophin-releasing hormone** (**GnRH**, also referred to as **luteinizing hormone-releasing hormone, LHRH**) from a large precursor. GnRH is secreted from axon terminals in the median eminence into the hypothalamic–pituitary portal circulation. GnRH stimulates gonadotrophs of the anterior pituitary to secrete two gonadotrophins, **follicle-stimulating hormone** (**FSH**) and **luteinizing hormone** (**LH**), into the systemic circulation which carries them to the gonads.

The gonadotrophins are large glycoproteins each consisting of two peptides, an α chain and a β chain. The α chains of FSH and LH are identical (and very similar to the α chain of TSH) but the β chains are distinct and confer the specificity of the hormones. Gonadotrophins stimulate the gonads to produce **sex steroids** and have effects on gamete development. Gonadotrophin secretion is cyclical in females but not in males. The secretion of gonadotrophins is pulsatile, as with other anterior lobe hormones, and is driven by bursts of GnRH from the hypothalamus. In males, the pulses are regular, spaced about 3 hours apart, but in females the period varies between 1 and 12 hours depending on the phase of her reproductive cycle. Experimentally replacing pulsatile with continuous GnRH delivery in female rhesus monkeys abolishes gonadotrophin secretion, showing that *pulsatile* GnRH output is essential for proper HPG axis function.

Feedback in males

At the testis, LH stimulates **Leydig cells** to synthesize and secrete androgens, principally **testosterone**. FSH, together with testosterone, acts on **Sertoli cells** to organize the development of spermatozoa and secrete a glycoprotein, **inhibin**. Gonadotrophin secretion in males is subject to negative feedback control by both of the secretions from the testis. Testosterone acts at the hypothalamus, decreasing the frequency of the episodic GnRH bursts, and the anterior pituitary, making it less responsive to GnRH. Inhibin specifically suppresses only FSH secretion and inhibin must act at the anterior lobe, since the hypothalamus has no inhibin receptors. Humans have a circadian rhythm in testosterone secretion but show no longer-term cycles in testosterone output.

Feedback in female reproductive cycles

In females, the situation is more complicated since here the role of the HPG axis is firstly to stimulate the growth of a group of ovarian follicles (one of which develops to maturity), secondly to produce cyclical changes in sex steroid output, which prepares the reproductive tract for fertilization and implantation, and thirdly to trigger ovulation at the appropriate time, which in humans occurs on day 14 of the 28-day cycle.

The first half of the cycle (days 1–14) is called the **follicular phase**, since it is dominated by the growth of the ovarian follicle which secretes **estradiol** and **inhibin**. The second half of the cycle is **the luteal phase** (days 15–28), since after ovulation the follicle becomes a **corpus luteum** which secretes **progesterone**.

The extent to which feedback by steroids acts at the anterior pituitary or hypothalamus depends on species. In primates, the anterior lobe is the predominant mediator. What follows relates to humans. Feedback in females depends on the phase of the menstrual cycle. For most of the follicular phase low or moderate levels of estradiol and inhibin exert negative feedback effect on gonadotrophin secretion (*Figs 1 and 2*).

However, by about day 14, levels of estradiol become high enough to flip the HPG axis into a positive feedback mode. Now the estrogen stimulates a *rise* in LH and FSH secretion, which triggers ovulation, and switches the steroid metabolism of the post-ovulatory follicle so that it synthesizes and secretes

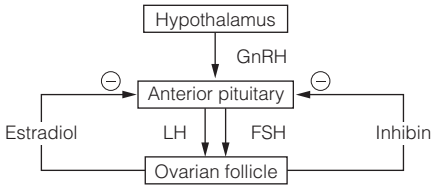


Fig. 1. Negative feedback inhibition during the follicular phase of the menstrual cycle. Negative feedback in males is very similar except that luteinizing hormone (LH) is regulated by testosterone from the testis. FSH, follicle stimulating hormone; GnRH, gonadotrophin releasing hormone.

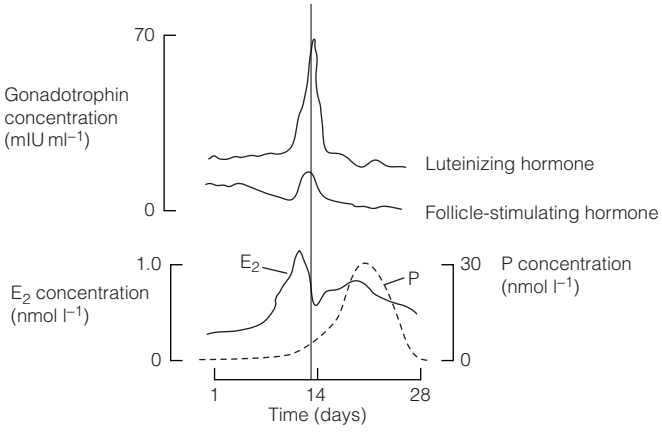


Fig. 2. Pattern of hormone secretion during the human menstrual cycle. Ovulation, triggered by the estradiol (E_2) evoked surge in luteinizing hormone occurs around day 14. P, progesterone; mIU, milli-international units.

progesterone (*Fig. 2*). The rise in progesterone secretion at the start of the luteal phase terminates the positive feedback LH surge and the system reverts to negative feedback mode.

How steroid feedback works

Early in the menstrual cycle estradiol causes the anterior pituitary to become less sensitive to GnRH so that the amplitude of the GnRH pulses is reduced. The reduced GnRH sensitivity may be due to down-regulation of GnRH receptors on the gonadotrophs. Late in the follicular phase the high estradiol concentration acts in the arcuate nucleus to cause GnRH release to occur in high-frequency, low-amplitude, pulses (*Fig. 3a, left*). This pattern allows up-regulation of GnRH receptors on the anterior pituitary gonadotrophs which become exquisitely sensitive to the releasing hormone. Consequently there is a rapid rise in LH and FSH secretion (*Fig. 3b*). The estradiol induces the expression of

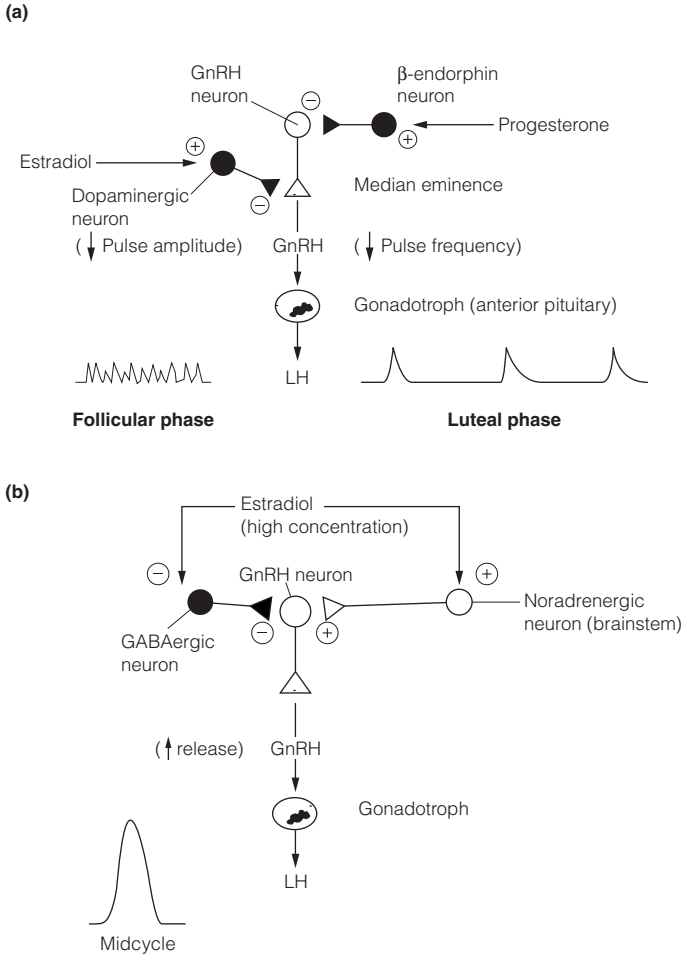


Fig. 3. Model for steroid feedback regulation of gonadotrophin releasing hormone (GnRH) producing neurons: (a) negative feedback by estradiol (left) and progesterone (right); (b) positive feedback by estradiol works by inhibiting γ-aminobutyrate (GABA) inhibition and stimulating noradrenergic excitation of GnRH release. Excitatory neurons (○) inhibitory neurons (●). LH, luteinizing hormone.

progesterone receptors in the arcuate nucleus. Rising progesterone concentrations after ovulation act on these to switch the secretion of GnRH into a low-frequency, high-amplitude, pattern of release (*Fig. 3a, right*). This down-regulates GnRH receptors in the anterior pituitary so that gonadotrophin secretion plummets in the luteal phase.

However, in humans the anterior pituitary is able to generate a normal pattern of gonadotrophin output even if exposed to pulses of GnRH of *constant* frequency and amplitude. Evidence for this comes from women with **Kallmann's syndrome** who have a deficiency in GnRH secretion caused by the failure of GnRH neurons to migrate properly during development. These women do not release gonadotrophins, do not experience puberty and are infertile. However, they can have normal menstrual cycles induced, exhibiting both negative and positive feedback, if given constant-size GnRH pulses. The implication is that the number of GnRH receptors on anterior pituitary cells (and hence their sensitivity to GnRH) can be regulated directly by steroids.

Puberty and menopause

Except for a brief postnatal period in primates, the HPG axis is quiescent until puberty so circulating levels of gonadal steroids are low. Inactivity of the HPG axis results from a tonic GABAergic inhibition on the GnRH neurons, and the onset of puberty is accompanied by decreased GABA inhibition. During puberty there is a dramatic circadian rhythmicity in gonadotrophin output; LH pulses are much larger during sleep. The precise trigger for puberty is not known, but a metabolic signal reflecting growth or body mass is probably involved. In girls, a critical mass of 30 kg appears necessary for puberty to commence, and needs to reach about 47 kg before menstrual cycles begin. Female dancers, athletes and anorexics fail to menstruate if their body mass is too low.

The end of reproductive life, the menopause, is characterized by ovarian failure; the hypothalamus and pituitary continue to function.

Prolactin

Prolactin (**PRL**) is secreted by lactotrophs of the anterior lobe. PRL is a glycoprotein with a similar amino acid sequence to growth hormone. It is one of several hormones that stimulates the growth of the mammary glands during pregnancy. PRL concentrations are highest during pregnancy and lactation. Suckling produces a reflex secretion of prolactin which stimulates the synthesis and secretion of milk. Like growth hormone, exercise and stress stimulate prolactin release.

PRL secretion is regulated by the hypothalamus, by dopaminergic neurons which are inhibitory, and by a number of peptides which activate secretion, a pattern similar to control of GH secretion. The dopaminergic neurons (A12 group) have their cell bodies in the arcuate nucleus of the hypothalamus and their axons run in the tuberoinfundibular pathway to terminate in the median eminence. Here the dopamine is released directly into the hypothalamic–pituitary portal circulation, carrying it to the anterior lobe. Lactotrophs express D2 dopamine receptors that are negatively coupled to adenylyl cyclase. The fall in cAMP concentration reduces transcription of the prolactin gene.

PRL release is stimulated by:

- thyrotrophin-releasing hormone (TRH);
- vasoactive intestinal peptide (VIP) secreted into the hypothalamic–pituitary portal circulation;

- oxytocin from the posterior pituitary, which gains access to the anterior lobe via tiny blood vessels called the **short portal vessels.**

However, there must be other prolactin-releasing factors to account for secretion under all physiological conditions. With frequent suckling (every 2–3 hours) the amount of prolactin released is sufficient to block ovulation by suppressing LH secretion, so lactating women are relatively infertile.

Excessive secretion of prolactin (**hyperprolactinemia**) may occur with some pituitary tumors or as a result of drugs which interfere with dopamine inhibition of prolactin secretion, e.g. dopamine receptor antagonists. Hyperprolactinemia causes infertility in males and females because of its suppression of LH secretion, and inappropriate secretion of milk in both sexes.

L5 Autonomic nervous system function

Key Notes

Overview of the autonomic nervous system (ANS)

The autonomic nervous system acts on smooth muscle, cardiac muscle and glands to keep key physiological variables at a level appropriate to an animals' activity and its environment. Much ANS regulation is by negative feedback in that it operates to defend a set point in some variable (e.g. mean arterial blood pressure). In other circumstances, autonomic mechanisms are homeostatic in the sense that variables change to meet altered demands. The sympathetic division is activated in stress to produce 'fright, fight or flight' responses, whereas the parasympathetic division is more active in 'rest and digest' situations. The two divisions often exert opposite effects on an organ (e.g. the heart). In a few instances ANS mechanisms are positive feedback in that they drive a physiological system away from its normal stable state (e.g. sexual responses).

ANS physiology

Acetylcholine (ACh) is the transmitter of autonomic ganglia and acts on both nicotinic (nAChR) and muscarinic (mAChR) receptors. The nAChR mediate fast ACh transmission. Activation of the mAChR greatly prolongs the time for which the postganglionic cells fire in response to the nicotinic receptor stimulation. Postganglionic sympathetic axons are usually noradrenergic and postganglionic parasympathetic terminals are invariably cholinergic. The adrenal medulla chromaffin cells are effectively postganglionic sympathetic cells and secrete mostly adrenaline (epinephrine). Postganglionic terminals co-release a variety of peptides transmitters which modulate, or have additional actions to, the primary transmitters.

An example of coordinated autonomic control: the urinary bladder

The urinary bladder has a dual autonomic innervation. Tonic sympathetic activity permits the bladder to fill. Urination is achieved by a spinal reflex in which bladder wall stretch receptors activate parasympathetic relaxation of the internal sphincter of the bladder. This reflex is normally under conscious control exerted via neurons in the pons. After severing the spinal cord, urination becomes entirely reflexive.

Enteric nervous system (ENS)

The myenteric and submucosal plexus coordinate gut motility and secretion respectively. Myenteric motor neurons exert tonic excitatory or inhibitory actions on gut smooth muscle. They are driven, via cholinergic interneurons, by sensory neurons that respond to stretch or chemical signals in the gut lumen. Usually sensory neuron firing stimulates motility in the oral direction, but inhibits it in the anal direction. Submucosal reflexes increase glandular secretion both directly and by increasing local blood flow.

Related topics

Organization of the peripheral nervous system (A4)
Control of the autonomic function (L6)

Overview of the autonomic nervous system (ANS)

The autonomic nervous system adjusts the contraction of smooth muscle and heart muscle and controls glandular secretion so that key physiological variables (e.g. core temperature, cardiac output, blood pressure, blood glucose) are maintained at levels appropriate to an animal's activity or the environment in which it finds itself. The term autonomic (self-governing) is apt since the ANS usually operates without conscious awareness and has no cognitive component.

To a first approximation much of the activity of the ANS is concerned with homeostatic regulation of physiological variables. For example, mean arterial blood pressure is kept fairly constant despite changes in posture which cause large swings in the hydrostatic pressure of blood. When suddenly standing from a lying position the tendency for blood to pool in the legs due to the force of gravity is offset by autonomic reflexes which monitor the drop in pressure and elicit constriction of arterioles and venules in the legs. This is a classic negative feedback mechanism.

Many autonomic adjustments are not negative feedback, since they do not defend a set point in some variable, but are homeostatic in that they change physiological variables so as to cope with altered demands. In response to a wide variety of stressors, activation of the sympathetic nervous system (SNS) to targets such as the heart, blood vessels, airways and liver results in increased cardiac output, regional alterations in blood flow, raised airflow through the lungs and elevations in blood glucose concentrations, all adaptations which improve the chances of surviving the stress unscathed. In general the SNS mediates the response of 'fright, fight and flight'. In contrast, parasympathetic nervous system (PNS) activation is seen when the body is in 'rest and digest' mode; the PNS generally stimulates exocrine gland secretion and promotes anabolic processes.

The sympathetic and parasympathetic divisions of the ANS often have opposing effects on a system, e.g. on pupil diameter or heart rate, and it is the balance of activities in the two divisions that achieves the appropriate outcome; that is, the SNS and PNS work in concert.

In a few situations the ANS works by positive feedback. Sexual responses in humans require autonomic reflexes (both sympathetic and parasympathetic) in which the motor response (enlargement of the penis or clitoris by vasocongestion) increases the firing of the same visceral afferents which drive the reflex response. This is positive feedback because it carries the system away from its usual stable state.

The classical view of the ANS is that it is exclusively a visceral motor system, the activities of which can be altered by sensory input. An alternative view is that the ANS should include visceral afferents since these run in the same nerve trunks as visceral efferents, they can be distinguished by their neuropeptide transmitters from somatic afferents and it brings both sensory and motor components into the same part of the nervous system.

ANS physiology

The distribution of the ANS to target organs is shown in *Fig. 1*. Acetylcholine (ACh) is the major neurotransmitter at all autonomic ganglia. ACh released from preganglionic neurons acts on nicotinic cholinergic receptors to produce a fast excitatory postsynaptic potential (epsp) which, if sufficiently large, makes the postganglionic cell fire. In addition ACh acts on M1 muscarinic receptors producing a slow epsp – prolonging the firing of the postganglionic cell by many seconds – by closing a population of potassium (K_M) channels. In this way long-lasting autonomic responses are generated by brief stimuli.

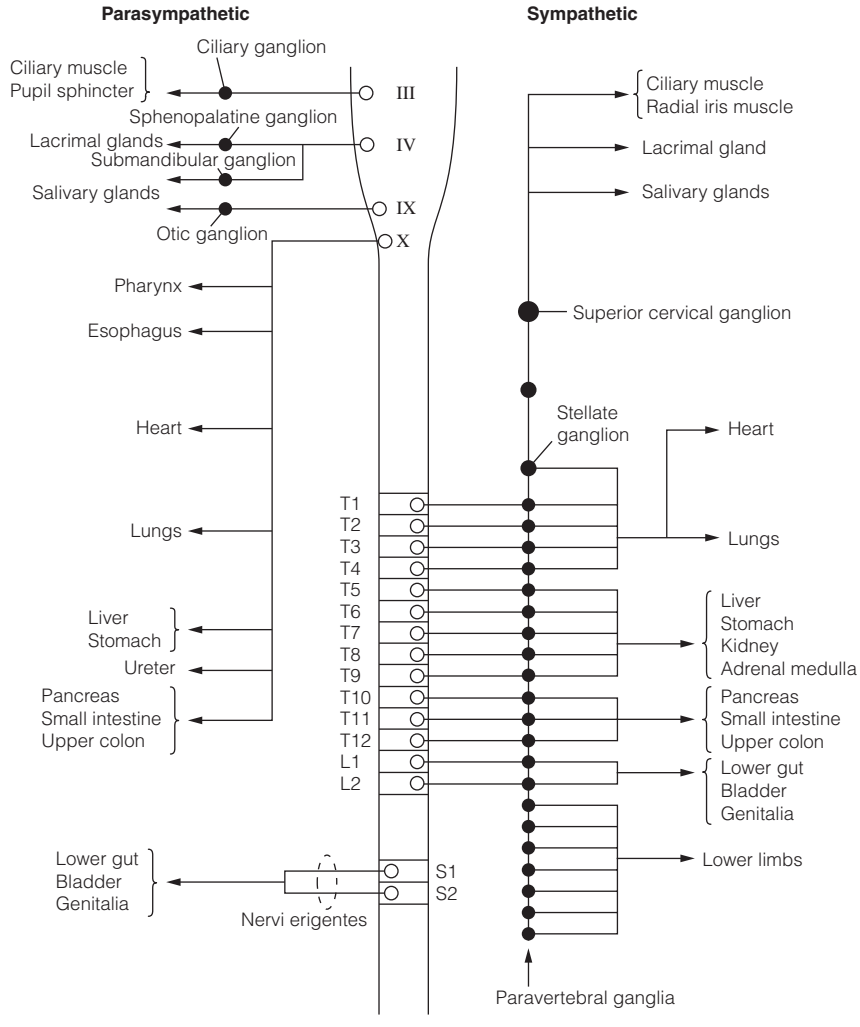


Fig. 1. Distribution of the ANS to target organs.

Almost all terminals of the **sympathetic postganglionic** axons secrete noradrenaline (norepinephrine), the only important exception in humans being the *cholinergic* sympathetic supply to sweat glands. The adrenal medulla chromaffin cells are regarded as a postganglionic component of the SNS and secrete adrenaline (epinephrine) (and noradrenaline (norepinephrine)) directly into the blood as a result of activity in the preganglionic sympathetic fibers which supply them. There are four major types of adrenoreceptors that mediate the various effects of sympathetic stimulation. They are all G-protein linked receptors and their properties are summarized in *Table 1*.

All **parasympathetic postganglionic** axons release ACh, the effects of which are brought about by muscarinic cholinergic (mAChR) receptors. There are several subtypes of mAChR, all G-protein linked, which account for the diverse effects of parasympathetic activity (*Table 1*).

Table 1. Properties of adrenoceptors and muscarinic receptors and their principal actions at ANS targets

Receptor	G protein	Second messenger	Major tissues	Effect
α1	Gq	IP_3/DAG	Vascular smooth muscle Sphincters[a]	Contraction
α2	Gi	↓ cAMP	Adrenergic terminals (presynaptic)	↓ NA release
β1	Gs	↑ cAMP	Cardiac muscle	↑ Force of contraction ↑ Rate
β2	Gs	↑ cAMP	Airway smooth muscle Gut smooth muscle[b]	Relaxation
			Liver	Gluconeogenesis Glycogenolysis
β3	Gs	↑ cAMP	Fat cells	Lipolysis
M1	Gq	IP_3/DAG	Autonomic ganglia	Close K_m channels
M2	Gi and Go	↓ cAMP opens K^+ channels	Cardiac muscle Sphincters[a] Gut smooth muscle[b] Airway smooth muscle	↓ Rate Relaxation Contraction
M3	Gq	IP_3/DAG	Exocrine glands Endothelium	↑ Secretion NO release

[a] Includes gut and genito-urinary sphincters. [b] Except sphincters. IP_3, inositol triphosphate; DAG, diacylglycerol; NA, norepinephrine; cAMP, cyclic adenosine monophosphate; NO, nitric oxide.

Autonomic nerve terminals, in addition to secreting noradrenaline (norepinephrine) or ACh, also release ATP and peptides as cotransmitters. ATP, for example, acts on the smooth muscle of blood vessels to produce fast excitatory postsynaptic potentials and rapid contraction. This is followed by a slower response due to noradrenaline (norepinephrine). Peptide cotransmitters include neuropeptide Y (NPY) and vasoactive intestinal peptide (VIP). They prolong and modulate the effects of the primary transmitter. For example, NPY in sympathetic terminals enhances the vasoconstrictor response to noradrenaline (norepinephrine). VIP from parasympathetic terminals on salivary glands causes vasodilation, which enables ACh to produce a greater salivary secretion.

An example of coordinated autonomic control: the urinary bladder

The urinary bladder receives a dual autonomic innervation (*Fig. 2*). The sympathetic supply to the bladder wall causes relaxation of the detrusor smooth muscle via β2 adrenoceptors. By contrast the internal sphincter smooth muscle has α1 adrenoceptors and is contracted by sympathetic activity. Hence tonic sympathetic activity allows the bladder to fill and prevents voiding. Contraction of the striated muscle of the external sphincter by somatic motor neurons in the sacral spinal cord also helps to maintain continence. Urination (**micturition**) requires the activation of a parasympathetic reflex. Stretch receptors in the bladder wall signal via Aδ and C visceral afferents which enter the sacral spinal cord to synapse with preganglionic parasympathetic neurons. The parasympathetic supply causes contraction of the detrusor muscle but relaxation of the internal sphincter. In addition, collaterals of the afferents establish a long reflex arc via the pons which inhibits both the sympathetic input to the bladder and the motor neurons to the external sphincter. Conscious control over micturation is exerted by descending pathways acting on the pontine neurons. This is lost in patients in whom the spinal cord is severed; in these people micturition is entirely controlled by the spinal reflex and this can be trained to produce urination 'on demand'.

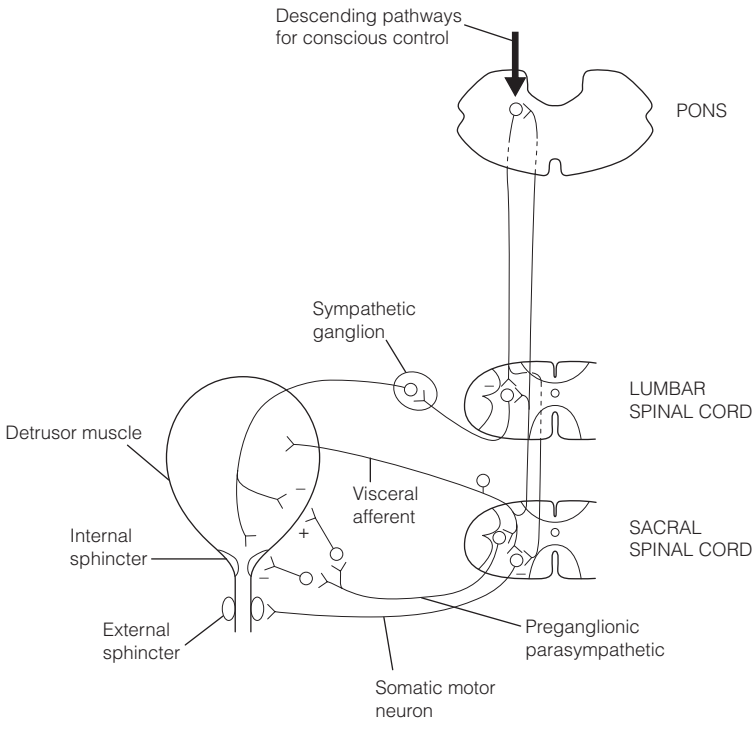


Fig. 2. Control of the urinary bladder. +, Excitatory; –, inhibitory; synapses without a symbol are excitatory.

Enteric nervous system (ENS)

The enteric nervous system (ENS), often regarded as the third division of the autonomic nervous system, is responsible for coordinating gut motility (myenteric plexus), secretion and absorption (submucosal plexus). Most myenteric neurons are motor neurons, unipolar cells supplying the smooth muscle of the longitudinal and circular muscle layers. Excitatory motor neurons are cholinergic and corelease tachykinins and neuropeptide Y. Inhibitory motor neurons release vasoactive intestinal peptide (VIP) and nitric oxide (NO). The motor neurons are tonic and fire volleys of action potentials for as long as they are stimulated.

The sensory neurons are multipolar cells that respond either to stretch, or have neurites in the mucosa that can respond to chemical signals in the intestinal lumen or distortion of the mucosa, associated with the presence of partially digested food (**chyme**). The sensory neurons are normally phasic. In response to sustained stimulus they fire just a few action potentials, because each action potential has a large after-hyperpolarization caused by the activation of Ca^{2+}-dependent K^+ channels. However, their activity can be modulated by input from other neurons, the transmitter of which (VIP, substance P or serotonin) causes a slow epsp by closing K^+ channels. This renders the sensory neurons more sensitive to incoming stimulation.

Most myenteric interneurons which couple sensory and motor neurons are cholinergic, and the receptors on motor neurons (and other interneurons) are nicotinic. Hence transmission between interneuron and motor neuron is fast. Typically, activation of a sensory neuron by stretch, caused by a bolus of chyme, stimulates contraction of circular-layer smooth muscle on the oral side, but

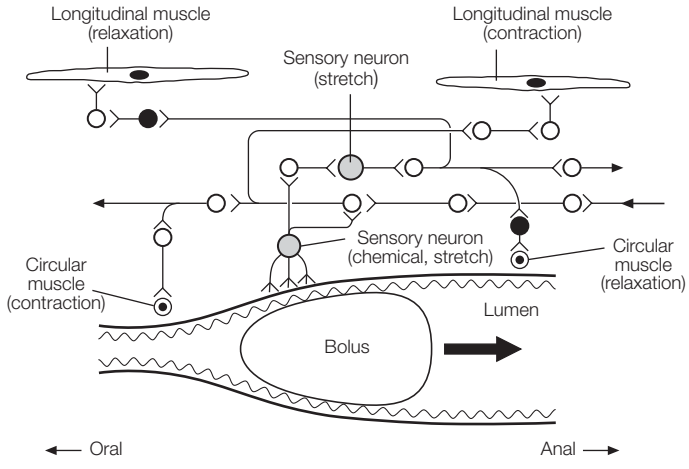


Fig. 3. Elementary gut motility reflex circuitry of the myenteric plexus.

inhibits it on the anal side. Furthermore, in the longitudinal muscle layers smooth muscle is relaxed on the oral side but contracted on the anal side. The effect of this coordinated smooth muscle activity, **peristalsis**, is to move chyme in the anal direction. *Fig. 3* shows the circuitry involved.

The submucosal plexus contains sensory neurons sensitive to chemical stimuli or mechanical distortion of the mucosa. These connect via cholinergic interneurons with stimulatory **secretomotor neurons** which release ACh and VIP. They increase secretion by a direct effect on the glands, and by relaxing smooth muscle of gut arterioles, produce vasodilation that raises local blood flow. Inhibition of the secretomotor neurons by enkephalinergic neurons within the ENS, or by noradrenergic sympathetic nervous system input, acts to promote absorption.

L6 Control of autonomic function

Key Notes

Thermoregulation

Both behavior and physiological mechanisms allow core temperature to be held at a roughly constant 37ºC. Outside a narrow thermoneutral window in which an individual is comfortable, heat gain or heat loss is initiated by negative feedback processes. Small shifts in ambient temperature cause altered sympathetic tone to skin arterioles, producing either cutaneous vasodilation (in the warm), or vasoconstriction (in the cold). Bigger changes in the temperature of the surroundings also trigger sweating, mediated by sympathetic cholinergic stimulation of sweat glands, or shivering, a rapid contraction of muscles driven by the somatic motor system. The posterior hypothalamus integrates signals from internal warm receptors in the hypothalamus and the spinal cord, and signals from skin thermoreceptors. The result of this integration is an appropriate thermoregulatory response. The set point signal ('thermostat') is provided by temperature-insensitive hypothalamic interneurons. The set point falls at night, and is raised by progesterone and immune responses to infection (fever).

Cardiovascular regulation

The ANS is vital for short-term negative feedback control of mean arterial blood pressure (MAP). This is achieved by a combination of modifying the cardiac output and peripheral resistance. Increasing either raises MAP. Adjustments to the tonic output of sympathetic and parasympathetic supply to the heart alters its rate and force; increasing them when the sympathetic dominates. Tonic sympathetic discharge to vascular smooth muscle controls blood vessel diameter and so peripheral resistance. Increased activity causes vasoconstriction which elevates the resistance. Mean arterial pressure is monitored by baroreceptors in the aorta and carotid arteries, afferents from which go to the nucleus of the solitary tract (NST) which, by way of other medullary nuclei, controls preganglionic autonomic neurons. A rise in MAP excites baroreceptors, reflexly activating parasympathetic, but inhibiting sympathetic, neurons. The consequent fall in heart rate and force, and peripheral resistance, restores the blood pressure. Numerous inputs to this circuitry from other brain regions are responsible for alterations to the cardiovascular system in exercise and emotions.

Control of breathing

Although breathing uses the somatic motor system and skeletal muscles, the circuitry involved gets inputs from visceral afferents and is interconnected with central autonomic neurons. Respiratory muscles (e.g. the diaphragm) are driven by rhythmic activity in motor neurons in the cervical spinal cord. Input to these respiratory motor neurons comes from the ventral respiratory group (VRG) in the medulla. A network of VRG cells, some of which have intrinsic pacemaker activity, acts as a central pattern generator to produce the respiratory rhythm. Inputs to the VRG

modify breathing. Several types of receptors in the airways bring about reflex inhibition of inspiration, as does baroreceptor stimulation. Peripheral chemoreceptors in arteries that respond to reduced blood oxygen concentration, and central chemoreceptors in the brain, responsive to a rise in CO_2 or H^+ concentration, stimulate breathing. The basic respiratory rhythm is modified in many activities by other brain regions.

Related topics

Autonomic nervous system function (L5)
Emotion (M1)

Thermoregulation

A core temperature of around 37°C is homeostatically defended by behavioral and physiological mechanisms. Behavior, which includes seeking sun or shade, curling up into the fetal position when cold (to minimize the surface area for radiation), wearing clothes and building shelter, is particularly useful for reducing the effects of extreme environmental temperatures.

Physiological heat loss or heat gain mechanisms are activated whenever the ambient temperature moves outside the **thermoneutral zone**, a window about 1°C wide in which an individual feels comfortable. The position of the thermoneutral zone depends on humidity, wind velocity and clothing. For naked humans in still air at 50% relative humidity it is 28°C. The first response to ambient temperature moving outside the thermoneutral zone is adjustment in sympathetic tone to smooth muscles of skin arterioles. In heat stress, reduced tone results in a fall in noradrenaline (norepinephrine)-evoked vascular smooth muscle contraction and so **cutaneous vasodilation** occurs. This warms the skin, increasing heat loss by radiation. In the cold, increased sympathetic activity causes **cutaneous vasoconstriction**. Larger excursions from the thermoneutral zone evoke either sweating or shivering. Sweat glands are innervated by sympathetic neurons that are atypical in secreting acetylcholine rather than noradrenaline (norepinephrine). ACh acts on muscarinic receptors to trigger sweat production which causes skin cooling by evaporation. Shivering is the almost simultaneous contraction of agonist–antagonist muscle pairs. It starts in masseter (jaw) muscles in humans and spreads to the trunk and proximal limb muscles. Shivering is brought about by activation of brainstem reticular neurons that synapse with γ-fusimotor neurons. The contraction of intrafusal fibers excites stretch reflexes. So, shivering is mediated peripherally by the somatic *not* the autonomic nervous system. Muscle contraction, both in shivering or exercise, generates heat.

A further heat-gain mechanism, **non-shivering thermogenesis**, is particularly important in human babies. It is caused by increased sympathetic activity to brown adipose tissue (BAT), mostly located in the neck and between the shoulder blades. Released noradrenaline (norepinephrine) acts on β3 adrenoceptors to stimulate a rise in cAMP. This activates lipolysis, liberating free fatty acids which are metabolized by β-oxidation in BAT mitochondria, and at the same time uncoupling oxidative phosphorylation in the mitochondria, generating heat (see *Instant Notes in Biochemistry*).

Thermoregulation depends on the integration of signals from two classes of thermoreceptors. Cutaneous warm and cold thermoreceptor afferents, conveying information about skin temperature, run in the spinothalamic tract.

Internal warm thermoreceptors monitor core temperature and are located in the preoptic area of the hypothalamus and the cervical spinal cord. Both cutaneous and internal thermoreceptor afferents go to the posterior hypothalamus, the region of the CNS responsible for driving thermoregulatory responses. Thresholds for activating either sweating or shivering depend on both core temperature and skin temperature. For example during exercise, sweating, triggered by internal thermoreceptors as core temperature rises, is reduced in linear fashion the colder the skin temperature.

The core temperature maintained by thermoregulation is the **set point**. It is defined as the temperature at which neither heat loss nor heat gain mechanisms are activated. The neural signal that acts as the set point 'thermostat' is provided by the integrated activity of interneurons in the hypothalamus that are *not* temperature sensitive. These interneurons are regulated by catecholaminergic neurons in the pontine reticular formation, and the set point is not constant. It shows a circadian rhythm, falling about 0.5°C during sleep and is increased by progesterone during the luteal phase of the menstrual cycle by about the same amount. Chronic exposure to hot or cold environments causes gradual long-term shifts (adaptation) of the set point. During infections, bacterial endotoxins stimulate macrophages to secrete **interleukin-1 (IL-1)** while virus-infected cells produce **interferons**. These **cytokines** (signaling molecules of the immune system) act at the hypothalamus to raise the set point, causing fever.

Cardiovascular regulation

Long-term regulation of blood pressure does not require the ANS and relies on control of blood volume and osmolality via vasopressin and the renin–angiotensin–aldosterone cascade. However, the ANS is crucial for the short-term regulation of **mean arterial blood pressure** (**MAP**). At rest the ANS operates to maintain a roughly constant MAP by negative feedback. Mean arterial pressure is the product of cardiac output (Q), the volume output of the left ventricle per minute, and the peripheral resistance (R), which is related to the radius of the arterioles.

Now, the cardiac output is in turn a product of stroke volume**,** the volume ejected from the left ventricle per beat, determined by the contractile force of the heart, and heart rate. Hence cardiac output can be raised (or lowered) either by increasing (or decreasing) stroke volume or heart rate or both. The ANS regulates cardiac output via both sympathetic and parasympathetic supply to the heart. Both are tonically active at rest and increases (or decreases) in cardiac output are achieved by raising (or lowering) sympathetic activity and reducing (or elevating) parasympathetic activity. In humans sympathetic activity raises both force and rate, whereas parasympathetic activity lowers rate, but has little effect on force since few parasympathetic fibers innervate the ventricles.

Peripheral resistance is controlled solely by altering the tonic firing rates of sympathetic neurons going to vascular smooth muscle. Increased firing frequency causes vasoconstriction which raises peripheral resistance.

Circuitry for the negative feedback regulation of MAP resides in the medulla (*Fig. 1*). Baroreceptors are stretch receptors located in the carotid sinus and the aortic arch that are sensitive to rapid alterations in MAP. Their afferents run in the glossopharyngeal (IX) and vagus (X) cranial nerves respectively and terminate in the nucleus of the solitary tract (NST), a structure involved in a wide variety of visceral reflexes (e.g. swallowing, chemoreceptor responses). The NST projects to the dorsal vagal nucleus (DNX) and to the **nucleus ambiguus (NA),** both of which give rise to preganglionic parasympathetic axons that project via

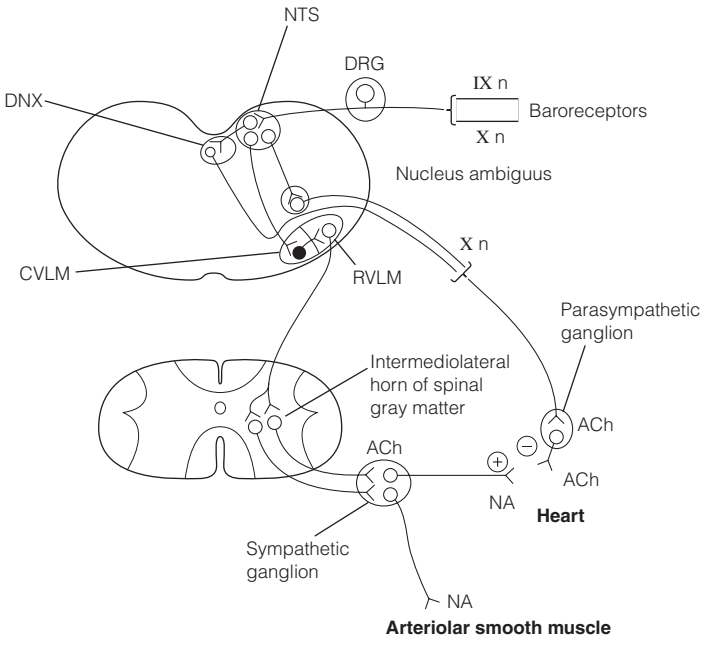


Fig. 1. Brainstem circuits controlling mean arterial blood pressure. DRG, Dorsal root ganglion; DNX, dorsal vagal nucleus; CVLM, caudal ventrolateral medulla; RVLM, rostral ventrolateral medulla; NST, nucleus of the solitary tract.

the vagus nerve to the heart. The NST controls sympathetic outflow to heart and blood vessels by input to the **caudal ventrolateral medulla** (**CVLM**). This contains GABAergic inhibitory neurons which synapse in the **rostral ventrolateral medulla**, axons of which run down the spinal cord, terminating on preganglionic sympathetic neurons. A rise in MAP increases the firing rate of baroreceptor afferents, and this directly activates the parasympathetic innervation to the heart, slowing its rate. However, the presence of inhibitory neurons in the CVLM means that baroreceptor discharge suppresses sympathetic outflow to the heart, reducing its rate and force of contraction, and arterioles, which reduces peripheral resistance. The net effect is a fall in blood pressure back to the set point. Responses occur in the opposite direction to an initial drop in MAP.

Arterial pressure is altered to match circumstances. The **defense reaction**, a stereotyped autonomic response seen in animals faced with sudden danger organized by a **defense area** in the anterior hypothalamus, includes increased heart rate, widespread vasoconstriction and a sharp rise in MAP. Stimulation of the defense area results in inhibition of those neurons in the NST that are driven by baroreceptor afferents. Cardiovascular changes that occur during exercise involve the cerebellar and cerebral cortex which act to modify hypothalamic autonomic regulation. Similarly cardiovascular responses seen in emotional states require elements in the limbic system such as the amygdala and the cingulate cortex.

Control of breathing

The diaphragm and muscles of the chest wall used in breathing are skeletal muscles supplied by motor neurons of the somatic nervous system. However,

the central circuitry which regulates breathing gets sensory input from visceral afferents, and is interconnected with central autonomic circuits controlling the cardiovascular system. This is illustrated by sinus arrhythmia, a change in heart rate with the phase of respiration: during inspiration heart rate rises, during expiration it falls. It is caused by neurons responsible for inspiration inhibiting the preganglionic parasympathetic neurons in the nucleus ambiguus that supply the heart.

Breathing results from the rhythmic discharge of spinal motor neurons which supply ventilatory muscles. Axons of motor neurons in spinal segments C3–C5 run in the phrenic nerves to the **diaphragm**, contraction of which increases chest volume during **inspiration**. Motor neurons in C4–L3 supply neck muscles and external intercostal muscles that aid inspiration and internal intercostal muscles and abdominal muscles responsible for **expiration**. Most of the muscles used for breathing are also important in other functions; for example abdominal muscles are needed to increase intra-abdominal pressure for defecation and vomiting, and during locomotion.

The spinal motor neurons are driven by premotor neurons located in the **ventral respiratory group** (**VRG**) of the ventrolateral medulla. It is here that the respiratory rhythm is generated by a network of cells called the **pre-Botzinger complex** which acts as a central pattern generator. It contains several populations of neurons, each of which fires a burst of action potentials at a specific phase of a respiratory cycle. Some of these cells have intrinsic pacemaker properties. Both inspiratory and expiratory VRG neurons are found and their axons make excitatory (glutamatergic) connections with the motor neurons.

Neurons in the NST receive sensory input, largely via the vagus nerve, from:

- pulmonary stretch receptors which inhibit inspiration;
- nociceptors (lung irritant receptors) which trigger the cough reflex;

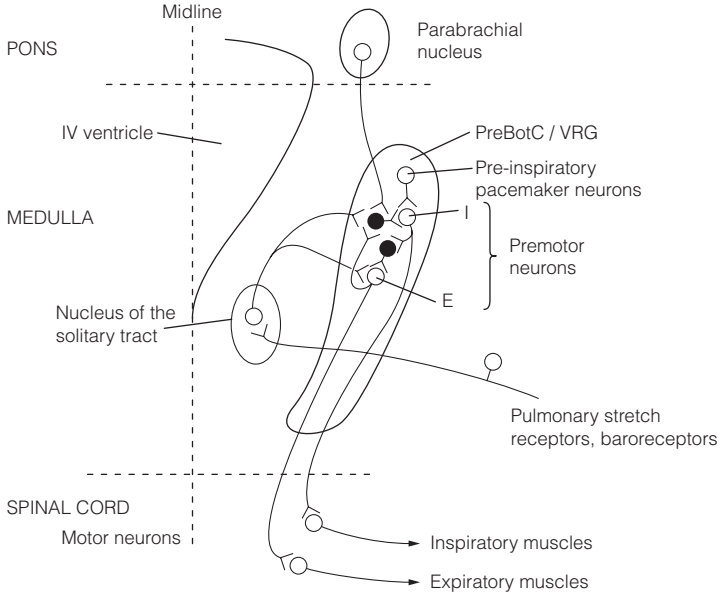


Fig. 2. A highly simplified model of the central control of respiration. E, Expiratory neurons; I, inspiratory neuron; PreBotC, pre-Botzinger Complex; VRG, ventral respiratory group.

- baroreceptors;
- peripheral chemoreceptors in the carotid body and aortic arch which monitor blood concentrations of O_2.

In addition there are central chemoreceptors, sensitive to brain extracellular fluid concentrations of CO_2 and H^+, located in the NST itself. Inputs from all these sources modify the basic rhythm of breathing (*Fig. 2*).

Baroreceptor discharge inhibits inspiration. So if blood pressure falls (as a result of hemorrhage, for example) depth of inspiration increases. Peripheral chemoreceptors activated principally by a reduced partial pressure of O_2 and central chemoreceptors stimulated by raised partial pressure of CO_2, and fall in pH, drive increased depth of breathing via neurons in the NST. The respiratory rhythm is modified by input from the pons and altered in many situations (e.g. sleep, exercise, emotional states and speech) so it is clear that many other brain regions are involved in the control of breathing.

M1 EMOTION

Key Notes

Functions of emotion

Emotions are the unconscious evaluation of a situation as beneficial or harmful. Emotional responses may be hard-wired but many are learnt. Emotions have a visceral sensory component produced by autonomic and endocrine responses, a motor component and a cognitive component. The subjective experience of emotion (feeling) includes the perception of internal changes and the cognitive analysis of the situation, and these two aspects are self-reinforcing. Emotions have obvious survival value, in part as non-verbal communication to guide social interactions.

Affective basal ganglia loop

At the heart of the limbic system are the affective basal ganglia loop and amygdala. The affective basal ganglia loop consists of ventral striatum (nucleus accumbens) which projects to the ventral pallidum of the globus pallidus that relays via the mediodorsal thalamus to limbic cortex. The loop is closed by connections from the cortex back to the nucleus accumbens. The affective loop has reciprocal connections with the amygdala, responsible for fear learning, and is modulated by the brain reward system. Nucleus accumbens output goes to the substantia nigra and hence to motor and cognitive basal ganglia loops, providing for some of the motor and cognitive aspects of emotional states.

Affective motor pathways

Output of the amygdala goes to the hypothalamus and a variety of brainstem structures that organize motor, arousal and visceral aspects of emotions. Facial expressions engendered by emotions, e.g. smiling, are brought about by extrapyramidal motor pathways in the brainstem reticular system, not the corticobulbar fibers used for intentional movement.

Papez circuit and learning

The limbic cortex projects via the entorhinal cortex to the hippocampus, which in turn sends output to the hypothalamus. Pathways from hypothalamus to anterior thalamus and then limbic cortex close a Papez circuit. This is thought to be concerned with explicit learning that occurs during emotional states.

Amygdala and fear learning

The amygdala is a cluster of nuclei in the temporal lobe. Olfactory input enters the corticomedial nucleus. Other input (sensory, arousal state and cognitive information) is relayed to the basolateral nuclei. The output of the amygdala comes from its central nucleus and goes either to the hypothalamus, septum, and several brainstem nuclei via the stria terminalis, or the nucleus accumbens via the ventral amygdalofugal pathway. The amygdala is largely concerned with innate and learned fear responses. It evaluates threats on the basis of prior learning and organizes the appropriate avoidance behavior. The amygdala is necessary for recognizing facial expressions of emotion.

Neocortex and emotion

Connections between different regions of association cortex are the routes by which cognition influences emotions and vice versa. A number of cortical areas are implicated in emotion including the orbital prefrontal cortex and anterior cingulate cortex. Surgical disconnection of the orbital frontal cortex to relieve emotional distress in patients with depression, while effective, produces undesirable personality changes and deficits in executive functions such as planning.

Related topics

Anatomy of the basal ganglia (K7)
Types of learning (O1)
Physiological psychology of memory (O2)
Cell physiology of learning (O3)

Functions of emotion

Emotion may be defined as the unconscious evaluation of a situation as potentially harmful or beneficial. A **feeling** is the conscious subjective experience that corresponds to an emotion. **Mood** is the background flux of feelings experienced at a particular time.

Emotions arise in response to changes in the state of the world that could have important consequences. These consequences may be immediate (e.g. being confronted by a mugger) or delayed (e.g. anticipating an examination). Some emotional responses are hard-wired – executed by neural circuits that are genetically specified during development – such as the universal aversive reaction of infants to bitter-tasting (potentially toxic) foods, but most are probably learnt. Note that apparently hard-wired responses need not be forever fixed; most adults come to like the bitter foods their culture teaches them is safe (e.g. coffee).

Emotional states have three components:

- a visceral sensory component caused by autonomic and endocrine events;
- a motor component, particularly involving the facial muscles;
- a conscious analytical (cognitive) component.

The subjective experience of emotion (feeling) includes the perception of changes in the viscera and any conscious analysis of the situation. It is likely that these two aspects of the emotional state are self-reinforcing because there are learned associations between visceral sensations and emotional states. Realization of just how bad or good a situation is drives visceral changes, while conscious efforts to stem visceral sensation (e.g. controlled breathing) lessen emotional intensity.

Emotions have survival value for several reasons.

1 They are arousing and direct attention to important aspects of a situation so that it can be assessed as threatening or beneficial.
2. Emotions are goads to useful action. We usually avoid snakes but try to get close to those we love.
3. Motor components of emotions (e.g. laughing or crying) communicate our emotional state to others, altering *their* behavior. This is crucial for social interactions. Before it acquires language an infant can only communicate its needs and desires by expressing its emotions.

Affective basal ganglia loop

The core of the limbic system consists of the **affective striato-thalamo-cortical circuit** and its connections with the amygdala (*Fig. 1*). The affective basal ganglia loop is laid out in much the same pattern as the motor loop, except that the striatal component is the **ventral striatum** (**nucleus accumbens, nAc**) which projects to the **ventral pallidum** (ventral part of the globus pallidus). The ventral pallidum relays via the **mediodorsal thalamus** to the **anterior cingulate cortex** and **medial orbital prefrontal cortex**. The loop is closed by connections from the cortex back to the nucleus accumbens. The affective loop has reciprocal connections with the amygdala, which is responsible for fear learning, and is modulated by the dopaminergic mesolimbic system which is concerned with reward learning.

Output of the nucleus accumbens goes to the compact part of the substantia nigra. This allows the activity of the motor and cognitive striato-thalamo-cortical circuits to be modified, providing for some of the motor and the cognitive aspects of emotional states.

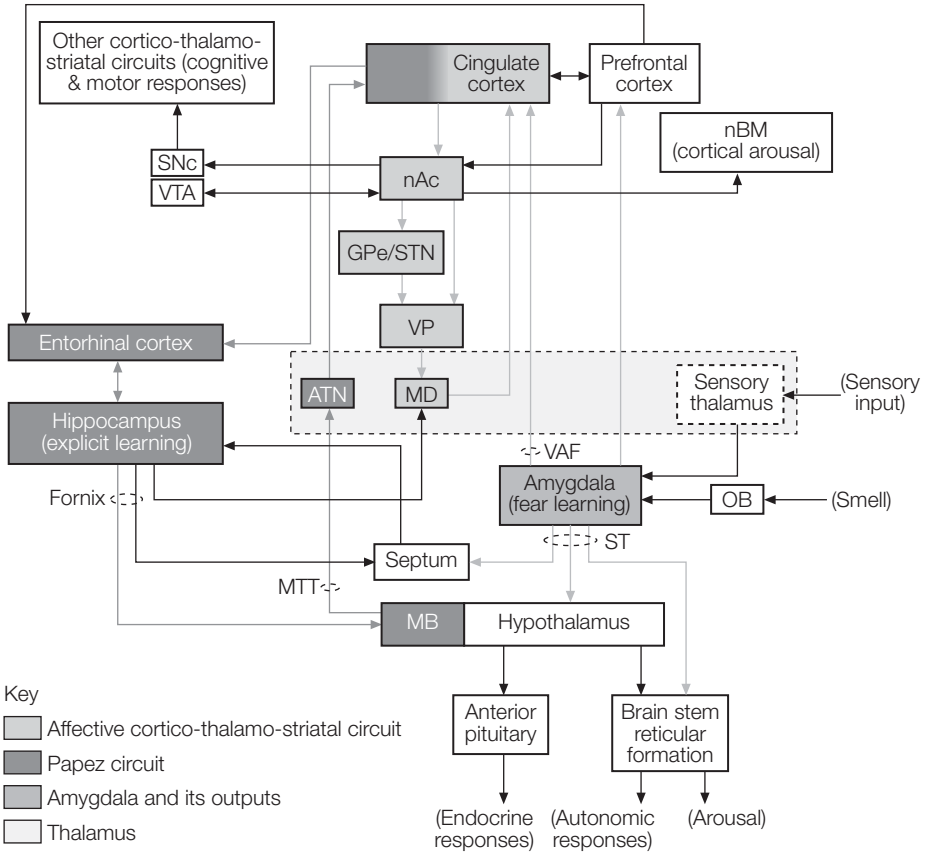


Fig. 1. Circuitry implicated in emotions. MB, mamillary bodies; ATN, anterior thalamic nucleus; MD, mediodorsal nucleus; nAc, nucleus accumbens; GPe/STN, globus pallidus (external part)/subthalamic nucleus; VP, ventral pallidum; SNc, substantia nigra (compact part); VTA, ventral tegmental area; nBM, basal nucleus of Meynert; VAF, ventral amygdalofugal pathway; ST, stria terminalis; MTT, mamillothalamic tract; OB, olfactory bulb.

Affective motor pathways

Output of the amygdala goes to the hypothalamus and a variety of brainstem structures that organize motor, arousal and visceral aspects of emotions. Facial expressions engendered by emotions, e.g. smiling, frowning, crying, etc., are brought about by extrapyramidal motor pathways that run in the brainstem reticular formation. Patients with unilateral damage to corticobulbar fibers descending from the motor cortex have voluntary motor paresis on the opposite side. When asked to smile on demand their smile is lop-sided. However, when genuinely amused their smile is natural and bilateral (**Duchenne smile**) because different emotion-driven motor pathways are engaged.

Papez circuit and learning

The affective loop is wired into a second circuit. The anterior cingulate cortex projects to the entorhinal cortex that acts as a gateway for all neocortical input to the hippocampus. Efferents leave the hippocampus by way of the fornix for the hypothalamus. Output from the mammillary bodies (by way of the mammillothalamic tract) goes via the **anterior thalamic nuclei** back to the anterior cingulate cortex. The hypothalamus also has connections with the prefrontal cortex. This **Papez circuit** was originally thought to be *the* circuit for emotion, but the well-established function of the hippocampus in memory consolidation shows its role to be more restricted, being largely involved in **explicit learning** (learning that can be consciously recalled) during emotional states (e.g. remembering the location of a hornet's nest so as to avoid being stung again).

Amygdala and fear learning

The **amygdala** is a cluster of nuclei in the white matter of the temporal lobe lying anterior to the tail of the caudate nucleus. Olfactory input runs from the olfactory bulb to the **corticomedial nucleus**. Sensory information from other modalities (vision, hearing, somatosensory) enters the **basolateral nuclei** from specific thalamic nuclei and their corresponding areas of sensory cortex. For example, auditory input is relayed from medial geniculate nucleus and auditory cortex, and pain from the posterior insula. The basolateral nuclei also receive information about:

- the state of the viscera from the hypothalamus;
- arousal status from the locus ceruleus and nucleus basalis of Meynert;
- cognitive processing by the orbital prefrontal cortex.

The output of the amygdala is from its **central nucleus** and follows two anatomical pathways. Efferents to the hypothalamus, septum, and several brainstem nuclei go via the **stria terminalis**, while the **ventral amygdalofugal pathway** conveys connections to the nucleus accumbens. The functions of these outputs in the expression of emotions are summarized in *Table 1*.

The amygdala is implicated in innate fear responses. It has access to hardwired neural representations of universal scary things (e.g. crocodiles) to drive automatic fear responses. It is also essential for aversive learning. Fear conditioning occurs when a neutral stimulus (CS) such as a tone is paired with a noxious stimulus (US) such as a brief electric foot shock. After several tone–shock pairings the tone becomes a negative reinforcer and it elicits conditioned fear responses (CR), including autonomic, endocrine and behavioral signs of fear. Neurons in the **central nucleus** of the amygdala fire in a way that correlates with the development of the fear responses. Lesions of the amygdala prevent acquisition of new conditioned fear responses or expression of preexisting ones, although they do not affect *autonomic* responses to aversive stimuli (e.g. the defense reaction) which are organized by the hypothalamus.

Table 1. Effects of amygdala central nucleus efferents.

Target nucleus/pathway	Effect
Periaqueductal gray → raphe nuclei → medullary reticular nuclei	Anti-nociception Freezing
Locus ceruleus	Arousal
Noradrenergic medullary neurons → pre-ganglionic sympathetic neurons	↑ Heart rate, vasoconstriction
Hypothalamus → dorsal nucleus of vagus	↓ Heart rate (mediates vaso-vagal syncope)
Hypothalamus	Corticotrophin releasing hormone secretion → activation of HPA axis (stress response)
Parabrachial nucleus → respiratory CPG	Hyperventilation

Electrical stimulation of the amygdala in humans during surgery evokes feelings of apprehension and fear. The functional connectivity of the amygdala (*Table 1* and *Fig.1*) supports its role in aversive learning since it activates the cholinergic attentional system, the sympathetic nervous system and the release of stress hormones.

Armed with representations of both innate and learned fear-inducing stimuli to compare with the ongoing data stream, the amygdala evaluates the significance of (i.e. the threat posed by) the current situation, and organizes the appropriate visceral and avoidance responses. Several characteristics of amygdala function are noteworthy.

1. The evaluation of a stimulus by the amygdala begins *earlier* than any conscious cognitive appraisal of the situation.
2. Amygdala fear learning is **implicit learning**, which means that the fear responses cannot be consciously generated (although explicit aspects of the fearful situation may be recalled), however any subsequent presentation of the appropriate stimulus does generate fear and escape responses. In infancy and early childhood the amygdala develops more rapidly than the hippocampus (responsible for explicit memory). During this time fearful memories may be acquired which cannot later be consciously accounted for. This could underlie specific **phobias**. By presenting the phobic stimulus in an environment where fear responses are reduced (e.g. with the use of anxiolytic drugs) phobias can often be completely abolished.
3. It seems likely that emotional memories are not stored in the amygdala but in the cerebral cortex with which it is connected. Presentation of fear-evoking stimuli activates visual association cortex and orbital prefrontal cortex as well as the amygdala. **Post-traumatic stress disorder**, in which an image or a sound triggers inappropriate fear responses, often accompanied by recall of some terrible event the afflicted individual has suffered, is presumably due to cortical activation of the amygdala.
4. An important role of the amygdala is in recognizing emotions in human facial expressions. Measurements of regional cerebral blood flow using PET and fMRI show increased activity in the amygdala in subjects shown fearful faces. In the left amygdala the extent of the response increases the more fearful the expression. This is impaired in people in whom the amygdala is

calcified, even though they are still able to identify individual faces. This confirms that the neural system for emotional memory is distinct from that for explicit memory of faces.

5. At the cellular level fear learning by the amygdala involves NMDA receptor-dependent long-term potentiation remarkably similar to that underlying hippocampal learning.

Neocortex and emotion

Extensive interconnections between different parts of association neocortex allows cognition to influence emotional states (e.g. although feeling very anxious we do sit the examination, because we recognize that in the long term doing so will bring benefits), and *vice versa* (e.g. we know that the chance of our child being harmed on walking to school is negligible, but we feel sufficiently anxious to drive her anyway). Lesions of the orbital prefrontal cortex reduce emotional responses (e.g. aggression) in primates, while in humans lesions of the anterior cingulate cortex reduce the emotional distress of chronic intractable pain.

Disconnection of the orbital frontal cortex (**prefrontal leucotomy**), first done in 1935 in an attempt to relieve emotional distress in patients with depression, does reduce anxiety, but produces undesirable personality changes including inappropriate social behaviors, increased impulsiveness and risk taking, together with deficits in executive functions such as planning. The unwanted effects, the advent of more effective psychopharmacology and ethical concerns means that this type of **psychosurgery** (surgery to treat psychiatric disorders) is rarely done today.

M2 MOTIVATION AND ADDICTION

Key Notes

Motivated behavior

Motivated (goal-directed) behavior is aimed at achieving a specific outcome, and is driven by internal states and external cues. Some are homeostatic. Physiological deficits produce internal states (e.g. hunger) that motivate appetitive behavior (e.g. the search for food) and consummatory behavior (e.g. eating) to correct the deficit. Most motivated behaviors are not homeostatic and are motivated by internal states that are poorly understood. A stimulus that increases the chance of a motivated response is a positive reinforcer; if it decreases the probability it is a negative reinforcer.

Dopaminergic reward system

The dopaminergic mesolimbic pathway is a brain reward system which motivates goal-directed behaviors. Evidence for this comes from studying the contexts in which mesolimbic neurons fire, the effect on behavior of pharmacological manipulation of the mesolimbic system and that electrical stimulation of the mesolimbic system appears inherently highly rewarding in rats. The nucleus accumbens (ventral striatum), a limbic structure that integrates dopaminergic motivational input and information from the amygdala about reinforcers, probably determines whether motivation gives rise to action.

Drug addiction

Addictive drugs are positive reinforcers that take the place of natural reinforcers (e.g. food, sex) in driving the brain dopaminergic reward system. There are four aspects to addiction: tolerance (repeated doses of a drug become progressively less effective), dependence (normal functioning is only possible in the drugged state), withdrawal (unpleasant effects result if the drug is not taken) and craving (an intense desire for the drug, which long outlasts the other components on cessation of drug taking) which is based on the learned association between the pleasure produced by the drug and the context in which it is taken. Tolerance can occur without dependence, but is a necessary prerequisite for dependence.

Neurobiology of addiction

Many addictive drugs (cocaine, nicotine, cannabinoids) stimulate the release of dopamine from the mesolimbic system and various measures of addiction are prevented by destroying mesolimbic neurons or blocking dopamine D1 receptors. Tolerance to cocaine is due to the down-regulation of D1 receptors in mesolimbic system targets such as the nucleus accumbens, so withdrawal results from reduced activity of the reward system. Opiates also activate the mesolimbic system, but opiate addiction cannot be accounted for by this alone. Addiction to ethanol, benzodiazepines and barbiturates does not rely on the mesolimbic system. Sensitization of the ventral tegmental area to a drug may account for heightened learning that contributes to craving. It may involve a mechanism like long-term potentiation. The end result is an enduring reduction of dopamine transmission.

Related topics

Dopamine (D4)
Anatomy of the basal ganglia (K7)
Control of feeding (M3)
Types of learning (O1)

Motivated behavior

Behavior that is driven by internal states or external events and which is aimed at achieving a particular outcome is **motivated** or **goal-directed** behavior.

Some motivated behavior occurs in order to satisfy physiological needs. Deficits in energy substrates (e.g. glucose or lipids) generate neural signals which give rise to an internal state, hunger. This drives **appetitive** or goal-seeking behavior, i.e. foraging for food, and subsequently **consummatory** behavior, eating. These goal-directed behaviors are clearly homeostatic, and self-limiting in the sense that the internal states that drive them are sated by consumption.

Much motivated behavior is not so straightforward since it occurs in the absence of any obvious physiological deficit. Sexual behavior leading to copulation, although not homeostatic, is driven by an internal state (libido) that can be sated. However, parenting behavior is neither homeostatic nor self-limiting and little is understood about the internal state that motivates it, although hormones (e.g. oxytocin) seem to be important. What motivates other goal-directed activities, such as listening to music, exploring a forest path, engaging in a sport or academic study, is currently a mystery.

Any stimulus that increases the probability of a motivated response occurring is a **positive reinforcer**. An animal will work to get access to a positive reinforcer. By contrast a reinforcer is said to be negative if the animal works to *avoid* the stimulus, in which case it is displaying **aversive behavior**. The reinforcing quality of a stimulus depends on context. For example, food is a powerful reward to a hungry person but its positive reinforcing quality diminishes with satiety. However, a particular food may still be a positive reinforcer if it is novel and sufficiently delicious even if the person is not hungry. Hence the motivation to eat is a complex interplay of internal state, external cues and memory. Similar qualifications apply to other goal-directed behaviors.

Dopaminergic reward system

Considerable evidence suggests that ascending dopaminergic neurons are responsible for activating motivated behavior. Dopaminergic neurons in the **ventral tegmental area** (**VTA**) project to the nucleus accumbens (ventral striatum) as the **mesolimbic system** and to the frontal cortex as the **mesocortical system**. The mesolimbic system is thought to signal the hedonic (pleasurable or positively reinforcing) qualities of a stimulus such as food or drink, on the basis of internal states and external reinforcers, activating the appropriate goal-seeking behavior. Evidence for this includes:

1. Firing of neurons in the mesolimbic system is increased by the presence of natural reinforcers such as food.
2. Injection of amphetamine (which elevates synaptic concentrations of dopamine) into the nucleus accumbens increases motivated behaviors.
3. Surgical or chemical (6-hydroxy dopamine) lesions of the mesolimbic pathway, reduces appetitive behavior.
4. **Intracranial self-stimulation** (**ICSS**), in which electrodes are chronically implanted into rat brains and the animals learn to press a lever to deliver a small stimulating current through the electrode. When the electrode is in the **medial forebrain bundle** (**MFB**), which contains the mesolimbic axons, the rats lever-press up to 100 times per minute. Given the choice between food or sex and ICSS, suitably deprived rats choose ICSS, which implies that it is a very powerful reinforcer.
5. ICSS, natural rewards and addictive drugs all increase the release of

dopamine from mesolimbic terminals in the nucleus accumbens, and the reinforcing properties of all three are blocked by dopamine D1 receptor antagonists.

The nucleus accumbens forms part of the anterior cingulate (affective) basal ganglia circuit which is thought to translate motivation into the motor activity appropriate to the goal-directed behavior, with the motivation signal provided by the mesolimbic system. So, for example, when hungry we are more likely to find, prepare and consume food than if we have recently eaten. However, the mesolimbic system is more sophisticated than this. Firing of mesolimbic neurons in response to a natural reward (e.g. fruit juice) or a conditional stimulus – light or tone – with which this has been paired, depends on the *predictability* of the reward. Unexpected or novel rewards elicit a strong response which declines with repeated presentation. Perhaps this explains why we eat more of a meal that consists of six courses of gourmet foods than a single bowl of rice. Predicted rewards have little effect, but omission of a predicted reward *reduces* mesolimbic activity. The immediate response to omission of an expected reward is to persevere with the operation that usually provides it. So, low activity of mesolimbic neurons when a predicted reward is missed could be a signal *not* to switch from the current focus of attention.

Drug addiction

All addictive drugs have positively reinforcing properties and these are primarily responsible for drug-seeking behavior in addicts. Animal operant learning studies have proved useful in investigating both behavioral and physiological aspects of addiction. The positive reinforcing properties of a drug can be assessed by **self-stimulation studies** which measure the extent to which animals (usually rats or monkeys) will press a lever to obtain an oral or intravenous dose of the drug. The harder animals work to get it, the more positively the drug is reinforced.

Drug discrimination studies work by training animals to learn that of two levers, one is associated with the delivery of a food reward plus an intravenous drug dose while the other is associated with a food reward alone. Animals can discriminate these two conditions with greater than 90% accuracy. It is then possible to test whether other substances are like the drug on which the animals have been trained by substituting it and examining the effects on rates of lever pressing.

The context in which drugs are taken is important. This is illustrated by **conditioned place preference** in animals. Animals are first exposed to one environment when drugged and to a different environment when non-drugged. Next, the animals are given a choice between the two environments (they can now move freely between them) and the time spent in each is recorded. With positively reinforcing drugs, animals spend more time in the environment they experienced in the drugged state. This also shows that tolerance depends on context; that is, tolerance is greater when a drug is given in an environment or situation where it is normally administered than in a novel situation. This is an example of **context-dependent learning** and illustrates the importance of learning in addictive behavior.

Drug addiction has four components.

1. **Tolerance**. On repeated administration a drug becomes progressively less effective, so the dose has to be increased if the original action is to be maintained. The precise mechanism for tolerance depends on the drug, but

includes enzyme induction, changes in receptor numbers and alterations to second messengers. Acquiring tolerance to a drug does not necessarily lead to addiction; e.g. moderate drinkers (one or two units per day) have tolerance for ethanol, but they are not alcoholics. For *some* addictive drugs **sensitization** (reverse tolerance) is seen to their rewarding effects.

2. **Dependence**. This occurs when biological changes brought about by the drug are such that normal functioning is only possible when the drug is present.
3. **Abstinence** (**withdrawal**) **syndrome**. If a drug is withheld after dependence is established an abstinence syndrome results which is extremely unpleasant and lasts until the long-term biological changes that brought about dependence have abated. Hence, **addiction** (the need to take the drug repeatedly) can be driven as much by the aversion to withdrawal as by the positive reinforcing qualities of the drug.
4. **Craving**. The intense longing for a drug felt by addicts is a learned response that long outlasts the abstinence syndrome. Addicts form memories which associate the pleasure produced by the drug with the environment and cues that accompany the drug taking. Subsequent exposure to the same context causes craving. Functional brain imaging shows that cocaine administration in addicts causes transient activity in the ventral tegmental area (VTA) and nucleus accumbens. However, a stronger nucleus accumbens signal is seen in addicts who are craving the drug. Craving is also accompanied by increased activity in the dorsolateral prefrontal cortex (implying involvement of cognition circuitry) and medial temporal cortex (explicit learning). It may be related to sensitization. Craving is the major barrier to the successful permanent rehabilitation of addicts.

Neurobiology of addiction

Current theories suggest that many addictive drugs act on the brain reward system, initially enhancing dopamine transmission much like natural rewards. **Cocaine** has served as a model for trying to fathom how addiction works. *In vivo* **microdialysis**, which allows continuous sampling of brain extracellular fluid from specific locations in the brain and on-line measurement of neurotransmitter output, shows that dopamine is released by the nucleus accumbens during self-administration of cocaine. Prior destruction of dopamine terminals in the nucleus accumbens by injecting the toxic dopamine analog, 6-hydroxy dopamine (6-OHDA), or dopamine D1 antagonists, curtails cocaine self-administration.

Cocaine acts by blocking the **dopamine transporter** (**DAT**) in the presynaptic terminals of mesolimbic neurons, limiting dopamine reuptake, so the concentration of transmitter in the synaptic cleft is raised. Tolerance develops because the increase in synaptic dopamine concentration causes down-regulation of postsynaptic dopamine receptors. This means that higher amounts of transmitter, and hence drug, are needed to achieve the same level of dopamine transmission.

When drug use stops, dopamine transmission in the mesolimbic system is decreased. This is because:

- there are fewer postsynaptic receptors;
- there is a decrease in the dopamine release from the nucleus accumbens (as shown by microdialysis).

This rebound drop in the effectiveness of the brain reward system is probably responsible for the **anhedonia** (lack of pleasure given by natural rewards) seen

with drug withdrawal. The stress responses seen on withdrawal are probably related to the increased secretion of corticotrophin-releasing hormone from the hypothalamus and amygdala.

Addictive drugs other than cocaine also act on the mesolimbic system. **Nicotine** enhances the synthesis and release of dopamine by acting on nicotinic cholinergic receptors on the dopamine cell bodies in the VTA, while **cannabinoids** act on excitatory **cannabinoid receptors** on dopamine terminals to increase dopamine release. With opiates (e.g. morphine and heroin) the situation is more complicated. Opiates activate the mesolimbic system by hyperpolarizing VTA GABAergic interneurons that normally tonically inhibit the dopaminergic cells. Tolerance arises not because of a change in the *number* of μ-opioid receptors, but because they couple less effectively to Gi proteins so failing to inhibit adenylyl cyclase as much as normal. Withdrawal from opiates leads to a rebound activation of adenylyl cyclase. However, destruction of the mesolimbic system by 6-OHDA does not prevent the self-administration of opiates (although μ-opioid receptor antagonists do), so other mechanisms must also be at work.

Ethanol, benzodiazepines and barbiturates are all positively reinforcing at least partly by reducing anxiety. They all increase GABAergic inhibition by binding to $GABA_A$ receptors and enhancing the Cl^- ion influx. $GABA_A$ receptor antagonists injected into the amygdala decrease ethanol self-stimulation in rats. With repeated administration ethanol decreases the activity of GABAergic neurons but the mechanism is unclear. The mesolimbic system seems to be needed for the acquisition of ethanol addiction, but not its maintenance, while barbiturates and benzodiazepines do not affect dopamine transmission.

There is evidence that even a single dose of cocaine can make the neurons of the ventral tegmental area more sensitive to the drug. This may be important in the early stages of addiction by enhancing the learning that leads to craving. This sensitization is also seen with morphine and alcohol, and with stress. Moreover, brain imaging shows that cocaine causes an enduring *reduction* of dopamine synthesis in mesolimbic neuron terminals. A hypothesis that links these findings is that sensitization is akin to long-term potentiation, the cellular process thought to underlie learning in the hippocampus. Elevated dopamine (caused by the drug) acts on D1 receptors on glutamatergic terminals in the VTA, enhancing glutamate release. This activates NMDA receptors in the VTA dopamine neuron, and the resulting Ca^{2+} entry recruits 'silent' AMPA receptors into the postsynaptic membrane. Many of the AMPA receptors are of the calcium-permeable variety constructed from GluR1 subunits. The effect is to render the VTA dopamine neurons more sensitive to glutamate and produce a sustained rise in Ca^{2+} entry that initiates pathophysiological changes in the VTA dopamine neurons. Amongst these are decreased numbers of neurofilaments, which reduces axoplasmic transport. The result is impaired trafficking of tyrosine hydroxylase to the nerve terminal and so a reduction in dopamine synthesis. Interestingly sensitization does not occur with intracranial self-stimulation.

M3 Control of feeding

Key Notes

Two-center hypothesis

Lesions of the ventrolateral hypothalamus (VLH) inhibit feeding, while lesions of the ventromedial hypothalamus (VMH) cause overeating. From this the two-center hypothesis proposed that the VLH responds to hunger signals to initiate feeding, whereas the VMH responds to satiety signals and terminates feeding. This hypothesis is no longer accepted because VLH lesions cause many other deficits in addition to that on feeding, and they occur because of the destruction of the nigrostriatal tract that passes through the VLH. Furthermore, VMH lesions produce their effects by altering the autonomic nervous system regulation of insulin secretion. Oversecretion of insulin does elevate food intake, but only by increasing the frequency of meals, and the obesity arises by insulin-stimulated fat synthesis.

Satiety factors

Meal size is controlled by neural and blood-borne satiety signals detected by the nucleus of the solitary tract (NST) and the area postrema (a circumventricular organ). Eating is inhibited by stomach distension, cholecystokinin secreted by the duodenum in response to the products of digestion, and by the rise in plasma osmolality that occurs on feeding. The inhibition of feeding is mediated by oxytocinergic neurons in the paraventricular hypothalamus that project to the NST.

Adiposity signals

Adiposity signals released in proportion to the size of the body fat store exert homeostatic control over body mass in the long term by balancing food intake and energy expenditure. Leptin is an important adiposity signal; its plasma concentrations reflect the amount of body fat and it crosses the blood–brain barrier. Leptin is released from fat cells and reduces food intake and increases energy expenditure. Hence it regulates fat stores by negative feedback.

Central regulation of feeding and satiety

Food intake is controlled by two parallel brain pathways which originate in the hypothalamus. An orexigenic (anabolic) pathway promotes feeding whilst an anorexigenic (catabolic) pathway reduces feeding. The first cells in the orexigenic pathway use neuropeptide Y as a transmitter, the most powerful stimulant of feeding known. The orexigenic pathway projects to the nucleus of the solitary tract, decreasing the effectiveness satiety factors. The anorexigenic pathway uses a number of peptide transmitters including oxytocin and corticotrophin-releasing hormone as transmitters. (Many molecules have dual roles, released as a hormone at one site and as a neurotransmitter at another.) These molecules suppress feeding by increasing the sensitivity of the nucleus of the solitary tract to satiety factors. Leptin reduces food intake by inhibiting the orexigenic pathway and stimulating the anorexigenic pathway.

Obesity

Obesity is produced by a mismatch between food intake and energy expenditure. Obese individuals probably respond more to external cues

than internal (physiological) cues for feeding and have poorer capacity for energy expenditure. Mutations of the leptin gene, or its receptor causes obesity in mice. Most human obesity is acquired. Obese humans have high plasma leptin, reflecting their high body fat, but have leptin resistance triggered initially by a high-fat diet.

Anorexia nervosa

Anorexics self-starve and indulge in excessive physical activity. Many have a misperception of their own weight. One hypothesis argues that when eating, anorexics have elevated secretion of corticotrophin-releasing hormone (CRH) which both promotes anxiety and inhibits feeding.

Related topics

Anatomy and connections of the hypothalamus (L1)
Neuroendocrine control of metabolism and growth (L3)
Posterior pituitary function (L2)
Motivation and addiction (M2)

Two-center hypothesis

Lesions of the **ventrolateral hypothalamus (VLH)** result in **aphagia** (refusal to feed) whereas similar lesions of the **ventromedial hypothalamus (VMH)** cause **hyperphagia** (overeating) and obesity. These findings led to the notion that the VLH was a 'hunger center', the destruction of which prevented the animal from responding to hunger signals, whereas the VMH was a 'satiety center' which once lesioned resulted in an animal that could not respond to satiety signals. In this **two-center hypothesis**, the VLH normally initiated feeding whilst the VMH terminated it. This hypothesis is no longer tenable for the following reasons:

1. Lesions of the VLH produce, in addition to aphagia, **adipsia** (failure to drink), **akinesia** (greatly reduced motor activity) and **sensory neglect**, failure to respond to sensory stimulation. Both akinesia and sensory neglect are seen in Parkinson's disease, caused by a loss of nigrostriatal dopaminergic neurons that run through the VLH. Destruction of this pathway by the injection of 6-OHDA into the substantia nigra produces effects similar to those seen in VLH lesions. Hence VLH lesions destroy dopaminergic neurons and so have general effects on motivation *not* specific effects on feeding mechanisms.
2. Animals with VMH lesions become hyperphagic and obese by eating more often, but the size of individual meals stays the same. This is because VMH lesions damage the central control of the autonomic nervous system which in turn alters the endocrine regulation of metabolism. Normally the secretion of insulin from the β cells of the endocrine pancreas is under autonomic control, being excited by the parasympathetic and inhibited by the sympathetic system. In VMH-lesioned animals the activity of the parasympathetic nervous system is increased but that of the sympathetic system is decreased. This results in oversecretion of insulin, promoting excessive synthesis and storage of triglycerides in fat cells. Moreover, the elevated insulin concentration causes rapid clearance of glucose and amino acids from the circulation. According to the **glucostatic hypothesis** a transient fall in blood glucose (always seen a few minutes before eating in rats) is responsible for initiating feeding. This fall occurs earlier in VMH-lesioned animals because of their higher insulin concentration, and so they feed sooner; in other words the

intervals between meals are shortened. In summary, normally the VMH is responsible for the *timing* of meals.

Satiety factors

Over short timescales the brainstem regulates food intake on the basis of inhibitory **satiety signals** that may be either humoral or neural (*Fig. 1*). The nucleus of the solitary tract (NST) and neighboring circumventricular organ, the **area postrema** (**AP**), act as a functional unit to integrate input from the periphery. Afferent input from taste buds and pharynx, stomach distension, **cholecystokinin** (**CCK**) secreted by the duodenum in response to products of digestion (e.g. fatty acids), and neural signals related to energy metabolism from the liver all inhibit eating. Stretch receptor afferents conveying gastric distension signals from the gut run in the vagus nerve to the NST and CCK inhibits feeding in part by exciting these afferents. (CCK also inhibits feeding by acting centrally as a neurotransmitter.)

The area postrema contains chemosensory neurons which are able to respond to blood-borne factors. Ablation of the AP results in rats eating larger than normal meals, but their total food intake remains normal because they eat less frequently, suggesting that longer-term controls on feeding are unimpaired.

Ingestion of food increases plasma osmolality which normally stimulates drinking. If water is unavailable the rise in osmolality is detected by osmoreceptors in the walls of the hepatic portal vein and feeding is curtailed. This is a mechanism for controlling plasma osmolality by restricting ingestion of osmotically active solutes. Note that it uses a different set of osmoreceptors than those involved in drinking; hypo-osmolality inhibits drinking, it does not stimulate feeding.

The inhibition of food intake by gastric distention, CCK and hyperosmolality is mediated by the paraventricular nucleus (PVN) of the hypothalamus which receives input from the NST. In turn, parvocellular (small) neurons in the PVN which use oxytocin as a transmitter, project back to the NST which controls reflex aspects of food intake such as swallowing. This control is very specific. In decerebrate rats (lacking a functional forebrain) the brainstem triggers swallowing of food, but not water, in food-deprived animals. This response is inhibited by a full stomach, intravenous CCK or hypertonic saline.

Meal termination by these satiety factors occurs in animals in which all

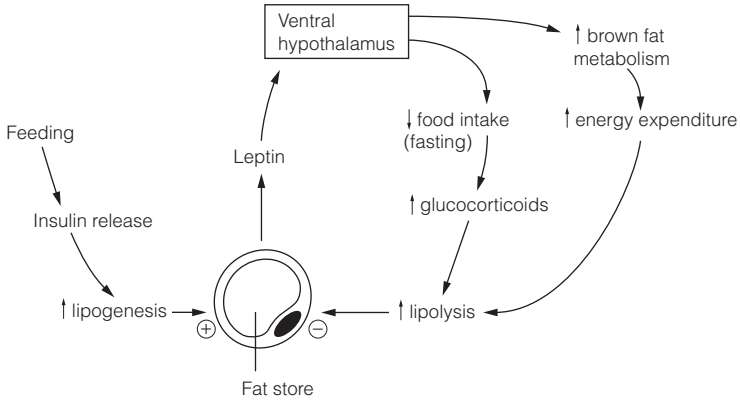


Fig. 1. Negative feedback control of fat stores by leptin. An increase in the size of the fat store causes a rise in leptin release. This activates processes which result in increased fat breakdown.

communication between the hindbrain and forebrain has been cut, although normally reciprocal connections between the NST and the hypothalamus allow for much more sophisticated control over feeding.

Adiposity signals

Despite day-to-day imbalances between food intake and energy expenditure, people are able to match energy input and output to high precision over periods of years. The most important variable that is defended by feedback mechanisms to achieve this is the size of the body fat store (*Fig. 1*). **Adiposity signals** reflect the size of the fat store and regulate it homeostatically by means of mechanisms that balance food intake and energy expenditure. **Leptin**, secreted mostly by adipose cells, is an important adiposity signal. It has a plasma concentration that correlates well with body fat content and crosses the blood–brain barrier. Leptin inhibits feeding and increases energy expenditure.

Leptin secretion is enhanced by the insulin-stimulated lipogenesis that occurs on feeding and is suppressed by the glucocorticoid-stimulated lipolysis that accompanies fasting, so leptin regulates fat mass by means of a negative feedback loop. Leptin secretion is also regulated in anticipation of fat store changes, because food deprivation lowers leptin concentration faster than predicted by shrinking of the fat store. In starvation, insulin levels are low so the transport of glucose into fat cells is minimal and this results in a metabolic signal for leptin secretion. This is an example of feedforward.

Leptin raises energy expenditure by increasing the expression of uncoupling proteins in mitochondria of fat and skeletal muscle, and by acting in the hypothalamus to increase sympathetic activity to brown adipose tissue. Both effects generate metabolic heat.

Central regulation of feeding and satiety

Food intake is regulated in the CNS by two parallel brain pathways which originate in the **arcuate nucleus** of the hypothalamus. An **orexigenic** (anabolic) pathway promotes feeding whilst an **anorexigenic** (catabolic) pathway reduces feeding (*Fig. 2*).

Arcuate nucleus cells that use **neuropeptide Y** (**NPY**) and **Agouti-related protein** (**AGRP**) as co-transmitters are the first order neurons in the orexigenic pathway. NPY is the most potent stimulant of feeding known. NPY neurons project to the lateral hypothalamus, synapsing with second-order neurons that use peptides termed **orexins** (**hypocretins**) as neurotransmitters. These relay to neurons in the nucleus of the solitary tract (NST), decreasing their sensitivity to satiety factors (e.g. glucose).

First-order cells in the anorexigenic pathway contain **pro-opiomelanocortin** (**POMC**), the precursor protein for several biologically active peptides, one of which, **melanocortin**, is a neurotransmitter of these neurons. Other first-order cells in the anorexigenic pathway express **cocaine- and amphetamine-related transcript** (**CART**), one of a growing list of peptides that promote negative energy balance. POMC/CART cell axons run to the paraventricular nucleus of the hypothalamus which contains neurons that use oxytocin, thyrotrophin-releasing hormone or corticotrophin-releasing hormone as transmitters. (Many molecules have dual roles, released as a hormone at one site and as a neurotransmitter at another.) These molecules suppress feeding by increasing the sensitivity of NST neurons to satiety factors.

Leptin receptors exist on first-order neurons of both orexigenic and anorexigenic pathways. However, leptin inhibits the orexigenic pathway and stimulates the anorexigenic pathway. The effect is that leptin reduces food intake.

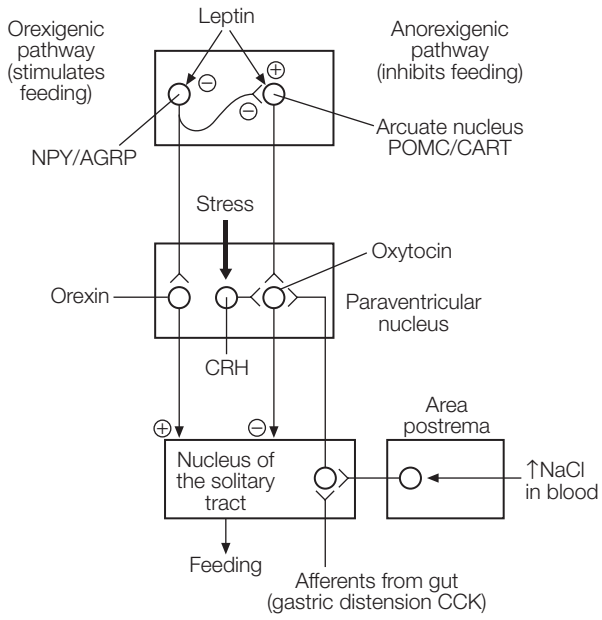


Fig. 2. Central pathways controlling feeding. CRH, corticotrophin-releasing hormone; POMC, pro-opiomelanocortin neuron (releases melanocortin); NPY, neuro-peptide Y; AGRP, Agouti-related protein; CART, cocaine- and amfetamine-related transcript.

Obesity

Over half of USA and UK citizens are overweight, and about one sixth are obese. In a few instances obesity is genetic; mutant mice that lack either functional leptin or leptin receptors are obese. However, the overwhelming majority of people are overweight because over the long term their energy expenditure is lower than their energy intake. Experiments in which obese individuals must work to get food suggest they respond more to external cues (how appetizing food seems) than to internal cues (hunger and satiety) than lean people. Individuals vary by up to 30% in 'basal metabolic rate' and those with low metabolic rates may gain weight despite *apparently* not overeating.

Humans with acquired obesity have elevated plasma leptin concentrations – reflecting the size of their fat stores – but fail to respond to it. This leptin resistance means that despite having leptin concentrations that reflect their weight, the set point of body fat they defend by energy homeostasis is raised. Leptin resistance is associated with genetic defects in the leptin receptor or in its signaling pathway and by failure of leptin transport across the blood–brain barrier. Of most relevance to the current epidemic of obesity in the Western world, leptin resistance is triggered by a high-fat diet, which may act by suppressing leptin receptor signal transduction. As fat stores and leptin levels rise, leptin resistance is further reinforced by down-regulation of leptin receptors. Moreover, as blood leptin concentrations rise they eventually exceed the capacity of the leptin transporter at the blood–brain barrier. Hence, the CNS leptin concentration reaches a steady state and the hormone becomes ineffective in preventing further weight gain.

Recent discoveries relating to the central control of feeding suggest several potential treatments for acquired obesity including:

- NPY receptor antagonists;
- orexin receptor antagonists;
- CART receptor agonists;
- melanocortin (MC4) receptor agonists;
- CRH receptor agonists.

Some of these interventions are likely to be problematic; e.g. orexins are involved in sleep as well as feeding, and CRH receptor agonists are anxiogenic. A potentially promising molecule is **ciliary neurotrophic factor** (**CNTF**), normally produced by glial cells. This reduces food intake and the weight of mutant mice lacking functional leptin, possibly by activating the signal transduction pathway in leptin receptor-expressing cells in the anorexigenic pathway. CNTF analogs may thus prove useful in treating obesity.

Anorexia nervosa

Anorexia nervosa is an eating disorder, mostly of young, white, middle class women, characterized by pathological weight loss due to extremely low food intake, often accompanied by very high levels of physical activity. Sufferers often have a distorted perception of body image, regarding themselves as overweight even when they are not, or may be denying their sexuality by retaining or acquiring a pre-pubertal physique. Several hypotheses have been proposed to explain it. Any hypothesis is blighted by the difficulty in deciding whether any neurochemical findings are the cause, or an effect, of the starvation.

The starting point for one model is that for anorexics eating is very stressful. Normally stress is thought to reduce appetite by exciting the corticotrophin-releasing hormone (CRH) neurons of the paraventricular nucleus in the anorexigenic pathway. The CRH acts locally, stimulating the oxytocinergic neurons that inhibit feeding. In addition, feeding (via the secretion of cholecystokinin) excites CRH neurons and, moreover, CRH is known to be anxiogenic. Hence there is a potential link between feeding and anxiety. Anorexics have heightened secretion of corticotrophin-releasing hormone and glucocorticoids, a stress response perhaps caused initially by anxiety about eating, and subsequently reinforced by the normal stress response to starvation. Whatever the reason, a raised CRH secretion in the anorexigenic pathway could lead to both heightened anxiety and suppression of appetite, though it is not clear why this is not overridden by the low leptin concentrations that are seen in untreated anorexics. The orexigenic pathway also appears to be defective in anorexia nervosa because the low leptin concentrations do not result in increases in NPY. Leptin concentrations rise in weight-recovering anorexics in an apparently physiological fashion, but are higher than in normal women with the same body mass for their height.

M4 Brain biological clocks

Key Notes

Intrinsic rhythms

Many physiological parameters (e.g. sleep/waking, core temperature, secretion of anterior pituitary hormones) show circadian (about a day) variation. In the complete absence of external time cues, the most important of which is light, circadian rhythms free run initially with a period of 25 rather than 24 hours. After a couple of weeks, sleep–waking cycles decouple from other functions to free run with a period of 30 hours. Hence, circadian clocks must exist that are entrained by external events.

Suprachiasmatic nucleus

Stimulation, lesion and transplant experiments show that the circadian clock regulating sleep–waking cycles is located in the suprachiasmatic nucleus (SCN) of the anterior hypothalamus. An intrinsically active neural oscillator in the SCN fires with a frequency that varies sinusoidally, peaking in the daytime. The SCN outputs to other parts of the hypothalamus to regulate sleep–waking cycles, autonomic and endocrine functions. Light signals arrive at the SCN by way of the retinohypothalamic tract, to synchronize it to a 24-hour cycle.

Pineal gland

The pineal gland, a circumventricular organ, secretes a hormone, melatonin, during the hours of darkness. Melatonin in the blood crosses the blood–brain barrier to gain access to the SCN. The duration of the melatonin pulse is a direct measure of the length of the night, and so signals the time of year for animals living at latitudes away from the equator. For seasonal breeders, melatonin secretion acting at the SCN controls reproductive cycles. Light inhibits melatonin synthesis by means of a pathway from the SCN (which receives retinal input) via the sympathetic nervous system to the pineal. There are melatonin receptors in the SCN and melatonin will reduce symptoms of jet-lag.

Related topics

- Anatomy and connections of the hypothalamus (L1)
- Neuroendocrine control of metabolism and growth (L3)
- Neuroendocrine control of reproduction (L4)
- Sleep (M5)

Intrinsic rhythms

Many body functions show a **circadian** rhythm, that is they vary cyclically with a period of about a day. Functions regulated in this way include sleep/wakefulness, core temperature and the secretion of anterior pituitary hormones. Humans isolated from all external time cues – for example living in deep caves, sleeping, eating, switching lights on and off whenever they wish – show intrinsic circadian rhythms with a period of about 25 hours initially. This decoupling of circadian rhythms from the normal 24-hour period is called **free running** and shows that there exist intrinsic **circadian clocks** that are usually entrained by environmental cues called **zeitgebers** (the German word for time-giver). Zeitgebers include

light, exercise, social interactions and work schedules. Light is the most powerful zeitgeber. A powerful light pulse given during subjective night, produces shifts in the circadian cyclicity. In humans with normal sleep patterns the nadir in core temperature occurs at about 5 a.m. A light pulse given during the night before this time causes circadian rhythms to be delayed (**phase delay**) whereas a light pulse after this time causes **phase advance**.

After 1–2 weeks free running, physiological variables often desynchronize from each other. For example, typically, fluctuations in core temperature, secretion of ACTH and glucocorticoids, and rapid eye movement (REM) sleep continue with a period of about 25 hours, but the cycles of sleep–wakefulness and secretion of growth hormone (GH) lengthen to over 30 hours. This suggests that there are two circadian clocks. Both are normally entrained by light–dark cycles of day and night.

Suprachiasmatic nucleus

The circadian clock that regulates sleep–waking cycles resides in the **suprachiasmatic nucleus** (**SCN**) of the anterior hypothalamus. Electrical stimulation of the SCN shifts circadian rhythms in predictable ways, destruction of the SCN produces permanent abolition of circadian cyclicity (but does not prevent sleep), and transplanting a fetal SCN into a host animal with its own SCN destroyed restores circadian rhythmicity.

The pacemaker (clock) function of the SCN is due to the presence of individual neurons in the ventrolateral core division of the SCN, which fire with a frequency that varies in a circadian fashion. These neurons are circadian oscillators and maintain their rhythmic activity even when isolated from the rest of the nervous system in cell culture. The firing frequency of SCN neurons varies sinusoidally with a period of 24 hours, peaking during the day, and dropping to its lowest rate during the night. The SCN projects largely to other hypothalamic structures to regulate sleep–wake cycles, autonomic and endocrine functions, but also send output to the thalamus and basal forebrain (e.g. septal nucleus) which probably accounts for circadian variation in memory and cognitive functions. Most SCN neurons are GABAergic and co-release peptides, and are assumed to be inhibitory on their targets.

Light signals encoding total luminance are relayed to the SCN by the **retinohypothalamic tract** (**RHT**). This pathway consists of the axons of a population of small retinal ganglion cells driven by cone photoreceptors over a wide area, which synapse directly with neurons in the core of the SCN. The RHT uses glutamate as a transmitter.

Pineal gland

The **pineal gland** is a circumventricular organ which secretes **melatonin** into the blood during the hours of darkness, and the duration of the pineal melatonin pulse is a direct measure of the length of night, and hence also of day length, the **photoperiod**. Melatonin secreted into the blood is transported across the blood–brain barrier to act on the SCN. For animals living at latitudes other than the equator, day length varies during the year, so melatonin secretion acts as a signal which codes for the time of the year. For seasonal breeders, the length of the melatonin pulse regulates the hypothalamic–pituitary–gonadal (HPG) axis of both males and females via its action on the SCN. For example, in sheep the longer melatonin signals produced in the shorter photoperiods of November (in the Northern hemisphere) activate estrous cycles in ewes, and testicular growth – with a consequent rise in testosterone secretion and spermatogenesis – in the rams. Although photoperiodic control of melatonin secretion is not important for

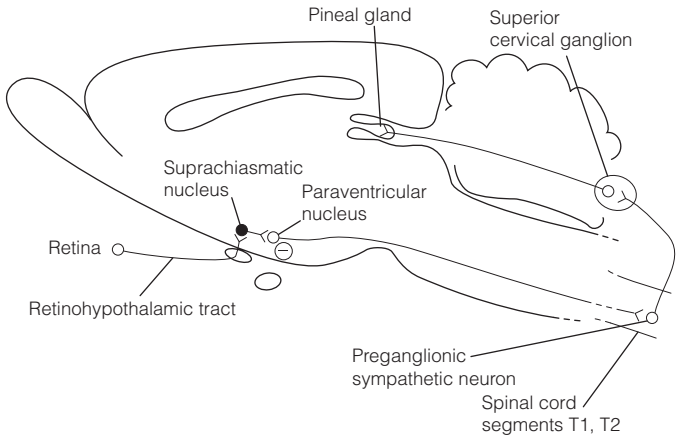


Fig. 1. The brain circuitry by which light inhibits the secretion of melatonin from the pineal gland.

reproduction in humans, it does exert feedback effects on the function of the SCN circadian clock and so affects sleep–wake cycles.

The pathway by which light inhibits melatonin synthesis is circuitous (*Fig. 1*). Neurons in the core of the SCN (which get retinal input from the RHT) inhibit the autonomic division of the paraventricular nucleus (PVN) of the hypothalamus. The PVN sends axons through the brainstem to synapse with preganglionic sympathetic neurons in the intermediolateral horn of spinal cord segments, T1 and T2. These autonomic neurons project to the superior cervical ganglion (SCG), the postganglionic cells of which innervate the pineal gland. At night the activity of SCG neurons is increased, and the secretion of noradrenaline (norepinephrine) from sympathetic terminals acts on β adrenoceptors of **pinealocytes** to synthesize melatonin. The biosynthetic pathway is illustrated in *Fig. 2*.

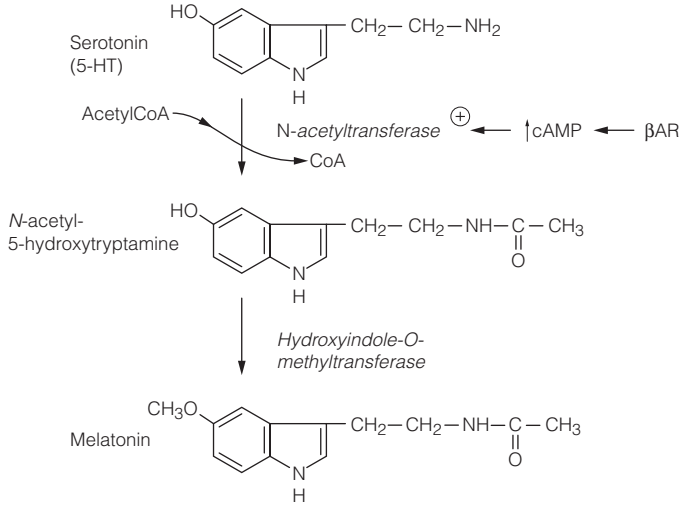


Fig. 2. Synthesis of melatonin in the pineal. N*-acetyl transferase is activated by norepinephrine, secreted from sympathetic terminals, acting at β adrenoceptors (βAR). cAMP, cyclic adenosine monophosphate.*

Melatonin interacts both with metabotropic receptors, linked to Gi proteins which bring about inhibition of adenylyl cyclase, and with one of the steroid receptor superfamily members, in the SCN. Melatonin can entrain the circadian clock in the SCN, reset sleep–wake cycles in animals and humans, and reduce the symptoms of **jet-lag**, the sleep disturbance that arises when light–dark cycles and circadian rhythms are suddenly desynchronized by air travel over several time zones.

M5 SLEEP

Key Notes

Electroencephalography

Electroencephalography (EEG) records the electrical activity of the brain via scalp electrodes. The frequency bands of EEG signals correlate with behavioral state; alpha (8–13 Hz) is seen in relaxed wakefulness, beta (13–30 Hz) with intense mental activity, and lower frequencies are seen in sleep.

States of sleep

EEG and other physiological measures distinguish two states of sleep. Slow wave (or non-rapid eye movement, NREM) sleep has low frequency, high-voltage EEG is characterized by a drop in cerebral blood flow and brain glucose utilization. Muscle tone is retained. As a person falls asleep they drift through four stages of NREM sleep, each with a lower frequency than the last. After about 90 minutes they enter rapid eye movement (REM) sleep which has a similar EEG waveform to the awake state. Muscle tone is absent apart from rapid eye movements and brief twitches of limb muscles. Dreaming occurs mostly in REM sleep. Sleep is homeostatically regulated, with the amount of slow-wave activity reflecting the length of the prior waking period. During a night's sleep the proportion of REM sleep increases and the amount of slow-wave activity falls. Total sleep time, and the proportion of time in REM sleep drops with age.

Sleep pathways

Early studies established that the midbrain was required for wakefulness, but there must be hindbrain regions that actively generate sleep. The ascending reticular formation forms an arousal system. It splits into two branches at the level of the diencephalon. One branch, containing monoaminergic neurons, goes to the cerebral cortex from brainstem, hypothalamus and basal forebrain. The second branch projects to the thalamus and includes axons of pontine cholinergic neurons; these wake-on/REM-on cells are responsible for EEG desynchronization wakefulness and REM sleep. The other neurons (wake-on/REM-off cells) fire at the highest rate in alert animals but are silent during REM sleep. Thalamic relay neurons fire single action potentials in the awake state when depolarized by arousal system input. This allows thalamocortical transmission. However, during NREM sleep when the relay cells are hyperpolarized by a loss of input from the arousal system they go into burst firing mode. This prevents transmission from thalamus to cortex. Cortical neurons now fire in isolation with their own intrinsic (delta) rhythm.

Physiology of NREM sleep

Activity in GABAergic neurons in the preoptic hypothalamus hyperpolarizes thalamic relay neurons so they go into burst firing, NREM sleep, mode. They also inhibit all the monoaminergic neurons of the cortical branch of the ascending reticular formation. The triggers for NREM sleep, which act through the preoptic hypothalamus, include the brain biological clock, elevated core temperature and possibly a number of endogenous sleep-promoting substances.

Physiology of REM sleep

REM sleep may be brought about by GABAergic inhibition from the periaqueductal gray matter. This shuts down the noradrenergic and serotonergic (wake-on/REM-off) neurons, but *not* the pontine cholinergic wake-on/REM-on cells. These cholinergic neurons depolarize the thalamic relay neurons allowing EEG desynchronization, inhibit (via connections with the reticular formation) sensory input and motor output, effectively disconnecting the brain from the external world during REM sleep, and generate spikes of activity that originate in the pons and spread to the lateral geniculate nucleus and visual cortex. The cells that trigger these pontine geniculate-occipital spikes are also responsible for the rapid eye movements.

Disorders of sleep–wakefulness circuitry

Neurons using a peptide neurotransmitter, orexin, are thought to stabilize the brain into waking or REM sleep once it has made the transition into one or other state. Orexin transmission is defective in animals and humans with narcolepsy, in which frequent episodes of REM sleep interrupt the awake state. Loss of cerebral cortical arousal can arise either from metabolic disturbance to the cortex itself or lesions affecting the arousal system.

Functions of sleep

Metabolic hypotheses suggest that sleep is a homeostatic mechanism to correct a metabolic energy deficit that accrues during waking hours. This theory is based on the fact that sleep-deprived rats die of failure of their thermoregulatory and immune systems.

Related topics

- Noradrenaline (norepinephrine) (D5)
- Serotonin (D6)
- Acetylcholine (D7)
- Physiological psychology of memory (O2)

Electroencephalography

Recording the net electrical activity of the brain by means of surface electrodes attached to the scalp is termed **electroencephalography** (**EEG**). Large numbers of cerebral cortical cells fire in synchrony and consequently their summed activity produces potentials large enough that they can be recorded with scalp electrodes. By using an array of electrodes, activity of different brain areas can be examined simultaneously. The recording may be monopolar – each scalp electrode measures the potential with respect to a distant indifferent electrode – or bipolar, in which the potential is measured between a pair of scalp electrodes. The EEG waveform varies in frequency and the frequency ranges are conventionally grouped: alpha (8–13 Hz), beta (13–30 Hz), delta (0.5–2 Hz), theta (4–7 Hz). Alpha is typically seen in relaxed wakefulness and beta when an individual is alert or engaged in intense mental activity. The delta waveforms (often referred to as slow-wave activity) are seen during the deepest stages of NREM sleep. Theta rhythms are thought to be important for learning.

States of sleep

When awake, the EEG waveforms are of low amplitude and high frequency (*Fig. 1*) and described as **desynchronized**. On the basis of the EEG and other physiological measures, two sleep states can be distinguished, **non-rapid eye movement** (**NREM**) sleep and **rapid eye movement** (**REM**) sleep.

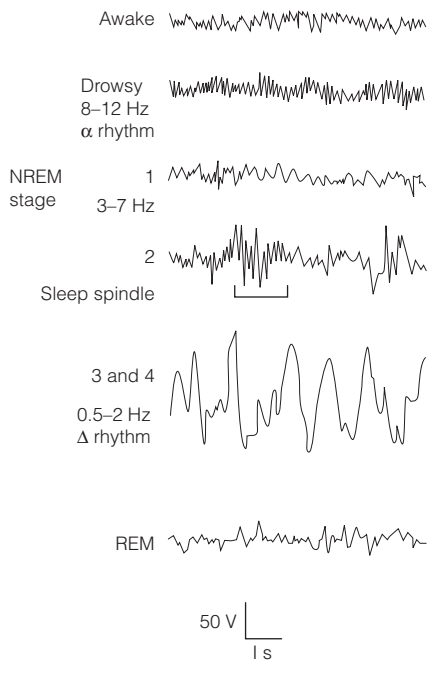


Fig. 1. EEG waveforms recorded from the human brain when awake and asleep.

NREM sleep, has high-amplitude, low-frequency (synchronized) EEG waveforms. During NREM sleep, muscle tone is retained and postural adjustments (turning over in bed) are occasionally made. Respiration rate, heart rate, and mean arterial blood pressure all fall, though gastrointestinal motility increases. Most growth hormone secretion occurs during NREM sleep. At the beginning of a period of sleep the EEG changes from the fully awake to the NREM sleep state by a progressive decrease in EEG frequency (and increase in amplitude) through four stages of NREM sleep (stages 1–4). Stage 2 is interspersed by higher-frequency bursts called **sleep spindles**. Stages 3 and 4 are often collectively referred to as **slow-wave sleep** because it is characterized by slow-wave activity (delta waves). In passing from stage 1 through to 4 it becomes increasingly difficult to arouse the sleeper; if awakened they are confused, and rapidly go back to sleep if left. PET scanning shows that cerebral blood flow and glucose utilization fall by as much as 40% in NREM sleep.

REM sleep is sometimes described as **paradoxical** sleep since its EEG resembles that seen in the awake state. Muscle tone is absent, except for transient contractions of extraocular eye muscles which produce the rapid movements of the eyes for which this stage of sleep is named, and brief contractions of middle ear muscles and distal limb muscles. Respiration rate, heart rate, mean arterial blood pressure and core temperature become irregular. Penile erections occur during REM sleep, and their absence distinguishes physiological from psychological causes of impotence. People aroused from REM sleep usually report that they were dreaming. At this time the dream content is in short-term memory since unless the details are immediately rehearsed they are forgotten within a minute or two. Less frequently dreaming occurs in NREM sleep. In REM sleep the brain is as metabolically active as it is when awake.

Sleep is homeostatically regulated. Prolonged wakefulness results in an increased tendency to fall asleep and this correlates with a rise in the amount of slow-wave activity in NREM sleep, in proportion to the time spent awake. During a typical night's sleep (*Fig. 2*) adults drop rapidly into deep (stage 4) NREM sleep and then REM and NREM sleep alternate about every 90 minutes with increasingly longer periods of REM sleep as the night progresses; i.e. the amount of slow-wave activity falls to baseline.

The proportion of time spent asleep changes dramatically during development. Human fetuses sleep (mostly REM) almost all the time. This falls to 17–18 hours sleep (50% REM) for babies born at term. In infants, sleep is distributed throughout the 24-hour day, but by 3–4 years a child's ten or so hours of sleep is restricted to night-time. The amount of time spent in stage 4 sleep falls exponentially with age, with most of the decline happening over the first 20 years. Between 10 and 70 years the proportion of REM sleep is constant at about 25% of total sleep time and declines in the elderly. Most vertebrates sleep but only homeotherms have REM sleep.

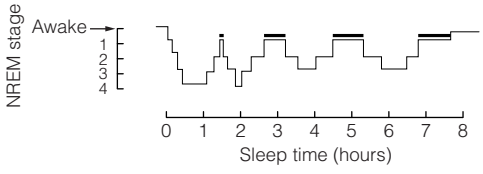


Fig. 2. Distribution of sleep stages during a typical nights sleep. Dark bars are rapid eye movement sleep periods. NREM, Nonrapid eye movement sleep.

Sleep pathways

Early studies established that the brainstem regulates the sleep–wake cycle. The brain of a cat in which the spinal cord is cut at C1 continues to show EEGs characteristic of normal sleep–wake cycles. However, transecting the midbrain results in an isolated forebrain in which the EEG is permanently synchronized in slow-wave sleep activity. More selective lesions showed that the loss of wakefulness was caused by severing the thalamic projections of the midbrain reticular formation, not by cutting the classical sensory pathways to the thalamus. High-frequency electrical stimulation of the midbrain in intact cats causes arousal and desynchronized EEG. Hence the midbrain has neurons that are required for the awake state. However, injecting barbiturate into the artery supplying the caudal brainstem causes EEG desynchronization in cats previously asleep. Hence the caudal brainstem (medulla and back part of the pons) contains neurons which *actively* generate sleep.

Monoaminergic neurons in the **ascending reticular formation** required to generate the awake state, and to increase arousal and the responsiveness of the cortex to sensory input, constitute an ascending arousal system. It splits into two branches at the level of the diencephalon (*Fig. 3*). One branch projects through the lateral hypothalamus to the cerebral cortex from a variety of sources:

- noradrenergic neurons of the locus ceruleus;
- serotonergic cells in the raphe nuclei;
- histaminergic neurons in the tuberomammillary nucleus (TMN) of the hypothalamus.

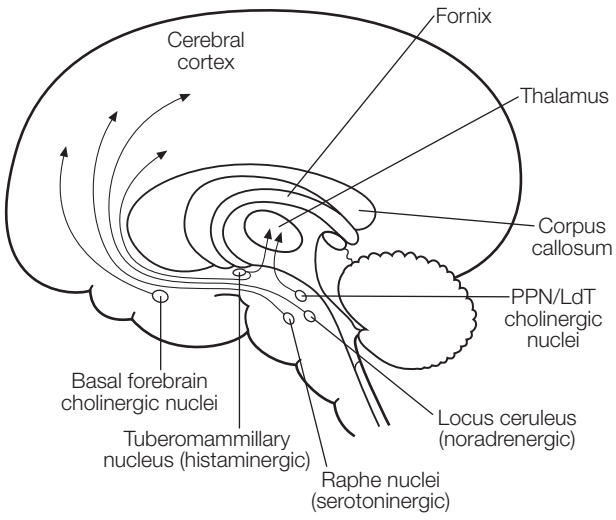


Fig. 3. Ascending arousal system.

This branch is augmented by cholinergic neurons in the basal forebrain.

The second branch projects to the thalamus and contains:

- histaminergic neurons in the TMN which make excitatory synapses with thalamic relay neurons;
- cholinergic neurons in the pons (**pedunculopontine nucleus** (**PPN**) and **lateral dorsal tegmental nucleus** (**LdT**)) which inhibit GABAergic thalamic reticular neurons and hence excite thalamic relay cells.

The *pontine* cholinergic cells are active during wakefulness and REM sleep – and are responsible for the desynchronization of the EEG in these states – but become quiescent during NREM sleep, and hence are described as **wake-on/REM-on** cells. The other neurons fire at the highest rate in alert animals, have low firing rates during NREM sleep and go silent during REM sleep. Because of this firing behavior the noradrenergic and serotonergic neurons in particular are often called **wake-on/REM-off** cells.

Neurons which use **orexins**, located in the tuberal region of the hypothalamus excite all the groups of neurons listed above and are active during the awake state and REM sleep. Orexigenic neurons probably stabilize whichever state (awake or REM) the brain is in.

Only 5–10% of synapses in the thalamus are made by afferent terminals. Most of the thalamus is concerned with controlling which sensory information is sent to the cortex (attention). **Thalamic relay cells** project to the cortex and get reciprocal connections from the cortex (*Fig. 4*). Both relay cells and cortical projection neurons use glutamate and are excitatory. Specific sensory input (e.g. from the retina) directly excites the relay cells. The relay cells are subject to inhibition from GABAergic interneurons and from the **thalamic reticular nucleus** (**TRN**), a sheet of GABAergic neurons covering the thalamus.

Thalamic relay neurons have two modes of firing.

1. Tonic firing of single-action potentials occurs in the awake state when the relay cells are depolarized by input from the ascending arousal system. In this mode, transmission of sensory input from thalamus to cortex takes place.

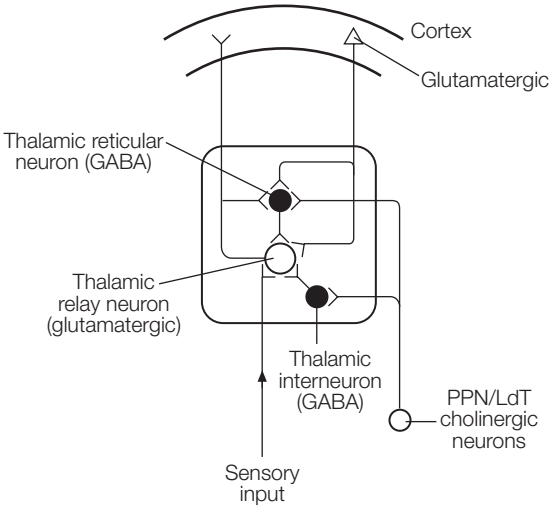


Fig. 4. Neural circuitry of the thalamus.

2. Burst firing is seen during NREM sleep when the relay cells are hyperpolarized by a loss of input from the ascending arousal system. Burst firing occurs because relay cells have T-type Ca^{2+} channels that are activated by hyperpolarization. Opening these channels causes a calcium depolarization that triggers a burst of 4–5 conventional action potentials. Burst firing of the relay cells drives synchronized bursting of cortical cells. The conventional argument is that this limits information transmission between thalamus to cortex. With deepest NREM sleep the thalamocortical neurons become so hyperpolarized they go silent and cortical neurons, now completely decoupled from the thalamus, fire with their own intrinsic (delta) rhythm.

Physiology of NREM sleep

The start of NREM sleep is organized by hypothalamic nuclei and involves suppressing both branches of the ascending reticular formation (*Fig. 5*). GABAergic neurons in the **ventrolateral preoptic area** (**VLPO**) inhibit histaminergic neurons in the tuberomammillary nucleus. The loss of excitation on thalamic relay neurons causes them to hyperpolarize so they go into burst firing, NREM sleep, mode. (The sedative effect of antihistamines which cross the blood–brain barrier is due to reduced excitation of thalamic relay neurons.) VLPO cells suppress the cortical branch of the ascending reticular formation by inhibiting all the monoaminergic and the orexigenic neurons. Inhibition of the orexin neurons:

- shuts off the cholinergic basal forebrain cortical arousal system;
- reduces activity of the noradrenergic and serotonergic (wake-on/REM-off) cells and of the pontine cholinergic (wake-on/REM-on) cells.

Although the loss of activity in the wake-on/REM-off cells lifts their inhibition on the pontine cholinergic cells, the lack of tonic excitation from the orexin neurons keeps them quiet.

The triggers for NREM sleep, which activate the GABAergic VLPO cells, are not well understood but include the biological clock in the suprachiasmatic nucleus, and elevated core temperature detected by warm receptors in the

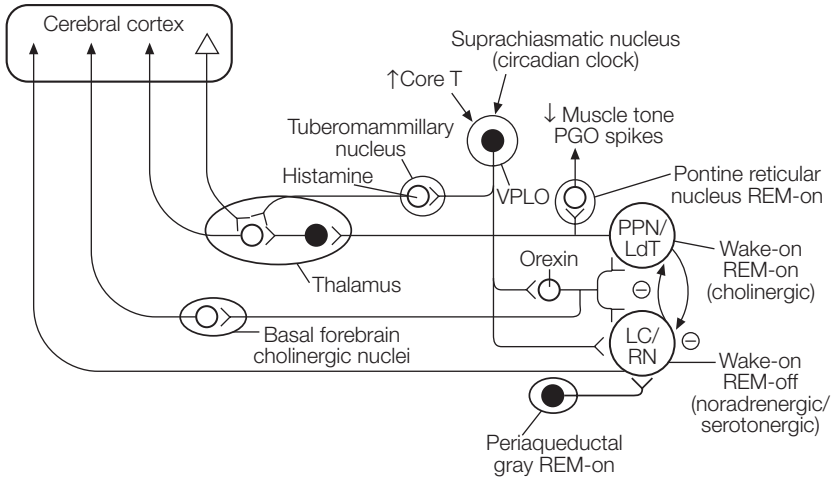


Fig. 5. A model for the neural circuitry regulating sleep and wakefulness. LC/RN, locus ceruleus/Raphe nuclei; PPT/LdT, pedunculo-pontine nucleus/lateral dorsal tegmental nucleus; VLPO, ventral lateral pre-optic hypothalamus. Open circles, excitatory neurons; filled circles, GABAergic inhibitory neurons.

hypothalamus. A number of molecules have also been proposed as candidate endogenous **hypnogenic** (sleep-producing) substances identified in sleep-deprived animals, including the cytokine **interleukin-1** (which may explain the tiredness felt during infective illnesses) and melatonin.

Physiology of REM sleep

During REM sleep (*Fig. 5*) GABAergic inhibition comes from a different source, REM-on cells, possibly located in the periaqueductal gray matter. These shut down the noradrenergic and serotonergic (wake-on/REM-off) neurons, but *not* the orexin cells. Now, pontine cholinergic wake-on/REM-on cells are activated by the combination of continued excitation by orexinergic cells plus disinhibition from the aminergic neurons. In consequence:

- high levels of activity in these cholinergic neurons cause depolarization of the thalamic relay neurons which go into tonic firing mode;
- the continual activity of the orexin neurons keeps the basal forebrain cholinergic cells responsive.

Both of the above effects conspire to desynchronize the EEG.

The pontine cholinergic neurons also organize – via relays through pontine and medullary reticular nuclei – two major features of REM sleep. The first is the powerful inhibition of sensory input and motor output. Presynaptic GABAergic inhibition by reticular neurons on afferent terminals is responsible for the blockade of sensory input, while glycinergic postsynaptic inhibition of motor neurons is the route by which muscle **atonia** (loss of muscle tone) is brought about. Thus, during REM sleep the brain is effectively uncoupled from the external world; it is 'off-line'. Lesions of the pons which prevent the muscle atonia produce animals which express stereotyped behaviors during REM sleep. This suggests that during normal REM sleep motor patterns are generated but not executed; we cannot act out our dreams!

A second major feature of REM sleep are periodic **pontine–geniculate–occipital** (**PGO**) **spikes** in the EEG. These originate from cholinergic (**PGO-on**) cells in the pons. They drive vestibular and reticular neurons to excite oculomotor neurons (causing the rapid eye movements) and other cells to produce the phasic alterations in respiration, heart rate, blood flow and muscle twitches seen in REM sleep. PGO-on cells also initiate the spread of activity to the lateral geniculate nucleus and visual cortex recorded as PGO spikes. During wakefulness PGO-on cells are usually inhibited by serotonergic cells, however PGO spikes can be produced in awake subjects by sudden stimuli, so they may underlie startle responses.

Disorders of sleep–wakefulness circuitry

It is not known what causes the brain to switch from NREM to REM sleep and back several times during the night, and there are several candidates for the GABAergic neurons that trigger REM sleep. However, it is thought that orexigenic neurons stabilize the waking or REM state once the brain has made the transition. This idea is supported by the discovery of deficits in orexin neurotransmission in animals and humans with **narcolepsy**. In this condition patients experience frequent and undesired slips into REM sleep during the day and fragments of the REM state (muscle atonia, dreaming) intrude into wakefulness.

Cerebral cortical arousal can be impaired either by dysfunction of the cortex itself or interruption of either of the two branches of the ascending arousal system. Usually cortical dysfunction is metabolic whereas impairment of the ascending arousal system is due to localized damage. The most serious disturbance to consciousness is **coma**, in which an individual cannot be aroused and does not make any purposive responses to stimuli. Loss of large numbers of cortical neurons can result in **persistent vegetative state**. Afflicted individuals have sleep–wake cycles, may eat food placed in their mouth and show some emotional responses. However, they show no evidence of self-awareness and only minimal responses to their environment.

Functions of sleep

Remarkably, the function of sleep is not known. Among the numerous ideas that have been proposed two have emerged as frontrunners, the metabolic hypothesis and the memory hypothesis. Discussion of the possible role of sleep in memory is deferred until Section O.

The **metabolic hypothesis** postulates that NREM sleep provides a period of low metabolic demand needed to replenish neural energy resources depleted during waking. One possibility is that wakefulness decreases the ATP:AMP ratio, resulting in accumulation of extracellular adenosine which, acting through adenosine A1 receptors, inhibits neuronal activity and lowers brain metabolic demand. However, the need for NREM sleep seems to extend way beyond the nervous system. Sleep-deprived rats suffer anorexia (even though their food intake increases), lose the ability to thermoregulate so they become hypothermic, and die as a consequence of immune system failure after about 4 weeks. Hence, sleep may have global anabolic functions that conserve energy stores and core temperature. Some credence for this is that there is a correlation between the size of a mammal and the time it spends asleep. Smaller species (with the highest metabolic rates) sleep the most. In this context it is interesting that the preoptic area (POA) of the hypothalamus is involved in both thermoregulation and triggering NREM sleep, that local warming of the POA in behaving animals can trigger or prolong NREM sleep, and that heat loading during waking hours causes a rise in the amount of delta sleep taken during the next night.

N1 Early patterning of the nervous system

Key Notes

Neural tube formation

The early human embryo consists of two layers, ectoderm and endoderm. Ectodermal cells in the midline of the posterior end form the mesoderm and notochord, which lie between the other layers. Ectoderm above the notochord becomes the neural plate, which rolls up to become a closed neural tube by 28 days. Cells of the dorsal neural tube migrate away to become the neural crest which forms the peripheral nervous system. The neural tube becomes the central nervous system.

Neural induction

The signals for inducing ectoderm to become the neural plate arise from an organizer in the early mesoderm. In mammalian embryos this is Hensen's node. It secretes an inducer (chordin) which inhibits the action of bone morphogenic proteins that would otherwise drive the differentiation of ectoderm to epidermis. Early development organizes the neural tube along three axes, anteroposterior, dorsoventral and radial.

Formation of the anteroposterior axis

The hindbrain comes to be divided into a series of segments called rhombomeres. This is brought about by the activation of specific genes in narrow bands along the anteroposterior axis. These include the highly conserved *Hox* genes which code for homeodomain transcription factors. Their expression is controlled by a gradient of retinoic acid secreted from Hensen's node. The unique pattern of gene expression in each rhombomere precisely specifies the fates of its cells. The midbrain is induced by two signals (Wnt-1 and FGF-8) produced from cells at the hindbrain–midbrain junction. These inducers operate by switching on expression of two homeodomain transcription factors (En-1 and En-2). In a similar manner inductive signals divide the forebrain into six prosomeres each with their distinctive repertoire of homeobox transcription factor expression.

Formation of the dorsoventral axis

Sonic hedgehog (SHH) from the notochord induces the floor plate in the ventral neural tube. The floor plate then secretes its own SHH which induces the differentiation of ventral interneurons and motor neurons. SHH also induces oligodendrocytes, has a role in rostrocaudal patterning in the forebrain, and in cerebellar and eye development. Epidermal ectoderm induces a roof plate in the dorsal neural tube by secreting bone morphogenetic proteins (BMPs). The roof plate then generates its own BMPs which promote the differentiation of dorsal horn cells. SHH and BMPs act on cell surface receptors causing dorsoventral patterning by activating genes that encode transcription factors. These then modulate gene expression to produce differentiation.

Related topics

Organization of the peripheral nervous system (A4)
Organization of the central nervous system (A5)
Cell determination (N2)

Neural tube formation

At 11 days after fertilization, the human embryo is a two-layered circular disc. The **endoderm** is a single layer of flat cells forming the roof of the yolk sac. The **ectoderm** is 2–4 cells thick and continuous at its edges with the amniotic membrane. In the caudal midline, ectodermal cells proliferate to form a **primitive streak**, at the front end of which a knot of cells forms called **Hensen's node** (*Fig. 1*). The primitive streak gives rise to mesoderm which spreads laterally and forward between the endoderm and ectoderm, transforming the embryo into a three-layered structure, a process termed **gastrulation**. From Hensen's node, on day 5, a rod of cells grows forward in the midline between the ectoderm and endoderm, the **notocord**, the forerunner of the vertebral column.

Hensen's node and the notocord induce the formation, within the ectoderm above, of a **neural plate**. A groove develops along the midline of the neural plate which deepens, and the neural folds at the margins of the groove close over to form the **neural tube**. Closure of the neural tube starts in the middle and progresses in both rostral and caudal directions to be complete by the end of the fourth week (*Fig. 2*). Formation of the neural tube is called **neurulation**.

Cells on the dorsal side of the tube migrate laterally to form the **neural crest** which eventually gives rise to the peripheral nervous system. The neural tube becomes the central nervous system. Failure of the neural tube to close causes **anencephaly** (the fetus lacks much of the forebrain and cranium, dying shortly after birth) and **spina bifida**, in which the lumbosacral tube fails to close. In the most severe cases, **meningomyelocele**, elements of the spinal cord, cauda equina and meninges herniate through a defect in the vertebral column. More commonly, less serious varieties occur in which the defect is restricted to meninges or bone.

Neural induction

In all vertebrates an **organizer** region releases a chemical signal responsible for **neural induction**, the differentiation of ectoderm into neural tissue. In bird and mammalian embryos the organizer is Hensen's node and the most likely candidate for the signal molecule is the secreted protein **chordin**. Remarkably, the

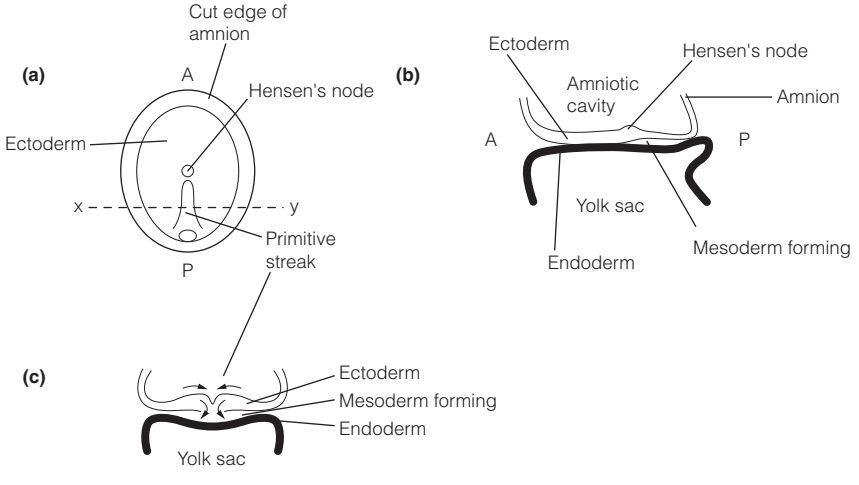


Fig. 1. Human embryo at 13 days gestation. A, Anterior; P, posterior. (a) Plan view; (b) longitudinal section; (c) transverse plane section at level x–y in (a), arrows depict migration of cells to form mesoderm.

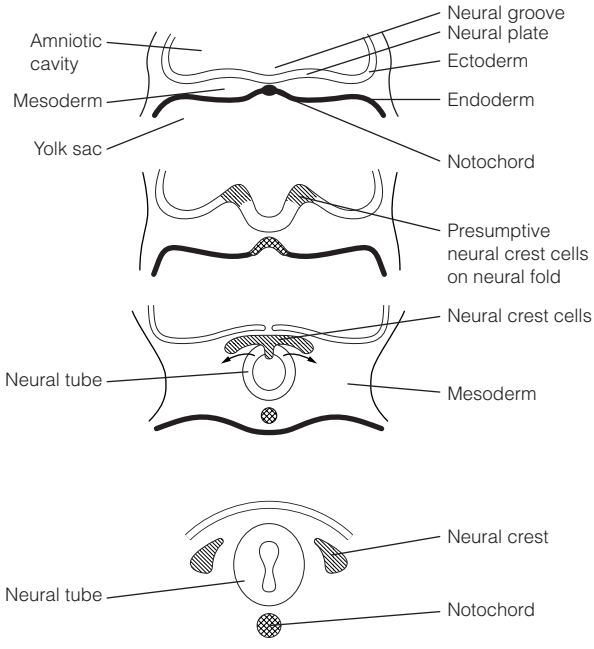


Fig. 2. Neurulation shown in midtransverse sections of the human embryo from day 20 to day 24. Closure of the rostral and caudal ends of the neural tube is not complete until day 28.

default state of ectoderm is to differentiate into neural tissue. This process is inhibited by a subfamily of **transforming growth factor** β (**TGF-**β) molecules called **bone morphogenetic proteins** (**BMPs**). These are produced endogenously in the ectoderm and normally act to transform it into epidermis. By binding BMPs, chordin inhibits their normal epidermalizing effect, with the result that ectoderm becomes neural tissue.

Subsequent development of the neural tube is organized in three directions:

- the anteroposterior (rostrocaudal) axis which is aligned along the long axis of the body;
- the dorsoventral (top to bottom) axis;
- the radial axis, which extends from the ventricular system out to the pial surface.

Formation of the anteroposterior axis

Chordin induces the expression of genes characteristic of forebrain. **Retinoic acid** produced by Hensen's node, together with a number of peptide growth factors which act as inducers, control gene expression in neural tissue at the posterior end of the neural tube so as to generate spinal cord and hindbrain. The effect is to activate several sets of genes, each of which is expressed in a restricted band along the neural tube. This divides the early embryo into a series of anteroposterior segments, a process termed **segmentation**.

Segments in the hindbrain are termed **rhombomeres** and the fates of cells in individual rhombomeres (and hence the distinctive character of each rhombomere) are determined by precisely which genes are expressed in its cells. As a result of a unique pattern of gene expression played out in space and time, the cells within a rhombomere come to express specific cell-surface proteins (cell

adhesion molecules), and secrete signaling molecules, which code their positions and specify how they interact with their neighbors. The interactions determine how the cells migrate, what connections they make, and even which of them survive. In this way, for example, rhombomere 3 gives rise to the trigeminal nerve while the facial nerve develops from rhombomere 5. The same general processes shape development throughout the nervous system.

One group of genes involved in hindbrain and spinal cord segmentation belongs to the ***Hox*** family. First identified in *Drosophila* by studying mutant flies with serious body plan errors (e.g. the *Antennapedia* mutant, which develops legs where it should have antennae), many of these genes turned out to have vertebrate homologs. *Hox* genes code for transcription factors which control the expression of other genes. All *Hox* genes contain a highly conserved region which encodes a sequence of 60 amino acids, the **homeodomain**, a part of which is responsible for binding to specific regions in the DNA of the genes they regulate. *Hox* genes are expressed in narrow bands across the AP axis and their effect is to help specify rhombomeres.

Extraordinarily, the sequence of *Hox* genes along the chromosome exactly matches the order in which they are expressed along the neuraxis. The closer the *Hox* gene is to the 3′ end of the coding strand of DNA the closer to the anterior end it is expressed. This arrangement allows *Hox* genes to be activated sequentially in a particular order by retinoic acid. Retinoic acid receptors (RXRs) are members of the intracellular steroid receptor superfamily. These are regulators of gene expression and promoters of *Hox* genes contain retinoic acid response elements which bind the RXRs.

Cells move from the primitive streak into the node where they proliferate and migrate forward as the **head process** (*Fig. 3*). The first cells to move through the node in this way get exposed to a brief, low-concentration pulse of retinoic acid which activates *Hox* genes at the 3′ end of the DNA. These genes determine the most anterior structures. Later, as retinoic acid production by the node increases, cells moving through the node are exposed to high concentrations of retinoic acid for longer. This activates progressively more posterior determining (i.e. towards the 5′ end) *Hox* genes. Inducers which direct different cell fates

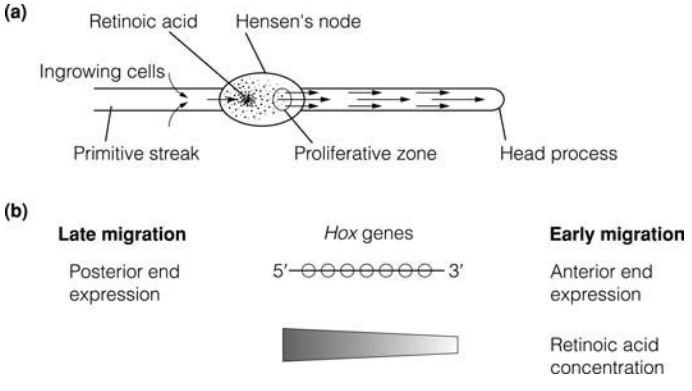


Fig. 3. Retinoic acid (RA) as a positional signal. (a) Cells are exposed to RA as they migrate through Hensen's node. (b) Early migration exposes cells to low RA concentrations which switches on anterior Hox *genes.*

depending on their concentration, like retinoic acid, are termed **morphogens**. Manipulating retinoic acid concentrations alters the expression of *Hox* genes; adding it to embryos *in utero* shifts the anterior limits of the *Hox* gene expression more rostrally so that hindbrain development expands at the expense of forebrain. Either deficiency or excess of retinoic acid causes severe craniofacial and brain malformations.

Whereas the *Hox* gene expression that causes compartmentation of the hindbrain occurs in the cells of the hindbrain itself, the rostrocaudal patterning of the spinal cord is brought about by *Hox* gene activation in the **paraxial mesoderm** that runs alongside the cord, which consequently secretes signals to affect cord development. *Hox* genes are not the only ones involved in segmentation of the hindbrain and spinal cord. The selective expression of *Hox* genes in rhombomeres is itself regulated by other transcription factors (e.g. Krox-20). In addition other proteins are differentially expressed along the hindbrain such as a number of Eph family tyrosine kinase receptors and their ligands (Ephrins).

The embryonic midbrain is not obviously segmented. Its development is under the control of a second organizer at the junction of the embryonic hindbrain and midbrain termed the **isthmus** (*Fig. 4*). The isthmus secretes two signals **Wnt-1** and **fibroblast growth factor-8** (**FGF-8**) which specify that the surrounding cells form the midbrain. The FGF-8 determines the rostrocaudal polarity of the midbrain by modulating the expression of two homeodomain proteins termed **En-1** and **En-2**. Loss of function mutations of *wnt-1* or *en-1/en-2* genes produces mice embryos that lack midbrain, pons and cerebellum.

The embryonic forebrain is divided into six **prosomeres**, with 1–3 corresponding to the caudal diencephalon and 4–6 to rostral diencephalon and telencephalon. As in the other regions, the rostrocaudal axis is produced by inductive signals establishing patterns of expression of homeodomain protein transcription factors, with prosomere boundaries corresponding to boundaries in expression of these molecules. However, prosomeres do not remain isolated developmental compartments. For example, many GABAergic neurons in the neocortex migrate from the striatum, even though early in development there is a clear boundary in gene expression between cortex and striatum.

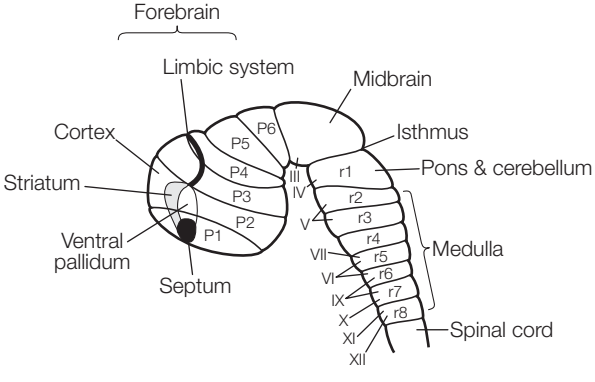


Fig. 4. Segmentation in 10-day mouse brain showing origin of cranial nerve motor outflows: r, Rhombomeres; P, prosomeres.

Formation of the dorsoventral axis

A top–bottom differentiation of the neural tube is particularly obvious in the hindbrain and spinal cord. The notocord which lies immediately beneath the infolding neural tube expresses a gene called *sonic hedgehog* (*shh*) which encodes a protein of the same name (SHH) that is a member of the transforming growth factor-β family of growth factors. SHH induces the formation of the **floor plate**, a narrow strip of glial cells along the ventral midline of the neural tube. The floor plate then becomes self-inducing by expressing *shh* itself. The SHH protein, first from the notocord and later from the floor plate, induces the differentiation of interneurons (at low concentrations) and motor neurons (at higher concentrations) in the ventral spinal cord (*Fig. 5*); i.e. SHH is a morphogen.

Although the floor plate secretes SHH along its whole length, in the hindbrain and midbrain regions of the neural tube it induces differentiation of serotinergic and dopaminergic neurons respectively. Hence the type of cell produced in response to the SHH signal depends on its AP position along the neuraxis. The cell type produced also depends on timing. After differentiation of motor neurons is complete a burst of SHH triggers the induction of oligodendrocyte progenitors in two narrow ventral columns either side of the floor plate and in the entopeduncular area of the forebrain. Moreover SHH is secreted from the zona limitans intrathalamica between prosomeres 2 and 3, where it determines the local rostrocaudal axis, and is also required for development of the cerebellum and the eyes. This illustrates economy in the organization of nervous system development; one molecule can play many parts depending on concentration, time, place and context.

Although the details of the signaling mechanism by SHH are not yet completely clear, the general theme is similar to that of inducers of the rostrocaudal axis. SHH binds to one subunit (patched) of a heterodimeric transmembrane receptor, freeing the other subunit (smoothed) to activate a number of zinc finger transcription factors. These activate transcription of some homeodomain proteins while suppressing the transcription of others. Mutations affecting components of the SHH signaling pathway produce severe neurological defects; e.g. **holoprosencephaly**, in which ventral brain structures fail to develop properly, or **cyclopia**, in which there is just a single eye in the middle of the forehead.

As the neural tube closes, **bone morphogenetic proteins** (**BMPs**) in the epidermal ectoderm induce the formation of a dorsal roof plate, which then produces its own BMPs. These locally generated BMPs in turn induce the

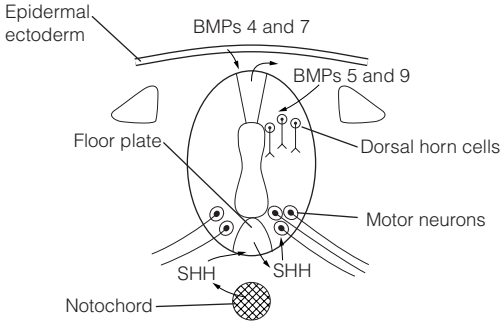


Fig. 5. Dorsoventral patterning of the neural tube. SHH, Sonic hedgehog; BMP, bone morphogenetic protein.

differentiation of the interneurons that will become dorsal horn cells. The receptors for bone morphogenic proteins are transmembrane serine-threonine kinases. Binding of BMP activates the intrinsic catalytic activity of the receptor which phosphorylates proteins termed **SMADs**. Unphosphorylated SMADs reside in the cytoplasm but on phosphorylation they enter the nucleus where they act as transcription factors to control the expression of proteins that regulate cell fate.

In the interface region between epidermal and neural ectoderm the neural crest forms. Once the neural tube has closed, cells from here migrate laterally. Those at the caudal end will become sensory neurons of the dorsal root ganglia, sympathetic neurons (including the chromaffin cells of the adrenal medulla), enteric neurons and Schwann cells. At the rostral end, neural crest cells form the sensory nuclei of the cranial nerves and parasympathetic neurons. The identity of neural crest cells appears to be predetermined by their expression of *Hox* genes, so when they migrate away from the neural tube they go to their appropriate destinations.

N2 Cell determination

Key Notes

Overview of cell determination

Cell differentiation is driven by an interaction between extracellular signals to which the cell is exposed and the pattern of gene expression within the cell. Extracellular signals can alter gene expression and some genes code for extracellular signals. Cells in particular lineages can be recognized by the marker antigens they carry.

Stem cells

Neural stem cells are pluripotent. They can undergo repeated division to produce neurons, astrocytes and oligodendrocytes and some of their daughter cells will remain as stem cells. Initially all neuroepithelial cells around the ventricles are stem cells and give rise mostly to projection neurons. Later, stem cells are concentrated in the subventricular germinal zone where interneurons are generated until the second postnatal year in humans. Stem cells migrate to the olfactory bulb and hippocampus where neurogenesis continues throughout adulthood. Adult stem cells are also found in the ependyma but usually give rise to oligodendrocytes.

Neuron fate determination

Epithelial cells lining the neural tube give rise to neuroblasts which proliferate and eventually give rise to postmitotic cells which differentiate into neurons. In *Drosophila*, neuroblast fate of epithelial cells requires expression of proneural genes, then neurogenic genes. Those cells which express neurogenic ones a little earlier become neuroblasts and their neurogenic gene products inhibit proneural gene expression in surrounding cells, which thus become epidermal. Vertebrate homologs of the *Drosophila* proneural and neurogenic genes probably have similar roles.

Neuron lineages

The position of a neuroblast in the neural tube determines its pattern of gene expression and the extracellular signals to which it is exposed. This specifies the type of neuron it will become. Motor neuron differentiation is initiated by sonic hedgehog (SHH) protein. Motor neurons then induce neighboring cells to become interneurons. Shortly after motor neurons are born the destination of their axons is unspecified by the position of the cell body in the neural tube, but within a few hours the motor neurons become committed so that their axons grow to particular muscles.

Glial cell lineages

In response to sonic hedgehog, stem cells produce glia progenitors capable of giving rise to either astrocytes or oligodendrocytes depending on the mix of growth factors it is exposed to. Astrocyte differentiation needs bone morphogenic protein-4 and ciliary neurotrophic factor (CNTF) whereas oligodendrocyte differentiation requires fibroblast growth factor-2 and thyroid hormone. Platelet-derived growth factor stimulates the oligodendrocyte progenitor to divide repeatedly so that large number of oligodendrocytes are made. In the presence of CNTF, oligodendrocyte progenitors will become astrocytes. Neural crest cells exposed to glial growth factor become Schwann cells.

Neurogenesis or gliogenesis

The CNS is divided into niches that control whether stem cells give rise to neurons or glia. In most of the adult brain, astrocyte activities support gliogenesis not neurogenesis. They do this by secreting a number of growth factors that stimulate glial differentiation. The olfactory bulb and the hippocampus continue to produce neurons into adult life because local ependymal cells produce a member of the TGF-β family of proteins, noggin, which blocks the stimulatory effect of some astrocyte growth factors on gliogenesis.

Related topics

Glial cells and myelination (A3)
Early patterning of the nervous system (N1)
Neurotrophic factors (N6)

Overview of cell determination

Neuroectoderm gives rise to a large number of cell types, both neurons and glia, each with their own identity or **phenotype**. In development, the future phenotype of an undifferentiated cell is its **fate**. The process by which cell precursors are transformed into mature cells is **differentiation** and the sequence of cell types that lead from precursor to mature cell is the **cell lineage**. Differentiation occurs through an interplay between:

- extrinsic signaling molecules (diffusible, or bound to the cell surface or extracellular matrix) in the cells' surroundings that have been generated by other cells;
- a timed sequence of gene expression mediated by intracellular signals, usually transcription factors. Some of these intrinsic signals are inherited from the cell lineage, others are generated in response to second messenger cascades activated by the extrinsic signals.

As cells differentiate they express different antigens. These can be assayed and used as markers for particular cell lineages. For example, astrocytes express **glial fibrillary acidic protein** (**GFAP**) and this can also be used to identify cells which are destined to become astrocytes (astrocyte progenitors) even when they have yet to acquire the characteristic astrocyte morphology.

Stem cells

Early in development all the cells of the neuroepithelium are **stem cells**, that is they have the ability to become any cell type within the nervous system; neuron, astrocyte or oligodendrocyte. Neural stem cells:

- are multipotential;
- have an unlimited capacity for cell division;
- are self-renewing, that is after each mitotic division at least one of their daughters is a stem cell.

There is considerable interest in neural stem cells because of their potential for treating diseases in which specific cell types are lost (e.g. Parkinson's disease) or in neural repair, such as spinal cord injuries. Unfortunately, while stem cells isolated from adult spinal cord can be made to differentiate into neurons *in vitro* they do not participate in neural regeneration *in vivo*. With better understanding of how they respond to their neural environment it may be possible to encourage them to do so.

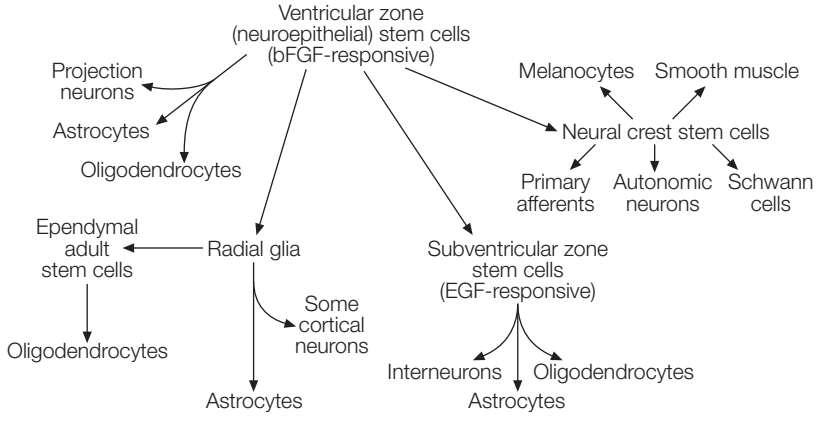


Fig. 1. Neural stem cells and their progeny.

There are several populations of stem cells (*Fig. 1*). Neuroepithelial stem cells are located in the **ventricular zone** (**VZ**) which surrounds the ventricles. These cells require **basic fibroblast growth factor** (**bFGF**) for their survival and proliferation, and the neurons which they spawn are projection neurons. Neuroepithelial stem cells give rise to:

1. Stem cells in the **subventricular zone** (**SVZ**) that encircles the VZ. These are responsive to **epidermal growth factor** (**EGF**) and are so extraordinarily prolific that the SVZ is described as a **germinal zone**. The nerve cells generated by the SVZ stem cells are interneurons and in humans neurogenesis by the SVZ continues into the second postnatal year. SVZ stem cells migrate to the olfactory bulb and into the hippocampus. Currently these are thought to be the only brain regions where neurogenesis from adult stem cells continues throughout adult life.
2. Neural crest stem cells. These remarkable cells give rise not only to primary afferents, sympathetic, parasympathetic and enteric neurons, but also to Schwann cells, smooth muscle and melanocytes.
3. Radial glia. Traditionally regarded as a type of astrocyte that guides the migration of cortical neurons, **radial glia** also appear to give rise to conventional astrocytes, cortical neurons and possibly adult stem cells in the ependyma. All that remains of the VZ in the adult CNS is a single layer of ciliated cells, the **ependyma**, on which sit subendymal cells. These are adult stem cells even though they look like astrocytes histologically, resemble radial glia antigenically and express GFAP. Although GFAP has been regarded as a marker for astrocytes, recent work shows that most CNS neurons are derived from progenitors that once expressed GFAP. Adult neural stem cells in the ependyma almost invariably differentiate into oligodendrocytes if they migrate into the surrounding white matter.

Neuron fate determination

Initially in vertebrate development, most of the cells generated by neural stem cells lining the neural tube become **neuroblasts**, which divide mitotically. After a time, which varies depending on the fate of the cell, neuroblasts produce daughter cells that are no longer able to undergo mitosis. These migrate away

from the epithelium towards the surface of the neural tube, where they continue to differentiate into neurons.

The first steps in making neuroblasts from undifferentiated ectoderm has been worked out in *Drosophila*. Initially **proneural genes** are switched on in a cluster of cells termed the proneural cluster. Proneural genes code for basic helix-loop-helix transcription factors (see *Instant Notes in Molecular Biology*) that bring about a pattern of gene expression that makes the cells competent to become neuroblasts. However, proneural genes also trigger expression of **neurogenic genes**. These code for cell surface signaling proteins that mediate interactions between neighboring cells, the outcome of which is that only a few become neuroblasts and these inhibit the remainder from doing so, forcing them to become epidermoblasts. Two neurogenic genes that have been extensively studied are *notch* (*N*) and *delta* (*Dl*) which code for a receptor, **Notch**, and its ligand, **Delta**. Delta transcription is controlled by the **Achaete-scute** family of proneural transcription factors. This means that any proneural cluster cell that just happens to be slightly ahead of its neighbors on the path to being a neuroblast is likely to express a few more Delta molecules on its surface. This is enough to trigger a positive feedback process that will lead to it becoming a neuroblast whilst inhibiting its neighbors from doing the same. It works in the following way (*Fig. 2*): on binding Delta, the intracellular domain of Notch is proteolytically cleaved from the rest of the receptor and moves into the nucleus where it activates the transcription factor **Suppressor of hairless**. This initiates transcription of a *repressor* of proneural genes termed **Enhancer of split**, that will shut down transcription of *achaete-scute* genes, so reducing Delta expression. Hence, those cells producing the highest amount of Delta will cause the greatest suppression of their neighbors' proneural genes; they will win the 'race' to become neuroblasts.

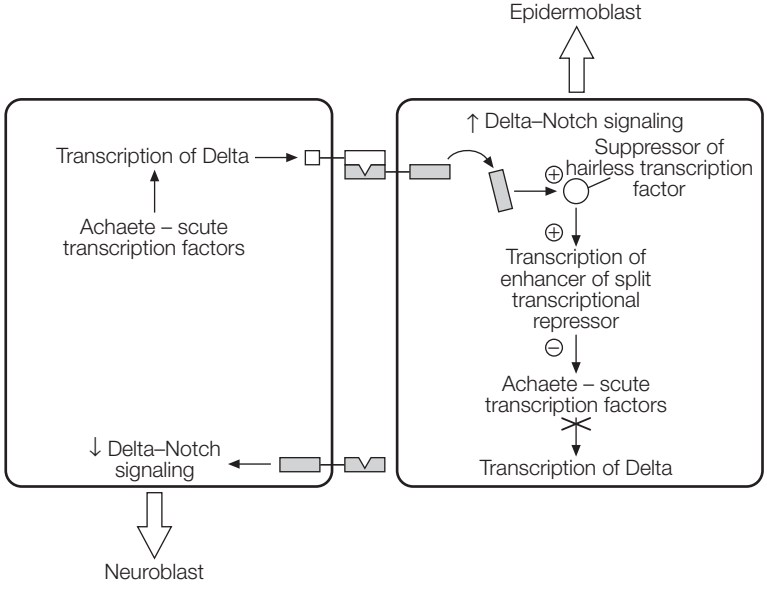


Fig. 2. Delta–Notch signaling in neural determination. Cells producing the highest amount of Delta inhibit Delta transcription in neighboring cells by Notch signaling. Cells experiencing high levels of Delta–Notch signaling become epidermoblasts, those with low levels of Delta–Notch signaling become neuroblasts.

Many of the genes extensively studied in *Drosophila,* including the *achaete-scute* complex, and *notch* and *delta,* have homologs in vertebrates, where they seem to have similar functions.

Neuron lineages

A single cell in each proneural cluster eventually becomes a fully committed neuroblast. Its subsequent fate is now determined by the activation of additional subsets of genes, but which subset depends on the history of the neuroblast and its surroundings (i.e. its position). Many of these genes code for transcription factors that contain homeodomains. The particular combination expressed in a neuroblast specifies what sort of neuron it will become. Although many of these genes were first identified in invertebrates, homologous genes occur in vertebrates which appear to have similar roles. At present, the mechanisms by which a neuroblast turns into a particular type of neuron are understood (rather poorly) in only a few cases, for example, the motor neuron.

When sonic hedgehog (SHH) from the floor plate induces motor neurons in the ventral neural tube it does so by directly activating the expression of two transcription factors and repressing two others. When this was investigated in the zebrafish it was found that the *Drosophila* hedgehog (*hh*) gene can substitute perfectly for zebrafish *shh* gene in motor neuron induction. This illustrates how close the homologies can be between invertebrate and vertebrate development.

In zebrafish, each spinal segment initially has just three primary motor neurons so their fates are easily followed (*Fig. 3*). The rostral primary (rp) motor neuron axon grows to the lateral muscle, the middle primary (mp) motor neuron innervates the dorsal muscle, while the caudal primary (cp) projects to the ventral muscle. When these motor neurons are transplanted to different locations in the spinal segment at the time when their axons are beginning to grow, they project to their original muscles. Hence, at this time the neurons are

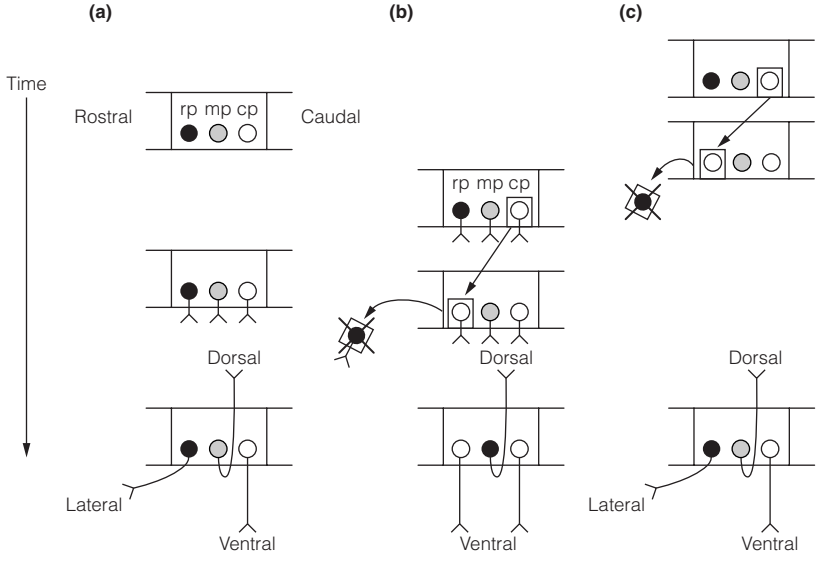


Fig. 3. Motor neuron differentiation: (a) normal development; (b) a caudal primary neuron (cp) is transplanted into the position normally occupied by the rostral primary neuron (rp) at the time of axonogenesis; (c) a cp is transplanted to the rp position before axonogenesis. mp, Middle primary neuron.

committed to their particular fates and continue to develop independently. If the neurons are transplanted earlier, shortly after they are born, they take on the characteristics expected of the cell at their *new* position. Thus a cp transplanted to the rostral end of a spinal segment comes to innervate lateral (not ventral) muscle. This suggests that early development is dictated by the position of the cell along the spinal segment.

Glial cell lineages

Early in development most stem cell progeny become neurons but as development proceeds this gives way to the generation of glia. In response to sonic hedgehog (SHH) stem cells produce glia progenitors capable of giving rise to either astrocytes or oligodendrocytes depending on the mix of growth factors it is exposed to (*Fig. 4*). BMP-4 and ciliary neurotrophic factor (CNTF) stimulates differentiation to astrocytes whereas fibroblast growth factor-2 (FGF-2) and thyroid hormone drive the progenitor towards an oligodendrocyte fate. In the absence of platelet-derived growth factor (PDGF) the progenitor immediately differentiates into an oligodendrocyte. However, in the presence of PDGF the progenitor is stimulated to divide repeatedly so that large numbers of oligodendrocytes are made. However, even at this stage it is not too late, in the presence of CNTF, for these cells to become astrocytes. Indeed if oligodendrocyte progenitor cells are isolated in the absence of astrocytes they will give rise to neurons in the appropriate environment.

In the peripheral nervous system, once neural crest cells have migrated to form early ganglia, neuroblasts express **glial growth factor-2** (**GGF-2**) which promotes the emergence of a Schwann cell phenotype by suppressing neuronal differentiation.

Microglia are derived from mesodermal stem cells which invade the developing nervous system.

Neurogenesis or gliogenesis

Whether a multipotential cell gives rise to neurons or glia depends on local signals it is exposed to. For example, stem cells destined for the olfactory bulb, where they become neurons, do not differentiate into neurons when diverted into the cortex. Hence the CNS is divided into niches which control cell fates, not the other way round. Intriguingly astrocytes play a major role in this.

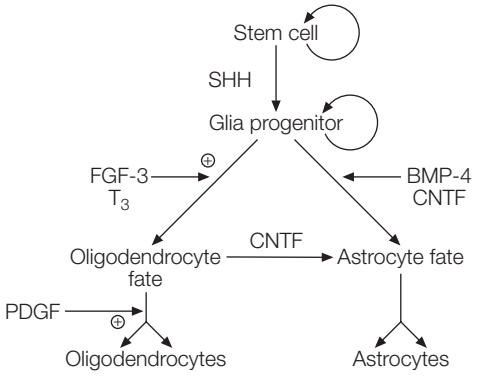


Fig. 4. Gliogenesis is driven by astrocytes. FGF-3, fibroblast growth factor-3; BMP-4, bone morphogenic protein-4; CNTF, ciliary neurotrophic factor; PDGF, platelet-derived growth factor are all secreted by astrocytes. SHH, sonic hedgehog; T_3, tri-iodothyronine

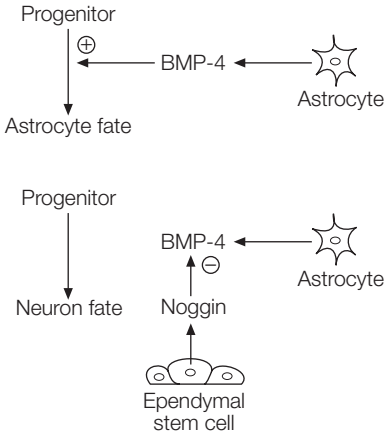


Fig. 5. Control of neurogenesis by ependymal cell-derived noggin block on BMP-4, bone morphogenetic protein-4 action.

Firstly, some astrocytes appear to be stem cells in their own right. Secondly, astrocytes have an instructive role in determining whether a local niche is neurogenic or gliogenic by virtue of their ability to secrete a large variety of growth factors and cytokines. For example, astrocytes release BMPs, CNTF, FGF-2, and PDGF, all of which are involved in glial differentiation (see *Fig. 4*).

In most of the adult brain, astrocyte activities support gliogenesis not neurogenesis. However, two niches, the olfactory bulb and the hippocampus, remain neurogenic into adult life. One explanation for this (*Fig. 5*) is that ependymal cells secrete **noggin**, a member of the TGF-β family of proteins similar to chordin. Noggin binds to BMPs secreted by astrocytes, blocking their effect so as to redirect the fates of local adult stem cells to spawn nerve cells rather than glia.

N3 Cortical development

Key Notes

Cerebral cortex development

The ventricular zone (VZ) of the neural tube contains stem cells which continually divide, giving rise to neurons and glial cells. Neurons migrate away from the VZ in such a way that the neural tube develops six radially organized layers. The two most superficial layers eventually become the neocortex, the VZ becomes the ependyma and the overlying subventricular zone retains the capacity to generate glial cells into adult life. Neuron migration occurs along radial glial cells, the processes of which extend the full thickness of the neural tube. Large projection neurons are born first and migrate the shortest distance. Smaller cells, born later, migrate through zones containing older cells to more superficial positions.

Cerebellar cortex development

The ventricular zone (VZ) gives rise to an internal granular layer which generates the neurons of the deep cerebellar nuclei and Purkinje cells which migrate superficially. VZ cells of the anterior rhombic lip migrate to form an external granular layer which generates huge numbers of granule cells that migrate down radial glial cells to lie deep to the Purkinje cells.

Related topics

Organization of the central nervous system (A5)

Cerebellar cortical circuitry (K5)

Cerebral cortex development

Early on, the walls of the three vesicles (*Fig. 1*) at the anterior end of the neural tube are a single layer of pseudostratified columnar neuroepithelium. These cells give rise to neurons and glia and initially the neuroblasts divide more frequently than the glia. In humans, up to about 6 weeks gestation cell division in the neuroepithelium causes it to increase in area but not in thickness. After 6 weeks, the walls of the neural tube go from being a single layer to several layers. This represents **radial patterning** of the neural tube. At first the neural tube becomes differentiated into an inner **ventricular zone** (**VZ**) in which cells proliferate and a **marginal zone** (**MZ**) containing radially projected cell processes.

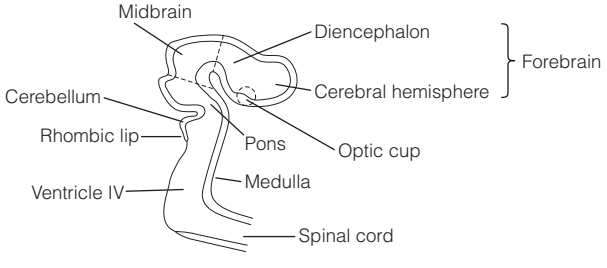


Fig. 1. Neural tube of a human embryo at about 5 weeks gestation; midsagittal section.

Cells that are not actively dividing (i.e., those in the interphase of the cell cycle) are bipolar and extend their processes through the full thickness of the neuroepithelium. Progenitor cells undergoing mitosis retract their processes and sit on the wall of the ventricle. These cells divide symmetrically so that the parent produces one **stem cell** (which subsequently continues to divide) and a post-mitotic cell which migrates away from the ventricle (*Fig. 2*). The periodic changes in shape and positions of progenitor cells as they go through successive cell cycles are called **interkinetic movements**.

Migrating neurons leave the VZ to form a single layer called the **preplate** (*Fig. 3*). Between the preplate and VZ is an **intermediate zone** (**IZ**) in which is found axons of the preplate neurons and axons growing into the cortex from developing subcortical structures, such as the thalamus. As neurons continue to be generated in the VZ they migrate through the IZ to form the **cortical plate** (**CP**) which splits the preplate into the marginal zone above and a **subplate** (**SP**) region that lies below. The subplate is the site of intense **synaptogenesis** (synapse formation). As neurons migrate through the subplate they await the arrival of afferents from other areas of cortex and subcortical structures to make contact, then continue into the cortical plate. Neurons that remain in the subplate suffer apoptosis, so this layer is transient.

During the later phases of neurogenesis a second region of dividing cells develops, the **subventricular zone**. This generates mostly small interneurons, a process which in humans continues into the second postnatal year. The mature cerebral cortex is formed mostly from the cortical plate, with the marginal zone being layer 1. The intermediate zone becomes the subcortical white matter. The

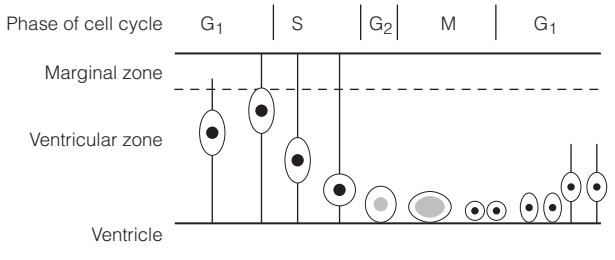


Fig. 2. Interkinetic movements of proliferating cells in the neural tube. G_1, Resting; S, DNA synthesis; G_2, preparation for mitosis; M, mitosis. (For further details of the cell cycle see Instant Notes in Biochemistry, *2nd edn.)*

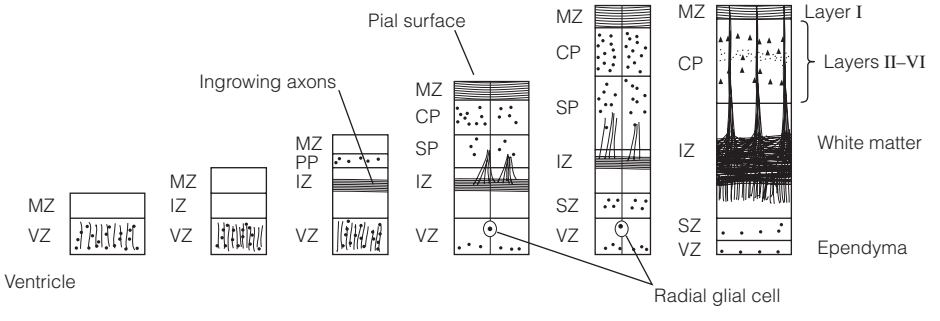


Fig. 3. Development of the neocortex. CP, Cortical plate; IZ, intermediate zone; MZ, marginal zone; SP, subplate; SZ, subventricular zone; VZ, ventricular zone.

original ventricular zone becomes the ependymal layer, above which a thin subventricular zone contains cells which continue to proliferate to provide glia. Until recently it was thought that new neurons are not produced in the mature CNS. Recent evidence, however, shows that the ependyma of the adult brain retains stem cells capable of dividing to generate neurons.

Neurons migrate out of the VZ along **radial glial cells** by amoeboid movements. Radial glial cells differentiate early and have their cell bodies in the VZ. They are bipolar and one end of their long process is attached to the ventricular cell while the other end is fixed to the pial cell basement membranes. This arrangement provides a scaffold for neuron migration. Neurons moving along a single glial cell generally remain close together after migration and this forms the columnar structure of the cortex. When migration is complete, most radial glia differentiate into astrocytes. In two brain regions they remain virtually unaltered; as Muller cells in the retina and Bergmann glia in the cerebellum.

The birthday of a neuron is the day on which it loses the ability to undergo further mitotic division. Postmitotic neurons subsequently migrate. The birthday of a neuron and its migration can be tracked by autoradiography of embryos at different times after giving a brief pulse of [^{3}H] thymidine. The isotope labels all cells synthesizing DNA (i.e. in the S phase of the cell cycle) at the time of the pulse because these cells will incorporate the label into their growing DNA molecules. This type of experiment reveals general features of neuron development. Firstly, larger projection neurons are born earlier than small neurons or interneurons. Secondly, younger neurons migrate through zones occupied by older neurons. Hence neurons with late birthdays migrate furthest and come to occupy a more superficial position. In consequence, the mature cerebral cortex has small cells in superficial layers and larger ones in deeper layers. In addition, the birthday of a neuron determines its subsequent fate; i.e. what type of cell is formed and in which layer it ends up. Young progenitor cells transplanted into host embryos of a different age switch their fates to be appropriate to the developmental age of the host. However, cells about to divide for the last time (i.e. just about to give rise to a specific neuron type) retain their donor fate. This shows that extrinsic cues, operating at about the time when a neuron is born, determine its fate. The extrinsic cues are thought to come from the cell type born previously. For example, layer VI cortical pyramidal cells are born first and as they differentiate they produce signals which switch the VZ neuroblasts into generating layer V precursors.

Cerebellar cortex development

The primordial cerebellum consists of pseudostratified neuroepithelium showing interkinetic movements which soon develop into ventricular, intermediate and marginal zones, as elsewhere in the neural tube. As the cerebellum thickens, the VZ gives rise to uncommitted progenitors, which form the **internal granular layer (IGL)** and **external granular layer (EGL)**. Cells which form the EGL migrate from the VZ in the anterior rhombic lip (the rostrodorsal edge of ventricle IV) (*Fig. 4*).

The IGL gives rise to nuclear neuroblasts and Purkinje neuroblasts. The nuclear neuroblasts form the neurons of the deep cerebellar nuclei and remain embedded in the IZ, the future white matter of the cerebellum. The Purkinje neuroblasts migrate superficially towards their final position, trailing one neurite (the future axon), which lengthens as the cell body migrates. Finally, the IGL produces Golgi neuroblasts which migrate a short distance. The EGL initially gives rise to basket neuroblasts and subsequently its cells undergo an

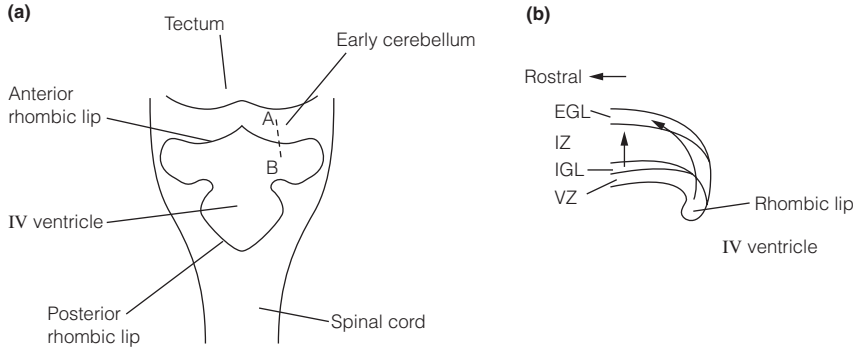


Fig. 4. Early cerebellar development: (a) Dorsal view of early hindbrain. The rhombic lip surrounds an opening in the neural tube which will eventually be covered by the cerebellum to form the IV ventricle. The posterior rhombic lip gives rise to the choroid plexus. (b) Section A–B through the rhombic lip and early cerebellum. The anterior rhombic lip generates cells which migrate to the external granule layer (EGL) to become granule neuroblasts. The internal granule layer (IGL) produces Purkinje neuroblasts (straight arrow). IZ, Intermediate zone.

intense and prolonged proliferation generating vast numbers of granule cell neuroblasts and, at about the same time, stellate neuroblasts.

Radial glial cells (**Bergmann glia**) provide a scaffold for the migration of granule neuroblasts from the EGL through the Purkinje cell layer to the IGL. As the granule cells migrate they trail their axons behind. These axons bifurcate, giving rise to the parallel fibers (*Fig. 5*).

The importance of glial cells is illustrated by *weaver* mutant mice. Animals homozygous for the *weaver* mutation suffer disastrous motor incoordination. Bergmann cells are misaligned, and granule cell migration stops at the molecular layer, disrupting further cerebellar cortex development. Impaired signaling between glial and granule cells causes the defect.

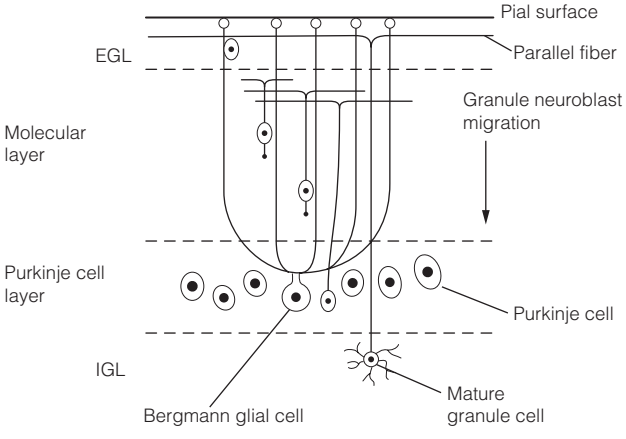


Fig. 5. Migration of granule neuroblasts through cerebellar cortex.

N4 Axon pathfinding

Key Notes

Formation of topographic maps

Studies of the manner in which retinal ganglion cells form precisely ordered connections with the tectum, in frogs and chicks, suggest that relative position on the tectum is specified by gradients of signaling molecules across anteroposterior and dorsoventral axes of the tectum, that are detected by incoming axons.

Growth cones

The tips of growing axons form growth cones. These express cell surface receptors which recognize short- or long-range signals that guide their axon, either by attraction or repulsion. Growth cones crawl with amoeboid motion over the surface. Binding of growth cone receptors to molecules in the surface on which it rests causes polymerization of actin at the leading edge of the cone, which consequently advances. Polymerization of microtubules in the body of the growth cone is needed for sustained progress, partly by facilitating fast axoplasmic transport.

Cell adhesion molecules

Cell adhesion molecules are integral membrane proteins, or extracellular matrix proteins (e.g. laminin, fibronectin, collagens) that mediate interaction between cells, or between a cell and the extracellular matrix. The interactions can be either attractive or repulsive. There are three groups of cell surface cell adhesion molecules; integrins, the immunoglobulin superfamily and cadherins. Integrins are the growth cone receptors for extracellular matrix molecules. The immunoglobulin superfamily and cadherins mediate cell adhesiveness either by binding to identical molecules or different superfamily members in adjacent cells. Cell adhesion molecules guide axons and specify which cells they can make contacts with.

Axon guidance molecules

Netrins are responsible for long-range axon guidance. Netrins are secreted, but then bind to the extracellular matrix where they interact with netrin receptors (Unc-5H, neurogenin) in the growth cone. Netrins attract axons of dorsal neural tube cells to grow ventrally into the floor plate, but can also be repulsive to other neurons, e.g. trochlear nerve axons. Semaphorins, interacting with their cognate receptors (neuropilins and plexins) cause growth cones to collapse and for axons to be redirected. Ephrins and their corresponding Eph tyrosine kinase receptors control axon guidance by inhibitory interactions. They, together with semaphorins, are responsible for specifying how the retinotectal pathway wires up in the chick.

Related topics

Organization of the central nervous system (A5)
Eye and visual pathways (G2)
Neurotrophic factors (N6)

Formation of topographic maps

During development axons need to grow along particular paths in the nervous system so that they make connections with the appropriate target cells. This process must be particularly well regulated for those axons which establish topographic connections; neighboring axons must synapse with neighboring target cells in a very precise way in order to produce a smooth map. The process by which neural connections form topographic maps has been extensively studied in the **retinotectal pathway** of the frog and chick. This pathway maps visual input from the retina to the optic tectum of the midbrain. If the optic nerve of a frog is cut and the eye rotated through 180°, the optic nerve regenerates (*Fig. 1a*). However, the frog now behaves as if its visual world has been rotated (*Fig. 1b*), and since it never learns to compensate for this deficit, the retinotectal pathway must be specified by developmental mechanisms to be immutable. If, in a normal frog, half of the retina is removed, axons from the remaining half eventually expand to cover the whole tectum (*Fig. 1c*). If half of the tectum is removed (*Fig. 1d*) retinal axons re-sort so that eventually a complete retinal map exists on the remaining hemitectum. These experiments show that it is *relative*, not absolute, position that is encoded, and suggested that gradients of signaling molecules across the anteroposterior and dorsoventral axes of the tectum, detected by receptors in the arriving growth cones account for the positional information. Since these pioneering experiments were done some of the signaling molecules that guide axons have been identified.

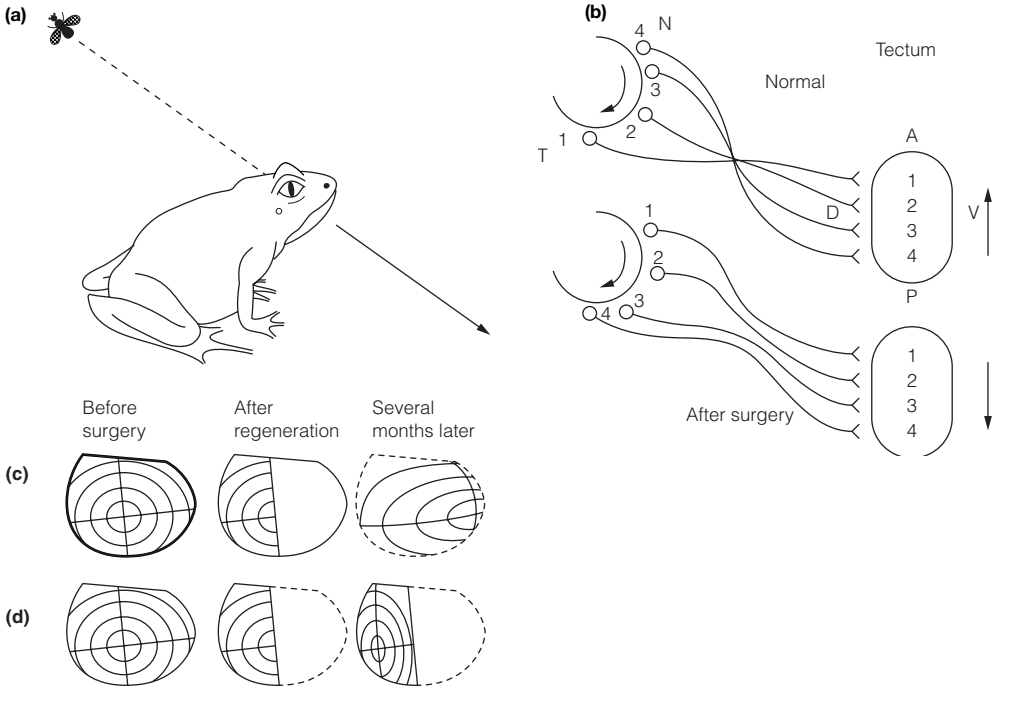


Fig. 1. Retinotectal specificity. (a) Behavior of a frog after cutting the optic nerve and 180° rotation of the eye. (b) rewiring of the tectum maintains the original position information. A retinal image, ↑, now has a tectal representation that is misaligned by 180° (i.e. it is upside down). N, nasal retina; T, temporal retina; A, anterior; P, posterior; D, dorsal; V, ventral. (c) Expansion of the topographic map after removal of half the retina. (d) Restoration of a complete map after removal of half of the tectum. In (c) and (d) the contours are isospatial lines at intervals 20° from the center of the retina.

Growth cones

Axons emerge from neuroblasts, elongate and grow towards their proper targets. The tip of the growing axon is termed the **growth cone**. It expresses a sequence of cell surface receptors which recognize either short-range (local) cues, such as specific marker molecules on the extracellular matrix or on the surface of particular cell types, or long-range (secreted) cues, which can be either attractive or repulsive on axon growth. Interactions between the growth cone and the markers guide the axon to its correct destination.

The growth cone (*Fig. 2*) consists of a flattened, fan-shaped cytoplasmic extension called a **lamellipodium**, rich in mitochondria and stacks of membrane-bound vesicles, the membranes of which are incorporated into the surface of the advancing growth cone. The plus ends of microtubules extend from the axon into the lamellipodium. Tubulin polymerization occurs at the plus end to lengthen the microtubule. At the leading edge of the lamellipodium is a dense meshwork of **filamentous actin (F-actin)**. Projecting from the lamellipodium are numerous spiky **filipodia**. These contain bundles of F-actin.

Growth cones advance by crawling with an amoeboid motion over the surface (**substratum**) until receptors in the growth cone membrane find and bind to molecules on the substratum. This interaction causes actin, which is coupled to the receptor, to polymerize so the leading edge of the cone extends. F-actin further back depolymerizes, shrinking proximal regions of the cone. Actin monomer released by the depolymerization is recycled to the leading edge. The advance of the tip of the growth cone occurs in the following way: polymerized stretches of actin at the leading edge are driven in a retrograde direction by interaction with myosin in a process reminiscent of muscle contraction. This generates tension, and results in filipodia poorly attached to the substratum to retract. Strong attachment of a filopodium to the substratum opposes this tension. Hence the growth cone progresses in the direction in which it can fix itself most avidly to its surroundings. Often, growth cones are guided into the appropriate place by molecules secreted either locally or at a distance, which are either attractive or repulsive to the cone. Polymerization of tubulin into microtubules is required for the sustained advance of growth cones. In part, this is because of the role microtubules play in fast axoplasmic transport which brings proteins synthesized in the cell body to the cone.

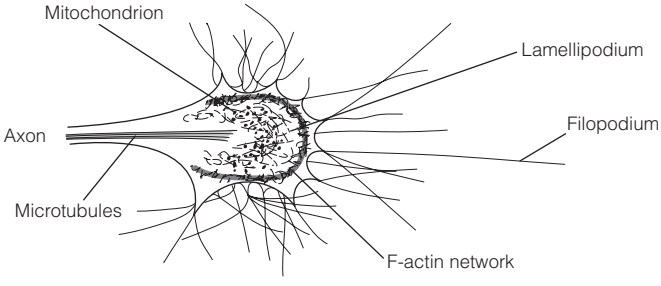


Fig. 2. The structure of a growth cone.

Cell adhesion molecules

Several large families of molecules, collectively referred to as **cell adhesion molecules**, mediate both cell adhesion and axon guidance (*Table 1*). Those that mediate short-range (local) interactions may be either integral membrane proteins or in the extracellular matrix. When the interaction is attractive the

term **contact guidance** is used, when repulsive, **contact inhibition**. Extracellular matrix molecules include **laminin**, **fibronectin**, and **collagens**. Laminin and fibronectin stimulate axon outgrowth and line corridors which channel axons to grow in the appropriate direction. Cell-surface molecules important in the nervous system fall into three broad classes, the integrins, the immunoglobulin superfamily and the cadherins.

Table 1. Classes of molecules involved in axon guidance

Ligand[a]	Receptor[b]	Effects
Laminins	Integrins	Interaction of growth cones with substrate Axon growth and guidance
IgG[c] superfamily cell adhesion molecules	IgG superfamily cell adhesion molecules (heterophilic and homophilic binding)	Cell-cell adhesion Axon growth and guidance
Cadherins	Cadherins (homophilic binding)	Cell-cell adhesion Axon growth and guidance
Netrins	Netrin receptors (e.g. Unc-5H, neurogenin)[c]	Attract/repel axons
Ephrins	Eph kinases (receptor tyrosine kinases)[c]	Inhibitory. Retinotectal specificity in chick
Semophorins	Neuropilins, plexins[c]	Inhibitory, growth cone collapse, axon re-directed

a All transmembrane proteins except netrins, which are secreted, and laminins which are extracellular matrix glycoproteins.
b All transmembrane proteins
c Immunoglobulin superfamily molecules

The growth cone receptors for extracellular matrix molecules belong to the **integrin** family. Integrins are integral membrane proteins made up of heterodimers consisting of an α and a β chain. There are several isoforms of each chain and different combinations of isoforms generate a large number of integrins, each with preferential binding affinities for particular extracellular matrix molecules (e.g. $\alpha 4\beta 1$ integrin binds to fibronectin). This gives the potential to encode a large number of axon guidance pathways.

The **immunoglobulin** superfamily is huge and includes not just the standard cell adhesion molecules, such as N-CAMs, LI and axonin-1, but also the receptors for cell surface or secreted axon guidance ligands, e.g. Eph kinase receptors. They have a single transmembrane segment and a small intracellular terminal. The extracellular region is large and consists of a variable number of immunoglobulin domains and fibronectin type III domains. Immunoglobulin–superfamily cell adhesion molecules are both receptor and ligand. By binding other copies of themselves (**homophilic binding**) or binding other family members (**heterophilic binding**) on adjacent cells, they are able to mediate cell adhesion. The adhesiveness of N-CAMs is regulated by long chains of sialic acid, a negatively charged sugar. The more sialic acid residues a N-CAM possesses the weaker the adhesiveness. Embryonic N-CAMS are highly sialylated so that cell–cell contacts are readily broken. This allows extensive sorting to occur before irrevocable contacts are forged by differentiated cells, which have N-CAMs that are much less sialylated, and are hence more adhesive.

An example of immunoglobulin–superfamily cell adhesion molecule axon guidance is that of commissural neurons in the dorsal neural tube. These cells extend their axons ventrally, whereupon they cross the floor plate to the opposite side. The commissural cells express axonin-1, while the floor plate cells have Nr-CAM on their surface. These molecules bind one another (heterophilic binding). Antibodies to either axonin-1 or Nr-CAM injected into the neural tube result in about half of commissural axons failing to cross the floor plate.

One class of cell adhesion molecule called **fasciculins** is responsible for the cell adhesion often seen between axons from several neurons (either of the same or different type) which causes them to aggregate into bundles or **fasciculi**. In this situation axons are acting as guides for each other. Axons with the incorrect molecules are excluded.

The cadherins are a group of over 100 cell adhesion molecules that engage in homophilic binding. Unlike the immunoglobulin superfamily, binding is Ca^{2+}-dependent. On binding, the intracellular domains of cadherins interact with the neuronal cytoskeleton, via a group of proteins called catenins, to promote axon outgrowth.

Axon guidance molecules

Long-range interactions are brought about by several families of molecules (see *Table 1*). **Netrins** are large, secreted proteins with N-terminal sequences that resemble laminins. Although soluble netrins bind to the extracellular matrix where they interact with receptors on the surface of the growth cone and can have either attractive or repulsive effects by encouraging axon growth either up or down netrin concentration gradients. Netrin receptors in vertebrates (e.g. Unc-5H, neurogenin), members of the immunoglobulin superfamily, are homologous to similar axon guidance molecules first identified as products of *unc* genes in the nematode *Caenorhabditis elegans*. This illustrates how highly conserved developmental processes are.

The ventral growth of the commissural axons in the dorsal neural tube (see above) is mediated by two netrins. Netrin 1 is expressed in the floor plate, while netrin 2 is expressed more weakly in ventral and lateral regions. This has been demonstrated in mice genetically engineered to be without functional genes for netrins. The generic term **knockout** is used for animals engineered to lack the function of specific genes. In mice lacking netrin genes (netrin knockouts), commissural axons fail to grow towards the floor plate. Thus netrins are chemoattractant for commissural neurons. Remarkably, as the axons grow through the floor plate they become desensitized to netrin, allowing the axons to continue to extend.

Netrins can be chemorepellent for some neurons. Motor neurons of the trochlear (cranial IV) nerve are located in the ventral neural tube. Their axons grow dorsally, away from the floor plate. Explants of neural tube containing trochlear motor neurons, when cultured with cells expressing netrins, show repulsion of the outgrowing axons. It has been demonstrated that a netrin can be both attractant and repellent by acting on different netrin receptors.

Two other major classes of receptor, together with their cognate receptors, have been shown to be crucial for axon guidance. Both are involved in sculpting the retinotectal map in the chick, in which nasal retina projects to the posterior tectum and the temporal retina projects to the anterior tectum (*Fig. 3*). Nasal growth cones are unaffected by incubation with anterior or posterior tectal membranes. Temporal growth cones are not influenced by anterior tectal membranes, but collapse and fail to advance when exposed to posterior tectal membranes.

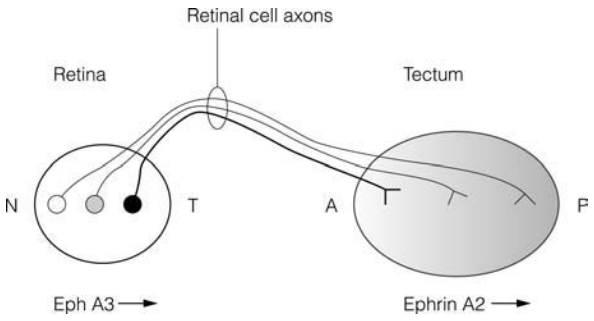


Fig. 3. Anteroposterior gradients for tectal ephrin A2 expression and retinal ephrin A3 receptor expression. These gradients provide positional information for specifying the retinotectal topographic map. N, Nasal retina; T, posterior retina; A, anterior tectum; P, posterior tectum.

This response is brought about by cell surface molecules termed **semaphorins** (e.g. collapsin-1) which bind to immunoglobulin superfamily receptors, **neuropilins** and **plexins**.

The direction in which retinotectal axons grow is specified by inhibitory interactions between **Eph receptors** on the retinotectal axons and their corresponding ligands, termed **Ephrins**, on tectal neurons. Eph receptors are members of both the tyrosine kinase family of receptors and the immunoglobulin superfamily of axon guidance molecules. Gradients for Ephrin A2 and A5 exist across the tectum with the highest concentration posteriorly. A gradient for Eph A3, exists on retinotectal axons with the greatest concentration on temporal retinal cells. Temporal cells expressing high amounts of Eph A3 will find increasingly posterior tectal regions repulsive and so tend to innervate the anterior tectum. In contrast nasal cells bearing little Eph A3 are not repelled by the posterior tectum and hence establish contacts there. At some sites in the nervous system the interaction of Ephrins with Eph receptors mediates cell adhesion and is attractive rather than repulsive.

N5 Synaptogenesis and developmental plasticity

Key Notes

Overview of synapse formation

Axons arriving at a target compete to establish synapses. Unsuccessful cells suffer apoptosis (programmed cell death). Successful cells are maintained by neurotrophic factors.

Synapse formation at the neuromuscular junction

The basal lamina in the synaptic cleft of the neuromuscular junction (nmj) can induce the formation of normal presynaptic terminals and postsynaptic structures on muscle. Motor nerve terminals secrete agrin, which binds to the basal lamina. Agrin activates mechanisms that result in clustering of nicotinic cholinergic receptors (nAChR) at the synapse. Nerve terminals also produce a protein which triggers an increase in the number of nAChRs. However, excitation of the muscle suppresses transcription of nAChR genes in muscle fiber nuclei a long way from the synapse. The net result of these processes is a high density of nAChRs at the synapse and few nAChRs extrajunctionally.

Synapse elimination at the nmj

Early in development, muscle fibers are innervated by several motor neurons, but this is pruned to single innervation by synapse elimination. The first event is a down-regulation in the number of nAChRs and this is followed by retraction of the axon. Synapse elimination is caused by voltage changes generated by the activity of neighboring axons.

Synapse formation in the CNS

The type of receptor expressed at central synapses is precisely matched to the transmitters secreted by incoming axons. Proteins implicated in clustering glutamate receptors have been identified.

Developmental plasticity in the visual system

Developmental plasticity is the modification of synaptic connections as a result of early sensory experience. It involves synapse elimination and is seen at several levels in the mammalian visual system. Ocular dominance columns in the primary visual cortex, important for binocular vision, are established by neural activity. In kittens, this occurs during a critical period after birth and the neural activity is caused by light stimulation. After the critical period no visual experience will alter ocular dominance. In primates, ocular dominance columns begin to form before birth and the activity is provided by spontaneous firing of retinal ganglion cells. Neurotrophic factors enhance cortical synaptic plasticity probably by acting in synergy with neuronal electrical activity to strengthen synapses. The timing of the critical period depends on the development of GABAergic inhibitory interneurons in the cortex.

Related topics

Ionotropic receptors (D1)
Touch (F2)
Early visual processing (G6)
Nerve–muscle synapse (J1)
Axon pathfinding (N4)
Neurotrophic factors (N6)
Cell physiology of learning (O3)

Overview of synapse formation

During nervous system development far more neurons are born than subsequently survive. When axons arrive at a target they compete with each other to form synapses; this is termed **synaptogenesis**. Cell adhesion molecules are crucial for this process. Cells which fail to establish synapses eventually suffer apoptosis. Neurotrophic factors released by the target tissue, and from other sources, ensure the survival of cells which successfully make contacts.

Synapse formation at the neuromuscular junction

Synapse formation has been most extensively studied at the neuromuscular junction (nmj). Differentiation of both presynaptic and postsynaptic structures requires signals from the **synaptic basal lamina (bl)**, an extracellular matrix within the synaptic cleft which binds collagen, acetylcholinesterase, laminin and other proteins. The signaling role of the bl is revealed by experiments in which both muscle fibers and motor nerve axons are damaged in adult frogs. Both muscle fibers and axons degenerate to leave the bl and the Schwann cells that encapsulated the axon. After a few days, myoblasts invade and differentiate into new muscle fibers, the axons regrow and synapses are formed at precisely the same places as before. If muscle regeneration is prevented by exposure to X-rays, the axon still forms a normal presynaptic terminal with active zones exactly in register with the basal lamina of the postjunctional folds; a feature of normal nmjs. If axon regrowth is prevented, the regenerating muscle fibers form normal postsynaptic structures at the original synaptic sites and nicotinic acetylcholine receptors (nAChR) cluster at the sites of the presumptive synapses as occurs normally.

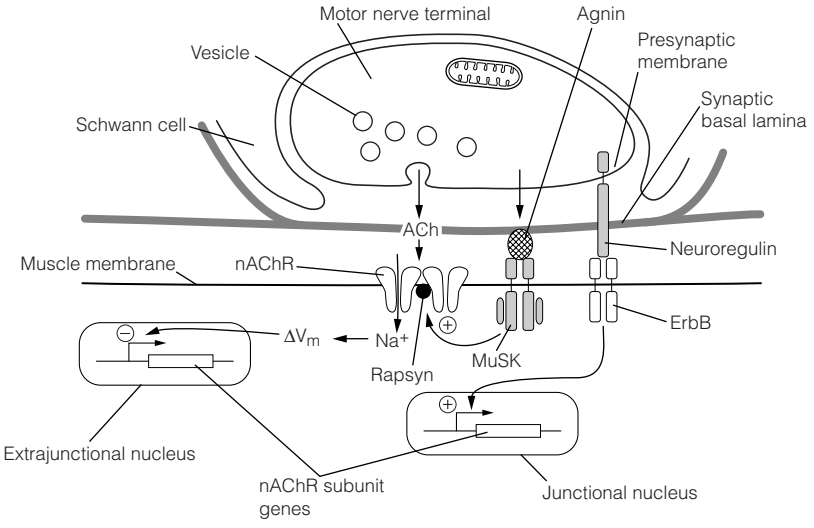


Fig. 1. Mechanisms regulating expression and clustering of nicotinic acetylcholine receptors (nAChR) at the neuromuscular junction. ErbB, Epidermal growth factor receptor; MuSK, muscle specific kinase.

Normally, developing muscle fibers express low levels of nAChR across their whole surface (about 1000 nAChR/μm^2). With the arrival of the nerve terminal, nAChRs cluster to a very high density immediately beneath the active zone (10 000 nAChR/μm^2), whereas in the extrajunctional region the density drops to 10 nAChR/μm^2. A relatively high density of nAChR is also seen over the entire surface of muscles that have been denervated for an appreciable time. This renders the muscle exquisitely sensitive to ACh, a state known as **denervation supersensitivity**.

Three mechanisms regulate nAChR at the developing synapse:

- Clustering of nAChR requires a large protein, **agrin**, that is secreted by motor neuron terminals and becomes bound to the synaptic basal lamina. The action of agrin is mediated by a receptor in the muscle membrane, one component of which is **MuSK** (**muscle specific kinase**), a receptor tyrosine kinase which promotes clustering of nAChR by activating a membrane protein, **rapsyn** (*Fig. 1* above). The agrin–MuSK–rapsyn cascade is also involved in the clustering of other synaptic proteins, e.g. acetylcholinesterase. Knockout mice lacking any of these proteins fail to form normal synapses, are immobile, cannot breathe and die shortly after birth.
- Not only do nAChRs cluster in the postsynaptic membrane, there is also an up-regulation of the receptor itself. This is due to a signaling mechanism that activates transcription of the nAChR subunit genes in nuclei that lie beneath the synapse. The nerve terminal expresses a transmembrane protein, **neuregulin** (also referred to as AChR-inducing activity, ARIA) which contains an epidermal growth factor (EGF)-like domain. It binds to EGF (ErbB) receptors to generate a transcriptional signal.
- ACh acting on nAChR causes depolarization of the muscle membrane. This activity-dependent change in membrane potential inhibits transcription of nAChR genes by nuclei that lie distant from the synaptic region.

The outcome of the processes above is that in normally innervated muscle the nicotinic receptors are clustered at the neuromuscular junction. In denervated muscle, the lack of activity causes expression of nicotinic receptors across the whole muscle surface. This gives rise to the denervation supersensitivity.

Synapse elimination at the nmj

Each motor neuron has many branches by which it innervates numerous muscle fibers. Early in development, branches of several motor neurons converge to form synapses on each muscle fiber. All but the synapse from a single motor neuron on each fiber is eliminated in mammals, as motor neurons withdraw branches from multiply innervated fibers. Eventually, each motor neuron restricts its branches to a set of fibers over which it has sole control; the motor unit (*Fig. 2*). If a muscle is denervated, exactly the same phenomenon is observed; multiple innervation occurs initially and is pruned after 2–3 weeks to single innervation.

The earliest event in synapse elimination is a down-regulation in the number of nAChRs. This is followed by a retraction of the presynaptic terminal. Clustering of nAChRs is important in maintaining the synapses, since mice lacking agrin or MuSK show loss of normal nerve terminals. The ultimate cause of synapse elimination is competition from other synapses, brought about by their activity. In adult animals, a motor neuron branches making several neuromuscular junctions (nmjs) on a particular muscle fiber. Blocking one of these with the irreversible antagonist of nAChR, α-bungarotoxin (α-BTX), causes

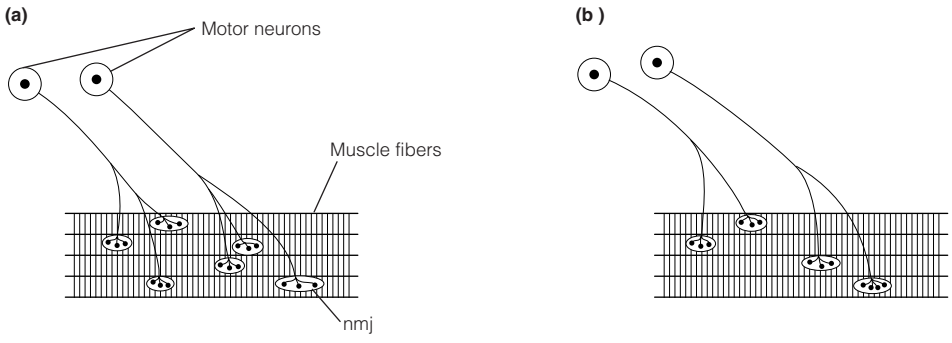


Fig. 2. Synapse elimination at the neuromuscular junction (nmj). (a) Normally each motor neuron would contact many muscle fibers. Initially, the territories of the two motor neurons overlap. Pruning of synapses leads to each neuron having exclusive control over a set of muscle fibers (b).

elimination of the blocked synapse. If several neighboring nmjs are blocked by α-BTX no synapses are eliminated. This implies that a local activity signal from unblocked synapses is responsible for elimination of blocked synapses. Molecular details of this process are not known.

Synapse formation in the CNS

Central synapses have poorly developed basal lamina and although agrin is present neither MuSK nor rapsyn are present in brain. However, appropriate receptors are clustered beneath the active zone, so mechanisms must exist which match transcription of the correct receptors in the postsynaptic cell to the arrival of terminals containing particular transmitters (i.e. glutamatergic terminals induce the expression of appropriate glutamate receptor subtypes). These receptors must also be clustered at the synapse. How this is achieved is not understood, but a family of proteins have been identified at glutamatergic synapses that are implicated in the clustering of receptors and anchoring them to cytoskeletal proteins.

Developmental plasticity in the visual system

During development, synaptic connections can be modified as a result of early sensory experience. This is termed developmental plasticity. The rewiring usually involves synapse elimination and this comes about either because of apoptosis (programmed cell death) of neurons or because inappropriate axon branches are retracted. Developmental plasticity has been extensively studied in the visual system.

In lower vertebrates (e.g. frogs), the retinotectal pathways are not much refined after being laid down. By contrast, in the rat axons from nasal and temporal parts of the retina overlap considerably when they first reach the superior colliculus (tectum). However, by the end of the second postnatal week synapse elimination, largely by retraction of spurious axon terminals, results in temporal retinal axons being restricted to the anterior part, with nasal retinal axons being confined to the posterior part of the superior colliculus.

Plasticity in the primary visual cortex, V1, is responsible for the formation of ocular dominance columns and hence for the development of binocular vision. Kittens reared with one eye sutured closed have much smaller ocular dominance columns for the sutured eye. This occurs because the normal exuberant connections established early on are largely eliminated, if they are from the closed eye, by the excessively powerful competition exerted by the open eye.

The competition exerted by the open eye is due to light-evoked electrical activity. Injecting tetrodotoxin (TTX) into the open eye to block all action potentials in its retinal ganglion cells shifts ocular dominance to the sutured eye. Exactly the same shift is seen when activation of cortical cells by the open eye is blocked by enhancing GABAergic inhibition with GABA receptor agonists.

NMDA receptors are implicated in visual system plasticity. They are expressed in greater numbers and are activated at lower threshold in very young animals than in adults. The contribution they make to visual responses in deep layers of V1 decreases from 3–6 weeks and this decline can be delayed by rearing animals in the dark. This matches the **critical period** for developmental plasticity in the kitten visual system, the timing of which is delayed if animals are kept in darkness. A critical period is a narrow time interval during development in which experience can result in long-lasting changes in behavior. During the critical period for binocular wiring of the visual system, the effect of monocular deprivation can be partly reversed by restoring binocular vision. Outside the critical period, however, no amount of manipulation of visual experience will alter ocular dominance.

In primates, unlike kittens, ocular dominance columns are established before birth. Activity is still required, but in this case is provided by spontaneous firing of retinal ganglion cells, in the absence of light, that spreads in waves across the retina. The bursts of activity in the two retinas are out of phase and this is thought to be responsible for segregation of inputs from the left and right eyes. The burst firing may work by activating NMDA receptors, since blocking NMDA receptor action disrupts the segregation of inputs. Although apparent before birth, ocular dominance columns in primates remain plastic until 6 weeks postnatally.

Neurotrophins also control visual system plasticity. Apart from their well-documented effects on gene expression, neurotrophins can modulate electrical activity of neurons by directly depolarizing them, or increasing transmitter release. Overexpression of brain-derived neurotrophic factor (BDNF) in mice accelerates synaptic plasticity and the development of ocular dominance columns. BDNF enhances cortical synaptic plasticity probably because it operates in synergy with electrical activity of neurons to produce long-term potentiation (LTP), including increasing CREB-mediated gene expression.

GABA transmission by cortical interneurons is crucial for cortical plasticity and seems to be a determinant of the critical period. Precocious enhancement of GABA inhibition by early administration of diazepam to the visual cortex brought forward the opening of the critical period. The development of GABAergic inhibition in the cortex is accelerated in mice over-expressing BDNF. These animals had earlier opening and closure of the critical period.

N6 Neurotrophic factors

Key Notes

Neurotrophic factors

Neurotrophic factors promote the survival of neurons by preventing apoptosis. They fall into three classes, neurotrophins, growth factors and cytokines.

Neurotrophins

Nerve growth factor (NGF) promotes the survival of sensory neurons in dorsal root ganglia and sympathetic neurons. Other neurotrophins, such as brain-derived neurotrophic factor, neurotrophin-3 and neurotrophin-4/5, promote the survival of a variety of central and peripheral neurons. They have a common structure.

Neurotrophin signaling

Neurotrophins bind to low-affinity and high-affinity receptors. The high-affinity receptors are tyrosine kinase receptors (trks). Binding of ligand causes the receptors to dimerize, activating tyrosine kinase domains on their cytoplasmic side. This forms binding sites for proteins, activating three signal transduction pathways that block apoptosis and promote growth by altering gene expression.

Actions of neurotrophins

Different classes of neuron require different neurotrophins for their survival. The requirement of a neuron for neurotrophins also changes during development. Neurotrophins may be secreted by the tissue through which axons must grow, their targets and even the neurons which depend on them. In the brain the highest expression of neurotrophins is in the hippocampus where they may be important in cellular mechanisms that underlie learning.

Apoptosis

Half of all neurons are estimated to suffer apoptosis (programmed cell death) during development. Apoptosis differs from necrosis, the mode of cell death that follows acute injury, by not stimulating immune responses. It is activated, in the absence of neurotrophic factors binding to trk receptors, by a biochemical cascade that switches on proteases.

Related topics

Metabotropic receptors (D2)
Axon pathfinding (N4)
Synaptogenesis and developmental plasticity (N5)

Neurotrophic factors

In some instances, synaptic connections are eliminated by the death (apoptosis) of the neurons forming them. The remaining synapses are maintained by neurotrophic factors, molecules released from the target, or other source, which ensure the survival of neurons by preventing apoptosis. Neurotrophic factors fall into three major classes.

1. **Neurotrophins**. These were first identified because of their role in promoting neuron differentiation and survival (see below).
2. **Growth factors**. Growth factors were originally recognized by their actions in stimulating proliferation and differentiation of numerous cell types. Many

of those with neurotrophic actions belong to the transforming growth factor-β class, that includes bone morphogenetic proteins (BMPs), and glial-derived neurotrophic factor (GDNF) which is required for the survival of many motor neurons.

3. **Cytokines**. Cytokines are a large and diverse group of secreted molecules associated with regulation of the immune system. In recent years it has become clear that these molecules have important roles beyond the immune system. Many interleukins are secreted by glial cells and exert neurotrophic actions, e.g. ciliary neurotrophic factor, a molecule related to interleukin-6.

Neurotrophins

The first neurotrophic factor to be discovered, **nerve growth factor (NGF)** promotes the survival of sensory neurons in the dorsal root ganglia (DRG) and sympathetic neurons during development, and maintains the adult phenotypes. NGF-treated chick embryos have greater numbers of neurons in their DRG, since fewer have undergone apoptosis, their cell bodies are larger and their axons are longer; similar changes are seen in sympathetic neurons. Chick embryos treated with anti-NGF antibodies lose most of their sympathetic neurons. NGF is secreted by the targets of the sensory and sympathetic neurons and by some neurons. Secretion of NGF from neurons is regulated by activity and depends on mobilization of Ca^{2+} from internal stores. Several other neurotrophins, e.g. **brain-derived neurotrophic factor (BDNF)**, **neurotrophin-3** (**NT3**) and **neurotrophin-4/5** (**NT4/5**) promote the survival of a variety of central and peripheral neurons (*Table 1*). Neurotrophins share a common structure. Each consists of two identical peptides about 120 amino acids long, linked by several hydrophobic residues. Variable regions containing exposed basic amino acid residues provide for the specificity shown by individual neurotrophins.

Table 1. Neurotrophins, neurotrophin receptors (trks) and targets

Neurotrophin	Receptor	Exemplar targets
Nerve growth factor (NGF)	trk A	Dorsal root ganglion cells (DRGs) Sympathetic neurons Basal forebrain cholinergic neurons
Brain-derived neurotrophic factor (BDNF)	trk B	Mechanoreceptor DRGs Motor neurons Retinal ganglion cells
Neurotrophin-3 (NT3)	trk C > trk A & trk B	Large diameter DRGs γ-motor neurons
Neurotrophin-4/5 (NT4/5)	trk B	Cranial nerve afferents Sympathetic neurons Midbrain neurons

Neurotrophin signaling

Neurotrophins bind two classes of receptor, high-affinity receptors and low-affinity receptors. The low-affinity ($p75^{LNTR}$) receptor enhances the binding of neurotrophins to the high-affinity receptors. High-affinity receptors are members of the **tyrosine kinase receptor** (**trk**, pronounced 'track') superfamily. These are large transmembrane receptors with single transmembrane segments, a tyrosine kinase domain on the intracellular face and immunoglobulin-like domains on the extracellular side. Three distinct trk receptors mediate the responses to neurotrophins, trk A–C, with the specificities shown in *Table 1*. Their expression is restricted to neurons responsive to the particular neurotrophins.

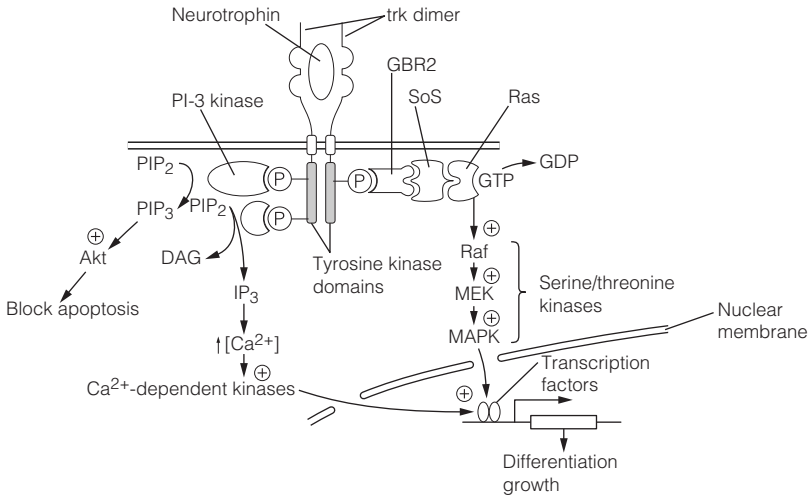


Fig. 1. Signaling by neurotrophins. Binding of SH2 domain proteins to phosphotyrosines on activated tyrosine kinase receptor (trk) dimers switches on three signaling pathways. DAG, Diacylglycerol; IP_3, inositol 1,4,5-trisphosphate; PI-3 kinase, phosphatidylinositol-3-OH kinase; PIP_2, phosphatidylinositol-4,5-bisphosphate; PIP_3, phosphatidylinositol-3,4,5-trisphosphate; SoS, Son of Sevenless.

The trk superfamily of receptors is large and apart from receptors for neurotrophins, includes receptors for growth factors (e.g. epidermal growth factor receptor, EGFR; platelet-derived growth factor receptor, PDGFR), insulin and the Eph receptors. They share common signaling mechanisms. Binding of ligand causes receptors to dimerize (*Fig. 1*) which activates the intrinsic tyrosine kinase domains. This catalyzes autophosphorylation (each receptor phosphorylating its opposite number) of several tyrosine residues to create binding sites for **src homology domain 2 (SH2)** regions on three proteins. Binding of each of the SH2 proteins initiates three second messenger cascades.

1. Binding and phosphorylation of **phospholipase C** (**PLC-γ**) generates diacylglycerol and inositol 1,4,5-trisphosphate from phosphatidylinositol-4,5-bisphosphate (PIP_2). This mobilizes Ca^{2+} from intracellular stores to have effects on assembly of the cytoskeleton and on gene transcription.
2. Activation of **phosphotidylinositol-3-OH kinase** (**PI-3K**) generates phosphatidylinositol-3,4,5-trisphosphate from PIP_2 which activates a kinase, **akt**, that promotes neuron survival by inhibiting apoptosis.
3. Binding of an adaptor protein, **GBR2**, couples the receptor to a complex containing a small GTP binding protein, **ras**, which has some homology with the Gα subunit of G proteins, and a guanine nucleotide exchange protein called **Son of Sevenless** (**SoS**). In its inactive state, ras binds GDP. SoS causes GDP to dissociate from ras so that GTP can spontaneously bind. In its GTP-bound form, ras is activated and is released from the trk receptor to switch on the MAP kinase cascade, a set of serine/threonine protein kinases that are phosphorylated in succession. The end result is activation of transcription factors and expression of genes responsible for differentiation and growth.

Actions of neurotrophins

Neurotrophins are defined by the fact that they promote the survival of neurons. However, at least some appear to act as neuromodulators influencing neuronal signaling.

Neurotrophin actions are specific. Different classes of sensory neurons in the dorsal root ganglion (DRG) require different neurotrophins for their survival. This has been demonstrated in neurotrophin knockout mice which die at birth, but primary afferents differentiate and make connections before birth. Mutant mice lacking either the gene for NGF or trkA lose the small-diameter primary afferents responsible for transmitting nociceptor and thermoreceptor information. Knockout mice lacking NT-3 or trkC genes lose large DRG cells that innervate muscle spindles and Golgi tendon organs. Mechanoreceptor afferents are unaffected by these gene deletions.

The neurotrophin requirements of neurons change during development. Very early on, sensory neurons in the DRG and trigeminal ganglia are independent of neurotrophins, but subsequently require both BDNF and NT-3 for survival. On arrival at the target, sensory epithelium trigeminal afferents lose their dependence on BDNF and NT-3 and instead require NGF, which is secreted by the epithelium. The BDNF and NT-3 needed initially is secreted locally, either by the mesenchyme through which the axons grow or by the neurons themselves. Many neurons express both the genes for BDNF and NT-3 *and* their receptors at the same time. This is an example of **autocrine** secretion in which a cell produces a factor which stimulates its own growth.

Expression of neurotrophin genes in the brain is regionally specific. Cholinergic neurons of the basal forebrain and in the striatum require NGF and trkA receptors are quite specifically restricted to these regions. The highest level of neurotrophin expression is in the hippocampus. Neurotransmission regulates neurotrophin expression; excitatory glutamate input to hippocampal neurons increases BDNF and NGF expression *in vivo*, while GABAergic input to the same cells decreases this expression. The link between neural excitation and activation of neurotrophic genes is Ca^{2+} entry. In long-term potentiation (LTP), a process thought to underlie some types of learning and which is triggered by calcium entry, there is up-regulation of neurotrophin expression. Hence neurotrophins probably contribute to the morphological changes that occur at synapses in learning and memory.

Apoptosis

About half of all neurons that are born undergo programmed cell death, **apoptosis**, during development. They do so because, in competing with other neurons to innervate their targets, they fail to establish connections in the appropriate time and, insufficiently stimulated by neurotrophic factors, they die. Glial cells also experience apoptosis, though to what extent is not clear. It has been argued that apoptosis is the default mode for cells, and only the presence of appropriate trophic factors rescue them from apoptosis. That nervous system development should proceed by overproliferation of cells followed by apoptosis, which seems uneconomic, probably allows for optimal modeling of neural connections.

Apoptosis is very different from necrosis, the cell death that follows acute injury. In necrosis, cells swell, the cytoplasm vacuolates, subcellular organelles break down and the plasma membrane ruptures to release the cell contents. This activates immune cells, microglia and macrophages, triggering an inflammatory response. By contrast, in apoptosis (*Fig. 2*) cells initially shrink, the nucleus becomes **pycnotic** (that is, chromatin condenses onto the nuclear membrane and

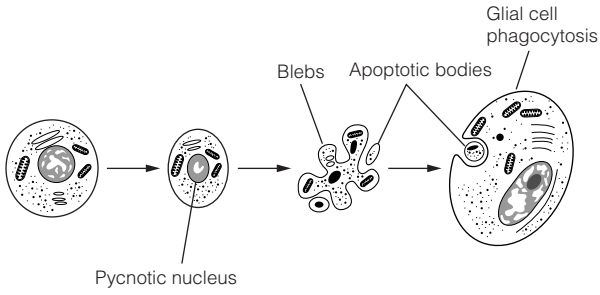


Fig. 2. Ultrastructural appearance of apoptosis. The nucleus becomes pycnotic, the cell blebs, fragments and is phagocytosed.

the nucleus fragments) and the cell blebs, breaking into membrane-bound apoptotic bodies. Subcellular organelles (other then the nucleus) remain intact and because the cell contents are not released during apoptosis there is no inflammatory reaction. Local glial cells (e.g. Schwann cells) that are not normally phagocytotic, engulf the apoptotic bodies.

Numerous genes, many of which are involved in the regulation of the cell cycles are implicated in the control of apoptosis. A simplified model of the process is illustrated in *Fig. 3*.

The sequence of events thought to be crucial to apoptosis is as follows. In the absence of trophic factor, the cytoplasmic protein **BAD** is not phosphorylated and hence binds the antiapoptotic **bcl** proteins in the outer mitochondrial membrane. This prevents the bcl proteins from interacting with **bax** proteins,

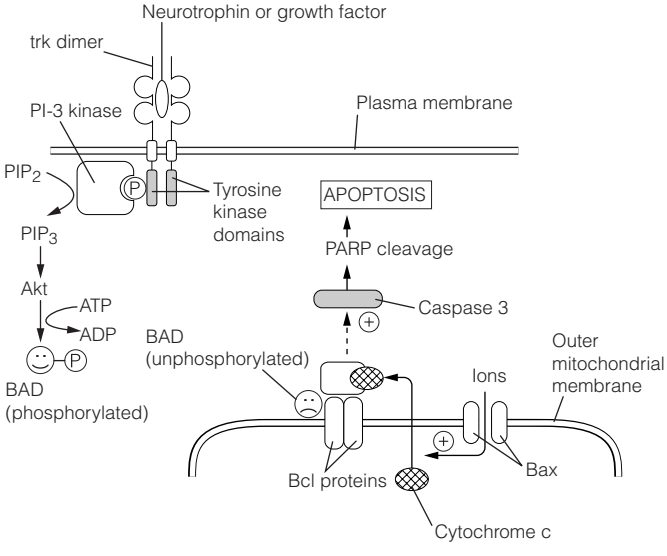


Fig. 3. A highly simplified diagram of events which inhibit or promote apoptosis. In the presence of neurotrophic factor, a soluble cytoplasmic protein, BAD, is phosphorylated and apoptosis inhibited. In the absence of neurotrophic factors BAD is unphosphorylated and initiates apoptosis (see the text for details).

which consequently form homomeric ion channels. The influx of ions through the bax channels translocates cytochrome c from the mitochondrion to the cytosol, where it activates cysteine proteases called **caspases**. One of these, caspase-3 cleaves **poly [ADP-ribose] polymerase** (**PARP**), a key enzyme in DNA repair. Brains of caspase-3 knockout mice have many more cells (neurons and glia) than normal mice, and far fewer cells show apoptotic changes.

N7 Brain sexual differentiation

Key Notes

Sexual dimorphism

Differences in brain structure and physiology and in behavior or cognition between the sexes is sexual dimorphism. The rat hypothalamus is sexually dimorphic. The preoptic area (POA) of males is concerned with copulation and tonic output of gonadotrophin, but in females is responsible for cyclical output of gonadotrophins that control the estrus cycle. The ventromedial nucleus in females organizes lordosis, a receptive sexual behavior that requires estrogens and progesterone for its expression. Other brain regions (e.g. amygdala, hippocampus and orbitofrontal cortex) show sexual dimorphism in the rat and might explain difference in cognitive skills between the sexes.

The rodent model

The rat brain is sexually differentiated by differences in hormone exposure of the two sexes during a critical period around the time of birth. High testosterone output by the male testes between E15 and P10 is responsible for the anatomical, physiological and behavioral masculinization of the male brain. Testosterone is converted to estradiol by neuronal aromatase and the estradiol acts on receptors in the hypothalamus, other limbic structures and the orbitofrontal cortex.

Human brain sexual differentiation

There is no clear evidence that human brain sexual differentiation results from early exposure to hormones. Naturally occurring mutations that expose the prenatal human brain to high concentrations of sex steroids have little effect on psychosexual development. Human male fetal testis secretes testosterone between 12–18 weeks of gestation but there are no estrogen or androgen receptors in the brain at this time. A second period of testosterone secretion occurs during the perinatal period when sexual differentiation might occur via androgen receptors.

Related topics

Neuroendocrine control of reproduction (L4)

Language (O5)

Sexual dimorphism

Male and female brains differ in terms of structure and reproductive physiology and this is reflected in the distinctive reproductive behaviors and cognitive skills of the two sexes. This is sexual dimorphism. It has been studied particularly in the rat hypothalamus where it arises as a result of exposure to hormones during a critical period. A nucleus in the medial preoptic area (MPOA), called the **sexually dimorphic nucleus of the preoptic area (SDN-POA)** is larger in males than in females. This difference is established in the perinatal period (i.e. around the time of birth), when male rats have higher concentrations of testosterone than females, and once established does not depend on the continued

presence of gonadal hormones. The MPOA in male rats is involved in copulation and maintaining tonic output of reproductive hormones whereas in females it is concerned with regulating estrus cycles. In the rat, cells in the preoptic area (POA) produce the gonadotrophin-releasing hormone (GnRH) responsible for stimulating the secretion of luteinizing hormone and follicle-stimulating hormone from the anterior pituitary. In females, gonadotrophin release can be enhanced by the high concentrations of estrogens produced by mature follicles. This is mediated by neurons in the POA which express estrogen receptors and trigger GnRH secretion. In males, estrogens do not produce increased gonadotrophin release. In male rats and rhesus monkeys, MPOA lesions virtually abolish copulatory behavior, and firing of MPOA neurons correlates with copulation.

The ventromedial hypothalamus (VMH) is concerned with lordosis in female rats. **Lordosis** is receptive sexual behavior in which the rat raises her hindquarters and moves her tail out of the way to facilitate copulation. Ablation of the VMH abolishes lordosis, and recording from behaving female animals shows that VMH neuron activity correlates with lordosis. By contrast, in male rats VMH neurons are inhibited during copulation by the MPOA. Lordosis in female rats requires estrogen exposure for 24 hours, followed by progesterone acting for 1 hour. Estrogens cause the up-regulation of progesterone receptors in the VMH in female rats but not in male rats. Hypothalamic sexual dimorphism in rats is summarized in *Fig. 1.*

Other brain regions in the rat, amygdala, dorsal hippocampus and orbitofrontal cortex, are sexually dimorphic and might account for cognitive differences between the sexes. For example, while male rats are better at maze learning than females, the reverse is true for avoidance learning.

Brain sexual dimorphism occurs in primates, including humans. Reports that nuclei in the hypothalamus assumed to be equivalent to the SDN-POA of the rat are larger in human males than females have not been confirmed in other studies. The anterior commissure and another structure connecting the two hemispheres are larger in women than men. In rhesus monkeys, the orbitofrontal cortex is involved in certain spatial discrimination tasks in which

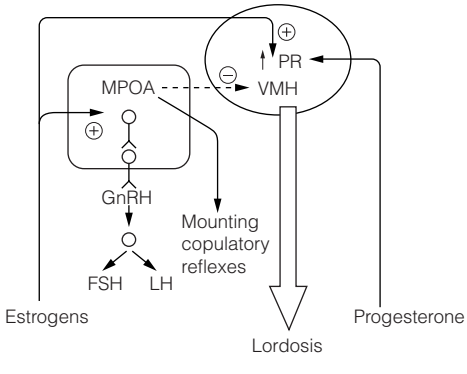


Fig. 1. A model for hypothalamic involvement in reproductive functions in the female rat. Typical female behavior, lordosis, requires progesterone and estrogens. The small medial preoptic area (MPOA) exerts only slight inhibition on the ventromedial hypothalamus (VMH). In the males the much larger MPOA organizes typical male sexual behavior and strongly inhibits the VMH.

adult males outperform adult females. Lesioning the orbitofrontal cortex at different times shows that its ability to mediate these tasks arises earlier in male than female monkeys.

In humans, lateralization of cognitive functions is seen in which in the great majority of people, the left hemisphere is specialized for language tasks, while the right hemisphere is specialized for non-verbal, visuospatial tasks. This functional asymmetry is weaker in females than males. PET scans during verbal language tasks show that women have some activity in the right as well as in the left hemisphere. In men there appears to be little involvement of the right hemisphere. Less hemispheric specialization of cognitive function in girls may allow their brains to retain greater plasticity for longer than boys, which confers advantages. Recovery of language function following left hemisphere damage in childhood is better in girls than boys, presumably because of the greater plasticity of the right hemisphere in the girls. Developmental dyslexia, aphasia and autism, conditions in which language deficits are predominant, are associated with left hemisphere dysfunction and are much more common in males.

The rodent model

Sexual differentiation of the rat brain is due to differences in hormone exposure during a critical period, the perinatal period. The testis secretes high concentrations of testosterone from embryonic day 15 (E15) to postnatal day 10 (P10) which is responsible for masculinizing the male brain; the gestation period in rats is 21 days.

This is confirmed in studies that exposed rats in the first four postnatal days (P1–P4) to inappropriate hormone environments. Female rats injected with testosterone during P1–P4, when adult, had anovulatory sterility, showed no enhanced gonadotrophin release in response to estrogens and had much reduced female typical sexual behavior (lordosis), together with male typical sexual behavior. These animals had large MPOAs. Castration of males on day P1 produced adult animals that would show enhanced gonadotrophin output when given estrogens. When given estrogens and progesterone they exhibited lordosis. Neonatal castration resulted in small (i.e. female-like) MPOAs. Hormone manipulation of rats beyond P10 had no effect on their brain sexual development.

Studies of hypothalamic cells in culture show that testosterone promotes neurite outgrowth but it does so by first being converted, by the enzyme **aromatase,** to estradiol which acts on estrogen receptors (*Fig. 2*). High levels of neuronal aromatase are expressed perinatally in the same cells that express estrogen receptors, i.e. hypothalamus, amygdala and other limbic structures and orbitofrontal cortex.

Hence, masculinization of the male rat brain is brought about by aromatization of the testosterone (secreted between E15 and P10) to estradiol, which then acts on estrogen receptors to promote neuron growth. Maternal blood

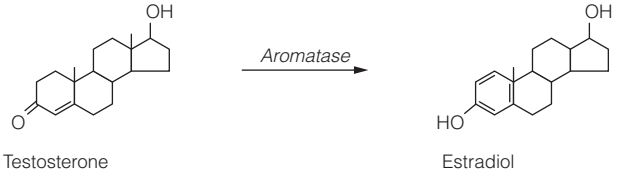


Fig. 2. Aromatization of testosterone to estradiol.

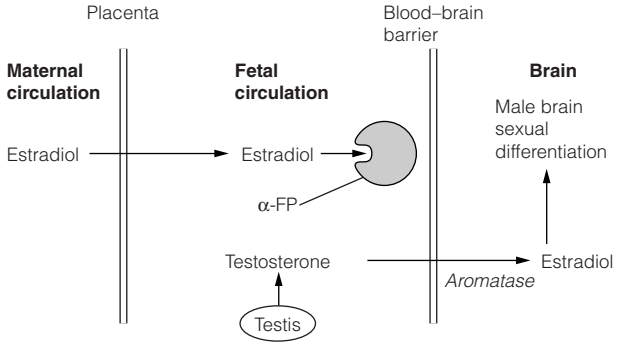


Fig. 3. A model for brain sexual differentiation in rats.

concentrations of estradiol are high in late pregnancy. However, female fetal rats are not masculinized because the fetal circulation contains α-fetoprotein which binds estrogens so that they do not cross the blood–brain barrier (*Fig. 3*).

Human brain sexual differentiation

Although the human brain shows some sexual dimorphism, unlike rats, there is no good evidence that this is due to early differences in hormone exposure. Girls exposed to very high androgen concentrations as a result of an inborn error of metabolism, **congenital adrenal hyperplasia**, do not generally show shifts towards male sexual behavior, although there is a slightly higher incidence of homosexuality in this group. **Androgen insensitivity syndrome** is due to loss of function mutations in the gene encoding androgen receptors. Affected genetic (XY) males have a short vagina and female secondary sexual characteristics because their testes secrete testosterone which is aromatized to estradiol. These individuals look like normal women and they behave as females, forming sexual relationships with men, *despite* their brain having been exposed to high concentrations of estrogens throughout their development. This implies that the rodent model is not applicable in humans.

In human male fetuses, the testis secretes testosterone between 12–18 weeks of gestation. Aromatase is present in the hypothalamus during this time. However, no estrogen or androgen receptors are found in the human brain between 12 and 24 weeks. At two other times during development, in the perinatal period and at puberty, human males have much greater concentrations of testosterone. However, since human brain sexual dimorphism is apparent from about 2 years of age, if it is organized by hormone exposure, sexual differentiation must occur during the perinatal period. Furthermore, the androgen insensitivity syndrome implies that if hormones are responsible it must be androgens acting at androgen receptors, not estrogens.

01 Types of learning

Key Notes

Definition of learning

Learning is the acquisition of altered behavior as a result of experience and occurs by rewiring of neural pathways (plasticity). Storing the changes over time is memory. Prior learning is tested by recall, elicited by the appropriate stimuli.

Declarative and procedural memory

There are two broad categories of memory, declarative and procedural. Declarative memory is memory for facts, is fast and consciously recalled. It can be for a set of facts that are associated together because they all relate to a single event (episodic memory), or for facts in isolation (semantic memory) that are stored in a manner that reflects similarities between them. Procedural memory is memory for motor skills. It is slow and not recalled consciously. Many learning situations include elements of both categories.

Short-term and long-term memory

Declarative memory has at least two temporal phases. Short-term memory (STM) is brief, of limited capacity and requires continual rehearsal of items to retain them in STM. Long-term memory (LTM) is long lasting and apparently of unlimited capacity. Information may enter STM and LTM sequentially or in parallel. Amnesias (loss of memory) following brain trauma usually afflict LTM and leave STM untouched. Loss of memory for events before the trauma is retrograde amnesia, the inability to form memories subsequent to the trauma is anterograde amnesia.

Consolidation

Consolidation is the process that makes memories increasingly resistant to disruption. For declarative learning, serial models postulate that consolidation transfers selected material from STM to LTM, whereas parallel models have input going into both STM and LTM, with consolidation saving selected material in LTM. When attempting to learn two similar motor tasks the second task can impair subsequent performance of the first one if it follows quickly but not if delayed by 6 hours, showing that consolidation is complete by this time. In some instances recalling a consolidated memory returns it to a fragile state and it must be reconsolidated for the memory to persist. Both consolidation and reconsolidation require protein synthesis.

Non-associative and associative learning

Procedural memory is either non-associative or associative. Only a single type of stimulus is needed for non-associative learning. In habituation, repetitive delivery of a weak stimulus causes the loss of a motor response. Sensitization is the enhancement of a response to innocuous stimuli seen after an unpleasant stimulus. Associative learning requires pairing of two events within a short time. In classical conditioning, animals learn an association between one stimulus (the conditioned stimulus) and the appearance of a second that may be rewarding or unpleasant (unconditioned stimulus). The conditioned stimulus must

always be presented immediately before the unconditioned stimulus. In operant conditioning, animals learn an association between some action they perform and the arrival of a stimulus which may be either rewarding or aversive.

Related topics

Synaptogenesis and developmental plasticity (N5)
Physiological psychology of memory (O2)

Definition of learning

Some neural pathways establish connections during development that subsequently remain unaltered. These pathways are often said to be **hard-wired** and the generic term for these processes that ensure the pathway is properly connected is **specificity**. However, pathways subject to continual re-wiring, either during development or as a result of experience, are referred to as **plastic**, and the re-wiring processes described as **plasticity**. **Developmental plasticity** is wiring that is conditional on early sensory experience and shapes subsequent perception. **Learning** is also plasticity and is the *acquisition* of reproducible alterations in behavior as a result of particular experiences. The storage of the altered behavior over time is **memory**. In animals, learning and memory can only be tested operationally by **recall**, in which the previously learned behavior is elicited by the appropriate stimuli.

Declarative and procedural memory

Learning occurs in a variety of distinct situations, differing in time course, stimulus requirements and outcomes, so is classified in several ways to reflect these features (*Fig. 1*). A major distinction is between declarative and procedural memory. Some authors argue for a third emotional category of learning.

Declarative (explicit) memory is memory for facts. Declarative learning is fast, it requires few trials, requires conscious recall and may be readily forgotten. It has two components that are dissociable in patients with cortical damage. **Episodic memory** is memory for specific events, in which associations are established at a specific time and place (e.g. going to a gamelan concert

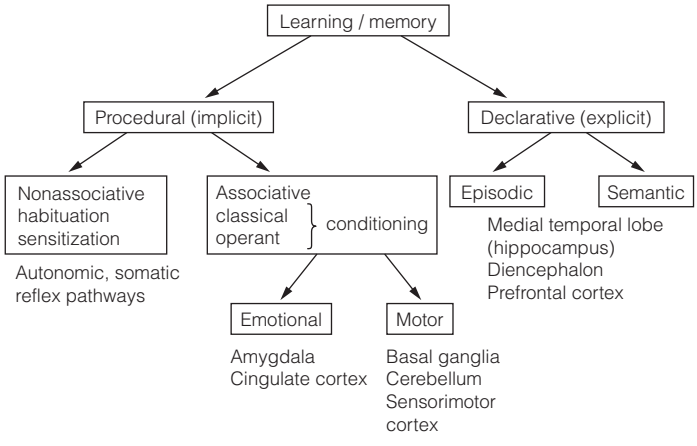


Fig. 1. Types of memory. Structures responsible for a particular category are shown in square brackets. Episodic learning is also associative.

whilst on holiday on Bali). The ability of rats to navigate through a maze in which they must learn to associate their positions in the maze with cues in their surroundings, **spatial navigation learning**, is a special case of episodic learning. In humans, PET scans reveal the involvement of the hippocampus, medial temporal lobe and prefrontal areas in episodic learning.

The second component is **semantic memory** which is memory of facts unrelated to events; that Bali is an Indonesian island can be recalled without ever having been there, so semantic memory is about 'knowing that'. PET scans of patients with defects of semantic memory show reduced metabolic activity in the anterior temporal lobes, particularly in the left hemisphere. Studies of brain-damaged patients show that semantic memories are sorted into categories (sets of related objects), which appear to be stored in different areas of brain. However, recall of specific items seems to need activation of multiple brain sites, each of which codes for a given attribute (e.g. color, function, name) of the item. Patients with bizarrely specific deficits (e.g. losing the ability to name fruit, while still being able to name vegetables) have been reported.

Procedural (**motor**) **memory** is memory for skills, such as learning to walk, swim, ride a bike or play a musical instrument. It is 'knowing how' memory. Procedural memory is slow, it needs many trials – in other words, a lot of rehearsal – and it is incremental in that improvement occurs gradually over time. Performance of procedural tasks does not involve conscious recall. For this reason procedural memory (like emotional memory) is described as **implicit memory**. Once established, procedural memories are not forgotten even after many years without rehearsal.

Many tasks have both factual and skill memory components. Playing the flute requires declarative memory for the musical notation of the score and procedural memory for the sequence of finger movements and the breathing pattern needed to create the sounds.

Short-term and long-term memory

Declarative memory has at least two (and probably more) phases categorized by their time course. **Short-term** (**recent**) **memory** (**STM**) is temporary, limited in capacity, requires continuous rehearsal to keep it and is easily disrupted by conflicting input. It is continually decaying, as the oldest unrehearsed items are lost, and updated by addition of new input. STM is often tested in humans by the ability to recall random strings of digits after a single presentation. If the subject successfully remembers five digits the next trial has six digits; each successive trial has one extra digit. If the subject makes an error, the previous sequence is repeated until the subject gets it right. The number of trials to success is plotted against the number of digits. Normal subjects show **primacy**, greater recall of material at the beginning of the list, because they are rehearsed most often, and **recency**, better recall for items at the end of the list, because STM for these has decayed the least.

Long-term (**remote**) **memory** (**LTM**) is, if not permanent, at least long lasting, has a capacity which is so great that there appears to be no obvious limit in a human lifetime, and does not require continual rehearsal. LTM and STM can be distinguished physiologically in that STM is disrupted by anesthesia or temporary cooling of the brain. LTM is unaffected by these procedures. In addition, **amnesias** (loss of memory) due to brain damage show that STM and LTM are separable. Usually amnesia affects LTM but leaves STM untouched. A few patients, however, have intact LTM but impaired STM, showing that information can enter LTM even if STM is compromised. Amnesias of LTM are of two

types, depending on whether memories are lost for events and facts acquired before, **retrograde amnesia**, or after, **anterograde amnesia**, the brain damage. Brain trauma often results in elements of both.

Consolidation

Consolidation is the process that makes both declarative and procedural memories increasingly resistant to disruption or interference from similar learning over time. How consolidation operates in declarative learning is uncertain because the exact relationship between STM and LTM is not clear. Serial models postulate that before STM decays, elements may be selected by attention and arousal mechanisms for consolidation into LTM (*Fig. 2*). Alternatively input may go in parallel to both STM and LTM, with consolidation acting to save selected material in LTM.

Interference in procedural tasks is seen when attempting to acquire two similar motor tasks in quick succession. Learning the second task impairs subsequent performance of the first one. Interference does not occur (in humans) if the second task is delayed by 6 hours. This implies that consolidation is complete by this time. Experiments in which animals are injected shortly after training on a novel task with antibiotics which inhibit protein synthesis (e.g. anisomycin, cyclohexamide) shows that consolidation requires protein synthesis.

In at least some cases recalling a consolidated memory returns it to a labile state that is once again sensitive to interference. It must now be **reconsolidated** for the memory to persist. For example, if established auditory fear memories are reactivated in rats, protein synthesis is required for the memory to be retained; injection of anisomycin into the basolateral amygdala at the time of reactivation hinders retrieval in a subsequent test session. The biochemistry of consolidation and reconsolidation appears to be different.

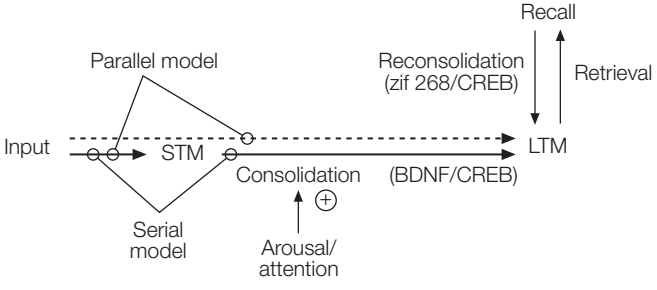


Fig. 2. Possible functional relationships between short-term memory (STM) and long-term memory (LTM). Different patterns of gene expression imply consolidation and reconsolidation are biochemically distinct: BDNF, brain-derived neurotrophic factor; CREB, cyclic AMP response element binding protein; zif 268, zinc finger transcription protein 268

Non-associative and associative learning

Procedural memory is learning to produce a motor response to a particular input. It is divided into two types; non-associative and associative. **Non-associative learning** occurs in response to only a single kind of stimulus. Two examples are **habituation**, in which repeated exposure to a weak stimulus results in a reduction or a loss of the response normally seen with occasional presentation of the stimulus, and **sensitization**, which is an exaggerated response to innocuous stimuli following a strong noxious (unpleasant) stimulus.

Associative learning needs the pairing of two different types of stimulus within a short time and in the correct order. It enables animals to behave as if they can predict relationships of the kind: if A then B. **Classical conditioning** was first investigated in dogs that learned to associate a sound with a subsequent food reward. Hungry dogs salivate at the sight or smell of food. The food is the **unconditioned stimulus** (**US**) and the salivation an **unconditioned response** (**UR**), so called because the nervous system is hard-wired in such a way that salivation occurs as an autonomic reflex response to food. If, in a series of training trials, a sound, the **conditioned stimulus** (**CS**), is presented shortly before the arrival of the food, then in a subsequent test, presentation of the sound *alone* will elicit salivation. The salivation response is now a **conditioned response** (**CR**) since the animals have learnt to salivate when the CS (sound) is presented. Classical conditioning is characterized by **temporal contiguity**, the requirement that the CS must be presented *before* the US, and **contingency**, that animals learn that a predictive relationship exists between the CS and the US. **Extinction** of the conditioned response occurs if the CS is repeatedly presented without the US or if the temporal pairing of the CS and US is disrupted (i.e. if they are presented randomly). Extinction is *not* the same as forgetting. If after extinction the pairing of CS and US is restored the CR returns much more rapidly than it does in naive animals. Classical conditioning can allow perceptual capabilities to be explored. For example, if the CS is a red light, a response will be conditioned only if the animal can distinguish red light from light of other wavelengths. Classical conditioning in which the US is noxious and which results in fear responses to normally neutral stimuli is **aversive conditioning**.

In **operant (instrumental) conditioning** an animal learns an association between a motor activity it performs (e.g. pressing a lever) and the arrival of a stimulus, termed the **reinforcer** (e.g. a food pellet). Reinforcers may be positive, in which case they increase the probability that an animal will act to obtain it, or negative (an aversive stimulus, such as an electric foot shock) in which case the animal will work to avoid it. Operant conditioning is used to investigate motivated behaviors.

02 Physiological psychology of memory

Key Notes

Hippocampus

Damage to the structures in the medial temporal lobe, particularly the hippocampus in humans, monkeys and rats, causes both retrograde and anterograde amnesia. Procedural learning is unaffected. The hippocampus is particularly involved in the consolidation of new episodic memories into LTM. In primates, these memories are very general but in rats the hippocampus seems primarily concerned with place learning by which animals are able to find their way around. The rat hippocampus is the site of cognitive maps which are updated as the animal explores its surroundings. The maps allow the animal to navigate its way through the world. Evidence for the cognitive map hypothesis comes from the presence in the hippocampus of place cells which fire only when the rat is in a particular location, which constitutes the place field of a cell. Place cells use a combination of sensory and locomotor cues to encode the spatial relations between objects in the rat's environment.

Diencephalon

Medial temporal lobe structures are connected to nuclei in the hypothalamus and thalamus. Lesions of these diencephalic structures in monkeys or humans (either through trauma or disease) cause severe amnesias.

Working memory

For complex tasks sensory input, current STM and items recalled from LTM can be held in a temporary store, working memory, for processing. Much of the neural machinery for working memory is shared with STM. Components of both include a phonological loop which holds verbal information and requires the left cerebral hemisphere, and a visuospatial sketch pad that holds information about spatial relations and requires the right hemisphere.

Prefrontal cortex

The connections of the prefrontal cortex with temporal lobe and diencephalic structures involved in learning and the effects of lesioning it, both in monkeys and humans, implies that it is concerned with tasks requiring working memory.

Memory modulation

The degree of arousal determines the probability that specific memories will be consolidated. Arousal, signaled by release of catecholamines from the sympathetic nervous system, stimulates afferents in the vagus nerve. This activates the brain noradrenergic arousal system, which projects to the amygdala and hippocampus to bring about consolidation. Hormones secreted by every level of the hypothalamic–pituitary–adrenal axis that is activated in stress also have actions on the amygdala and hippocampus. Optimal learning occurs with moderate catecholamine or glucocorticoid concentration. High levels of these hormones are detrimental. The cholinergic attentional system in the forebrain is necessary for memory. It is activated by the amygdala in response to salient stimuli.

Memory and sleep

The memory hypothesis of sleep postulates that sleep is a favorable time for consolidation and reorganization of memories. The hypothesis is based on four results: learning increases the duration of REM sleep the following night, performance in memory tasks is enhanced during the next sleep period, sleep deprivation impairs memory, and physiological studies imply that memory processing occurs during sleep. Some studies suggest that REM sleep is involved in procedural and emotional (implicit) memory, whereas slow-wave sleep may have a role in declarative (explicit) memory. PGO spikes in REM sleep may be needed to erase unwanted associations. REM sleep appears to be a playback mode. Firing patterns of hippocampal cells seen during training are recapitulated during REM sleep as information flows from neocortex to hippocampus. This consolidation is favored by activation of the cholinergic septohippocampal pathway. During slow-wave sleep information flow is in the reverse direction – hippocampus to cortex – thought to correspond to the transfer of memory into the cortex and its gradual loss from the hippocampus. A problem with the memory hypothesis of sleep is drugs that almost irradicate REM sleep have no effect on memory even with long-term use. During REM sleep brain activity is reduced in the prefrontal cortex (involved in planning and problem solving), but increased in the limbic cortex and the amygdala. This may account for dreams having little episodic memory, but appreciable emotional content.

Related topics

Noradrenaline (norepinephrine) (D5)
Acetylcholine (D7)
Neuroendocrine control of metabolism and growth (L3)
Sleep (M5)
Types of learning (O1)
Cell physiology of learning (O3)

Hippocampus

Much of the evidence implicating brain structures in memory is from studies of humans with brain damage or lesions in animals. Two clinical cases illustrate the involvement of the medial temporal lobe, particularly the hippocampus (*Fig. 1*).

The first, HM, had neurosurgery to treat intractable epilepsy. The surgery removed an 8 cm length of the medial temporal lobe including the amygdala, the anterior two thirds of the hippocampus and the overlying cortex from both hemispheres. Whilst the surgery successfully alleviated the seizures it produced devastating deficits in his declarative memory. While his short-term memory and very remote memory were normal he had a partial retrograde amnesia for some years before the surgery, and an anterograde amnesia so extreme he could not consolidate any new long-term memories; he was unable to retain the memory of events, places or people that he had experienced since his surgery for longer than the time he could hold them in STM. HM could learn new procedural tasks, though he could not remember having acquired them. The second, RB, suffered brain hypoxia following heart surgery, and subsequently had a memory deficit of the same character, though not as severe, as HM. Postmortem examination 5 years later revealed only a bilateral, highly selective, loss of pyramidal cells from one region of the hippocampus.

Bilateral medial temporal lobe lesions in macaque monkeys provide an animal model for human amnesias. Typically, the animals are trained on a

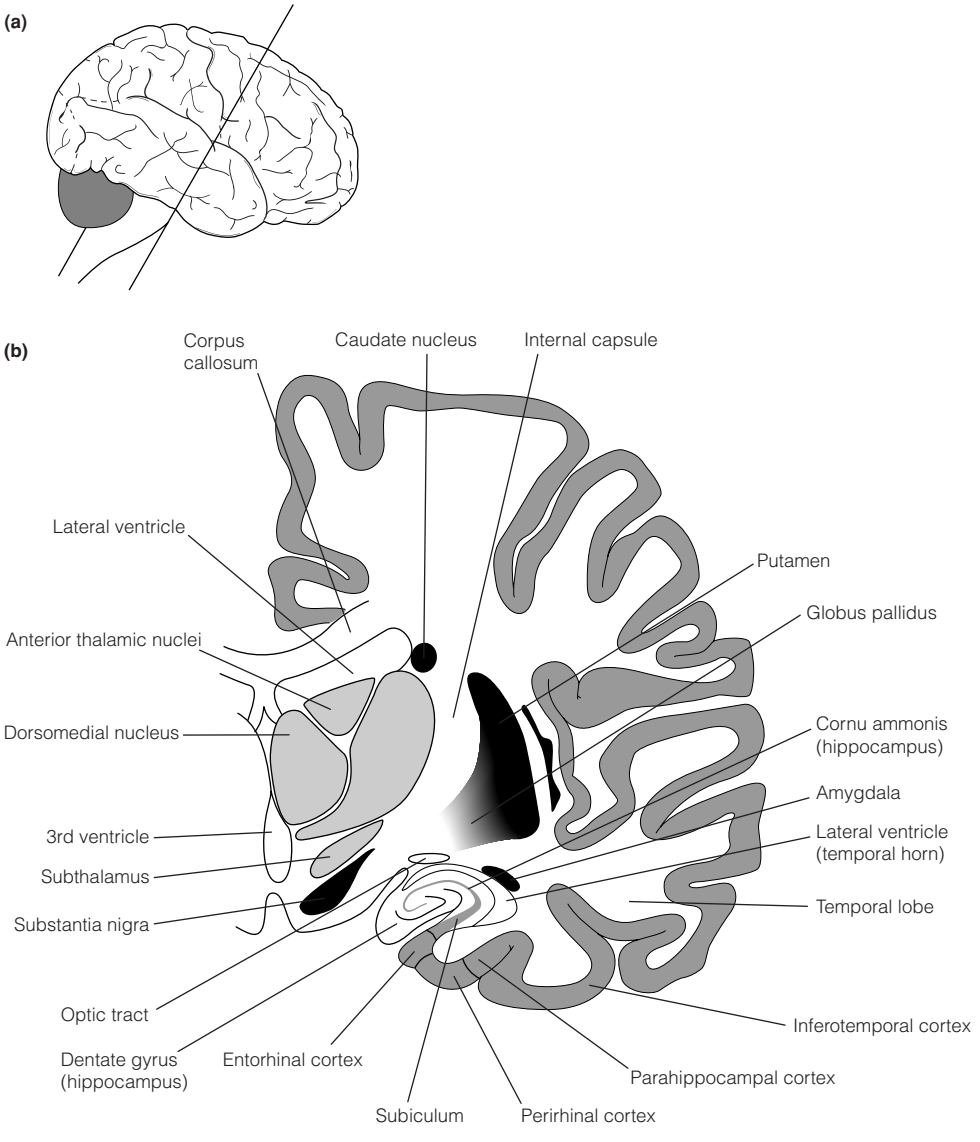


Fig. 1. Section (a) through the human brain to show (b) the gross anatomy of the medial temporal lobe.

delayed non-matching to sample task. In this, the monkey is trained to select an object to get a food reward. Following a variable delay during which the animal cannot see any manipulations, it is given a choice between the same object and a novel object and is required to select the novel object to get the reward, i.e. it needs to remember which object it saw first. Lesioned animals show an anterograde amnesia and selective lesioning shows that the most severe deficit occurs with damage to the perirhinal and parahippocampal cortex of the temporal lobe.

In rats, pure hippocampal lesions cause proportionally greater learning deficits than in primates. The hippocampus of rats is particularly important for

spatial navigation (**place**) **learning**, by which animals acquire a memory for their location. One widely used way of investigating this is the **Morris water maze**. This is a circular pool filled with opaque warm water. Hidden just below the surface is a small platform. During learning trials rats swim in the pool, discover the platform and learn its position in the pool relative to cues in the laboratory. The motivation is that the platform provides an escape from the water. Learning takes several trials and can be tested by measuring the time taken for a rat to reach the platform or the length of the path it swims to reach it as recorded on a video camera. Cued control experiments in which the platform is raised just above the level of the water ensure that any differences in behavior are not attributable to any motivational, perceptual or locomotor factors.

Rats with hippocampal, but not selective neocortical, lesions are seriously compromised in the learning, but not the cued, versions of this task. Rats given a microinjection of colchicine to destroy a specific population of cells (dentate granule cells) in their hippocampus either 1, 4, 8 or 12 weeks after learning the location of a submerged platform, and tested 2 weeks later (*Fig*. 2), reveal that the hippocampus is not a permanent site for the spatial memory. The 12-week group remembered the location as well as control (uninjected) animals, but performance got progressively worse for 8-, 4- and 1-week groups. This study shows that the hippocampus is needed for consolidation of spatial learning, but that over successive weeks the site of the memory store is transferred elsewhere, probably the neocortex. Retrograde amnesia in humans with medial temporal lobe damage probably results from the loss of memory not yet transferred from hippocampus to neocortex. Although the hippocampus may be predominantly for spatial learning in rats, in primates it has a broader role in consolidating all episodic memories.

The hippocampus of the rat and associated cortex is thought to provide the rat with a representation of the space around it and its location within it. This is the **cognitive map hypothesis** and it has several postulates. Firstly, the map allows the animal to find its way through the environment. Secondly, it is constructed by episodic learning as specific locations come to be associated with particular sensory and motor cues. Thirdly, it does not require reinforcers and finally, the map is continually updated by exploration.

Evidence for the hypothesis is the existence of **place cells**, pyramidal cells in the hippocampus that fire when the rat is in a particular position in the environment. In a typical experiment, rats with electrodes chronically implanted in the hippocampus for extracellular recording, are allowed to explore a plus-shaped maze. The animals learn the spatial relationships between the maze and visual cues in the surrounding laboratory so as to find a food reward located in one of the four arms of the maze. The maze can be rotated with respect to the constellation of cues in the laboratory. The location in the maze which causes the place cell to fire is the cell's **place field** (analogous to the sensory field of sensory

Fig. 2. Protocol for investigating the time course of place learning by rats in the Morris water maze.

neurons). The properties of place fields imply that they encode spatial relations between features of the rat's surroundings:

- An array of place fields represents a particular environment.
- There is no correspondence between the locations of place fields in the world and the positions of the place cells in the brain.
- A given place cell may have several place fields, each in a different context.
- Place fields move in concert with rotation of visual cues in the laboratory, but remain when the lights are turned off, so they are not *explicitly* coupled to sensory input.
- They enlarge or shrink with manipulations in the size or shape of an enclosed space in a way that implies the rat uses locomotor cues (e.g. number of footsteps) to work out where it is.
- New place fields arise as an animal explores a novel environment.
- Altering familiar surroundings disrupts pre-existing place fields.
- In maze tasks in which a rat must know its location in order to get a food reward, if the rat makes an error, the place fields correspond to the incorrect location (i.e. where the rat 'thinks' it is).
- Old rats, which consistently show deficits in spatial learning, have place fields with poorer spatial selectivity and reliability (stability from trial to trial).

When rats explore a maze, the EEG shows a theta (θ) rhythm with a frequency of 4–10 Hz. This reflects periodic firing of hippocampal neurons that is driven by the cholinergic septohippocampal pathway. During θ discharge, place cells encoding new information about the environment fire in phase (i.e. during the peaks of the θ wave), but all other pyramidal cells are silenced by inhibition from GABAergic interneurons. Thus, θ activity ensures that only cells involved in learning a particular environment are active.

Diencephalon

Three diencephalon structures are extensively connected with the temporal lobe and play a role in memory. A major output of the hippocampus is the **fornix** which projects largely to the **mammillary bodies** of the hypothalamus, output

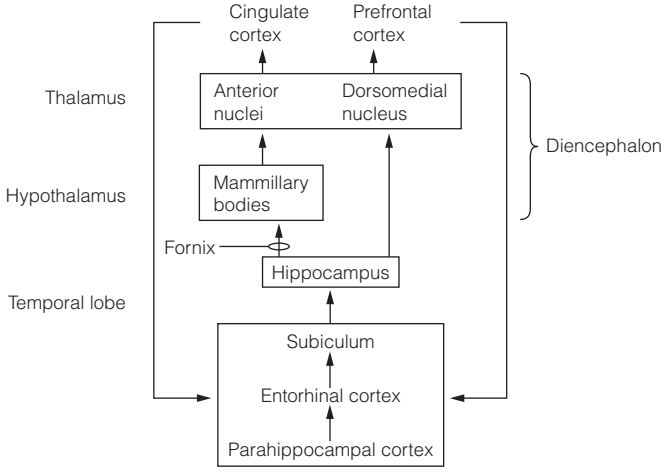


Fig. 3. Mammalian forebrain memory circuitry.

from which goes to the **anterior thalamus**. Furthermore, areas of the temporal cortex and amygdala make connections with the **dorsomedial nucleus** of the thalamus. Bilateral lesions confined to just one of these diencephalonic structures in monkeys modestly impairs performance in the delayed non-matching to samples tasks, but larger lesions affecting all three produce very severe deficits. For example one individual, NA, sustained damage to his left dorsomedial thalamus in a fencing accident. NA suffered amnesia very similar to HM, but less severe. Moreover, severe anterograde and retrograde amnesia are symptoms in **Korsakoff's syndrome**, in which damage occurs to the dorsomedial nucleus, mammillary bodies and other brain regions due to the vitamin B1 deficiency of chronic alcoholism. Hence interconnected medial temporal lobe and diencephalon structures form components of a brain memory system. The proposed circuitry for declarative memory is summarized in *Fig. 3*.

Working memory

For complicated tasks (e.g. speaking or listening to speech, planning ahead, problem solving) information from several sources, i.e. ongoing sensory input, items in STM and material recalled from LTM, must be available for processing. This is made available in a temporary store termed **working memory** that is distributed in multiple sites in the brain. For example, when driving it is necessary simultaneously to hold in working memory the position and speed of other traffic (sensory input), the current speed limit (STM), both of which need to be continually updated throughout the journey, and the intended route (accessed from LTM).

Working memory uses at least some of the same neural machinery as short-term memory. Humans with cortical lesions show that there are at least two independent subsystems for STM, also regarded as components of working memory. The **phonological loop** (**verbal sketch pad**) allows speech sounds to be held for long enough to give continuity to spoken language, so that phrases and sentences can be comprehended. It requires the left cerebral hemisphere. The **visuospatial sketch pad** is a temporary store for visual and spatial input that PET scans indicate involves several regions in the right hemisphere.

Prefrontal cortex

The **prefrontal cortex** (**PFC**) is involved in complex problem solving and planning future actions and there are good reasons for supposing that these executive tasks require working memory. The connectivity of the prefrontal cortex argues for its role in working memory. Firstly, association fibers make reciprocal connections between the PFC and other cortical areas, so the PFC receives visual, auditory and somatosensory information. Secondly, because the PFC is interconnected with the medial temporal lobe and dorsomedial thalamus that have a well-documented role in learning and memory.

In **spatial delayed response tasks**, monkeys see a food reward placed in one of several covered locations. After a delay, which can be varied over trials, the animals are tested to see if they remember the location of the food. Monkeys with lesions of the prefrontal cortex have deficits in these tasks, and performance degrades progressively as the delay is lengthened. Recording from the PFC in alert behaving monkeys reveals cells that fire in predictable ways during delayed-response tasks. For example, many cells fire throughout the delay period, others fire when the food is placed in the location and when the animal is allowed to choose the location. Particular regions of the PFC seem to be modality-specific and so responsible for specific types of working memory.

Humans with prefrontal lesions also show deficits on working memory tasks

in which they are required to use recent data to make correct decisions. Such individuals have great difficulty in tracing a path through a drawing of a maze. They will make the same errors repeatedly, and start right from the beginning of the maze after making a mistake, rather than from the position in the maze just before they made the error.

Memory modulation

The arousal levels associated with an event modulate the likelihood of specific memories being consolidated. The arousal signals to which the brain memory circuits respond are adrenal hormones (both catecholamines and steroids), and several CNS peptide neurotransmitters released in response to stress.

Evidence for the involvement of catecholamines includes:

- Enhanced recall of emotionally neutral learning tasks by noradrenaline (norepinephrine) or adrenaline (epinephrine) given within a short time of the learning trials.
- No better recall of an emotionally charged version of a story compared to the neutral version after administration of the β-adrenoceptor antagonist propranolol.
- People with higher levels of sympathetic activity are more likely to suffer from post-traumatic stress disorder after a traumatic experience.

The catecholamine dose–response curve has an inverted U shape; moderate concentrations are more effective enhancers of memory than either high or low levels. As neither of these hormones crosses the blood–brain barrier their actions on the CNS must be exerted peripherally. The catecholamines act at β-adrenoceptors on visceral afferents that run in the vagus (X) nerve to the nucleus of the solitary tract. This results in activation of noradrenergic neurons of the locus ceruleus that are part of a brain arousal system. This system projects to the amygdala and hippocampus to modulate learning. Electrically stimulating the vagus nerve immediately after training improves recall in an inverted U relationship with firing frequency. Cutting the vagus nerves or lesioning the nucleus of the solitary tract blocks the effects of systematically administered catecholamines on memory.

Glucocorticoids released by activation of the hypothalamic–pituitary–adrenal axis in stress also have effects on learning and memory. These hormones readily cross the blood–brain barrier to act on steroid receptors that are located in high density in the amygdala and hippocampus. Low doses of glucocorticoids enhance, while high doses (or the chronic exposure in long-term stress) impair memory. Low concentrations occupy the high-affinity mineralocorticoid receptors (MR) and this facilitates strengthening of synapses thought to be crucial for learning. In contrast, high glucocorticoid concentrations fully saturate the low-affinity glucocorticoid receptors (GR) and this blocks the synaptic strengthening necessary for learning.

The anterior pituitary corticotrophs manufacture, from a single precursor, adrenocorticotrophic hormone (ACTH) and the opioid peptide, β endorphin, both of which impair learning by direct action on the CNS. Enkephalins, also opioid peptides, are co-released from the adrenal medulla along with catecholamines and impair memory by a peripheral action. Naloxone, an antagonist of opioid receptors facilitates memory.

Cholinergic enhancement of memory is well documented. Muscarinic receptor antagonists impair memory, while inhibitors of acetylcholinesterase improve it. Acetylcholine modulation of memory is mediated by the septohippocampal

pathway and the cholinergic nuclei of the basal forebrain. The amygdala may enhance consolidation by activating the cholinergic attentional system in the basal forebrain. At least part of the catastrophic impairment of learning and memory that occurs in Alzheimer's disease is probably due to the loss of cholinergic neurons in the forebrain.

Memory and sleep

The memory hypothesis of sleep postulates that during sleep adjustments to synaptic strengths are made in the hippocampus and neocortex so that consolidation and reorganization of memories occurs. The hypothesis does not claim that sleep is needed for consolidation of all memories; indeed substantial memory consolidation happens during normal waking. Most studies have concentration on REM sleep but more recent work has also implicated slow-wave sleep. The evidence for the memory hypothesis is of four general types:

- Learning results in an increase in the duration of REM sleep in the following night.
- Performance in memory tasks is enhanced by a period of sleep within a few hours of the original learning.
- Sleep deprivation impairs memory.
- Physiological studies imply that memory processing occurs during sleep.

The first three of these approaches imply that REM sleep is involved in procedural memory, whereas slow-wave sleep may have a role in declarative (explicit) memory. Some of the strongest evidence for human learning being sleep-dependent comes from a visual texture discrimination task. Improvement on this is not seen until after post-training sleep, and sleep deprivation on the night after training completely eliminates the effect of training, even after two subsequent nights of recovery sleep. Selective sleep deprivation shows that this task seems to require both slow-wave and REM sleep for consolidation.

The enhanced recall of emotionally salient memories after particularly REM-rich sleep in humans suggests that REM sleep is also important for emotional memory. In this light it is intriguing that REM sleep is increased in stress generally and specifically in major depression, bereavement and post-traumatic stress disorder. PET studies show that during REM sleep while brain activity is reduced in the dorsolateral prefrontal cortex (involved in planning and problem solving), it is increased in the limbic cortex and the amygdala.

Evidence for memory processing during sleep is quite persuasive. Human volunteers trained on a complicated eye–hand coordination task before they went to sleep showed increased EEG slow-wave activity in the right parietal cortex (the region involved in the original learning) during the post-training NREM sleep, and moreover the amount of slow-wave activity predicted their subsequent improvement in performance of the task. PET scans show that the same brain regions that are active when learning a task are also active during subsequent REM sleep. In rats, post-training REM shows a greater number of PGO spikes and there is a correlation between spike activity and improvement in task performance. One idea about how this works is that although cortical networks are off-line during REM sleep they are excited periodically by the PGO spikes, which act as random noise. Any activity generated by this will not be specific memories but random associations and the synaptic links which allow them are weakened. In this view REM sleep is needed to prevent neural networks from becoming saturated.

In rats undergoing spatial navigation learning the firing patterns seen in place

cells in the CA1 region of the hippocampus (see *Fig. 1, Topic O3*) during waking periods are replayed during slow-wave sleep about 30 minutes after training and during REM sleep about 24 hours later. This is interpreted as a playback mode needed for consolidation. Theta activity generated in the septum is seen in the hippocampus during the REM replay, just as it is during the original learning. Place cells responding to novel aspects of the environment fire in phase with the θ waves (during the peaks) and over 4–7 days these same cells drift out of phase so they end up firing during the troughs. This is the same time course that it takes for hippocampal-dependent learning to become independent of the hippocampus. *In vitro* studies which record the activity of neurons in hippocampal slices in response to different stimulation protocols show that synapses are strengthened between cells that fire in phase with the θ wave, but weakened by out-of-phase firing. Hence at first specific hippocampal synapses increase in strength in response to the novel environment, but decrease as the memory is transferred to the neocortex and eradicated from the hippocampus.

In the rat hippocampus θ waves are seen during the waking state when the animal is actively exploring and during REM sleep. Theta waves reflect activity in the cholinergic septohippocampal pathway. In addition, in REM sleep, cholinergic neurons in the basal forebrain that project to the neocortex are highly active. Acetylcholine increases the excitability of pyramidal cells in cortex and hippocampus. This cholinergic modulation is thought to facilitate the traffic of information from the cortex to the hippocampus *(Fig. 4)*. In contrast, **sharp waves** are initiated in the CA3 region and spread through CA1 to the entorhinal cortex during slow-wave sleep. This is believed to allow hippocampal information to be played back via the entorhinal cortex to the neocortex. During the early part of the night information flow is mostly in the hippocampus to cortex direction because that is when most of the slow-wave sleep happens. Presumably this is when memories are transferred out of the hippocampus to the cortex. Firing of neocortical neurons at about 10 Hz during the long-lasting depolarizations of slow-wave sleep activity probably induces long-term potentiation that allows cortical connections to be re-specified. Later in the night, as the proportion of REM sleep rises, the traffic goes mostly from cortex to hippocampus. During this time new memories are consolidated in the hippocampus while old ones are erased. One problem with the memory hypothesis of sleep is that monoamine oxidase inhibitors used for the treatment of depression severely reduce the proportion of REM sleep. However, there is no

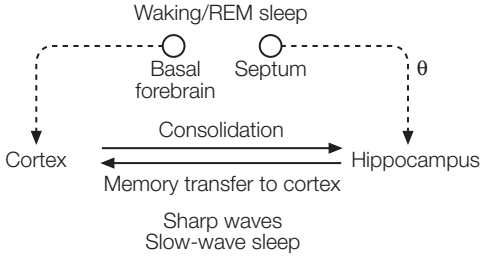


Fig. 4. Putative roles of sleep in memory processing ○--> facilitatory cholinergic pathways

evidence that patients treated with these agents have any memory deficits even with long-term use.

REM sleep is the period when the most intense dreams are reported in individuals who are woken. Dreams generally bear rather little relation to waking experiences and very few of them resemble episodic memories. On the other hand many have significant emotional content. These characteristics seem to reflect the pattern of activity in REM sleep, namely little information flow from hippocampus to cortex, low activity in the prefrontal cortex, and high activity in the limbic cortex.

03 Cell physiology of learning

Key Notes

Hippocampal circuitry

Excitatory input to the hippocampus comes from the entorhinal cortex by way of the perforant pathway, axons of which synapse with granule cells of the dentate gyrus. Granule cell mossy fiber axons synapse with pyramidal cells in CA3. Branches of CA3 cell axons go to three possible destinations, the contralateral hippocampus, the hypothalamus, or to CA1 pyramidal cells via recurrent collaterals (Schaffer collaterals). Axons of CA1 cells go to the entorhinal cortex. All these principal neurons use glutamate and are excitatory. In addition, the hippocampus contains GABAergic inhibitory interneurons. Modulatory neurons using acetylcholine, noradrenaline (norepinephrine) or serotonin also provide input.

Long-term potentiation (LTP)

Learning is thought to be due to changes in the strength (weight) of synapses. Hebb's rule proposes that a synapse between two neurons is strengthened when the neurons are activated together. Synapses that obey this rule are said to be Hebbian. LTP is a long-term increase in the strength of synapses and may be associative (Hebbian) or non-associative. Associative LTP has been extensively studied at synapses between CA3 and CA1 cells in the hippocampus, where it is produced by applying a high-frequency (tetanic) stimulus to the CA3 axons. Subsequent single stimuli elicit a larger excitatory postsynaptic potential in the CA1 cell than before the tetanic stimulus. LTP can last for many hours in brain slices.

Cellular physiology of associative LTP

There are three phases to LTP, induction, expression and maintenance. Induction depends on the activation of NMDA receptors by the tetanic stimulation which causes both increased glutamate release from CA3 axons and a large depolarization of the CA1 cell. This represents coactivation of pre- and postsynaptic cells (satisfying Hebb's rule) and fulfills the conditions for opening NMDA receptors. It is the Ca^{2+} entry through these glutamate receptors that is one of the necessary requirements for LTP induction. Expression of LTP involves second messenger mediated increases in the sensitivity of AMPA glutamate receptors, acquisition of AMPA receptors by previously silent synapses or perforation of synapses to give new active sites. In addition, an increase in presynaptic glutamate release may occur, triggered by a retrograde messenger liberated from the postsynaptic cell to affect the presynaptic terminal. Maintenance of LTP beyond two hours requires transcription and translation to synthesize proteins that allow the synaptic modification to persist in the face of continual turnover of molecules.

LTP and learning

The optimal protocol for generating LTP is essentially identical to theta (θ) activity seen in rats learning a spatial task. The θ rhythm is caused by regular firing of hippocampal neurons driven by the cholinergic pathway from the septum. LTP *in vivo* may last for many months. Pharmacological or genetic engineering manipulations which impair LTP often produce deficits in spatial learning.

CREB and long-term memory

The cyclic AMP second messenger system is able to modify gene expression. The activated catalytic subunit of protein kinase A translocates to the nucleus where it phosphorylates a transcription factor that binds to cAMP response elements (CREs) in genes regulated by cAMP. The transcription factor is termed cAMP response element binding protein (CREB). When phosphorylated CREB binds to CREs, initiating transcription. This process can be inhibited by proteins related to CREBs called cAMP response element modulators. CREB is implicated in long-term memory in a number of different learning models; procedural learning in a marine snail, odor discrimination learning in the fruit-fly, and spatial navigation learning and fear conditioning in mice.

The Marr–Albers–Ito model

Motor learning in the cerebellum comes about by a weakening of the strength of synapses between parallel fibers (pf) and Purkinje cells (PC) that are active at the same time as error signals arrive at the PC via climbing fibers. The synaptic weakening is long-term depression (LTD).

Classical conditioning of the eye blink reflex

A puff of air delivered to the eye normally causes a reflex eye blink. This can be classically conditioned by pairing a tone with the air puff. The tone activates pf–PC synapses just before the air puff signal arrives at the PC via the climbing fibers, and so the pf–PC synapses suffer LTD. Subsequent occurrence of the tone causes reduced PC excitation which translates into a larger cerebellar output to motor neurons driving the eye blink.

Long-term depression (LTD)

LTD is seen in the cortex and hippocampus as well as the cerebellum. In the cerebellar cortex, LTD requires simultaneous Ca^{2+} input into the PC (caused by climbing fibers) and activation of glutamate receptors at pf–PC synapses. The effect is to desensitize the AMPA receptors at the synapses.

Related topics

Ionotropic receptors (D1)
Cerebellar cortical circuitry (K5)
Types of learning (O1)
Physiological psychology of memory (O2)

Hippocampal circuitry

The hippocampal formation is folded archaecortex (ancient cortex) consisting of the **dentate gyrus** and the **cornu ammonis** (**CA**) – collectively termed the **hippocampus** – plus the subiculum. The cortex of the dentate gyrus and CA have three layers, while the subiculum is transitional cortex between the hippocampus proper and the six-layered neocortex of the **entorhinal** area. A major input to the hippocampus from the entorhinal cortex comes via the **perforant pathway**, axons of which synapse with **granule cells** of the dentate

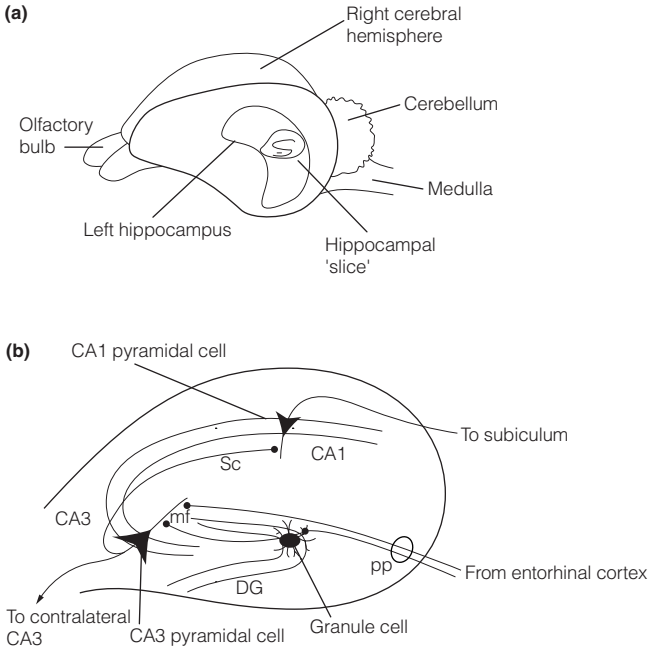


Fig. 1. Hippocampus. (a) Location of the left hippocampus in rat brain; a hippocampal slice is at right angles to the long axis of the hippocampus. (b) Structure of a hippocampal slice showing the principal excitatory neurons. From Revest, P. and Longstaff, A. (1998) Molecular Neuroscience. *© BIOS Scientific Publishers, Oxford. DG, dentate gyrus; pp, perforant pathway; Sc, Schaffer collateral; mf, mossy fibers.*

gyrus or pyramidal cells in the CA3 region of the CA (*Fig. 1*). Axons of the granule cells (mossy fibers) also synapse with CA3 pyramidal cells.

The CA3 pyramidal cell axons branch, forming:

- **commissural fibers** which pass to the opposite hippocampus;
- efferents which leave the hippocampus via the **fornix** to terminate largely in the hypothalamus or thalamus;
- collaterals which turn back to form synapses on the same and neighboring CA3 cells (**recurrent collaterals**), or which synapse with cells in the CA1 region of the CA (**Schaffer collaterals**).

CA1 cell axons go to the subiculum and entorhinal cortex. The perforant pathway, granule and pyramidal cells are glutamatergic and excitatory. The hippocampus also harbors inhibitory interneurons that are GABAergic. Other inputs to the hippocampus include a cholinergic pathway from the septum and noradrenergic and serotinergic axons from the brainstem reticular system. These inputs are modulatory.

Long-term potentiation (LTP)

A core idea of contemporary neuroscience is that learning occurs by changes in the strength of synapses. A mechanism to account for how this might occur was proposed in 1949 and is called **Hebb's rule**. This states that all synapses between two neurons become stronger if both of the neurons are activated at the

same time. Synapses which show this type of plasticity are said to be **Hebbian**, and can mediate associative learning since they act as coincidence detectors that associate firing of the presynaptic *and* postsynaptic cell. Hebb's rule is summarized by the aphorism 'what fires together, wires together'. Several mechanisms to bring about synaptic modifications are now known which either increase synaptic weighting, **long-term potentiation (LTP)**; or decrease synaptic weighting, **long-term depression (LTD)**. Both occur in the hippocampus, LTP is also seen in the neocortex, amygdala and at other sites in the nervous system, while LTD also occurs in the cerebellum and spinal cord. LTP and LTD are regarded as cellular substrates of learning.

LTP can be either associative (Hebbian) or non-associative. LTP at the synapses between CA3 Schaffer collaterals (Scs) and CA1 cells in the hippocampus is Hebbian. It can be studied in hippocampal brain slices by intracellular or extracellular recording from CA1 neurons whilst electrically stimulating a bundle of Scs (*Fig. 2*). In response to brief, low-frequency stimulation of the Scs, the CA1 cells show a brief epsp due to glutamate release. If a brief tetanic burst of high-frequency stimulation is given (typically 100 Hz for 0.5 s), subsequent low-frequency pulses now elicit a larger epsp. This is LTP, it may last for as long as the brain slice survives (many hours) and it has three properties:

1. **Input specificity**. Delivery of low-frequency stimuli to the CA1 cell via a different untetanized bundle of Scs does not elicit the enhanced epsp.
2. **Cooperativity**. The probability of producing LTP increases with the number of *afferent* fibers (Scs) tetanically stimulated. While weak (i.e. low-current) high-frequency stimuli often fail to generate LTP, because they excite only a few afferents, strong tetanic stimuli are successful because they recruit many afferents.
3. **Associativity**. A given CA1 cell receives Scs from CA3 cells on the same side and commissural axons that come from CA3 cells in the contralateral hippocampus. A weak tetanic stimulus to either pathway that fails to generate LTP will do so if it is paired with strong tetanic stimulus in the other pathway.

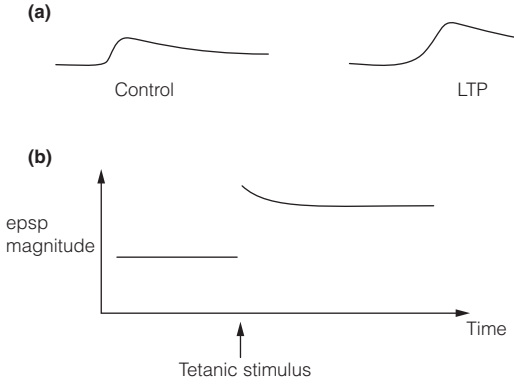


Fig. 2. LTP in a hippocampal slice. (a) Excitatory postsynaptic potentials (epsps) recorded from CA1 pyramidal cells before (control) and after tetanic stimulus (LTP). (b) Epsp magnitude remains elevated over several hours.

Cellular physiology of associative LTP

CA1 cells have AMPA and NMDA glutamate receptors. In order to be activated, NMDA receptors must bind glutamate *and* experience a depolarization big enough to remove Mg^{2+} ions from the channel (voltage-dependent blockade). This condition is not provided by low-frequency Sc stimulation. The amount of glutamate released is low, few AMPA receptors are activated and the resulting epsp is too small to open NMDA receptors. However, high-frequency stimulation opens numerous AMPA receptors and so depolarizes the cell sufficiently to activate NMDA receptors. Cooperativity arises because the more afferents activated the greater the depolarization of CA1 cells. Associativity is a property of the NMDA receptor itself: to activate it requires the coincidence of presynaptic glutamate release and postsynaptic neuron depolarization (Hebb's rule).

In fact, tetanic stimulation is not mandatory for LTP; any manipulation that depolarizes CA1 cells enough, such as antagonizing the inhibitory actions of GABA, will allow LTP to an excitatory input. It is Ca^{2+} entry through NMDA receptors that triggers the induction of LTP. Antagonists of NMDA receptors, for example the competitive antagonist, D-2-amino-5-phosphonovalerate, or the open channel blocker, dizocilpine, prevent the induction of LTP. Metabotropic glutamate receptors are also necessary for LTP induction, but only on the first occasion in which the synapse is tetanically stimulated; metabotropic glutamate receptors are said to act as 'molecular switches' for LTP.

Both presynaptic and postsynaptic sites are implicated in the expression of LTP, but the specifics depend very much on the type of LTP and the synapses involved. At the CA3–CA1 synapses postsynaptic changes include:

1. An increase in responsiveness of the postsynaptic membrane as AMPA receptors become more sensitive to glutamate. The biochemistry of this has been worked out. **Calcium-calmodulin dependent protein kinase II** (**CaMKII**), a major protein of the postsynaptic density, is thought to be activated by Ca^{2+} entry through NMDA receptors, and it then phosphorylates AMPA receptors (GluR1 subunits), enhancing their response to glutamate.
2. Activation of **silent synapses**. These are synapses that harbor only NMDA receptors. They are normally silent because low-frequency stimulation does not activate NMDA receptors). Calcium entry during LTP induction causes them to acquire AMPA receptors and so become responsive to low-frequency input.
3. The formation of new synapses. Synapses perforate and eventually form two synapses where previously there was one.

A presynaptic component to CA3–CA1 synapse LTP has been proposed in which there is increased glutamate release. This requires that the NMDA Ca^{2+} signal generates a **retrograde messenger** in the postsynaptic cell that travels the 'wrong' way across the synaptic cleft. One molecule proposed for this role is **nitric oxide** (**NO**). In neurons NO is synthesized by a Ca^{2+}-dependent **nitric oxide synthase** (**NOS**). As a small, freely diffusible molecule, NO rapidly diffuses out of the postsynaptic cell, across the cleft and into the presynaptic cell where it stimulates guanylyl cyclase, thereby enhancing the probability of glutamate release.

With brief tetanic stimulation, LTP lasts for only about 2 hours. However, with several tetanic stimuli, LTP seems to last indefinitely. This long-lasting LTP depends on transcription and translation since it is blocked by drugs that inhibit mRNA or protein synthesis. Several mechanisms probably have a role in this **maintenance** of LTP. They must explain how synaptic alterations are retained,

even while individual molecules are being turned over, long after the original signals that brought about the alterations have gone. Two well-documented processes are as follows:

1. CaMKII consists of four subunits. Ca^{2+} activation phosphorylates them and once the Ca^{2+} concentration has fallen to resting levels they remain phosphorylated. This is because if a subunit becomes dephosphorylated it will immediately become autonomously phosphorylated by one of the other subunits. In this way CaMKII remains persistently active.
2. Ca^{2+} stimulates an isoform of adenylyl cyclase and consequently cAMP concentrations increase in LTP. Protein kinase A becomes persistently activated and has effects on gene expression.

Key events in LTP are summarized diagrammatically in *Fig. 3*. Somewhat different mechanisms are responsible for non-associative LTP that occur at some other synapses.

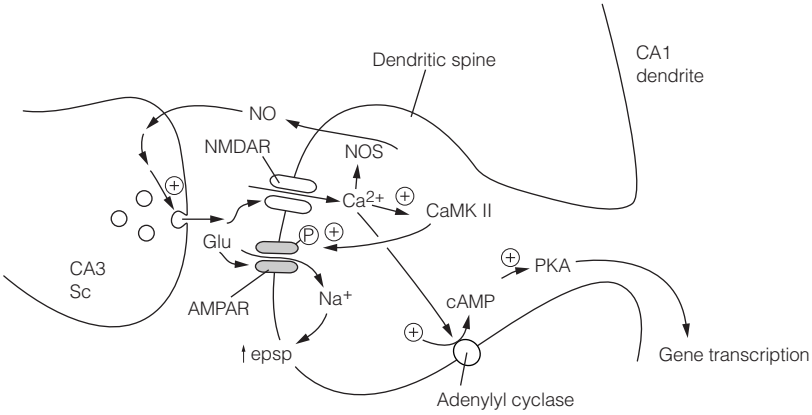


Fig. 3. Key events in LTP. Sc, Schaffer collateral; glu, glutamate; AMPAR, AMPA receptor; NMDAR, NMDA receptor; NO, nitric oxide; NOS, nitric oxide synthase; PKA, protein kinase A; CaMKII, Ca^{2+}–calmodulin-dependent kinase II.

LTP and learning

Two sorts of evidence support the view that LTP is learning at the level of individual neurons. Firstly, tetanic stimulation arises physiologically, secondly, manipulations that impair LTP cause learning deficits and *vice versa*. When rats explore a maze, the EEG shows a theta (θ) rhythm with a frequency of 4–10 Hz. This reflects periodic firing of hippocampal neurons that is driven by the cholinergic septohippocampal pathway. During θ discharge, place cells fire in phase, but all other pyramidal cells are silenced by the increased discharge of inhibitory cells. Thus θ activity ensures that only cells involved in learning a particular environment are active. Moreover, one of the most effective stimulus protocols for producing LTP *in vitro* is just like θ discharge. Hence θ rhythm may be the brain's 'natural' tetanic stimulus for learning. When LTP is generated in intact behaving rats via chronically implanted electrodes it may last for many months.

Preventing NMDA receptors from functioning not only blocks induction of LTP, it also prevents some types of learning. Blocking NMDA receptors with APV or dizocilpine impairs learning of the Morris water maze by rats. Mice

genetically engineered so that NMDA receptor subunits are deleted from the CA1 region in the third postnatal week (after hippocampal development is complete) show no NMDA-receptor-dependent LTP in CA1, have impaired spatial learning in the Morris water maze, and show a loss of the correlated firing of place cells that is seen in normal animals.

CREB and long-term memory

The cyclic AMP second messenger system is able to modify gene expression. The activated catalytic subunit of protein kinase A translocates to the nucleus where it phosphorylates a transcription factor that binds to cAMP response elements (CREs) in upstream regions of genes regulated by cAMP. The transcription factor is termed **cAMP response element binding protein** (**CREB**). When it is phosphorylated CREB binds to CRE and this engages the other components needed for transcription (*Fig. 4*). CREB can also be phosphorylated by calcium-calmodulin-dependent kinases (e.g. CaMKIV) which will be activated by Ca^{2+} entry via NMDA receptors, or any other mechanism that increases cytoplasmic calcium concentration. CREB-mediated gene transcription is controlled in at least two ways:

- by protein phosphatases that are activated by the calcium-dependent protein calcineurin;
- by repressor proteins termed **cAMP response element modulators** (**CREMs**) that bind to CREs and so prevent CREBs from doing so.

Considerable evidence suggests that CREB is implicated in long-term memory in a number of different learning models; procedural learning in a marine snail (*Aplysia*), odor discrimination learning in the fruit-fly *Drosophila* and several types of learning in mammals. For example, transgenic mice that lack two of the three CREB isoforms are impaired in three different tasks that depend on the hippocampus (including the Morris water maze) and fear conditioning that requires the amygdala. Interestingly, long-lasting LTP is impaired by protein kinase A inhibitors and in the CREB-deficient mice.

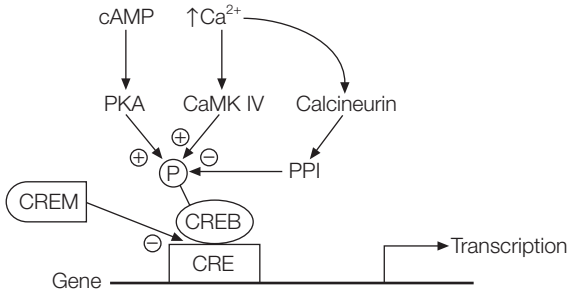


Fig. 4. Cyclic AMP response element binding protein (CREB) modulation of gene transcription. CRE, cyclic AMP response element; CaMK IV, calcium-calmodulin-dependent kinase IV; PKA, protein kinase A; PPI, protein phosphatase I.

The Marr–Albers–Ito model

Motor learning in the cerebellum involves alterations in the strengths of synapses between parallel fibers (pf) and Purkinje cells (PC). Those synapses that are active at exactly the same time that there is climbing fiber input to the Purkinje cell, experience a reduction in the synaptic strength, a type of plasticity called **long-term depression (LTD)**.

In the **Marr–Albers–Ito model** of motor learning, the frontal cortex (via the corticopontine cerebellar tract) provides the mossy fiber–parallel fiber inputs and the climbing fibers from the inferior olive are thought to transmit error signals. All the pf–PC synapses that happen to be activated by a pattern of mossy fiber inputs at the same time as climbing fiber error signals arrive will show LTD. Synapses not concurrently active are unchanged. Subsequently, parallel fiber activity at the depressed synapses excites Purkinje cells less, thereby reducing their inhibitory output on deep cerebellar nuclei. The overall effect is that synapses at which LTD occurs enhance cerebellar output.

Classical conditioning of the eye blink reflex

Motor learning that occurs during the classical conditioning of the eye blink reflex has been extensively studied. In the eye blink reflex, a puff of air delivered to the eye (US) will produce an eye blink (UR). The eye blink reflex can be conditioned if the air puff is paired with a tone (CS). The circuitry involved in motor learning in this reflex is shown in *Fig. 5*.

The air puff (US) is sensed by neurons in the spinal nucleus of the trigeminal (5th) cranial nerve. The eye blink reflex (UR) is executed by connections between these cells and motor neurons in the facial (7th cranial) nerve. Conditioning of the reflex requires the cerebellum. The US signal is transmitted via climbing fibers that arise from the inferior olivary nucleus. The tone (CS) signal goes by way of the ventral cochlear nucleus and pontine nucleus arriving at the cerebellum in mossy fibers. Activation of the pf–PC synapse by the CS, 250 ms before the arrival of the US via the climbing fiber, results in LTD of the

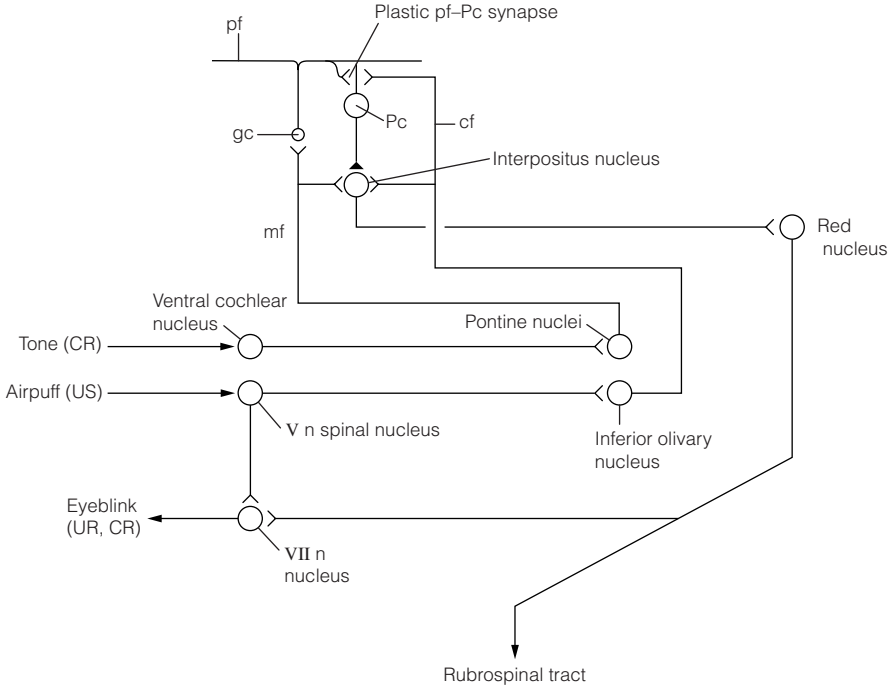


Fig. 5. Motor learning in the cerebellum; circuitry implicated in the conditioned eye blink reflex. pf, Parallel fiber; Pc, Purkinje cell; CR, conditioned response; UR, unconditioned response; US, unconditioned stimulus; cf, climbing fiber; gc, granule cell; mf, mossy fiber.

pf–PC synapse. The effect of the LTD is that any subsequent arrival of the CS produces a smaller excitation of the PC. Hence PC inhibition of the interpositus neurons is diminished, so these cerebellar nucleus cells drive the eye blink via their connections with the red nucleus.

Long-term depression (LTD)

Long-term depression (LTD) is seen in the hippocampus and cerebral cortex where it can occur alongside LTP, and in the cerebellum (in which LTP is never seen). Induction of LTD in the cerebellum at the pf–PC synapse requires coincident Ca^{2+} influx into the Purkinje cell and activation of AMPA and metabotropic glutamate receptors at the synapse. The Ca^{2+} influx is provided by the large depolarization due to climbing fiber activity, which opens P-type voltage-dependent Ca^{2+} channels. The receptors are activated by the release of glutamate from the parallel fibers. The final cause of the synaptic depression is desensitization of the AMPA receptors (*Fig. 6*) brought about by their phosphorylation by protein kinase C and possibly also by protein kinase G activated as a result of nitric oxide synthesis.

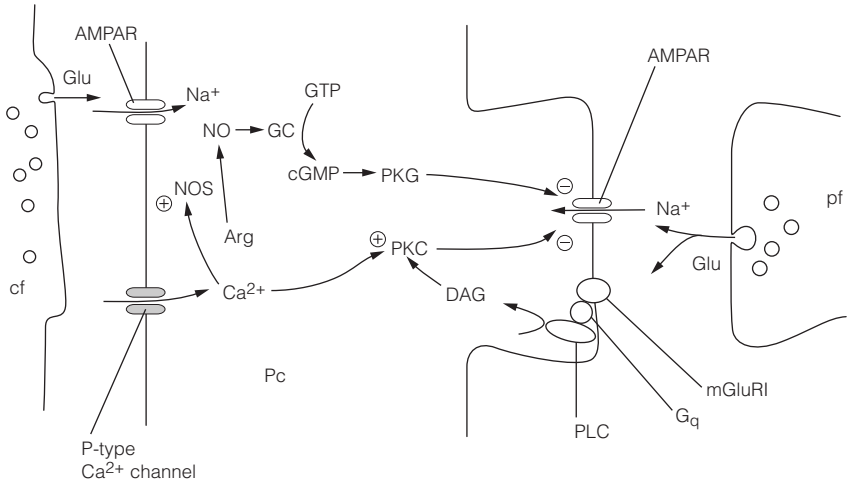


Fig. 6. Cellular events in cerebellar LTD. glu, Glutamate; AMPAR, AMPA receptor; mGluR1, type 1 metabotropic glutamate receptor; G_q, G protein; NO, nitric oxide; NOS, nitric oxide synthase; PKG, protein kinase G; PKC, protein kinase C; GC, guanylyl cyclase; PLC, phospholipase C; DAG, diacylglycerol; Pc, Purkinje cell; pf, parallel fiber; cf, climbing fiber.

04 Arousal and attention

Key Notes

Attention and its link to arousal

Attention is the selective filtering of sensory input (sensory attention) or thoughts (executive attention) so that only a fraction is subject to conscious perception or manipulation. A necessary condition for attention is a general state of arousal. This is generated by the diffuse noradrenergic system. Sensory attention studies show that it is possible to respond to stimuli that are not being attended to and that attention is not continuous but frequently and briefly interrupted.

Arousal

Noradrenergic neurons in the locus ceruleus fire in a manner that correlates with an animal's level of arousal and in response to the appearance of stimuli they have been trained to attend to. Higher firing rates are seen during visual scanning (labile attention) than during highly focused attention. The effect of noradrenaline (norepinephrine) secreted throughout the brain is to enhance the response of neurons to specific excitatory and inhibitory inputs.

Visual attention

Visual attention can be overt – orienting the eyes towards a stimulus – or covert, in which the internal state is altered to attend to a stimulus. Neural activity associated with covert attention can be averaged as event-related potentials in the EEG, recorded over many stimulus presentations. How easy it is to attend to and thus select a target stimulus from a sea of distractors depends on how similar the target is to the distractors, what sensory dimensions (e.g. color, movement) the target and distractors have, and how many of each there are. These constraints on the ease of attention are imposed by the receptive field (RF) properties of visual neurons.

Neuroanatomy of visual attention

Only cells with large RFs have an attentional role and these are found in areas involved in late visual processing. Brain scans imply that the superior parietal lobe, part of the 'where' stream of visual processing is concerned with attention. Damage to this region causes neglect syndrome that is interpreted as an attention deficit. Non-cortical areas implicated in visual attention include the superior colliculus and pulvinar. The pulvinar gets information about eye movements from the superior colliculus and motion input from the cortex, and filters out retinal signals produced by eye movements. Additionally the pulvinar has a part in selective attention in which target objects must be identified when surrounded by distractors.

Related topics

Noradrenaline (norepinephrine) (D5)

Parallel processing in the visual system (G7)

Attention and its link to arousal

Attention is the way in which sensory input is selectively filtered so that only a fraction is consciously perceived at any time (such as concentrating on one voice at a noisy party) or in which specific mental representations of objects and ideas are selected for conscious manipulation (i.e. control of thought processes).

A prerequisite for attention is maintenance of an appropriate state of alertness or **arousal**. This is achieved by activity in the diffuse noradrenergic system that projects from the brainstem into the forebrain, but since this operates globally (i.e. across most of the brain) attending to particular objects (either external or internal) must use more specific circuits.

Sensory attention has been most extensively investigated for vision and hearing. Psychology studies show that people can behave in response to visual or auditory stimuli to which they are not attending and for which they have no conscious awareness. Attention is not continuous but frequently interrupted by short lapses termed **Bills' blocks**, which is interpreted as being due to multitasking by the circuitry involved. Normal individuals have the subjective impression that they exert intentional control over their thought processes. An attention system involving the anterior cingulate and prefrontal cortex is postulated to be responsible for this **executive attention**.

Arousal

The necessary brain state for attention is regulated by noradrenergic neurons in the locus ceruleus (LC) and other hindbrain nuclei, axons of which ascend to widely distributed parts of the forebrain. Activity of these cells causes a globally synchronized release of NA throughout much of the brain which acts to modulate neuron responses to input from other transmitters. Recording from the LC in behaving animals shows that the firing rate is low during sleep and increases with the level of arousal. In wakefulness the frequency of firing rises when animals switch from low vigilance behavior (i.e. grooming) to orienting towards a stimulus. LC neurons will fire in response to the presentation of stimuli that animals have been trained to pay attention to. The firing is not related to the sensory attributes or motor requirements of the task.

Optimal performance of tasks in which animals are required to attend to stimuli is associated with intermediate firing frequency of LC neurons and is thought to represent highly focused attention. The highest frequencies are associated with a greater number of errors in discriminating between target objects (to which animals have been trained to attend) and distractors (which the animals should ignore). This is labile attention and is seen when animals are scanning their surroundings.

The neuromodulatory effect of NA on cells in the cerebral or cerebellar cortex is to augment the effects of excitatory (glutamatergic) or inhibitory (GABAergic) inputs relative to the basal firing rate. This is interpreted as an improvement in the **signal-to-noise ratio** of neurons by noradrenergic inputs.

Visual attention

Visual stimuli can be attended to by moving the head or eyes (**orientation**) but attention to any stimuli can also be made without moving, by alterations to the internal state (**covert attention**). Covert visual attention shifts independently of eye movements. This is shown by spatial cueing, a technique that can be used in both humans and monkeys. The subject is required to press a button whenever a stimulus appears on either the left or right side of a computer monitor. Before the stimulus is presented, the subject is cued (by some visual signal) as to which side the stimulus is likely to appear. Subjects respond faster when the stimulus appears on the cued side compared with the uncued side, even though the eyes

do not move. The difference is attributed to a covert shift of attention to the cued side. Averaging the electrical activity from the brain (recorded via scalp electrodes) over many stimulus presentations gives characteristic waveforms called **event-related potentials** (**ERPs**). ERP components occurring after about 80 ms are altered in size for stimuli presented to the cued side compared with the uncued side. The difference is attributed to neural activity needed for visual attention. While saccades follow about 200 ms after stimulus presentation, attentional shifts can occur as early as 50 ms, so they must be covert.

The ease with which a target object, selected for attention, can be singled out from among a sea of distractors, depends on the sensory attributes of target and distractor objects and their numbers. This has been investigated in visual attention by measuring the time taken to search for particular targets displayed on a computer monitor, along with distractors. In general, the ability to discriminate target from distractor deteriorates (in other words, attention becomes less reliable) the more similar the target and distractor. A single visual target that differs in color or motion from the distractors is identified in a time that does not depend on the number of distractors. If there are a number of such targets displayed, the time taken to find them depends on how many there are. However, in more complicated cases, such as trying to locate a red spot moving to the left in an array of blue, green, red and yellow spots all moving in random directions, the time taken rises with the number of distractors.

These limits to visual attention are thought to arise from the nature of the receptive fields (RFs) of visual neurons. For example, target objects that differ in apparent depth and movement from the surrounding distractors 'pop out' because there are visual system neurons that encode both binocular disparity *and* movement. In contrast, cells that respond both to movement and color have never been identified and the visual system is presumably forced to analyze each object in turn, for each of these stimulus features, in order to identify the one with the target combination; this takes time.

Neuroanatomy of visual attention

Visual attention is accompanied by alterations in the receptive field properties of some visual neurons. During attentional tasks these cells become transiently biased to ignore distractor stimuli that would, in the non-attending state, elicit a response. It is as if in attention the RFs shrink to encompass only the target stimulus, so that the cell responds to it alone. This is essentially a raising of the signal-to-noise ratio. Not all visual system neurons are modulated in visual attention. Only cells with large enough RFs to respond to several objects (target plus distractor) are involved. These are generally not found in regions responsible for early processing (V1, V2) but predominate in late-processing visual areas (e.g. inferotemporal and medial temporal cortex). In monkeys trained to attend to a particular region of the visual field in a spatial cueing task the baseline firing rate of neurons serving the region (except those in V1) increased before the presentation of the stimulus. This is thought to be due to an attention signal.

The parietal cortex has long been thought to play a part in visual attention and is a major component of the 'where' stream of visual processing. PET and fMRI studies show bilateral activation of the superior parietal lobe when attending to an object in the right visual field, but activation of only the right parietal lobe when attending to the left visual field. This is an example of lateralization of brain function; the right hemisphere predominates in visual attention tasks. These brain scans also showed activation of the thalamus contralateral to

the side attended to. Unilateral parietal lobe lesions cause difficulty in switching attention from objects in the ipsilateral visual field to those in the contralateral field. This is a component of the **neglect syndrome** which commonly follows a stroke affecting the right posterior parietal lobe. Afflicted patients ignore objects, people, even parts of their own body, to one side of their center of gaze. Neglect syndrome is thought to be an attention deficit not only for visual stimuli but also for other modalities (e.g. somatosensory).

Several non-cortical brain regions contribute to visual attention. The superior colliculus and **pulvinar** (a thalamic nucleus having extensive connections with visual cortical areas) are implicated in generating saccades to visual targets, directing gaze at interesting (salient) objects. A further role of the pulvinar in visual attention is to filter out irrelevant information. Many pulvinar neurons fire in response to the motion of stimuli in the real world but are silent for the apparent motion of a stimulus generated by eye movements. These neurons compare input from the superior colliculus about eye movements with motion input from the visual cortex and are thought to filter out retinal signals produced by eye movements. In addition the pulvinar has a role in selective attention. Monkeys given unilateral microinjections of GABA agonist into the pulvinar had difficulty in shifting their attention to objects in the contralateral visual field in a spatial cueing task. Precisely the same problem has been reported in three patients with thalamic damage. Further support for the pulvinar in attention comes from PET studies. Normal humans were asked to perform a task requiring identification of objects that were either presented alone or with distractors. Increased activity was seen in the contralateral pulvinar only when the task included the distractors.

05 Language

Key Notes

Development of language

It is thought that children have neural circuits dedicated to learning language. The errors that children make when learning to speak shows that they are not just mimicking the sounds they hear but are learning their local grammar. Language acquisition and other types of semantic learning are separate. There is a critical period for learning language which ends by about 7 years old. Before this time brain regions for learning other languages overlap extensively with those for the native language. Later in life brain regions involved in learning a second language do not overlap those for the first.

Lateralization of language

Language functions are lateralized, with the left hemisphere being dominant in 90% of people, and the right hemisphere in 7.5%. Cerebral dominance for language and motor function are independent. Language appears to be rather less lateralized in women than men. Damage to the left hemisphere in infancy produces mild defects in language acquisition whereas the same damage in adults is devastating.

Neuroanatomy of language

The temporal cortex adjacent to the lateral sulcus which includes Wernicke's area is involved in sensory functions associated with speech and reading. The frontal and parietal cortex adjacent to the lateral sulcus is concerned with the motor aspects of reading and writing. It includes Broca's area that is reciprocally interconnected to Wernicke's area. Broca's area is the premotor area for speech, sending its output to the face and tongue area of the neighboring motor cortex. The insula, the floor of the lateral sulcus, plans the articulatory movements needed for speech and projects to Broca's area. Disorders of spoken language are termed aphasias. They generally arise as a result of focal brain lesions following stroke or head trauma and they fall into several types. The right cerebral hemisphere does have some language functions, in particular communicating or interpreting the emotional quality of speech.

Active listening

Brain regions implicated in listening to the spoken word include Wernicke's area which discriminates verbal from non-verbal material, the angular gyrus which identifies phonemes, the middle temporal gyrus which identifies words from phonemes, and the dorsolateral prefrontal cortex required for the meaning of speech.

Reading

The angular gyrus projects through Wernicke's area to area 21 to transform written syllables to phonemes. The meaning and pronunciation of the words represented by the phonemes are retrieved by a pathway that activates the dorsolateral prefrontal cortex. Reading aloud activates Broca's area, the medial supplementary motor area, motor areas subserving face and tongue, and the contralateral cerebellar hemisphere.

Developmental dyslexia

Difficulty in reading, writing and spelling in people of normal intelligence with no sensory deficits is dyslexia. It is probably not a single disorder. Impaired ability to process fast visual stimuli, difficulty decomposing speech into its constituent phonemes, or a defect of the cerebellum which prevents learning of the associations between sounds and the articulations needed to make them, have all been considered as primary causes of dyslexia.

Related topics Brain imaging (A6) Brain sexual differentiation (N7)

Development of language

Initially infants learn to discriminate **phonemes**, the fundamental sound units of their native tongue. Children speak their first words at about 15 months, utter phrases at 2 years, and are speaking fluently with quite low grammatical error rates by 3 years. The nature of the errors shows that children are not just mimicking the sounds they hear but are learning their local grammar, the set of rules (**syntax**) for manipulating words into larger units which communicate meaning (**semantics**).

It is thought that children have neural circuits dedicated to learning language. Evidence for this is that early language acquisition requires no formal training and a number of pathologies show that language acquisition and other types of semantic learning are separate; e.g. some patients with hydrocephalus have compromised general intelligence but can construct completely grammatical speech and have normal speech comprehension.

Children growing up in multilingual households are soon able to discriminate individual languages and become multilingual, speaking each with a perfect accent. There is a critical period for learning language which ends by about 7 years old. Before this time brain imaging shows that the brain loci for learning other languages overlap extensively with those for the native language. Later in life brain regions involved in learning a second language do not overlap those for the first. Evidence for this is that in adults focal lesions can selectively disrupt either the native language or the foreign one while leaving the other relatively intact. Adults learning a foreign language need extensive training and do not acquire a perfect accent in the new tongue even after many years immersion.

Lateralization of language

Language functions are lateralized, with the left hemisphere being dominant in 90% of people, and the right hemisphere in 7.5%. Brain-imaging studies of normal individuals show that language appears to be somewhat less lateralized in women than men. Cerebral dominance for language and motor function are independent; many left-handed individuals have left hemispheric language dominance. Infants that have their left cerebral hemisphere excised to treat neurological disease learn to speak fluently (although they do have deficits in speech comprehension), whereas adults treated in the same way suffer permanent, near total, loss of language. This reflects the huge plasticity of which the infant brain is capable.

Neuroanatomy of language

A considerable volume of brain is involved in language (*Fig. 1*).

1. The temporal cortex adjacent to the lateral sulcus is involved in sensory functions associated with speech and reading. It includes **Wernicke's area** (area

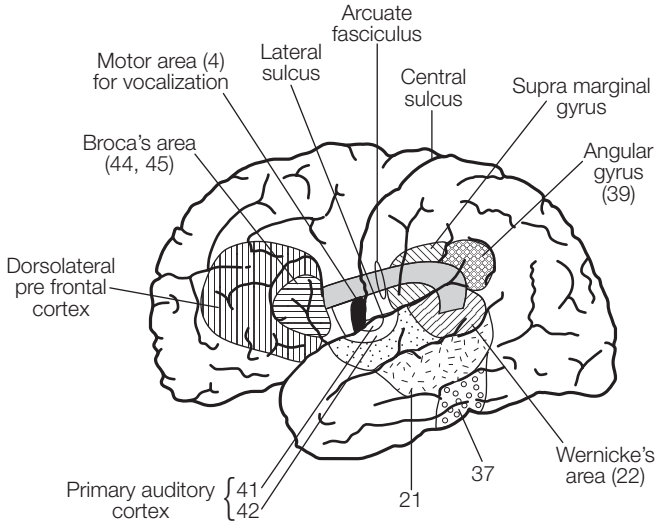


Fig. 1. Regions of the left cerebral cortex involved in language

22) the upper part of which, the **temporal plane**, is larger on the left side in 60% of people. Bordering Wernicke's area lies the **angular gyrus** (area 39), which receives extensive visual input from the extrastriate cortex (area 19) and projects to the temporal plane. Interestingly, abnormal development of the left temporal plane is a feature of schizophrenia.

2. The frontal and parietal cortex adjacent to the lateral sulcus is concerned with the motor aspects of reading and writing, and contains **Broca's area** (Brodmann's areas 44 and 45 of the left inferior frontal gyrus) that is reciprocally interconnected by the **arcuate fasciculus** to Wernicke's area.

Two brain regions, usually located in the left hemisphere, are particularly active during speech:

- Broca's area is the premotor area for speech, sending its output to the face and tongue area of the neighboring primary motor cortex. The **insula**, which forms the floor of the lateral sulcus, is responsible for planning the articulatory movements needed for speech on the basis of verbal working memory input from the prefrontal cortex, and projects to Broca's area.
- Wernicke's area is a sensory area concerned with understanding speech.

Disorders of spoken language are termed **aphasias**. They generally arise as a result of focal brain lesions following stroke or head trauma and they fall into several types.

1. **Motor** (**expressive**) **aphasia** due to a lesion of the left posterior frontal cortex including Broca's area renders speech labored and lacking in fluency and may result in inability to express thoughts in writing.
2. **Sensory** (**receptive**) **aphasia** results from lesions to Wernicke's area and is a deficit in auditory comprehension. Not only does the patient fail to understand the speech of others, they cannot monitor their own speech, which consequently may become unintelligible, albeit fluent. A major sign is frequent errors in which an incorrect phoneme is substituted for the correct one. If the angular gyrus is included in the lesion reading is impaired.

3. **Conduction aphasia** patients have less-severe defects in speech production or auditory comprehension than motor or sensory aphasias but cannot assemble phonemes into strings well. It is caused by lesions which disconnect corticocortical connections, probably including the arcuate fasciculus.
4. **Global aphasia** is caused by large infarcts in the territory of the middle cerebral artery and is essentially a combination of the three aphasias above.

The language non-dominant hemisphere (usually the right) does have some language function, in communicating or interpreting the emotional quality of speech. Lesions of areas 44 or 22 render an individual unable to add emotional color to their speech, or to interpret it in the speech of others, respectively. These are termed **aprosodias**.

Active listening

Brain regions *specifically* implicated in listening to the spoken word (active listening) have been identified on MRI scans by subtracting the signal from regions (such as the auditory cortex) that are engaged when listening to random tones (passive listening) from the total signal produced by listening to speech. Listening to speech activates:

- Wernicke's area on the left side, which is thought to permit discrimination of verbal from non-verbal material;
- the angular gyrus which identifies phonemes;
- the middle temporal gyrus (area 21) and area 37 identify words from phoneme strings and tap into semantic networks located in the left dorsolateral prefrontal cortex (areas 9 and 46), that must be searched to traduce the meaning of speech;
- Broca's area is activated, because when listening to speech we covertly rehearse the articulatory commands needed to pronounce the words, a process referred to as **subvocal articulation**.

Reading

Clearly reading requires visual processing. Subsequently, in novice readers, the parieto-temporal region (angular gyrus and Wernicke's area) dismantles words into phonemes so that they can be identified. However, in experienced readers the extra-striate occipito-temporal cortex (area 19) recognizes entire words instantly. Activation of a network that links the **supramarginal gyrus** (area 40), and area 37, to the anterior part of Broca's area (area 45), via the insula, allows access to semantic networks in the dorsolateral prefrontal cortex so that the meaning and pronunciation of the words can be retrieved.

Finally, either subvocal articulation or reading aloud is accompanied by activation of the whole of Broca's area, the medial supplementary motor area (area 6), motor areas subserving face and tongue (area 4), and the contralateral cerebellar hemisphere.

Developmental dyslexia

Difficulty in reading, writing and spelling in educated people, who match their peers in other intellectual fields and have no apparent sensory deficits, is termed **developmental dyslexia**. Dyslexia is probably a cluster of disorders. It affects some 3–4% of literate populations. Three hypotheses are currently in vogue.

1. The **magnocellular deficit** hypothesis is predicated on the finding that dyslexics have impaired ability to process fast visual stimuli and some have abnormally small cells in the magnocellular layers of the lateral geniculate nuclei. In addition, the magnocellular cells of the medial geniculate nucleus

that project to the primary auditory cortex are smaller than normal and this impairs detection of frequency and amplitude of sounds, leading to poorer speech discrimination. An autoimmune response directed at antigens on the surface of magnocellular cells has been postulated to be responsible for their abnormalities.

2. The **cerebellar deficit** hypothesis takes the idea that normally phonemes come to be recognized not by their sounds but the articulations that are needed to produce them and argues that the cerebellum is important in learning this association. The notion is supported by finding behavioral evidence for impaired cerebellar motor functions in dyslexic children, and PET scans show reduced activation of the cerebellum during a motor learning task in dyslexic adults. A cerebellar involvement would also account for the execrable handwriting of many dyslexics and the link between dyslexia and clumsiness.
3. The **phonological deficit** hypothesis is that dyslexics have difficulty decomposing speech into its constituent phonemes and hence find it hard to learn the specific associations between phonemes and letters. The incidence of dyslexia is not the same for all languages. It is less common in native Italian speakers than native English speakers, probably because Italian has a much simpler phonetic structure than English. Neuroanatomical findings lend credence to the phonological deficit idea. There are histological abnormalities in the left temporal cortex of dyslexics (implying a developmental defect) and PET scans show that the insula is inactive, suggesting that the link between the supramarginal and angular gyri (the presumed locus of the phonological representation) and Broca's area, is not engaged in dyslexics. Functional MRI shows that dyslexic individuals have much lower activity in the parieto-temporal region implicated in phonological analysis than normal readers, and increased activity in the Broca's area, especially in older people, which is interpreted as an attempt to compensate for the defect. An intensive study of 16 dyslexics and an equal number of control subjects showed that all dyslexics have phonological deficits whereas five had none of the problems predicted by the other hypotheses.

P1 SCHIZOPHRENIA

Key Notes

The nature of schizophrenia

Two syndromes are recognized; type I includes hallucinations, delusions, disorders of thought and language; type II is a reduction in psychomotor functions, blunting of emotions and social withdrawal.

Genes or environment?

The incidence of schizophrenia is about 1%, with onset usually during the third decade. It afflicts all cultures similarly, regardless of geographical location, making risk factors hard to discern. Both genes and environment are thought to contribute to schizophrenia, but no specific genes or environmental factors have yet been unambiguously identified.

Anatomical abnormalities

Enlargement of the ventricles, disruption to cortical architecture with a modest loss of cortical mass, and a loss of cerebral asymmetry are evidence for a developmental defect, namely a failure of neurons to migrate into the cortex and consequently inadequate dopaminergic innervation of the cortex. The loss of cerebral asymmetry might account for the disorders of language and thinking.

Neurochemistry of schizophrenia

Because dopamine agonists can cause psychosis and dopamine antagonists are antipsychotic, and because there is evidence for increased numbers of dopamine receptors in the brains of untreated schizophrenics, it is argued that overactivity of dopamine transmission underpins schizophrenia. This dopamine hypothesis has been modified in the light of contradictory evidence. Type I schizophrenia is now attributed to hypoactivity of the mesocortical system, coupled with increased activation of the mesolimbic pathway. Conversely, type II results from increased mesocortical, but reduced mesolimbic dopamine transmission. There are also deficits in NMDA receptor-mediated glutamate transmission.

Cognitive deficits in schizophrenia

Schizophrenics show disorders in language, thinking and a decreased working memory, which are probably interrelated. For language and thinking, an executive attention system in the prefrontal cortex is presumed to access material from a long-term word memory store (lexicon) and put it into working memory. The lexicon is regarded as a semantic network in which related items are more likely to be recalled than unrelated items. In schizophrenia, the access to the semantic network is faster and more extensive than in normal subjects, so inappropriate or bizarre associations are made.

Dopamine acts on D1 receptors in the cortex to increase the signal-to-noise ratio for excitatory input; amplifying NMDA receptor transmission, but suppressing AMPA receptor input. This constrains access to semantic networks so that associations are only made between closely related items. In normal subjects, high motivation reinforces this by increasing dopaminergic transmission, while a low-motivation state, corresponding to low dopamine transmission, favors wider activation of material from semantic networks. In schizophrenia the deficit in mesocortical dopamine transmission results in the abnormal spread of activation through the semantic network. Since working memory is thought to be maintained by reverberating networks using NMDA receptors this is also impaired by lack of mesocortical input.

Pharmacology of schizophrenia

Typical antipsychotics are quite effective in treating positive symptoms by acting as antagonists of dopamine D2 receptors in the limbic system, but produce unwanted extrapyramidal effects because there are D2 receptors in the striatum. Atypical antipsychotics have the advantage that they do not cause extrapyramidal effects. They tend to be more effective in treating the negative symptoms than typical antipsychotics. Some act on dopamine D4 receptors while others are serotonin $5HT_2$ and $5HT_3$ antagonists.

Autism

In autism, after 18–20 months of normal development, children become socially withdrawn, exhibit pointless repetitive activity, have language problems and usually intellectual impairment. Autism is probably due to faulty brain development before 30 weeks of neonatal life that results in abnormal dorsal prefrontal cortex and limbic system functions. There is absolutely no evidence for the claim that the measles–mumps–rubella triple vaccine is a risk factor.

Related topics

Neurotransmitter release (C5)
Dopamine (D4)
Basal ganglia function (K8)
Cortical development (N3)
Language (O5)

The nature of schizophrenia

Two distinct sets of symptoms are recognized in **schizophrenia**.

1. **Positive symptoms** are principally delusions (irrational beliefs held despite confounding evidence), hallucinations (usually auditory) and thought disorder characterized by bizarre use of language. These imply dysfunction of medial temporal and lateral prefrontal cortices.
2. **Negative symptoms** include social withdrawal, low emotional arousal, lack of motivation, short attention span, and impaired judgment, planning and problem solving. This constellation of symptoms suggests limbic and prefrontal cortex dysfunction.

It is unclear whether these two syndromes represent different disorders or two phases of the same illness. Certainly both can be present in the same individual at different times. Most patients present with type I syndrome and progression, when it occurs, involves development of type II symptoms.

A period when the positive symptoms predominate is termed a **psychotic episode.** Psychotic episodes are not confined to schizophrenia but are also induced by a variety of drugs, such as those which potentiate dopamine neurotransmission (amphetamine, L-DOPA) and NMDA receptor antagonists (phencyclidine, ketamine). Interestingly, the young brain seems incapable of becoming schizophrenic; symptoms rarely occur until after puberty, and NMDA receptor antagonists do not produce psychosis in children.

Genes or environment?

The lifetime risk of schizophrenia is about 1% and for 40% of those afflicted, psychotic episodes recur throughout life. The usual age of onset is in the 20s, appearing on average 2–3 years earlier in men than women. Its incidence and presentation are much the same across all cultures and throughout the world, making it difficult to tease out risk factors.

Both genetic predisposition and environment play a part in the etiology of

schizophrenia. Adoption and twin studies have implied a 40% risk for children with two schizophrenic parents and a 50% concordance rate for monozygotic twins. However, many of these studies have been heavily criticized on the basis that they failed to properly factor-out environmental influences. No genetic linkage has been reliably identified up to the time of writing. In any case it is virtually certain that any genetic predisposition will involve many genes. The role of the environment is highlighted by the 63% of patients who have no family history of the disease. Epidemiological studies have suggested an increased risk of schizophrenia in individuals exposed to viral infections (there are seasonal differences in onset), or hypoxia, or high lead concentrations *in utero*. Stressful events commonly precede onset.

Anatomical abnormalities

The pathological anatomy seen in some schizophrenics, particularly those in whom negative symptoms predominate, are suggestive of defective brain development. They include:

- enlargement of the ventricles;
- reduction in the size of the thalamus;
- a small (~6%) decrease in frontal and medial temporal cortical mass, especially in the left hemisphere, accompanied by disruption of cellular architecture, particularly in the hippocampus and parahippocampal gyrus, where pyramidal cells are orientated abnormally;
- a reduction of the normal cerebral asymmetry in which (among other differences) the temporal plane, a region of the temporal lobe implicated in language, is usually larger in the left hemisphere than that on the right;
- smaller change in cerebral blood flow in the frontal cortex during working memory tasks than normal individuals.

None of these changes is accompanied by gliosis (scarring) so they are not the result of trauma. Furthermore, cerebral asymmetry is seen in the fetal brain, which suggests a developmental defect. The abnormalities are thought to reflect:

1. Deranged development in which neurons destined for the cortex instead migrate abnormally into the white matter, resulting in aberrant patterns of connections being made in the cortex.
2. A loss of thalamocortical axons and dopaminergic projections in the cortex, and reduced number of distal dendrites on cortical pyramidal cells.

One hypothesis is that the failure to develop cerebral asymmetry in the second trimester of pregnancy is the primary defect. The other changes follow, and since language is predominantly a left hemisphere function in most individuals, the lack of asymmetry accounts for the disorders of language and thought seen in schizophrenia.

Neurochemistry of schizophrenia

The original dopamine hypothesis, that schizophrenia results from overactivity of dopamine transmission was founded on pharmacological observations:

- High doses of amphetamines, which increase dopamine release, cause psychosis.
- Many dopamine receptor antagonists cause extrapyramidal motor disorders reminiscent of Parkinson's disease, which is caused by underactivity of dopamine transmission. Both L-DOPA (in Parkinson's disease patients) and D2 agonists can precipitate a psychosis that resembles type I schizophrenia.

- **Neuroleptic** (**antipsychotic**) drugs are antagonists of dopamine receptors. There is a high correlation between the therapeutic dose of neuroleptic drugs and their ability to bind D2 dopamine receptors.

This original dopamine hypothesis is contradicted by several findings including:

- Dopamine receptor antagonists cause Parkinson-like symptoms almost immediately but their antipsychotic action is not seen for 2–3 weeks.
- Type II symptoms are unaffected or even improved by amphetamines. This, together with the lack of motivation seen in the type II syndrome implies that type II symptoms are caused by a *lack* of dopamine function.

The contradictions are resolved by recognizing that two dopaminergic pathways are involved. Both originate in the ventral tegmental area of the midbrain. The mesocortical pathway goes to the prefrontal cortex, while the mesolimbic pathway goes to limbic structures such as the amygdala, hippocampus and anterior cingulate cortex, and to parts of the frontal cortex. There is a reciprocal relationship between the activity of these two pathways because the mesocortical pathway inhibits the mesolimbic pathway (*Fig. 1*). It is argued that increased activity in the mesolimbic pathway but a *decreased* activity of the mesocortical system is responsible for the positive symptoms.

Some positive symptoms improve with neuroleptics because they bind to D2 autoreceptors on mesolimbic neurons. At first this activates them but after 2 weeks or so the cells become persistently depolarized and fall silent so mesolimbic transmission is suppressed. This explains the lag in the onset of the antipsychotic action of D2 antagonists. Because the mesocortical neurons have few D2 autoreceptors they are not much affected by neuroleptics.

In contrast, negative (type II) symptoms may be due to increased mesocortical activity which *reduces* activity in the mesolimbic reward system. These symptoms may be improved by amphetamine stimulation of mesolimbic transmission. The resistance of negative symptoms to conventional neuroleptics occurs because they silence mesolimbic neurons that are already hypoactive, and have little effect on suppressing mesocortical cells, because these lack D2 autoreceptors.

However, schizophrenia is not just about alterations in dopamine transmission. There is abundant evidence that lack of glutamate transmission at NMDA receptors might be involved in schizophrenia:

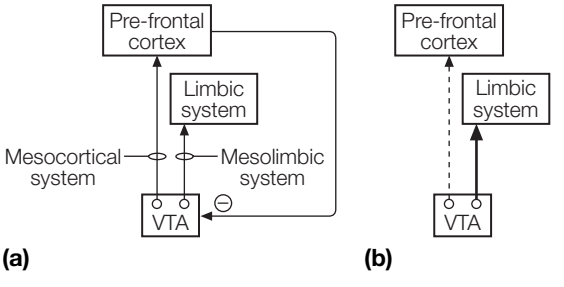


Fig. 1. Pre-frontal cortex modulation of dopaminergic pathways from the ventral tegmental area (VTA). (a) In health feedback inhibition operates facilitated by mesocortical input. (b) In schizophrenia loss of feedback inhibition occurs due to lack of mesocortical input. The mesolimbic system is disinhibited.

- Some NMDA receptor antagonists (phencyclidine or ketamine) cause psychosis and the cognitive problems seen in schizophrenia, except when administered to pre-pubertal children.
- Rats treated with NMDA receptor antagonists *in utero* suffer neurodegenerative changes in the limbic cortex, probably by interfering with the role of NMDA receptors in cortical plasticity.
- Lesions of specific regions of the frontal cortex or hippocampus (in which most excitatory transmission is glutamatergic) in the perinatal period result in cognitive and behavior deficits, usually attributed to hyperactivity of dopamine transmission, that becomes most apparent after puberty. This reflects the time course of schizophrenia.

Changes in glutamate transmission are possible candidates for explaining the cognitive deficits in schizophrenia.

Cognitive deficits in schizophrenia

Language, thought and working memory dysfunction are the principal cognitive signs in schizophrenia and they are probably all interrelated. Speech becomes repetitive and without clear direction, dominated by a succession of ideas linked by remote associations. The term **thought disorder** describes a state in which the patient feels that they have no control over their own thoughts and may be disturbed by their content. Schizophrenics have a decrease in working memory and show reduced activation of the prefrontal lobes in imaging studies. The inability to hold items in working memory might account for the failure to use ongoing cues (external and internal) as guides for behavior, which hence seems bizarre and inappropriate.

Schizophrenics are more likely to have had reading difficulties as children than non-schizophrenics, which is taken to reflect a lack of left hemisphere lateralization for language. This is a functional correlate of the lack of anatomical cerebral asymmetry seen in schizophrenics that could account for the language dysfunction they have. Moreover, as much thinking (making plans, rehearsing arguments) in humans is done by internal speech, the thought disorders in schizophrenia could relate to the language deficit. A cognitive model for thinking is that an executive attention system manipulates sequences of words (phrases, sentences) in working memory, and this occurs in the left frontal cortex. Word sequences are recruited from a long-term word memory store (**lexicon**). The lexicon is one of a class of neural networks which store **semantic maps** that are hypothesized to exist in the brain. In semantic maps, objects (e.g. words), with similar meanings are represented by similar sets of synaptic connections. Activating recall of a specific object from the network is more likely to generate recall of related than unrelated objects.

Computer simulations indicate that for a neural network to store a semantic map it needs to be associative and self-organizing (capable of constructing itself without knowledge of the correct associations in advance), have Hebbian synapses and show lateral inhibition. These are likely properties of networks of nerve cells in the brain, which supports the view that the brain has semantic maps and provides clues as to how they might work.

The existence of semantic maps in the brain can also be inferred from **semantic priming**. Less time is taken to recognize a word (bee) flashed onto a computer screen if it follows a related **primer** word (honey) than an unrelated primer word (train). This is **direct semantic priming** and is seen to a greater extent in schizophrenics than controls. In addition, **indirect semantic priming**,

in which a response can only be understood in terms of an interpolated association (honey, [sweet], sour) is also different in schizophrenics, While it takes normal individuals about 700 ms to think of an indirectly primed word (sour) after the primer (honey), schizophrenics take only about 200 ms. In addition, normal subjects are less likely to make the indirect associations than schizophrenics. From these data it is argued that the semantic networks storing the semantic maps are activated faster and more extensively (that is, in a less-focused manner) in schizophrenics than in normal individuals.

How does the failure of proper dopamine modulation of cortex cause type I schizophrenia? Excessive activation of a striato–thalamo–cortical circuit involving the inferotemporal cortex has been proposed to underlie hallucinations on the basis of functional brain-imaging studies. The thought disorder may be explained by a joint dysfunction of associative semantic memory and working memory. A model to account for how access to semantic maps might be influenced by alterations in dopamine transmission is as follows. The mesocortical dopaminergic neurons terminate mostly in the frontal cortex, forming synapses with the pyramidal cells that give rise to cortico-cortical association fibers. This dopaminergic innervation is established at about 2–4 years of age. The dopamine receptors here are mostly D1. Dopamine acts as a neuromodulator of excitatory input. It *attenuates* depolarizing responses of glutamate by AMPA receptors, but *enhances* the action of glutamate at NMDA receptors. In this way, dopamine increases the signal-to-noise ratio, amplifying only the strongest (NMDA receptor-mediated) excitatory signals, at the expense of the weaker AMPA receptor activation. This is postulated to keep activation of the semantic networks sharply focused.

L-DOPA in normal subjects causes a reduction in indirect priming, consistent with a more focused activation of semantic networks. Hence a reduced mesocortical dopamine transmission in schizophrenia might be expected to result in a less focused (i.e. wider) activation of semantic networks; items that are only tenuously related may come to be spuriously and bizarrely associated in language and thinking. In normal subjects, high motivation, with consequent enhanced dopamine modulation, will produce sharply focused activation of semantic networks. However, in a lower motivational state there would be a wider activation, with a greater tendency of making faster and more unusual associations; an attribute of creativity. In addition, working memory is thought to be maintained by reverberating circuits stabilized by recurrent excitation via NMDA receptors. These circuits are likely to be less stable in schizophrenics because the lack of mesocortical dopamine means NMDA receptor-mediated activity is less amplified.

Pharmacology of schizophrenia

Typical antipsychotics, which are quite effective in ameliorating positive symptoms, all have high affinity for dopamine D2 receptors. D2 receptors located in limbic structures and parts of the cortex are presumably responsible for the antipsychotic actions, but unfortunately D2 receptors are also found in the striatum. Blockade of striatal D2 receptors inhibits the indirect pathway for basal ganglia movement control resulting in the unwanted extrapyramidal effects (e.g. tardive dyskinesia) of the typical antipsychotics.

Atypical antipsychotics do not cause extrapyramidal effects and tend to be more effective in treating the negative symptoms than typical antipsychotics. Some act on dopamine receptors that belong to the same family as D2 receptors. Termed D4 receptors, they are found in the limbic system and cortex but not in

the striatum explaining why atypical antipsychotics such as clozapine – which acts selectively at the D4 receptor – are free of extrapyramidal side effects. Other atypical antipsychotics are antagonists at serotonin 5-HT_2 and 5-HT_3 receptors. How they work to alleviate schizophrenia is not known.

Autism

In autism, after 18–20 months of normal development, children become socially withdrawn, show **stereotypic behavior** (pointless repetitive activity), and language abnormalities. This pattern implies some similarity to schizophrenia but the primary problem in autism is thought to be the failure of the child to develop a theory of mind, an internal representation of how other people think and feel. Hence, autistic patients cannot make normal social attachments. Intellectual impairment is seen in 80% of autistic children.

Twin studies imply a genetic predisposition to the disorder but it is likely that many genes are involved. Autism most probably arises from a defect in brain development before 30 weeks of neonatal life. Autistics have poor executive functions (e.g. inflexible thinking, difficulty in switching attention) and this is reflected in abnormal dorsal prefrontal cortex function as shown by fMRI. Abnormalities are also seen in limbic structures (e.g. amygdala, anterior cingulate cortex) and cerebellum, where there is a loss of Purkinje cells. There is evidence for dopaminergic and serotinergic dysfunction in autism, but attempts to treat it with antipsychotics, dopamine agonists or antidepressants have not shown consistent improvements. Mercury neurotoxicity has been postulated on the basis of similarities between neurochemical findings in autism and mercury exposure. Although rubella infection is a risk factor for autism, several large retrospective studies provide strong evidence *against* the claim that the measles–mumps–rubella triple vaccine is a risk factor.

P2 DEPRESSION

Key Notes

The nature of depression

Reactive depression is the normal response to very stressful life events and is common. Endogenous unipolar depression, and bipolar affective (manic-depressive) disorder, characterized by cycles of depression and mania, have no obvious triggers but tend to recur throughout life. Depression not only affects mood, it also produces physiological dysfunction. Twin studies indicate a genetic component, while the younger age of onset and increased incidence since 1940 implies environmental changes are important. Brain imaging shows decreased activity of the prefrontal cortex in depression.

Neurochemistry in depression

All treatments for depression increase serotonin transmission in the long term. They either act on presynaptic inhibitory autoreceptors to enhance serotonin release (monoamine oxidase inhibitors, selective serotonin reuptake inhibitors or autoreceptor agonists) or increase postsynaptic serotonin receptors (tricyclic antidepressants, electroconvulsive therapy). Long-term antidepressant treatment down-regulates β-adrenoceptors, but only if serotonin transmission is intact, showing the importance of interactions between serotonin and noradrenaline (norepinephrine) systems. Dysfunction of the mesolimbic and mesocortical reward system is thought to underlie the clear swings in motivation seen in bipolar disorder. This is supported by the effects of dopamine receptor agonists (which precipitate mania) or antagonists (which reduce mania) and by the structural changes to the targets of the dopamine reward pathways.

Depression and the hypothalamic–pituitary–adrenal (HPA) axis

Enhanced activity of the HPA axis and defects in its negative feedback system are seen in depression. Corticotrophin-releasing hormone reproduces many of the features of depression in rats. HPA anomalies in depression may reflect defects in its regulation by monoaminergic transmitter systems.

Related topics

Neurotransmitter release (C5)
Neurotransmitter inactivation (C7)
Dopamine (D4)
Noradrenaline (norepinephrine) (D5)
Serotonin (D6)
Neuroendocrine control of metabolism and growth (L3)

The nature of depression

Depression is an **affective** disorder since it alters mood. There are three broad types of depression. **Reactive depression** is the most common and is seen in response to life events such as the death of a loved one, disease, or severe and inescapable stress. Given time individuals can be expected to recover fully. **Endogenous unipolar depression**, which occurs spontaneously, or to a life event which would not normally be thought severe enough to have such a profound effect on mood is at least two distinct diseases (*Table 1*) and probably more. Depression interspersed with periods of mania is termed **bipolar**

(manic-depressive) disorder. Both unipolar and bipolar depression tend to recur throughout life.

Table 1. Features of unipolar depression

Feature	Melancholic	Atypical
Incidence (%)	~60	~15
Usual time-course	Recurrent	Chronic
Depression	Worse a.m.	Worse p.m.
Sleep pattern	Insomnia with early a.m. waking	Excessive sleep
Eating pattern	Undereating, anorexia	Overeating, weight gain
Typical mental state	Anhedonia	Anxiety

All depressed patients have altered mood in which they feel miserable, apathetic, pessimistic, stressed, anxious, inadequate or guilty in some proportion. In those with clinical depression these feelings are usually more intense, protracted and accompanied by indecisiveness, or diminished ability to think, and often suicidal thoughts. In addition they show evidence of physiological dysfunction; altered motor activity, loss of motivation and pleasure sense, altered (usually loss of) appetite and libido, and altered sleep patterns. Mania is the opposite of depression and is characterized by euphoria, restlessness, insomnia, rapid thought processes and, when severe, disorganized and irrational thinking.

There is circumstantial evidence for a genetic component to both disorders though no genes have yet been identified (*Table 2*). There is about a 50% chance that both members of a pair of monozygotic (identical) twins will be manic-depressive, but for dizygotic twins this concordance rate drops to 10%.

Table 2. Familial patterns in depressive disorders

	Lifetime risk for first degree relatives (%)		
	Healthy relative	Relative with unipolar disorder	Relative with bipolar disorder
Unipolar	5.4	15.0	12.0
Bipolar	0.8	2.6	6.0

The importance of environment is shown by two trends in depression seen since 1940:

- the average age of onset has decreased from 35 to 28;
- the incidence of depression has increased.

One explanation for this is that susceptible individuals are becoming more likely to succumb because of the increasing stress of everyday life.

PET scans show decreased activity of the part of the prefrontal cortex in patients during depressed periods. Lesions of the same region make it difficult for people to express emotions and they have abnormal autonomic responses to emotive stimuli. The prefrontal cortex has extensive connections with the limbic structures implicated in emotion.

Neurochemistry in depression

At present there is no coherent neurochemical model for depression but most evidence points to abnormalities in monoamine transmission.

All treatments for depression increase serotonin neurotransmission. However, in no case is this directly brought about by their immediate pharmacodynamic action. Instead, delayed adaptive changes are thought to be responsible because therapeutic effects are not seen for several weeks. Three classes of drugs appear to work by decreasing the sensitivity of 5-HT_{1A} and 5-HT_{1B} inhibitory autoreceptors, so that serotonin release is enhanced:

- **monoamine oxidase inhibitors** (**MAOIs**) which increase stores of NA and 5-HT by suppressing their catabolism
- **selective serotonin reuptake inhibitors** (**SSRIs**), which block the uptake of serotonin
- **5-HT_{1A} receptor agonists** which directly down-regulate the presynaptic autoreceptors.

Tricyclic antidepressants which unselectively inhibit the transporters responsible for the reuptake of both noradrenaline (norepinephrine) and serotonin, and **electroconvulsive therapy** (**ECT**), increase postsynaptic serotonin 5-HT_{1A} and 5-HT_{2A} receptors. Long-term antidepressant therapy also results in changes in noradrenergic transmission, for example, the down-regulation of β-adrenoceptors (βAR). However, in rats, destruction of serotonergic neurons with a selective toxin prevents the down-regulation of βAR receptors seen with chronic antidepressant treatment. This shows that there are interactions between the serotonin and noradrenaline (norepinephrine) transmitter system neurons (e.g. 5-HT neurons in the raphe nuclei make inhibitory connections with noradrenergic neurons in the locus ceruleus), which complicates the interpretation of neurochemical experiments. Indeed the significance of the antidepressant-induced reduction in βAR is unclear. Although increased numbers of βARs are seen in the brains of depressed patients postmortem, β-adrenoceptor antagonists are not antidepressant.

Dysfunction of the mesolimbic and mesocortical dopaminergic reward system is implicated in bipolar disorder because of the wide swings in **hedonia** (ability to experience pleasure), motivation and psychomotor activity seen in this condition. It is postulated that underactivity is responsible for depression and overactivity for the mania. Evidence for this is as follows:

- Parkinson's disease, in which dopamine is severely depleted, is frequently accompanied by depression.
- Dopamine antagonists reduce mania, while dopamine agonists can precipitate mania in both normal individuals and in patients with bipolar disorder who are in a depressed phase.

A major target of the dopaminergic system is the medial prefrontal cortex. Brain-imaging studies show that this region has a smaller volume in patients with bipolar disorder. PET scans of 2-deoxyglucose uptake show that neural activity of this region is reduced during depression and elevated in mania. Furthermore, the medial thalamus and amygdala which have extensive connections with the medial prefrontal cortex show abnormal reductions in volume in bipolar disorder.

The mania of bipolar disorder responds to lithium, though why it is effective is unknown. It has a number of actions on second messenger systems including the phosphoinositide and adenylyl cyclase systems, any combination of which might be antidepressant.

Depression and the hypothalamic–pituitary–adrenal (HPA) axis

One of the best-documented findings in depression is an increased activity of the HPA axis, as shown by raised concentrations of ACTH and glucocorticoids in unmedicated patients with depression. Moreover, the normal negative feedback effect of glucocorticoids on the hypothalamus and pituitary is blunted in depressed patients. Postmortem studies reveal a greater number of corticotrophin-releasing hormone (CRH)-secreting cells in the hypothalamus and elevated expression of the CRH gene. Microinjection of CRH into the brains of rats produces insomnia, anxiety and a fall in appetite, all of which are cardinal features of depression in humans. Depriving rat pups of maternal contact for brief periods over the first 21 days postnatally makes their HPA axis lastingly more sensitive to stress; these animals show greater rises in CRH, ACTH and glucocorticoid secretion and a higher density of CRH receptors. Treating these rats with an SSRI normalized their HPA axis and made them less fearful. Hence the HPA may play a key part in depression. As there are monoaminergic projections to CRH-secreting neurons, dysfunction in noradrenergic and serotonergic systems and the HPA axis in depression may be linked. CRH receptor antagonists have antidepressant potential.

P3 STROKES AND EXCITOTOXICITY

Key Notes

Cellular events in strokes

Most strokes are caused by blockage of cerebral arteries by blood clots produced locally or circulating from the heart. In the core of the ischemic region, cells starved of oxygen and glucose suffer rapid necrotic cell death. Death of cells in the surrounding penumbra is delayed and occurs because excessive glutamate release overexcites neurons, triggering apoptosis (programmed cell death).

Excitotoxic cell death

Deprived of oxygen, cellular ATP concentrations fall and the consequent reduced activity of the sodium pump facilitates Na^+ and Cl^- influx into cells (producing osmotic swelling) and K^+ efflux which depolarizes cells. Excessive glutamate release follows from the depolarization, stimulating NMDA, AMPA and metabotropic glutamate receptors to promote Ca^{2+} entry into neurons. Inappropriate Ca^{2+} entry can trigger either necrosis or (via free radical production or transcription) apoptosis. Excess secretion of Zn^{2+} from excitatory nerve terminals may act in a similar manner to Ca^{2+}. Excitotoxicity due to excessive glutamate is a final common path for cell death in a number of neurodegenerative conditions.

Possible treatments in stroke

Treatment strategies aim to rescue cells in the penumbra. One successful approach in humans with acute ischemic stroke is the use of clot-dissolving agents. Drug development involves using animal models of stroke, but some approaches that are successful in animals have not proven clinically useful. NMDA receptor antagonists, Ca^{2+} channel blockers and free radical scavengers have not proved efficacious in clinical trials. AMPA receptor antagonists, apoptosis-inhibiting agents and Zn^{2+} chelators are currently being considered.

Related topics

Resting potentials (B1)
Neurotransmitter inactivation (C7)
Ionotropic receptors (D1)
Metabotropic receptors (D2)
Amino acid transmitters (D3)
Neurotrophic factors (N6)

Cellular events in strokes

Most strokes (cerebrovascular accidents, CVAs) result from occlusion of a brain artery by a thrombus (blood clot) that forms *in situ* or travels there from the heart. The loss of blood flow (**ischemia**) starves brain tissue of oxygen and glucose. In the region of brain entirely dependent on the blocked blood supply, neurons and glia will die of **hypoxia** (lack of oxygen). This region is the **core** and the mode of cell death here is termed **necrosis**. Surrounding the core is the **penumbra**, this region suffers some ischemia but receives blood from other arteries. Cells in this region may survive or suffer delayed neuron death (DND). DND occurs not *directly* from the lack of oxygen but from **excitotoxicity**. Here, the reduced oxygen concentration causes neurons to secrete excessive amounts of glutamate, activating glutamate receptors to produce an influx of Ca^{2+} that is thought to kill cells by triggering **apoptosis** (programmed cell death).

Excitotoxic cell death

The cascade of events which leads to the death of cells from excitotoxicity is as follows. During ischemia, ATP concentrations are depleted so that energy-requiring processes are compromised, including the Na^+/K^+-ATPase. Decreased activity of this cation pump increases the intracellular concentration of Na^+, and consequently Cl^- and water shift into the cells, which swell (*Fig. 1*). In addition, the extracellular K^+ concentration rises making the potassium equilibrium potential (E_K) more positive (see the Nernst equation for E_K, Topic B1), causing membrane depolarization. This depolarization activates voltage-dependent Ca^{2+} channels provoking neurotransmitter release. Since a large number of neurons in the brain use glutamate, this transmitter is released in excessive quantities. Glutamate release is exacerbated by the action of the Na^+/K^+-dependent glutamate transporter which normally removes the neurotransmitter from the cleft. This depends indirectly on metabolic energy since it is driven by the gradients for Na^+ and K^+ ions that are maintained by Na^+/K^+-ATPase. In ischemia the intracellular Na^+ and extracellular K^+ are thought to become so high as to force the transporter to work in reverse, extruding glutamate from the depolarized axons and astrocytes into the extracellular space.

The combination of a large depolarization and synaptic glutamate are precisely the conditions needed to activate NMDA receptors, resulting in Ca^{2+} influx. However, there are other routes by which intracellular Ca^{2+} is raised that may be as important:

- Ischemia-evoked up-regulation of the Ca^{2+}-permeable subtypes of the AMPA receptor.
- Stimulation of metabotropic glutamate receptors that are coupled to the phosphoinositide second messenger system (Type I mGluRs) since inositol trisphosphate (IP_3) causes the liberation of Ca^{2+} from internal stores.
- Operation of the Na^+/Ca^{2+} exchanger in reverse because of the high intracellular Na^+ concentration. Normally, this transport system uses sodium influx to expel calcium ions from the cell. If the sodium gradient falls however, the transport goes in the opposite direction.

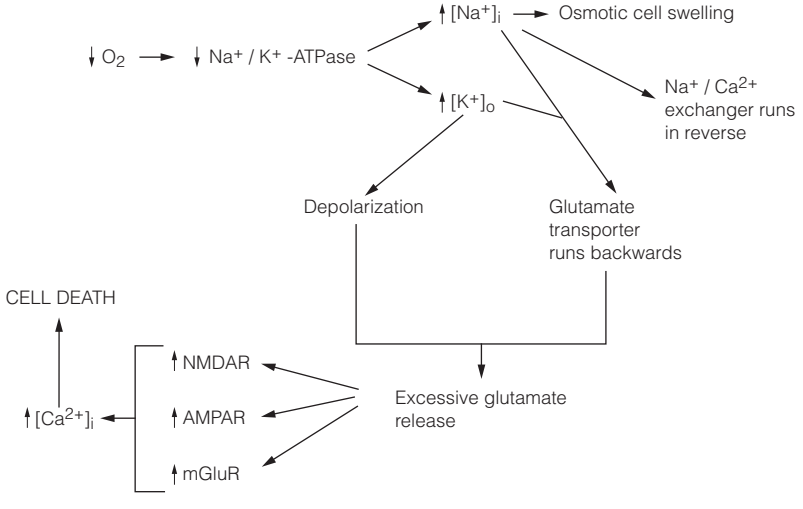


Fig. 1. Key events in ischemic stroke that lead to excitotoxic cell death. NMDAR, NMDA receptor; AMPAR, AMPA receptor; mGluR, metabotropic glutamate receptor.

- The uncontrolled Ca^{2+} influx overloads the transport systems and buffers, which normally act to reduce the concentration of free Ca^{2+} in the cell. The Ca^{2+} triggers processes that eventually kill the cell.

Very high cytoplasmic Ca^{2+} concentrations are clearly cytotoxic and cause necrosis. The Ca^{2+} levels needed to trigger apoptosis are not well delineated. Although elevated Ca^{2+} concentrations seem to be associated with apoptosis in some cases, paradoxically, apoptosis of neurons in culture can be attenuated by modest increases in intracellular Ca^{2+} (e.g. by activating voltage-dependent Ca^{2+} channels).

Zinc ions (Zn^{2+}) which are released from some excitatory nerve terminals (e.g. mossy fiber terminals of the hippocampus) acts as a neurotransmitter or neuromodulator in the CNS, having actions at several types of receptor. Zn^{2+} is released from nerve terminals in ischemia, enters neurons via the same routes as Ca^{2+} (i.e. via NMDA and AMPA receptors and voltage-dependent Ca^{2+} channels) and is neurotoxic at the concentrations seen in ischemic brain. Administering agents which chelate Zn^{2+} before transient global ischemia in animals reduced subsequent cell death. Hence, Zn^{2+} may be, with Ca^{2+}, an important contributor to ischemic brain cell death in strokes.

There are several ways in which Ca^{2+} may kill neurons. Firstly, Ca^{2+} activates a number of enzymes, endonucleases and proteases, the uncontrolled action of which would terminally disrupt cell function. Secondly, Ca^{2+} switches on processes that generate **free radicals**, including superoxide anions ($^{\bullet}O_2^-$), which are highly reactive, initiating extremely damaging chemical reactions, such as peroxidation of lipids in cell membranes. Free radical damage can lead to apoptosis. Thirdly, Ca^{2+} acting via a number of kinases can activate the transcription of genes that trigger apoptosis.

Excitotoxic cell death is also implicated in neurological disorders that arise as a result of genetic mutations that cause specific, regional cell loss; for example:

- **Huntington's disease** in which medium spiny neurons in the striatum which receive glutamatergic inputs from the cortex die;
- **Amylotropic lateral sclerosis (motor neuron disease)** which results from the death of motor neurons in the brain stem and spinal cord.

Possible treatments in stroke

Since cells in the core are killed within a very short time of the occlusion, short-term treatment strategies must focus on increasing the survival rate of cells in the penumbra and this means limiting excitotoxicity. The endogenous thrombolytic (clot-dissolving) protein, **tissue plasminogen activator (tPA)** has proved effective in clinical trials and is now used for the treatment of acute ischemic

Table 1. Classes of neuroprotective drug that have failed phase III clinical trials in occlusive strokes*

Na^+ channel antagonists
Ca^{2+} channel antagonists
NMDA receptor competitive antagonists
$GABA_A$ agonist
Nitric oxide synthase inhibitor
Free-radical scavengers

* Out of 19 trials, no improvement was seen in 17, and in two there was an unfavorable risk–benefit ratio.

stroke. This approach cannot be used in hemorrhagic strokes caused by the rupture of blood vessels.

Development of novel drugs for the treatment of stroke uses animal models in which ischemic lesions are generated by occluding the cerebral arteries. Unfortunately, some strategies that seemed promising in animal studies have not been vindicated in clinical trials (*Table 1*). The use of AMPA receptor antagonists, Zn^{2+} chelators or caspase-3 inhibitors to prevent apoptotic cell death may yet prove useful.

P4 EPILEPSY

Key Notes

Types of epilepsy

Epilepsy is characterized by recurrent seizures, brief periods of abnormal synchronized neural firing. Epilepsy can be acquired (head trauma, brain cancer and neurodegenerative disease are risk factors) or inherited. Generalized (tonic-clonic, absence) seizures are widespread and consciousness is always lost, whereas partial seizures originate from a single focus and may (complex) or may not (simple) be accompanied by loss of consciousness. A particular drug may not be effective in all types of epilepsy.

Neurobiology of epilepsy

Normally in the hippocampus CA3 pyramidal cells fire bursts of action potentials and via their Schaffer collaterals drive CA1 cells. However, extensive inhibition from GABAergic interneurons prevents burst firing of CA1 cells. In epileptic hippocampal slices however, CA3 cells do trigger bursting of CA1 cells, and burst firing of CA3 cells is itself exaggerated by paroxysmal depolarizing shifts (PDS) due to Ca^{2+} influx. PDSs cause abnormal (interictal) spikes seen in the EEG of the epileptic cortex between seizures. The development of the hyperexcitable state that predisposes to epilepsy seems to require NMDA receptors. However, individual seizures are initiated by AMPA receptors (NMDA receptors are subsequently activated) and terminated by adenosine produced from ATP as a result of the high neural activity in seizures.

Possible causes of hyperexcitability

Rare familial epilepsies caused by mutations in voltage-dependent Na^+ or K^+ channels or nicotinic cholinergic receptors have been discovered. Hyperexcitability of acquired epilepsy may initially be produced in a manner akin to long-term potentiation. Over longer periods neurotrophic factors stimulate excessive sprouting of dentate granule cell axons. This reinforces the hyperexcitability by forging of an increased number of recurrent excitatory connections. In an animal model of epilepsy, hyperexcitability is caused by increased sensitivity of NMDA receptors coupled to a reduction of inhibition mediated by $GABA_A$ receptors. In temporal lobe epilepsy, release of neurotrophins enhances NMDA receptor activation and initiates long-term potentiation. Neurotrophins also trigger sprouting of the mossy fiber axons of dentate gyrus granule cells, which consequently become hyperexcitable.

Pharmacology of epilepsy

Barbiturates and benzodiazepines enhance inhibition by GABA at $GABA_A$ receptors but may also be anticonvulsant by actions on voltage-dependent ion channels or AMPA receptors. Blockade of Na^+ channels is the mode of action of phenytoin and carbamazepine. Absence seizures are caused by the synchronized firing of thalamocortical circuits, caused by activation of T-type Ca^{2+} channels. These channels are blocked by ethosuximide, a drug effective in absence seizures.

Related topics

Amino acid transmitters (D3)
Sleep (M5)
Neurotrophic factors (N6)
Cell physiology of learning (O3)

Types of epilepsy

Epileptic seizures are caused by an abnormal, synchronized firing of large populations of neurons that is usually self limiting. **Epilepsy** is defined as a disease in which such seizures recur. The incidence is about 1% and despite the variety of drugs currently available, control of seizures is acceptably good in only about three-quarters of individuals. Head trauma sufficient to cause coma, brain cancer, withdrawal from ethanol dependence and neurodegenerative brain disease are common risk factors for epilepsy. In addition, rarely epilepsies are inherited. Epilepsy is classified according to clinical presentation.

Generalized seizures are widespread. They include **tonic-clonic seizures** (grand mal), characterized by loss of consciousness and convulsions, and **absence seizures** (petit mal) in which the individual loses consciousness for just a few seconds (they appear to be 'day dreaming'), and which is accompanied by a 3-Hz EEG signal originating from the thalamus. **Partial (focal) seizures** are initiated from one region of the cortex, typically the motor cortex (in **Jacksonian** epilepsy) or the temporal lobe. In simple partial seizures consciousness remains, in complex partial seizures consciousness is lost. Drugs effective on some subtypes of epilepsy fail to be effective on others, which implies that different mechanisms are at work. However, some people have more than one type of fit, or progress from one type to another, which suggests a common underlying abnormality.

Neurobiology of epilepsy

Cellular and molecular mechanisms that might underlie epilepsy have been studied particularly in the hippocampus. Brief, high-frequency electrical stimulation (1 s, 60 Hz) delivered to the hippocampus or amygdala of a rat once or twice each day, through chronically implanted electrodes, results (after about 2 weeks) in animals having seizures that resemble complex partial seizures in humans. This animal model is called **kindling** and epileptiform activity is investigated in brain slices removed from kindled animals. Epileptiform activity can also be provoked in hippocampal slices from normal (unkindled) animals by a variety of manipulations, e.g. applying NMDA receptor agonists or $GABA_A$ receptor antagonists.

Spontaneous burst firing is part of the normal repertoire of CA3 pyramidal cells, but not of CA1 pyramidal cells. When CA1 cells are driven physiologically by CA3 cells they are prevented from burst firing by inhibition delivered by GABAergic interneurons. However, epileptogenesis in hippocampal slices alters the firing behavior of pyramidal cells in several ways. Firstly, CA3 cells trigger burst firing in CA1 cells via Schaffer collaterals, which suggests a weakening of inhibition. Secondly, computer modeling of the hippocampus shows that if normal levels of inhibition are reduced, the connections between neighboring CA3 pyramidal cells would allow them to fire in synchrony. Thirdly, the burst activity of pyramidal neurons in slices made epileptic is itself abnormal. The cells display **paroxysmal depolarizing shifts** (**PDS**), each a long-lasting depolarization, due largely to Ca^{2+} influx, which generates a burst of action potentials. These events underlie **interictal spikes**, abnormal EEG activity produced by epileptic cortex between seizures (*Fig. 1*).

Epileptogenesis refers to the development of the hyperexcitable state that predisposes to seizures. It is blocked by **antiepileptic** drugs, and it is distinct from the processes that trigger individual seizures that can be inhibited by **anticonvulsants**. NMDA receptor antagonists are good antiepileptics in that they completely prevent hippocampal kindling, but they are not very effective at stopping epileptiform activity in slices already kindled; i.e. they are poor

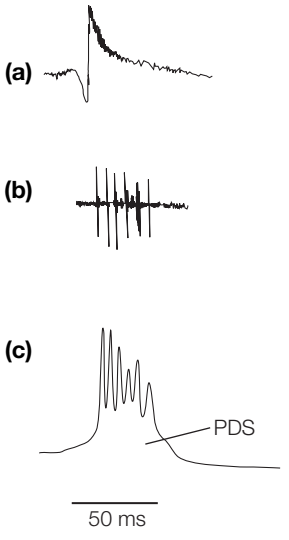


Fig. 1. Electrophysiology of epileptogenic activity. (a) Interictal spike recorded with surface electrodes. (b) Extracellular recording. (c) Intracellular recording of burst firing of hippocampal pyramidal cell. PDS, paroxysmal depolarizing shift.

anticonvulsants. The implication is that NMDA receptors are necessary for epileptogenesis.

By contrast, AMPA receptor activity is thought to initiate the generation of *individual* seizures. NMDA receptors and L-type Ca^{2+} channels are subsequently recruited which leads to prolonged burst firing. Adenosine is probably a key player in the termination of individual seizures. Neural activity is very high during seizures and adenosine derived from extensive ATP catabolism is transported across the plasma membrane to act on neural adenosine receptors. The rise in adenosine concentration that peaks about 30–60 seconds after the onset of a seizure is part of a normal physiological response that increases blood flow to match metabolic demand. Adenosine acts at A1 receptors that, by coupling to G_i proteins, open K^+ channels and close Ca^{2+} channels. Seizure termination is brought about by the resulting hyperpolarization. Adenosine A1 receptor agonists are anticonvulsants.

Possible causes of hyperexcitability

In three, rare, inherited epilepsies a single mutant gene is responsible in each case. In one, a point mutation occurs in a component of voltage-dependent Na^+ channels. In a second, the mutated gene codes for the subunit of a voltage-dependent K^+ channel. Those channels that include the mutant subunit have a 20–40% reduction in K^+ current and this accounts for the neuronal hyperexcitability in this epilepsy. The third familial epilepsy is caused by a mutation of the α4 nicotinic acetylcholine receptor (nAChR) subunit. Brain nAChRs have a significant permeability to Ca^{2+} and some are located presynaptically where they promote GABA release. The mutant nAChRs have a lower Ca^{2+} permeability so may produce epilepsy by preventing proper synaptic inhibition.

Epileptogenesis in acquired epilepsies is far more difficult to account for. High-frequency stimulation of hippocampal pathways that mimics the hyperactivity of epilepsy triggers a cascade of gene expression that leads to neurite growth in hippocampal cells which may be similar to the events that cause

remodeling of neural circuits in LTP. Hence epileptic activity may involve inappropriate strengthening or formation of synapses. Increased recurrent excitatory synapses also develop in the CA3 region of cultured hippocampal slices after cutting the axons of the pyramidal cells. In these slices seizure activity is normally kept in check by inhibition from GABAergic interneurons. Applying low doses of bicuculline, a $GABA_A$ receptor antagonist, triggers seizures in these slices, but not in control slices. This suggests that individual seizures in epilepsy might be triggered by brief lapses of synaptic inhibition.

The modes of action of some drugs that successfully prevent seizures support the long-held contention that epilepsy is a matter of too much excitation and too little inhibition. In kindling, NMDA receptors become more sensitive to agonists and there is a reduction of inhibition by $GABA_A$ receptors that might be caused by NMDA receptor activation. The link between NMDA and $GABA_A$ receptors is Ca^{2+}. Influx of Ca^{2+} through the NMDA receptor activates a phosphatase, **calcineurin**. This dephosphorylates the $GABA_A$ receptor which consequently has a much reduced Cl^- current upon binding GABA. In summary, in this hypothesis, epileptogenesis results from greater NMDA receptor activation coupled with lower $GABA_A$-mediated inhibition.

The increase in NMDA receptor sensitivity may be due to the action of neurotrophins. In temporal lobe epilepsy seizure activity is accompanied by the expression of brain-derived neurotrophic factor (BDNF). In conjunction with neuron electrical activity, BDNF enhances NMDA receptor activation, initiating long-term potentiation at glutamatergic synapses. Such a mechanism would explain why seizure activity makes future seizures more likely; i.e. seizures are themselves epileptogenic, as is demonstrated by kindling. Presumably it can be a fairly long-term process because acquired epilepsy typically follows brain injury only after a delay ranging from weeks to years. This time course suggests growth of neurites and rewiring is happening.

Indeed, neurotrophic factors are responsible for the **mossy fiber sprouting** (**MFS**) seen in temporal lobe epilepsy. This is extensive sprouting of the axons (mossy fibers) of dentate gyrus granule cells, which establish aberrant recurrent excitatory connections between the granule cells that consequently become uncharacteristically hyperexcitable. (Normally it is extremely difficult to induce seizure-like activity in dentate granule cells.) MFS occurs in numerous animal models of temporal lobe epilepsy and in **Ammon's horn sclerosis**, in which extensive death of neurons and injury-evoked proliferation of glial cells occurs, a pathology that is commonly seen in the hippocampus in drug-resistant human epilepsies. Surgical resection of sclerotic hippocampus usually cures the epilepsy.

Pharmacology of epilepsy

Barbiturates and benzodiazepines act on $GABA_A$ receptors, increasing the Cl^- current caused when GABA binds. The anticonvulsive action of these drugs is usually attributed to this enhancement of GABA inhibition. However, in addition, barbiturates may owe some of their antiepileptic action to blocking L-and N-type Ca^{2+} channels and by non-competitive inhibition of AMPA receptors. Moreover, blockade of adenosine uptake and block of voltage-dependent Na^+ channels may contribute to the anticonvulsant profile of benzodiazepines. Vigabatrin, an anticonvulsant useful in treating partial seizures, potentiates GABAergic inhibition indirectly, by inhibiting the enzyme, GABA-transaminase, which normally breaks down GABA.

It has long been recognized that two widely used agents, phenytoin and

carbamazepine, are anticonvulsant by binding inside the pores of voltage-dependent Na^+ channels, stabilizing them in an inactivated state.

During absence seizures, thalamic neurons go into the burst-firing mode seen during slow-wave sleep. The bursting is synchronized and sustained by the reciprocal connections between the thalamus and cortex, and gives rise to the characteristic 3-Hz EEG signal. Burst firing is caused by activation of T-type Ca^{2+} channels, and the brevity of absence seizures (just a few seconds) is presumably because T-type Ca^{2+} channels are rapidly inactivated by depolarization. Ethosuximide and related drugs that are particularly effective in the treatment of absence seizures produce partial block of T-type Ca^{2+} channels at therapeutic concentrations. Curiously another compound valproate, used in absence seizures, has no effect on T-type Ca^{2+} channels; it one of several anticonvulsant drugs for which the mode of action is unknown.

P5 PARKINSON'S DISEASE

Key Notes

Symptoms and etiology of Parkinson's disease (PD)

PD is a hypokinetic disorder characterized by tremor, rigidity, akinesia (difficulty in initiating movements) and bradykinesia (slowness of movements). Rare familial PD has been linked to mutations in α-synuclein, a component of Lewy bodies. Acquired PD is associated with head trauma, brain cancers and possibly environmental toxins.

Neuropathology

The major pathology in PD is the death of large numbers of dopaminergic neurons in the substantia nigra which project to the striatum. Other monoamine neurons are also lost. Lewy bodies are cytoplasmic inclusions containing α-synuclein, found in afflicted cells in PD, and other neurodegenerative diseases. Cell death in PD is caused by free radical reactions to which the substantia nigra is particularly vulnerable.

The MPTP model of PD

The pyridine, MPTP, causes rapid, full-blown PD in humans and monkeys. It has proved useful in studying how movement deficits arise in PD. MPTP crosses the blood–brain barrier and is oxidized to the toxic metabolite MPP^+ which enters dopaminergic neurons, killing them via production of free radicals. In monkeys, MPTP-evoked PD reduces activity in the thalamocortical circuits that enable movement. The tremor arises from oscillations in the activity of neurons in the thalamus.

Treatment of PD

The key drug in the treatment of PD is L-DOPA, which crosses the blood–brain barrier and is converted to dopamine by dopamine-β-carboxylase. Although effective early on, after several years it becomes less useful and the majority of patients develop dyskinesia, a disabling hyperkinetic disorder. Currently, dopamine receptor antagonists, monoamine oxidase inhibitors and muscarinic cholinergic receptor antagonists also have a role in PD treatment. Future pharmacological strategies include the use of glutamate or adenosine receptor antagonists. Surgical approaches involve selective lesions of the globus pallidus pars interna, subthalamus or ventrolateral thalamus, and the transplant of dopaminergic cells, harvested from the midbrain of human fetuses, into the striatum.

Related topics

- Dopamine (D4)
- Cortical control of voluntary movement (K1)
- Anatomy of the basal ganglia (K7)
- Basal ganglia function (K8)

Symptoms and etiology of Parkinson's disease (PD)

The most common hypokinetic disorder, **Parkinson's disease (PD)**, causes a 4–7-Hz **tremor**, especially of limbs, which reduces with intentional activity, an increase in muscle tone, **rigidity** of all limb muscles (unlike the selective rigidity of spasticity), a difficulty in initiating movements (**akinesia**) and movements that are made are slow (**bradykinesia**). The sufferer has a mask-like facial

expression with a very low blink rate, walks with a bent back and shuffling gait and if unbalanced may not recruit righting reflexes quickly enough to prevent themselves from falling.

Although PD is generally **idiopathic** (of unknown cause), rare familial cases occur. By studying family pedigrees several genes linked to familial PD have been identified. One of these families has point mutations, in afflicted individuals, in the gene coding for **α-synuclein**, a component of Lewy bodies (see below). There is also likely to be a genetic component to idiopathic PD. There is an increased incidence of the disease in close relatives of patients with PD, including a 53% concordance rate for monozygotic (identical) twins of PD patients showing dopaminergic dysfunction on PET scan. Sporadic PD occurs with head trauma or tumors that damage the midbrain and there is epidemiological evidence that it may result from exposure to environmental toxins; very severe PD follows ingestion of the pyridine compound, MPTP, which is chemically related to some herbicides.

Neuropathology

The defining pathology of PD is the death of large numbers of neurons in the substantia nigra pars compacta (SNpc) that give rise to the nigrostriatal pathway. Other dopaminergic neurons in the midbrain also die, but not to the same extent, and loss of noradrenergic cells in the locus ceruleus and cholinergic neurons in the basal forebrain is also seen. Bilateral destruction of the SNpc in monkeys with the neurotoxin 6-OHDA causes rigidity and bradykinesia but not tremor. The cell death in PD is accompanied by the appearance of **Lewy bodies** in the cytoplasm of neurons, particularly in the SNpc. Lewy bodies are 5–25 μm in diameter and consist of a core of an abnormally folded protein, α-synuclein, surrounded by a halo of **ubiquitin**, a small protein present in all eukaryotic cells that tags proteins for destruction. Lewy bodies also feature in other neurodegenerative diseases and are probably a vehicle for sequestering abnormal proteins.

There is a normal age-related loss of neurons from the SN, at a rate of about 5% per decade, but a 50% loss (associated with a 70–80% fall in striatal dopamine) is necessary to account for the onset of symptoms, so normal losses could only account for a very late onset of the disease. PET scans show that the rate of cell loss in PD is massively accelerated (up to 12% per year). This suggests that the disorder causing PD starts only about 5 years before symptoms appear.

The vulnerability of cells in PD correlates with their neuromelanin content (a dark pigment that accumulates with age). The SNpc has the greatest number of pigmented cells of any nucleus (about 90%) and hence suffers most. The significance of neuromelanin is that it binds iron and this metal contributes to the mechanism of cell death in PD.

Cell death in PD is caused by reactive oxygen species (*Fig. 1*). Normally the **superoxide anion** ($^{\bullet}O_2^-$) is converted by **superoxide dismutase** (**SOD**) to hydrogen peroxide which is subsequently reduced to water by **glutathione peroxidase** (**GP**). However, in the SNpc of PD concentrations of glutathione are less than half that in the normal SN while the amount of iron bound to neuromelanin is greater. This promotes the **Fenton reaction** which converts hydrogen peroxide to give the highly toxic **hydroxyl radical** ($^{\bullet}OH$).

The MPTP model of PD

The chance discovery in 1982 that an illicitly manufactured pyridine **MPTP** caused very severe PD in a group of heroin addicts who ingested it, has led to a useful animal model of PD. In monkeys, MPTP causes a full-blown PD.

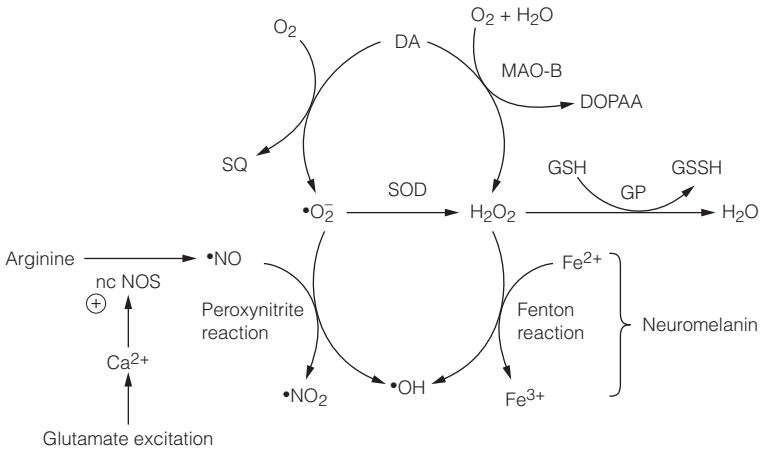


Fig. 1. Free radical reactions in Parkinson's disease. DA, dopamine; DOPAA, 3,4-dihydrophenyl-acetaldehyde; GP, glutathione peroxidase; GSH and GSSH, reduced and oxidized glutathione; ncNOS, neuronal isoform of nitric oxide synthase; SOD, superoxide dismutase; SQ, semiquinone. See text for details.

Following injection, MPTP crosses the blood–brain barrier and is taken up by astrocytes where it is oxidized to a metabolite MPP^+ by glial monoamine oxidase (MAO-B). MPP^+ is subsequently taken up via the specific dopamine transporter into dopaminergic neurons where it is responsible for the toxic effects of MPTP.

MPP^+ inhibits mitochondrial respiration, which depletes ATP levels, and generates superoxide anions, both of which kill the dopamine cells. Monkeys with MPTP-induced PD show increased firing of neurons in the globus pallidus pars interna (GPi), and the subthalamic nuclei, and reduced firing of neurons in the globus pallidus pars externa (GPe). The effect of this is to reduce the activation of the thalamocortical circuits which enable movement. This accounts for the akinesia and bradykinesia.

In monkeys, rigidity is probably due to abnormally active long loop cortical stretch reflexes. Patients with PD are unable to suppress stretch reflexes when attempting to change posture. For example, when instructed to sit they show inappropriate co-contraction of back and limb muscles, which is normally the required postural set for standing. The neurophysiological cause of the tremor is unclear, but the tremor correlates with a 3–6-Hz oscillation of neural activity in the ventrolateral thalamus. This provides the rationale for surgically lesioning this region of the thalamus to alleviate the tremor of PD (see below).

Treatment of PD

Both pharmacological and surgical treatments for PD exist. L-DOPA is the mainstay of drug therapy. It is the immediate precursor for dopamine, crosses the blood–brain barrier (BBB) and probably works by being taken up by the remaining dopaminergic terminals where it is converted to dopamine. Since L-DOPA is rapidly metabolized by DOPA decarboxylase in the periphery it is usual to administer it with a DOPA decarboxylase inhibitor that does not cross the BBB. This improves uptake into the brain and reduces peripheral side effects. L-DOPA is effective in the early years of treatment but becomes less so with long-term treatment and 80% of patients develop dyskinesia, a hyperkinetic disorder, which resembles the motor dysfunction of Huntington's disease.

Other current pharmacological approaches include:

- dopamine receptor agonists;
- MAO-B inhibitors, which retard the catabolism of dopamine;
- muscarinic acetylcholine receptor antagonists, which presumably work by reducing excitation of the GABAergic striatal neurons by the large aspiny cholinergic interneurons.

Other approaches might be fruitful in the future include the following.

- Antioxidants have been used, on the premise that PD is due to free radical damage, though a large multicenter clinical trial has provided no evidence that high doses of vitamin E are beneficial.
- The finding of increased firing of glutamatergic neurons in the subthalamic nucleus in MPTP-induced PD has prompted clinical trials of glutamate receptor antagonists.
- The activity of the GABA-enkephalin containing striatal neurons of the indirect pathway is enhanced in PD by the lack of dopamine inhibition. However these cells also experience recurrent inhibition by GABAergic interneurons (*Fig. 2*). Release of GABA from these interneurons is reduced by presynaptic adenosine ($A2_A$) receptors. Hence antagonists of $A2_A$ receptors could have a role in treatment of PD by increasing the recurrent inhibition of the indirect pathway. An $A2_A$ antagonist improves the motor disability of marmosets with MPTP-induced PD.

Surgical approaches include lesions and tissue transplants. Discrete lesions of the subthalamic nucleus, which reduces the excessive excitatory drive of the indirect pathway on the globus pallidus (GPi), or lesions of the GPi itself, ameliorate the cardinal signs of Parkinson's disease in MPTP-treated monkeys. In humans, pallidotomy – removal of the posterior (sensorimotor) portion of the GPi – is proving highly effective. Lesions of the ventrolateral thalamus are particularly successful in reducing tremor, but less good at relieving rigidity or bradykinesia. In transplant therapy, human fetal midbrain dopaminergic cells are injected into the striatum. PET scans using [^{18}F] fluorodopa, which is taken

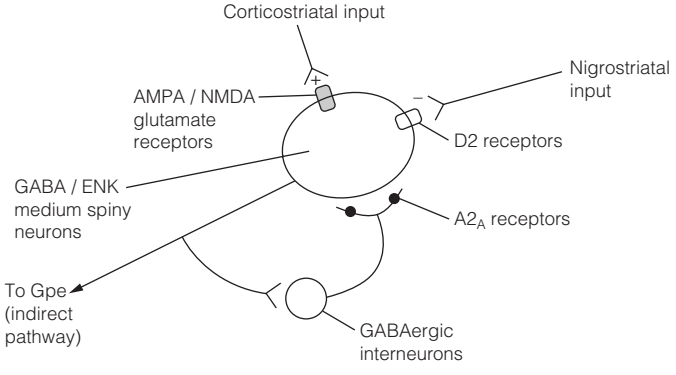


Fig. 2. Recurrent inhibition on indirect pathway striatal neurons is suppressed by adenosine ($A2_A$) receptors. Nigrostriatal input is lost in Parkinson's disease.

up by dopaminergic cells, show that the transplanted cells survive and even increase their dopaminergic activity with time. The majority of patients show partial improvement. Ethical concerns (material from 8–10 fetuses is needed for each transplant) are driving research into the use of dopamine cells from other sources, and the development of appropriate stem cell lines.

P6 Alzheimer's disease

Key Notes

Features of Alzheimer's disease (AD)

AD is the commonest dementia of the old and its incidence increases with age. AD is insidious, progressive and causes defects in memory, cognition, attention and motivation. Most AD sufferers also have symptoms of Parkinson's disease and neurodegeneration of the substantia nigra.

Neuropathology of AD

Patients with AD have massive shrinkage of the cortex and subcortical structures. Two characteristic lesions are found, particularly in the neocortex, hippocampus and amygdala. Neuritic plaques are extracellular deposits of β-amyloid peptide surrounded by degenerating neurites and glial cells activated by inflammatory processes. Diffuse plaques lacking damaged neurites and reactive glia are found in AD in brain regions not implicated in the disease, and in the brains of normal old people. Neurofibrillary tangles are cytoplasmic bundles of paired helical filaments made from an abnormal, hyperphosphorylated state of tau, a protein normally associated with microtubules. The density of tangles correlates with the severity of the dementia. Death of glutamatergic cells in the cortex and loss of several transmitter systems that project to the cortex (e.g. cholinergic pathways from the basal forebrain to cortex and septum to hippocampus) dominate the pathology.

Familial AD

Rare, familial AD is linked to mutations in four genes and much idiopathic AD might be accounted for by polymorphisms of an apolipoprotein gene. Mutations of the gene for amyloid precursor protein (located on chromosome 21) which gives rise to Aβ are linked to early-onset AD in several families and in Down syndrome. Two genes that code for presenilins have numerous mutations linked to early-onset AD, and *tau* gene mutations are seen in dementia with Parkinson's disease. There are three common alleles of apolipoprotein E (apoE), one of which is a clear risk factor for developing late-onset idiopathic AD. Polymorphisms of the apoE gene could account for up to 90% of idiopathic AD.

Amyloid cascade hypothesis

Two pathways cleave amyloid precursor protein (APP), one of which (sequential cleavage by β- then γ-secretase) leads to the formation of β-amyloid (Aβ). Aβ is normally produced in the brain where it appears to be an antioxidant involved in metal ion homeostasis. APP is over-expressed in Down syndrome, following brain ischemia or head injury, and by mutations of the APP or presenilin (γ-secretase) genes. Aβ aggregates and is deposited as amyloid in plaques. Whether amyloid is itself pathogenic or is sequestered as a protective response is not clear. Soluble Aβ may be neurotoxic by liberating H_2O_2 on binding Cu^{2+}. Although some evidence suggests that Aβ causes hyperphosphorylation of the microtubule-associated protein tau, mutations of the *tau* gene are linked to dementias in which neurofibrillary tangles are found extensively in the frontal and temporal cortex, but in which there are few plaques. Hence the relationship between Aβ and tangles is unclear.

Pharmacological intervention in AD

Acetylcholinesterase inhibitors, which potentiate the effects of ACh released in the cortex from cholinergic neurons, can produce modest improvements in some patients. Anti-inflammatory drugs can arrest cognitive decline, and a chelating agent which reduces amyloid and prevents the resulting soluble Aβ from being toxic may prove useful. However, real progress will require novel strategies, for example, the discovery of drugs to inhibit the enzymes that process APP.

Related topics Acetylcholine (D7) Parkinson's disease (P5)

Features of Alzheimer's disease

Alzheimer's disease (AD) is the commonest cause of **dementia** (insanity due to the loss of cognitive abilities) in the elderly. Its incidence increases with age so that while about 5% of people over 70 are sufferers, this becomes 20% of those over 80. It is a disease predominantly of the nervous system (although deficits of gut function are also seen) and should be distinguished from **multi-infarct dementia**, the second commonest cause of dementia, that is usually due to a succession of small cerebrovascular accidents.

AD is insidious, progresses unevenly, and at different rates in different individuals. Sudden deterioration may follow a stressful event. The earliest impairment is in semantic memory (particularly verbal), followed by failure to consolidate new long-term memories. Remote memory recall tends to be preserved until later. Deficits occur in cognition, attention and motivation (decreased appetite and libido) and AD sufferers are often depressed or frustrated. Two-thirds of AD patients also have symptoms of Parkinson's disease, with neurodegeneration of the substantia nigra and Lewy bodies. Eventually, AD sufferers can no longer care for themselves. Average life expectancy after diagnosis is about 5 years.

Neuropathology of AD

CAT scans of patients with AD show severe atrophy of cortical and subcortical regions with enlargement of the cerebral ventricles and brain sulci. Brain weight is reduced by 30–40%. Two lesions, characteristic of AD, are found predominantly in the neocortex (often concentrated in frontal and temporal regions), hippocampus and amygdala.

1. **Neuritic plaques** are spherical extracellular lesions 5–150 μm in diameter. They consist largely of deposits of an insoluble form of **β-amyloid peptide** (**A**β) called **amyloid**, which also contains reduced metal ions (e.g. Cu^{2+}), apolipoprotein E (apoE), and components of the complement cascade (proteins of the immune system). Surrounding and within this core are dystrophic (swollen, damaged and degenerating) neurites, together with microglia and astrocytes activated by cytokines resulting from inflammatory processes. In addition, there are diffuse plaques which contain Aβ but lack dystrophic neurites and reactive glial cells. These are located in regions not implicated in AD (e.g. thalamus and cerebellum) and are seen in the brains of old, normal humans. The significance of these diffuse plaques to AD is not known.
2. **Neurofibrillary tangles** are cytoplasmic bundles of paired 10 nm filaments that are twisted into a helix to form **paired helical filaments** (**PHFs**). Tangles are found in large numbers in neurons (often pyramidal cells) in the

entorhinal cortex, hippocampus, amygdala and in many areas that project to them. They are often seen in the dystrophic neurites of plaques. There is a correlation between the density of tangles and severity of the dementia. Tangles appear in other neurodegenerative diseases where plaques are absent, so these two lesions are independent of each other. PHFs are composed of **tau**, normally a soluble cytoplasmic microtubule-associated protein, present in an abnormal highly phosphorylated, insoluble form.

The cells that die in greatest numbers in AD are glutamatergic pyramidal cells of the cortex (including the hippocampus), cholinergic cells in the septohippocampal pathway and in the nucleus basalis of Meynert (NBM) and noradrenergic and serotonergic projections from the locus ceruleus (LC) and raphe nuclei (RN) respectively. Nigrostriatal dopaminergic cell loss is responsible for the Parkinson's disease associated with AD. GABAergic and peptide neurotransmitter systems are spared. It has been suggested that the neurodegenerative changes start in the olfactory bulb (AD sufferers have a defective sense of smell), where neurogenesis normally continues throughout life, and spreads to the entorhinal cortex and hippocampus. Subsequently, cortico-cortical association axons die, so regions of cortex become disconnected from each other. Subcortical regions (NBM, LC, RN) deprived of their normal cortical targets degenerate.

Familial AD

The great majority of AD cases are idiopathic, but there are a number of rare familial types of AD linked to mutations of four genes. These provide important clues to causes of the disease. Idiopathic AD might also have a genetic predisposition since it is linked to polymorphisms of an apolipoprotein gene.

Although the majority of AD cases are late onset (>60 years), rarely the illness may begin as early as 30 years old. This **early-onset AD** is linked to mutations on three genes. **Amyloid precursor protein** (**APP**) is a membrane glycoprotein that gives rise to the β-amyloid peptides deposited in plaques. Its gene is on chromosome 21. Linkage between the APP gene and rare early-onset AD in several families has been established. Notably, early-onset AD is seen in **Down syndrome**, which is caused by the presence of an extra copy of chromosome 21 (**trisomy 21**); this is evidence that excess APP leads to AD. Transgenic mice engineered to have copies of mutant APP genes are proving useful animal models of AD since they express aspects of the abnormal pathology.

Two genes (on chromosome 14 and 1) linked to early-onset AD code for the closely related **presenilin 1** and **presenilin 2**, large membrane proteins in the endoplasmic reticulum and Golgi. Over 50 missense mutations have been identified in the presenilins that cause an aggressive early-onset AD. Missense mutations in the gene (on chromosome 17) coding for tau is linked to dementia with Parkinson's disease in over 12 families.

Late-onset idiopathic AD is linked with polymorphisms of apolipoprotein E (apoE), a molecule involved in recycling cholesterol during membrane repair. The apoE gene, on chromosome 19, has three common alleles, ε2, ε3 and ε4. The ε4 allele is a risk factor for AD in that there is a high association between AD and possession of the ε4 allele. In contrast, the ε2 allele confers some protection. Individuals homozygous for ε4 are not only more likely to develop AD than other genotypes but if they do they become sick earlier. It is estimated that 60–90% of total AD cases can be accounted for by apoE genotype. The disease-promoting effect of inheriting the ε4 isoform seems to be related to the fact that

it enhances the aggregation of β-amyloid that accumulates in plaques (see below).

Amyloid cascade hypothesis

β-amyloid peptides are derived from amyloid precursor protein (APP) by the action of proteases. Since all the mutations in APP that give rise to AD are clustered around the cleavage sites in APP where these enzymes act, it is postulated that abnormal processing of the APP molecule leads to inappropriate β-amyloid production and this underlies Alzheimer's disease.

Amyloid precursor protein is processed by two pathways, only one of which gives rise to the two β-amyloid peptides, with 40 and 42 amino acids, that are deposited in plaques. Only a minority of APP molecules are processed under normal circumstances. APP contains a 23-residue hydrophobic region near its C terminal that anchors it in membranes of the endoplasmic reticulum (ER), Golgi and the cell membrane. In the first pathway, a protease called **α-secretase** cuts APP (see *Fig. 1*) producing a large soluble N terminal product, sAPPα that is released into the lumens of the ER and Golgi or from the cell surface. This leaves behind a C terminal fragment (αCTF). In the second pathway, **β-secretase** cuts APP at a site closer to the N terminal than α-secretase to release sAPPβ. The remaining membrane bound C terminal component (βCTF) is now cleaved by **γ-secretase** to produce the β-amyloid peptides (Aβ).

One way to understand what is amiss in AD is to know what roles APP and β-amyloid serve. Unfortunately the function of APP is not clear; it may be a protease inhibitor. Moreover β-amyloid peptides are produced by healthy brain so presumably it has a normal function. One possibility is that Aβ is an antioxidant involved in copper homeostasis. It has a high-affinity binding site for Cu^{2+}, Fe^{2+} and Zn^{2+} that resembles the metal binding site on superoxide dismutase, and transgenic mice that over-express APP have constitutively lower brain Cu^{2+} levels while APP knockout mice have elevated Cu^{2+} concentrations.

Whatever their functions, the prevailing assumption is that *excessive* APP expression and Aβ production is a necessary condition for AD, and that in excess Aβ is converted to amyloid in plaques. Plaque formation is viewed as either pathogenic itself, or an attempt to ameliorate pathology that results from the presence of excess soluble Aβ.

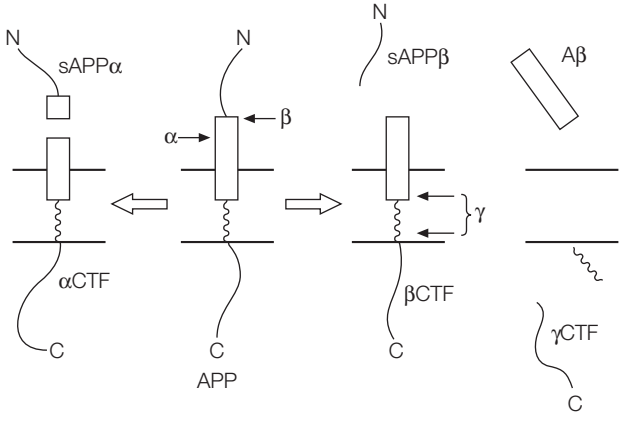


Fig. 1. Cleavage of amyloid precursor protein (APP) by secretases.

Several factors favor the aggregation of soluble Aβ into amyloid including, higher Aβ concentrations, Zn^{2+}, and the ε4 isoform of apoE. For example, crossing transgenic mice that have mutant forms of the APP gene with knockout mice that lack the apoE gene results in offspring which develop far less plaque formation than ordinary APP transgenic mice.

The mutations of the APP gene in familial early-onset AD enhance the cleavage of APP by β-secretase or γ-secretase, but are also likely to increase the tendency for the mutant peptides to aggregate. Hence, APP mutations increase the production and deposition of Aβ.

Brains of individuals with presenilin mutations, or animals genetically engineered to over-express presenilins, have higher concentrations of Aβ. This led to the discovery that presenilins are components of γ-secretases.

How might abnormal amounts of β-amyloid lead to AD? One possibility is that it is neurotoxic. When soluble Aβ binds Cu^{2+} or Fe^{2+} it produces H_2O_2. By contrast when the peptide binds Zn^{2+} it does not generate H_2O_2 and is not toxic. Studies that have measured the level of oxidative damage in AD brain *in vitro* show an inverse relationship with the amount of amyloid. This implies that deposition of amyloid could be a protective response to oxidative damage

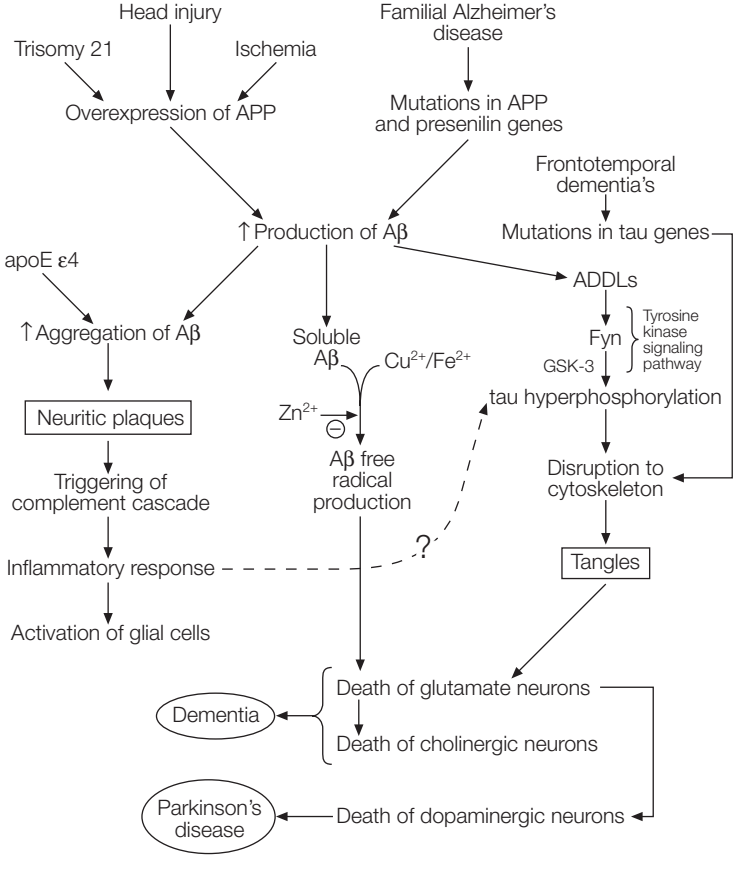


Fig. 2. A summary of a possible progression of events in Alzheimer's disease. APP, amyloid precursor protein; Aβ, β-amyloid peptides; apoE, apolipoprotein.

caused by metal-bound soluble Aβ. There is a marked rise in brain Cu^{2+} and Fe^{2+} in later life and it has been suggested that this can disrupt a normal role of Aβ in the homeostatic regulation of metal ion concentrations.

A major problem in AD research is working out the relationship between Aβ and tangles. There is evidence that overproduction of Aβ leads to hyperphosphorylation of tau. In this state tau cannot bind microtubules but instead aggregates to form tangles which kill neurons. Although the relationship between tau, tangles and dementia is clear in the **frontotemporal dementias** (**FTDs**), in which mutations of the *tau* gene are associated with extensive tangle formation restricted to the frontotemporal cortex, the link between Aβ and tangles generally is far from clear. Patients with FTDs do not have plaques, the cerebellum of AD patients have plaques but no tangles and mice genetically engineered to have the same APP mutations seen in FAD have more Aβ and form plaques, but they do not have abnormally phosphorylated tau or tangles, neither do they lose neurons. In animal studies it has not proved possible to establish unambiguously that Aβ induces tangles. The culprit may not be Aβ itself but **Aβ-derived diffusible ligands** (**ADDLs**), small soluble oligomers that are derived from Aβ. The concentration of ADDLs correlates well with synapse loss, the most reliable measure of cognitive decline in AD, and ADDLs inhibit long-term potentiation. ADDLs might impair memory and kill neurons by activating a tyrosine kinase signaling pathway. *Fig. 2* summarizes the progression of events that might occur in AD.

Pharmacological intervention in AD

Acetylcholinesterase inhibitors (e.g., donezepil) are presumed to exert their therapeutic effect by potentiating the action of acetylcholine released from cholinergic neurons in the basal forebrain and septum, and of course depends on there being some functioning cholinergic neurons remaining. They have not proved particularly successful, probably because most of the cognitive deficits result from the loss of glutamatergic cortical pyramidal cells. A double-blind trial of the non-steroidal anti-inflammatory drug, indomethacin, showed that cognitive decline is halted in treated patients, which implies that the inflammatory component to AD is worth targeting. Active vaccination against Aβ clears plaques in animals and humans, but whether this improves cognition has yet to be shown, and patient trials have been stopped because brain inflammation occurred in 5% of cases. The chelating agent clioquinol has produced a 100-fold decrease in amyloid burden in transgenic mice expressing a human mutant APP, with consequent improvements in alertness and motor activity. Clioquinol is thought to work by causing the dissociation of Cu^{2+} and Zn^{2+} from Aβ. Loss of Cu^{2+} prevents it from forming neurotoxic H_2O_2, while loss of Zn^{2+} renders it soluble. Clioquinol is in phase II clinical trials. Other therapeutic approaches to try include:

- antioxidants and free radical scavengers to target the downstream neurotoxic effects of Aβ;
- neurotrophins to promote the survival of injured cells;
- inhibitors of β- or γ-secretases to reduce Aβ production.

Further reading

There are many comprehensive textbooks of neuroscience and no single volume is likely to satisfy all needs. Different readers will prefer different textbooks depending, for example, on their prior learning, so I do not think it helpful to recommend one over another. Rather I have listed some of the leading books which, experience shows, students find useful.

General reading

Bear, M.F., Connors, B.W., Paradiso, M.A. (1996) *Neuroscience: Exploring the Brain.* Williams and Wilkins, Baltimore.

Kandel, E.R., Schwartz, J.H., Jessel, T.M. (2000) *Principles of Neural Science*, 4th Edn. Elsevier, New York.

Purves, D. *et al* (eds) (2001) *Neuroscience*, 2nd Edn. Sinauer Associates Inc., Sunderland, MA.

Rosenzweig, M.R., Breedlove, S.M., Watson, N.V. (2004) *Biological Psychology: An Introduction to Behavioural and Cognitive Neuroscience*, 4th edn. Sinauer, Sunderland, MA.

Zigmond, M.J., Bloom, F.E., Landis, S.C., Roberts, J.L., Squire, L.R. (1999) *Fundamental Neuroscience.* Academic Press, San Diego, CA.

More advanced reading

The following selected articles and books are recommended for those who wish to know more about specific subjects. Some of the papers discuss ideas reviewed in this book, others deal with material which could not be included for lack of space. Inevitably these sources vary in difficulty; some will be accessible to first year students, most will be more suitable for students later in their studies.

Section A

Attwell, D., Iadecola, C. (2002) The neural basis of functional brain imaging signals. *Trends Neurosci*. **25**, 621–625.

Berns, G.S. (1999) Functional neuroimaging. *Life Sci*. **65**, 2531–2540

Fitzgerald, M.J.T., Folan-Curran, J. (2002) *Clinical Neuroanatomy*, 4th edn. Saunders, London.

Nedergaard, M., Ransom, B., Goldman, S.A. (2003) New roles for astrocytes: redefining the functional architecture of the brain. *Trends Neurosci*. **26**, 523–530.

Pappiat, C., Attwell, D. (2004) Feeding the brain. *Nature* **431**, 137–138.

Pardridge, W.M. (1998) *Introduction to the blood–brain barrier*. Cambridge University Press, Cambridge.

Ransom, B., Behar, T., Nedergaard, M. (2003) New roles for astrocytes (stars at last). *Trends Neurosci*. **26**, 520–522.

Section B

Hille, B. (2001) *Ionic Channels of Excitable Membranes*, 3rd edn. Sinauer Associates, Sunderland, MA.

Johnston, D., Wu, S.M-S. (1995) *Foundations of Cellular Neurophysiology*. The MIT Press, Cambridge, MA.

Yellen, G. (2002) The voltage-gated potassium channels and their relatives. *Nature* **419**, 35–42.

Section C

Ehlers, M.D. (2002) Molecular morphogens for dendritic spines. *Trends Neurosci.* **25**, 64–67.

Jonas, P., Bischofberger, J., Fricker, D., Miles, R. (2004) Fast in, fast out – temporal and spatial signal processing in hippocampal interneurons. *Trends Neurosci*. **27**, 30–40.

Ludwig, M., Pittman, Q.J. (2003) Talking back: dendritic neurotransmitter release. *Trends Neurosci*. **26**, 255–261.

Newman, E.A. (2003) New roles for astrocytes: regulation of synaptic transmission. *Trends Neurosci*. **26**, 536–542.

Revest, P., Longstaff, A. (1998) *Molecular Neuroscience.* Bios Scientific Publications, Oxford, UK.

Seeburg, P.H., Osten, P. (2003) A thorny issue. *Nature* **424**, 627–628.

Tong-Wey, K., Bellen, H.J. (2003) Synaptotagmin I, a calcium sensor for neurotransmitter release. *Trends Neurosci*. **26**, 413–422.

Walmsey, B., Alvarez, F.J., Fyffe, R.E.W. (1998) Diversity of structure and function at mammalian central synapses. *Trends Neurosci*. **21**, 81–88.

Section D

Baraban, S.C., Tallent, M.K. (2004) Interneuronal neuropeptides – endogenous regulators of neuronal excitability. *Trends Neurosci*. **27**, 135–142.

Cooper, J.R., Bloom, F.E., Roth, R.H. (2002) *The Biochemical Basis of Neuropharmacology*, 8th edn. Oxford University Press, Oxford, UK.

Kuner, T., Seeburg, P.H., Guy, R.H. (2003) A common architecture for K^+ channels and ionotropic glutamate receptors? *Trends Neurosci*. **26**, 27–32.

Perry, E., Walker, M., Grace, J., Perry, R. (1999) Acetylcholine in mind: a neurotransmitter correlate of consciousness. *Trends Neurosci*. **22**, 273–280.

Section E

Abeles, M. (1991) *Corticonics: Neural Circuits of the Cerebral Cortex.* Cambridge University Press, Cambridge, UK.

Singer, W. (2001) Consciousness and the binding problem. *Ann. NY Acad. Sci*. **929**, 123–146.

Section F

Berlucchi, G., Aglioti, S. (1997) The body in the brain: neural bases of corporeal awareness. *Trends Neurosci*. **20**, 560–564.

Craig, A.D. (2003) A new view of pain as a homeostatic emotion. *Trends Neurosci*. **26**, 302–307.

Derbyshire, S.W.G. (2002) Measuring our natural painkiller *Trends Neurosci*. **25**, 67–69.

Ji-Sheng, H. (2003) Acupuncture: neuropeptide release produced by electrical stimulation of different frequencies. *Trends Neurosci*. **26**, 17–22.

McClesky, E.W. (2003) New player in pain. *Nature* **424**, 729–730.

Wall, P. (1999) *Pain: The Science of Suffering*. Weidenfeld and Nicholson, London.

Section G

Billock, V.A., Tsou, B.H. (2004) What do catastrophic visual binding failures look like? *Trends Neurosci*. **27**, 84–89.

Bullier, J., Novak, L.G. (1995) Parallel versus serial processing: new vistas on the distributed organization of the visual system. *Curr. Opin. Neurobiol*. **5**, 497–503

Crick, F., Koch, C. (1995) Are we aware of neural activity in primary visual cortex? *Nature* **375**, 121–123.
Sharpe, L.T., Stockman, A. (1999) Rod pathways: the importance of seeing nothing. *Trends Neurosci*. **22**, 497–504.
Zeki, S. (2003) Improbable areas in the visual brain. *Trends Neurosci*. **26**, 23–26.

Section H

Cohen, Y.E., Knudsen, E.I. (1999) Maps versus clusters: different representations of auditory space in the midbrain and forebrain. *Trends Neurosci*. **22**, 128–135.
Hudspeth, A.J. (1997) Mechanical amplification of stimuli by hair cells. *Curr. Opin. Neurobiol.* **7**, 480–486.
McAlpine, D., Grothe, B. (2003) Sound localization and delay lines – do mammals fit the model? *Trends Neurosci*. **26**, 347–350.
Pollak, G.D., Burger, R.M., Klug, A. (2003) Dissecting the circuitry of the auditory system. *Trends Neurosci*. **26**, 33–39.
Scott, S.K., Johnsrude, I.S. (2003) The neuroanatomical and functional organization of speech perception. *Trends Neurosci*. **26**, 100–107.

Section I

Friedrich, F.R. (2002) Real time odor representations. *Trends. Neurosci*. **25**, 487–489.
Mombaerts, P. (1999) 7TM proteins as odorant and chemosensory receptors. *Science* **286**, 707–711.
Mori, K., Nagao, H., Yoshihara, Y. (1999) The olfactory bulb: coding and processing of odor molecule information. *Science* **286**, 711–715.
Schoppa, N.E., Urban, N.N. (2004) Dendritic processing within olfactory bulb circuits. *Trends Neurosci*. **27**, 501–506.
Smith, D.V., Margolis, F.L. (1999) Taste processing: whetting our appetites. *Curr. Biol*. **9**, 453–455.
Smith, D.V., St John, S.J. (1999) Neural coding of gustatory information. *Curr. Opin. Neurobiol*. **9**, 427–435.

Section J

Blake, D.J., Kroger, S. (2000) The neurobiology of Duchenne muscular dystrophy: learning lessons from muscle? *Trends Neurosci*. **23**, 92–99.
Clarac, F., Cattaert, D., Le Ray, D. (2000) Central control components of a 'simple' stretch reflex. *Trends Neurosci*. **23**, 199–208.
Poppele, R., Bosco, G. (2003) Sophisticated spinal contributions to motor control. *Trends Neurosci*. **26**, 269–276.
Rowe, J.B., Frackowiak, R.S. (1999) The impact of brain imaging technology on our understanding of motor function and dysfunction. *Curr. Opin. Neurobiol*. **9**, 728–734.

Section K

Botvinick, M. (2004) Probing the neural basis of body ownership. *Science* **305**, 782–783.
Capaday, C. (2003) The special nature of human walking and its neural control. *Trends Neurosci*. **26**, 370–376.
Grieve, K.L., Acuna, C., Cudeiro, J. (2000) The primate pulvinar nuclei: vision and action. *Trends Neurosci*. **23**, 35–39.
Ohyama, T., Nores, W.J., Murphy, M., Mauk, M.D. (2003) What the cerebellum computes. *Trends Neurosci*. **26**, 222–227.
Swanson, G. (ed.) (1998) Special issue: cerebellum – development, physiology and plasticity. *Trends Neurosci*. **21**, 367–418.

Section L

Hadley, M.E. (2000) *Endocrinology*, 5th Edn. Prentice-Hall International, New Jersey.

Herman, J.P., Cullinan, W.E. (1997) Neurocircuitry of stress: central control of the hypothalamo-pituitary-adrenocortical axis. *Trends Neurosci.* **20**, 78–84.

Johnson, M., Everitt, B. (2000) *Essential Reproduction*, 5th edn. Blackwell Scientific Publications, Oxford, UK.

Jordon, D. (ed.) (1997) *Central Nervous Control of Autonomic Function*. Harwood Academic, Amsterdam.

Kalin, N.H. (1993) The neurobiology of fear. *Sci. Amer.* **208**, 54–60.

Levay, S. (1993) *The Sexual Brain*. MIT Press, Cambridge, MA.

Section M

Balaban, E. (2004) Why voles stick together. *Nature* **429**, 711–712.

Berridge, K.C., Robinson, T.E. (2004) Parsing reward. *Trends Neurosci.* **27**, 507–513.

Berson, D.M. (2003) Strange vision: ganglion cells as circadian photoreceptors. *Trends Neurosci.* **26**, 314–320.

Carlezon, W.A., Nestler, E.J. (2002) Elevated levels of GluR1 in the midbrain: a trigger for sensitization to drugs of abuse. *Trends Neurosci.* **25**, 610–615.

Damasio, A. (1999) *The Feeling of What Happens*. Heinemann, London.

Elmkuist, J.K., Flier, J.S. (2004) The fat–brain axis enters a new dimension. *Science* **304**, 63–64.

Kalivas, P.W., Nakamura, M. (1999) Neural systems for behavioral activation and reward. *Curr. Opin. Neurobiol.* **9**, 223–227.

Maren, S. (1999) Long-term potentiation in the amygdala: a mechanism for emotional learning and memory. *Trends Neurosci.* **22**, 561–566.

Section N

Campbell, K., Gotz, M. (2003) Radial glia: multi-purpose cells for vertebrate brain development. *Trends Neurosci.* **25**, 235–238.

Edelman, G. (1989) *Neural Darwinism.* Oxford University Press, Oxford, UK.

Eisen, J.S. (1999) Patterning motoneurons in the vertebrate nervous system. *Trends Neurosci.* **22**, 321–326.

Goldman, S. (2003) Glia as neural progenitor cells. *Trends Neurosci.* **26**, 590–596.

Horner, P.J., Palmer, T.D. (2003) New roles for astrocytes: the nightlife of an astrocyte. *Trends Neurosci.* **26**, 597–603.

Knoll, B., Drescher, U. (2002) Ephrin-As as receptors in topographic projections. *Trends Neurosci.* **25**, 145–149.

Marti, E., Boloventa, P. (2002) Sonic hedgehog in CNS development: one signal, multiple outputs. *Trends Neurosci.* **25**, 89–96.

Parnavelas, J.G. (2000) The origin and migration of cortical neurones: new vistas. *Trends Neurosci.* **23**, 126–131.

Pevny, L., Rao, M.S. (2003) The stem-cell menagerie. *Trends Neurosci.* **26**, 351–359.

Slezak, M., Pfrieger, F.W. (2003) New role for astrocytes: regulation of CNS synaptogenesis. *Trends Neurosci.* **26**, 531–535.

Wu, J., Saint-Jeannet, J-P., Klein, P.S. (2003) Wnt-frizzled signaling in neural crest formation. *Trends Neurosci.* **26**, 40–45.

Wurmser, A.E., Palmer, T.D., Gage, F.H. (2004) Cellular interactions in the stem cell niche. *Science* **304**, 1253–1255.

Section O

Bilkey, D.K. (2004) In the place space. *Science* **305**, 1245–1246.

Bliss, T.V.P., Collingridge, G.L. (1993) A synaptic model of memory: long-term potentiation in the hippocampus. *Nature* **361**, 31–39.

Lee, J.L.C., Everitt, B.J., Thomas, K.L. (2004) Independent cellular processes for hippocampal memory consolidation and reconsolidation. *Science* **304**, 839–843.

Murphy, G. (2003) Lost for words. *Nature* **425**, 340–342.

Nader, K. (2003) Memory traces unbound. *Trends Neurosci*. **26**, 65–72.

Soderling, T.R., Derkach, V.A. (2000) Postsynaptic protein phosphorylation and LTP. *Trends Neurosci*. **23**, 75–80.

Stern, P. (ed.) (2001) Sleep, dreams, and memory. *Science* **294**, 1047–1063.

Stern, P., Chin, G., Travis, J. (eds) (2004) Neuroscience: higher brain functions. *Science* **306**, 431–452.

Wilson, M.A., Tonagawa, S. (1997) Synaptic plasticity, place cells and spatial memory: study with second generation knockouts. *Trends Neurosci*. **20**, 102–106.

Section P

Bush, A. (2003) The metallobiology of Alzheimer's disease. *Trends Neurosci*. **26**, 207–214.

Campbell, P. (ed.) (1999) Neurological disorders. *Nature* **399** (suppl to no. 6738): A3–A45.

Dirnagl, U., Iadecola, C., Moskowitz, M.A. (1999) Pathobiology of ischemic stroke: an integrated approach. *Trends Neurosci*. **22**, 391–397.

Fortini, M.E. (2003) Double trouble for neurons. *Nature* **425**, 565–566.

Mattson, M.P. (2004) Pathways towards and away from Alzheimer's disease. *Nature* **430**, 631–639.

Nemeroff, C.B. (1998) The neurobiology of depression. *Sci. Amer*. **278**, 28–35.

Rao, S.D., Weiss, J.H. (2004) Excitotoxic and oxidative cross-talk between motor neurons and glia in ALS pathogenesis. *Trends Neurosci*. **27**, 17–23.

Staley, K. (2004) Epileptic neurons go wireless. *Science* **305**, 482–483.

INDEX