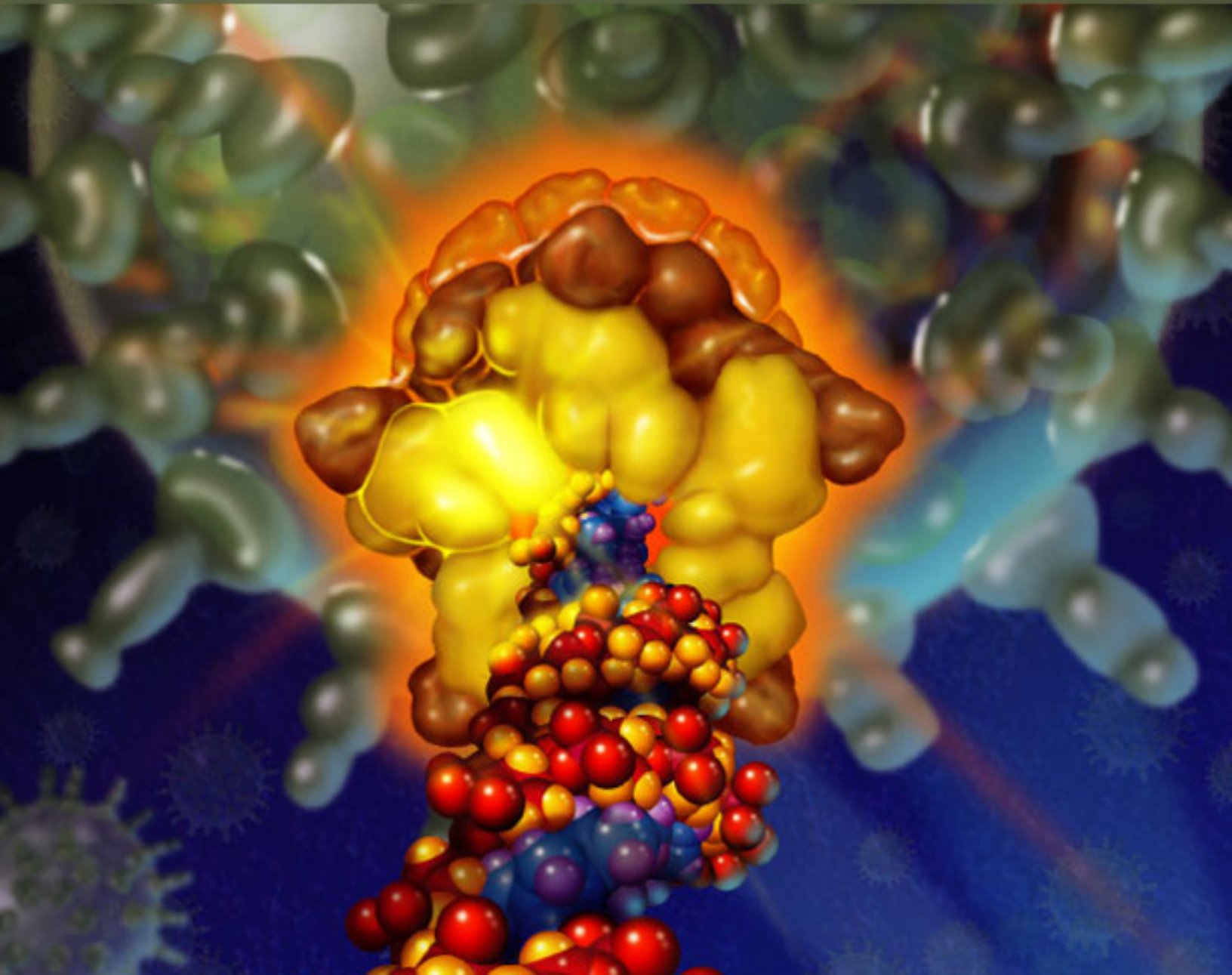
Biological Physics

Sally Challenor



First Edition, 2009

ISBN 978 93 80168 83 8

Published by:

Global Media
1819, Bhagirath Palace,
Chandni Chowk, Delhi-110 006
Email: globalmedia@dkpd.com

Table of Contents

Perception and Systems

Systems theory, in its broadest sense, is the interdisciplinary study of organisation within the context of a definitive system. More specifically, it is a framework by which one can analyse and/or describe any group of objects that work in concert to produce some result. This could be a single organism, any organisation or society, or any electro-mechanical, or informational artifact. While systems concepts had long been used in sociology and the area is often associated with cybernetics, systems theory as a technical and general academic area of study predominantly refers to the science of systems that resulted from Bertalanffy's General System Theory (GST) among others in initiating what became a project of systems research and practice. It was Margaret Mead and Gregory Bateson that developed interdisciplinary perspectives in systems theory (such as positive and negative feedback in the social sciences).

Ideas from systems theory have grown with diversified areas, exemplified by the work of Bela H. Banathy, ecological systems with Howard T. Odum, Eugene P Odum and Fritjof Capra, organisational theory and management with individuals such as Peter Senge, interdisciplinary study with areas like Human Resource Development from the work of Richard A. Swanson,

and insights from educators such as Debora Hammond. As a transdisciplinary, interdisciplinary and multiperspectival domain, the area brings together principles and concepts from ontology, philosophy of science, physics, computer science, biology, and engineering as well as geography, sociology, political science, psychotherapy (within family systems therapy) and economics among others. Systems theory thus serves as a bridge for interdisciplinary dialogue between autonomous areas of study as well as within the area of systems science itself.

In this respect, with the possibility of misinterpretations, Bertalanffy believed a general theory of systems "should be an important regulative device in science," to guard against superficial analogies that "are useless in science and harmful in their practical consequences."

Others remain closer to the direct systems concepts developed by the original theorists. For example, Ilya Prigogine, of the Centre for Complex Quantum Systems at the University of Texas, Austin, has studied emergent properties, suggesting that they offer analogues for living systems. The theories of Autopoiesis of Francisco Varela and Humberto Maturana are a further development in this field. Important names in contemporary systems science at the dusk of the Cold War include Russell Ackoff, Bela Banathy, Stanford Beer, Mandy Brown, Peter Checkland, Robert Flood, Fritjof Capra, Michael Jackson, and Werner Ulrich, among others.

History

Whether considering the first systems of written communication with Phoenician cuneiform to Mayan numerals, or the feats of engineering with the Egyptian pyramids, systems thinking in essence dates back to antiquity. Differentiated from Western rationalist traditions of philosophy, C. West Churchman often identified with the I Ching as a systems approach sharing a frame of reference similar to pre-Socratic philosophy and Heraclites (Hammond 2003). Bertalanffy traced systems concepts to the philosophy of G.W. von Leibniz and Nicholas of Cusa's *Coincidentia Oppositorum*. While modern systems are considerably more complicated, today's systems are embedded in history.

Systems theory as an area of study specifically developed following the World Wars from the work of Ludwig von Bertalanffy, Anatol Rapoport, Kenneth E. Boulding, William Ross Ashby, Margaret Mead, Gregory Bateson, C. West Churchman and others in the 1950s, specifically catalyst from the Macy conferences. Cognizant of advances in science that questioned classical assumptions in the organisational sciences, Bertalanffy's idea to develop a theory of systems began as early as the interwar period, publishing "An Outline for General Systems Theory" in the *British Journal for the Philosophy of Science*, by 1950. Where assumptions in Western science from Greek thought with Plato and Aristotle to Newton's Principia have historically influenced all areas from the social to hard sciences, the original theorists explored the implications of twentieth century advances in terms of systems.

Subjects like complexity, self-organisation, connectionism and adaptive systems had already been studied in the 1940s and 1950s. In fields like cybernetics, researchers like Norbert Wiener, William Ross Ashby, John von Neumann and Heinz von Foerster examined complex systems using mathematics and no more than pencil and paper. John von Neumann discovered cellular automata and self-reproducing systems, again with only pencil and paper. Aleksandr Lyapunov and Jules Henri Poincare worked on the foundations of chaos theory without any computer at all.

At the same time Howard T. Odum, the radiation ecologist recognised that the study of general systems required a language that could depict energetics and kinetics at any system scale. Odum developed a general systems, or Universal language, based on the circuit language of electronics to fulfil this role, known as the Energy Systems Language. Between 1929-1951, Robert Maynard Hutchins at the University of Chicago had undertaken efforts to encourage innovation and interdisciplinary research in the social sciences, aided by the Ford Foundation with the interdisciplinary Division of the Social Sciences established in 1931 (Hammond 2003). Numerous scholars had been actively engaged in ideas before, but in 1937 Bertalanffy presented the general theory of systems for a conference at the University of Chicago.

The systems view was based on several fundamental ideas. First, all phenomena can be viewed as a web of relationships among elements, or a system. Second, all systems, whether electrical, biological, or social, have common patterns, behaviours, and properties that can be understood and used to develop greater insight into the behaviour of complex phenomena and to move closer towards a unity of science. System philosophy, methodology and application are complementary to this science (Laszlo 1974).

By 1956, the Society for General Systems Research was established, renamed the International Society for Systems Science in 1988. The Cold War had its affects upon the research project for systems theory in ways that sorely disappointed many of the seminal theorists. Some began to recognise theories defined in association with systems theory had deviated from the initial General Systems Theory (GST) view (Hull 1970). The economist Kenneth Boulding, an early researcher in systems theory, had concerns over the manipulation of systems concepts. Boulding concluded from the effects of the Cold War that abuses of power always prove consequential and that systems theory might address such issues (Hammond 2003). Since the end of the Cold War, there has been a renewed interest in systems theory with efforts to strengthen an ethical view.

Overview

With the modern foundations for a general theory of systems following the World Wars, Ervin Laszlo, in the preface for Bertalanffy's book *Perspectives on General System Theory*, maintains that the translation of "general system theory" from German into English has "wroth a certain amount of Havoc" (Laszlo 1974). The preface explains that the original concept of a general system theory was "Allgemeine Systemtheorie (or Lehre)," pointing out the fact that "Theorie" (or Lehre) just as "Wissenschaft" (translated Scholarship), "has a much broader meaning in German than the closest English words 'theory' and 'science'" (Laszlo 1974).

With these ideas referring to an organised body of science and "any systematically presented set of concepts, whether they are empirical, axiomatic, or philosophical," "Lehre" is associated with theory and science in the etymology of general systems, but also

does not translate from the German very well; "teaching" is the closest equivalent (Laszlo 1974). While many of the root meanings for the idea of a "general systems theory" might have been lost in the translation and many were led to believe that the theorists had nothing but articulated a pseudoscience, systems theory became a nomenclature that early investigators used to describe the interdependence of relationships in organisation by defining a new way of thinking about science and scientific paradigms.

A system from this frame of reference is composed of regularly interacting or interrelating groups of activities. For example, in noting the influence in organisational psychology as the field evolved from "an individually oriented industrial psychology to a systems and developmentally oriented organisational psychology," it was recognised that organisations are complex social systems; reducing the parts from the whole reduces the overall effectiveness of organisations (Schein 1980). This is at difference to conventional models that centre on individuals, structures, departments and units separate in part from the whole instead of recognising the interdependence between groups of individuals, structures and processes that enable an organisation to function.

Laszlo (1972) explains that the new systems view of organised complexity went "one step beyond the Newtonian view of organised simplicity" in reducing the parts from the whole, or in understanding the whole without relation to the parts. The relationship between organisations and their environments became recognised as the foremost source of complexity and interdependence. In most cases the whole has properties that cannot be known from analysis of the constituent elements in isolation. Bela H. Banathy, who argued - along with the founders of the systems society—that "the benefit of humankind" is the purpose of science, has made significant and far-reaching contributions to the area of systems theory. For the Primer Group at ISSS, Banathy defines a perspective that iterates this view:

> The systems view is a world-view that is based on the discipline of *System Enquiry*, Central to systems enquiry is the concept of SYSTEM. In the most general sense, system means a configuration of parts connected and joined together by a web of relationships. The Primer

> group defines system as a family of relationships among the members acting as a whole. Bertalanffy defined system as "elements in standing relationship."

Similar ideas are found in learning theories that developed from the same fundamental concepts, emphasising that understanding results from knowing concepts both in part and as a whole. In fact, Bertalanffy's organismic psychology paralleled the learning theory of Jean Piaget (Bertalanffy 1968). Interdisciplinary perspectives are critical in breaking away from industrial age models and thinking where history is history and math is math segregated from the arts and music separate from the sciences and never the twain shall meet.

The influential contemporary work of Peter Senge (2000) provides detailed discussion of the commonplace critique of educational systems grounded in conventional assumptions about learning, including the problems with fragmented knowledge and lack of holistic learning from the "machine-age thinking" that became a "model of school separated from daily life." It is in this way that systems theorists attempted to provide alternatives and an evolved ideation from orthodox theories with individuals such as Max Weber, Emile Durkheim in sociology and Frederick Winslow Taylor in scientific management, which were grounded in classical assumptions. The theorists sought holistic methods by developing systems concepts that could be integrated with different areas.

The contradiction of reductionism in conventional theory (which has as its subject a single part) is simply an example of changing assumptions. The emphasis with systems theory shifts from parts to the organisation of parts, recognising interactions of the parts are not "static" and constant but "dynamic" processes.

Conventional closed systems were questioned with the development of open systems perspectives. The shift was from absolute and universal authoritative principles and knowledge to relative and general conceptual and perceptual knowledge (Bailey 1994), still in the tradition of theorists that sought to provide means in organising human life. Meaning, the history of ideas that preceded where rethought not lost. Mechanistic thinking was

particularly critiqued, especially the industrial-age mechanistic metaphor of the mind from interpretations of Newtonian mechanics by Enlightenment philosophers and later psychologists that laid the foundations of modern organisational theory and management by the late 19th century (Bailey 1994; Flood 1997; Checkland 1999; Laszlo 1972). Classical science had not been overthrown, but questions arose over core assumptions that historically influenced organised systems, within both social and technical sciences.

Types of Systems

Evidently, there are many types of systems that can be analysed both quantitatively and qualitatively. For example, with an analysis of urban systems dynamics, Steiss (1967) defines five intersecting systems, including the physical subsystem and behavioural system. For sociological models influenced by systems theory, where Bailey (1994) defines systems in terms of conceptual, concrete and abstract systems (either isolated, closed, or open), Buckley (1967) defines social systems in sociology in terms of mechanical, organic, and process models. Banathy (1997) cautions that with any enquiry into a system that understanding the type of system is crucial and defines Natural and Designed systems. In offering these more global definitions, the author maintains that it is important not to confuse one for the other.

The theorist explains that natural systems include subatomic systems, living systems, the solar system, the galactic system and the Universe. Designed systems are our creations, our physical structures, hybrid systems which include natural and designed systems, and our conceptual knowledge. The human element of organisation and activities are emphasised with their relevant abstract systems and representations. A key consideration in making distinctions among various types of systems is to determine how much freedom the system has to select purpose, goals, methods, tools, etc. and how widely is the freedom to select distributed (or concentrated) in the system.

Klin (1969: 69-72) maintains that no "classification is complete and perfect for all purposes," and defines systems in terms of abstract, real, and conceptual physical systems, bounded and

unbounded systems, discrete to continuous, pulse to hybrid systems, et cetera. The interaction between systems and their environments are categorised in terms of absolutely closed systems, relatively closed, and open systems. The case of an absolutely closed system is a rare, special case.

Important distinctions have also been made between hard and soft systems (Checkland 1999; Flood 1997). Hard systems are associated with areas such as systems engineering, operations research and quantitative systems analysis. Soft systems are commonly associated with concepts developed by Checkland through Soft Systems Methodology (SSM) involving methods such as action research and emphasising participatory designs. Where hard systems might be identified as more "scientific," the distinction between them is actually often hard to define.

General Systems Theory, Software and Computing

In the 1960s, systems theory was adopted by the post-John Von Neumann computing and information technology field, and, in fact formed the basis of structured analysis and structured design. It was also the basis for early software engineering and computer-aided software engineering principles.

By the 1970s, General Systems Theory (GST) was the fundamental underpinning of most commercial software design techniques, and by the 1980, W. Vaughn Frick and Albert F. Case, Jr. had used GST to design the "missing link" transformation from system analysis (defining what's needed in a system) to system design (what's actually implemented) using the Yourdon/Demarco notation. These principles were incorporated into computer-aided software engineering tools delivered by Nastec Corporation, Transform Logic, Inc., KnowledgeWare, Texas Instruments, Arthur Andersen and ultimately IBM Corporation.

System Theory Applications in Biology

Ideas from System Theory lie behind many of the new powerful methods being developed in the burgeoning field of Systems Biology. In this paper, we show two examples of this: one in the area of stochastic chemical kinetics, and the other in biological model invalidation. Stochastic chemical kinetics has gained a lot

of attention in the last few years. In order to capture certain important dynamics in the subcellular environment, it is necessary to model molecular interactions at the gene level as discrete stochastic events. The dynamics of such processes is typically described by probability distributions, which evolve according to the set of linear ordinary differential equations known as the chemical master equation (CME). Until recently, it has been believed that the CME could not be solved analytically except in the most trivial of problems, and the CME has been analysed almost exclusively with Monte Carlo (MC) algorithms.

However, concepts from linear systems theory have enabled the Finite State Projection (FSP) approach and have significantly enhanced our ability to solve the CME without resorting to MC simulations. In this paper we review the FSP approach as well as a variety of systems theory based modifications to the FSP algorithm that dramatically improve the computational efficiency of the algorithm and expand the class of solvable problems. Notions such as observability, controllability and minimal realisations enable large reductions in the order of models and increase efficiency with little to no loss in accuracy.

Model reduction techniques based upon linear perturbation theory allow for the systematic projection of multiple time scale dynamics onto a slowly varying manifold. Our second example shows the application of systems ideas in the area of biological model invalidation. As a specific case study, we use a dynamic model of the bacterial heat-shock response to demonstrate the approach. Using recent sum-of-squares techniques we show that the heat-shock model, when stripped from a certain protein-protein interaction that implements a certain feedback loop, cannot account for the input-output data regardless of the parameter choice for the model. In essence, such a deficient model is invalidated. Such conclusions are essential for pointing out the likelihood of missing components or interactions, thereby guiding new biological experiments.

With new experimental tools, contemporary molecular biologists are discovering complex gene regulatory networks, which control the expression of diverse biological traits. As more data becomes available, these networks become so complicated

that it is becoming increasingly difficult to understand their function without the aid of clear quantitative models. With such models at hand, it becomes possible to capture the various interactions among the known players and to provide a holistic system level understanding of the underlying biology.

The new understanding that emerges also provides insight into how one may alter these networks and affect their function, making it possible to design and construct new biological parts and systems, or to redesign existing, natural biological systems for useful purposes. The latter is the main aim of the new field of synthetic biology.

In modelling biological phenomena at the cellular level, one might be initially tempted to propose deterministic models to describe the concentrations of key proteins and other molecules within the cell as a function of time. This is feasible so long as there are huge numbers of each reactant species, such that fluctuations in these concentrations are inconsequential. At the sub cellular level, however, processes can, and often do, depend upon individual molecular interactions.

A single transcription factor binding to a single gene may initiate production of a key protein and eventually result in a cascade of events that affects the whole cell.

Here random fluctuations have great impact on the cell, one that cannot be captured with a deterministic model. Instead, these processes must be modelled on the mesoscopic scale using discrete and stochastic models. This is the subject of stochastic chemical kinetics, an area that has received a lot of attention recently and one where notions from system theory are beginning to have a big impact. In the first part of this paper, we describe a new and very promising direction in the field of stochastic chemical kinetics that relies on finite state projections and ideas from system theory to provide probability densities of important biological states within pre-specified errors.

In the second part of this paper, we show the application of systems ideas in the area of biological model invalidation. As a specific example, we apply recent sum-of-squares techniques to show that a dynamic model of the bacterial heat-shock response

must include a certain protein-protein interaction. With out the feedback loop enacted by that reaction, there exists no set of model parameters for which the system will exhibit the observed input-output data. This type of model invalidation is essential for the discovery of missing components or interactions and can guide new biological experiments.

Categorisation

Genes encoding for the CYP enzymes, and the enzymes themselves, are designated with the abbreviation "CYP", followed by an Arabic numeral indicating the gene family, a capital letter indicating the subfamily, and another numerals for the individual gene. The convention is to italicise when referring to the gene. For example, *CYP2E1* is the gene that encodes for the enzyme CYP2E1– one of the enzymes involved in paracetamol (acetaminophen) metabolism.

The current nomenclature guidelines suggest that members of new CYP families share >40 per cent amino acid identity, while members of subfamiles must share >55 per cent amino acid identity. There is a Nomenclature Committee that keeps track of and assigns new names.

Various Methods

The active site of cytochromes P450 contain a haem iron centre. The iron is tethered to the protein via a thiolate ligand derived from a cysteine residue. This cysteine and several flanking residues (RXCXG) are absolutely conserved over all known CYPs. Because of the vast variety of reactions catalysed by CYPs, a generalised description of the enzyme mechanism by necessity will not detail many of the known aspects of different CYPs. However, in general:

1. The resting state of the protein is as oxidised Fe^{3+}.
2. Binding of a substrate initiates electron transport and oxygen binding.
3. Electrons are supplied to the CYP by another protein, either cytochrome P450 reductase, ferredoxins, or cytochrome b5 to reduce the haem iron.

4. Molecular oxygen is bound and split by the now reduced iron.
5. An iron-bound oxidant, in some cases an iron(IV) oxo, oxidises the substrate to an alcohol or an epoxide, regenerating the resting state of the CYP.

Because most CYPs require a protein partner to deliver one or more electrons to reduce the iron (and eventually molecular oxygen), CYPs are properly speaking part of P450-containing systems of proteins. Five general schemes are known:

- CPR/cyb5/P450 systems employed by most eukaryotic microsomal (i.e. not mitochondrial) CYPs involve the reduction of cytochrome P450 reductase (variously CPR,POR, or CYPOR) by NADPH, and the transfer of reducing power to the CYP. Cytochrome b5 (cyb5) can also contribute reducing power to this system after being reduced by cytochrome b5 reductase (CYB5R).
- FR/Fd/P450 systems which are employed by mitochondrial and some bacterial CYPs.
- CYB5R/cyb5/P450 systems in which both electrons required by the CYP come from cytochrome b5.
- FMN/Fd/P450 systems originally found in *Rhodococcus* sp. in which a FMN domain containing reductase is fused to the CYP.
- P450 only systems, which do not require external reducing power. Notably these include CYP5 (thromboxane synthase), CYP8, prostacyclin synthase, and CYP74A (allene oxide synthase).

P450s in Bacteria

Bacterial cytochromes P450 are often soluble enzymes and are involved in critical metabolic processes of bacteria.

Three examples that have contributed significantly to structural and mechanistic studies are listed here, but many different families exist.

- Cytochrome P450cam (CYP101) originally from Pseudomonas putida has been used as a model for many cytochrome P450s and was the first cytochrome P450 three

dimensional protein structure solved by X-ray crystallography. This enzyme is part of a camphor-hydroxylating catalytic cycle comprised of two electron transfer steps from putidaredoxin, a 2Fe-2S cluster-containing protein cofactor.

- Cytochrome P450 eryF (CYP107A1) originally from the actinomycete bacterium Saccharopolyspora erythraea is responsible for the biosynthesis of the antibiotic erythromycin by C6-hydroxylation of the macrolide 6-deoxyerythronolide B.
- Cytochrome P450 BM3 (CYP102A1) from the soil bacterium Bacillus megaterium catalyses the NADPH-dependent hydroxylation of several long-chain fatty acids at the $\omega-1$ through $\omega-3$ positions. Unlike almost every other known CYP (except CYP505A1, cytochrome P450 foxy), it constitutes a natural fusion protein between the CYP domain and an electron donating cofactor. Thus, BM3 is potentially very useful in biotechnological applications.

P450s in Plants

Plant cytochrome P450s are involved in a wide range of biosynthetic reactions, leading to various fatty acid conjugates, plant hormones, defensive compounds, or medically important drugs. Terpenoids, which represent the largest class of characterised natural plant compounds, are often metabolic substrates for plant CYPs.

P450s in Animals

Animal CYPs are primarily membrane-associated proteins, located either in the inner membrane of mitochondria or in the endoplasmic reticulum of cells. CYPs metabolise thousands of endogenous and exogenous compounds. Most CYPs can metabolise multiple substrates, and many can catalyse multiple reactions, which accounts for their central importance in metabolising the potentially endless variety of endogenous and exogenous molecules. In the liver, these substrates include drugs and toxic compounds as well as metabolic products such as bilirubin (a

breakdown product of haemoglobin). Cytochromes P450 are present in many other tissues of the body including the mucosa of the gastrointestinal tract, and play important roles in hormone synthesis and breakdown (including estrogen and testosterone synthesis and metabolism), cholesterol synthesis, and vitamin D metabolism. In most animals, including humans, hepatic cytochromes P450 are the most widely studied of the P450 enzymes.

The Human Genome Project has identified more than 63 human genes (57 full genes and 5 pseudogenes) coding for the various cytochrome P450 enzymes.

Existence of Drugs

In drug metabolism, cytochrome P450 is probably the most important element of oxidative metabolism (a part of Phase I metabolism) in animals (metabolism in this context being the chemical modification or degradation of chemicals including drugs and endogenous compounds). Many drugs may increase or decrease the activity of various CYP isozymes in a phenomenon known as enzyme induction and inhibition.

This is a major source of adverse drug interactions, since changes in CYP enzyme activity may affect the metabolism and clearance of various drugs. For example, if one drug inhibits the CYP-mediated metabolism of another drug, the second drug may accumulate within the body to toxic levels, possibly causing an overdose.

Hence, these drug interactions may necessitate dosage adjustments or choosing drugs which do not interact with the CYP system. In addition, naturally occurring compounds may also cause a similar effect. For example, bioactive compounds found in grapefruit juice and some other fruit juices, including bergamottin, dihydroxybergamottin, and paradisin-A, have been found to inhibit CYP3A4-mediated metabolism of certain medications, leading to increased bioavailability and thus the strong possibility of overdosing. Because of this risk, avoiding grapefruit juice and fresh grapefruits entirely while on drugs is usually advised.

Other Specific CYP Functions in Animals

A subset of cytochrome P450 enzymes play important roles in the synthesis of steroid hormones by the adrenals, gonads, and peripheral tissue:

- CYP11A1 (also known as P450scc or P450c11a1) in adrenal mitochondria effects "the activity formerly known as 20,22-desmolase" (steroid 20 α -hydroxylase, steroid 22-hydroxylase, cholesterol side chain scission).
- CYP11B1 (encoding the protein P450c11 β) found in the inner mitochondrial membrane of adrenal cortex has steroid 11 β -hydroxylase, steroid 18-hydroxylase, and steroid 18-methyloxidase activities.
- CYP11B2 (encoding the protein P450c11AS), found only in the mitochondria of the adrenal zona glomerulosa, has steroid 11 β -hydroxylase, steroid 18-hydroxylase, and steroid 18-methyloxidase activities.
- CYP17A1, in endoplasmic reticulum of adrenal cortex has steroid 17 α -hydroxylase and 17,20-lyase activities.
- CYP21A1 (P450c21) in adrenal cortex conducts 21-hydroxylase activity.
- CYP19A (P450arom, aromatase) in endoplasmic reticulum of gonads, brain, adipose tissue, and elsewhere catalyses aromatisation of androgens to estrogens.

CYP Families in Humans

Humans have 57 genes and more than 59 pseudogenes divided among 18 families of cytochrome P450 genes and 43 subfamilies. This is a summary of the genes.

Family	*Function*	*Members*	*Names*
CYP1	drug and steroid (especially estrogen) metabolism	3 subfamilies, 3 genes, 1 pseudogene	CYP1A1, CYP1A2, CYP1B1
CYP2	drug and steroid metabolism	13 subfamilies, 16 genes,	CYP2A6, CYP2A7, CYP2A13,
		16 pseudogenes	CYP2B6, CYP2C8, CYP2C9, CYP2C18, CYP2C19, CYP2D6, CYP2E1, CYP2F1, CYP2J2, CYP2R1, CYP2S1, CYP2U1, CYP2W1
CYP3	drug and steroid (including testosterone) metabolism	1 subfamily, 4 genes, 2 pseudogenes	CYP3A4, CYP3A5, CYP3A7, CYP3A43

Contd..

Family	*Function*	*Members*	*Names*
CYP4	arachidonic acid or fatty acid metabolism	6 subfamilies, 11 genes, 10 pseudogenes	CYP4A11, CYP4A22, CYP4B1, CYP4F2, CYP4F3, CYP4F8, CYP4F11, CYP4F12, CYP4F22, CYP4V2, CYP4X1, CYP4Z1
CYP5	thromboxane A_2 synthase	1 subfamily, 1 gene	CYP5A1
CYP7	bile acid biosynthesis 7-alpha hydroxylase of steroid nucleus	2 subfamilies, 2 genes	CYP7A1, CYP7B1
CYP8	*varied*	2 subfamilies, 2 genes	CYP8A1(prostacyclin synthase), CYP8B1 (bile acid biosynthesis)
CYP11	steroid biosynthesis	2 subfamilies, 3 genes	CYP11A1, CYP11B1, CYP11B2
CYP17	steroid biosynthesis, 17-alpha hydroxylase	1 subfamily, 1 gene	CYP17A1
CYP19	steroid biosynthesis: aromatase synthesises estrogen	1 subfamily, 1 gene	CYP19A1
CYP20	unknown function	1 subfamily, 1 gene	CYP20A1
CYP21	steroid biosynthesis	2 subfamilies, 2 genes, 1 pseudogene	CYP21A2
CYP24	vitamin D degradation	1 subfamily, 1 gene	CYP24A1
CYP26	retinoic acid hydroxylase	3 subfamilies, 3 genes	CYP26A1, CYP26B1, CYP26C1
CYP27	*varied*	3 subfamilies, 3 genes	CYP27A1 (bile acid biosynthesis), CYP27B1 (vitamin D3 1-alpha hydroxylase, activates vitamin D3), CYP27C1 (unknown function)
CYP39	7-alpha hydroxylation of 24-hydroxycholesterol	1 subfamily, 1 gene	CYP39A1
CYP46	cholesterol 24-hydroxylase	1 subfamily, 1 gene	CYP46A1
CYP51	cholesterol biosynthesis	1 subfamily, 1 gene, 3 pseudogenes	CYP51A1 (lanosterol 14-alpha demethylase)

Area under the Curve: In the field of pharmacokinetics, the area under the curve (AUC) is the area under the curve in a plot of concentration of drug in plasma against time.

Interpretation and Usefulness of AUC Values

In real-world terms the AUC (from zero to infinity) represents the total amount of drug absorbed by the body, irrespective of the rate of absorption. This is useful when trying to determine whether two formulations of the same dose release the same dose of drug to the body. Another use is in the therapeutic monitoring of toxic drugs. For example, gentamicin is an antibiotic which displays nephro and ototoxicities; measurement of gentamicin concentrations in a patient's plasma and calculation of the AUC

is used to guide the dosage of this drug. AUC becomes useful for knowing the average concentration over a time interval, AUC/t. Also, AUC is referenced when talking about elimination. The amount eliminated by the body = clearance (volume/time)* AUC (mass*time/volume).

Capacity of Binding

In pharmacokinetics and receptor-ligand kinetics the binding potential (BP) is a combined measure of the density of "available" neuroreceptors and the affinity of a drug to that neuroreceptor.

Description

Consider a ligand receptor binding system. Ligand with a concentration L associates with a receptor of concentration or availability R to form a ligand-receptor complex with concentration RL. The binding potential is then the ratio ligand-receptor complex to free ligand at equilibrium and in the limit of L tending to 0, and is given symbol BP:

$$BP = \left.\frac{RL}{L}\right|_{L \approx 0}$$

This quantity, originally defined by Mintun, describes the capacity of a receptor to bind ligand. It is a limit (L << Ki) of the general receptor association equation:

$$RL = \frac{R.L}{L+Ki}$$

and is thus also equivalent to:

$$BP = \frac{R}{K_i}$$

These equations apply equally when measuring the total receptor density or the residual receptor density available after binding to second ligand—availability.

BP in Positron Emission Tomography

BP is a pivotal measure in the use of positron emission tomography (PET) to measure the density of "available" receptors, e.g. to assess the occupancy by drugs or to characterise

neuropsychiatric diseases (yet, one should keep in mind that binding potential is a combined measure that depends on receptor density as well as on affinity). An overview of the related methodology is, e.g. given in Laruelle *et al.* (2002). Estimating BP with PET usually requires that a reference tissue is available.

A reference tissue has negligible receptor density and its distribution volume should be the same as the distribution volume in the target region if all receptors were blocked. Although the BP can be measured in a relatively unbiased way by measuring the whole time course of labelled ligand association and blood radioactivity, this is practically not always necessary. Two other common measures have been derived, which involve assumptions, but result in measures that should correlate with BP: BP_1 and BP_2.

- BP_2: The "specific to non-specific equilibrium partition coefficient, in the literature also denoted as V_3". This is the ratio of specifically bound to non-displaceable tracer in brain tissue at true equilibrium. It can be calculated without arterial blood sampling. In the two-tissue compartment model: $BP_2 = k_3 / k_4$ and $BP_2 = f_2 BP$ where f_2 is the free fraction of the tracer in the first tissue compartment, i.e. a measure that depends on the non-specific binding of the ligand in brain tissue
- BP_1: The ratio of specifically bound tracer to tracer in plasma at true equilibrium, in the literature also denoted BP'. Measuring BP_1 includes measurements of radioactivity in plasma, including metabolite correction. From the two-tissue compartment model and by assuming there is only passive diffusion across the blood brain barrier, one obtains: $BP_1 = f_1 BP$ where f_1 is the free fraction of the tracer in arterial plasma, i.e. a measure that depends on plasma binding. Measuring and dividing by f_1 finally allows to obtain BP.

Definitions and Symbols

While BP_1 and BP_2 are non-ambiguous symbols, BP is not. There are many publications in which BP denotes BP_2. Generally, if there were no arterial samples (non-invasive imaging), BP denotes BP_2.

B_{max}**:** Total density of receptors = $R + RL$. In PET imaging, the amount of radioligand is usually very small, thus $B_{max} \approx R$

k_3 **and** k_4**:** Transfer rate constants from the two tissue compartment model.

Bioavailability

In pharmacology, bioavailability is used to describe the fraction of an administered dose of unchanged drug that reaches the systemic circulation, one of the principal pharmacokinetic properties of drugs. By definition, when a medication is administered intravenously, its bioavailability is 100 per cent.

However, when a medication is administered via other routes (such as orally), its bioavailability decreases (due to incomplete absorption and first-pass metabolism). Bioavailability is one of the essential tools in pharmacokinetics, as bioavailability must be considered when calculating dosages for non-intravenous routes of administration.

Definition

Bioavailability is a measurement of the extent of a therapeutically active drug that reaches the systemic circulation and is available at the site of action.

It is expressed as the letter *F*.

Absolute Bioavailability

Absolute bioavailability measures the availability of the active drug in systemic circulation after non-intravenous administration (i.e., after oral, rectal, transdermal, subcutaneous administration).

In order to determine absolute bioavailability of a drug, a pharmacokinetic study must be done to obtain a *plasma drug concentration vs time* plot for the drug after both intravenous (IV) and non-intravenous administration. The absolute bioavailability is the dose-corrected area under curve (AUC) non-intravenous divided by AUC intravenous.

For example, the formula for calculating *F* for a drug administered by the oral route (po) is given below.

$$F = \frac{[\mathrm{AUC}]_{po} * \mathrm{dose}_{IV}}{[\mathrm{AUC}]_{IV} * \mathrm{dose}_{po}}$$

Therefore, a drug given by the intravenous route will have an absolute bioavailability of 1 (F = 1) while drugs given by other routes usually have an absolute bioavailability of less than one.

Relative Bioavailability

This measures the bioavailability of a certain drug when compared with another formulation of the same drug, usually an established standard, or through administration via a different route.

When the standard consists of intravenously administered drug, this is known as absolute bioavailability.

$$relative\ bioavailability = \frac{[AUC]_A * dose_B}{[AUC]_B * dose_A}$$

Factors Influencing Bioavailability

The absolute bioavailability of a drug, when administered by an extravascular route, is usually less than one (i.e. $F<1$). Various physiological factors reduce the availability of drugs prior to their entry into the systemic circulation.

Such factors may include, but are not limited to:

- poor absorption from the gastrointestinal tract.
- degradation or metabolism of the drug prior to absorption.
- hepatic first pass effect.

Each of these factors may vary from patient to patient, and indeed in the same patient over time.

Whether a drug is taken with or without food will affect absorption, other drugs taken concurrently may alter absorption and first-pass metabolism, intestinal motility alters the dissolution of the drug and may affect the degree of chemical degradation of the drug by intestinal microflora. Disease states affecting liver metabolism or gastrointestinal function will also have an effect.

Bioequivalence

Bioequivalence is a term in pharmacokinetics used to assess the expected *in vivo* biological equivalence of two proprietary preparations of a drug. If two products are said to be bioequivalent it means that they would be expected to be, for all intents and purposes, the same.

Birkett (2003) defined bioequivalence by stating that, "two pharmaceutical products are bioequivalent if they are pharmaceutically equivalent and their bioavailabilities (rate and extent of availability) after administration in the same molar dose are similar to such a degree that their effects, with respect to both efficacy and safety, can be expected to be essentially the same. Pharmaceutical equivalence implies the same amount of the same active substance(s), in the same dosage form, for the same route of administration and meeting the same or comparable standards".

The United States Food and Drug Administration (FDA) has defined bioequivalence as, "the absence of a significant difference in the rate and extent to which the active ingredient or active moiety in pharmaceutical equivalents or pharmaceutical alternatives becomes available at the site of drug action when administered at the same molar dose under similar conditions in an appropriately designed study."

Bioequivalence Testing

In determining bioequivalence, for example, between two products such as a commercially-available Brand product and a potential to-be-marketed Generic product, pharmacokinetic studies are conducted whereby each of the preparations are administered in a cross-over study to volunteer subjects, generally healthy individuals but occasionally in patients. Serum plasma samples are obtained at regular intervals and assayed for parent drug (or occasionally metabolite) concentration. Occasionally, blood concentration levels are neither feasible or possible to compare the two products (e.g. inhaled corticosteroids), then pharmacodynamic endpoints rather than pharmacokinetic endpoints are used for comparison. For a pharmamacokinetic comparison, the plasma concentration data are used to assess key pharmacokinetic

parameters such as area under the curve (AUC), peak concentration (C_{max}), time to peak concentration (T_{max}), and absorption lag time (t_{lag}). Testing should be conducted at several different doses, especially when the drug displays non-linear pharmacokinetics.

In addition to data from bioequivalence studies, other data may need to be submitted to meet regulatory requirements for bioequivalence. Such evidence may include:

- Analytical method validation
- *In vitro-in vivo* correlation studies

Supervisory Definition

Australia

In Australia, the Therapeutics Goods Administration (TGA) considers preparations to be bioequivalent if the 90 per cent confidence intervals (90% CI) of the transformed natural log ratios, between the two preparations, of C_{max} and AUC lie in the range 0.80-1.25. T_{max} should also be similar between the products. There are tighter requirements for drugs with a narrow therapeutic index and/or saturable metabolism – thus no generic products exist on the Australian market for digoxin or phenytoin for instance.

Europe

According to European regulations EMEA-CPMP, Note for Guidance on the investigation of Bioavailability and Bioequivalence, London, July 2001 CPMP/EWP/QWP/1401/98 two medicinal products are bioequivalent if they are pharmaceutically equivalent or pharmaceutical alternatives and if their bioavailabilities after administration in the same molar dose are similar to such a degree that their effects, with respect to both efficacy and safety, will be essentially the same.

United States

The FDA considers two products bioequivalent if the 90 per cent CI of the relative mean C_{max}, $AUC_{(0-t)}$ and $AUC_{(0-\infty)}$ of the test (e.g. generic formulation) to reference (e.g. innovator brand formulation) should be within 80.00 per cent to 125.00 per cent

in the fasting state. Although there are a few exceptions, generally a bioequivalent comparison of Test to Reference formulations also requires administration after an appropriate meal at a specified time before taking the drug, a so-called "fed" or "food-effect" study. A food-effect study requires the same statistical evaluation as the fasting study.

Importance of Biological Half-life

The biological half-life of a substance is the time required for half of that substance to be removed from an organism by either a physical or a chemical process. Biological half-life is an important pharmacokinetic parameter and is usually denoted by the abbreviation $t_{1/2}$.

While a radioactive substance decays perfectly according to first order kinetics where the rate constant is fixed, the elimination of a substance from a living organism follows more complex kinetics.

Examples of Biological Half-lives

Water: The biological half-life of water in a human is about 7 to 10 days. It can be altered by behaviour. Drinking large amounts of beer will reduce the biological half-life of water in the body. This has been used to decontaminate humans who are internally contaminated with tritiated water (tritium). Drinking the same amount of water would have a similar effect, but many would find it difficult to drink a large volume of water. The basis of this decontamination method (used at Harwell) is to increase the rate at which the water in the body is replaced with new water.

Alcohol: The removal of ethanol (alcohol) through oxidation by alcohol dehydrogenase in the liver from the human body is limited. Hence the removal of a large concentration of alcohol from blood may follow zero-order kinetics.

Also the rate-limiting steps for one substance may be in common with other substances. For instance, the blood alcohol concentration can be used to modify the biochemistry of methanol and ethylene glycol. In this way the oxidation of methanol to the

toxic formaldehyde and formic acid in the human body can be prevented by giving an appropriate amount of ethanol to a person who has ingested methanol. Note that methanol is very toxic and causes blindness and death. A person who has ingested ethylene glycol can be treated in the same way.

Prescription for Cure

Fluoxetine: Some substances migrate slowly from the brain to the blood. The active metabolite of fluoxetine, a prescription anti-depressant, remains a long time in the brain because it is lipophilic. Biological half life of 4 to 16 days. The biological half life of the parent drug is 1 to 6 days.

Methadone: Methadone 15 to 60 hours, in rare cases up to 190 hours.

Salbutamol: 7 hours

Digoxin: Digoxin 24 to 36 hours

Amiodarone: Amiodarone 25 days

Cisplatin: Cisplatin 20 to 30 minutes

Chlorambucil: Chlorambucil 1.53 hours

Oxaliplatin: Oxaliplatin 14 minutes.

Metals: The biological half-life of caesium in humans is between one and four months. This can be shortened by feeding the person prussian blue. The prussian blue in the digestive system acts as a solid ion exchanger which absorbs the caesium while releasing potassium ions.

For some substances, it is important to think of the human or animal body as being made up of several parts, each with their own affinity for the substance, and each part with a different biological half-life.

Attempts to remove a substance from the whole organism may have the effect of increasing the burden present in one part of the organism.

For instance, if a person who is contaminated with lead is given EDTA in a chelation therapy, then while the rate at which

lead is lost from the body will be increased, the lead within the body tends to relocate into the brain where it can do the most harm.

- Polonium in the body has a biological half-life of about 30 to 50 days.
- Cesium in the body has a biological half-life of about one to four months.
- Lead in bone has a biological half-life of about ten years.
- Cadmium in bone has a biological half-life of about 30 years.
- Plutonium in bone has a biological half-life of about 100 years.
- Plutonium in the liver has a biological half-life of about 40 years.

A substance can have an effect on the health of a person long after the substance has left the body. For example, a car crash under the influence of chemicals may have consequences long into the future. A carcinogenic substance may cause cancer cells to appear, which may continue to multiply even after exposure to the carcinogen has stopped.

Maths

Zero-order Elimination: There are circumstances where the half-life varies with the concentration of the drug. For example, ethanol may be consumed in sufficient quantity to saturate the metabolic enzymes in the liver, and so is eliminated from the body at an approximately constant rate (zero-order elimination). Thus the half-life, under these circumstances, is proportional to the initial concentration of the drug A_0 and inversely proportional to the zero-order rate constant k_0 where:

$$t_{1/2} = \frac{0.5 A_0}{k_0}$$

First-order Elimination

This process is usually a first-order logarithmic process - that is, a constant proportion of the agent is eliminated per unit time (Birkett, 2002). Thus the fall in plasma concentration after the

administration of a single dose is described by the following equation:

$$C_t = C_0 e^{-kt}$$

- C_t is concentration after time t
- C_0 is the initial concentration (t=0)
- k is the elimination rate constant

The relationship between the elimination rate constant and half-life is given by the following equation:

$$k = \frac{\ln 2}{t_{1/2}}$$

Half-life is determined by clearance (CL) and volume of distribution (V_D) and the relationship is described by the following equation:

$$t_{1/2} = \frac{\ln 2 . V_D}{CL}$$

In clinical practice, this means that it takes just over 4.7 times the half-life for a drug's serum concentration to reach steady state after regular dosing is started, stopped, or the dose changed. So, for example, digoxin has a half-life (or $t_{½}$) of 24-36 hours; this means that a change in the dose will take the best part of a week to take full effect. For this reason, drugs with a long half-life (e.g. amiodarone, elimination $t_{½}$ of about 90 days) are usually started with a loading dose to achieve their desired clinical effect more quickly.

Definition of Blood-brain Barrier

The blood-brain barrier (abbreviated BBB) is a membranic structure that acts primarily to protect the brain from chemicals in the blood, while still allowing essential metabolic function. It is composed of endothelial cells, which are packed very tightly in brain capillaries. This higher density restricts passage of substances from the bloodstream much more than endothelial cells in capillaries elsewhere in the body. Processes (outgrowths) from astrocytes called astrocytic feet surround the endothelial

cells of the BBB, providing biochemical support to those cells. The BBB is distinct from the similar blood-cerebrospinal fluid barrier, a function of the choroid plexus.

History

The existence of such a barrier was first noticed in experiments by Paul Ehrlich in the late-19th century. Ehrlich was a bacteriologist who was studying staining, used for many studies to make fine structures visible. When injected, some of these dyes (notably the aniline dyes that were then popular) would stain all of the organs of an animal except the brain. At the time, Ehrlich attributed this to the brain simply not picking up as much of the dye.

However, in a later experiment in 1913, Edwin Goldmann (one of Ehrlich's students) injected the dye into the spinal fluid of the brain directly. He found that in this case the brain would become dyed, but the rest of the body would not. This clearly demonstrated the existence of some sort of barrier between the two. At the time, it was thought that the blood vessels themselves were responsible for the barrier, as no obvious membrane could be found. The concept of the blood-brain (then termed *hematoencephalic*) barrier was proposed by Lina Stern in 1921. It was not until the introduction of the scanning electron microscope to the medical research fields in the 1960s that the actual membrane could be demonstrated.

It was once believed that astrocytes rather than epithelial cells were the basis of the blood-brain barrier because of the densely packed astrocyte processes that surround the epithelial cells of the BBB.

Physiology

Throughout the body, the walls of the capillaries (the smallest of the blood vessels) are made up of fenestrated endothelial cells, separated by small gaps, or windows known as fenestrations. Soluble chemicals within the various tissues pass through these gaps into the blood stream, to be carried throughout the body and into different tissues.

In the brain, however, these endothelial cells are packed more tightly together, due to the existence of *zonulae occludentes* (tight

junctions) between them. This blocks the passage of most molecules. The blood-brain barrier blocks all molecules except those that cross cell membranes by means of lipid solubility (such as oxygen, carbon dioxide, ethanol, and steroid hormones) and those that are allowed in by specific transport systems (such as sugars and some amino acids). Substances with a molecular weight higher than 500 daltons (500 u) generally cannot cross the blood-brain barrier, while smaller molecules often can. In addition, the endothelial cells metabolise certain molecules to prevent their entry into the central nervous system. For example, L-DOPA, the precursor to dopamine, can cross the BBB, whereas dopamine itself cannot. (As a result, L-DOPA is administered for dopamine deficiencies (e.g., Parkinson's disease) rather than dopamine).

In addition to tight junctions acting to prevent transport in between epithelial cells, there are two mechanisms to prevent passive diffusion through the cell membranes. Glial cells surrounding capillaries in the brain pose a secondary hindrance to hydrophilic molecules, and the low concentration of interstitial proteins in the brain prevent access by hydrophilic molecules.

The blood-brain barrier protects the brain from the many chemicals flowing within the blood. However, many bodily functions are controlled by hormones in the blood, and while the secretion of many hormones is controlled by the brain, these hormones generally do not penetrate the brain from the blood. This would prevent the brain from directly monitoring hormone levels. In order to control the rate of hormone secretion effectively, there exist specialised sites where neurons can "sample" the composition of the circulating blood. At these sites, the blood-brain barrier is 'leaky'; these sites include three important 'circumventricular organs', the subfornical organ, the area postrema and the organum vasculosum of the lamina terminalis (OVLT).

The blood-brain barrier acts very effectively to protect the brain from many common infections. Thus, infections of the brain are very rare. However, since antibodies are too large to cross the blood-brain barrier, infections of the brain which do occur are often very serious and difficult to treat.

Targeting the Brain by Drugs

Overcoming the difficulty of delivering therapeutic agents to specific regions of the brain presents a major challenge to treatment of most brain disorders. In its neuroprotective role, the blood-brain barrier functions to hinder the delivery of many potentially important diagnostic and therapeutic agents to the brain. Therapeutic molecules and genes that might otherwise be effective in diagnosis and therapy do not cross the BBB in adequate amounts.

Mechanisms for drug targeting in the brain involve going either "through" or "behind" the BBB. Modalities for drug delivery through the BBB entail its disruption by osmotic means, biochemically by the use of vasoactive substances such as bradykinin, or even by localised exposure to high intensity focused ultrasound (HIFU). Other strategies to go through the BBB may entail the use of endogenous transport systems, including carrier-mediated transporters such as glucose and amino acid carriers; receptor-mediated transcytosis for insulin or transferrin; and blocking of active efflux transporters such as p-glycoprotein. Strategies for drug delivery behind the BBB include intracerebral implantation and convection-enhanced distribution.

Nanotechnology may also help in the transfer of drugs across the BBB. Recently, researchers have been trying to build nanoparticles loaded with liposomes to gain access through the BBB. More research is needed to determine which strategies will be most effective and how they can be improved for patients with brain tumours. The potential for using BBB opening to target specific agents to brain tumours has just begun to be explored.

Diseases

Meningitis: Meningitis is inflammation of the membranes which surround the brain and spinal cord (these membranes are also known as meninges). Meningitis is most commonly caused by infections with various pathogens. When the meninges are inflamed, the blood-brain barrier may be disrupted. This disruption may increase the penetration of various substances (including antibiotics) into the brain. Treatment with third generation or fourth generation cephalosporin is usually preferred.

Equation Derivation

Equation *1* is derived from a mass balance:

$$\Delta m_{body} = \left(-\dot{m}_{out} + \dot{m}_{in} + \dot{m}_{gen.} \right) \Delta t \tag{2}$$

where:

- Δt is a period of time
- Δm_{body} the change in mass of the toxin in the body during $"t$
- $\dot{m}_{in}$ is the toxin intake rate
- $\dot{m}_{out}$ is the toxin removal rate
- $\dot{m}_{gen.}$ is the toxin generation rate

In words, the above equation states:

The change in the mass of a toxin within the body ("m) during some time "t is equal to the toxin intake plus the toxin generation minus the toxin removal.

Since

$$m_{body} = C \cdot V \tag{3}$$

and

$$\dot{m}_{out} = K \cdot C \tag{4}$$

Equation A1 can be re-written as:

$$\Delta (C \cdot V) = \left(-K \cdot C + \dot{m}_{in} + \dot{m}_{gen.} \right) \Delta t \tag{5}$$

If one lumps the *in* and *gen.* terms together, i.e. $\dot{m} = \dot{m}_{in} + \dot{m}_{gen.}$ and divides by Δt the result is a difference equation:

$$\frac{\Delta (C \cdot V)}{\Delta t} = -K \cdot C + \dot{m} \tag{6}$$

If one applies the limit $\Delta t \to 0$ one obtains a differential equation:

$$\frac{d(C \cdot V)}{dt} = -K \cdot C + \dot{m} \tag{7}$$

Using the chain rule this can be rewritten as:

$$C\frac{dV}{dt} + V\frac{dC}{dt} = -K \cdot C + \dot{m} \tag{8}$$

If one assumes that the volume change is not significant, i.e. $C\frac{dV}{dt} = 0$, the result is Equation *1*:

$$V\frac{dC}{dt} = -K \cdot C + \dot{m} \tag{1}$$

Resolution of Differential Equation

The general solution of the above differential equation (*1*) is:

$$C = \frac{\dot{m}}{K} + \left(C_0 - \frac{\dot{m}}{K} \right) e^{-\frac{K \cdot t}{V}} \tag{9}$$

Where:

- C_o is the concentration at the beginning of dialysis *or* the initial concentration of the substance/drug (after it has distributed) [mmol/L] or [mol/m^3].
- e is the base of the natural logarithm.

Balanced Solution

The solution to the above differential equation (*9*) at time infinity (steady state) is:

$$C_{\infty} = \frac{\dot{m}}{K} \tag{10a}$$

The above equation (*10a*) can be rewritten as:

$$K = \frac{\dot{m}}{C_{\infty}} \tag{10b}$$

The above equation (*10b*) makes clear the relationship between mass removal and clearance. It states that (with a constant mass generation) the concentration and clearance vary inversely with one another.

If applied to creatinine (i.e. creatinine clearance), it follows from the equation that if the serum creatinine doubles the clearance halves and that if the serum creatinine quadruples the clearance is quartered.

Capacity of Renal Clearance

Renal clearance can be measured with a timed collection of urine and an analysis of its composition with the aid of the following equation (which follows directly from the derivation of (*10b*)):

$$K = \frac{C_U \cdot Q}{C_B} \tag{11}$$

Where:

- K is the clearance [mL/min]
- C_U is the urine concentration [mmol/L] (in the USA often [mg/mL])
- Q is the urine flow (volume/time) [mL/min] (often [mL/24 hours])
- C_B is the plasma concentration [mmol/L] (in the USA often [mg/mL])

Note: the above equation (*11*) is valid *only* for the steady-state condition. If the substance being cleared is *not* at a constant plasma concentration (i.e. *not* at steady-state) *K* must be obtained from the (full) solution of the differential equation (*9*).

Drug Endurance

Drug metabolism is the metabolism of drugs, their biochemical modification or degradation, usually through specialised enzymatic systems.

Drug metabolism often converts lipophilic chemical compounds into more readily excreted polar products. Its rate is

an important determinant of the duration and intensity of the pharmacological action of drugs.

Drug metabolism can result in toxication or detoxication- the activation or deactivation of the chemical. While both occur, the major metabolites of most drugs are detoxication products.

Drugs are almost all xenobiotics. Other commonly used organic chemicals are also drugs, and are metabolised by the same enzymes as drugs. This provides the opportunity for drug-drug and drug-chemical interactions or reactions.

Phase I vs. Phase II

Phase I and Phase II reactions are biotransformations of chemicals that occur during drug metabolism.

Phase I reactions—usually precedes Phase II, though not necessarily. During these reactions, polar bodies are either introduced or unmasked, which results in (more) polar metabolites of the original chemicals.

In the case of pharmaceutical drugs, Phase I reactions can lead either to activation or inactivation of the drug.

Phase I reactions (also termed Non-synthetic reactions) may occur by oxidation, reduction, hydrolysis, cyclisation, and decyclisation reactions.

Oxidation involves addition of oxygen (forming a negatively charged radical) or removal of hydrogen (forming a positively charged radical).

The process of oxidation takes place in the presence of mixed function oxidases and monooxygenases in the liver. These oxidative reactions typically involve a cytochrome P450 haemoprotein, NADPH and oxygen.

The classes of pharmaceutical drugs that utilise this method for their metabolism include phenothiazines, paracetamol, and steroids. If the metabolites of phase I reactions are sufficiently polar, they may be readily excreted at this point.

However, many phase I products are not eliminated rapidly and undergo a subsequent reaction in which an endogenous

substrate combines with the newly incorporated functional group to form a highly polar conjugate.

Phase II reactions — usually known as conjugation reactions (e.g., with glucuronic acid, sulphonates (commonly known as sulphation) , glutathione or amino acids) — are usually detoxication in nature, and involve the interactions of the polar functional groups of phase I metabolites.

Sites

Quantitatively, the smooth endoplasmic reticulum of the liver cell is the principal organ of drug metabolism, although every biological tissue has some ability to metabolise drugs. Factors responsible for the liver's contribution to drug metabolism include that it is a large organ, that it is the first organ perfused by chemicals absorbed in the gut, and that there are very high concentrations of most drug-metabolising enzyme systems relative to other organs.

If a drug is taken into the GI tract, where it enters hepatic circulation through the portal vein, it becomes well-metabolised and is said to show the first pass effect.

Other sites of drug metabolism include epithelial cells of the gastrointestinal tract, lungs, kidneys, and the skin. These sites are usually responsible for localised toxicity reactions.

Major Enzymes and Pathways

Several major enzymes and pathways are involved in drug metabolism, and can be divided into Phase I and Phase II reactions:

Phase I: Oxidation

- Cytochrome P450 monooxygenase system.
- Flavin-containing monooxygenase system.
- Alcohol dehydrogenase and aldehyde dehydrogenase.
- Monoamine oxidase.
- Co-oxidation by peroxidases.

Reduction

- NADPH-cytochrome P450 reductase.
- Reduced (ferrous) cytochrome P450.

It should be noted that during reduction reactions, a chemical can enter *futile cycling*, in which it gains a free-radical electron, then promptly loses it to oxygen (to form a superoxide anion).

Process of Hydrolysis

- Esterases and amidases
- Epoxide hydrolase

Phase II

- Glutathione S-transferases
  - — Mercapturic acid biosynthesis
- UDP-Glucuron(os)yltransferases
- N-Acetyltransferases
- Amino acid N-acyl transferases
- Sulphotransferases

Factors that Affect Drug Metabolism

The duration and intensity of pharmacological action of most lipophilic drugs are determined by the rate they are metabolised to inactive products. The Cytochrome P450 monooxygenase system is the most important pathway in this regard. In general, anything that increases the rate of metabolism (*e.g.*, enzyme induction) of a pharmacologically active metabolite will *decrease* the duration and intensity of the drug action. The opposite is also true (*e.g.*, enzyme inhibition).

Various physiological and pathological factors can also affect drug metabolism. Physiological factors that can influence drug metabolism include age, individual variation (e.g., pharmacogenetics), enterohepatic circulation, nutrition, intestinal flora, or sex differences.

In general, drugs are metabolised more slowly in fetal, neonatal and elderly humans and animals than in adults.

Genetic variation (polymorphism) accounts for some of the variability in the effect of drugs. With N-acetyltransferases (involved in Phase II reactions), individual variation creates a group of people who acetylate slowly (*slow acetylators*) and those

who acetylate quickly, split roughly 50:50 in the population of Canada. This variation may have dramatic consequences, as the slow acetylators are more prone to dose-dependent toxicity.

Cytochrome P450 monooxygenase system enzymes can also vary across individuals, with deficiencies occurring in 1 - 30 per cent of people, depending on their ethnic background.

Pathological factors can also influence drug metabolism, including liver, kidney, or heart diseases.

Enzyme Induction and Inhibition

Enzyme induction is a process in which a molecule (*e.g.* a drug) induces (*i.e.* initiates or enhances) the expression of an enzyme.

Enzyme inhibition can refer to:

- The inhibition of the expression of the enzyme by another molecule.
- Interference at the enzyme-level, basically with how the enzyme works. This can be competitive inhibition, uncompetitive inhibition, non-competitive inhibition or partially competitive inhibition.

If the molecule induces enzymes that are responsible for its own metabolism, this is called auto-induction (or auto-inhibition if there is inhibition). These processes are particular forms of gene expression regulation.

These terms are of particular interest to pharmacology, and more specifically to drug metabolism and drug interactions. They also apply to molecular biology.

In the late 1950s and early 1960s, the French molecular biologists Francois Jacob and Jacques Monod became the first to explain enzyme induction, in the context of the lac operon of Escherichia coli. In the absence of lactose, the constitutively expressed lac repressor protein binds to the operator region of the DNA and prevents the transcription of the operon genes. When present, lactose binds to the lac repressor, causing it to separate from the DNA and thereby enabling transcription to occur. Monod and Jacob generated this theory following 15 years of work by

them and others (including Joshua Lederberg), partially as an explanation for Monod's observation of diauxie. Previously, Monod had hypothesised that enzymes could physically adapt themselves to new substrates; a series of experiments by him, Jacob, and Arthur Pardee eventually demonstrated this to be incorrect and led them to the modern theory, for which he and Jacob shared the 1965 Nobel Prize in Physiology or Medicine (together with Andre Lwoff).

Cytochrome P450

One class of key enzymes for drug metabolism belong to the family of cytochrome P450 oxidases, like CYP3A4, CYP2D6, CYP1A2, *etc*. They reside in the endoplasmatic reticulum (ER), and prolonged usage of substances inducing enzymes here may cause proliferation of the ER. They are responsible for phase I reactions.

Enzyme induction and inhibition are important processes to take in account when using drugs of vital importance to the patient, drugs with important side effects and drugs with small therapeutic windows, but any drug may be subject to an altered plasma concentration due to altered drug metabolism.

A classical example includes anti-epileptic drugs. Phenytoin, for example, induces CYP1A2, CYP2C9, CYP2C19 and CYP3A4. Substrates for the latter may be drugs with critical dosage, like amiodarone or carbamazepine, whose blood plasma concentration may decrease because of enzyme induction.

Not only drugs may alter drug metabolism. Cigarette smoke induces CYP1A2 (example substrates are clozapine/olanzapine), Saint-John's wort (a common herbal remedy) induces CYP3A4, which is inhibited by grapefruit juice. There are known examples of situations where this may produce clinical effects.

First Pass Effect

The first-pass effect (or first-pass metabolism) is a phenomenon of drug metabolism.

After a drug is swallowed, it is absorbed by the digestive system and enters the hepatic portal system. The absorbed drug is carried through the portal vein into the liver.

The liver is responsible for metabolising many drugs. Some drugs are so extensively metabolised by the liver that only a small amount of unchanged drug may enter the systemic circulation, so the bioavailability of the drug is reduced.

Alternative routes of administration (e.g., intravenous, intramuscular, sublingual) avoid the first-pass effect.

Several Sclerosis

Multiple sclerosis (MS) is considered an autoimmune disorder in which the immune system attacks the myelin protecting the nerves in the central nervous system. Normally, a person's nervous system would be inaccessible for the white blood cells due to the blood-brain barrier. However, it has been shown using Magnetic Resonance Imaging that, when a person is undergoing an MS "attack," the blood-brain barrier has broken down in a section of his/her brain or spinal cord, allowing white blood cells called T lymphocytes to cross over and destroy the myelin. It has been suggested that, rather than being a disease of the immune system, MS is a disease of the blood-brain barrier. However, current scientific evidence is inconclusive.

There are currently active investigations into treatments for a compromised blood-brain barrier. It is believed that oxidative stress plays an important role into the breakdown of the barrier; anti-oxidants such as lipoic acid may be able to stabilise a weakening blood-brain barrier.

Neuromyelitis Optica

Neuromyelitis optica, also known as Devic's disease, is similar and often confused with multiple sclerosis. Among other differences from MS, the target of the autoimmune response has been identified. Patients with neuromyelitis optica have high levels of antibodies against a protein called aquaporin 4 (a component of the astrocytic foot processes in the blood-brain barrier).

Late-stage Neurological Trypanosomiasis (Sleeping Sickness)

Late-stage neurological trypanosomiasis, or sleeping sickness, is a condition in which trypanosoma protozoa are found in brain tissue. It is not yet known how the parasites infect the brain from

the blood, but it is suspected that they cross through the choroid plexus, a circumventricular organ.

Progressive Multifocal Leukoencephalopathy

Progressive multifocal leukoencephalopathy (PML) is a demyelinating disease of the central nervous system caused by reactivation of a latent papovavirus (the JC polyomavirus) infection, that can cross the BBB. It affects immune-compromised patients and is usually seen with patients having AIDS.

De Vivo Problem

De Vivo disease (also known as GLUT1 deficiency syndrome) is a rare condition caused by inadequate transport of glucose across the barrier, resulting in mental retardation and other neurological problems. Genetic defects in glucose transporter type 1 (GLUT1) appears to be the main cause of De Vivo disease.

Alzheimer's Problem

New evidence indicates that disruption of the blood brain barrier in AD patients allows beta amyloid containing blood plasma to enter the brain where the A beta adheres preferentially to the surface of astrocytes.

These findings have led to hypothesise that (1) breakdown of the blood-brain barrier allows access of neuron-binding auto-antibodies and soluble exogenous Aβ42 to brain neurons and (2) binding of these auto-antibodies to neurons triggers and/or facilitates the internalisation and accumulation of cell surface-bound Aβ42 in vulnerable neurons through their natural tendency to clear surface-bound auto-antibodies via endocytosis. Eventually the astrocyte is overwhelmed, dies, ruptures, and disintegrates, leaving behind the insoluble Aβ42 plaque. Thus, in some patients, Alzheimer's disease may be caused (or more likely, aggravated) by a breakdown in the blood brain barrier.

Clearance

In medicine, the clearance, also renal clearance or renal plasma clearance (when referring to the function of the kidney), of a

substance is the inverse of the time constant that describes its removal rate from the body divided by its volume of distribution (or total body water).

In steady-state, it is defined as the mass generation rate of a substance (which equals the mass removal rate) divided by its concentration in the blood.

It is considered to be the amount of liquid filtered out of the blood that gets processed by the kidneys or the amount of blood cleaned per time because it has the units of a volumetric flow rate [volume/time].

However, it does not refer to a real value; "the kidney does not completely remove a substance from the total renal plasma flow." From a mass transfer perspective and physiologically, volumetric blood flow (to the dialysis machine and/or kidney) is only one of several factors that determine blood concentration and removal of a substance from the body. Other factors include the mass transfer coefficient, dialysate flow and dialysate recirculation flow for haemodialysis, and the glomerular filtration rate and the tubular reabsorption rate, for the kidney. A physiologic interpretation of clearance (at steady-state) is that clearance is a ratio of the mass generation and blood (or plasma) concentration.

Its definition follows from the differential equation that describes exponential decay and is used to model kidney function and haemodialysis machine function:

$$V\frac{dC}{dt} = -K \cdot C + \dot{m} \quad (1)$$

Where:

- $\dot{m}$ is the mass generation rate of the substance - assumed to be a constant, i.e. not a function of time (equal to zero for foreign substances/drugs) [mmol/min] or [mol/s].
- t is dialysis time or time since injection of the substance/drug [min] or [s].
- V is the volume of distribution or total body water [L] or [m^3].
- K is the clearance [mL/min] or [m^3/s].

- C is the concentration [mmol/L] or [mol/m^3] (in the USA often [mg/mL]).

From the above definitions it follows that $\frac{dC}{dt}$ is the first derivative of concentration with respect to time, i.e. the change in concentration with time.

It is derived from a mass balance.

Different Models

Multi-compartment Model: A multi-compartment model is a type of mathematical model used to describe the way materials or energies are transmitted among the *compartments* of a system. Each compartment is assumed to be a homogenous entity within which the entities being modelled are equivalent. For instance, in a pharmacokinetic model, the compartments may represent different sections of a body within which the concentration of a drug is assumed to be constant. Hence a multi-compartment model is a lumped parameters model.

Multi-compartment models are used in many fields including pharmacokinetics, epidemiology, biomedicine, systems theory, complexity theory, engineering, physics, information science and social science. The circuits systems can be viewed as a multi-compartment model as well.

In systems theory, it involves the description of a network whose components are compartments that represent a population of elements that are equivalent with respect to the manner in which they process input signals to the compartment.

Assumptions

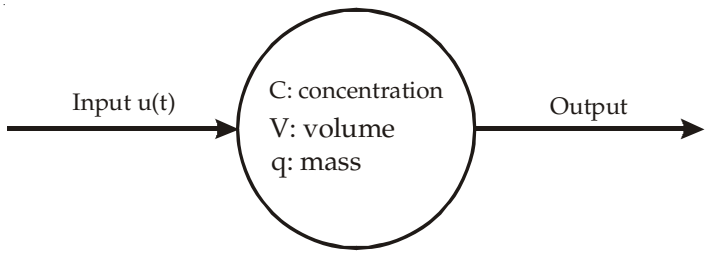
Multi-compartment modelling requires the adoption of several assumptions, such that systems in physical existence can be modelled mathematically:

- Instant homogeneous distribution of materials or energies within a "compartment."
- The exchange rate of materials or energies among the compartments is related to the densities of these compartments.

- Usually, it is desirable that the materials do not undergo chemical reactions while transmitting among the compartments.
- When concentration of the cell is of interest, typically the volume is assumed to be constant over time, though this may not be totally true in reality.

Most commonly, the mathematics of multi-compartment models is simplified to provide only a single parameter—such as concentration—within a compartment.

Single Compartment Model



Possibly the simplest application of multi-compartment model is in the single cell concentration monitoring.

If the volume of a cell is V, the mass of solute is q, the input is $u(t)$ and the secretion of the solution is proportional to the density of it within the cell, then the concentration of the solution C within the cell over time is given by:

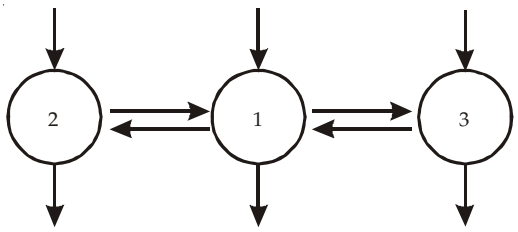
$$\frac{dq}{dt} = u(t) - kq$$

$$C = \frac{q}{V}$$

where k is the proportionality.

Multi-compartment Model

As the number of compartments increases, the model can be very complex and the solutions usually beyond ordinary calculation. Below shows a three-cell model with interlinks among each other.



The formula for n-cell multi-compartment models become:

$$\dot{q}_1 = q_1 k_{11} + q_2 k_{12} + ... + q_n k_{1n} + u_1(t)$$

$$\dot{q}_2 = q_1 k_{21} + q_2 k_{22} + ... + q_n k_{2n} + u_2(t)$$

$$\dot{q}_n = q_1 k_{n1} + q_2 k_{n2} + ... + q_n k_{nn} + u_n(t)$$

where

$$k_{ii} = \sum_{j=0;i\neq j}^{n} k_{ji} \text{ for } i=1, 2, ..., n$$

Or in matrix forms:

$$\dot{q} = Kq + u$$

where

$$K = \begin{bmatrix} k_{11} & k_{12} & L & k_{1n} \\ k_{21} & k_{22} & L & k_{2n} \\ M & M & O & M \\ k_{n1} & k_{n2} & ... & k_{nn} \end{bmatrix}_{and}$$

$$q = \begin{bmatrix} q_1 & q_2 & L & q_n \end{bmatrix}^T \text{ and } u = \begin{bmatrix} u_1(t) & u_2(t) & L & u_n(t) \end{bmatrix}^T$$

Topological Model

Generally speaking, as the number of compartment increases, it is challenging both to find the algebraic and numerical solutions of the model. However, there are special cases of models, which really exist in nature, when the topologies exist certain regularities that the solutions become easier to find. The model can be classified according to the interconnection of cells and input/output characteristics:

1. *Closed Model:* No sinks or source, lit. all $k_{oi} = 0$ and $u_i = 0$;
2. *Open Model:* There're sinks or/and sources among cells;
3. *Catenary Model:* All compartments arranges in a chain, with each pool connecting only to its neighbours. This model has 2 or more cells;
4. *Cyclic Model:* It's a special case of the catenary model, with 3 or more cells, in which the first and last cell are connected, i.e. $k_{1n} \neq 0$ or/and $k_{n1} \neq 0$.
5. *Mammillary Model:* Consists of a central compartment with peripheral compartments connecting to it. There're no interconnections among other compartments;
6. *Reducible Model:* It's a set of unconnected models. It bears great resemblance to the computer concept of forest as against trees.

Systems of Biochemicals

The Biochemical Systems Theory is a mathematical modelling framework for biochemical systems, based on ordinary differential equations (ODE), in which biochemical processes are represented using power-law expansions in the variables of the system. The dynamics of a specie is represented by a differential equation with the structure:

$$\frac{dX_i}{dt} = \sum_j \sigma_{ij} \cdot \gamma_j \prod_k X_k^{g_{jk}}$$

where X_i represents one of the n_d variables of the model (metabolite concentrations, protein concentrations or levels of gene expression). j represents the n_f biochemical processes affecting the dynamics of the specie. On the other hand, σ_{ij} (stoichiometric coefficient), γ_j (rate constants) and g_{ik} (kinetic orders) are two different kinds of parameters defining the dynamics of the system.

The principal difference of power-law models with respect to other ODE models used in biochemical systems is that the kinetic orders can be non-integer numbers. A kinetic order can have even negative value when inhibition is modelled. In this way, power-law models have a higher flexibility to reproduce the non-linearity of biochemical systems.

Models using power-law expansions have been used during the last 35 years to model and analyse several kinds of biochemical systems including metabolic networks, genetic networks and recently in cell signalling.

Definition of Pharmacokinetics

Pharmacokinetics (in Greek: "pharmacon" meaning drug, and "kinetikos" meaning putting in motion) is a branch of pharmacology dedicated to the determination of the fate of substances administered externally to a living organism. In practice, this discipline is applied mainly to drug substances, though in principle it concerns itself with all manner of compounds ingested or otherwise delivered externally to an organism, such as nutrients, metabolites, hormones, toxins, etc. Pharmacokinetics is often divided into several areas including, but not limited to, the extent and rate of Absorption, Distribution, Metabolism and Excretion. This sometimes is referred to as the ADME scheme.

Absorption is the process of a substance entering the body. Distribution is the dispersion or dissemination of substances throughout the fluids and tissues of the body. Metabolism is the transformation of the substances and its daughter metabolites. Excretion is the elimination of the substances from the body. In rare cases, some drugs irreversibly accumulate in a tissue in the body.

Pharmacokinetics is often studied in conjunction with pharmacodynamics. So while pharmacodynamics explores what a drug does to the body, pharmacokinetics explores what the body does to the drug.

Pharmacokinetics is sometimes abbreviated as "PK".

Role of Cytochrome

Cytochrome P450 (abbreviated CYP, P450, infrequently CYP450) is a diverse superfamily of haemoproteins found in bacteria, archaea and eukaryotes. Cytochromes P450 are involved in metabolism of a plethora of both exogenous and endogenous compounds.

Usually they form part of multicomponent electron transfer chains, called P450-containing systems. The most common reaction catalysed by cytochrome P450 is a monooxygenase reaction, i.e. insertion of one atom of oxygen into an organic substrate (RH) while the other oxygen atom is reduced to water:

$$RH + O_2 + 2H^+ + 2e^- \rightarrow ROH + H_2O$$

CYP homologs have been sequenced from all lineages of life, including mammals, birds, fish, insects, worms, sea squirts, sea urchins, plants, fungi, slime moulds, bacteria and archaea. More than 6400 distinct CYP sequences are known.

- The name P450 refers to the "pigment at 450 nm", so named for the characteristic Soret peak formed by absorbance of light at wavelengths near 450 nm when the haem iron is reduced (with sodium dithionite) and complexed to carbon monoxide.

Biophysics and Environment

In physics, the exchange interaction is a quantum mechanical effect which increases or decreases the energy of two or more electrons when their wave functions overlap. Arising from the Pauli exclusion principle, this energy change is the result of the identity of particles, exchange symmetry, and the electrostatic force. Exchange interaction effects were discovered independently by Heisenberg and Dirac in 1926.

The exchange interaction is also called the *exchange force*, but is not the same as the exchange forces produced by the exchange of force carriers, such as the electromagnetic force produced between two electrons by the exchange of a photon, or the strong force between two quarks produced by the exchange of a gluon.

Quantum mechanical particles are classified as bosons or fermions. The spin-statistics theorem of quantum field theory demands that all particles with half-integer spin behave as fermions and all particles with integer spin behave as bosons. Multiple bosons may occupy the same quantum state; by the Pauli exclusion principle, however, no two fermions can occupy the same state.

Since electrons have spin 1/2, they are fermions. This means that the overall wavefunction of a system must be anti-symmetric when two electrons are exchanged.

Taking a system with two electrons, we may attempt to model the state of each electron by first assuming the electrons behave independently, and taking wavefunctions in position space of $\Psi_1(r_1)$ for the first electron and $\Psi_2(r_2)$ for the second electron.

We assume that Ψ_1 and Ψ_2 are orthogonal, and that each corresponds to an energy eigenstate of its electron.

Now, if the overall system has spin 1, the spin wave function is symmetric, and we may construct a wavefunction for the overall system in position space by anti-symmetrising the product of these wavefunctions in position space:

$$\Psi_A(r_1, r_2) = \left(\Psi_1(r_1)\Psi_2(r_2) - \Psi_2(r_1)\Psi_1(r_2)\right)/\sqrt{2}. \quad 7$$

On the other hand, if the overall system has spin 0, the spin wave function is anti-symmetric, and we may therefore construct the overall position-space wavefunction by symmetrising the product of the wavefunctions in position space:

$$\Psi_S(r_1, r_2) = \left(\Psi_1(r_1)\Psi_2(r_2) + \Psi_2(r_1)\Psi_1(r_2)\right)/\sqrt{2}.$$

If we assume that the interaction energy between the two electrons, $V_I(r_1,r_2)$, is symmetric, and restrict our attention to the vector space spanned by Ψ_A and Ψ_S, then each of these wavefunctions will yield eigen states for the system energy, and the difference between their energies will be:

$$J = 2\int \Psi_1^*(r_1)\Psi_2^*(r_2)V_I(r_1, r_2)\Psi_2(r_1)\Psi_1(r_2)\,dr_1\,dr_2.$$

Taking into account the different joint spins of these eigen states, we may model this difference by adding a spin-spin interaction term:

$$-J S_1 \cdot S_2,$$

to the Hamiltonian, where S_1 and S_2 are the spin operators of the two electrons. This is one form of the exchange interaction. Despite its form, it is not magnetic in nature. In materials such as iron, this effect favours electrons with parallel spins and is thus a cause of ferromagnetism.

Energy Conservation

In physics, the conservation of energy states that the total amount of energy in an isolated system remains constant, although it may change forms, e.g. friction turns kinetic energy into thermal energy.

In thermodynamics, the first law of thermodynamics is a statement of the conservation of energy for thermodynamic systems, and is the more encompassing version of the conservation of energy. In short, the law of conservation of energy states that energy can not be created or destroyed, it can only be changed from one form to another, such as when electrical energy is changed into heat energy.

From a mathematical point of view, the energy conservation law is a consequence of the shift symmetry of time; energy conservation is implied by the empirical fact that the laws of physics do not change with time itself.

Ancient philosophers as far back as Thales of Miletus had inklings of the conservation of some underlying substance of which everything is made. However, there is no particular reason to identify this with what we know today as "mass-energy" (for example, Thales thought it was water). In 1638, Galileo published his analysis of several situations — including the celebrated "interrupted pendulum"—which can be described (in modernised language) as conservatively converting potential energy to kinetic energy and back again. However, Galileo did not state the process in modern terms and again cannot be credited with the crucial insight.

It was the German Gottfried Wilhelm Leibniz during 1676-1689 who first attempted a mathematical formulation of the kind of energy which is connected with *motion* (kinetic energy). Leibniz noticed that in many mechanical systems (of several masses, m_i each with velocity v_i):

$$\sum_i m_i v_i^2$$

was conserved so long as the masses did not interact. He called this quantity the *vis viva* or *living force* of the system. The

principle represents an accurate statement of the approximate conservation of kinetic energy in situations where there is no friction.

However, many physicists were influenced by the prestige of Sir Isaac Newton in England and of Renee Descartes in France, both of whom had set great store by the conservation of momentum (which holds even in systems with friction), as a guiding principle. Thus the momentum:

$$\sum_i m_i v_i$$

was held by the rival camp to be the conserved *vis viva*. It was largely engineers such as John Smeaton, Peter Ewart, Karl Hotzmann, Gustave-Adolphe Hirn and Marc Seguin who objected that conservation of momentum alone was not adequate for practical calculation and who made use of Leibniz's principle. The principle was also championed by some chemists such as William Hyde Wollaston.

Academics such as John Playfair were quick to point out that kinetic energy is clearly not conserved. This is obvious to a modern analysis based on the second law of thermodynamics but in the 18th and 19th centuries, the fate of the lost energy was still unknown.

Gradually it came to be suspected that the heat inevitably generated by motion under friction, was another form of *vis viva*. In 1783, Antoine Lavoisier and Pierre-Simon Laplace reviewed the two competing theories of *vis viva* and caloric theory. Count Rumford's 1798 observations of heat generation during the boring of cannons added more weight to the view that mechanical motion could be converted into heat, and (as importantly) that the conversion was quantitative and could be predicted (allowing for a universal conversion constant between kinetic energy and heat). *Vis viva* now started to be known as *energy*, after the term was first used in that sense by Thomas Young in 1807.

The recalibration of *vis viva* to:

$$\frac{1}{2}\sum_i m_i v_i^2$$

which can be understood as finding the exact value for the kinetic energy to work conversion constant, was largely the result of the work of Gaspard-Gustave Coriolis and Jean-Victor Poncelet over the period 1819-1839. The former called the quantity *quantite de travail* (quantity of work) and the latter, *travail mecanique* (mechanical work), and both championed its use in engineering calculation.

In a paper *Uber die Natur der Warme*, published in the Zeitschrift fur Physik in 1837, Karl Friedrich Mohr gave one of the earliest general statements of the doctrine of the conservation of energy in the words: "besides the 54 known chemical elements there is in the physical world one agent only, and this is called *Kraft* [energy or work]. It may appear, according to circumstances, as motion, chemical affinity, cohesion, electricity, light and magnetism; and from any one of these forms it can be transformed into any of the others."

A key stage in the development of the modern conservation principle was the demonstration of the mechanical equivalent of heat. The caloric theory maintained that heat could neither be created nor destroyed but conservation of energy entails the contrary principle that heat and mechanical work are interchangeable.

The mechanical equivalence principle was first stated in its modern form by the German surgeon Julius Robert von Mayer. Mayer reached his conclusion on a voyage to the Dutch East Indies, where he found that his patients' blood was a deeper red because they were consuming less oxygen, and therefore less energy, to maintain their body temperature in the hotter climate. He had discovered that heat and mechanical work were both forms of energy, and later, after improving his knowledge of physics, he calculated a quantitative relationship between them.

Meanwhile, in 1843 James Prescott Joule independently discovered the mechanical equivalent in a series of experiments. In the most famous, now called the "Joule apparatus", a descending weight attached to a string caused a paddle immersed in water to rotate.

He showed that the gravitational potential energy lost by the weight in descending was equal to the thermal energy (heat) gained by the water by friction with the paddle.

Over the period 1840-1843, similar work was carried out by engineer Ludwig A. Colding though it was little-known outside his native Denmark.

Both Joule's and Mayer's work suffered from resistance and neglect but it was Joule's that, perhaps unjustly, eventually drew the wider recognition.

Drawing on the earlier work of Joule, Sadi Carnot and Emile Clapeyron, in 1847, Hermann von Helmholtz postulated a relationship between mechanics, heat, light, electricity and magnetism by treating them all as manifestations of a single force (*energy* in modern terms).

He published his theories in his book *Uber die Erhaltung der Kraft* (*On the Conservation of Force*, 1847). The general modern acceptance of the principle stems from this publication.

In 1877, Peter Guthrie Tait claimed that the principle originated with Sir Isaac Newton, based on a creative reading of propositions 40 and 41 of the Philosophiae Naturalis Principia Mathematica. This is now generally regarded as nothing more than an example of Whig history.

Thermodynamic Law

For a thermodynamic system with a fixed number of particles, the first law of thermodynamics may be stated as:

$$\delta Q = \mathrm{d}U + \delta W, \text{ or equivalently, } \mathrm{d}U = \delta Q - \delta W,$$

where δQ is the amount of energy added to the system by a heating process, δW is the amount of energy lost by the system due to work done by the system on its surroundings and $\mathrm{d}U$ is the increase in the internal energy of the system.

The δ's before the heat and work terms are used to indicate that they describe an increment of energy which is to be interpreted somewhat differently than the $\mathrm{d}U$ increment of internal energy. Work and heat are *processes* which add or subtract energy, while

the internal energy U is a particular *form* of energy associated with the system.

Thus the term "heat energy" for δQ means "that amount of energy added as the result of heating" rather than referring to a particular form of energy. Likewise, the term "work energy" for δW means "that amount of energy lost as the result of work".

The most significant result of this distinction is the fact that one can clearly state the amount of internal energy possessed by a thermodynamic system, but one cannot tell how much energy has flowed into or out of the system as a result of its being heated or cooled, nor as the result of work being performed on or by the system.

The first law can be written exclusively in terms of system variables. For a simple compressible system, the work performed by the system may be written:

$$\delta W = P \mathrm{d} V,$$

where P is the pressure and dV is a small change in the volume of the system, each of which are system variables. The heat energy may be written:

$$\delta Q = T \mathrm{d} V,$$

where T is the temperature and $\mathrm{d}S$ is a small change in the entropy of the system. Temperature and entropy are also system variables.

Significance of Mechanics

In mechanics, conservation of energy is usually stated as:

$$E = T + V$$

Actually this is the particular case of the more general conservation law:

$$\sum_{i=1}^{N} p_i \dot{q}_i - L = \text{const and } p_i = \frac{\partial L}{\partial \dot{q}_i}$$

where L is the Lagrangian function. For this particular form to be valid, the following must be true:

- The system is scleronomous (neither kinetic nor potential energy are explicit functions of time).
- The kinetic energy is a quadratic form with regard to velocities.
- The potential energy doesn't depend on velocities.

Contemporary Physics

Noether's Theorem: The conservation of energy is a common feature in many physical theories. It is understood as a consequence of Noether's theorem, which states every symmetry of a physical theory has an associated conserved quantity; if the theory's symmetry is time invariance then the conserved quantity is called "energy".

In other words, if the theory is invariant under the continuous symmetry of time translation then its energy (which is canonical conjugate quantity to time) is conserved. Conversely, theories which are not invariant under shifts in time (for example, systems with time dependent potential energy) do not exhibit conservation of energy — unless we consider them to be exchanging energy with another, external system so that the theory of the enlarged system becomes time invariant again. Since any time-varying theory can be embedded within a time invariant meta-theory energy conservation can always be recovered by a suitable redefinition of what energy is. Thus conservation of energy is valid in all modern physical theories, such as special and general relativity and quantum theory (including QED).

Advantages of Relativity

With the discovery of special relativity by Albert Einstein, energy was found to be one component of an energy-momentum 4-vector. Each of the four components (one of energy and three of momentum) of this vector is separately conserved in any given inertial reference frame. Also conserved is the vector length (Minkowski norm), which is the rest mass. The relativistic energy of a single massive particle contains a term related to its rest mass in addition to its kinetic energy of motion. In the limit of zero kinetic energy (or equivalently in the rest frame of the massive

particle, or the centre-of-momentum frame for objects or systems), the total energy of particle or object (including internal kinetic energy in systems) is related to its rest mass via the famous equation $E = mc^2$.

Thus, the rule of *conservation of energy in special relativity* was shown to be a special case of a more general rule, alternatively called the conservation of mass and energy, the conservation of mass-energy, the conservation of energy-momentum, the conservation of invariant mass or now usually just referred to as conservation of energy.

In general relativity conservation of energy-momentum is expressed with the aid of a stress-energy-momentum pseudotensor.

Quantum Theory

In quantum mechanics, energy is defined as proportional to the time derivative of the wave function. Lack of commutation of the time derivative operator with the time operator itself mathematically results in an uncertainty principle for time and energy: the longer the period of time, the more precisely energy can be defined (energy and time become a conjugate Fourier pair). However, quantum theory in general, and the uncertainty principle specifically, do not violate energy conservation.

Atmospheric Thermodynamics

In the physical sciences, atmospheric thermodynamics is the study of heat and energy transformations in the earth's atmospheric system. Following the fundamental laws of classical thermodynamics, atmospheric thermodynamics studies such phenomenon as properties of moist air, formation of clouds, atmospheric convection, boundary layer meteorology, and vertical stabilities in the atmosphere. Atmospheric thermodynamics forms a basis for cloud microphysics and convection parameterisations in numerical weather models, and is used in many climate considerations, including convective-equilibrium climate models.

Atmospheric thermodynamics focuses on water and its transformations. Areas of study include the law of energy conservation, the ideal gas law, specific heat capacities, adiabatic processes (in which entropy is conserved), and moist adiabatic

processes. Most of tropospheric gases are treated as ideal gases and water vapour is considered as one of the most important trace components of air.

Advanced topics are phase transitions of water, homogeneous and inhomogeneous nucleation, effect of dissolved substances on cloud condensation, role of supersaturation on formation of ice crystals and cloud droplets. Considerations of moist air and cloud theories typically involve various temperatures, such as equivalent potential temperature, wet-bulb and virtual temperatures. Connected areas are energy, momentum, and mass transfer, turbulence interaction between air particles in clouds, convection, dynamics of tropical cyclones, and large scale dynamics of the atmosphere.

The major role of atmospheric thermodynamics is expressed in terms of adiabatic and diabatic forces acting on air parcel included in primitive equations of air motion either as grid resolved or subgrid parameterisations. These equations form a basis for the numerical weather and climate predictions.

History

In the early 1800s, thermodynamicists such as Sadi Carnot, Rudolf Clausius, and Emile Clapeyron worked to developed and to build mathematical models on the dynamics of bodies fluids and vapours related to the combustion and pressure cycles of atmospheric steam engines; one example is the Clausius-Clapeyron equation. In 1873, thermodynamicist Willard Gibbs published "Graphical Methods in the Thermodynamics of Fluids."

These sorts of foundations naturally began to be applied towards the development of theoretical models of atmospheric thermodynamics which drew the attention of the best minds. Papers on atmospheric thermodynamics appeared in the 1860s that treated such topics as dry and moist adiabatic processes. In 1884, Heinrich Hertz devised first atmospheric thermodynamic diagram (emagram). Pseudo-adiabatic process was coined by von Bezold describing air as it is lifted, expands, cools, and eventually precipitates its water vapour; in 1888 he published voluminous work entitled "On the thermodynamics of the atmosphere".

In 1911, von Alfred Wegener published a book "Thermodynamik der Atmosphare", Leipzig, J. A. Barth. From here the development of atmospheric thermodynamics as a branch of science began to take root. The term "atmospheric thermodynamics", itself, can be traced to Frank W. Verys 1919 publication: "The radiant properties of the earth from the standpoint of atmospheric thermodynamics" (Occasional scientific papers of the Westwood Astrophysical Observatory). By the late 1970s various textbooks on the subject began to appear. Today, atmospheric thermodynamics is an integral part of weather forecasting.

Chronology

- 1751 Charles Le Roy recognised dew point temperature as point of saturation of air.
- 1782 Jacques Charles made hydrogen balloon flight measuring temperature and pressure in Paris.
- 1784 Concept of variation of temperature with height was suggested.
- 1801-1803 John Dalton developed his laws of pressures of vapours.
- 1804 Joseph Luis Guy-Lussac made balloon ascent to study weather.
- 1805 Pierre Simon Laplace developed his law of pressure variation with height.
- 1841 James Pollard Espy publishes paper on convection theory of cyclone energy.
- 1889 Herman von Helmholtz and John William von Bezold used the concept of potential temperature, von Bezold used adiabatic lapse rate and pseudoadiabat.
- 1893 Richard Asman constructs first aerological sonde (pressure-temperature-humidity).
- 1894 John Wilhelm von Bezold used concept of equivalent temperature.
- 1926 Sir Napier Shaw introduced tephigram.
- 1933 Tor Bergeron published paper on "Physics of Clouds and Precipitation" describing precipitation from

supercooled (due to condensational growth of ice crystals in presence of water drops).

- 1946 Vincent J. Schaeffer and Irving Langmuir performed the first cloud-seeding experiment.
- 1986 K. Emanuel conceptualises tropical cyclone as Carnot heat engine.

Applications

Tropical Cyclone Carnot Cycle: The thermodynamic structure of the hurricane can be modelled as a heat engine running between sea temperature of about 300K and tropopause which has temperature of about 200K. Parcels of air travelling close to the surface take up moisture and warm, asceding air expands and cools releasing moisture (rain) during the condensation. This release of latent heat energy during the condensation provides mechanical energy for the hurricane.

Both decreasing temperature of upper troposphere or increasing temperature of atmosphere close to the surface will increase on maximum winds observed in hurricanes. When applied to hurricane dynamics it defines Carnot heat engine cycle and predicts maximum hurricane intensity.

Clausius-Clapeyron Equation

The Clausius Clapeyron equation governs the water-holding capacity of the atmosphere, which increases by about 7 per cent per degree Celsius increase in temperature. Saturation water vapour pressure is given by:

$$e_s(T) = 6.112 \exp\left(\frac{17.67\,T}{T + 243.5}\right)$$

where $e_s(T)$ is in hPa, and T is in Celsius. Neglecting weak dependence of the denominator on temperature one notices that saturation water vapour pressure changes approximately exponentialy with T. Therefore, when temperature increases in the atmosphere due to greenhouse gases the absolute humidity should go up. This results is one of the most celebrated statements in the debate about global climate change. However, this purely thermodynamic argument is subject of considerable debate because

convective processes might cause extensive drying due to increased areas of subsidence, efficiency of precipitation could be influenced by the intensity of convection, and because cloud formation is related to relative humidity.

Process of Adiabatic

In thermodynamics, an adiabatic process or an isocaloric process is a process in which no heat is transferred to or from the working fluid. The term "adiabatic" literally means impassable (from a dia bainein), corresponding here to an absence of heat transfer. For example, an adiabatic boundary is a boundary that is impermeable to heat transfer and the system is said to be adiabatically (or thermally) insulated; an insulated wall approximates an adiabatic boundary. Another example is the adiabatic flame temperature, which is the temperature that would be achieved by a flame in the absence of heat loss to the surroundings. An adiabatic process that is reversible is also called a isentropic process.

The opposite extreme — of maximum heat transfer with the surroundings, causing the temperature to remain constant— is known as an isothermal process. Since temperature is thermodynamically conjugate to entropy, the isothermal process is conjugate to the adiabatic process for reversible transformations.

A transformation of a thermodynamic system can be considered adiabatic when it is quick enough that no significant heat is transferred between the system and the outside. The adiabatic process can also be called quasi-static. At the opposite, a transformation of a thermodynamic system can be considered isothermal if it is slow enough so that the system's temperature remains constant by heat exchange with the outside.

Adiabatic Heating and Cooling

Adiabatic heating and cooling are processes that commonly occur due to a change in the pressure of a gas. Adiabatic heating occurs when the pressure of a gas is increased. An example of this is what goes on in a bicycle pump. After using a pump to inflate a pneumatic tire or ball, the barrel of the pump is found to have

heated up as a result of adiabatic heating. Diesel engines rely on adiabatic heating during their compression stroke to elevate the temperature sufficiently to ignite the fuel. Adiabatic heating also occurs in the Earth's atmosphere when an air mass descends, for example, in a katabatic wind or Foehn wind flowing downhill.

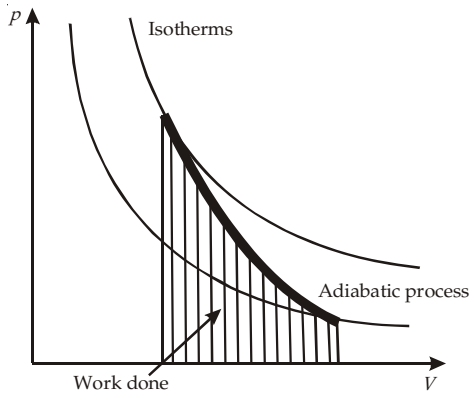
Adiabatic cooling occurs when the pressure of a substance is decreased, such as when it expands into a larger volume. An example of this is when the air is released from a pneumatic tire; the outlet air will be noticeably cooler than the tire. Adiabatic cooling does not have to involve a fluid.

One technique used to reach very low temperatures (thousandths and even millionths of a degree above the theory of absolute zero) is adiabatic demagnetisation, where the change in magnetic field on a magnetic material is used to provide adiabatic cooling. Adiabatic cooling also occurs in the Earth's atmosphere with orographic lifting and lee waves, and this can form pileus or lenticular clouds if the air is cooled below the dew point.

Such temperature changes can be quantified using the ideal gas law, or the hydrostatic equation for atmospheric processes.

It should be noted that no process is truly adiabatic. Many processes are close to adiabatic and can be easily approximated by using an adiabatic assumption, but there is always some heat loss. There is no such thing as a perfect insulator.

Ideal Gas



For a simple substance, during an adiabatic process in which the volume increases, the internal energy of the working substance must necessarily decrease.

The mathematical equation for an ideal fluid undergoing an adiabatic process is:

$$PV^{\gamma} = \text{constant}$$

where P is pressure, V is volume, and

$$\gamma = \frac{C_P}{C_V} = \frac{\alpha + 1}{\alpha},$$

C_P being the specific heat for constant pressure and C_V being the specific heat for constant volume. α comes from the number of degrees of freedom divided by 2 (3/2 for monatomic gas, 5/2 for diatomic gas). For a monatomic ideal gas, γ = 5/3, and for a diatomic gas (such as nitrogen and oxygen, the main components of air) γ = 7/5. Note that the above formula is only applicable to classical ideal gases and not Bose-Einstein or Fermi gases.

For adiabatic processes, it is also true that:

$$VT^{\alpha} = \text{constant}$$

where T is temperature in kelvins. This can also be written as:

$$TV^{\gamma-1} = \text{constant}$$

Derivation of Continuous Formula

The definition of an adiabatic process is that heat transfer to the system is zero, δQ = 0. Then, according to the first law of thermodynamics:

$$(1) \qquad dU + \delta W = \delta Q = 0$$

where dU is the change in the internal energy of the system and δW is work done *by* the system. Any work (δW) done must be done at the expense of internal energy U, since no heat δQ is being supplied from the surroundings. Pressure-volume work δW done *by* the system is defined as:

$$(2) \qquad \delta Q = P\,dV.$$

However, P does not remain constant during an adiabatic process but instead changes along with V.

It is desired to know how the values of dP and dV relate to each other as the adiabatic process proceeds. For an ideal gas the internal energy is given by,

$$(3) \qquad U = \alpha nRT$$

where R is the universal gas constant and n is the number of moles in the system (a constant).

Differentiating Equation (3) and use of the ideal gas law yields

$$(4) \qquad \mathrm{d}U = \alpha nRdT = \alpha d(PV) = \alpha\left(P\,dV + V\,dP\right).$$

Equation (4) is often expressed as $\mathrm{d}U = \mathrm{n}C_v\,dT$ because $C_V = \pm R$. Now substitute equations (2), (3), and (4) into equation (1) to obtain

$$-P\,dV = \alpha P\,dV + \alpha V\,dP$$

simplify,

$$-(\alpha+1)\,P\,dV = \alpha\,V\,dP$$

divide both sides by PV,

$$-(\alpha+1)\frac{dV}{V} = \alpha\frac{dP}{P}.$$

After integrating the left and right sides from Vo to V and from Po to P and changing the sides respectively,

$$\mathrm{In}\left(\frac{P}{P_0}\right) = -\frac{\alpha+1}{\alpha}\mathrm{In}\left(\frac{V}{V_0}\right).$$

Exponentiate both sides,

$$\left(\frac{P}{P_0}\right) = \left(\frac{V}{V_0}\right)^{-\frac{\alpha+1}{\alpha}},$$

eliminate the negative sign,

$$\left(\frac{P}{P_0}\right) = \left(\frac{V_0}{V}\right)^{\frac{\alpha+1}{\alpha}}.$$

Therefore

$$\left(\frac{P}{P_0}\right)\left(\frac{V_0}{V}\right)^{\frac{\alpha+1}{\alpha}} = 1$$

and

$$PV^{\frac{\alpha+1}{\alpha}} = P_0 V_0^{\frac{\alpha+1}{\alpha}} = PV^{\eta} = \text{constant}.$$

Derivation of Discrete Formula

The change in internal energy of a system, measured from state 1 to state 2, is equal to

$$(1) \qquad \delta U = \alpha R n_2 T_2 - \alpha R_1 T_1 = \alpha R\left(n_2 T_2 - n_1 T_1\right)$$

At the same time, the work done by the pressure-volume changes as a result from this process, is equal to

$$(2) \qquad \delta W = P_2 V_2 - P_1 V_1$$

Since we require the process to be adiabatic, the following equation needs to be true

$$(3) \qquad \delta U + \delta W = 0$$

Substituting (1) and (2) in (3) leads to

$$\alpha R(n_2 T_2 - n_1 T_1) + \left(P_2 V_2 - P_1 V_1\right) = 0$$

or

$$\frac{\left(P_2 V_2 - P_1 V_1\right)}{-\left(n_2 T_2 - n_1 T_1\right)} = \alpha R$$

If it's further assumed that there are no changes in molar quantity (as often in practical cases), the formula is simplified to this one:

$$\frac{\left(P_2 V_2 - P_1 V_1\right)}{-\left(T_2 - T_1\right)} = \alpha n R$$

Adiabatic Graphing

Properties of adiabats on a P-V diagram are:

1. Every adiabat asymptotically approaches both the V axis and the P axis (just like isotherms).
2. Each adiabat intersects each isotherm exactly once.
3. An adiabat looks similar to an isotherm, except that during an expansion, an adiabat loses more pressure than an isotherm, so it has a steeper inclination (more vertical).

4. If isotherms are concave towards the "northeast" direction (45), then adiabats are concave towards the "east northeast" (31).
5. If adiabats and isotherms are graphed severally at regular changes of entropy and temperature, respectively (like altitude on a contour map), then as the eye moves towards the axes (towards the southwest), it sees the density of isotherms stay constant, but it sees the density of adiabats grow. The exception is very near absolute zero, where the density of adiabats drops sharply and they become rare.

Air Parcel

Definition: An air parcel is an imaginary volume of air used by meteorologists to conceptualise the thermodynamic fluid motions of the atmosphere for use in weather forecasting. For mathematical simplicity, an air parcel is usually considered a rigid cube which has limited interactions with surrounding environmental air. The dimensions of this parcel are determined by the atmospheric scale under study.

The importance of the concept of air parcels in meteorology lies in its ability to assist meteorologists in conceptualising how areas of an atmosphere will feature rising and sinking motions, the magnitude of those motions, and the possibility of ensuing clouds and precipitation. Many older numerical weather prediction models used the conceptual models of air parcels.

Buoyancy: The primary concern of air parcels for meteorologists is whether they will rise, sink or remain steady. To understand this motion, it is helpful to consider thermodynamic buoyancy equations.

One such equation can be derived as follows.

Let the weight of the parcel be Vg, while the pressure gradient force is $\frac{\mathrm{dp}}{\mathrm{dz}}$. Using the hydrostatic assumption (both terms equal zero) and observing that $\rho V\frac{dw}{dt}=0$ in this case gives $\rho V\frac{dw}{dt}=-V\frac{dp}{dz}-V\rho g$ Allowing for the pressure to be the sum of

the environmental and parcel pressure $p = \bar{p} + p'$ yields the following as the buoyancy equation: $\frac{dw}{dt} = -\frac{1}{\rho}\frac{dp'}{dz} + g\left(\frac{\bar{p} + p'}{p}\right)$ (where $\frac{dw}{dt}$ is the buoyancy and w is the vertical motion $\frac{dp}{dz}$).

The first term of this equation details the effects of pressure perturbation, wherein the parcel must literally "push" surrounding air out of the way to move through the environment. This effect tends to retard parcel acceleration, and is an important consideration in thunderstorm updrafts.

From the second term, it is easily seen that if the parcel is denser than its environment (i.e. $\bar{p} > p'$), it will have a negative buoyancy and will thus sink relative to the environment.

The opposite case is also true ($\bar{p} < p'$ causes rising), while neutral buoyancy is achieved by the two terms being equal. Alternatively, the density can be replaced by the virtual temperature (virtual temperature is necessary because of the density differences between air and water vapour). Basically, an air parcel which is warmer than its environment will rise. Conversely, if it is cooler, it will sink.

In addition to the basic buoyancy equation, there are two other factors which govern parcel movement; precipitation loading, amount of parcel water vapour, and entrainment.

If moisture precipitates within the parcel, the amount of precipitation in the parcel versus the amount in the environment will affect the parcel's buoyancy (every drop of precipitation adds extra weight to the parcel). This is described mathematically as a third term in the buoyancy equation $\bar{q}_c + q_c$. Lastly, if mixing of environmental air with parcel air is considered, the properties of the parcel will change to reflect that of the environment. This process is known as entrainment and is mathematically represented as a coefficient lambda $\lambda = \frac{1}{m}\frac{dm}{dz}$ (m is the mass of air entrained into the parcel).

Combining all the factors yields an equation which suitably expresses the buoyancy of an air parcel

$$\frac{dw}{dt} = -\frac{1}{\rho}\frac{dp'}{dz} + g\left(\left(\frac{\bar{\rho}+\rho}{\rho}\right) + \left(\bar{q}_c + q_c\right)\right) - \lambda * w^2$$

To summarise in simpler terms, there are four major factors which affect the buoyancy of air parcels. Pressure perturbation decelerates parcels because moving parcels have to "push" surrounding air out of the way.

Density differences between the air parcel and its environment can accelerate or decelerate the parcel. Precipitation within the parcel can act a drag on upward motions. Lastly, mixing of the parcel air with environmental air acts to decrease buoyancy.

Applications

When using a Skew-T chart (a Tephigram chart can be used for this as well) to diagnose the atmosphere, meteorologists assume that air interacting with the environment will behave as an air parcel (that individual parcels of air will not significantly interact with the environment). A meteorologist can then trace a parcel up or down the appropriate adiabat to determine the desired variable, which can be temperature, pressure, moisture content, and so forth.

The notion of air parcels is partially confirmed by examining convective plumes within cumulus congestus clouds. Individual columns of rising air can be observed in such clouds. In time lapse photography, areas of convection can be observed, and the individual terms of the buoyancy equation can be observed in action (e.g. warmer air rising, entrainment of environmental air causing decay, etc.).

Equation of Arden Buck

The Arden Buck Equation is an equation describing the saturation vapour pressure of water at various temperatures. It is based on more recent experiments than the Goff-Gratch equation.

A set of several equations were developed, each of which is applicable in a different situation.

The one suggested for common use in 1996 is:

$$p_{u.} = 6.1121 \exp\left(\frac{(18.678 - t/234.5)\,t}{257.14 + t}\right)$$

where:

p_w is the saturation vapour pressure (hPa)

exp is the exponential function

t is the absolute air temperature in degrees Celsius.

Significance of Atmospheric Pressure

Atmospheric pressure is the pressure at any point in the Earth's atmosphere. In most circumstances atmospheric pressure is closely approximated by the hydrostatic pressure caused by the weight of air above the measurement point. Low pressure areas have less atmospheric mass above their location, whereas high pressure areas have more atmospheric mass above their location. Similarly, as elevation increases there is less overlying atmospheric mass, so that pressure decreases with increasing elevation. A column of air 1 square inch in cross-section, measured from sea level to the top of the atmosphere, would weigh approximately 14.7 lbf. A 1 m column of air would weigh about 100 kilonewtons (equivalent to a mass of 10.2 tonnes at the surface).

Standard Atmospheric Pressure

The standard atmosphere (symbol: atm) is a unit of pressure and is defined as being precisely equal to 101.325 kPa. This value is intended to represent the mean sea level pressure at the latitude of Paris, France, and as a practical matter, approximates the mean sea level pressure for many of the industrialised nations (those with latitudes similar to Paris). One standard atmosphere is standard pressure used for pneumatic fluid power (ISO R554), and in the aerospace (ISO 2533) and petroleum (ISO 5024) industries.

In 1982, the International Union of Pure and Applied Chemistry (IUPAC) recommended that for the purposes of specifying the physical properties of substances, "*the standard pressure*" should be defined as precisely 100 kPa ($\approx$ 760.062 torr)

or 29.9230 inHg rather than the 101.325 kPa value of "one standard atmosphere". This value is used as the standard pressure for the compressor and the pneumatic tool industries (ISO 2787).

In the United States, compressed air flow is often measured in "standard cubic feet" per unit of time, where the "standard" means the equivalent quantity of air at standard temperature and pressure. However, this standard atmosphere is defined slightly differently: temperature = 68 F (20 C), air density = 0.0765 lb/ft (1.225 kg/m), altitude = sea level, and relative humidity = 0 per cent. In the air conditioning industry, the standard is often temperature = 32 F (0 C) instead.

For natural gas, the petroleum industry uses a standard temperature of 60 F (15.6 C), pressure 14.73 psia.

Mean Sea Level Pressure

Mean sea level pressure (MSLP or QFF) is the pressure at sea level or (when measured at a given elevation on land) the station pressure reduced to sea level assuming an isothermal layer at the station temperature.

This is the pressure normally given in weather reports on radio, television, and newspapers or on the Internet. When barometers in the home are set to match the local weather reports, they measure pressure reduced to sea level, not the actual local atmospheric pressure.

The reduction to sea level means that the normal range of fluctuations in pressure is the same for everyone. The pressures which are considered high pressure or low pr*essure* do not depend on geographical location. This makes isobars on a weather map meaningful and useful tools.

The altimeter setting in aviation, set either QNH or QFE, is another atmospheric pressure reduced to sea level, but the method of making this reduction differs slightly.

- QNH—atmospheric pressure (Q) at nautical (N) height (H) or sea-level pressure. Barometric altimeter setting which will cause the altimeter to read airfield elevation when on the airfield.

In ISA temperature conditions, the altimeter will read altitude above mean sea level in the vicinity of the airfield.

- QFE barometric altimeter setting which will cause an altimeter to read zero when at the reference datum of a particular airfield (generally a runway threshold). In ISA temperature conditions the altimeter will read height above the datum in the vicinity of the airfield.

Average sea-level pressure is 101.325 kPa (1013.25 mbar) or 29.921 inches of mercury (inHg). In aviation weather reports (METAR), QNH is transmitted around the world in millibars or hectopascals (1 millibar = 1 hectopascal), except in the United States where it is reported in inches (or hundredths of inches) of mercury.

The United States also reports sea-level pressure SLP, which is reduced to sea level by a different method, in the remarks section, not an internationally transmitted part of the code, in hectopascals or millibars.

In Canada's public weather reports, sea level pressure is reported in kilopascals. In the weather code, three digits are all that is needed; decimal points and the one or two most significant digits are omitted: 1013.2 mbar or 101.32 kPa is transmitted as 132; 1000.0 mbar or 100.00 kPa is transmitted as 000; 998.7 mbar or 99.87 kPa is transmitted as 987; etc.

The highest *sea-level pressure* on Earth occurs in Siberia, where the Siberian High often attains a *sea-level pressure* above 1032.0 mbar. The lowest measurable *sea-level pressure* is found at the centres of hurricanes (typhoons, baguios).

Altitude Atmospheric Pressure Variation

Pressure varies smoothly from the earth's surface to the top of the mesosphere. Although the pressure changes with the weather, NASA has averaged the conditions for all parts of the earth year-round.

The following is a list of air pressures (as a fraction of one atmosphere) with the corresponding average altitudes. The table gives a rough idea of air pressure at various altitudes.

fraction of 1 atm	*average altitude*	
	(m)	*(ft)*
1	0	0
1/2	5,486	18,000
1/3	8,376	27,480
1/10	16,132	52,926
1/100	30,901	101,381
1/1000	48,467	159,013
1/10000	69,464	227,899
1/100000	96,282	283,076

Calculating Variation with Altitude

There are two different equations for computing pressure at various height regimes below 86 km (or 278,400 feet). Equation 1 is used when the value of standard temperature lapse rate is not equal to zero and equation 2 is used when standard temperature lapse rate equals zero.

Equation 1:

$$P = P_b \cdot \left[\frac{T_b}{T_b + L_b \cdot (h - h_b)} \right]^{\frac{g_0 \cdot M}{R^* \cdot L_b}}$$

Equation 2:

$$P = P_b \cdot \exp\left[\frac{-g_0 \cdot M \cdot (h - h_b)}{R^* \cdot T_b} \right]$$

where

P = Static pressure (pascals)

T = Standard temperature (kelvins) =\

L = Standard temperature lapse rate (kelvins per metre)

h = Height above sea level (metres)

R^* = Universal gas constant: 8.31432 10^3 N m / (kmol K)

g_0 = Gravitational constant (9.80665 m/s)

M = Molar mass of Earth's air (28.9644 g/mol)

Or converted to English units: where

P = Static pressure (inches of mercury)

T = Standard temperature (kelvins)

L = Standard temperature lapse rate (kelvins per foot)

h = Height above sea level (feet)

R^* = Universal gas constant (using feet and kelvins and gram moles: $8.9494596\ 10^4$ kg ft^2 s^{-2} K^{-1} kmol^{-1})

g_0 = Gravitational constant (32.17405 ft/s)

M = Molar mass of Earth's air (28.9644 g/mol)

The value of subscript b ranges from 0 to 6 in accordance with each of seven successive layers of the atmosphere shown in the table below. In these equations, g_0, M and R^* are each single-valued constants, while P, L, T, and h are multivalued constants in accordance with the table below. It should be noted that the values used for M, g_0, and R^* are in accordance with the US Standard Atmosphere, 1976, and that the value for R^* in particular does not agree with standard values for this constant.

The reference value for P_b for $b = 0$ is the defined sea level value, P_0 = 101,325 pascals or 29.92126 inHg. Values of P_b of b = 1 through $b = 6$ are obtained from the application of the appropriate member of the pair equations 1 and 2 for the case when $h = h_{b+1}$.:

Subscript b	*Height Above Sea Level*		*Static Pressure*		*Standard Temperature*	*Temperature Lapse Rate*	
	(m)	*(ft)*	*(pascals)*	*(inHg)*	*(K)*	*(K/m)*	*K/ft)*
0	0	0	101325	29.92126	298.15	–0.0065	–0.0019812
1	11,000	36,089	22632.1	6.683245	216.65	0.0	0.0
2	20,000	65,617	5474.89	1.616734	216.65	0.001	0.0003048
3	32,000	104,987	868.019	0.2563258	228.65	0.0028	0.00085344
4	47,000	154,199	110.906	0.0327506	270.65	0.0	0.0
5	51,000	167,323	66.9389	0.01976704	270.65	–0.0028	–0.00085344
6	71,000	232,940	3.95642	0.00116833	214.65	–0.002	–0.0006096

Sample Calculation:

- Find the pressure at 30,000 metres.

First note that 30,000 metres is above 20,000 but below 32,000 so it therefore falls in the range of subscript b = 2 in the chart above. Also note that the temperature lapse rate for that region is not equal to zero; therefore, equation 1 is appropriate.

$$P = P_2 \cdot \left[\frac{T_2}{T_2 + L_2 \cdot (h - h_2)} \right]^{\frac{g_0 \cdot M}{R^* \cdot L_2}}$$

Or

$$P = 5474.89 \cdot \left[\frac{216.65}{216.65 + 0.001 \cdot (30{,}000 - 20{,}000)} \right]^{\frac{9.80665 \cdot 28.9644}{8314.32 \cdot 0.001}}$$

$$P = 5474.89 \cdot \left[\frac{216.65}{226.65} \right]^{34.163195}$$

P = 5474.89 · 0.214044

P = 1171.867 Pascals at 30,000 metres

Atmospheric Pressure based on Height of Water

Atmospheric pressure is often measured with a mercury barometer, and a height of approximately 760 mm (30 inches) of mercury is often used to teach, make visible, and illustrate (and measure) atmospheric pressure. However, since mercury is not a substance that humans commonly come in contact with, water often provides a more intuitive way to conceptualise the amount of pressure in one atmosphere.

One atmosphere (101.325 kPa or 14.7 lbf/in) is the amount of pressure that can lift water approximately 10.3 m (33.9 feet). Thus, a diver at a depth 10.3 metres under water in a freshwater lake experiences a pressure of about 2 atmospheres (1 atm for the air and 1 atm for the water).

Non-professional barometers are generally aneroid barometer or strain gauge based.

Ratio of Bowen

A Bowen ratio is the ratio of energy fluxes from one medium to another by sensible and latent heating respectively. It is calculated by the equation

$$B = \frac{Q_h}{Q_e,}$$

where Q_h is sensible heating and Q_e is latent heating. The quantity was named by Harald Sverdrup after Ira Sprague Bowen (1898-1973), an astrophysicist whose theoretical work on evaporation to air from water bodies made first use of it, and it is used most commonly in meteorology and hydrology. In this context, when the magnitude of B is less than one, a greater proportion of the available energy at the surface is passed to the atmosphere as latent heat than as sensible heat, and the converse is true for values of B greater than one. As $Q_e \to 0$, however, B becomes unbounded making the Bowen ratio a poor choice of variable for use in formulae, especially for arid surfaces. For this reason, the evaporative fraction is sometimes a more appropriate choice of variable representing the relative contributions of the turbulent energy fluxes to the surface energy budget.

The Bowen ratio is related to the evaporative fraction, *EF*, through the equation:

$$EF = \frac{Q_e}{Q_e + Q_h} = \frac{1}{1 + B}.$$

Significance of Carnot Cycle

The Carnot cycle is a particular thermodynamic cycle, modelled on the Carnot heat engine, studied by Nicolas Leonard Sadi Carnot in the 1820s and expanded upon by Benoit Paul Emile Clapeyron in the 1830s and 40s.

Every thermodynamic system exists in a particular state. A thermodynamic cycle occurs when a system is taken through a series of different states, and finally returned to its initial state. In the process of going through this cycle, the system may perform work on its surroundings, thereby acting as a heat engine.

A heat engine acts by transferring energy from a warm region to a cool region of space and, in the process, converting some of that energy to mechanical work. The cycle may also be reversed. The system may be worked upon by an external force, and in the process, it can transfer thermal energy from a cooler system to a warmer one, thereby acting as a refrigerator rather than a heat engine.

The Carnot cycle is a special type of thermodynamic cycle. It is special because it is the most efficient cycle possible for converting a given amount of thermal energy into work or, conversely, for using a given amount of work for refrigeration purposes.

The Carnot cycle when acting as a heat engine consists of the following steps:

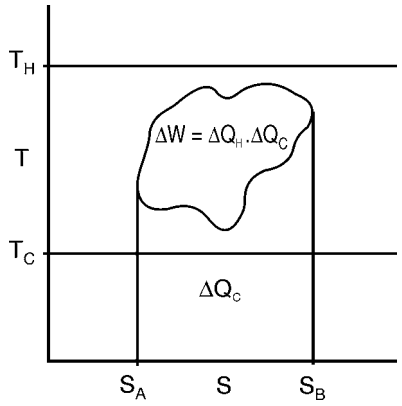
1. Reversible isothermal expansion of the gas at the "hot" temperature, T_H (isothermal heat addition): During this step the expanding gas causes the piston to do work on the surroundings. The gas expansion is propelled by absorption of heat from the high temperature reservoir.
2. Isentropic (Reversible adiabatic) expansion of the gas: For this step we assume the piston and cylinder are thermally insulated, so that no heat is gained or lost. The gas continues to expand, doing work on the surroundings. The gas expansion causes it to cool to the "cold" temperature, T_C.
3. Reversible isothermal compression of the gas at the "cold" temperature, T_C (isothermal heat rejection): Now the surroundings do work on the gas, causing heat to flow out of the gas to the low temperature reservoir.
4. Isentropic compression of the gas: Once again we assume the piston and cylinder are thermally insulated. During this step, the surroundings do work on the gas, compressing it and causing the temperature to rise to T_H. At this point the gas is in the same state as at the start of step 1.

Properties and Significance

Temperature-entropy Diagram: The behaviour of a Carnot engine or refrigerator is best understood by using a temperature-

entropy (TS) diagram, in which the thermodynamic state is specified by a point on a graph with entropy (S) as the horizontal axis and temperature (T) as the vertical axis.

For a simple system with a fixed number of particles, any point on the graph will represent a particular state of the system. A thermodynamic process will consist of a curve connecting an initial state (A) and a final state (B). The area under the curve will be:



$$\Delta Q = \int_A^B T\, ds \tag{1}$$

which is the amount of thermal energy transferred in the process. If the process moves to greater entropy, the area under the curve will be the amount of heat absorbed by the system in that process. If the process moves towards lesser entropy, it will be the amount of heat removed. For any cyclic process, there will be an upper portion of the cycle and a lower portion.

For a clockwise cycle, the area under the upper portion will be the thermal energy absorbed during the cycle, while the area under the lower portion will be the thermal energy removed during the cycle. The area inside the cycle will then be the difference between the two, but since the internal energy of the system must have returned to its initial value, this difference must be the amount of work done by the system over the cycle. Mathematically, for a reversible process we may write the amount of work done over a cyclic process as:

$$\Delta W = \oint PdV = \oint (TdS - dU) \quad (2)$$

Since dU is an exact differential, its integral over any closed loop is zero and it follows that the area inside the loop on a T-S diagram is equal to the total work performed if the loop is traversed in a clockwise direction, and is equal to the total work done on the system as the loop is traversed in a counterclockwise direction.

Evaluation of the above integral is particularly simple for the Carnot cycle. The amount of energy transferred as work is

$$\Delta W = \oint PdV = (T_H - T_C)(S_B - S_A)$$

The total amount of thermal energy transferred between the hot reservoir and the system will be

$$\Delta Q_H = T_H(S_B - S_A)$$

and the total amount of thermal energy transferred between the system and the cold reservoir will be

$$\Delta Q_C = T_C(S_B - S_A).$$

The efficiency is defined to be:

$$\eta = \frac{\Delta W}{\Delta Q_H} = 1 - \frac{T_C}{T_H} \quad (3)$$

where

"W is the work done by the system (energy exiting the system as work),

"Q_H is the heat put into the system (heat energy entering the system),

T_C is the absolute temperature of the cold reservoir, and

T_H is the temperature of the hot reservoir.

This efficiency makes sense for a heat engine, since it is the fraction of the heat energy extracted from the hot reservoir and converted to mechanical work. It also makes sense for a refrigeration cycle, since it is the ratio of energy input to the refrigerator divided by the amount of energy extracted from the hot reservoir.

Carnot's Theorem

For any cycle operating between temperatures T_H and T_C, none can exceed the efficiency of a Carnot cycle.

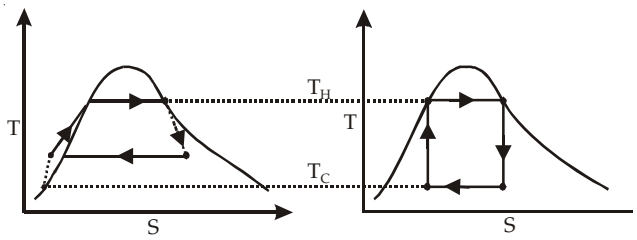


Fig. A real engine (left) compared to the Carnot cycle (right)

The entropy of a real material changes with temperature. This change is indicated by the curve on a T-S diagram. For this figure, the curve indicates a vapour-liquid equilibrium. Irreversible systems and losses of heat (for example, due to friction) prevent the ideal from taking place at every step.

Carnot's theorem is a formal statement of this fact: *No* engine operating between two heat reservoirs can be more efficient *than a Carnot engine operating between the same reservoirs.* Thus, Equation 3 gives the maximum efficiency possible for any engine using the corresponding temperatures. A corollary to Carnot's theorem states that: *All reversible engines operating between the same heat reservoirs are equally efficient.* Rearanging the right side of the equation gives what may be a more easily understood form of the equation.

Namely that the theoretical maximum efficiency of a heat engine equals the difference in temperature between the hot and cold reservoir divided by the absolute temperature of the hot reservoir. To find the absolute temperature in Kelvins, add 273 degrees to the Celsius temperature. Looking at this formula an interesting fact becomes apparent. Lowering the temperature of the cold reservoir will have more effect on the ceiling efficiency of a heat engine than raising the temperature of the hot reservoir by the same amount.

In the real world, this may be difficult to achieve since the cold reservoir is often an existing ambient temperature.

In other words, maximum efficiency is achieved if and only if no new entropy is created in the cycle. Otherwise, since entropy

is a state function, the required dumping of heat into the environment to dispose of excess entropy leads to a reduction in efficiency. So Equation 3 gives the efficiency of any reversible heat engine.

Efficiency of Real Heat Engines

Carnot realised that in reality it is not possible to build a thermodynamically reversible engine, so real heat engines are less efficient than indicated by Equation 3. Nevertheless, Equation 3 is extremely useful for determining the maximum efficiency that could ever be expected for a given set of thermal reservoirs.

Although Carnot's cycle is an idealisation, the expression of Carnot efficiency is still useful. Consider the average temperatures,

$$\langle T_H \rangle = \frac{1}{\Delta S}\int_{Q_{in}} T\, ds$$

$$\langle T_C \rangle = \frac{1}{\Delta S}\int_{Q_{out}} T\, ds$$

at which heat is input and output, respectively. Replace T_H and T_C in Equation (3) by $<T_H>$ and $<T_C>$ respectively.

For the Carnot cycle, or its equivalent, $<T_H>$ is the highest temperature available and $<T_C>$ the lowest. For other less efficient cycles, $<T_H>$ will be lower than T_H, and $<T_C>$ will be higher than T_C. This can help illustrate, for example, why a reheater or a regenerator can improve thermal efficiency.

Exchange Arrangement

Exchange symmetry is derived from a fundamental postulate of quantum statistics, which states that no observable physical quantity should change after exchanging two identical particles.

Because all observables are proportional to the square of the modulus of the wave function the wave function of a system of identical particles must either remain the same, or must change sign upon such an exchange.

Because the exchange of two identical particles is mathematically equivalent to the rotation of each particle by 360

degrees, the symmetricity of the wave function depends on the particle's spin after the rotation operator is applied to it. Integer spin particles do not change the sign of their wave function upon a 360 degree rotation—therefore the sign of the wave function of the entire system does not change. Semi-integer spin particles change the sign of their wave function upon a 360 degree rotation.

Particles for which the wave function does not change sign upon exchange are called bosons, or particles with a symmetric wave function. The particles for which the wave function of the system changes sign are called fermions, or particles with an anti-symmetric wave function.

Fermions therefore obey different statistics (called Fermi-Dirac statistics) than bosons (which obey Bose-Einstein statistics). One of the consequences of Fermi-Dirac statistics is the exclusion principle for fermions—no two fermions can share the same quantum state (in other words, the wave function of two fermions in the same state is zero).

This in turn results in degeneracy pressure for fermions—the strong resistance of fermions to compression into smaller volume. This resistance gives rise to the "stiffness" or "rigidity" of ordinary atomic matter (as atoms contain electrons which are fermions).

Motion and Mass of Transport

Momentum: In classical mechanics, momentum (pl. momenta; SI unit kg m/s) is the product of the mass and velocity of an object. In general, the momentum of an object can be conceptually thought of as how difficult it is to stop the object, as determined by multiplying two factors: its mass and its velocity. As such, it is a natural consequence of Newton's first and second laws of motion. Having a lower speed or having less mass (how we measure inertia) results in having less momentum. Momentum is a conserved quantity, meaning that the total momentum of any closed system (one not affected by external forces, and whose internal forces are not dissipative in nature) cannot be changed.

The concept of momentum in classical mechanics was originated by a number of great thinkers and experimentalists. The first of these was Ibn Sina (Avicenna) *circa* 1000, who referred

to impetus as proportional to weight times velocity. Rene Descartes later referred to *mass times velocity* as the *fundamental force of motion*. Galileo in his Two New Sciences used the term "impeto" (Italian), while Newton's Laws of Motion uses *motus* (Latin), which has been interpreted by subsequent scholars to mean momentum.

Newtonian Mechanics Momentum

If an object is moving in any reference frame, then it has momentum in that frame. It is important to note that momentum is frame dependent. That is, the same object may have a certain momentum in one frame of reference, but a different amount in another frame. For example, a moving object has momentum in a reference frame fixed to a spot on the ground, while at the same time having 0 momentum in a reference frame attached to the object's centre of mass.

The amount of momentum that an object has depends on two physical quantities: the mass and the velocity of the moving object in the frame of reference. In physics, the symbol for momentum is usually denoted by a small bold *p* (bold because it is a vector); so this can be written:

$\mathbf{p} = m\mathbf{v}$

where:

p is the momentum

m is the mass

v the velocity

(using bold text for vectors).

The origin of the use of p for momentum is unclear. It has been suggested that, since *m* had already been used for "mass", the p may be derived from the Latin petere ("to go") or from "progress" (a term used by Leibniz).

The velocity of an object at a particular instant is given by its speed and the direction of its motion at that instant. Because momentum depends on and includes the physical quantity of velocity, it too has a magnitude and a direction and is a vector quantity. For example the momentum of a 5-kg bowling ball

would have to be described by the statement that it was moving westward at 2 m/s. It is insufficient to say that the ball has 10 kg m/s of momentum because momentum is not fully described unless its direction is given.

Motion for a System

Relating to Mass and Velocity: The momentum of a system of objects is the vector sum of the momenta of all the individual objects in the system.

$$\mathrm{p} = \sum_{i=1}^{n} m_i \vec{\mathrm{v}}_i = m_1 \vec{\mathrm{v}}_1 + m_2 \vec{\mathrm{v}}_2 + m_3 \vec{\mathrm{v}}_3 + \mathrm{K} + m_n \vec{\mathrm{v}}_n$$

where,

p is the momentum

m_i is the mass of object i

$\vec{\mathrm{v}}_i$ the vector velocity of object i

n is the number of objects in the system

Referring to Force

Force is equal to the rate of change of momentum:

$$\mathrm{F} = \frac{\mathrm{d}\mathrm{p}}{\mathrm{d}t}.$$

In the case of constant mass, and velocities much less than the speed of light, this definition results in the equation F = ma, commonly known as Newton's second law.

If a system is in equilibrium, then the change in momentum with respect to time is equal to 0:

$$\mathrm{F} = \frac{\mathrm{d}\mathrm{p}}{\mathrm{d}t} = m\mathbf{a} = 0$$

Motion Conservation

The law of conservation of momentum is a fundamental law of nature, and it states that the total momentum of a closed system of objects (which has no interactions with external agents) is

constant. One of the consequences of this is that the centre of mass of any system of objects will always continue with the same velocity unless acted on by a force outside the system.

In an isolated system (one where external forces are absent) the total momentum will be constant: this is implied by Newton's first law of motion. Newton's third law of motion, the law of reciprocal actions, which dictates that the forces acting between systems are equal in magnitude, but opposite in sign, is due to the conservation of momentum.

Since position in space is a vector quantity, momentum (being canonical conjugate of position) is a vector quantity as well—it has direction. Thus, when a gun is fired, although overall movement has increased compared to before the shot was fired, the momentum of the bullet in one direction is equal in magnitude, but opposite in sign, to the momentum of the gun in the other direction. These then sum to zero which is equal to the zero momentum that was present before either the gun or the bullet was moving.

Conservation of Momentum and Collisions

Momentum has the special property that, in a closed system, it is always conserved, even in collisions. Kinetic energy, on the other hand, is not conserved in collisions if they are inelastic. Since momentum is conserved it can be used to calculate unknown velocities following a collision.

A common problem in physics that requires the use of this fact is the collision of two particles. Since momentum is always conserved, the sum of the momenta before the collision must equal the sum of the momenta after the collision:

$$m_1u_1 + m_2u_2 = m_1v_1 + m_2v_2$$

where:

u signifies vector velocity before the collision

v signifies vector velocity after the collision.

Usually, we either only know the velocities before or after a collision and would like to also find out the opposite. Correctly solving this problem means you have to know what kind of

collision took place. There are two basic kinds of collisions, both of which conserve momentum:

- Elastic collisions conserve kinetic energy as well as total momentum before and after collision.
- Inelastic collisions don't conserve kinetic energy, but total momentum before and after collision is conserved.

Relation between Clausius-Clapeyron

In thermodynamics, The Clausius-Clapeyron relation, named after Rudolf Clausius and Emile Clapeyron, is a way of characterising the phase transition between two states of matter, such as solid and liquid.

On a pressure-temperature (P – T) diagram, the line separating the two phases is known as the coexistence curve. The Clausius-Clapeyron relation gives the slope of this curve. Mathematically,

$$\frac{dP}{dT} = \frac{L}{T\Delta V}$$

where $\mathrm{d}P/\mathrm{d}T$ is the slope of the coexistence curve, L is the latent heat, T is the temperature, and ΔV is the volume change of the phase transition.

Derivation

Suppose two phases, I and II, are in contact and at equilibrium with each other. Then the chemical potentials are related by $\mu_I = \mu_{II}$. Along the coexistence curve, we also have $\mathrm{d}\mu_I = \mathrm{d}\mu_{II}$. We now use the Gibbs-Duhem relation $\mathrm{d}\mu = -s\mathrm{d}T + \nu\,\mathrm{d}P$, where s and ν are, respectively, the entropy and volume per particle, to obtain

$$-(s_I - s_{II})\mathrm{d}T + (\nu_I - \nu_{II})\mathrm{d}P = 0.$$

Hence, rearranging, we have

$$\frac{\mathrm{d}P}{\mathrm{d}T} = \frac{s_I - s_{II}}{v_I - v_{II}}.$$

From the relation between heat and change of entropy in a reversible process $\delta Q = T\,\mathrm{d}S$, we have that the quantity of heat added in the transformation is

$$L = T(s_I - s_{II}).$$

Combining the last two equations we obtain the standard relation.

Applications

Chemistry: This equation gives the quantitative dependence of the vapour pressure of a liquid on its temperature. It can be used to predict the temperature at a certain pressure, given the temperature at another pressure, or vice versa. If the corresponding temperature and pressure is known at two points, the enthalpy of vaporisation can also be determined from this equation.

$\ln (P2/P1) = (\Delta Hvap/R) * (1/T1 - 1/T2)$

where

T1 and P1 are a corresponding temperature (K) and pressure

T2 and P2 are the corresponding temperature and pressure at another point

Δ Hvap is the enthalpy of vaporisation

R is the gas constant (8.314)

Chemical Engineering: A specific derivation of the equation is used in chemistry and chemical engineering to estimate the vapour pressure of a substance based on the heat of vaporisation of that substance, and on the temperature of the system under consideration. The equation is as follows:

$$\ln p^* = -\frac{\Delta \tilde{H}_v}{RT} + B$$

where

p^* is the vapour pressure (mmHg)

$\Delta \tilde{H}_v$ is the enthalpy of vaporisation (kJ/mole)

R is the gas constant

T is the temperature (Kelvin)

B is a variable based on the substance and the system parameters.

Meteorology: In meteorology, a specific derivation of the Clausius-Clapeyron equation is used to describe dependence of saturated water vapour pressure on temperature. This is similar to its use in chemistry and chemical engineering.

It plays crucial role in the current debate on climate change because its solution predicts exponential behaviour of saturation water vapour pressure (and, therefore water vapour concentration) as a function of temperature. In turn, because water vapour is a greenhouse gas, it might lead to further increase in the sea surface temperature leading to runaway greenhouse effect. Debate on iris hypothesis and intensity of tropical cyclones dependence on temperature depends in part on "Clausius-Clapeyron" solution.

Clausius-Clapeyron equations is given for typical atmospheric conditions as

$$\frac{\mathrm{d}e_s}{\mathrm{d}T} = \frac{L_v e_s}{R_v \mathrm{T}^2}$$

where e_s is saturation water vapour pressure, T is a temperature, L_v is latent heat of evaporation, R_v is water vapour gas constant. One can solve this equation to give

$$e_s(T) = 6.112 \exp\left(\frac{17.67\,T}{T + 243.5}\right)$$

where $e_s(T)$ is in hPa, and T is in Celsius. Thus, neglecting the weak variation of (T+243.5) at normal temperatures, one observes that saturation water vapour pressure changes exponentially with T.

Example

One of the uses of this equation is to determine if a phase transition will occur in a given situation. Consider the question of how much pressure is needed to melt ice at a temperature ΔT below 0 C. We can assume

$$\Delta P = \frac{L}{T \Delta V} \Delta T$$

and substituting in

L = 3.34 10^5 J/kg (latent heat of water),

T = 273K (absolute temperature), and

ΔV = –9.05 10^{-5} m /kg (change in volume from solid to liquid),

we obtain

$$\frac{\Delta P}{\Delta T} = -13.1 \text{ MPa/ C.}$$

To provide a rough example of how much pressure this is, to melt ice at –7 C (the temperature many ice skating rinks are set at) would require balancing a small car (mass = 1000 kg) on a thimble (area = 1 cm).

Role of Convective Potential Energy

In meteorology, convective available potential energy (CAPE), sometimes, simply, available potential energy (APE), is the amount of energy a parcel of air would have if lifted a certain distance vertically through the atmosphere. CAPE values are valuable in predicting severe weather.

Mechanics

CAPE exists within the conditionally unstable layer of the troposphere, the free convective layer (FCL), where an ascending air parcel is warmer than the ambient air. CAPE is measured in joules per kilogram of air (J/kg). Any value greater than 0 J/kg indicates instability and the possibility of thunderstorms. Generic CAPE is calculated by integrating vertically the local buoyancy of a parcel from the level of free convection (LFC) to the equilibrium level (EL):

$$CAPE = \int_{Z_f}^{Z_n} g\left(\frac{T_{parcel} - T_{env}}{T_{env}}\right) dz$$

Where z_f and z_n are, respectively, the heights of the levels of free convection and equilibrium (neutral buoyancy), T_{parcel} and T_{env} are the air temperatures, respectively, of the specific parcel and the environment, and g is the acceleration due to gravity. CAPE for a given region is most often calculated from a thermodynamic or sounding diagram (e.g., a Skew-T log-P diagram) using air

temperature and dew point data usually measured by a weather balloon.

CAPE is effectively positive buoyancy, expressed B+ or simply B; the opposite of convective inhibition (CIN), which is expressed as B-, and can be thought of as "negative CAPE". As with CIN, CAPE is usually expressed in J/kg but may also be expressed as m^2/s^2, as the values are equivalent.

In fact, CAPE is sometimes referred to as positive buoyant energy (PBE). This type of CAPE is the maximum energy available to an ascending parcel and to moist convection. When a layer of CIN is present, the layer must be eroded by surface heating or mechanical lifting, so that convective boundary layer parcels may reach their LFC.

CAPE is the positive area above the LFC, the area between the parcel temperature line and the environmental temperature line. The ascending parcel is warmer the environment. CAPE may also exist below the LFC, but if a layer of CIN (subsidence) is present, it is unavailable to deep, moist convection until CIN is exhausted. When there is mechanical lift to saturation, cloud begins at the lifted condensation level (LCL); absent forcing, cloud base begins at the convective condensation level (CCL) where heating from below causes spontaneous buoyant lifting to the point of condensation when the convective temperature is reached. When CIN is absent or is overcome, saturated parcels at the LCL or CCL, which had been small cumulus clouds, will rise to the LFC, and then spontaneously rise until hitting the stable layer of the equilibrium level. The result is deep, moist convection (DMC), or simply, a thunderstorm.

When a parcel is unstable, it will continue to move vertically, in either direction, dependent on whether it receives upward or downward forcing, until it reaches a stable layer (though momentum, gravity, and other forcing may cause the parcel to continue). There are multiple types of CAPE, downdraft CAPE (DCAPE), estimates the potential strength of rain and evaporatively cooled downdrafts. Other types of CAPE may depend on the depth being considered. Other examples are surface based CAPE (SBCAPE), mixed layer or mean layer CAPE (MLCAPE), most

unstable or maximum usable CAPE (MUCAPE), and normalised CAPE (NCAPE).

Significance to Thunderstorms

Thunderstorms form when air parcels are lifted vertically. Deep, moist convection requires a parcel to be lifted to the LFC where it then rises spontaneously until reaching a layer of non-positive buoyancy. The atmosphere is warm at the surface and lower levels of the troposphere where there is mixing (the planetary boundary layer (PBL)), but becomes substantially cooler with height. The temperature profile of the atmosphere, the change in temperature, the degree that it cools with height, is the lapse rate. When the rising air parcel cools more slowly than the surrounding atmosphere, it remains warmer and relatively less dense. The parcel continues to rise freely (convectively; without mechanical lift) through the atmosphere until it reaches an area of air less dense (warmer) than itself.

The amount of and shape of the positive area modulates the speed of updrafts, extreme CAPE can result in explosive thunderstorm development; such rapid development usually occurs when CAPE stored by a capping inversion is released when the "lid" is broken by heating or mechanical lift. The amount of CAPE also modulates how low-level vorticity is entrained and then stretched in the updraft, with importance to tornadogenesis.

The most important CAPE for tornadoes is within the lowest 1 to 3 km (0.6 to 1.9 mi), whilst deep layer CAPE and the width of CAPE at mid-levels is important for supercells). Tornado outbreaks tend to occur within high CAPE environments. Large CAPE is required for the production of very large hail, owing to updraft strength, although a rotating updraft may be relatively stronger with less CAPE. Large CAPE also promotes lightning activity.

Two notable days for severe weather exhibited CAPE values over 5,000 J/kg. Two hours before the Oklahoma Tornado Outbreak occurred on May 3, 1999, the CAPE value sounding at Oklahoma City was at 5,885 J/kg. A few hours later, an F5 tornado ripped through the southern suburbs of the city. Also on May 4, 2007 CAPE values of 5,500 J/kg were reached and an EF5 tornado tore through Greensburg, Kansas.

On these days, it was apparent that conditions were ripe for tornadoes and CAPE wasn't a crucial factor. However, extreme CAPE, by modulating the updraft (and downdraft), can allow for exceptional events, such as the deadly F5 tornadoes that hit Plainfield, Illinois on August 28, 1990 and Jarrell, Texas on May 27, 1997 on days which weren't readily apparent as conducive to large tornadoes. CAPE was estimated to exceed 8,000 J/kg in the environment of the Plainfield storm and was around 7,000 J/kg for the Jarrell storm.

Severe weather and tornadoes can develop in an area of relatively low CAPE values. The surprise severe weather event that occurred in Illinois and Indiana on April 20, 2004 is a good example. Importantly in that case, was that although overall CAPE was weak, there was strong CAPE in the lowest levels of the troposphere which enabled an outbreak of minisupercells producing large, long-track, intense tornadoes.

Equation of Hypsometric

The hypsometric equation relates the atmospheric pressure ratio to the thickness of an atmospheric layer under the assumptions of constant temperature and gravity. It is derived from the hydrostatic equation and the ideal gas law.

It is expressed as:

$$h = z_2 - z_1 = \frac{R \cdot T}{g} . \mathrm{In}\left[\frac{P_1}{P_2}\right]$$

where:

h = thickness of the layer [m]

z = geopotential height [m]

R = gas constant for dry air

T = temperature in kelvins [K]

g = gravitational acceleration [m/s^2]

P = pressure [Pa]

In meteorology P_1 and P_2 are isobaric surfaces and T is the average temperature of the layer between them. In altimetry with

the International Standard Atmosphere the hypsometric equation is used to compute pressure at a given geopotential height in isothermal layers in the upper and lower stratosphere.

Derivation

The hydrostatic equation:

$$P = \rho \cdot g \cdot z$$

where ρ is the density [kg/m^3], is used to generate the equation for hydrostatic equilibrium, written in differential form:

$$dP = -\rho \cdot g \cdot \mathrm{d}z.$$

This is combined with the ideal gas law:

$$P = -\rho \cdot R \cdot T$$

to eliminate ρ:

$$\frac{\mathrm{d}P}{P} = \frac{-\mathrm{g}}{R \cdot T}\mathrm{d}z.$$

This is integrated from z_1 to z_2:

$$\int_{P(z_1)}^{P(z_2)} \frac{\mathrm{d}P}{P} = \int_{z_1}^{z_2} \frac{-g}{R \cdot T}\mathrm{d}z.$$

Integration gives:

$$\mathrm{In}\left(\frac{P(z_2)}{P(z_1)}\right) = \frac{-g}{R \cdot T}(z_2 - z_1)$$

simplifying to:

$$\mathrm{In}\left(\frac{P_1}{P_2}\right) = \frac{g}{R \cdot T}(z_2 - z_1).$$

Rearranging:

$$(z_2 - z_1) = \frac{R \cdot T}{g}\mathrm{In}\left(\frac{P_1}{P_2}\right)$$

or, eliminating the logarithm:

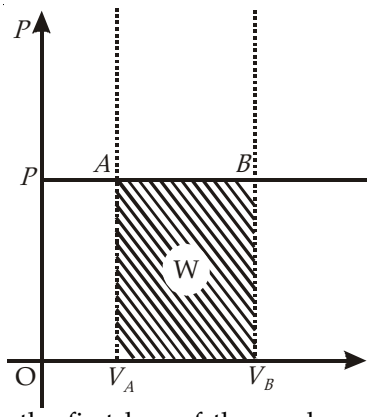
$$\frac{P_1}{P_2} = \mathrm{e}^{\frac{g}{R \cdot T} \cdot (z_2 - z_1)}.$$

Process of Isobar

An isobaric process is a thermodynamic process in which the pressure stays constant: $\Delta p = 0$.

The heat transferred to the system does work but also changes the internal energy of the system:

$$Q = \Delta U + W$$



According to the first law of thermodynamics, where W is work done by the system, U is internal energy, and Q is heat. Pressure-volume work (by the system) is defined as:

$$W = \Delta(pV)$$

but since pressure is constant, this means that

$$W = p\Delta V.$$

Applying the ideal gas law, this becomes

$$W = nR\Delta T$$

assuming that the quantity of gas stays constant (e.g. no phase change during a chemical reaction).

Since it is generally true that:

$$\Delta U = nc_V \Delta T$$

then substituting the last two equations into the first equation produces:

$$Q = nc_V \Delta T + nR\Delta T$$
$$= n(c_V + R)\Delta T.$$

The quantity in parentheses is equivalent to the molar specific heat for constant pressure:

$$c_p = c_V + R$$

and if the gas involved in the isobaric process is monatomic then

$$c_V = \frac{3}{2}R \text{ and } c_p = \frac{5}{2}R.\text{s}$$

An isobaric process is shown on a P-V diagram as a straight horizontal line, connecting the initial and final thermostatic states. If the process moves towards the right, then it is an expansion. If the process moves towards the left, then it is a compression.

Denominating Enthalpy

An isochoric process is described by the equation $Q = \Delta U$. It would be convenient to have a similar equation for isobaric processes. Substituting the second equation into the first yields:

$$Q = \Delta U + \Delta(pV) = \Delta(U + pV)$$

The quantity $U + p\,V$ is a state function so that it can be given a name.

It is called enthalpy, and is denoted as *H*. Therefore an isobaric process can be more succinctly described as:

$$Q = \Delta H.$$

Elastic Smash

A collision between two pool or snooker balls is a good example of an almost totally elastic collision. In addition to momentum being conserved when the two balls collide, the sum of kinetic energy before a collision must equal the sum of kinetic energy after:

$$\frac{1}{2}m_1 v_{1,i}^2 + \frac{1}{2}m_2 v_{2,i}^2 = \frac{1}{2}m_1 v_{1,f}^2 + \frac{1}{2}m_2 v_{2,f}^2$$

Since the 1/2 factor is common to all the terms, it can be taken out right away.

Head-on Collision

In the case of two objects colliding head on we find that the final velocity:

$$V_{1,f} = \left(\frac{m_1 - m_2}{m_1 + m_2} \right) V_{1,i} + \left(\frac{2m_2}{m_1 + m_2} \right) V_{2,i}$$

$$V_{2,f} = \left(\frac{2m_2}{m_1 + m_2} \right) V_{1,i} + \left(\frac{m_2 - m_1}{m_1 + m_2} \right) V_{2,i}$$

which can then easily be rearranged to:

$$m_{1,f} \cdot V_{1,f} + m_{2,f} \cdot V_{2,f} = m_{1,i} \cdot V_{1,i} + m_{2,i} \cdot V_{2,i}$$

Special Case: $m_1 >> m_2$

Now consider if the mass of one body, say m_1, is far greater than that of m_2 ($m_1 >> m_2$). In that case $m_1 + m_2$ is approximately equal to m_1. And $m_1 - m_2$ is approximately equal to m_1.

Put these values in the above equation to calculate the value of V_2 after collision. The expression changes to V_2 final is $2V_1 - V_2$. Its physical interpretation is in case of collision between two body one of which is very heavy, the lighter body moves with twice the velocity of the heavier body less its actual velocity but in opposite direction.

Special Case: $m_1 == m_2$

Another special case is when the collision is between two bodies of equal mass. Say body m_1 moving at velocity V_1 strikes body M_2 that is at rest (V_2). Putting this case in the equation derived above we will see that after the collision, the body that was moving (M_1) will start moving with velocity V_2 and the mass M_2 will start moving with velocity V_1. So there will be an exchange of velocities.

Now suppose one of the masses, say M_2, was at rest. In that case after the collision the moving body, M_1, will come to rest and the body that was at rest, M_2, will start moving with the velocity that M_1 had before the collision.

All of these observations are for an elastic collision. This phenomenon called Newton's cradle, one of the most well-known examples of conservation of momentum, is a real life example of this special case.

Multi-dimensional Collisions

In the case of objects colliding in more than one dimension, as in oblique collisions, the velocity is resolved into orthogonal components with one component perpendicular to the plane of collision and the other component or components in the plane of collision. The velocity components in the plane of collision remain unchanged, while the velocity perpendicular to the plane of collision is calculated in the same way as the one-dimensional case.

For example, in a two-dimensional collision, the momenta can be resolved into *x* and *y* components. We can then calculate each component separately, and combine them to produce a vector result. The magnitude of this vector is the final momentum of the isolated system.

Inelastic Collisions

A common example of a perfectly inelastic collision is when two snowballs collide and then *stick* together afterwards. This equation describes the conservation of momentum:

$$m_1 \mathrm{v}_{1,i} + m_2 \mathrm{v}_{2,i} = (m_1 + m_2)\,\mathrm{v}_f$$

It can be shown that a perfectly inelastic collision is one in which the maximum amount of kinetic energy is converted into other forms. For instance, if both objects stick together after the collision and move with a final common velocity, one can always find a reference frame in which the objects are brought to rest by the collision and 100 per cent of the kinetic energy is converted.

This is true even in the relativistic case and utilised in particle accelerators to efficiently convert kinetic energy into new forms of mass-energy (i.e. to create massive particles).

Contemporary Definitions of Momentum

Momentum in Relativistic Mechanics: In relativistic mechanics, in order to be conserved, momentum must be defined as:

$$\mathrm{p} = \gamma m_0 \mathrm{v}$$

where,

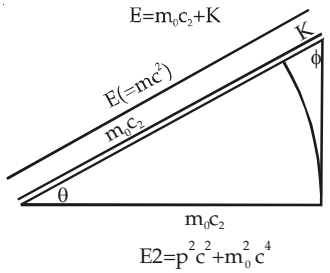
m_0 is the invariant mass of the object moving,

$\gamma = \dfrac{1}{\sqrt{1 - \dfrac{v^2}{c^2}}}$ is the Lorentz factor

v is the relative velocity between an object and an observer

c is the speed of light.

Relativistic momentum becomes Newtonian momentum: mv at low speed $\left(\mathrm{v/c} \to 0\right)$.



Relativistic four-momentum as proposed by Albert Einstein arises from the invariance of four-vectors under Lorentzian translation.

The four-momentum is defined as:

$$\left(\frac{E}{c}, p_x, p_y, p_z\right)$$

where

p_x is the x component of the *relativistic* momentum,

E is the total energy of the system:

$$E = \gamma mc^2$$

The "length" of the vector is the mass times the speed of light, which is invariant across all reference frames:

$$\left(E/c\right)^2 - p^2 = \left(mc\right)^2$$

Significance of Momentum

Momentum of Massless Objects: Massless objects such as photons also carry momentum. The formula is:

$$p = \frac{h}{\lambda} = \frac{E}{c}$$

where

h is Planck's constant,

λ is the wavelength of the photon,

E is the energy the photon carries and

c is the speed of light.

Generalisation of Momentum

Momentum is the Noether charge of translational invariance. As such, even fields as well as other things can have momentum, not just particles.

However, in curved space-time which is not asymptotically Minkowski, momentum isn't defined at all.

Momentum in Quantum Mechanics

In quantum mechanics, momentum is defined as an operator on the wave function. The Heisenberg uncertainty principle defines limits on how accurately the momentum and position of a single observable system can be known at once. In quantum mechanics, position and momentum are conjugate variables.

For a single particle with no electric charge and no spin, the momentum operator can be written in the position basis as:

$$p = \frac{\mathrm{h}}{i}\nabla = -i\mathrm{h}\nabla$$

where:

∇ is the gradient operator

h is the reduced Planck constant.

$i = \sqrt{-1}$ is the imaginary unit.

This is a commonly encountered form of the momentum operator, though not the most general one.

Momentum in Electromagnetism

When electric and/or magnetic fields move, they carry momenta. Light (visible, UV, radio) is an electromagnetic wave and also has momentum. Even though photons (the particle aspect of light) have no mass, they still carry momentum. This leads to applications such as the solar sail.

Momentum is conserved in an electrodynamic system (it may change from momentum in the fields to mechanical momentum of moving parts).

The treatment of the momentum of a field is usually accomplished by considering the so-called energy-momentum tensor and the change in time of the Poynting vector integrated over some volume. This is a tensor field which has components related to the energy density and the momentum density.

The definition canonical momentum corresponding to the momentum operator of quantum mechanics when it interacts with the electromagnetic field is, using the principle of least coupling:

$P = mv + qA,$

instead of the customary,

$p = mv,$

where:

A is the electromagnetic vector potential

m the charged particle's invariant mass

v its velocity

q its charge.

Descriptive Use

A process may be said to gain momentum. The terminology implies that it requires effort to start such a process, but that it is relatively easy to keep it going. Alternatively, the expression

can be seen to reflect that the process is adding adherents, or general acceptance, and thus has more mass at the same velocity; hence, it gained momentum.

Transfer in Mass

Mass transfer is the phrase commonly used in engineering for physical processes that involve molecular and convective transport of atoms and molecules within physical systems. Mass transfer includes both fluid flow and separation unit operations.

Some common examples of mass transfer processes are the evaporation of water from a pond to the atmosphere; the diffusion of chemical impurities in lakes, rivers, and oceans from natural or artificial point sources; mass transfer is also responsible for the separation of components in an apparatus such as a distillation column. In heating, ventilating, and air-conditioning (HVAC), examples of a heat and mass exchangers are cooling towers and evaporative coolers where evaporation of water cools that portion which remains as a liquid, as well as cooling and humidifying the air passing through.

The driving force for mass transfer is a difference in concentration; the random motion of molecules causes a net transfer of mass from an area of high concentration to an area of low concentration. The amount of mass transfer can be quantified through the calculation and application of mass transfer coefficients. Mass transfer finds extensive application in chemical engineering problems, where material balance on components is performed.

In astronomy, mass transfer is the process by which matter gravitationally bound to a body, usually a star, fills its Roche lobe and becomes gravitationally bound to a second body, usually a compact object (white dwarf, neutron star or black hole), and is eventually accreted onto it. It is a common phenomenon in binary systems, and may play an important role in some types of supernovae, and pulsars.

Analogies between Heat, Mass and Momentum Transfer

It is important to note that in molecular transport, heat, or mass there are many similarities. The molecular diffusion equations

of Newton for momentum, Fourier for heat, and Fick for mass are very similar. Therefore many analogies among these three molecular transport process.

A great deal of effort has been devoted in the literature to developing analogies among these three transport processes for turbulent transfer so as to allow prediction of one from any of the others. Reynolds analogy assumes that the turbulent diffusivities are all equal and that the molecular diffusivities mu/ro and Dab are negligible compared to the turbulent diffusivities.

When liquids are present and/or drag is present the analogy is not valid. Other analogies, such as Von Karman's and Prandtl's, usually results in poor relations. The most successful and most widely used analogy is the Chilton and Colburn J-factor analogy.

This analogy is based on experimental data for gases and liquids in both the laminar and turbulent regions. Although it is based on experimental data, it can be shown to satisfy the exact solution derived from laminar flow over a flat plate.

This page intentionally left blank

Molecular Biophysics

Molecular biophysics is a rapidly evolving interdisciplinary area of research that combines concepts in physics, chemistry, engineering, mathematics and biology.

It seeks to understand biomolecular systems and explain biological function in terms of molecular structure, structural organisation, and dynamic behaviour at various levels of complexity (from single molecules to supramolecular structures, viruses and small living systems). The technical challenges are formidable, and the discipline has required development of specialised equipment and procedures capable of imaging and manipulating minute living structures, as well as novel experimental approaches.

Role of Nucleotide

A nucleotide is a chemical compound that consists of 3 portions: a heterocyclic base, a sugar, and one or more phosphate groups. In the most common nucleotides the base is a derivative of purine or pyrimidine, and the sugar is the pentose (five-carbon sugar) deoxyribose or ribose. Nucleotides are the monomers of nucleic acids, with three or more bonding together in order to form a nucleic acid.

Nucleotides are the structural units of RNA, DNA, and several cofactors-CoA, flavin adenine dinucleotide, flavin mononucleotide, adenosine triphosphate and nicotinamide adenine dinucleotide phosphate. In the cell they have important roles in metabolism, and signalling.

Synthesis

Salvage synthesis refers to the reuse of parts of nucleotides in resynthesising new nucleotides. Salvage synthesis requires both breakdown and synthesis reactions in order to exchange the useful parts.

Natural

Purine Ribonucleotides: By using a variety of isotopically labelled compounds it was demonstrated that the sources of the atoms in of purines are as follows:

> The de novo synthesis of purine nucleotides by which these precursors are incorporated into the purine ring, proceeds by a 10 step pathway to the branch point intermediate IMP, the nucleotide of the base hypoxanthine. AMP and GMP are subsequently synthesised from this intermediate via separate, two step each, pathways. Thus purine moieties are initially formed as part of the ribonucleotides rather than as free bases.

Six enzymes take part in IMP synthesis. Three of them are multifunctional:

- GART (reactions 2, 3, and 5)
- PAICS (reactions 6, and 7)
- ATIC (reactions 9, and 10)

Reaction 1: The pathway starts with the formation of PRPP. PRPS1 is the enzyme that activates R5P, which is primarily formed by the pentose phosphate pathway, to PRPP by reacting it with ATP. The reaction is unusual in that a pyrophosphoryl group is directly transferred from ATP to C1 of R5P and that the product has the α configuration about C1. This reaction is also shared with the pathways for the synthesis of the pyrimidine nucleotides, Trp, and His. As a result of being on (a) such (a) major metabolic crossroad and the use of energy, this reaction is highly regulated.

Reaction 2: In the first reaction unique to purine nucleotide biosynthesis, PPAT catalyses the displacement of PRPP's pyrophosphate group (PP_i) by Gln's amide nitrogen. The reaction occurs with the inversion of configuration about ribose C1, thereby forming β -5-phosphorybosylamine (5-PRA) and establishing the anomeric form of the future nucleotide. This reaction which is driven to completion by the subsequent hydrolysis of the released PP_i, is the pathway's flux generating step and is therefore regulated too.

Role of Phosphate

A phosphate, in inorganic chemistry, is a salt of phosphoric acid. In organic chemistry, a phosphate, or organophosphate, is an ester of phosphoric acid. Phosphates are important in biochemistry and biogeochemistry.

Chemical Properties

The phosphate ion is a polyatomic ion with the empirical formula PO_4^{3-} and a molar mass of 94.973 g/mol; it consists of one central phosphorus atom surrounded by four identical oxygen atoms in a tetrahedral arrangement.

The phosphate ion carries a negative three formal charge and is the conjugate base of the hydrogenphosphate ion, HPO_4^2, which is the conjugate base of $H_2PO_4^-$, the dihydrogen phosphate ion, which in turn is the conjugate base of H_3PO_4, phosphoric acid. It is a hypervalent molecule (the phosphorus atom has 10 electrons in its valence shell). Phosphate is also an organophosphorus compound with the formula $OP(OR)_3$.

A phosphate salt forms when a positively charged ion attaches to the negatively charged oxygen atoms of the ion, forming an ionic compound. Many phosphates are insoluble in water at standard temperature and pressure, except for the alkali metal salts.

In dilute aqueous solution, phosphate exists in four forms. In strongly basic conditions, the phosphate ion (PO_4^{3-}) predominates, while in weakly basic conditions, the hydrogen phosphate ion (HPO_4^{2-}) is prevalent.

In weakly acid conditions, the dihydrogen phosphate ion ($H_2PO_4^-$) is most common. In strongly acid conditions, aqueous phosphoric acid (H_3PO_4) is the main form.

Phosphates are the naturally occurring form of the element phosphorus, found in many phosphate minerals. Elemental phosphorus and phosphides are not found (rare phosphide minerals may be found in meteorites). In mineralogy and geology, phosphate refers to a rock or ore containing phosphate ions.

The largest rock phosphate deposits in North America lie in the Bone Valley region of central Florida, United States, the Soda Springs region of Idaho, and the coast of North Carolina. Smaller deposits are located in Montana, Tennessee, Georgia and South Carolina near Charleston along Ashley Phosphate road. The small island nation of Nauru and its neighbour Banaba Island, which used to have massive phosphate deposits of the best quality, have been mined excessively. Rock phosphate can also be found on Navassa Island. Morocco, Tunisia, Israel, Togo, and Jordan have large phosphate mining industries as well.

In biological systems, phosphorus is found as a free phosphate ion in solution and is called inorganic phosphate, to distinguish it from phosphates bound in various phosphate esters. Inorganic phosphate is generally denoted P_i and can be created by the hydrolysis of pyrophosphate, which is denoted PP_i:

$$P_2O_7^{4-} + H_2O \rightarrow 2HPO_4^{2-}$$

However, phosphates are most commonly found in the form of adenosine phosphates, (AMP, ADP and ATP) and in DNA and RNA and can be released by the hydrolysis of ATP or ADP. Similar reactions exist for the other nucleoside diphosphates and triphosphates. Phosphoanhydride bonds in ADP and ATP, or other nucleoside diphosphates and triphosphates, contain high amounts of energy which give them their vital role in all living organisms. They are generally referred to as high energy phosphate, as are the phosphagens in muscle tissue. Compounds such as substituted phosphines, have uses in organic chemistry but do not seem to have any natural counterparts.

In ecological terms, because of its important role in biological systems, phosphate is a highly sought after resource. Consequently,

it is often a limiting reagent in environments, and its availability may govern the rate of growth of organisms. Addition of high levels of phosphate to environments and to micro-environments in which it is typically rare can have significant ecological consequences.

For example, booms in the populations of some organisms at the expense of others, and the collapse of populations deprived of resources such as oxygen. In the context of pollution, phosphates are a principal component of total dissolved solids, a major indicator of water quality.

Uses

Phosphates were once commonly used in laundry detergent in the form trisodium phosphate (TSP), but because of algae boom-bust cycles tied to emission of phosphates into watersheds, phosphate detergent sale or usage is restricted in some areas.

In agriculture, phosphate refers to one of the three primary plant nutrients, and it is a component of fertilizers. Rock phosphate is quarried from phosphate beds in sedimentary rocks. In former times, it was simply crushed and used as is, but the crude form is now used only in organic farming. Normally it is chemically treated to make superphosphate, triple superphosphate, or ammonium phosphates, which have higher concentration of phosphate and are also more soluble, therefore, more quickly usable by plants.

Fertilizer grades normally have three numbers; the first is the available nitrogen, the second is the available phosphate (expressed on a P_2O_5 basis), and the third is the available potash (expressed on a K_2O basis). Thus a 10-10-10 fertilizer would contain ten per cent of each, with the remainder being filler.

Surface runoff of phosphates from excessively fertilized farmland can be a cause of phosphate pollution leading to eutrophication (nutrient enrichment), algal bloom and consequent oxygen deficit. This can lead to anoxia for fish and other aquatic organisms in the same manner as phosphate-based detergents.

Phosphate compounds are occasionally added to the public drinking water supply to counter plumbosolvency.

Monosaccharide

Monosaccharides are the simplest form of carbohydrates. They consist of one sugar and are usually colourless, water-soluble, crystalline solids. Some monosaccharides have a sweet taste. Examples of monosaccharides include glucose (dextrose), fructose, galactose, and ribose. Monosaccharides are the building blocks of disaccharides like sucrose (common sugar) and polysaccharides (such as cellulose and starch). Further, each carbon atom that supports a hydroxyl group (except for the first and last) is chiral, giving rise to a number of isomeric forms all with the same chemical formula. For instance, galactose and glucose are both aldohexoses, but they have different chemical and physical properties.

Structure

Fructose, a monosaccharide, as a Haworth projection. With few exceptions (e.g., deoxyribose), monosaccharides have the chemical formula $(CH_2O)_{n+m}$ with the chemical structure $H(CHOH)_nC = O(CHOH)_mH$. If n or m is zero, it is an aldehyde and is termed an aldose, otherwise it is a ketone and is termed a ketose. Monosaccharides contain either a ketone or aldehyde functional group, and hydroxyl groups on most or all of the non-carbonyl carbon atoms.

Cyclic Structure

Most monosaccharides form cyclic structures, which predominate in aqueous solution, by forming hemiacetals or hemiketals (depending on whether they are aldoses or ketoses) between an alcohol and the carbonyl group of the same sugar. Glucose, for example, readily forms a hemiacetal linkage between its carbon-1 and the hydroxyl group of its carbon-5. Since such a reaction introduces an additional stereogenic centre, two anomers are formed (α-isomer and β-isomer) from each distinct straight-chain monosaccharide. The interconversion between these two forms is called mutarotation.

A common way of representing the cyclic structure of monosaccharides is the Haworth projection.

In Haworth projection, the α-isomer has the OH- of the anomeric carbon under the ring structure, and the β-isomer, has the OH- of the anomeric carbon on top of the ring structure. In chair conformation, the α-isomer has the OH- of the anomeric carbon in an axial position, whereas the β-isomer has the OH- of the anomeric carbon in equatorial position.

Isomerism

The total number of possible stereoisomers of one compound (n) is dependent on the number of stereogenic centres (c) in the molecule. The upper limit for the number of possible stereoisomers is $n = 2^c$. The only carbohydrate without an isomer is dihydroxyacetone or DHA.

Monosaccharide Nomenclature

Monosaccharides are classified by the number of carbon atoms they contain:

- Triose, 3 carbon atoms
- Tetrose, 4 carbon atoms
- Pentose, 5 carbon atoms
- Hexose, 6 carbon atoms
- Heptose, 7 carbon atoms
- Octose, 8 carbon atoms
- Nonose, 9 carbon atoms
- Decose, 10 carbon atoms

Monosaccharides are classified the type of keto group they contain:

- Aldose, -CHO (aldehyde)
- Ketose, C=O (ketone)

Monosaccharides are classified according to their molecular configuration at carbon 2:

- D, configuration as in D-glyceraldehyde
- L, configuration as in L-glyceraldehyde

All these classifications can be combined, resulting in names like *D-aldohexose* or *ketotriose*.

List of Monosaccharides

This is a list of some common monosaccharides, not all are found in nature—some have been synthesised:

- *Trioses:*
  - — *Aldotriose:* glyceraldehyde
  - — *Ketotriose:* dihydroxyacetone
- *Tetroses:*
  - — *Aldotetrose:* erythrose and threose
  - — *Ketotetrose:* erythrulose
- *Pentoses:*
  - — *Aldopentoses:* arabinose, lyxose, ribose and xylose
  - — *Ketopentoses:* ribulose and xylulose
- *Hexoses:*
  - — *Aldohexoses:* allose, altrose, galactose, glucose, gulose, idose, mannose and talose
  - — *Ketohexoses:* fructose, psicose, sorbose and tagatose
- *Heptoses:*
  - — *Keto-heptoses:* mannoheptulose, sedoheptulose
- *Octoses:* octolose, 2-keto-3-deoxy-manno-octonate
- *Nonoses:* sialose

Reactions

1. Formation of acetals.
2. Formation of hemiacetals and hemiketals.
3. Formation of ketals.

Role of Peptides

Peptides are the family of short molecules formed from the linking, in a defined order, of various α -amino acids. The link between one amino acid residue and the next is an amide bond and is sometimes referred to as a peptide bond.

Proteins are polypeptide molecules (or consist of multiple polypeptide subunits). The distinction is that peptides are short and polypeptides

proteins are long. There are several different conventions to determine these, all of which have flaws.

One convention is that those peptide chains that are short enough to be made synthetically from the constituent amino acids are called peptides rather than proteins. However, with the advent of better synthetic techniques, peptides as long as hundreds of amino acids can be made, including full proteins like ubiquitin. Native chemical ligation has given access to even longer proteins, so this convention seems to be outdated.

Another convention places an informal dividing line at approximately 50 amino acids in length (some people claim shorter lengths). However, this definition is somewhat arbitrary. Long peptides, such as the amyloid beta peptide linked to Alzheimer's disease, can be considered proteins; and small proteins, such as insulin, can be considered peptides.

Because of the arbitrary nature of this definition, there is considerable movement within the scientific community to ascribe the more-specific definition that "a peptide is an amino acid molecule without secondary structure; on gaining defined structure, it is a protein".

Thus the same molecule can be either a peptide or a protein depending on its environment, though there are peptides that cannot be proteins.

Peptide Classes

Here are the major classes of peptides, according to how they are produced:

Ribosomal Peptides: Are synthesised by translation of mRNA. They are often subjected to proteolysis to generate the mature form. These function, typically in higher organisms, as hormones and signalling molecules.

Some lower organisms produce peptides as biotics, such as microcin J25. Since they are translated, the amino acid residues involved are restricted to those utilised by the ribosome, and post-translational modifications thereof, such as phosphorylation, hydroxylation, sulfonation, disulfide formation, etc. In general, they are linear, although lariat structures are common. More exotic

manipulations do occur, however, such as racemisation (as in platypus venom) or usage of non-ribosmal peptide modules.

Nonribosomal Peptides: Are synthesised using a modular enzyme complex (which functions much like a conveyor belt on a factory). Non-ribosomal peptides are confined primarily to unicellular organisms, plants, and fungi.

All of these complexes are laid out in a similar fashion, and they can contain many different modules to perform a diverse set of chemical manipulations on the developing product. In general, these peptides are cyclic (often with highly-complex cyclic structures), although linear non-ribosomal peptides are common. Since the system is modular and closely related to the machinery for building fatty acids and polyketides, hybrid compounds are often found. Oxazoles, thiazoles, and their reduced counterparts often indicate that the compound was synthesised in this fashion.

Digested Peptides: Are the result of non-specific proteolysis as part of the digestive cycle. It has also been documented that, when certain food proteins such as gluten, casein, egg protein, and spinach protein are broken down, opioid peptides are formed. These peptides mimic the effects of morphine, and those individuals that are unable to break them down will experience mental illness.

These peptides are quite short and are given names such as casomorphine, gluten exorphine, and dermorphine. Ultimately digested peptides are ribosomal peptides, although they aren't made on the ribosome of the organism that contains them.

Peptones: Are derived from animal meat digested by proteolases. The resulting material is used as a source of proteins in nutrient media for growing bacteria and fungi.

Peptide Fragments: Refer to fragments of proteins which used to identify or quantify the source protein. Often these are the products of enzymatic degradation performed in the laboratory on a controlled sample, but can also be forensic or paleontological samples which have been degraded by natural effects.

Peptides in Molecular Biology

Peptides have received prominence in molecular biology in recent times for several reasons. The first and most important is

that peptides allow the creation of antibodies in animals without the need to purify the protein of interest. One can simply make antigenic peptides of sections of the protein of interest. These will suffice in making antibodies in a rabbit or mouse against the protein.

Another reason is that peptides have become instrumental in mass spectrometry, allowing the identification of proteins of interest based on peptide masses and sequence.

Peptides have recently been used in the study of protein [structure] and function. Synthetic peptides can be used as probes to see where protein-peptide interactions occur.

Inhibitory peptides are also used in clinical research to examine the effects of peptides on the inhibition of cancer proteins and other diseases.

Well-known Peptide Families

The peptide families in this section are all ribosomal peptides, usually with hormonal activity. All of these peptides are synthesised by cells as longer "propeptides" or "proproteins" and truncated prior to exiting the cell. They are released into the bloodstream where they perform their signalling functions.

The Tachykinin Peptides

- Substance P
- Kassinin
- Neurokinin A
- Eledoisin
- Neurokinin B

Vasoactive Intestinal Peptides

- VIP *Vasoactive intestinal peptide*
- PACAP Pituitary adenylate cyclase activating peptide
- PHI 27
- PHM 27
- GHRH 1-24 Growth hormone releasing hormone 1-24
- Glucagon
- Secretin

Pancreatic Polypeptide-related Peptides

- NPY
- PYY Peptide YY
- APP Avian pancreatic polypeptide
- HPP Human pancreatic polypeptide

Opioid Peptides

- Proopiomelanocortin (POMC) Peptides
- The Enkephalin pentapeptides
- The Prodynorphin peptides

Calcitonin Peptides

- Calcitonin
- Amylin
- AGG01

Uses of Proteins

Proteins are large organic compounds made of amino acids arranged in a linear chain and joined together by peptide bonds between the carboxyl and amino groups of adjacent amino acid residues. The sequence of amino acids in a protein is defined by a gene and encoded in the genetic code.

Although this genetic code specifies 20 "standard" amino acids, the residues in a protein are often chemically altered in post-translational modification: either before the protein can function in the cell, or as part of control mechanisms. Proteins can also work together to achieve a particular function, and they often associate to form stable complexes.

Like other biological macromolecules such as polysaccharides and nucleic acids, proteins are essential parts of organisms and participate in every process within cells. Many proteins are enzymes that catalyse biochemical reactions, and are vital to metabolism.

Proteins also have structural or mechanical functions, such as actin and myosin in muscle, and the proteins in the cytoskeleton,

which forms a system of scaffolding that maintains cell shape. Other proteins are important in cell signalling, immune responses, cell adhesion, and the cell cycle. Protein is also a necessary part of our diet, since animals cannot synthesise all the amino acids and must obtain essential amino acids from food. Through the process of digestion, animals break down ingested protein into free amino acids that can be used for protein synthesis.

The word *protein* comes from the Greek ("prota"), meaning "*of primary importance*" and these molecules were first described and named by Jöns Jakob Berzelius in 1838. However, proteins' central role in living organisms was not fully appreciated until 1926, when James B. Sumner showed that the enzyme urease was a protein.

The first protein to be sequenced was insulin, by Frederick Sanger, who won the Nobel Prize for this achievement in 1958. The first protein structures to be solved included haemoglobin and myoglobin, by Max Perutz and Sir John Cowdery Kendrew, respectively, in 1958. Both proteins' three-dimensional structures were first determined by x-ray diffraction analysis; the structures of myoglobin and haemoglobin won the 1962 Nobel Prize in Chemistry for their discoverers.

Proteins are linear polymers built from 20 different L-α-amino acids. All amino acids share common structural features including an α carbon to which an amino group, a carboxyl group, and a variable side chain are bonded. Only proline differs from this basic structure, as it contains an unusual ring to the N-end amine group, which forces the CO-NH amide moiety into a fixed conformation.

The side chains of the standard amino acids, detailed in the list of standard amino acids, have different chemical properties that produce proteins' three-dimensional structure and are therefore, critical to protein function. The amino acids in a polypeptide chain are linked by peptide bonds formed in a dehydration reaction. Once linked in the protein chain, an individual amino acid is called a *residue* and the linked series of carbon, nitrogen, and oxygen atoms are known as the *main chain* or *protein backbone*.

The peptide bond has two resonance forms that contribute some double bond character and inhibit rotation around its axis, so that the alpha carbons are roughly coplanar. The other two dihedral angles in the peptide bond determine the local shape assumed by the protein backbone.

Due to the chemical structure of the individual amino acids, the protein chain has directionality. The end of the protein with a free carboxyl group is known as the C-terminus or carboxy terminus, while the end with a free amino group is known as the N-terminus or amino terminus.

There is some ambiguity between the usage of the words *protein*, polypeptide, and peptide. *Protein* is generally used to refer to the complete biological molecule in a stable conformation, while *peptide* is generally reserved for a short amino acid oligomers often lacking a stable 3-dimensional structure. However, the boundary between the two is ill-defined and usually lies near 20-30 residues. *Polypeptide* can refer to any single linear chain of amino acids, usually regardless of length, but often implies an absence of a single defined conformation.

Synthesis

Proteins are assembled from amino acids using information encoded in genes. Each protein has its own unique amino acid sequence that is specified by the nucleotide sequence of the gene encoding this protein. The genetic code is a set of three-nucleotide sets called codons and each three-nucleotide combination stands for an amino acid, for example AUG stands for methionine. Because DNA contains four nucleotides, the total number of possible codons is 64; hence, there is some redundancy in the genetic code and some amino acids are specified by more than one codon.

Genes encoded in DNA are first transcribed into premessenger RNA (mRNA) by proteins such as RNA polymerase. Most organisms then process the pre-mRNA (also known as a *primary transcript*) using various forms of post-transcriptional modification to form the mature mRNA, which is then used as a template for protein synthesis by the ribosome. In prokaryotes the mRNA may either be used as soon as it is produced, or be bound by a ribosome after having moved away from the nucleoid. In contrast, eukaryotes

make mRNA in the cell nucleus and then translocate it across the nuclear membrane into the cytoplasm, where protein synthesis then takes place. The rate of protein synthesis is higher in prokaryotes than eukaryotes and can reach up to 20 amino acids per second.

The process of synthesising a protein from an mRNA template is known as translation. The mRNA is loaded onto the ribosome and is read three nucleotides at a time by matching each codon to its base pairing anti-codon located on a transfer RNA molecule, which carries the amino acid corresponding to the codon it recognises.

The enzyme aminoacyl tRNA synthetase "charges" the tRNA molecules with the correct amino acids. The growing polypeptide is often termed the *nascent chain*. Proteins are always biosynthesised from N-terminus to C-terminus.

The size of a synthesised protein can be measured by the number of amino acids it contains and by its total molecular mass, which is normally reported in units of *daltons* (synonymous with atomic mass units), or the derivative unit kilodalton (kDa). Yeast proteins are on average 466 amino acids long and 53 kDa in mass. The largest known proteins are the titins, a component of the muscle sarcomere, with a molecular mass of almost 3,000 kDa and a total length of almost 27,000 amino acids.

Chemical Synthesis

Short proteins can also be synthesised chemically by a family of methods known as peptide synthesis, which rely on organic synthesis techniques such as chemical ligation to produce peptides in high yield. Chemical synthesis allows for the introduction of non-natural amino acids into polypeptide chains, such as attachment of fluorescent probes to amino acid side chains. These methods are useful in laboratory biochemistry and cell biology, though generally not for commercial applications. Chemical synthesis is inefficient for polypeptides longer than about 300 amino acids, and the synthesised proteins may not readily assume their native tertiary structure. Most chemical synthesis methods proceed from C-terminus to N-terminus, opposite the biological reaction.

Structure of Proteins

Most proteins fold into unique 3-dimensional structures. The shape into which a protein naturally folds is known as its native state. Although many proteins can fold unassisted simply through the structural propensities of their component amino acids, others require the aid of molecular chaperones to efficiently fold to their native states. Biochemists often refer to four distinct aspects of a protein's structure:

- *Primary Structure:* The amino acid sequence.
- *Secondary Structure:* Regularly repeating local structures stabilised by hydrogen bonds. The most common examples are the alpha helix and beta sheet. Because secondary structures are local, many regions of different secondary structure can be present in the same protein molecule.
- *Tertiary Structure:* The overall shape of a single protein molecule; the spatial relationship of the secondary structures to one another. Tertiary structure is generally stabilised by non-local interactions, most commonly the formation of a hydrophobic core, but also through salt bridges, hydrogen bonds, disulphide bonds, and even post-translational modifications. The term "tertiary structure" is often used as synonymous with the term *fold*.
- *Quaternary Structure:* The shape or structure that results from the interaction of more than one protein molecule, usually called protein subunits in this context, which function as part of the larger assembly or protein complex.

Proteins are not entirely rigid molecules. In addition to these levels of structure, proteins may shift between several related structures in performing their biological function. In the context of these functional rearrangements, these tertiary or quaternary structures are usually referred to as "conformations," and transitions between them are called *conformational changes*. Such changes are often induced by the binding of a substrate molecule to an enzyme's active site, or the physical region of the protein that participates in chemical catalysis. In solution, all proteins also undergo variation in structure through thermal vibration and the collision with other molecules.

Proteins can be informally divided into three main classes, which correlate with typical tertiary structures: globular proteins, fibrous proteins, and membrane proteins. Almost all globular proteins are soluble and many are enzymes. Fibrous proteins are often structural; membrane proteins often serve as receptors or provide channels for polar or charged molecules to pass through the cell membrane.

A special case of intramolecular hydrogen bonds within proteins, poorly shielded from water attack and hence promoting their own dehydration, are called dehydrons.

Structure Determination

Discovering the tertiary structure of a protein, or the quaternary structure of its complexes, can provide important clues about how the protein performs its function. Common experimental methods of structure determination include X-ray crystallography and NMR spectroscopy, both of which can produce information at atomic resolution.

Cryoelectron microscopy is used to produce lower-resolution structural information about very large protein complexes, including assembled viruses; a variant known as electron crystallography can also produce high-resolution information in some cases, especially for two-dimensional crystals of membrane proteins. Solved structures are usually deposited in the Protein Data Bank (PDB), a freely available resource from which structural data about thousands of proteins can be obtained in the form of Cartesian coordinates for each atom in the protein.

There are many more known gene sequences than there are solved protein structures. Further, the set of solved structures is biased towards those proteins that can be easily subjected to the experimental conditions required by one of the major structure determination methods. In particular, globular proteins are comparatively easy to crystallise in preparation for X-ray crystallography, which remains the oldest and most common structure determination technique.

Membrane proteins, by contrast, are difficult to crystallise and are underrepresented in the PDB. Structural genomics initiatives have attempted to remedy these deficiencies by systematically

solving representative structures of major fold classes. Protein structure prediction methods attempt to provide a means of generating a plausible structure for proteins whose structures have not been experimentally determined.

Cellular Functions

Proteins are the chief actors within the cell, said to be carrying out the duties specified by the information encoded in genes. With the exception of certain types of RNA, most other biological molecules are relatively inert elements upon which proteins act.

Proteins make up half the dry weight of an Escherichia coli cell, while other macromolecules such as DNA and RNA make up only 3 per cent and 20 per cent respectively. The set of proteins expressed in a particular cell or cell type is known as its proteome.

The chief characteristic of proteins that enables them to carry out their diverse cellular functions is their ability to bind other molecules specifically and tightly. The region of the protein responsible for binding another molecule is known as the binding site and is often a depression or "pocket" on the molecular surface. This binding ability is mediated by the tertiary structure of the protein, which defines the binding site pocket, and by the chemical properties of the surrounding amino acids' side chains.

Protein binding can be extraordinarily tight and specific; for example, the ribonuclease inhibitor protein binds to human angiogenin with a sub-femtomolar dissociation constant ($<10^{-15}$ M) but does not bind at all to its amphibian homolog onconase (>1 M). Extremely minor chemical changes such as the addition of a single methyl group to a binding partner can sometimes suffice to nearly eliminate binding; for example, the aminoacyl tRNA synthetase specific to the amino acid valine discriminates against the very similar side chain of the amino acid isoleucine.

Proteins can bind to other proteins as well as to small-molecule substrates. When proteins bind specifically to other copies of the same molecule, they can oligomerise to form fibrils; this process occurs often in structural proteins that consist of globular monomers that self-associate to form rigid fibres. Protein-protein interactions also regulate enzymatic activity, control progression

through the cell cycle, and allow the assembly of large protein complexes that carry out many closely related reactions with a common biological function. Proteins can also bind to, or even be integrated into, cell membranes. The ability of binding partners to induce conformational changes in proteins allows the construction of enormously complex signalling networks.

Significance of Enzymes

The best-known role of proteins in the cell is their duty as enzymes, which catalyse chemical reactions. Enzymes are usually highly specific catalysts that accelerate only one or a few chemical reactions. Enzymes effect most of the reactions involved in metabolism and catabolism as well as DNA replication, DNA repair, and RNA synthesis.

Some enzymes act on other proteins to add or remove chemical groups in a process known as post-translational modification. About 4,000 reactions are known to be catalysed by enzymes. The rate acceleration conferred by enzymatic catalysis is often enormous—as much as 10^{17}-fold increase in rate over the uncatalysed reaction in the case of orotate decarboxylase (78 million years without the enzyme, 18 milliseconds with the enzyme).

The molecules bound and acted upon by enzymes are known as substrates. Although enzymes can consist of hundreds of amino acids, it is usually only a small fraction of the residues that come in contact with the substrate and an even smaller fraction - 3-4 residues on average - that are directly involved in catalysis. The region of the enzyme that binds the substrate and contains the catalytic residues is known as the active site.

Cell Signalling and Ligand Transport

Many proteins are involved in the process of cell signalling and signal transduction. Some proteins, such as insulin, are extracellular proteins that transmit a signal from the cell in which they were synthesised to other cells in distant tissues. Others are membrane proteins that act as receptors whose main function is to bind a signalling molecule and induce a biochemical response in the cell.

Many receptors have a binding site exposed on the cell surface and an effector domain within the cell, which may have enzymatic activity or may undergo a conformational change detected by other proteins within the cell.

Antibodies are protein components of adaptive immune system whose main function is to bind antigens, or foreign substances in the body, and target them for destruction. Antibodies can be secreted into the extracellular environment or anchored in the membranes of specialised B cells known as plasma cells. While enzymes are limited in their binding affinity for their substrates by the necessity of conducting their reaction, antibodies have no such constraints. An antibody's binding affinity to its target is extraordinarily high.

Many ligand transport proteins bind particular small biomolecules and transport them to other locations in the body of a multicellular organism. These proteins must have a high binding affinity when their ligand is present in high concentrations but must also release the ligand when it is present at low concentrations in the target tissues.

The canonical example of a ligand-binding protein is haemoglobin, which transports oxygen from the lungs to other organs and tissues in all vertebrates and has close homologs in every biological kingdom.

Transmembrane proteins can also serve as ligand transport proteins that alter the permeability of the cell's membrane to small molecules and ions. The membrane alone has a hydrophobic core through which polar or charged molecules cannot diffuse.

Membrane proteins contain internal channels that allow such molecules to enter and exit the cell. Many ion channel proteins are specialised to select for only a particular ion; for example, potassium and sodium channels often discriminate for only one of the two ions.

Structural Proteins

Structural proteins confer stiffness and rigidity to otherwise fluid biological components. Most structural proteins are fibrous proteins; for example, actin and tubulin are globular and soluble

as monomers but polymerise to form long, stiff fibres that comprise the cytoskeleton, which allows the cell to maintain its shape and size. Collagen and elastin are critical components of connective tissue such as cartilage, and keratin is found in hard or filamentous structures such as hair, nails, feathers, hooves, and some animal shells.

Other proteins that serve structural functions are motor proteins such as myosin, kinesin, and dynein, which are capable of generating mechanical forces. These proteins are crucial for cellular motility of single-celled organisms and the sperm of many sexually reproducing multicellular organisms. They also generate the forces exerted by contracting muscles.

Methods of Study

As some of the most commonly studied biological molecules, the activities and structures of proteins are examined both *in vitro* and *in vivo*. *In vitro* studies of purified proteins in controlled environments are useful for learning how a protein carries out its function: for example, enzyme kinetics studies explore the chemical mechanism of an enzyme's catalytic activity and its relative affinity for various possible substrate molecules. By contrast, *in vivo* experiments on proteins' activities within cells or even within whole organisms can provide complementary information about where a protein functions and how it is regulated.

Protein Purification

In order to perform *in vitro* analyses, a protein must be purified away from other cellular components. This process usually begins with cell lysis, in which a cell's membrane is disrupted and its internal contents released into a solution known as a crude lysate. The resulting mixture can be purified using ultracentrifugation, which fractionates the various cellular components into fractions containing soluble proteins; membrane lipids and proteins; cellular organelles, and nucleic acids. Precipitation by a method known as salting out can concentrate the proteins from this lysate.

Various types of chromatography are then used to isolate the protein or proteins of interest based on properties such as molecular weight, net charge and binding affinity. The level of purification

can be monitored using gel electrophoresis if the desired protein's molecular weight is known, by spectroscopy if the protein has distinguishable spectroscopic features, or by enzyme assays if the protein has enzymatic activity.

For natural proteins, a series of purification steps may be necessary to obtain protein sufficiently pure for laboratory applications. To simplify this process, genetic engineering is often used to add chemical features to proteins that make them easier to purify without affecting their structure or activity. Here, a "tag" consisting of a specific amino acid sequence, often a series of histidine residues (a "His-tag"), is attached to one terminus of the protein. As a result, when the lysate is passed over a chromatography column containing nickel, the histidine residues ligate the nickel and attach to the column while the untagged components of the lysate pass unimpeded.

Cellular Localisation

The study of proteins *in vivo* is often concerned with the synthesis and localisation of the protein within the cell. Although many intracellular proteins are synthesised in the cytoplasm and membrane-bound or secreted proteins in the endoplasmic reticulum, the specifics of how proteins are targeted to specific organelles or cellular structures is often unclear. A useful technique for assessing cellular localisation uses genetic engineering to express in a cell a fusion protein or chimera consisting of the natural protein of interest linked to a "reporter" such as green fluorescent protein (GFP).

Through another genetic engineering application known as site-directed mutagenesis, researchers can alter the protein sequence and hence its structure, cellular localisation, and susceptibility to regulation, which can be followed *in vivo* by GFP tagging or *in vitro* by enzyme kinetics and binding studies.

Proteomics and Bioinformatics

The total complement of proteins present at a time in a cell or cell type is known as its proteome, and the study of such large-scale data sets defines the field of proteomics, named by analogy to the related field of genomics.

Key experimental techniques in proteomics include protein microarrays, which allow the detection of the relative levels of a large number of proteins present in a cell, and two-hybrid screening, which allows the systematic exploration of protein-protein interactions. The total complement of biologically possible such interactions is known as the interactome. A systematic attempt to determine the structures of proteins representing every possible fold is known as structural genomics.

The large amount of genomic and proteomic data available for a variety of organisms, including the human genome, allows researchers to efficiently identify homologous proteins in distantly related organisms by sequence alignment.

Sequence profiling tools can perform more specific sequence manipulations such as restriction enzyme maps, open reading frame analyses for nucleotide sequences, and secondary structure prediction. From this data phylogenetic trees can be constructed and evolutionary hypotheses developed using special software like ClustalW regarding the ancestry of modern organisms and the genes they express.

The field of bioinformatics seeks to assemble, annotate, and analyse genomic and proteomic data, applying computational techniques to biological problems such as gene finding and cladistics.

Structure Prediction and Simulation

Complementary to the field of structural genomics, protein structure prediction seeks to develop efficient ways to provide plausible models for proteins whose structures have not yet been determined experimentally. The most successful type of structure prediction, known as homology modelling, relies on the existence of a "template" structure with sequence similarity to the protein being modelled; structural genomics' goal is to provide sufficient representation in solved structures to model most of those that remain.

Although producing accurate models remains a challenge when only distantly related template structures are available, it has been suggested that sequence alignment is the bottleneck in

this process, as quite accurate models can be produced if a "perfect" sequence alignment is known. Many structure prediction methods have served to inform the emerging field of protein engineering, in which novel protein folds have already been designed. A more complex computational problem is the prediction of intermolecular interactions, such as in molecular docking and protein-protein interaction prediction.

The processes of protein folding and binding can be simulated using techniques derived from molecular dynamics, which increasingly take advantage of distributed computing as in the folding home project.

The folding of small alpha-helical protein domains such as the villin headpiece and the HIV accessory protein have been successfully simulated *in silico*, and hybrid methods that combine standard molecular dynamics with quantum mechanics calculations have allowed exploration of the electronic states of rhodopsins.

Nutrition

Most microorganisms and plants can biosynthesise all 20 standard amino acids, while animals must obtain some of the amino acids from the diet. Key enzymes in the biosynthetic pathways that synthesise certain amino acids—such as aspartokinase, which catalyses the first step in the synthesis of lysine, methionine, and threonine from aspartate - are not present in animals.

The amino acids that an organism cannot synthesise on its own are referred to as essential amino acids. If amino acids are present in the environment, microorganisms can conserve energy by taking up the amino acids from their surroundings and downregulating their biosynthetic pathways.

In animals, amino acids are obtained through the consumption of foods containing protein. Ingested proteins are broken down through digestion, which typically involves denaturation of the protein through exposure to acid and hydrolysis by enzymes called proteases.

Some ingested amino acids are used for protein biosynthesis, while others are converted to glucose through gluconeogenesis, or fed into the citric acid cycle. This use of protein as a fuel is particularly important under starvation conditions as it allows the body's own proteins to be used to support life, particularly those found in muscle. Amino acids are also an important dietary source of nitrogen.

History

Proteins were recognised as a distinct class of biological molecules in the eighteenth century by Antoine Fourcroy and others, distinguished by the molecules' ability to coagulate or flocculate under treatments with heat or acid. Noted examples at the time included albumen from egg whites, blood, serum albumin, fibrin, and wheat gluten.

Dutch chemist Gerhardus Johannes Mulder carried out elemental analysis of common proteins and found that nearly all proteins had the same empirical formula.

The term "protein" to describe these molecules was proposed in 1838 by Mulder's associate Jöns Jakob Berzelius. Mulder went on to identify the products of protein degradation such as the amino acid leucine for which he found a (nearly correct) molecular weight of 131 Da.

The difficulty in purifying proteins in large quantities made them very difficult for early protein biochemists to study. Hence, early studies focused on proteins that could be purified in large quantities, e.g., those of blood, egg white, various toxins, and digestive/metabolic enzymes obtained from slaughterhouses. In the late 1950s, the Armour Hot Dog Co. purified 1 kg (= one million milligrams) of pure bovine pancreatic ribonuclease A and made it freely available to scientists around the world.

Linus Pauling is credited with the successful prediction of regular protein secondary structures based on hydrogen bonding, an idea first put forth by William Astbury in 1933. Later work by Walter Kauzman on denaturation, based partly on previous studies by Kaj Linderstrom-Lang, contributed an understanding of protein folding and structure mediated by hydrophobic interactions.

In 1949, Fred Sanger correctly determined the amino acid sequence of insulin, thus conclusively demonstrating that proteins consisted of linear polymers of amino acids rather than branched chains, colloids, or cyclols. The first atomic-resolution structures of proteins were solved by X-ray crystallography in the 1960s and by NMR in the 1980s. As of 2006, the Protein Data Bank has nearly 40,000 atomic-resolution structures of proteins. In more recent times, cryo-electron microscopy of large macromolecular assemblies and computational protein structure prediction of small protein domains are two methods approaching atomic resolution.

DNA

Deoxyribonucleic acid, or DNA is a nucleic acid molecule that contains the genetic instructions used in the development and functioning of all known living organisms. The main role of DNA is the long-term storage of information and it is often compared to a set of blueprints, since DNA contains the instructions needed to construct other components of cells, such as proteins and RNA molecules. The DNA segments that carry this genetic information are called genes, but other DNA sequences have structural purposes, or are involved in regulating the use of this genetic information.

Fig. The Structure of part of a DNA Double Helix

Chemically, DNA is a long polymer of simple units called nucleotides, with a backbone made of sugars and phosphate atoms

joined by ester bonds. Attached to each sugar is one of four types of molecules called bases. It is the sequence of these four bases along the backbone that encodes information.

This information is read using the genetic code, which specifies the sequence of the amino acids within proteins. The code is read by copying stretches of DNA into the related nucleic acid RNA, in a process called transcription. Most of these RNA molecules are used to synthesise proteins, but others are used directly in structures such as ribosomes and spliceosomes.

Within cells, DNA is organised into structures called chromosomes and the set of chromosomes within a cell make up a genome. These chromosomes are duplicated before cells divide, in a process called DNA replication. Eukaryotic organisms such as animals, plants, and fungi store their DNA inside the cell nucleus, while in prokaryotes such as bacteria it is found in the cell's cytoplasm. Within the chromosomes, chromatin proteins such as histones compact and organise DNA, which helps control its interactions with other proteins and thereby control which genes are transcribed.

Physical and Chemical Properties

DNA is a long polymer made from repeating units called nucleotides. The DNA chain is 22 to 24 Angstroms wide (2.2 to 2.4 nanometres), and one nucleotide unit is 3.3 Angstroms (0.33 nanometres) long. Although each individual repeating unit is very small, DNA polymers can be enormous molecules containing millions of nucleotides. For instance, the largest human chromosome, chromosome number 1, is 220 million base pairs long.

In living organisms, DNA does not usually exist as a single molecule, but instead as a tightly-associated pair of molecules. These two long strands entwine like vines, in the shape of a double helix. The nucleotide repeats contain both the segment of the backbone of the molecule, which holds the chain together, and a base, which interacts with the other DNA strand in the helix. In general, a base linked to a sugar is called a nucleoside and a base linked to a sugar and one or more phosphate groups is called

a nucleotide. If multiple nucleotides are linked together, as in DNA, this polymer is referred to as a polynucleotide.

The backbone of the DNA strand is made from alternating phosphate and sugar residues. The sugar in DNA is 2-deoxyribose, which is a pentose (five carbon) sugar. The sugars are joined together by phosphate groups that form phosphodiester bonds between the third and fifth carbon atoms of adjacent sugar rings. These asymmetric bonds mean a strand of DNA has a direction. In a double helix the direction of the nucleotides in one strand is opposite to their direction in the other strand. This arrangement of DNA strands is called anti-parallel. The asymmetric ends of a strand of DNA bases are referred to as the 52 (*five prime*) and 32 (*three prime*) ends. One of the major differences between DNA and RNA is the sugar, with 2-deoxyribose being replaced by the alternative pentose sugar ribose in RNA.

The DNA double helix is stabilised by hydrogen bonds between the bases attached to the two strands. The four bases found in DNA are adenine (abbreviated A), cytosine (C), guanine (G) and thymine (T). These four bases are shown below and are attached to the sugar/phosphate to form the complete nucleotide, as shown for adenosine monophosphate.

These bases are classified into two types; adenine and guanine are fused five- and six-membered heterocyclic compounds called purines, while cytosine and thymine are six-membered rings called pyrimidines. A fifth pyrimidine base, called uracil (U), usually takes the place of thymine in RNA and differs from thymine by lacking a methyl group on its ring. Uracil is not usually found in DNA, occurring only as a breakdown product of cytosine, but a very rare exception to this rule is a bacterial virus called PBS1 that contains uracil in its DNA. In contrast, following synthesis of certain RNA molecules, a significant number of the uracils are converted to thymines by the enzymatic addition of the missing methyl group. This occurs mostly on structural and enzymatic RNAs like transfer RNAs and ribosomal RNA.

The double helix is a right-handed spiral. As the DNA strands wind around each other, they leave gaps between each set of phosphate backbones, revealing the sides of the bases inside.

There are two of these grooves twisting around the surface of the double helix: one groove, the major groove, is 22 Å wide and the other, the minor groove, is 12 Å wide. The narrowness of the minor groove means that the edges of the bases are more accessible in the major groove. As a result, proteins like transcription factors that can bind to specific sequences in double-stranded DNA usually make contacts to the sides of the bases exposed in the major groove.

Base Pairing

Each type of base on one strand forms a bond with just one type of base on the other strand. This is called complementary base pairing. Here, purines form hydrogen bonds to pyrimidines, with A bonding only to T, and C bonding only to G. This arrangement of two nucleotides binding together across the double helix is called a base pair. In a double helix, the two strands are also held together via forces generated by the hydrophobic effect and pi stacking, which are not influenced by the sequence of the DNA.

As hydrogen bonds are not covalent, they can be broken and rejoined relatively easily. The two strands of DNA in a double helix can therefore, be pulled apart like a zipper, either by a mechanical force or high temperature. As a result of this complementarity, all the information in the double-stranded sequence of a DNA helix is duplicated on each strand, which is vital in DNA replication. Indeed, this reversible and specific interaction between complementary base pairs is critical for all the functions of DNA in living organisms.

The two types of base pairs form different numbers of hydrogen bonds, AT forming two hydrogen bonds, and GC forming three hydrogen bonds.

The GC base pair is therefore, stronger than the AT base pair. As a result, it is both the percentage of GC base pairs and the overall length of a DNA double helix that determine the strength of the association between the two strands of DNA.

Long DNA helices with a high GC content have stronger-interacting strands, while short helices with high AT content have weaker-interacting strands. Parts of the DNA double helix that

need to separate easily, such as the TATAAT Pribnow box in bacterial promoters, tend to have sequences with a high AT content, making the strands easier to pull apart.

In the laboratory, the strength of this interaction can be measured by finding the temperature required to break the hydrogen bonds, their melting temperature (also called T_m value). When all the base pairs in a DNA double helix melt, the strands separate and exist in solution as two entirely independent molecules. These single-stranded DNA molecules have no single common shape, but some conformations are more stable than others.

Sense and Anti-sense

A DNA sequence is called "sense" if its sequence is the same as that of a messenger RNA (mRNA) copy that is translated into protein. The sequence on the opposite strand is complementary to the sense sequence and is therefore, called the "antisense" sequence.

Since RNA polymerases work by making a complementary copy of their templates, it is this anti-sense strand that is the template for producing the sense mRNA. Both sense and anti-sense sequences can exist on different parts of the same strand of DNA (i.e. both strands contain both sense and anti-sense sequences). In both prokaryotes and eukaryotes, anti-sense RNA sequences are produced, but the functions of these RNAs are not entirely clear. One proposal is that anti-sense RNAs are involved in regulating gene expression through RNA-RNA base pairing.

A few DNA sequences in prokaryotes and eukaryotes, and more in plasmids and viruses, blur the distinction made above between sense and anti-sense strands by having overlapping genes. In these cases, some DNA sequences do double duty, encoding one protein when read 5′ to 3′along one strand, and a second protein when read in the opposite direction (still 5'to 3') along the other strand. In bacteria, this overlap may be involved in the regulation of gene transcription, while in viruses, overlapping genes increase the amount of information that can be encoded within the small viral genome. Another way of reducing genome size is seen in some viruses that contain linear or circular single-stranded DNA as their genetic material.

Supercoiling

DNA can be twisted like a rope in a process called DNA supercoiling. With DNA in its "relaxed" state, a strand usually circles the axis of the double helix once every 10.4 base pairs, but if the DNA is twisted the strands become more tightly or more loosely wound. If the DNA is twisted in the direction of the helix, this is positive supercoiling, and the bases are held more tightly together. If they are twisted in the opposite direction, this is negative supercoiling, and the bases come apart more easily. In nature, most DNA has slight negative supercoiling that is introduced by enzymes called topoisomerases. These enzymes are also needed to relieve the twisting stresses introduced into DNA strands during processes such as transcription and DNA replication.

Alternative Double-helical Structures

DNA exists in several possible conformations. The conformations so far identified are:

A-DNA, B-DNA, C-DNA, D-DNA, E-DNA, H-DNA, L-DNA, P-DNA, and Z-DNA.

However, only A-DNA, B-DNA, and Z-DNA have been observed in naturally occurring biological systems. Which conformation DNA adopts depends on the sequence of the DNA, the amount and direction of supercoiling, chemical modifications of the bases and also solution conditions, such as the concentration of metal ions and polyamines. Of these three conformations, the "B" form described above is most common under the conditions found in cells. The two alternative double-helical forms of DNA differ in their geometry and dimensions.

The A form is a wider right-handed spiral, with a shallow and wide minor groove and a narrower and deeper major groove. The A form occurs under non-physiological conditions in dehydrated samples of DNA, while in the cell it may be produced in hybrid pairings of DNA and RNA strands, as well as in enzyme-DNA complexes.

Segments of DNA where the bases have been chemically-modified by methylation may undergo a larger change in

conformation and adopt the Z form. Here, the strands turn about the helical axis in a left-handed spiral, the opposite of the more common B form. These unusual structures can be recognised by specific Z-DNA binding proteins and may be involved in the regulation of transcription.

Quadruplex Structures

At the ends of the linear chromosomes are specialised regions of DNA called telomeres. The main function of these regions is to allow the cell to replicate chromosome ends using the enzyme telomerase, as the enzymes that normally replicate DNA cannot copy the extreme 3'ends of chromosomes. As a result, if a chromosome lacked telomeres it would become shorter each time it was replicated.

These specialised chromosome caps also help protect the DNA ends from exonucleases and stop the DNA repair systems in the cell from treating them as damage to be corrected. In human cells, telomeres are usually lengths of single-stranded DNA containing several thousand repeats of a simple TTAGGG sequence.

These guanine-rich sequences may stabilise chromosome ends by forming very unusual structures of stacked sets of four-base units, rather than the usual base pairs found in other DNA molecules. Here, four guanine bases form a flat plate and these flat four-base units then stack on top of each other, to form a stable G-quadruplex structure.

These structures are stabilised by hydrogen bonding between the edges of the bases and chelation of a metal ion in the centre of each four-base unit. The structure shown to the left is a top view of the quadruplex formed by a DNA sequence found in human telomere repeats. The single DNA strand forms a loop, with the sets of four bases stacking in a central quadruplex three plates deep. In the space at the centre of the stacked bases are three chelated potassium ions. Other structures can also be formed, with the central set of four bases coming from either a single strand folded around the bases, or several different parallel strands, each contributing one base to the central structure.

In addition to these stacked structures, telomeres also form large loop structures called telomere loops, or T-loops. Here, the single-stranded DNA curls around in a long circle stabilised by telomere-binding proteins. At the very end of the T-loop, the single-stranded telomere DNA is held onto a region of double-stranded DNA by the telomere strand disrupting the double-helical DNA and base pairing to one of the two strands. This triple-stranded structure is called a displacement loop or D-loop.

Chemical Modifications

Base Modifications: The expression of genes is influenced by the chromatin structure of a chromosome and regions of heterochromatin (low or no gene expression) correlate with the methylation of cytosine. For example, cytosine methylation, to produce 5-methylcytosine, is important for X-chromosome inactivation. The average level of methylation varies between organisms, with Caenorhabditis elegans lacking cytosine methylation, while vertebrates show higher levels, with up to 1 per cent of their DNA containing 5-methylcytosine. Despite the biological role of 5-methylcytosine, it is susceptible to spontaneous deamination to leave the thymine base, and methylated cytosines are therefore, mutation hotspots. Other base modifications include adenine methylation in bacteria and the glycosylation of uracil to produce the "J-base" in kinetoplastids.

DNA Damage

DNA can be damaged by many different sorts of mutagens. These include oxidising agents, alkylating agents and also high-energy electromagnetic radiation such as ultraviolet light and x-rays. The type of DNA damage produced depends on the type of mutagen.

For example, UV light mostly damages DNA by producing thymine dimers, which are cross-links between adjacent pyrimidine bases in a DNA strand. On the other hand, oxidants such as free radicals or hydrogen peroxide produce multiple forms of damage, including base modifications, particularly of guanosine, as well as double-strand breaks. It has been estimated that in each human cell, about 500 bases suffer oxidative damage per day. Of these

oxidative lesions, the most dangerous are double-strand breaks, as these lesions are difficult to repair and can produce point mutations, insertions and deletions from the DNA sequence, as well as chromosomal translocations.

Many mutagens intercalate into the space between two adjacent base pairs. Intercalators are mostly aromatic and planar molecules, and include ethidium, daunomycin, doxorubicin and thalidomide. In order for an intercalator to fit between base pairs, the bases must separate, distorting the DNA strands by unwinding of the double helix. These structural changes inhibit both transcription and DNA replication, causing toxicity and mutations. As a result, DNA intercalators are often carcinogens, with benzopyrene diol epoxide, acridines, aflatoxin and ethidium bromide being well-known examples. Nevertheless, due to their properties of inhibiting DNA transcription and replication, they are also used in chemotherapy to inhibit rapidly-growing cancer cells.

Overview of Biological Functions

DNA usually occurs as linear chromosomes in eukaryotes, and circular chromosomes in prokaryotes. The set of chromosomes in a cell makes up its genome; the human genome has approximately 3 billion base pairs of DNA arranged into 46 chromosomes. The information carried by DNA is held in the sequence of pieces of DNA called genes.

Transmission of genetic information in genes is achieved via complementary base pairing. For example, in transcription, when a cell uses the information in a gene, the DNA sequence is copied into a complementary RNA sequence through the attraction between the DNA and the correct RNA nucleotides. Usually, this RNA copy is then used to make a matching protein sequence in a process called translation which depends on the same interaction between RNA nucleotides. Alternatively, a cell may simply copy its genetic information in a process called DNA replication. The details of these functions are covered in other articles; here we focus on the interactions between DNA and other molecules that mediate the function of the genome.

Genome Structure

Genomic DNA is located in the cell nucleus of eukaryotes, as well as small amounts in mitochondria and chloroplasts. In prokaryotes, the DNA is held within an irregularly shaped body in the cytoplasm called the nucleoid. The genetic information in a genome is held within genes. A gene is a unit of heredity and is a region of DNA that influences a particular characteristic in an organism. Genes contain an open reading frame that can be transcribed, as well as regulatory sequences such as promoters and enhancers, which control the expression of the open reading frame.

In many species, only a small fraction of the total sequence of the genome encodes protein. For example, only about 1.5 per cent of the human genome consists of protein-coding exons, with over 50 per cent of human DNA consisting of non-coding repetitive sequences.

The reasons for the presence of so much non-coding DNA in eukaryotic genomes and the extraordinary differences in genome size, or C-value, among species represent a long-standing puzzle known as the "C-value enigma." However, DNA sequences that do not code protein may still encode functional non-coding RNA molecules, which are involved in the regulation of gene expression.

Some non-coding DNA sequences play structural roles in chromosomes. Telomeres and centromeres typically contain few genes, but are important for the function and stability of chromosomes. An abundant form of non-coding DNA in humans are pseudogenes, which are copies of genes that have been disabled by mutation. These sequences are usually just molecular fossils, although they can occasionally serve as raw genetic material for the creation of new genes through the process of gene duplication and divergence.

Transcription and Translation

A gene is a sequence of DNA that contains genetic information and can influence the phenotype of an organism. Within a gene, the sequence of bases along a DNA strand defines a messenger RNA sequence, which then defines a protein sequence.

The relationship between the nucleotide sequences of genes and the amino-acid sequences of proteins is determined by the rules of translation, known collectively as the genetic code. The genetic code consists of three-letter 'words' called *codons* formed from a sequence of three nucleotides (e.g. ACT, CAG, TTT).

In transcription, the codons of a gene are copied into messenger RNA by RNA polymerase. This RNA copy is then decoded by a ribosome that reads the RNA sequence by base-pairing the messenger RNA to transfer RNA, which carries amino acids. Since there are 4 bases in 3-letter combinations, there are 64 possible codons (4^3 combinations). These encode the twenty standard amino acids, giving most amino acids more than one possible codon. There are also three 'stop' or 'nonsense' codons signifying the end of the coding region; these are the TAA, TGA and TAG codons.

Replication

Cell division is essential for an organism to grow, but when a cell divides it must replicate the DNA in its genome so that the two daughter cells have the same genetic information as their parent. The double-stranded structure of DNA provides a simple mechanism for DNA replication. Here, the two strands are separated and then each strand's complementary DNA sequence is recreated by an enzyme called DNA polymerase. This enzyme makes the complementary strand by finding the correct base through complementary base pairing, and bonding it onto the original strand. As DNA polymerases can only extend a DNA strand in a 5′ to 3′ direction, different mechanisms are used to copy the anti-parallel strands of the double helix. In this way, the base on the old strand dictates which base appears on the new strand, and the cell ends up with a perfect copy of its DNA.

Interactions with Proteins

All the functions of DNA depend on interactions with proteins. These protein interactions can be non-specific, or the protein can bind specifically to a single DNA sequence. Enzymes can also bind to DNA and of these, the polymerases that copy the DNA base sequence in transcription and DNA replication are particularly important.

DNA-binding Proteins

Structural proteins that bind DNA are well-understood examples of non-specific DNA-protein interactions. Within chromosomes, DNA is held in complexes with structural proteins. These proteins organise the DNA into a compact structure called chromatin. In eukaryotes, this structure involves DNA binding to a complex of small basic proteins called histones, while in prokaryotes multiple types of proteins are involved. The histones form a disk-shaped complex called a nucleosome, which contains two complete turns of double-stranded DNA wrapped around its surface.

These non-specific interactions are formed through basic residues in the histones making ionic bonds to the acidic sugar-phosphate backbone of the DNA, and are therefore, largely independent of the base sequence. Chemical modifications of these basic amino acid residues include methylation, phosphorylation and acetylation.

These chemical changes alter the strength of the interaction between the DNA and the histones, making the DNA more or less accessible to transcription factors and changing the rate of transcription. Other non-specific DNA-binding proteins found in chromatin include the high-mobility group proteins, which bind preferentially to bent or distorted DNA. These proteins are important in bending arrays of nucleosomes and arranging them into more complex chromatin structures.

A distinct group of DNA-binding proteins are the single-stranded-DNA-binding proteins that specifically bind single-stranded DNA. In humans, replication protein A is the best-characterised member of this family and is essential for most processes where the double helix is separated, including DNA replication, recombination and DNA repair. These binding proteins seem to stabilise single-stranded DNA and protect it from forming stem loops or being degraded by nucleases.

In contrast, other proteins have evolved to specifically bind particular DNA sequences. The most intensively studied of these are the various classes of transcription factors, which are proteins that regulate transcription. Each one of these proteins bind to one

particular set of DNA sequences and thereby activates or inhibits the transcription of genes with these sequences close to their promoters.

The transcription factors do this in two ways. Firstly, they can bind the RNA polymerase responsible for transcription, either directly or through other mediator proteins; this locates the polymerase at the promoter and allows it to begin transcription. Alternatively, transcription factors can bind enzymes that modify the histones at the promoter; this will change the accessibility of the DNA template to the polymerase.

As these DNA targets can occur throughout an organism's genome, changes in the activity of one type of transcription factor can affect thousands of genes.

Consequently, these proteins are often the targets of the signal transduction processes that mediate responses to environmental changes or cellular differentiation and development. The specificity of these transcription factors' interactions with DNA come from the proteins making multiple contacts to the edges of the DNA bases, allowing them to "read" the DNA sequence. Most of these base-interactions are made in the major groove, where the bases are most accessible.

DNA-modifying Enzymes

Nucleases and Ligases: Nucleases are enzymes that cut DNA strands by catalysing the hydrolysis of the phosphodiester bonds. Nucleases that hydrolyse nucleotides from the ends of DNA strands are called exonucleases, while endonucleases cut within strands. The most frequently-used nucleases in molecular biology are the restriction endonucleases, which cut DNA at specific sequences. For instance, the EcoRV enzyme shown to the left recognises the 6-base sequence 5′ -GAT | ATC-3′ and makes a cut at the vertical line. In nature, these enzymes protect bacteria against phage infection by digesting the phage DNA when it enters the bacterial cell, acting as part of the restriction modification system. In technology, these sequence-specific nucleases are used in molecular cloning and DNA fingerprinting.

Enzymes called DNA ligases can rejoin cut or broken DNA strands, using the energy from either adenosine triphosphate or

nicotinamide adenine dinucleotide. Ligases are particularly important in lagging strand DNA replication, as they join together the short segments of DNA produced at the replication fork into a complete copy of the DNA template. They are also used in DNA repair and genetic recombination.

Topoisomerases and Helicases

Topoisomerases are enzymes with both nuclease and ligase activity. These proteins change the amount of supercoiling in DNA. Some of these enzyme work by cutting the DNA helix and allowing one section to rotate, thereby reducing its level of supercoiling; the enzyme then seals the DNA break. Other types of these enzymes are capable of cutting one DNA helix and then passing a second strand of DNA through this break, before rejoining the helix. Topoisomerases are required for many processes involving DNA, such as DNA replication and transcription.

Helicases are proteins that are a type of molecular motor. They use the chemical energy in nucleoside triphosphates, predominantly ATP, to break hydrogen bonds between bases and unwind the DNA double helix into single strands. These enzymes are essential for most processes where enzymes need to access the DNA bases.

Polymerases

Polymerases are enzymes that synthesise polynucleotide chains from nucleoside triphosphates. They function by adding nucleotides onto the 3′ hydroxyl group of the previous nucleotide in the DNA strand. As a consequence, all polymerases work in a 5′ to 3′direction. In the active site of these enzymes, the nucleoside triphosphate substrate base-pairs to a single-stranded polynucleotide template: this allows polymerases to accurately synthesise the complementary strand of this template. Polymerases are classified according to the type of template that they use.

In DNA replication, a DNA-dependent DNA polymerase makes a DNA copy of a DNA sequence. Accuracy is vital in this process, so many of these polymerases have a proofreading activity. Here, the polymerase recognises the occasional mistakes in the synthesis reaction by the lack of base pairing between the

mismatched nucleotides. If a mismatch is detected, a 3′ to 5′ exonuclease activity is activated and the incorrect base removed. In most organisms DNA polymerases function in a large complex called the replisome that contains multiple accessory subunits, such as the DNA clamp or helicases.

RNA-dependent DNA polymerases are a specialised class of polymerases that copy the sequence of an RNA strand into DNA. They include reverse transcriptase, which is a viral enzyme involved in the infection of cells by retroviruses, and telomerase, which is required for the replication of telomeres. Telomerase is an unusual polymerase because it contains its own RNA template as part of its structure.

Transcription is carried out by a DNA-dependent RNA polymerase that copies the sequence of a DNA strand into RNA. To begin transcribing a gene, the RNA polymerase binds to a sequence of DNA called a promoter and separates the DNA strands. It then copies the gene sequence into a messenger RNA transcript until it reaches a region of DNA called the terminator, where it halts and detaches from the DNA. As with human DNA-dependent DNA polymerases, RNA polymerase II, the enzyme that transcribes most of the genes in the human genome, operates as part of a large protein complex with multiple regulatory and accessory subunits.

Genetic Recombination

Usually interact with other segments of DNA, and in human cells the different chromosomes even occupy separate areas in the nucleus called "chromosome territories". This physical separation of different chromosomes is important for the ability of DNA to function as a stable repository for information, as one of the few times chromosomes interact is during chromosomal crossover when they recombine. Chromosomal crossover is when two DNA helices break, swap a section and then rejoin.

Recombination allows chromosomes to exchange genetic information and produces new combinations of genes, which increases the efficiency of natural selection and can be important in the rapid evolution of new proteins. Genetic recombination can also be involved in DNA repair, particularly in the cell's response to double-strand breaks.

The most common form of chromosomal crossover is homologous recombination, where the two chromosomes involved share very similar sequences. Non-homologous recombination can be damaging to cells, as it can produce chromosomal translocations and genetic abnormalities. The recombination reaction is catalysed by enzymes known as *recombinases*, such as RAD51.

The first step in recombination is a double-stranded break either caused by an endonuclease or damage to the DNA. A series of steps catalysed in part by the recombinase then leads to joining of the two helices by at least one Holliday junction, in which a segment of a single strand in each helix is annealed to the complementary strand in the other helix. The Holliday junction is a tetrahedral junction structure that can be moved along the pair of chromosomes, swapping one strand for another. The recombination reaction is then halted by cleavage of the junction and re-ligation of the released DNA.

Evolution of DNA-based Metabolism

DNA contains the genetic information that allows all modern living things to function, grow and reproduce. However, it is unclear how long in the 4-billion-year history of life DNA has performed this function, as it has been proposed that the earliest forms of life may have used RNA as their genetic material. RNA may have acted as the central part of early cell metabolism as it can both transmit genetic information and carry out catalysis as part of ribozymes. This ancient RNA world where nucleic acid would have been used for both catalysis and genetics may have influenced the evolution of the current genetic code based on four nucleotide bases. This would occur since the number of unique bases in such an organism is a trade-off between a small number of bases increasing replication accuracy and a large number of bases increasing the catalytic efficiency of ribozymes.

Unfortunately, there is no direct evidence of ancient genetic systems, as recovery of DNA from most fossils is impossible. This is because DNA will survive in the environment for less than one million years and slowly degrades into short fragments in solution. Although claims for older DNA have been made, most notably

a report of the isolation of a viable bacterium from a salt crystal 250-million years old, these claims are controversial and have been disputed.

Uses in Technology

Genetic Engineering: Modern biology and biochemistry make intensive use of recombinant DNA technology. Recombinant DNA is a man-made DNA sequence that has been assembled from other DNA sequences. They can be transformed into organisms in the form of plasmids or in the appropriate format, by using a viral vector. The genetically modified organisms produced can be used to produce products such as recombinant proteins, used in medical research, or be grown in agriculture.

Forensic Department

Forensic scientists can use DNA in blood, semen, skin, saliva or hair at a crime scene to identify a perpetrator. This process is called genetic fingerprinting, or more accurately, DNA profiling. In DNA profiling, the lengths of variable sections of repetitive DNA, such as short tandem repeats and minisatellites, are compared between people.

This method is usually an extremely reliable technique for identifying a criminal. However, identification can be complicated if the scene is contaminated with DNA from several people. DNA profiling was developed in 1984 by British geneticist Sir Alec Jeffreys, and first used in forensic science to convict Colin Pitchfork in the 1988 Enderby murders case. People convicted of certain types of crimes may be required to provide a sample of DNA for a database. This has helped investigators solve old cases where only a DNA sample was obtained from the scene. DNA profiling can also be used to identify victims of mass casualty incidents.

Role of Bioinformatics

Bioinformatics involves the manipulation, searching, and data mining of DNA sequence data. The development of techniques to store and search DNA sequences have led to widely-applied advances in computer science, especially string searching

algorithms, machine learning and database theory. String searching or matching algorithms, which find an occurrence of a sequence of letters inside a larger sequence of letters, were developed to search for specific sequences of nucleotides.

In other applications such as text editors, even simple algorithms for this problem usually suffice, but DNA sequences cause these algorithms to exhibit near-worst-case behaviour due to their small number of distinct characters. The related problem of sequence alignment aims to identify homologous sequences and locate the specific mutations that make them distinct. These techniques, especially multiple sequence alignment, are used in studying phylogenetic relationships and protein function. Data sets representing entire genomes' worth of DNA sequences, such as those produced by the Human Genome Project, are difficult to use without annotations, which label the locations of genes and regulatory elements on each chromosome. Regions of DNA sequence that have the characteristic patterns associated with protein- or RNA-coding genes can be identified by gene finding algorithms, which allow researchers to predict the presence of particular gene products in an organism even before they have been isolated experimentally.

DNA and Computation

DNA was first used in computing to solve a small version of the directed Hamiltonian path problem, an NP-complete problem. DNA computing is advantageous over electronic computers in power use, space use, and efficiency, due to its ability to compute in a highly parallel fashion. A number of other problems, including simulation of various abstract machines, the boolean satisfiability problem, and the bounded version of the travelling salesman problem, have since been analysed using DNA computing. Due to its compactness, DNA also has a theoretical role in cryptography, where in particular it allows unbreakable one-time pads to be efficiently constructed and used.

History and Anthropology

Because DNA collects mutations over time, which are then inherited, it contains historical information and by comparing

DNA sequences, geneticists can infer the evolutionary history of organisms, their phylogeny. This field of phylogenetics is a powerful tool in evolutionary biology.

If DNA sequences within a species are compared, population geneticists can learn the history of particular populations. This can be used in studies ranging from ecological genetics to anthropology; for example, DNA evidence is being used to try to identify the Ten Lost Tribes of Israel.

DNA has also been used to look at modern family relationships, such as establishing family relationships between the descendants of Sally Hemings and Thomas Jefferson. This usage is closely related to the use of DNA in criminal investigations detailed above. Indeed, some criminal investigations have been solved when DNA from crime scenes has matched relatives of the guilty individual.

History

DNA was first isolated by the Swiss physician Friedrich Miescher who, in 1869, discovered a microscopic substance in the pus of discarded surgical bandages. As it resided in the nuclei of cells, he called it "nuclein". In 1929 this discovery was followed by Phoebus Levene's identification of the base, sugar and phosphate nucleotide unit.

Levene suggested that DNA consisted of a string of nucleotide units linked together through the phosphate groups. However, Levene thought the chain was short and the bases repeated in a fixed order. In 1937 William Astbury produced the first X-ray diffraction patterns that showed that DNA had a regular structure.

In 1943, Oswald Theodore Avery discovered that traits of the "smooth" form of the *Pneumococcus* could be transferred to the "rough" form of the same bacteria by mixing killed "smooth" bacteria with the live "rough" form. Avery identified DNA as this transforming principle. DNA's role in heredity was confirmed in 1953, when Alfred Hershey and Martha Chase in the Hershey-Chase experiment showed that DNA is the genetic material of the T2 phage.

In 1953, based on X-ray diffraction images taken by Rosalind Franklin and the information that the bases were paired, James

D. Watson and Francis Crick suggested what is now accepted as the first accurate model of DNA structure in the journal *Nature*. Experimental evidence for Watson and Crick's model were published in a series of five articles in the same issue of *Nature*. Of these, Franklin and Raymond Gosling's paper saw the publication of the X-ray diffraction image, which was key in Watson and Crick interpretation, as well as another article, co-authored by Maurice Wilkins and his colleagues. Franklin and Gosling's subsequent paper identified the distinctions between the A and B structures of the double helix in DNA. In 1962 Watson, Crick, and Maurice Wilkins jointly received the Nobel Prize in Physiology or Medicine (Franklin didn't share the prize with them since she had died earlier).

In an influential presentation in 1957, Crick laid out the "Central Dogma" of molecular biology, which foretold the relationship between DNA, RNA, and proteins, and articulated the "adaptor hypothesis". Final confirmation of the replication mechanism that was implied by the double-helical structure followed in 1958 through the Meselson-Stahl experiment. Further work by Crick and co-workers showed that the genetic code was based on non-overlapping triplets of bases, called codons, allowing Har Gobind Khorana, Robert W. Holley and Marshall Warren Nirenberg to decipher the genetic code. These findings represent the birth of molecular biology.

RNA

Ribonucleic acid or RNA is a nucleic acid polymer consisting of nucleotide monomers that plays several important roles in the processes that translate genetic information from deoxyribonucleic acid (DNA) into protein products; RNA acts as a messenger between DNA and the protein synthesis complexes known as ribosomes, forms vital portions of ribosomes, and acts as an essential carrier molecule for amino acids to be used in protein synthesis.

RNA is very similar to DNA, but differs in a few important structural details: RNA nucleotides contain ribose sugars while DNA contains deoxyribose and RNA uses predominantly uracil instead of thymine present in DNA. RNA is transcribed

(synthesised) from DNA by enzymes called RNA polymerases and further processed by other enzymes. RNA serves as the template for translation of genes into proteins, transferring amino acids to the ribosome to form proteins, and also translating the transcript into proteins.

Nucleic acids were discovered in 1868 (some sources indicate 1869) by Johann Friedrich Miescher (1844-1895), who called the material 'nuclein' since it was found in the nucleus. It was later discovered that prokaryotic cells, which do not have a nucleus, also contain nucleic acids.

The role of RNA in protein synthesis had been suspected since 1939, based on experiments carried out by Torbjörn Caspersson, Jean Brachet and Jack Schultz. Hubert Chantrenne elucidated the messenger role played by RNA in the synthesis of proteins in ribosome.

The sequence of the 77 nucleotides of a yeast RNA was found by Robert W. Holley in 1964, winning Holley the 1968 Nobel Prize for Medicine. In 1976, Walter Fiers and his team at the University of Ghent determined the complete nucleotide sequence of bacteriophage MS2-RNA.

Chemical and Stereochemical Structure

RNA is a polymer with a ribose and phosphate backbone and four different bases: adenine, guanine, cytosine, and uracil. The first three are the same as those found in DNA, but in RNA thymine is replaced by uracil as the base complementary to adenine. This base is also a pyrimidine and is very similar to thymine.

Uracil is energetically less expensive to produce than thymine, which may account for its use in RNA. In DNA, however, uracil is readily produced by chemical degradation of cytosine, so having thymine as the normal base makes detection and repair of such incipient mutations more efficient. Thus, uracil is appropriate for RNA, where quantity is important but lifespan is not, whereas thymine is appropriate for DNA where maintaining sequence with high fidelity is more critical.

However, there are also numerous modified bases and sugars found in RNA that serve many different roles. Pseudouridine

(Ψ), in which the linkage between uracil and ribose is changed from a C–N bond to a C–C bond, and ribothymidine (T), are found in various places (most notably in the T Ψ C loop of tRNA). Thus, it is not technically correct to say that uracil is found in RNA in place of thymine. Another notable modified base is hypoxanthine (a deaminated Guanine base whose nucleoside is called Inosine). Inosine plays a key role in the Wobble Hypothesis of the Genetic Code. There are nearly 100 other naturally occurring modified nucleosides, of which pseudouridine and nucleosides with 2'–O–methylribose are by far the most common.

The specific roles of many of these modifications in RNA are not fully understood. However, it is notable that in ribosomal RNA, many of the post-translational modifications occur in highly functional regions, such as the peptidyl transferase centre and the subunit interface, implying that they are important for normal function.

Single stranded RNA exhibits a right handed stacking pattern that is stabilised by base stacking.

The most important structural feature of RNA, indeed the only consistent difference between the two nucleic acids, that distinguishes it from DNA is the presence of a hydroxyl group at the 2'-position of the ribose sugar.

The presence of this functional group enforces the C3'-endo sugar conformation (as opposed to the C2'-endo conformation of the deoxyribose sugar in DNA) that causes the helix to adopt the A-form geometry rather than the B-form most commonly observed in DNA. This results in a very deep and narrow major groove and a shallow and wide minor groove. A second consequence of the presence of the 2'-hydroxyl group is that in conformationally flexible regions of an RNA molecule (that is, not involved in formation of a double helix), it can chemically attack the adjacent phosphodiester bond to cleave the backbone.

Comparison with DNA

Unlike DNA, RNA is a single-stranded molecule in most of its biological roles and has a much shorter chain of nucleotides. RNA contains ribose, rather than the deoxyribose found in DNA

(there is no hydroxyl group attached to the pentose ring in the 2' position whereas RNA has two hydroxyl groups).

These hydroxyl groups make RNA less stable than DNA because it is more prone to hydrolysis. Several types of RNA (tRNA, rRNA) contain a great deal of secondary structure, which help promote stability.

Like DNA, most biologically active RNAs including tRNA, rRNA, snRNAs and other non-coding RNAs (such as the SRP RNAs) are extensively base paired to form double stranded helices. Structural analysis of these RNAs have revealed that they are not, "single-stranded" but rather highly structured. Unlike DNA, this structure is not just limited to long double-stranded helices but rather collections of short helices packed together into structures akin to proteins.

In this fashion, RNAs can achieve chemical catalysis, like enzymes. For instance, determination of the structure of the ribosome in 2000 revealed that the active site of this enzyme that catalyses peptide bond formation is composed entirely of RNA.

Synthesis

Synthesis of RNA is usually catalysed by an enzyme - RNA polymerase, using DNA as a template. Initiation of synthesis begins with the binding of the enzyme to a promoter sequence in the DNA (usually found "upstream" of a gene).

The DNA double helix is unwound by the helicase activity of the enzyme. The enzyme then progresses along the template strand in the $3' \rightarrow 5'$ direction, synthesising a complementary RNA molecule with elongation occurring in the $5' \rightarrow 3'$ direction. The DNA sequence also dictates where termination of RNA synthesis will occur.

There are also a number of RNA-dependent RNA polymerases as well that use RNA as their template for synthesis of a new strand of RNA. For instance, a number of RNA viruses (such as poliovirus) use this type of enzyme to replicate their genetic material. Also, it is known that RNA-dependent RNA polymerases are required for the RNA interference pathway in many organisms.

Biological Roles

Messenger RNA (mRNA): Messenger RNA is RNA that carries information from DNA to the ribosome sites of protein synthesis in the cell. In eukaryotic cells, once mRNA has been transcribed from DNA, it is "processed" before being exported from the nucleus into the cytoplasm, where it is bound to ribosomes and translated into its corresponding protein form with the help of tRNA. In prokaryotic cells, which have not partition into nucleus and cytoplasm compartments, mRNA can bind to ribosomes while it is being transcribed from DNA. After a certain amount of time the message degrades into its component nucleotides, usually with the assistance of ribonucleases.

Transfer RNA (tRNA): Transfer RNA is a small RNA chain of about 74-95 nucleotides that transfers a specific amino acid to a growing polypeptide chain at the ribosomal site of protein synthesis during translation. It has sites for amino-acid attachment and an anti-codon region for codon recognition that binds to a specific sequence on the messenger RNA chain through hydrogen bonding. It is a type of non-coding RNA.

Ribosomal RNA (rRNA): Ribosomal RNA is the catalytic component of the ribosomes, the protein synthetic factories in the cell. Eukaryotic ribosomes contain four different rRNA molecules: 18S, 5.8S, 28S, and 5S rRNA. Three of the rRNA molecules are synthesised in the nucleolus, and one is synthesised elsewhere. rRNA molecules are extremely abundant and make up at least 80 per cent of the RNA molecules found in a typical eukaryotic cell.

In the cytoplasm, ribosomal RNA and protein combine to form a nucleoprotein called a ribosome. The ribosome binds mRNA and carries out protein synthesis. Several ribosomes may be attached to a single mRNA at any time.

Non-coding RNA: RNA genes (sometimes referred to as non-coding RNA or small RNA) are genes that encode RNA that is not translated into a protein. The most prominent examples of RNA genes are transfer RNA (tRNA) and ribosomal RNA (rRNA), both of which are involved in the process of translation. However, since the late 1990s, many new RNA genes have been found, and thus RNA genes may play a much more significant role than previously thought.

In the late 1990s and early 2000, there has been persistent evidence of more complex transcription occurring in mammalian cells (and possibly others).

This could point towards a more widespread use of RNA in biology, particularly in gene regulation. A particular class of non-coding RNA, micro RNA, has been found in many metazoans (from Caenorhabditis elegans to *Homo sapiens*) and clearly plays an important role in regulating other genes.

First proposed in 2004 by Rassoulzadegan and published in Nature 2006, RNA is implicated as being part of the germline. If confirmed, this result would significantly alter the present understanding of genetics and lead to many questions on DNA-RNA roles and interactions.

Catalytic RNA: Although RNA contains only four bases, in comparison to the twenty-odd amino acids commonly found in proteins, certain RNAs are still able to catalyse chemical reactions. These include cutting and ligating other RNA molecules and also the catalysis of peptide bond formation in the ribosome.

Double-stranded RNA: Double-stranded RNA (or dsRNA) is RNA with two complementary strands, similar to the DNA found in all "higher" cells. dsRNA forms the genetic material of some viruses.

In eukaryotes, it acts as a trigger to initiate the process of RNA interference and is present as an intermediate step in the formation of siRNAs (small interfering RNAs). siRNAs are often confused with miRNAs; siRNAs are double-stranded, whereas miRNAs are single-stranded. Although initially single stranded, there are regions of intramolecular association causing hairpin structures in pre-miRNAs; immature miRNAs.

Very recently, dsRNA has been found to induce gene expression at transcriptional level, a phenomenon named "small RNA induced gene activation RNAa". Such dsRNA is called "small activating RNA (saRNA)".

RNA World Hypothesis

The RNA world hypothesis proposes that the earliest forms of life relied on RNA both to carry genetic information (like DNA

does now) and to catalyse biochemical reactions like an enzyme. According to this hypothesis, descendants of these early lifeforms gradually integrated DNA and proteins into their metabolism.

RNA Secondary Structures

The functional form of single stranded RNA molecules (like proteins) frequently requires a specific tertiary structure. The scaffold for this structure is provided by secondary structural elements which are hydrogen bonds within the molecule. This leads to several recognisable "domains" of secondary structure like hairpin loops, bulges and internal loops. The secondary structure of RNA molecules can be predicted computationally by calculating the minimum free energies (MFE) structure for all different combinations of hydrogen bondings and domains. There has been a significant amount of research directed at the RNA structure prediction problem.

Uses of Oligosaccharide

An oligosaccharide is a saccharide polymer containing a small number (typically three to ten) of component sugars, also known as simple sugars. They are generally found either *O*- or *N*-linked to compatible amino acid side chains in proteins or to lipid moieties.

Examples: Fructo-oligosaccharides (FOS) and inulin, which are found in many vegetables, consist of short chains of fructose molecules. Inulin has a much higher degree of polymerisation than FOS. Galacto-oligosaccharides (GOS), which also occur naturally, consist of short chains of galactose molecules. These compounds can be only partially digested by humans.

Oligosaccharides are often found as a component of glycoproteins or glycolipids and as such are often used as chemical markers, often for cell recognition.

An example is ABO blood type specificity. A and B blood types have two different oligosaccharide glycolipids embedded in the cell membranes of the red blood cells, AB-type blood has both, while O blood type has neither.

Mannan-oligosaccharides (MOS) are widely used in animal feed to encourage gastrointestinal health and performance. They

are normally obtained from the yeast cell walls of Saccharomyces cerevisiae. Some brand names are: Bio-Mos, SAF-Mannan, Y-MOS and Celmanax.

Therapeutic Effects

When oligosaccharides are consumed, the undigested portion serves as food for the intestinal microflora. Depending on the type of oligosaccharide, different bacterial groups are stimulated or suppressed.

Clinical studies have shown that administering FOS, GOS, or inulin can increase the number of these friendly bacteria in the colon while simultaneously reducing the population of harmful bacteria.

Other benefits noted with FOS, GOS, or inulin supplementation include increased production of beneficial short-chain fatty acids such as butyrate, increased absorption of calcium and magnesium, and improved elimination of toxic compounds.

Because FOS, GOS, and inulin improve colon function and may influence the bacterial composition, one might expect these compounds would help relieve the symptoms of irritable bowel syndrome. However, a double-blind trial found no clear benefit with FOS supplementation (2 grams three times daily) in patients with this condition. Experimental studies with FOS in animals suggest a possible benefit in lowering blood sugar levels in people with diabetes and in reducing elevated blood cholesterol and triglyceride levels.

In a double-blind trial of middle-aged men and women with elevated cholesterol and triglyceride levels, supplementation with inulin (10 grams per day for eight weeks) significantly reduced insulin concentrations, suggesting an improvement in blood-glucose control, and significantly lowered triglyceride levels.

In a preliminary trial, administration of FOS (8 grams per day for two weeks) significantly lowered fasting blood-sugar levels and serum total-cholesterol levels in patients with type 2 (non-insulin-dependent) diabetes.

However, in another trial, people with type 2 diabetes supplementing with FOS (15 grams per day) for 20 days found

no effect on blood-glucose or lipid levels. Because of these conflicting results, more research is needed to determine the effect of FOS and inulin on diabetes and lipid levels.

Several double-blind trials have looked at the ability of FOS or inulin to lower blood cholesterol and triglyceride levels. These trials have shown that in people with elevated total cholesterol or triglyceride levels, including people with type 2 (adult onset) diabetes, FOS or inulin (in amounts ranging from 8 to 20 grams daily) produced significant reductions in triglyceride levels. However, the effect on cholesterol levels was inconsistent. In people with normal or low cholesterol or triglyceride levels, FOS or inulin produced little effect.

Sources

FOS and inulin are found naturally in Jerusalem artichoke, burdock, chicory, leeks, onions, and asparagus. FOS products derived from chicory root contain significant quantities of inulin, a fibre widely distributed in fruits, vegetables and plants. Inulin is a significant part of the daily diet of most of the world's population. FOS can also be synthesised by enzymes of the fungus *Aspergillus niger* acting on sucrose.

GOS is naturally found in soybeans and can be synthesised from lactose (milk sugar). FOS, GOS, and inulin are available as nutritional supplements in capsules, tablets, and as a powder.

Not all natural oligosaccharides occur as components of glycoproteins or glycolipids. Some, such as the raffinose series, occur as storage or transport carbohydrates in plants. Others, such as maltodextrins or cellodextrins, result from the microbial breakdown of larger polysaccharides such as starch or cellulose.

Significance of Polysaccharide

Polysaccharides are relatively complex carbohydrates.

They are polymers made up of many monosaccharides joined together by glycosidic linkages. They are, therefore, very large, often branched, molecules. They tend to be amorphous, insoluble in water, and have no sweet taste.

When all the constituent monosaccharides are of the same type they are termed *homopolysaccharides*; when more than one type of monosaccharide is present they are termed *heteropolysaccharides*.

Examples include storage polysaccharides such as starch and glycogen and structural polysaccharides such as cellulose and chitin.

Polysaccharides have a general formula of $C_n(H_2O)_n$ where n is usually a large number between 200 and 2500. The general formula can also be represented as $(C_6H_{12}O_6)_n$ where n = 40 – 3000.

Starches

Starches are glucose polymers in which glucopyranose units are bonded by *alpha*-linkages. It is made up of a mixture of Amylose and Amylopectin. Amylose consists of a linear chain of several hundred glucose molecules and Amylopectin is a branched molecule made of several thousand glucose units.

Starches are insoluble in water. They can be digested by hydrolysis, catalysed by enzymes called amylases, which can break the *alpha*-linkages. Humans and other animals have amylases, so they can digest starches. Potato, rice, wheat, and maize are major sources of starch in the human diet.

Cellulose

The structural components of plants are formed primarily from cellulose. Wood is largely cellulose and lignin, while paper and cotton are nearly pure cellulose. Cellulose is a polymer made with repeated glucose units bonded together by *beta*-linkages.

Humans and many other animals lack an enzyme to break the *beta*-linkages, so they do not digest cellulose. Certain animals can digest cellulose, because bacteria possessing the enzyme are present in their gut. The classic example is the termite.

Acidic Polysaccharides

Acidic polysaccharides are polysaccharides that contain carboxyl groups, phosphate groups and/or sulphuric ester groups.

Bacterial Capsule Polysaccharides

Pathogenic bacteria commonly produce a thick, mucous-like, layer of polysaccharide. This "capsule" cloaks antigenic proteins on the bacterial surface that would otherwise provoke an immune response and thereby lead to the destruction of the bacteria. Capsular polysaccharides are water soluble, commonly acidic, and have molecular weights on the order of 100 – 1000 kDa. They are linear and consist of regularly repeating subunits of one to six monosaccharides. There is enormous structural diversity; nearly two hundred different polysaccharides are produced by *E. coli* alone. Mixtures of capsular polysaccharides, either conjugated or native are used as vaccines.

Bacteria and many other microbes, including fungi and algae, often secrete polysaccharides as an evolutionary adaptation to help them adhere to surfaces and to prevent them from drying out. Humans have developed some of these polysaccharides into useful products, including xanthan gum, dextran, gellan gum, and pullulan.

Prion

A prion—short for proteinaceous infectious particle (-on) that lacks nucleic acid (by analogy to virion) — is a type of infectious agent composed only of protein. They cause a number of diseases in a variety of animals, including bovine spongiform encephalopathy (BSE, also known as mad cow disease) in cattle and Creutzfeldt-Jakob disease in humans. All known prion diseases affect the structure of the brain or other neural tissue, and all are untreatable and fatal.

Prions are believed to infect and propagate by refolding abnormally into a structure which is able to convert normal molecules of the protein into the abnormally structured form. However, the term in itself does not preclude other mechanisms of transmission.

All known prions induce the formation of an amyloid fold, in which the protein polymerises into a fibre with a core consisting of tightly packed beta sheets. Other mechanisms may exist in yet undiscovered infectious protein particles. This altered structure

renders them quite resistant to denaturation by chemical and physical agents, although infectivity can be reduced by these treatments, making disposal and containment of these particles difficult.

Proteins showing prion behaviour are also found in some fungi. Some fungal prions may not be associated with any disease; it is unknown whether these prions represent an evolutionary advantage for their hosts.

Discovery

Radiation biologist Tikvah Alper and physicist J.S. Griffith developed the theory in the 1960s that some TSEs are caused by an infectious agent made solely of protein. This theory was developed to explain the discovery that the mysterious infectious agent causing the diseases scrapie and Creutzfeldt-Jakob Disease resisted ultraviolet radiation (which breaks down nucleic acids present in viruses and all living things) yet responded to agents that disrupt proteins.

Francis Crick recognised the potential importance of the Griffith protein-only hypothesis for scrapie propagation in the second edition of his famous "Central dogma of molecular biology". While asserting that the flow of sequence information from protein to protein, or from protein to RNA and DNA was "precluded" by the dogma, he noted that Griffith's hypothesis was a potential difficulty (although it was not so promoted by Griffith). As the revised "dogma" was formulated, in part, to accommodate the then recent discovery of reverse transcription by Howard Temin and David Baltimore (Nobel Prize, 1975), proof of the protein-only hypothesis might have been seen as a sure bet for a future Prize.

Stanley B. Prusiner of the University of California, San Francisco announced in 1982 that his team had purified infectious material and that the infectious agent consisted mainly of a specific protein, although they had not managed to satisfactory isolate the protein until two years after making his announcement.

Prusiner coined the word "prion" as a name for the infectious agent, by combining the first two syllables of the words *proteinaceous* and *infectious*. While the infectious agent was named a prion, the

specific protein that the prion was made of was named PrP, an abbreviation for "protease-resistant protein". Prusiner was awarded the Nobel Prize in Physiology or Medicine in 1997 for his research into prions.

Structure

Isoforms: The protein that prions are made of is found throughout the body, even in healthy people and animals. However, the prion protein found in infectious material has a different structure and is resistant to proteases, the enzymes in the body that can normally break down proteins. The normal form of the protein is called PrP^{C}, while the infectious form is called PrP^{Sc} — the *C* refers to 'cellular' or 'common' PrP, while the *Sc* refers to 'scrapie', a prion disease occurring in sheep. While PrP^{C} is structurally well-defined, PrP^{Sc} is certainly polydisperse and defined at a relatively poor level. PrP can be induced to fold into other more-or-less well-defined isoforms in vitro, and their relationship to the form(s) that are pathogenic *in vivo* is not yet clear.

PrP^{C}: PrP^{C} is a normal protein found on the membranes of cells. Several topological forms of it exist; one cell surface form anchored via glycolipid and two transmembrane forms, however its function has not been fully resolved. PrP^{C} is readily digested by proteinase K and can be liberated from the cell surface by the enzyme *phosphatidyl inositol-specific phospholipase C*, which cleaves the *phosphatidyl inositol* glycolipid anchor. A typical yeast prion protein contains a core region (domain) with many repeats of the amino acids glutamine and asparagine. Normal yeast prion domains are flexible and lack a defined structure.

PrP^{Sc}: The infectious isoform of PrP^{C}, known as PrP^{Sc}, is able to catalyse the formation of other normal PrP^{C} proteins into the infectious isoform by changing their conformation. Although the exact 3D structure of PrP^{Sc} is not known, there is increased β-sheet content in the diseased form of the molecule, replacing normal areas of α-helix. Aggregations of these abnormal isoforms forms a highly structured amyloid fibre. The end of the fibre acts as a template for the free protein molecules, causing the fibre to grow.

Small differences in the amino acid sequence of prion-forming regions lead to distinct structural features on the surface of prion fibres. As a result, only free protein molecules that are identical in amino acid sequence to the prion protein can be recruited into the growing fibre. The mammalian prion proteins do not resemble the prion proteins of yeast in their amino acid sequence, however, they are still known as PrP^{C} and PrP^{Sc} and share basic structural features.

Function

PrP and Long-term Memory: There is evidence that PrP may have a normal function in maintenance of long term memory. Maglio and colleagues have shown that mice without the genes for normal cellular PrP protein have altered hippocampal LTP.

Prion Disease

Prions cause neurodegenerative disease by aggregating extracellularly within the central nervous system to form plaques known as amyloids, which disrupt the normal tissue structure. This disruption is characterised by holes in the tissue with resultant spongy architecture due to the loss of neurons. Other histological changes include astrogliosis and the absence of an inflammatory reaction. While the incubation period for prion diseases is generally quite long, once symptoms appear the disease progresses rapidly, leading to brain damage and death. Neurodegenerative symptoms can include convulsions, dementia, ataxia (balance and coordination dysfunction), and behavioural or personality changes.

All known prion diseases, collectively called *transmissible spongiform encephalopathies* (TSEs), are untreatable and fatal. However, a vaccine has been developed in mice that may provide insight into providing a vaccine in humans to resist prion infections. Additionally, in 2006 scientists announced that they had genetically engineered cattle lacking a necessary gene for prion production—thus theoretically making them immune to BSE, building on research indicating that mice lacking normally-occurring prion protein are resistant to infection by scrapie prion protein.

Prions are able to affect a variety of different species, however the prions involved are somewhat species-specific: they are similar

but not identical. However overlap may occur; the human prion disease *variant Creutzfeldt-Jakob disease* is believed to be caused by a prion which typically infects cattle and is transmitted through infected meat.

Metal ion interactions with prion proteins might be relevant to the progression of prion-mediated disease, based on epidemiological studies of clusters of prion disease in locales with low soil concentrations of copper.

The following diseases are believed to be caused by prions.

- In animals:
  - — Scrapie in sheep.
  - — Bovine spongiform encephalopathy (BSE) in cattle (known as *mad cow disease*).
  - — Transmissible mink encephalopathy (TME) in mink.
  - — Chronic wasting disease (CWD) in elk and mule deer.
  - — Feline spongiform encephalopathy in cats.
  - — Exotic ungulate encephalopathy (EUE) in nyala, oryx and greater kudu.
- In humans:
  - — Creutzfeldt-Jakob disease (CJD) and its varieties: iatrogenic Creutzfeldt-Jakob disease (iCJD), variant Creutzfeldt-Jakob disease (vCJD), familial Creutzfeldt-Jakob disease (fCJD), and sporadic Creutzfeldt-Jakob disease (sCJD).
  - — Gerstmann-Straussler-Scheinker syndrome (GSS).
  - — Fatal familial insomnia (fFI).
  - — Sporadic fatal insomnia (sFI).
  - — Kuru.
  - — Alpers syndrome.

Transmission

Although the identity and general properties of prions are now well-understood, the mechanism of prion infection and propagation remains mysterious. It is often assumed that the diseased form directly interacts with the normal form to make it

rearrange its structure. One idea, the "Protein X" hypothesis, is that an as-yet unidentified cellular protein (Protein X) enables the conversion of PrP^{C} to PrP^{Sc} by bringing a molecule of each of the two together into a complex.

Sterilisation

Infectious particles possessing nucleic acid are dependent upon it to direct their continued replication. Prions, however, are infectious by their effect on normal versions of the protein. Therefore, sterilising prions involves the denaturation of the protein to a state where the molecule is no longer able to induce the abnormal folding of normal proteins. However, prions are generally quite resistant to denaturation by proteases, heat, radiation, and formalin treatments, although their infectivity can be reduced by such treatments.

Prions can be denatured by subjecting them to a temperatures of 134 degrees Celsius for 18 minutes in a pressurised steam autoclave. Ozone sterilisation is currently being studied as a potential method for prion deactivation. Renaturation of a completely denatured prion to infectious status has not yet been achieved, however partially denatured prions can be renatured to an infective status under certain artificial conditions.

Infectious Agent Debate

Protein-only Hypothesis: Prior to the discovery of prions, it was thought that all pathogens used nucleic acids to direct their replication. The "protein-only hypothesis" states that a protein structure can replicate without the use of nucleic acid. This was initially controversial as it contradicts the so-called "central dogma of modern biology," which describes nucleic acid as the central form of replicative information.

Evidence in favour of a protein-only hypothesis include:

- No virus particles have been conclusively associated with prion diseases.
- No nucleic acid has been conclusively associated with infectivity; agent is resistant to degradation by nucleases.
- No immune response to infection.

- PrP^{Sc} experimentally transmitted between one species and another results in PrP^{Sc} with the amino-acid sequence of the recipient species, suggesting that replication of the donor agent doesn't occur.
- Level of infectivity is associated with levels of PrP^{Sc}.
- PrP^{Sc} and PrP^{C} do not differ in amino-acid sequence, therefore a *PrP^{Sc}-specific* nucleic acid is a redundant concept.
- Familial prion disease occurs in families with a mutation in the PrP gene, and mice with PrP mutations develop prion disease despite controlled conditions where transmission is prevented.
- PrP^{Sc} has been shown to arise from exposure of PrP^{C} to molecules of PrP^{Sc}.

Viral Hypothesis: The protein-only hypothesis was initially met with skepticism and still has critics. For more than a decade, Yale University neuropathologist Laura Manuelidis has been proposing that prion diseases are caused instead by an unidentifiable "slow virus". In January 2007, she and her colleagues published a non-peer reviewed article in the Proceedings of the National Academy of Science claiming to have found the elusive virus in less than 10 per cent of their scrapie-infected cells in culture.

Evidence in favour of a viral hypothesis include:

- No bacteria or other living organisms have been found in prion-affected organisms, defaulting to the idea that a virus must be involved.
- The long incubation and rapid onset of symptoms resembles some viral infections, such as HIV-induced AIDS.
- Differences in prion infectivity, incubation, symptomology and progression among species resembles the "strain variation" seen between viruses, especially RNA viruses.
- Familial prion disease is proposed to be due to genetic predisposition to the viral agent.

Genetics

A gene for the normal protein has been isolated: the PRNP gene. Some prion diseases can be inherited, and in all inherited

cases there is a mutation in the *PRNP* gene. Many different *PRNP* mutations have been identified and it is thought that the mutations somehow make PrP^{C} more likely to spontaneously change into the abnormal PrP^{Sc} form. Prion diseases are the only known diseases that can be sporadic, genetic, or infectious.

It should be noted that the same gene is responsible for spongiform encephalopathies which are not known to be transmissible, as well as some non-neurological diseases. Some require a mutation for transmission to occur, and there are *respective mutations* which can prevent transmission for most of the TSEs. Whether or not the prion gene has a non-disease function is an area of considerable active research.

Prions in Yeast and Other Fungi

Prion-like proteins that behave in a similar way to PrP are found naturally in some fungi and non-mammalian animals. A group at the Whitehead Institute has argued that some of the fungal prions are not associated with any disease state and may have a useful role; however, researchers at the NIH have also provided strong arguments demonstrating that fungal prions should be considered a diseased state.

Research into fungal prions has given strong support to the protein-only hypothesis for mammalian prions, as it has been demonstrated that seeds extracted from cells with the prion state, can convert the normal form of the protein into the infectious form *in vitro*, and in the process, preserve the information corresponding to different strains of the prion state. It has also shed some light on prion domains, which are regions in a protein that promote the conversion. Fungal prions have helped to suggest mechanisms of conversion that may apply to all prions.

Significance of Neurotransmitter

Neurotransmitters are chemicals that are used to relay, amplify and modulate electrical signals between a neuron and another cell. According to the prevailing beliefs of the 1960s, a chemical can be classified as a neurotransmitter if it meets the following conditions:

- It is synthesised endogenously, that is, within the presynaptic neuron;
- It is available in sufficient quantity in the presynaptic neuron to exert an effect on the post-synaptic neuron;
- Externally administered, it must mimic the endogenously-released substance; and
- A biochemical mechanism for inactivation must be present.

However, there are other materials, such as the zinc ion, that are neither synthesised nor catabolised and are considered neurotransmitters by some. Thus, the old definitions are being revised.

Types of Neurotransmitters

There are many different ways to classify neurotransmitters. Often, dividing them into amino acids, peptides, and monoamines is sufficient for many purposes.

Some more precise divisions are as follows:

- Around 10 "small-molecule neurotransmitters" are known:
  - — Acetylcholine,
  - — Monoamines (norepinephrine NE, dopamine DA and serotonin 5-HT),
  - — 3 or 4 amino acids, depending on exact definition used: (primarily glutamic acid, GABA, aspartic acid and glycine),
  - — Purines, (Adenosine, ATP, GTP and their derivatives), and
  - — Fatty acids are also receiving attention as the potential endogenous cannabinoid.
- Over 50 neuroactive peptides (vasopressin, somatostatin, neurotensin, etc.) have been found, among them hormones such as LH or insulin that have specific local actions in addition to their long-range signalling properties.
- Single ions, such as synaptically-released zinc, are also considered neurotransmitters by some.

The major "workhorse" neurotransmitters of the brain are glutamic acid (=glutamate) and GABA.

Effects

Some examples of neurotransmitter action:

- Acetylcholine—voluntary movement of the muscles
- Norepinephrine—wakefulness or arousal
- Dopamine—voluntary movement and motivation, "wanting"
- Serotonin—memory, emotions, wakefulness, sleep and temperature regulation
- GABA (gamma aminobutyric acid)—inhibition of motor neurons
- Glycine—spinal reflexes and motor behaviour
- Neuromodulators—sensory transmission—especially pain

It is important to appreciate that it is the receptor that dictates the neurotransmitter's effect.

Mechanism of Action

Within the cells, small-molecule neurotransmitter molecules are usually packaged in vesicles. When an action potential travels to the synapse, the rapid depolarisation causes calcium ion channels to open.

Calcium then stimulates the transport of vesicles to the synaptic membrane; the vesicle and cell membrane fuse, leading to the release of the packaged neurotransmitter, a mechanism called exocytosis.

The neurotransmitters then diffuse across the synaptic cleft to bind to receptors. The receptors are broadly classified into ionotropic and metabotropic receptors. Ionotropic receptors are ligand-gated ion channels that open or close through neurotransmitter binding. Metabotropic receptors, which can have a diverse range of effects on a cell, transduct the signal by secondary messenger systems, or G-proteins.

Neuroactive peptides are made in the neuron's soma and are transported through the axon to the synapse. They are usually packaged into dense-core vesicles and are released through a

similar, but metabolically distinct, form of exocytosis used for small-molecule synaptic vesicles.

Post-synaptic Effect

A neurotransmitter's effect is determined by its receptor. For example, GABA can act on both rapid or slow inhibitory receptors (the GABA-A and GABA-B receptor respectively). Many other neurotransmitters, however, may have excitatory or inhibitory actions depending on which receptor they bind to.

Neurotransmitters may cause either excitatory or inhibitory post-synaptic potentials. That is, they may help the initiation of a nerve impulse in the receiving neuron, or they may discourage such an impulse by modifying the local membrane voltage potential. In the central nervous system, combined input from several synapses is usually required to trigger an action potential. Glutamate is the most prominent of excitatory transmitters; GABA and glycine are well-known inhibitory neurotransmitters.

Many neurotransmitters are removed from the synaptic cleft by neurotransmitter transporters in a process called reuptake (or often simply 'uptake'). Without reuptake, the molecules might continue to stimulate or inhibit the firing of the post-synaptic neuron. Another mechanism for removal of a neurotransmitter is digestion by an enzyme.

For example, at cholinergic synapses (where acetylcholine is the neurotransmitter), the enzyme acetylcholinesterase breaks down the acetylcholine. Neuroactive peptides are often removed from the cleft by diffusion, and eventually broken down by proteases.

Specifications

While some neurotransmitters (glutamate, GABA, glycine) are used very generally throughout the central nervous system, others can have more specific effects, such as on the autonomic nervous system, by both pathways in the sympathetic nervous system and the parasympathetic nervous system, and the action of others are regulated by distinct classes of nerve clusters which can be arranged in familiar pathways around the brain. For example, Serotonin is released specifically by cells in the brainstem,

in an area called the raphe nuclei, but travels around the brain along the medial forebrain bundle activating the cortex, hippocampus, thalamus, hypothalamus and cerebellum. Also, it is released in the Caudal serotonin nuclei, so as to have effect on the spinal cord. In the peripherial nervous system (such as in the gut wall) serotonin regulates vascular tone. Dopamine classically modulates two systems: the brain's reward mechanism, and movement control.

Neurotransmitters that have these types of specific actions are often targeted by drugs.

- Cocaine, for example, blocks the reuptake of dopamine, leaving these neurotransmitters in the synaptic gap longer.
- Prozac is a serotonin reuptake inhibitor, hence potentiating its effect.
- AMPT prevents the conversion of tyrosine to L-DOPA, the precursor to dopamine; reserpine prevents dopamine storage within vesicles; and deprenyl inhibits monoamine oxidase (MAO)-B and thus increases dopamine levels.

Some neurotransmitter/neuromodulators like zinc not only can modulate the sensitivity of a receptor to other neurotransmitters (allosteric modulation) but can even penetrate specific, gated channels in post-synaptic neurons, thus entering the post-synaptic cells. This "translocation" is another mechanism by which synaptic transmitters can affect post-synaptic cells.

Diseases may affect specific neurotransmitter pathways. For example, Parkinson's disease is at least in part related to failure of dopaminergic cells in deep-brain nuclei, for example the substantia nigra. Treatments potentiating the effect of dopamine precursors have been proposed and effected, with moderate success.

Basics of Carbohydrates

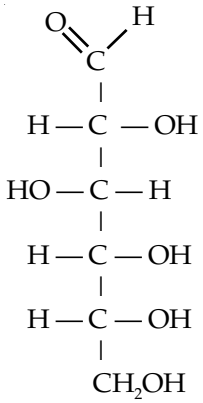
Carbohydrates or saccharides are simple molecules that are straight-chain aldehydes or ketones with many hydroxyl groups added, usually one on each carbon atom that is not part of the aldehyde or ketone functional group. Carbohydrates are the most

abundant biological molecules, and fill numerous roles in living things, such as the storage and transport of energy (starch, glycogen) and structural components (cellulose in plants, chitin in animals). Additionally, carbohydrates and their derivatives play major roles in the working process of the immune system, fertilization, pathogenesis, blood clotting, and development.

The basic carbohydrate units are called monosaccharides, such as glucose, galactose, and fructose. The general chemical formula of an unmodified monosaccharide is $(C H_2O)_n$, where n is any number of three or greater. Monosaccharides can be linked together in almost limitless ways.

Two joined monosaccharides are called disaccharides, such as sucrose and lactose. Carbohydrates containing between about three to six monosaccharide units are termed oligosaccharides; anything larger than this is a polysaccharide. Polysaccharides, such as starch, glycogen, or cellulose, can reach many thousands of units in length. Many carbohydrates contain one or more modified monosaccharide units that have had one or more groups replaced or removed. For example, deoxyribose, a component of DNA, is a modified version of ribose; chitin is composed of repeating units of N-acetylglucosamine, a nitrogen-containing form of glucose.

Categorisation of Monosaccharides



D-glucose is an aldohexose with the formula $(C.H_2O)_6$.

Monosaccharides are the simplest form of carbohydrates. They are aldehydes or ketones with many hydroxyl groups added, usually one on each carbon atom that is not part of the aldehyde or ketone functional group. They are called polyhydroxyl aldehydes or polyhydroxyl ketones respectively. The general chemical formula of an unmodified monosaccharide is $(C.H_2O)_n$, where n is any number of three or greater.

Classification of Monosaccharides

Monosaccharides are classified according to three different characteristics: the placement of its carbonyl group, the number of carbon atoms it contains, and its chiral handedness. If the carbonyl group is an aldehyde, the monosaccharide is an aldose; if the carbonyl group is a ketone, the monosaccharide is a ketose. The smallest possible monosaccharide, those with three carbon atoms, are called trioses.

Those with four are called tetroses, five are called pentoses, six are hexoses, and so on. These two systems of classification are often combined.

For example, glucose is an aldohexose (a six-carbon aldehyde), ribose is an aldopentose (a five-carbon aldehyde), and fructose is a ketohexose (a six-carbon ketone).

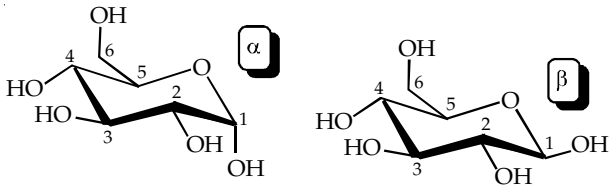


Fig. The α and β anomers of glucose. Note the position of the anomeric carbon (red or green) relative to the CH_2OH group bound to carbon 5: they are either on the opposite sides (α), or the same side (β).

Each carbon atom bearing a hydroxyl group (–OH), with the exception of the first and last carbons, are asymmetric, making them stereocenters with two possible configurations each (the –H and –OH may be on either side). Because of this asymmetry, a number of isomers may exist for any given monosaccharide formula.

The aldohexose D-glucose, for example, has the formula $(C H_2O)_6$, of which all but two of its six carbons atoms are chiral centres, making D-glucose one of $2^4 = 16$ possible stereoisomers. In the case of a triose, there is one pair of possible stereoisomers, which are enantiomers and epimers.

The assignment of D or L is made according to the orientation of the asymmetric carbon furthest from the carbonyl group: if the hydroxyl group is on the right the molecule is a D sugar, otherwise it is an L sugar. Because D sugars are biologically far more common, the D is often omitted.

Conformation

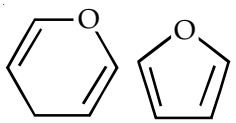


Fig. Pyran and furan, after which the pyranose and furanose configurations of monosaccharides are named.

The aldehyde or ketone group of a straight-chain monosaccharide will react reversibly with a hydroxyl group on a different carbon atom to form a hemiacetal or hemiketal, forming a heterocyclic ring with an oxygen bridge between two carbon atoms. Rings with five and six atoms are called furanose and pyranose forms, respectively, and exist in equilibrium with the straight-chain form.

During the conversion from straight-chain form to cyclic form, the carbon atom containing the carbonyl oxygen, called the anomeric carbon, becomes a chiral centre with two possible configurations: the oxygen atom may take a position either above or below the plane of the ring. The resulting possible pair of stereoisomers are called anomers. In the α *anomer*, the –OH substituent on the anomeric carbon rests on the opposite side of the ring from the CH_2OH attached to the asymmetric carbon furthest from the anomeric carbon. The alternative form, in which the CH_2OH and the anomeric hydroxyl are on the same side of the plane of the ring, is called the β *anomer*. Because the ring and straight-chain forms readily interconvert, both anomers exist in equilibrium.

Disaccharides

Disaccharides are the simplest polysaccharides. They are composed of two monosaccharide units bound together by a covalent glycosidic bond formed via a dehydration reaction, resulting in the loss of a hydrogen atom from one monosaccharide and a hydroxyl group from the other, so the formula of unmodified disaccharides is $C_{12}H_{22}O_{11}$. Although there are numerous kinds of disaccharides, a handful of disaccharides are particularly notable.

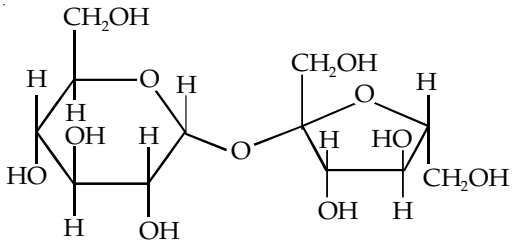


Fig. Sucrose, also known as table sugar, is a common disaccharide. It is composed of two linked monosaccharides: glucose (left) and fructose (right).

Sucrose, pictured to the right, is the most abundant disaccharide and the main form in which carbohydrates are transported in plants. It is composed of one glucose molecule and one fructose molecule. The systematic name for sucrose, O-α-D-glucopyranosyl-(1$\rightarrow$2)-D-fructofuranoside, indicates four things:

- Its monosaccharides: glucose and fructose.
- Their ring types: glucose is a pyranose, and fructose is a furanose.
- How they're linked together: the oxygen on the number 1 carbon (C1) of α-glucose is linked to the C2 of fructose.
- The *-oside* suffix indicates that the anomeric carbon of both monosacchaides participates in the glycosidic bond.

Lactose, a disaccharide composed of one galactose molecule and one glucose molecule, occurs naturally only in milk. The systematic name for lactose is *O*-β-D-galactopyranosyl-(1$\rightarrow$4)-D-glucopyranose. Other notable disaccharides include maltose (two glucoses linked α-1,4) and cellobiose (two glucoses linked β-1,4).

Oligosaccharides and Polysaccharides

Oligosaccharides and polysaccharides are composed of longer chains of monosaccharide units bound together by glycosidic bonds. The distinction between the two is based upon the number of monosaccharide units present in the chain. Oligosaccharides typically contain between two and nine monosaccharide units, and polysaccharides contain greater than ten monosaccharide units. Definitions of how large a carbohydrate must be to fall into each category vary according to personal opinion. Examples of oligosaccharides include the disaccharides mentioned above, the trisaccharide raffinose and the tetrasaccharide stachyose.

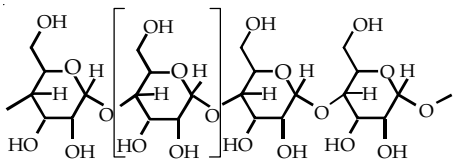


Fig. Amylose is a linear polymer of glucose mainly linked with $\alpha\,(1 \rightarrow 4)$ bonds. It can be made of several thousands of glucose units. It is one of the two components of starch, the other being amylopectin.

Oligosaccharides are found as a common form of protein post-translational modification. Such post-translational modifications include the Lewis oligosaccharides responsible for blood group incompatibilities, the alpha-Gal epitope responsible for hyperacute rejection in xenotransplanation, and O-GlcNAc modifications.

Polysaccharides represent an important class of biological polymer. Examples include starch, cellulose, chitin, glycogen, callose, laminarin, xylan, and galactomannan.

Nutrition

Carbohydrates require less water to digest than proteins or fats and are the most common source of energy. Proteins and fat are vital building components for body tissue and cells, and thus it could be considered advisable not to deplete such resources by necessitating their use in energy production.

Carbohydrates are not essential nutrients: the body can obtain all its energy from protein and fats. The brain cannot burn fat and

needs glucose for energy, but the body can make this glucose from protein. Carbohydrates, like proteins, contain 4 kilocalories per gram while fats contain 9 kilocalories and alcohol contains 7 kilocalories per gram.

Foods that are High in Carbohydrates: Breads, pastas, beans, potatoes, bran, rice and cereals.

Based on evidence for risk of heart disease and obesity, the Institute of Medicine recommends that American and Canadian adults get between 40-65 per cent of dietary energy from carbohydrates. The Food and Agriculture Organisation and World Health Organisation jointly recommend that national dietary guidelines set a goal of 55-75 per cent of total energy from carbohydrates, but only 10 per cent should be from Free sugars (their definition of simple carbohydrates).

This distinction between "good carbs" and "bad carbs" is a large part of the low-carbohydrate diets, which in general promotes a reduction in the consumption of grains and starches in favour of vegetables, thereby reducing the high insulin levels that cause fat storage, and increasing the burning of body fat for energy through ketosis.

Classification

Dietitians and nutritionists commonly classify carbohydrates as simple (monosaccharides and disaccharides) or complex (oligosaccharides and polysaccharides). The term *complex carbohydrate* was first used in the Senate Select Committee publication *Dietary Goals for the United States* (1977), where it denoted "fruit, vegetables and whole-grains". Dietary guidelines generally recommend that complex carbohydrates and nutrient-rich simple carbohydrates such as fruit and dairy products should make up the bulk of carbohydrate consumption. The USDA's *Dietary Guidelines for Americans 2005* dispenses with the simple/complex distinction, instead recommending fibre-rich foods and whole grains.

The glycemic index and glycemic load systems are popular alternative classification methods which rank carbohydrate-rich foods based on their effect on blood glucose levels. The insulin

index is a similar, more recent classification method which ranks foods based on their effects on blood insulin levels. This system assumes that high glycaemic index foods and low glycaemic index foods can be mixed to make the intake of high glycaemic foods more acceptable. The World Health Organisation and Food and Agriculture Organisation's joint expert report on Diet, Nutrition and the Prevention of Chronic Diseases (WHO Technical Report Series 916) advises carbohydrate consumption of 55-75 per cent carbohydrate, but restricts "Free sugar" intake to 10 per cent. Its definition is "The term "free sugars" refers to all monosaccharides and disaccharides added to foods by the manufacturer, cook or consumer, plus sugars naturally present in honey, syrups and fruit juices." This is their effective split between simple and complex carbohydrates.

Metabolism

Catabolism: Catabolism is the metabolic reaction cells undergo in order to extract energy. There are two major metabolic pathways of monosaccharide catabolism:

1. Glycolysis
2. Citric acid cycle

Oligo/polysaccharides are cleaved first to smaller monosaccharides by enzymes called Glycoside hydrolases. The monosaccharide units can then enter into monosaccharide catabolism.

Monomer

A monomer (from Greek *mono* "one" and *meros* "part") is a small molecule that may become chemically bonded to other monomers to form a polymer.

Examples of monomers are hydrocarbons such as the alkene and arene homologous series. Here hydrocarbon monomers such as phenylethene and ethene form polymers used as plastics like polyphenylethene (commonly known as polystyrene) and polyethene (commonly known as polyethylene or polythene). Other commercially important monomers include acrylic monomers such as acrylic acid, methyl methacrylate, and acrylamide.

Amino acids are natural monomers, and polymerise to form proteins. Glucose monomers can also polymerise to form starches, amylopectins and glycogen polymers. In this case, the polymerisation reaction is known as a dehydration or condensation reaction (due to the formation of water (H_2O) as one of the products) where a hydrogen atom and a hydroxyl (–OH) group are lost to form H_2O and an oxygen molecule bonds between each monomer unit.

The lower molecular weight compounds built from monomers are also referred to as dimers, trimers, tetramers, quadramers, pentamers, octamers, 20-mers, etc. if they have 2, 3, 4, 5, 8, or 20 monomer units, respectively. Any number of these monomer units may be indicated by the appropriate prefix, e.g., *deca*mer, being a 10-unit monomer chain or polymer. Larger numbers are often stated in English in lieu of Greek. Polymers with relatively low number of units are called oligomers.

Significance of Amino Acids

In chemistry, an amino acid is a molecule that contains both amine and carboxyl functional groups. In biochemistry, this term refers to alpha-amino acids with the general formula $NH_2CHRCOOH$. These are molecules where the amino and carboxylate groups are attached to the same carbon, which is called the α–carbon.

The various alpha amino acids differ in which side chain (R group) is attached to their alpha carbon. This can vary in size from just a hydrogen atom in glycine, through a methyl group in alanine, to a large heterocyclic group in tryptophan.

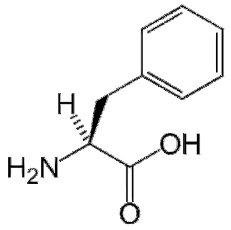


Fig. Phenylalanine is one of the standard amino acids.

Alpha-amino acids are the building blocks of proteins. A protein forms via the condensation of amino acids to form a chain of amino acid "residues" linked by peptide bonds. Each different protein has a unique sequence of amino acid residues; this sequence is the primary structure of the protein. Just as the letters of the alphabet can be combined to form an almost endless variety of words, amino acids can be linked in varying sequences to form a huge variety of proteins.

There are twenty standard amino acids used by cells in protein biosynthesis, and these are specified by the general genetic code. These twenty amino acids are biosynthesised from other molecules, but organisms differ in which ones they can synthesise and which ones must be provided in their diet. The ones that cannot be synthesised by an organism are called essential amino acids.

Functions in Proteins

Amino acids are the basic structural building units of proteins. They form short polymer chains called peptides or longer chains either called polypeptides or proteins. The process of such formation from an mRNA template is known as translation which is part of protein biosynthesis.

Twenty amino acids are encoded by the standard genetic code and are called proteinogenic or standard amino acids. Other amino acids contained in proteins are usually formed by post-translational modification, which is modification after translation in protein synthesis. These modifications are often essential for the function or regulation of a protein; for example, the carboxylation of glutamate allows for better binding of calcium cations, and the hydroxylation of proline is critical for maintaining connective tissues and responding to oxygen starvation. Such modifications can also determine the localisation of the protein, e.g., the addition of long hydrophobic groups can cause a protein to bind to a phospholipid membrane.

Non-protein Functions

The twenty standard amino acids are either used to synthesise proteins and other biomolecules, or oxidised to urea and carbon dioxide as a source of energy. The oxidation pathway starts with

the removal of the amino group by a transaminase, the amino group is then fed into the urea cycle. The other product of transamidation is a keto acid that enters the citric acid cycle. Glucogenic amino acids can also be converted into glucose, through gluconeogenesis.

Hundreds of types of non-protein amino acids have been found in nature and they have multiple functions in living organisms. Microorganisms and plants can produce uncommon amino acids. In microbes, examples include 2-aminoisobutyric acid and lanthionine, which is a sulfide-bridged alanine dimer. Both these amino acids are both found in peptidic lantibiotics such as alamethicin. While in plants, 1-Aminocyclopropane-1-carboxylic acid is a small disubstituted cyclic amino acid that is a key intermediate in the production of the plant hormone ethylene.

In humans, non-protein amino acids also have biologically-important roles. Glycine, gamma-aminobutyric acid and glutamate are neurotransmitters and many amino acids are used to synthesise other molecules, for example:

- Tryptophan is a precursor of the neurotransmitter serotonin,
- Glycine is a precursor of porphyrins such as haem,
- Arginine is a precursor of the hormone nitric oxide,
- Carnitine is used in lipid transport within a cell,
- Ornithine and S-adenosylmethionine are precursors of polyamines, and
- Homocysteine is an intermediate in S-adenosylmethionine recycling.

Also present are hydroxyproline, hydroxylysine, and sarcosine. The thyroid hormones are also alpha-amino acids.

Some amino acids have even been detected in meteorites, especially in a type known as carbonaceous chondrites. This observation has prompted the suggestion that life may have arrived on earth from an extraterrestrial source.

General Structure

In the structure shown to the right, the *R* represents a side chain specific to each amino acid. The central carbon atom called

C_α is a chiral central carbon atom (with the exception of glycine) to which the two termini and the R-group are attached. Amino acids are usually classified by the properties of the side chain into four groups. The side chain can make them behave like a weak acid, a weak base, a hydrophile if they are polar, and hydrophobe if they are non-polar. The chemical structures of the 20 standard amino acids, along with their chemical properties, are catalogued in the list of standard amino acids.

The phrase "branched-chain amino acids" or BCAA is sometimes used to refer to the amino acids having aliphatic side-chains that are non-linear, these are leucine, isoleucine and valine. Proline is the only proteinogenic amino acid whose side group links to the ±–amino group, and thus is also the only proteinogenic amino acid containing a secondary amine at this position. Proline has sometimes been termed an imino acid, but this is not correct in the current nomenclature.

Isomerism

Most amino acids can exist in either of two optical isomers, called D and L. The L-amino acids represent the vast majority of amino acids found in proteins. D-amino acids are found in some proteins produced by exotic sea-dwelling organisms, such as cone snails. They are also abundant components of the peptidoglycan cell walls of bacteria.

The L and D conventions for amino acid configuration do not refer to the optical activity, but rather to the optical activity of the isomer of glyceraldehyde having the same stereochemistry as the amino acid. *S*-Glyceraldehyde is levorotary, and *R*-glyceraldehyde is dexterorotary, and so *S*-amino acids are called L—even if they are not levorotary, and *R*-amino acids are likewise called D—even if they are not dexterorotary.

There are two exceptions to these general rules of amino acid isomerism. Firstly, glycine, where *R* = H, no isomerism is possible because the alpha-carbon bears two identical groups (hydrogen). Secondly, in cysteine, the L = *S* and D = *R* assignment is reversed to L = *R* and D = *S*. Cysteine is structured similarly (with respect to glyceraldehyde) to the other amino acids but the sulphur atom alters the interpretation of the Cahn-Ingold-Prelog priority rule.

Reactions

As amino acids have both a primary amine group and a primary carboxyl group, these chemicals can undergo most of the reactions associated with these functional groups. These include nucleophilic addition, amide bond formation and imine formation for the amine group and esterification, amide bond formation and decarboxylation for the carboxylic acid group. The multiple side chains of amino acids can also undergo chemical reactions.

Peptide Bond Formation

As both the amine and carboxylic acid groups of amino acids can react to form amide bonds, one amino acid molecule can react with another and become joined through an amide linkage. This polymerisation of amino acids is what creates proteins. This condensation reaction yields the newly formed peptide bond and a molecule of water. In cells, this reaction does not occur directly, instead the amino acid is activated by attachment to a transfer RNA molecule through an ester bond. This aminoacyl-tRNA is produced in an ATP-dependent reaction carried out by an aminoacyl tRNA synthetase.

This aminoacyl-tRNA is then a substrate for the ribosome, which catalyses the attack of the amino group of the elongating protein chain on the ester bond. As a result of this mechanism, all proteins are synthesised starting at their N-terminus and moving towards their C-terminus.

However, not all peptide bonds are formed in this way. In a few cases peptides are synthesised by specific enzymes. For example, the tripeptide glutathione is an essential part of the defences of cells against oxidative stress.

This peptide is synthesised in two steps from free amino acids. In the first step gamma-glutamylcysteine synthetase condenses cysteine and glutamic acid through a peptide bond formed between the side-chain carboxyl of the glutamate (the gamma carbon of this side chain) and the amino group of the cysteine. This dipeptide is then condensed with glycine by glutathione synthetase to form glutathione.

In chemistry, peptides are synthesised by a variety of reactions. One of the most used in solid-phase peptide synthesis, which uses the aromatic oxime derivatives of amino acids as activated units. These are added in sequence onto the growing peptide chain, which is attached to a solid resin support.

Zwitterions

As amino acids have both the active groups of an amine and a carboxylic acid they can be considered both acid and base (though their natural pH is usually influenced by the R group). At a certain pH known as the isoelectric point, the amine group gains a positive charge (is protonated) and the acid group a negative charge (is deprotonated).

The exact value is specific to each different amino acid. This ion is known as a zwitterion, which comes from the German word *Zwitter* meaning "hybrid". A zwitterion can be extracted from the solution as a white crystalline structure with a very high melting point, due to its dipolar nature. Near-neutral physiological pH allows most free amino acids to exist as zwitterions.

Hydrophilic and Hydrophobic Amino Acids

Depending on the polarity of the side chain, amino acids vary in their hydrophilic or hydrophobic character. These properties are important in protein structure and protein-protein interactions. The importance of the physical properties of the side chains comes from the influence this has on the amino acid residues' interactions with other structures, both within a single protein and between proteins.

The distribution of hydrophilic and hydrophobic amino acids determines the tertiary structure of the protein, and their physical location on the outside structure of the proteins influences their quaternary structure.

For example, soluble proteins have surfaces rich with polar amino acids like serine and threonine, while integral membrane proteins tend to have outer ring of hydrophobic amino acids that anchors them into the lipid bilayer, and proteins anchored to the membrane have a hydrophobic end that locks into the membrane. Similarly, proteins that have to bind to positively-charged

molecules have surfaces rich with negatively charged amino acids like glutamate and aspartate, while proteins binding to negatively-charged molecules have surfaces rich with positively charged chains like lysine and arginine. Recently a new scale of hydrophobicity based on the free energy of hydrophobic association has been proposed.

Hydrophilic and hydrophobic interactions of the proteins do not have to rely only on the sidechains of amino acids themselves. By various post-translational modifications other chains can be attached to the proteins, forming hydrophobic lipoproteins or hydrophilic glycoproteins.

Non-standard Amino Acids

Aside from the twenty standard

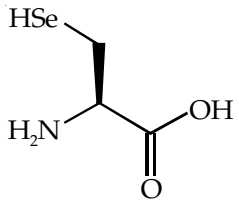


Fig. The Amino Acid Selenocysteine

Aside from the twenty standard amino acids, there are a vast number of "non-standard amino acids". Two of these can be encoded in the genetic code, but are rather rare in proteins. Selenocysteine is incorporated into some proteins at a UGA codon, which is normally a stop codon. Pyrrolysine is used by some methanogenic bacteria in enzymes that they use to produce methane. It is coded for with the codon UAG.

Examples of non-standard amino acids that are not found in proteins include lanthionine, 2-aminoisobutyric acid, dehydroalanine and the neurotransmitter gamma-aminobutyric acid. Non-standard amino acids often occur as intermediates in the metabolic pathways for standard amino acids - for example ornithine and citrulline occur in the urea cycle, part of amino acid catabolism.

Non-standard amino acids are usually formed through modifications to standard amino acids. For example, homocysteine

is formed through the transsulfuration pathway or by the demethylation of methionine via the intermediate metabolite S-adenosyl methionine, while dopamine is synthesised from tyrosine, and hydroxyproline is made by a post-translational modification of proline.

Nutritional Importance

Of the 20 standard proteinogenic amino acids, 10 are called essential amino acids because the human body cannot synthesise them from other compounds through chemical reactions, and they therefore, must be obtained from food. Cysteine, tyrosine, histidine and arginine are considered as semi-essential amino acids in children, because the metabolic pathways that synthesise these amino acids are not fully developed.

Essential	*Nonessential*
Isoleucine	Alanine
Leucine	Asparagine
Lysine	Aspartate
Methionine	Cysteine
Phenylalanine	Glutamate
Threonine	Glutamine
Tryptophan	Glycine
Valine	Proline
Arginine*	Serine
Histidine*	Tyrosine

(*) Essential only in certain cases

Several common mnemonics have evolved for remembering the essential amino acids. PVT TIM HALL ("Private Tim Hall") uses the first letter of each essential amino acid, including arginine. Another mnemonic that frequently occurs in student practice materials is "These ten valuable amino acids have long preserved life in man".

Forms of Biomolecules

A biomolecule is a chemical compound that naturally occurs in living organisms. Biomolecules consist primarily of carbon and

hydrogen, along with nitrogen, oxygen, phosphorus and sulphur. Other elements sometimes are incorporated but are much less common.

All known forms of life are comprised solely of biomolecules. For example, humans possess skin and hair. The main component of hair is keratin, an agglomeration of proteins which are themselves polymers built from amino acids.

Amino acids are some of the most important building blocks used, in nature, to construct larger molecules. Another type of building block is the nucleotides, each of which consists of three components: either a purine or pyrimidine base, a pentose sugar and a phosphate group. These nucleotides, mainly, form the nucleic acids.

Besides the polymeric biomolecules, numerous organic molecules are absorbed by living systems.

Types of Biomolecules

Lipid: Lipids are an amphiphilic class of hydrocarbon-containing organic compounds. Lipids are categorised by the fact that they have complicated solvation properties, giving rise to lipid polymorphism.

Lipid molecules have these properties because they consist largely of long hydrocarbon tails which are lipophilic in nature as well as polar head groups (e.g. phosphate-based functionality, and/or inositol based functionality). In living organisms, lipids are used for energy storage, serve as the structural components of cell membranes, and constitute important signalling molecules. Although the term *lipid* is often used as a synonym for fat, the latter is in fact a subgroup of lipids called triglycerides and should not be confused with the term fatty acid.

Structure

Fatty Acids and Glycerides: Chemically, fatty acids can be described as long-chain monocarboxylic acids the saturated examples of which have a general structure of $CH_3(CH_2)_nCOOH$. The length of the chain usually ranges from 12 to 24, always with an even number of carbon atoms. When the carbon chain contains

no double bonds, it is a saturated chain. If it contains one or more such bonds, it is unsaturated. The presence of double bonds reduces the melting point of fatty acids. Furthermore, unsaturated fatty acids can occur either in cis or trans geometric isomers. In naturally occurring fatty acids, the double bonds are in the cis-configuration.

Glycerides are lipids possessing a glycerol (a crude name for which is propan-1, 2, 3-triol) core structure with one or more fatty acyl groups, which are fatty acid-derived chains attached to the glycerol backbone by ester linkages. Glycerides with three acyl groups (triglycerides or neutral fats) are the main form of fatty energy storage in animals and plants.

An important type of glyceride-based molecule found in biological membranes, such as the cell's plasma membrane and the intracellular membranes of organelles, are the phosphoglycerides or glycerophospholipids. These are phospholipids that contain a glycerol core linked to two fatty acid-derived "tails" by ester or, more rarely, ether linkages and to one "head" group by a phosphate ester linkage.

The head groups of the phospholipids found in biological membranes are phosphatidylcholine (also known as PC, and lecithin), phosphatidylethanolamine (PE), phosphatidylserine and phosphatidylinositol (PI). These phospholipids are subject to a variety of functions in the cell: for instance, the lipophilic and polar ends can be released from particular phospholipids through enzyme-catalysed hydrolysis to generate secondary messengers involved in signal transduction.

In the case of phosphatidylinositol, the head group can be enzymatically modified by the addition of one, two or three phosphate groups, this constituting another mechanism of cell signalling. While phospholipids are the major component of biological membranes, other non-glyceride lipid components like sphingolipids and sterols (such as cholesterol in animal cell membranes) are also found in biological membranes.

A biological membrane is a form of lipid bilayer, as is a liposome. Formation of lipid bilayers is an energetically-favoured process when the glycerophospholipids described above are in an aqueous environment. In an aqueous system, the polar heads of

lipids orientate towards the polar, aqueous environment, while the hydrophobic tails minimise their contact with water.

The lipophilic tails of lipids (U) tend to cluster together, forming a lipid bilayer (1) or a micelle (2). Other aggregations are also observed and form part of the polymorphism of amphiphile behaviour. The polar heads (P) face the aqueous environment, curving away from the water. Phase behaviour is a complicated area within biophysics and is the subject of current academic research.

Micelles and bilayers form in the polar medium by a process known as the lipophilic effect. When dissolving a lipophilic or amphiphilic substance in a polar environment, the polar molecules (i.e. water in an aqueous solution) become more ordered around the dissolved lipophilic substance, since the polar molecules cannot form hydrogen bonds to the lipophilic areas of the amphiphile. So in an aqueous environment the water molecules form an ordered "clathrate" cage around the dissolved lipophilic molecule.

The self-organisation depends on the concentration of the lipid present in solution. Below the critical micelle concentration, the lipids form a single layer on the liquid surface and are (sparingly) dispersed in the solution. At the first critical micelle concentration (CMC-I), the lipids organise in spherical micelles, at given points above this concentration, other phases are observed.

Constraint Algorithm

In mechanics, a constraint algorithm is a method for satisfying constraints for bodies that obey Newton's equations of motion. There are three basic approaches to satisfying such constraints: choosing novel unconstrained coordinates ("internal coordinates"), introducing explicit constraint forces, and minimising constraint forces implicitly by the technique of Lagrange multipliers or projection methods.

Constraint algorithms are often applied to molecular dynamics simulations. Although such simulations are sometimes carried out in internal coordinates that automatically satisfy the bond-length and bond-angle constraints, they may also be carried out with explicit or implicit constraint forces for the bond lengths and

bond angles. Explicit constraint forces typically shorten the time-step significantly, making the simulation less efficient computationally; in other words, more computer power is required to compute a trajectory of a given length. Therefore, internal coordinates and implicit-force constraint solvers are generally preferred.

Periodic Wrapping

Instead of creating ghost cells, cell pairs that interact over a periodic boundary can also use a periodic correction vector $q_{\alpha\beta}$. This vector, which can be stored or computed for every cell pair (C_α, C_β) contains the correction which needs to be applied to "wrap" one cell around the domain to neighbour the other.

The pairwise distance between two particles $P_\alpha \in C_\alpha$ and $P_\beta \in C_\beta$ is then computed as:

$$r^2 = \left\| x[p_\alpha] - x[p_\beta] - q_{\alpha\beta} \right\|_2^2$$

This approach, although more efficient than using ghost cells, is less straightforward to implement (the cell pairs need to be identified over the periodic boundaries and the vector $q_{\alpha\beta}$ needs to be computed/stored).

Improvements

Despite reducing the computational cost of finding all pairs within a given cut-off distance from $O(N^2)$ to O(N), the cell list algorithm listed above still has some inefficiencies.

Consider a computational cell with edge length equal to the cut-off radius r_c. The pairwise distance between all particles in the cell and in one of the neighbouring cells is computed. The cell has 26 neighbours: 6 sharing a common face, 12 sharing a common edge and 8 sharing a common corner. Of all the pairwise distances computed, only about 16 per cent per cent will actually less or equal r_c. Otherwise put, 84 per cent of all pairwise distance computations are spurious.

One way of overcoming this inefficiency is to partition the domain into cells of edge length smaller than r_c. The pairwise

interactions are then not just computed between neighbouring cells, but between all cells within r_c of each other. This approach can be taken to the limit wherein each cell holds at most one single particle, therefore reducing the number of spurious pairwise distance evaluations to zero.

This gain in efficiency, however, is quickly offset by the number of cells C_β that need to be inspected for every interaction with a cell C_α, which grows cubically with the inverse of the cell edge length. Setting the edge length to $r_c/2$, however, already reduces the number of spurious distance evaluations to 63 per cent.

Another approach outlined in , where the particles are first sorted along the axis connecting the cell centres. This approach generates only about 40 per cent spurious pairwise distance computations, yet carries an additional cost due to sorting the particles.

Role of Glycolipids

Glycolipids are carbohydrate-attached lipids. Their role is to provide energy and also serve as markers for cellular recognition. They occur where a carbohydrate chain is associated with phospholipids on the exoplasmic surface of the cell membrane. The carbohydrates are found on the outer surface of all eukaryotic cell membranes.

They extend from the phospholipid bilayer into the aqueous environment outside the cell where it acts as a recognition site for specific chemicals as well as helping to maintain the stability of the membrane and attaching cells to one another to form tissues.

Types of Glycolipids

The following is an incomplete listing of glycolipid types.

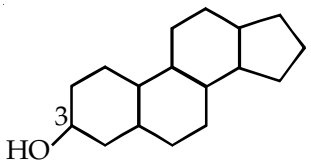
This list is incomplete; you can help by expanding it.

- Galactolipids
- Sulfolipids (SQDG)
- Glycosphingolipids

— Cerebrosides
  ⇒ Galactocerebrosides
  ⇒ Glucocerebrosides
— Globosides
— Gangliosides

Sterol

Sterols, or steroid alcohols are a subgroup of steroids with a hydroxyl group in the 3-position of the A-ring. They are amphipathic lipids synthetised from acetyl-coenzyme A. The overall molecule is quite flat. The hydroxyl group on the A ring is polar. The rest of the aliphatic chain is non-polar.



Sterols of plants are called phytosterols and sterols of animals are called zoosterols. The most important zoosterols are cholesterol and some steroid hormones; the most important phytosterols are campesterol, sitosterol, and stigmasterol.

Sterols play essential roles in the physiology of eukaryotic organisms. For example cholesterol forms part of the cellular membrane where its presence increases the cell membrane's fluidity and serves as secondary messenger in developmental signalling. Plant sterols are also known to block cholesterol absorption sites in the human intestine thus helping to reduce cholesterol in humans.

In humans sterols act to provide important signals and metabolic communications, e.g. circadium rhythms, blood clotting.

Vitamin

A vitamin is an organic compound required in tiny amounts for essential metabolic reactions in a living organism. The term *vitamin* does not include other essential nutrients such as dietary minerals, essential fatty acids, or essential amino acids, nor does it encompass the large number of other nutrients that promote health but that are not essential for life.

Vitamins are bio-molecules that act both as catalysts and substrates in chemical reactions. When acting as a catalyst, vitamins are bound to enzymes and are called cofactors. For example, vitamin K is part of the proteases involved in blood clotting. Vitamins also act as coenzymes to carry chemical groups between enzymes. For example, folic acid carries various forms of carbon group—methyl, formyl and methylene—in the cell.

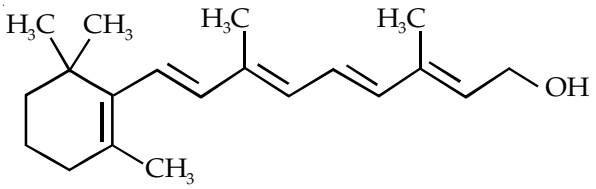


Fig. Retinol (Vitamin A)

Until the 1900s, vitamins were obtained solely through food intake, and changes in diet (which, for example, could occur during a particular growing season) can alter the types and amounts of vitamins ingested.

Vitamins have been produced as commodity chemicals and made widely available as inexpensive pills for several decades, allowing supplementation of the dietary intake.

History

The value of eating certain foods to maintain health was recognised long before vitamins were identified. The ancient Egyptians knew that feeding a patient liver would help cure night blindness, an illness now known to be caused by a vitamin A deficiency.

In 1747, the Scottish surgeon, James Lind discovered that citrus foods helped prevent scurvy, a particularly deadly disease in which collagen is not properly formed, and which is characterised by poor wound healing, bleeding of the gums, and severe pain. In 1753, Lind published his *Treatise on the Scurvy*, which recommended using lemons and limes to avoid scurvy, which was adopted by the British Royal Navy.

This led to the nickname, Limey for sailors of that organisation. Lind's discovery, however, was not widely accepted by individuals in the Royal Navy's Arctic expeditions in the 19th century, where

it was widely believed that scurvy could be prevented by practising good hygiene, regular exercise, and by maintaining the morale of the crew while on board, rather than by a diet of fresh food.

As a result, Arctic expeditions continued to be plagued by scurvy and other deficiency diseases. In the early 20th century, when Robert Falcon Scott made his two expeditions to the Antarctic, the prevailing medical theory was that scurvy was caused by "tainted" canned food.

In 1881, Russian surgeon Nikolai Lunin studied the effects of scurvy while at the University of Tartu in present-day Estonia. He fed mice an artificial mixture of all the separate constituents of milk known at that time, namely the proteins, fats, carbohydrates, and salts. The mice that received only the individual constituents died, while the mice fed by milk itself developed normally. He made a conclusion that "a natural food such as milk must therefore, contain, besides these known principal ingredients, small quantities of unknown substances essential to life". However, his conclusions were rejected by other researchers when they were unable to reproduce his results. One difference was that he had used table sugar (sucrose), while other researchers had used milk sugar (lactose) that still contained small amounts of vitamin B.

In 1897, Christiaan Eijkman discovered that eating unpolished rice instead of the polished variety helped to prevent the disease beriberi.

The following year, Frederick Hopkins postulated that some foods contained "accessory factors"—in addition to proteins, carbohydrates, fats, et cetera—that were necessary for the functions of the human body. Hopkins was awarded the 1929 Nobel Prize for Physiology or Medicine with Christiaan Eijkman for their discovery of several vitamins.

The Polish biochemist Kazimierz Funk was the first to isolate the water-soluble complex of micronutrients–the bioactivity of which Fletcher had identified–and proposed the complex be named "Vitamine" in 1912. The name soon became synonymous with Hopkins's "accessory factors", and by the time it was shown that not all vitamins were amines, the word was already ubiquitous. In 1920, Jack Cecil Drummond proposed that the final "e" be

dropped to deemphasise the "amine" reference after the discovery that vitamin C had no amine component.

Throughout the early 1900s, the use of deprivation studies allowed scientists to isolate and identify a number of vitamins. Initially, lipid from fish oil was used to cure rickets in rats, and the fat-soluble nutrient was called "antirachitic A". The irony here is that the first "vitamin" bioactivity ever isolated, which cured rickets, was initially called "vitamin A", the bioactivity of which is now called vitamin D, What we now call "vitamin A" was identified in fish oil because it was inactivated by ultraviolet light.

In 1931, Albert Szent-Györgyi and a fellow researcher Joseph Svirbely determined that "hexuronic acid" was actually vitamin C and noted its anti-scorbutic activity. In 1937, Szent-Györgyi was awarded the Nobel Prize for his discovery. In 1943 Edward Adelbert Doisy and Henrik Dam were awarded the Nobel Prize for their discovery of vitamin K and its chemical structure.

In Humans

Vitamins are classified as either water-soluble, meaning that they dissolve easily in water, or fat-soluble, and are absorbed through the intestinal tract with the help of lipids. Each vitamin is typically used in multiple reactions and, therefore, most have multiple functions.

Table: *In humans there are 13 vitamins: 4 fat-soluble (A, D, E and K) and 9 water-soluble (8 B vitamins and vitamin C).*

Vitamin name	*Chemical name*	*Solubility*	*Recommended dietary allowances (male, age 19–70)*	*Deficiency disease*	*Upper Intake Level (UL/day)*	*Overdose disease*
Vitamin A	Retinoids (retinol, retinoids and carotenoids)	Fat	900 µg	Night-blindness and Keratomalacia	3,000 µg	Hypervita-minosis A
Vitamin B_1	Thiamine	Water	1.2 mg	Beriberi	N/D	*?*
Vitamin B_2	Riboflavin	Water	1.3 mg	Ariboflavinosis	N/D	*?*
Vitamin B_3	Niacin	Water	16.0 mg	Pellagra	35.0 mg	"Niacin flush"
Vitamin B_5	Pantothenic acid	Water	5.0 mg	Paresthesia	N/D	*?*
Vitamin B_6	Pyridoxine	Water	1.3-1.7 mg	Anaemia	100 mg	Impairment of proprioceptor-tion

Contd..

Vitamin name	*Chemical name*	*Solubility*	*Recommended dietary allowances (male, age 19–70)*	*Deficiency disease*	*Upper Intake Level (UL/day)*	*Overdose disease*
Vitamin B_7	Biotin	Water	30.0 µg	None identified	N/D	?
Vitamin B_9	Folic acid	Water	400 µg	Deficiency during pregnancy is associated with birth defects, such as neural tube defects	1,000 µg	?
Vitamin B_{12}	Cyanoco-balamin	Water	2.4 µg	Megaloblastic anaemia	N/D	?
Vitamin C	Ascorbic acid	Water	90.0 mg	Scurvy	2,000 mg	Refer to Vitamin C megadosage
Vitamin D	Ergocalciferol and Cholecalciferol	Fat	5.0 µg-10 µg	Rickets and Osteomalacia	50 µg	Hypervita-minosis D
Vitamin E	Tocopherol and Tocotrienol	Fat	15.0 mg	Deficiency is very rare; mild hemolytic anaemia in newborn infants.	1,000 mg	?
Vitamin K	Naphtho-quinone	Fat	120 µg	Bleeding diathesis	N/D	?

In Nutrition and Disease

Vitamins are essential for the normal growth and development of a multicellular organism. Using the genetic blueprint inherited from its parents, a fetus begins to develop, at the moment of conception, from the nutrients it absorbs. It requires certain vitamins and minerals to be present at certain times. These nutrients facilitate the chemical reactions that produce among other things, skin, bone, and muscle. If there is serious deficiency in one or more of these nutrients, a child may develop a deficiency disease. Even minor deficiencies may cause permanent damage.

For the most part, vitamins are obtained with food, but a few are obtained by other means. For example, microorganisms in the intestine - commonly known as “gut flora” - produce vitamin K and biotin, while one form of vitamin D is synthesised in the skin with the help of natural ultraviolet in sunlight. Humans can produce some vitamins from precursors they consume. Examples include vitamin A, produced from beta carotene, and niacin, from the amino acid tryptophan.

Once growth and development are completed, vitamins remain essential nutrients for the healthy maintenance of the cells, tissues,

and organs that make up a multicellular organism; they also enable a multicellular life form to efficiently use chemical energy provided by food it eats, and to help process the proteins, carbohydrates, and fats required for respiration.

Deficiencies

Deficiencies of vitamins are classified as either primary or secondary. A primary deficiency occurs when an organism does not get enough of the vitamin in its food. A secondary deficiency may be due to an underlying disorder that prevents or limits the absorption or use of the vitamin, due to a "lifestyle factor", such as smoking, excessive alcohol consumption, or the use of medications that interfere with the absorption or use of the vitamin.

People who eat a varied diet are unlikely to develop a severe primary vitamin deficiency. In contrast, restrictive diets have the potential to cause prolonged vitamin deficits, which may result in often painful and potentially deadly diseases.

Because human bodies do not store most vitamins, humans must consume them regularly to avoid deficiency. Human bodily stores for different vitamins vary widely; vitamins A, D, and B_{12} are stored in significant amounts in the human body, mainly in the liver, and an adult human's diet may be deficient in vitamins A and B_{12} for many months before developing a deficiency condition. Vitamin B_3 is not stored in the human body in significant amounts, so stores may only last a couple of weeks.

Well-known human vitamin deficiencies involve thiamine (beriberi), niacin (pellagra), vitamin C (scurvy) and vitamin D (rickets). In much of the developed world, such deficiencies are rare; this is due to (1) an adequate supply of food; and (2) the addition of vitamins and minerals to common foods, often called fortification.

Side Effects and Overdose

In large doses some vitamins have documented side effects, that tend to be more severe with larger dosage. The likelihood of consuming too much of any vitamin from food is remote, but overdosing from vitamin supplementation does occur. At high

enough dosages some vitamins cause side effects such as nausea, diarrhoea, and vomiting. When side effects emerge, recovery is often accomplished by reducing the dosage. The concentrations of vitamins an individual can tolerate vary widely, and appear to be related to age and state of health. In the United States, overdose exposure to all formulations of vitamins was reported by 62,562 individuals in 2004 (nearly 80 per cent of these exposures were in children under the age of 6), leading to 53 "major" life-threatening outcomes and 3 deaths—a small number in comparison with the 19,250 people who died of unintentional poisoning of all kinds in the US in the same year (2004).

It is for these reasons that physicians and scientists carefully review the clinical data on supplement use in order to determine upper dosage thresholds for each vitamin that can be tolerated as a daily dose by the entire population without side effects. This dosage is known as the tolerable upper intake level (UL).

Supplements

Dietary supplements, often containing vitamins, are used to ensure that adequate amounts of nutrients are obtained on a daily basis, if optimal amounts of the nutrients cannot be obtained through a varied diet. Scientific evidence supporting the benefits of some dietary supplements is well established for certain health conditions, but others need further study.

In the United States, advertising for dietary supplements is required to include a disclaimer that the product is not intended to treat, diagnose, mitigate, prevent, or cure disease, and that any health claims have not been evaluated by the Food and Drug Administration. In some cases, dietary supplements may have unwanted effects, especially if taken before surgery, with other dietary supplements or medicines, or if the person taking them has certain health conditions. Vitamin supplements may also contain levels of vitamins many times higher, and in different forms, than one may ingest through food.

Intake of excessive quantities can cause vitamin poisoning, most commonly for iron, Vitamin A, and Vitamin D. For this reason, most common vitamins have recommended upper daily intake amounts.

Governmental Regulation of Vitamin Supplements

Most countries place dietary supplements in a special category under the general umbrella of "foods," not drugs. This necessitates that the manufacturer, and not the government, be responsible for ensuring that its dietary supplement products are safe before they are marketed. Unlike drug products, that must implicitly be proven safe and effective for their intended use before marketing, there are often no provisions to "approve" dietary supplements for safety or effectiveness before they reach the consumer. Also unlike drug products, manufacturers and distributors of dietary supplements are not generally required to report any claims of injuries or illnesses that may be related to the use of their products however, side effects have been reported for several types of vitamin supplements.

Names in Current and Previous Nomenclatures

The reason the set of vitamins seems to skip directly from E to K is that the vitamins corresponding to "letters" F-J were either reclassified over time, discarded as false leads, or renamed because of their relationship to "vitamin B", which became a "complex" of vitamins. The German-speaking scientists who isolated and described vitamin K (in addition to naming it as such) did so because the vitamin is intimately involved in the 'Koagulation' of blood following wounding. At the time, most (but not all) of the letters from F through I were already designated, so the use of the letter K was considered quite reasonable. The following table lists chemicals that had previously been classified as vitamins, as well as the earlier names of vitamins that later became part of the B-complex.

Previous name	*Chemical name*	*Reason for name change*
Vitamin B4	Adenine	DNA metabolite
Vitamin B8	Adenylic acid	DNA metabolite
Vitamin F	Essential fatty acids (doesnot fit the definition of a vitamin).	Needed in large quantities
Vitamin G	Riboflavin	Reclassified as Vitamin B2
Vitamin H	Biotin	Reclassified as Vitamin B7
Vitamin J	Catechol, Flavin	Protein metabolite

Contd..

Previous name	*Chemical name*	*Reason for name change*
Vitamin L_1	Anthranilic acid	Protein metabolite
Vitamin L_2	Adenylthiomethylpentose	RNA metabolite
Vitamin M	Folic acid	Reclassified as Vitamin B9
Vitamin O	Carnitine	Protein metabolite
Vitamin P	Flavonoids	No longer classified as a vitamin
Vitamin PP	Niacin	Reclassified as Vitamin B3
Vitamin U	S-Methylmethionine	Protein metabolite

Hormones

A hormone is a chemical messenger from one cell (or group of cells) to another. All multicellular organisms produce hormones.

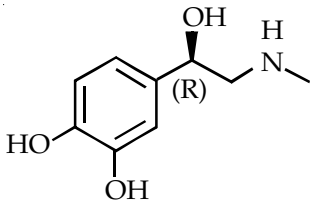


Fig. Epinephrine (adrenaline), a catecholamine-type hormone

The function of hormones is to serve as a signal to the target cells; the action of hormones is determined by the pattern of secretion and the signal transduction of the receiving tissue. The best-known animal hormones are those produced by endocrine glands of vertebrate animals, but hormones are produced by nearly every organ *system and* tissue type in a multicellular organism.

Endocrine hormone molecules are secreted (released) directly into the bloodstream, while exocrine hormones (or ectohormones) are secreted directly into a duct, and from the duct they either flow into the bloodstream or they flow from cell to cell by diffusion in a process known as paracrine signalling.

Hierarchical Nature of Hormonal Control

Hormonal regulation of some physiological activities involves a hierarchy of cell types acting on each other either to stimulate or modulate the release and action of a particular hormone. The

secretion of hormones from successive levels of endocrine cells is stimulated by chemical signals originating from cells higher up the hierarchical system.

The master coordinator of hormonal activity in mammals is the hypothalamus acting on input it receives from the central nervous system.

Other hormone secretion occurs in response to local conditions, such as the rate of secretion of parathyroid hormone by the parathyroid cells in response to fluctuations of ionised calcium levels in extracellular fluid.

Hormone Signalling

Hormonal signalling across this hierarchy involves the following:

1. Biosynthesis of a particular hormone in a particular tissue.
2. Storage and secretion of the hormone.
3. Transport of the hormone to the target cell(s).
4. Recognition of the hormone by an associated cell membrane or intracellular receptor protein.
5. Relay and amplification of the received hormonal signal via a signal transduction process. This then leads to a cellular response. The reaction of the target cells may then be recognised by the original hormone-producing cells, leading to a down-regulation in hormone production. This is an example of a homeostatic negative feedback loop.
6. Degradation of the hormone.

As can be inferred from the hierarchical diagram, hormone biosynthetic cells are typically of a specialised cell type, residing within a particular endocrine gland (e.g. the thyroid gland, ovaries or testes). Hormones may exit their cell of origin via exocytosis or another means of membrane transport. However, the hierarchical model is an over simplification of the hormonal signalling process. Typically cellular recipients of a particular hormonal signal may be one of several cell types that reside within a number of different tissues, as is the case for insulin,

which triggers a diverse range of systemic physiological effects. Different tissue types may also respond differently to the same hormonal signal. Because of this, hormonal signalling is elaborate and hard to dissect.

Interactions with Receptors

Most hormones initiate a cellular response by initially combining with either a specific intracellular or cell membrane associated receptor protein.

A cell may have several different receptors that recognise the same hormone and activate different signal transduction pathways, or alternatively different hormones and their receptors may invoke the same biochemical pathway.

For many hormones, including most protein hormones, the receptor is membrane associated and embedded in the plasma membrane at the surface of the cell.

The interaction of hormone and receptor typically triggers a cascade of secondary effects within the cytoplasm of the cell, often involving phosphorylation or dephosphorylation of various other cytoplasmic proteins, changes in ion channel permeability, or increased concentrations of intracellular molecules that may act as secondary messengers (e.g. cyclic AMP). Some protein hormones also interact with intracellular receptors located in the cytoplasm or nucleus by an intracrine mechanism.

For hormones such as steroid or thyroid hormones, their receptors are located intracellularly within the cytoplasm of their target cell. In order to bind their receptors these hormones must cross the cell membrane.

The combined hormone-receptor complex then moves across the nuclear membrane into the nucleus of the cell, where it binds to specific DNA sequences, effectively amplifying or suppressing the action of certain genes, and affecting protein synthesis. However, it has been shown that not all steriod receptors are located intracellularly, some are plasma membrane associated.

An important consideration, dictating the level at which cellular signal transduction pathways are activated in response to

a hormonal signal is the effective concentration of hormone-receptor complexes that are formed.

Hormone-receptor complex concentrations are effectively determined by three factors:

1. The number of hormone molecules available for complex formation,
2. The number of receptor molecules available for complex formation, and
3. The binding affinity between hormone and receptor.

The number of hormone molecules available for complex formation is usually the key factor in determining the level at which signal transduction pathways are activated. The number of hormone molecules available being determined by the concentration of circulating hormone, which is in turn influenced by the level and rate at which they are secreted by biosynthetic cells. The number of receptors at the cell surface of the receiving cell can also be varied as can the affinity between the hormone and its receptor.

Physiology of Hormones

Most cells are capable of producing one or more molecules, which act as signalling molecules to other cells, altering their growth, function, or metabolism. The classical hormones produced by endocrine glands mentioned so far in this article are cellular products, specialised to serve as regulators at the overall organism level. However, they may also exert their effects solely within the tissue in which they are produced and originally released.

The rate of hormone biosynthesis and secretion is often regulated by a homeostatic negative feedback control mechanism. Such a mechanism depends on factors which influence the metabolism and excretion of hormones. Thus, higher hormome concentration alone can not trigger the negative feedback mechanism. Negative feedback must be triggered by overproduction of an "effect" of the hormone.

Hormone secretion can be stimulated and inhibited by:

- Other hormones (*stimulating*- or *releasing*-hormones),
- Plasma concentrations of ions or nutrients, as well as binding globulins,

- Neurons and mental activity, and
- Environmental changes, e.g., of light or temperature.

One special group of hormones is the trophic hormones that stimulate the hormone production of other endocrine glands. For example, thyroid-stimulating hormone (TSH) causes growth and increased activity of another endocrine gland, the thyroid, which increases output of thyroid hormones.

A recently-identified class of hormones is that of the "hunger hormones"—ghrelin, orexin and PYY 3-36 - and "satiety hormones"—e.g., leptin, obestatin, nesfatin-1.

In order to release active hormones quickly into the circulation, hormone biosynthetic cells may produce and store biologically inactive hormones in the form of pre- or prohormones. These can then be quickly converted into their active hormone form in response to a particular stimulus.

Hormone Effects

Hormone effects vary widely, but can include:

- Stimulation or inhibition of growth,
- Induction or suppression of apoptosis (programmed cell death),
- Activation or inhibition of the immune system,
- Regulating metabolism,
- Preparation for a new activity (e.g., fighting, fleeing, mating),
- Preparation for a new phase of life (e.g., puberty, caring for offspring, menopause), and
- Controlling the reproductive cycle.

In many cases, one hormone may regulate the production and release of other hormones. Many of the responses to hormone signals can be described as serving to regulate metabolic activity of an organ or tissue.

Chemical Classes of Hormones

Vertebrate hormones fall into three chemical classes:

- Amine-derived hormones are derivatives of the amino acids tyrosine and tryptophan. Examples are catecholamines and thyroxine.

- Peptide hormones consist of chains of amino acids. Examples of small peptide hormones are TRH and vasopressin. Peptides composed of scores or hundreds of amino acids are referred to as proteins. Examples of protein hormones include insulin and growth hormone.

  More complex protein hormones bear carbohydrate side chains and are called glycoprotein hormones. Luteinising hormone, follicle-stimulating hormone and thyroid-stimulating hormone are glycoprotein hormones.
- Lipid and phospholipid-derived hormones derive from lipids such as linoleic acid and arachidonic acid and phospholipids. The main classes are the steroid hormones that derive from cholesterol and the eicosanoids. Examples of steroid hormones are testosterone and cortisol. Sterol hormones such as calcitriol are a homologous system.

  The adrenal cortex and the gonads are primary sources of steroid hormones. Examples of eicosanoids are the widely studied prostaglandins.

Pharmacology

Many hormones and their analogues are used as medication. The most commonly-prescribed hormones are estrogens and progestagens (as methods of hormonal contraception and as HRT), thyroxine (as levothyroxine, for hypothyroidism) and steroids (for autoimmune diseases and several respiratory disorders). Insulin is used by many diabetics. Local preparations for use in otolaryngology often contain pharmacologic equivalents of adrenaline, while steroid and vitamin D creams are used extensively in dermatological practice.

A "pharmacologic dose" of a hormone is a medical usage referring to an amount of a hormone far greater than naturally occurs in a healthy body. The effects of pharmacologic doses of hormones may be different from responses to naturally-occurring amounts and may be therapeutically useful. An example is the ability of pharmacologic doses of glucocorticoid to suppress inflammation.

Important Human Hormones

Spelling is not uniform for many hormones. Current North American and international usage is estrogen, gonadotropin, while British usage retains the Greek diphthong in oestrogen and the unvoiced aspirant h in gonadotrophin.

Structure	*Name*	*Abbreviation*	*Tissue*	*Cells*	*Mechanism*
amine-tryptophan	Melatonin (N-acetyl-5-methoxytryptamine)		pineal gland	pinealocyte	
amine-tryptophan	Serotonin	5-HT	CNS, GI tract	enterochromaffin cell	
amine-tyrosine	Thyroxine (thyroid hormone)	T4	thyroid gland	thyroid epithelial cell	direct
amine-tyrosine	Triiodothyronine (thyroid hormone)	T3	thyroid gland cell	thyroid epithelial	direct
amine - tyrosine (cat)	Epinephrine (or adrenaline)	EPI	adrenal medulla	chromaffin cell	
amine-tyrosine (cat)	Norepinephrine (or noradrenaline)	NRE	adrenal medulla	chromaffin cell	
amine-tyrosine (cat)	Dopamine	DPM	hypothalamus		
peptide	Antimullerian hormone (or mullerian inhibiting factor or hormone)	AMH	testes	Sertoli cell	
peptide	Adiponectin	Acrp30	adipose tissue		
peptide	Adrenocorticotropic hormone (or corticotropin)	ACTH	anterior pituitary	corticotrope	cAMP
peptide	Angiotensinogen and angiotensin	AGT	liver		IP3
peptide	Antidiuretic hormone (or vasopressin, arginine vasopressin)	ADH	posterior pituitary		varies
peptide	Atrial-natriuretic peptide (or atriopeptin)	ANP	heart		cGMP
peptide	Calcitonin	CT	thyroid gland	parafollicular cell	cAMP
peptide	Cholecystokinin	CCK	duodenum		
peptide	Corticotropin-releasing hormone	CRH	hypothalamus		cAMP
peptide	Erythropoietin	EPO	kidney		
peptide	Follicle-stimulating hormone	FSH	anterior pituitary	gonadotrope	cAMP
peptide	Gastrin	GRP	stomach, duodenum	G cell	
peptide	Ghrelin		stomach	P/D1 cell	
peptide	Glucagon	GCG	pancreas	alpha cells	cAMP
peptide	Gonadotropin-releasing hormone	GnRH	hypothalamus		IP3
peptide	Growth hormone -releasing hormone	GHRH	hypothalamus		IP3
peptide gonadotropin	Human chorionic	hCG	placenta phoblast cells	syncytiotro-	cAMP
peptide	Human placental lactogen	HPL	placenta		
peptide	Growth hormone	GH or hGH	anterior pituitary	somatotropes	

Contd...

Structure	*Name*	*Abbreviation*	*Tissue*	*Cells*	*Mechanism*
peptide	Inhibin		testes	Sertoli cells	
peptide	Insulin	INS	pancreas	beta cells	tyrosine kinase
peptide	Insulin-like growth factor (or somato-medin)	IGF	liver		tyrosine kinase
peptide	Leptin	LEP	adipose tissue		
peptide	Luteinising hormone	LH	anterior pituitary	gonadotropes	cAMP
peptide	Melanocyte stimulating hormone	MSH or α-MSH	anterior pituitary/ pars intermedia		cAMP
peptide	Oxytocin	OXT	posterior pituitary		IP3
peptide	Parathyroid hormone	PTH gland	parathyroid chief cell	parathyroid	cAMP
peptide	Prolactin	PRL	anterior pituitary	lactotrophs	
peptide	Relaxin	RLN	varies		
peptide	Secretin	SCT	duodenum	S cell	
peptide	Somatostatin	SRIF	hypothalamus, islets of Langerhans	delta cells	
peptide	Thrombopoietin	TPO	liver, kidney		
peptide hormone	Thyroid-stimulating	TSH pituitary	anterior	thyrotropes	cAMP
peptide	Thyrotropin-releasing hormone	TRH	hypothalamus		IP3
steroid-glu.	Cortisol		adrenal cortex (zona fasciculata)		direct
steroid-min.	Aldosterone		adrenal cortex (zona glomerulosa)		direct
steroid-sex (and)	Testosterone		testes	Leydig cells	direct
steroid-sex (and)	Dehydroepiandrosterone	DHEA	multiple		direct
steroid-sex (and)	Androstenedione		adrenal glands, gonads		direct
steroid-sex (and)	Dihydrotestosterone	DHT	multiple		direct
steroid-sex (est)	Estradiol	E2	ovary	granulosa cells	direct
steroid-sex (est)	Estrone		ovary	granulosa cells	direct
steroid-sex (est)	Estriol		placenta	syncytiotro-phoblast	direct
steroid-sex (pro)	Progesterone		ovary, adrenal glands, placenta	granulosa cells	direct
sterol	Calcitriol (Vitamin D3)		skin/proximal tubule of kidneys		direct
eicosanoid	Prostaglandins	PG	seminal vesicle		
eicosanoid	Leukotrienes	LT		white blood cells	
eicosanoid	Prostacyclin	PGI2	endothelium		
eicosanoid	Thromboxane	TXA2		platelets	

Role of Nutrition and Health

Lipids play diverse and important roles in nutrition and health. Many lipids are absolutely essential for life, however, there is also

considerable awareness that abnormal levels of certain lipids, particularly cholesterol (in hypercholesterolemia) and, more recently, fatty acids with trans fatty acids, are risk factors for heart disease amongst others. We need fats in our bodies and certain types in our diet. Animals in general use fat for energy storage because fat stores 9 KCal/g of energy. Plants, which do not require energy for movement, can afford to store food for energy in a less compact but more easily accessible form, so they have evolved use starch (a carbohydrate, not a lipid) for energy storage.

Carbohydrates and proteins store only 4 KCal/g of energy, so fat stores over twice as much energy/gram as other sources of energy.

Properties of Phospholipids

Phospholipids are a class of lipids, and a major component of all biological membranes, along with glycolipids, cholesterol and proteins. Understanding of the aggregation properties of these molecules is known as lipid polymorphism and forms part of current academic research.

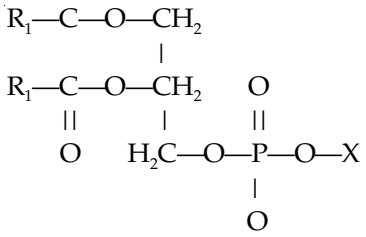


Fig. Phospholipids

Components

They are built upon to a nitrogen containing alcohol like ethanolamine or an organic compound such as choline.

Types

Phosphoglycerides: In phosphoglycerides, the carboxyl group of each fatty acid is esterified to the hydroxyl groups on carbon-1 and carbon-2 of the glycerol molecule. The phosphate group is attached to carbon-3 by an ester link. This molecule, known as a phosphatidate, is present in small quantities in membranes, but is also a precursor for the other phosphoglycerides.

In phosphoglyceride synthesis, phosphatidates must be activated first. Phospholipids can be formed from an activated diacylglycerol or an activated alcohol.

- Phosphatidyl serine and phosphatidyl inositol are formed from a phosphoester linkage between the hydroxyl of an alcohol (serine or inositol) and cytidine diphosphodiacylglycerol (CDP-diacylglycerol).
- In the synthesis of phospatidyl ethanolamine, the alcohol is phosphorylated by ATP first, and subsequently reacts with cytidine triphosphate (CTP) to form the activated alcohol (CDP-ethanolamine). The alcohol then reacts with a diacylglycerol to form the final product.
- In mammals, phosphatidyl choline can be synthesised via two separate pathways; a series of reactions similar to phosphatidyl ethanolamine synthesis, and the methylation of phosphatidyl ethanolamine, which is catalysed by phosphatidyl ethanolamine methyltransferase, an enzyme produced in the liver.

Formation and Character of Sphingomyelin

The backbone of sphingomyelin is sphingosine, an amino alcohol formed from palmitate and serine. The amino terminal is acylated with a long-chain acyl CoA to yield ceramide. Subsequent substitution of the terminal hydroxyl group by phosphatidyl choline forms sphingomyelin.

Sphingomyelin is also present in all eukaryotic cell membranes, especially the plasma membrane, and is particularly concentrated in the nervous system because sphingomyelin is a major component of myelin, the fatty insulation wrapped around nerve cells by Schwann cells or oligodendrocytes. Multiple sclerosis is a disease characterised by deterioration of the myelin sheath, leading to impairment of nervous conduction.

Amphipathic Character: Due to its polar nature, the head of a phospholipid is hydrophilic (attracted to water); the lipophilic tails are not attracted to water. When placed in water, phospholipids form one of a number of lipid phases. In biological systems this is restricted to bilayers, in which the lipophilic tails

line up against one another, forming a membrane with hydrophilic heads on both sides facing the water. This allows it to form liposomes spontaneously, or small lipid vesicles, which can then be used to transport materials into living organisms and study diffusion rates into or out of a cell membrane.

This membrane is partially permeable, capable of elastic movement, and has fluid properties, in which embedded proteins and phospholipid molecules are able to move laterally in it. Such movement can be described by the Fluid Mosaic Model, that describes the membrane as a mosaic of lipid molecules that act as a solvent for all the substances and proteins within it, so proteins and lipid molecules are then free to diffuse laterally through the lipid matrix and migrate over the membrane. However, this model has now been superseded, as through the study of lipid polymorphism it is now known that the behaviour of lipids under physiological (and other) conditions is not simple.

Supramolecular Assembly

A supramolecular assembly or "supermolecule" is a well defined complex of molecules held together by non-covalent bonds. While a supramolecular assembly can be simply composed of two molecules (e.g., a DNA double helix or an inclusion compound), it is more often used to denote larger complexes of molecules that form sphere-, rod-, or sheet-like species. The dimensions of supramolecular assemblies can range from nanometres to micrometers.

Thus they allow access to nanoscale objects using a bottom-up approach in much fewer steps than a single molecule of similar dimensions.

The process by which a supramolecular assembly forms is called molecular self-assembly. Some try to distinguish self-assembly as the process by which individual molecules form the defined aggregate.

Self-organisation, then, is the process by which those aggregates create higher-order structures. This can become useful when talking about liquid crystals and block copolymers.

Rotaxane

A rotaxane is a mechanically-interlocked molecular architecture consisting of a dumbbell-shaped molecule that is threaded through a macrocycle or ring-like molecule. The two components are kinetically trapped as the two end-groups of the dumbbell (often called stoppers) are larger than the internal diameter of the ring, and thus prevent dissociation (unthreading) since this would require significant distortion of the covalent bonds. The name, rotaxane, is derived from the Latin for wheel (rota) and axle (axis).

They are conceptually related to other mechanically-interlocked molecular architectures such as catenanes, molecular knots or borromean ring. The synthesis of such entangled architectures has been made efficient through the combination of supramolecular chemistry with traditional covalent synthesis, however mechanically-interlocked molecular architectures have properties that differ from both "supramolecular assemblies" and "covalently-bonded molecules". Recently the terminology "mechanical bond" has been coined to describe the connection between the two components of rotaxanes. Although mechanically-interlocked molecular architectures, such as rotaxanes, are an emerging area of research many examples have been found in biological systems including: cystine knots, cyclotides or lasso-peptides such as microcin J25 are protein, and a variety of peptides with rotaxane substructure.

Accepted nomenclature is to designate the number of components of the rotaxane in brackets as a prefix. Therefore the cartoon rotaxane displayed to the right would be a rotaxane as it consists of a single dumbbell and a single macrocycle.

Synthesis

Rotaxanes are most commonly constructed by pre-organising or threading the parts utilising hydrogen bonding, metal coordination, hydrophobic forces, covalent bonds, or coulombic interactions. Examples are crown ethers with a wide variety of structures, cyclodextrins with molecular wires and dyes and a rotaxane based on cucurbituril and hexamethylene diamine. Rotaxanes can also be synthesised by a clipping mechanism.

Potential Applications

Rotaxane-based molecular machines have been of initial interest for their potential use in molecular electronics as logic switching elements. These molecular machines are usually based on the movement of macrocycle on the dumbbell. The macrocycle can rotate around the axis of the dumbbell like a wheel and axle or it can slide along its axis from one site to another. Controlling the position of the macrocycle allows the rotaxane to function as molecular switch with each possible location of the macrocycle corresponding to a different state. Such systems have also been demonstrated as molecular muscles.

Potential application as long lasting dyes is based on the enhanced stability of the inner portion of the dumbbell shaped molecule. Studies with cyclodextrin protected rotaxane azo dyes established this characteristic. More reactive squaraine dyes have also been shown to have enhanced stability by preventing nucleophilic attack of the inner squaraine moiety. The enhanced stabilities of rotaxane dyes is attributed to the insulating effect of the macrocycle which is able to block interactions with other molecules.

In a nanorecording application a certain rotaxane is deposited as a Langmuir-Blodgett film on ITO-coated glass. When a positive voltage is applied with the tip of a scanning tunnelling microscope probe, the rotaxane rings in the tip area switch to a different part of the dumbbell and the resulting new conformation makes the molecules stick out from the surface by 0.3 nanometre and this height difference turns out to be sufficient for a memory dot. It is not yet possible to erase such a nanorecording film.

Catenane

A catenane is a mechanically-interlocked molecular architecture consisting of two or more interlocked macrocycles. The interlocked rings cannot be separated without breaking the covalent bonds of the macrocycles. Catenane is derived from the Latin *catena* meaning "chain". They are conceptually related to other mechanically-interlocked molecular architectures, such as rotaxanes, molecular knots or molecular Borromean rings. Recently

the terminology "mechanical bond" has been coined that describes the connection between the macrocycles of a catenane.

Synthesis

There are two primary approaches to the organic synthesis of catenanes. The first is to simply perform a ring closing reaction with the hope that some of the rings will form around other rings giving the desired catenane product. This so-called "statistical approach" led to the first successful synthesis of a catenane; however, the method is highly inefficient and is not usually used.

The second approach relies on supramolecular pre-organisation of the macrocyclic precursors utilising hydrogen bonding, metal coordination, hydrophobic forces, or coulombic interactions. These non-covalent interactions offset some of the entropic cost of association and help position the components to form the desired catenane upon the final ring-closing. This "template-directed" method has dramatically increased the yields that can be obtained for catenanes and thus has increased their potential for application. An example of this approach used bis-bipyridinium salts which form strong complexes threaded through crown ether bis (para-phenylene)-34-crown-10.

Families of Catenanes

There are a number of distinct methods of holding the precursors together prior to the ultimate ring-closing reaction in a template-directed catenane synthesis. Each non-covalent approach to catenane formation results in what can be considered different families of catenanes.

Another family of catenanes are called pretzelanes or bridged catenanes after their likeness to pretzels with a spacer linking the two macrocycles. In one such system one macrocycle is an electron deficient oligo Bis-bipyridinium ring and the other cycle is crown ether cyclophane based on paraphenylene or naphthalene. X-ray diffraction shows that due to pi-pi interactions the aromatic group of the cyclophane is held firmly inside the pyridinium ring. A limited number of (rapidly-interchanging) conformers exist for this type of compound.

In handcuff-shaped catenanes, two connected rings are threaded through the same ring. The bis-macrocycle (red) contains two phenanthroline units in a crown ether chain. The interlocking ring is self-assembled when two more phenanthroline units with alkene arms coordinate through a copper(I) complex followed by a metathesis ring closing step.

Families of Catenanes

Catenanes Pretzelanes Handcuff-shaped catenane

Applications

Supramolecular assemblies are being investigated as new materials in a variety of contexts. For instance, Samuel Stupp and co-workers at Northwestern University showed that a supramolecular assembly of peptide amphiphiles in the form of nanofibres could be used to promote the growth of neurons.

A great advantage to this supramolecular approach is that the nanofibres will degrade back into the individual peptide molecules that can be broken down by the body.

Another example with implications at the biology/materials science interface is of self-assembling dendritic dipeptides, which form hollow cylindrical supramolecular assemblies in solution and in bulk. The cylindrical assemblies possess internal helical order and self-organise into columnar liquid crystalline lattices. When inserted into vesicular membranes, the porous cylindrical assemblies mediate transport of protons across the membrane.

Self-assembling dendrons have also been used to generate arrays of nanowires. Electron donor-acceptor complexes comprise the core of the cylindrical supramolecular assemblies, which further self-organise into two-dimensional columnar liquid crystaline lattices. Each cylindrical supramolecular assembly functions as an individual wire. High charge carrier mobilities for holes and electrons were obtained.

Molecular Dynamics

Molecular dynamics (MD) is a form of computer simulation, wherein atoms and molecules are allowed to interact for a period of time under known laws of physics. Because molecular systems generally consist of a vast number of particles, it is impossible to find the properties of such complex systems analytically; MD simulation circumvents this problem by using numerical methods. It represents an interface between laboratory experiments and theory, and can be understood as a "virtual experiment".

Although we have known since Boltzmann's discoveries in the 19th century that matter consists of interacting particles in motion, many people still think of molecules as rigid museum models. Richard Feynman said in 1963 that "everything that living things do can be understood in terms of the jiggling and wiggling of atoms." One of MD's key contributions is creating awareness that molecules like proteins and DNA are machines in motion. MD probes the relationship between molecular structure, movement and function.

Molecular dynamics is a multidisciplinary field. Its laws and theories stem from mathematics, physics, and chemistry, and it employs algorithms from computer science and information theory.

It was originally conceived within theoretical physics in the late 1950's, but is applied today mostly in materials science and biomolecules.

Before it became possible to simulate molecular dynamics with computers, some undertook the hard work of trying it with physical models such as macroscopic spheres.

The idea was to arrange them to replicate the properties of a liquid. J.D. Bernal said, in 1962: "... I took a number of rubber balls and stuck them together with rods of a selection of different lengths ranging from 2.75 to 4 inch. I tried to do this in the first place as casually as possible, working in my own office, being interrupted every five minutes or so and not remembering what I had done before the interruption." Fortunately, now computers keep track of bonds during a simulation.

Molecular dynamics is a specialised discipline of molecular modelling and computer simulation based on statistical mechanics;

the main justification of the MD method is that statistical ensemble averages are equal to time averages of the system, known as the ergodic hypothesis.

MD has also been termed "statistical mechanics by numbers" and "Laplace's vision of Newtonian mechanics" of predicting the future by animating nature's forces and allowing insight into molecular motion on an atomic scale. However, long MD simulations are mathematically ill-conditioned, generating cumulative errors in numerical integration that can be minimised with proper selection of algorithms and parameters, but not eliminated entirely.

Furthermore, current potential functions are, in many cases, not sufficiently accurate to reproduce the dynamics of molecular systems. Nevertheless, molecular dynamics techniques allow detailed time and space resolution into representative behaviour in phase space.

Areas of Application

There is a significant difference between the focus and methods used by chemists and physicists, and this is reflected in differences in the jargon used by the different fields. In chemistry and biophysics, the interaction between the particles is either described by a "force field" (classical MD), a quantum chemical model, or a mix between the two. These terms are not used in physics, where the interactions are usually described by the name of the theory or approximation being used and called the potential energy, or just "potential".

Beginning in theoretical physics, the method of MD gained popularity in materials science and since the 1970s also in biochemistry and biophysics. In chemistry, MD serves as an important tool in protein structure determination and refinement using experimental tools such as X-ray crystallography and NMR.

It has also been applied with limited success as a method of refining protein structure predictions. In physics, MD is used to examine the dynamics of atomic-level phenomena that cannot be observed directly, such as thin film growth and ion-subplantation. It is also used to examine the physical properties of nanotechnological devices that have not or cannot yet be created.

In applied mathematics and theoretical physics, molecular dynamics is a part of the research realm of dynamical systems, ergodic theory and statistical mechanics in general. The concepts of energy conservation and molecular entropy come from thermodynamics. Some techniques to calculate conformational entropy such as principal components analysis come from information theory. Mathematical techniques such as the transfer operator become applicable when MD is seen as a Markov chain. Also, there is a large community of mathematicians working on volume preserving, symplectic integrators for more computationally efficient MD simulations.

MD can also be seen as a special case of the discrete element method (DEM) in which the particles have spherical shape (e.g. with the size of their van der Waals radii.) Some authors in the DEM community employ the term MD rather loosely, even when their simulations do not model actual molecules.

Design Constraints

Design of a molecular dynamics simulation should account for the available computational power. Simulation size (n = number of particles), timestep and total time duration must be selected so that the calculation can finish within a reasonable time period. However, the simulations should be long enough to be relevant to the time scales of the natural processes being studied. Most scientific publications about the dynamics of proteins and DNA use data from simulations spanning nanoseconds to microseconds. To obtain these simulations, several CPU-days to CPU-years are needed. Parallel algorithms allow the load to be distributed among CPUs; an example is the spatial decomposition in LAMMPS.

During a classical MD simulation, the most CPU intensive task is the evaluation of the potential (force field) as a function of the particles' internal coordinates. Within that energy evaluation, the most expensive one is the non-bonded or non-covalent part. In Big O notation, common molecular dynamics simulations scale by $O(n^2)$ if all pair-wise electrostatic and van der Waals interactions must be accounted for explicitly. This computational cost can be reduced by employing electrostatics methods such as Particle Mesh Ewald ($O(n\log(n))$) or good spherical cutoff techniques ($O(n)$).

Another factor that impacts total CPU time required by a simulation is the size of the integration timestep. This is the time length between evaluations of the potential.

The timestep must be chosen small enough to avoid discretisation errors (i.e. smaller than the fastest vibrational frequency in the system). Typical timesteps for classical MD are in the order of 1 femtosecond (1E-15 s). This value may be extended by using algorithms such as SHAKE and LINCS, which fix the fastest atoms (e.g. hydrogens) into place. Multiple time scale methods have also been developed, which allow for extended times between updates of slower long-range forces.

For simulating molecules in a solvent, a choice should be made between explicit solvent and implicit solvent. Explicit solvent particles (such as the TIP3P and SPC/E water models) must be calculated expensively by the force field, while implicit solvents use a mean-field approach.

The impact of explicit solvents on CPU-time can be 10-fold or more. But the granularity and viscosity of explicit solvent is essential to reproduce certain properties of the solute molecules.

In all kinds of molecular dynamics simulations, the simulation box size must be large enough to avoid boundary condition artifacts. Boundary conditions are often treated by choosing fixed values at the edges, or by employing periodic boundary conditions in which one side of the simulation loops back to the opposite side, mimicking a bulk phase.

Physical Principles

Microcanonical Ensemble (*NVE*)

In the microcanonical, or NVE ensemble, the system is isolated from changes in moles (N), volume (V) and energy (E). It corresponds to an adiabatic process with no heat exchange. A microcanonical molecular dynamics trajectory may be seen as an exchange of potential and kinetic energy, with total energy being conserved. For a system of N particles with coordinates X and velocities V, the following pair of first order differential equations may be written in Newton's notation as

$$F(X) = -\nabla U(X) = MV(t)$$

$$V(t) = X(t)$$

The potential energy function $U(X)$ of the system is a function of the particle coordinates X. It is referred to simply as the "potential" in Physics, or the "force field" in Chemistry. The first equation comes from Newton's laws; the force F acting on each particle in the system can be calculated as the negative gradient of $U(X)$.

For every timestep, each particle's position X and velocity V may be integrated with a symplectic method such as Verlet. The time evolution of X and V is called a trajectory. Given the initial positions (e.g. from theoretical knowledge) and velocities (e.g. randomised Gaussian), we can calculate all future (or past) positions and velocities.

One frequent source of confusion is the meaning of temperature in MD. Commonly we have experience with macroscopic temperatures, which involve a huge number of particles. But temperature is a statistical quantity. If there is a large enough number of atoms, statistical temperature can be estimated from the *instantaneous temperature*, which is found by equating the kinetic energy of the system to $nk_BT/2$ where n is the number of degrees of freedom of the system.

A temperature-related phenomenon arises due to the small number of atoms that are used in MD simulations. For example, consider simulating the growth of a copper film starting with a substrate containing 500 atoms and a deposition energy of 100 eV. In the real world, the 100 eV from the deposited atom would rapidly be transported through and shared among a large number of atoms (10^{10} or more) with no big change in temperature. When there are only 500 atoms, however, the substrate is almost immediately vaporised by the deposition. Something similar happens in biophysical simulations. The temperature of the system in NVE is naturally raised when macromolecules such as proteins undergo exothermic conformational changes and binding.

Canonical Ensemble (NVT)

In the canonical ensemble, moles (N), volume (V) and temperature (T) are conserved. It is also sometimes called constant

temperature molecular dynamics (CTMD). In NVT, the energy of endothermic and exothermic processes is exchanged with a thermostat. A variety of thermostat methods are required to add and remove energy from the boundaries of an MD system in a realistic way, approximating the canonical ensemble. Popular techniques to control temperature include the Nose-Hoover thermostat and Langevin dynamics.

Isothermal-Isobaric (NPT) Ensemble

In the isothermal-isobaric ensemble, moles (N), pressure (P) and temperature (T) are conserved. In addition to a thermostat, a barostat is needed. It corresponds most closely to laboratory conditions with a flask open to ambient temperature and pressure.

In the simulation of biological membranes, isotropic pressure control is not appropriate. For lipid bilayers, pressure control occurs under constant membrane area (NPAT) or constant surface tension "gamma" (NPγT).

Generalised Ensembles

The replica exchange method is a generalised ensemble. It was originally created to deal with the slow dynamics of disordered spin systems. It is also called parallel tempering. The replica exchange MD (REMD) formulation tries to overcome the multiple-minima problem by exchanging the temperature of non-interacting replicas of the system running at several temperatures.

Potentials in MD Simulations

A molecular dynamics simulation requires the definition of a potential function, or a description of the terms by which the particles in the simulation will interact. In chemistry and biology this is usually referred to as a force field. Potentials may be defined at many levels of physical accuracy; those most commonly used in chemistry are based on molecular mechanics and embody a classical treatment of particle-particle interactions that can reproduce structural and conformational changes but usually cannot reproduce chemical reactions.

When finer levels of detail are required, potentials based on quantum mechanics are used; some techniques attempt to create

hybrid classical/quantum potentials where the bulk of the system is treated classically but a small region is treated as a quantum system, usually undergoing a chemical transformation.

Empirical Potentials

Empirical potentials used in chemistry are frequently called force fields, while those used in materials physics are called just empirical or analytical potentials.

Most force-fields in chemistry are empirical and consist of a summation of bonded forces associated with chemical bonds, bond angles, and bond dihedrals, and non-bonded forces associated with van der Waals forces and electrostatic charge. Empirical potentials represent quantum-mechanical effects in a limited way through *ad hoc* functional approximations.

These potentials contain free parameters such as atomic charge, van der Waals parameters reflecting estimates of atomic radius, and equilibrium bond length, angle, and dihedral; these are obtained by fitting against detailed electronic calculations (quantum chemical simulations) or experimental physical properties such as elastic constants, lattice parameters and spectroscopic measurements.

Chemistry force fields are commonly only able to represent a single potential energy surface and thus are unable to model the process of chemical bond breaking and reactions explicitly. This is a direct consequence of using the Born-Oppenheimer approximation or adiabatic formalism.

If the potential is designed for the ground state, it cannot abandon it. On the other hand, many of the potentials used in physics, such as those based on the bond order formalism can describe several different coordinations of a system and bond breaking. Examples of such potentials include the Brenner potential for hydrocarbons and its further developments for the C-Si-H and C-O-H systems. The ReaxFF potential can be considered a fully reactive hybrid between bond order potentials and chemistry force fields.

The potential functions representing the non-bonded interactions are usually “pair potentials”, in which the total

potential energy of a system can be calculated from the sum of energy contributions from pairs of atoms. These non-bonded interactions, because they are non-local and involve at least weak interactions between every pair of particles in the system, are normally the bottleneck in the speed of MD simulations. In electrostatic interactions, solving the Poisson equation for complete systems is usually prohibitively slow; instead numerical approximations are used such as shifted cutoff radii, reaction field algorithms, particle mesh Ewald summation, or the newer Particle-Particle Particle Mesh (P3M).

An example of a calculated pair potential is the non-bonded Lennard-Jones potential (also known as the 6-12 potential), used for calculating van der Waals forces.

$$\mathrm{U(r)} = 4\varepsilon\left[\left(\frac{\sigma}{\mathrm{r}}\right)^{12} - \left(\frac{\sigma}{\mathrm{r}}\right)^{6}\right]$$

Another example is the Born (ionic) model of the ionic lattice. The first term in the next equation is Coulomb's law for a pair of ions, the second term is the short-range repulsion explained by Pauli's exclusion principle and the final term is the dispersion interaction term. Usually, a simulation only includes the dipolar term, although sometimes the quadrupolar term is as well.

$$U_{ij}(r_{ij}) = \sum \frac{z_i z_j}{4\pi \in_0} \frac{1}{r_{ij}} + \sum A_{l\exp} \frac{-r_{ij}}{pl} \sum C_{lr_{ij}}{}^{-n_l} + \ldots$$

Empirical potentials can be subcategorised into pair potentials and many-body potentials. In many-body potentials, the potential energy cannot be found by a sum over pairs of atoms. For example, the Tersoff potential, which was originally used to simulate Carbon, Silicon and Germanium and later on has been implemented for a wide range of other materials, involves a sum over groups of three atoms, with the angles between the atoms being an important factor in the potential. Other examples are the Embedded-atom method (EAM) and the Tight-Binding Second Moment Approximation (TBSMA) potentials , where the electron density of states in the region of an atom is calculated from a sum of contributions from surrounding atoms, and the potential energy contribution is then a function of this sum.

Semi-empirical Potentials

Semi-empirical potentials make use of the matrix representation from quantum mechanics. However, the values of the matrix elements are found through empirical formulae that estimate the degree of overlap of specific atomic orbitals. The matrix is then diagonalised to determine the occupancy of the different atomic orbitals, and empirical formulae are used once again to determine the energy contributions of the orbitals.

There are a wide variety of semi-empirical potentials, known as tight-binding potentials, which vary according to the atoms being modelled.

Ab-initio Methods

Ab-initio quantum-mechanical formula is used to calculate the potential energy of a system of atoms or molecules. Compared to classical potential function, which is represented by empirical functions, the properties of the system in *ab-intio* calculations are calculating the wavefunctions for electrons moving around the nucleus of atoms. This calculation is usually made "locally", i.e., for nuclei in the close neighbourhood of the reaction coordinate. Although various approximations may be used, these are based on theoretical considerations, not on empirical fitting.

Ab-Initio produce a large amount of information that is not available from the empirical methods, such as density of states information. Of course, the computational price paid is high. A significant advantage of using *ab-initio* methods is the ability to study reactions that involved breakage or formation of covalent bonds, this would correspond to multiple electronic states. Classical molecular dynamics is unable to simulate breakage and formation of covalent bonds, However, in recent years techniques such as thermodynamic integration and ghost particles have been introduced to overcome these limitations. The success, however, remains limited.

A popular package for *ab-initio* molecular dynamics is the Car-Parrinello Molecular Dynamics (CPMD) package based on the density functional theory.

Hybrid QM/MM

QM (quantum-mechanical) methods are very powerful, however they are computationally expensive, while the MM (classical or molecular mechanics) methods are fast but suffer from several limitations (require extensive parameterisation; energy estimates obtained are not very accurate; cannot be used to simulate reactions where covalent bonds are broken/formed; and are limited in their abilities for providing accurate details regarding the chemical environment). A new class of method has emerged that combines the good points of QM (accuracy) and MM (speed) calculations. These methods are known as mixed or hybrid quantum-mechanical and molecular mechanics methods (hybrid QM/MM).

The methodology for such techniques was introduced by Warshel and co-workers. In the recent years have been pioneered by several groups including: Arieh Warshel (University of Southern California), Weitao Yang (Duke University), Sharon Hammes-Schiffer (The Pennsylvania State University), Donald Truhlar and Jiali Gao (University of Minnesota) and Kenneth Merz (University of Florida). The most important advantage of hybrid QM/MM methods is the speed. The cost of doing classical molecular dynamics (MM) in the most straight forward case scales $O(n^2)$, where N is the number of atoms in the system.

This is mainly due to electrostatic interactions term (every particle interacts with everything else). However, use of cutoff radius, periodic pair-list updates and more recently the variations of the particle-mesh Ewald's (PME) method has reduced this between O(N) to $O(n^2)$. In other words, if a system with twice many atoms is simulated then it would take between twice to four times as much computing power. On the other hand the simplest *ab-initio* calculations typically scale $O(n^3)$ or worse [Restricted Hartree-Fock calculations have been suggested to scale $\sim O(n^{2.7})$]. To overcome the limitation, a small part of the system is treated quantum-mechanically (typically active-site of an enzyme) and the remaining system is treated classically.

In more sophisticated implementations, QM/MM methods exist to treat both light nuclei susceptible to quantum effects (such

as hydrogens) and electronic states. This allows generation of hydrogen wavefunctions (similar to electronic wavefunctions). This methodology has been useful in investigating phenomenon such as hydrogen tunnelling.

One example where QM/MM methods have provided new discoveries is the calculation of hydride transfer in the enzyme liver alcohol dehydrogenase. In this case, tunnelling is important for the hydrogen, as it determines the reaction rate.

Coarse-graining and Reduced Representations

At the other end of the detail scale are coarse-grained and lattice models. Instead of explicitly representing every atom of the system, one uses "pseudo-atoms" to represent groups of atoms. MD simulations on very large systems may require such large computer resources that they cannot easily be studied by traditional all-atom methods. Similarly, simulations of processes on long timescales (beyond about 1 microsecond) are prohibitively expensive, because they require so many timesteps. In these cases, one can sometimes tackle the problem by using reduced representations, which are also called coarse-grained models.

Examples for coarse graining (CG) methods are discontinuous molecular dynamics (CG-DMD) and Go-models. Coarse-graining is done sometimes taking larger pseudo-atoms. Such united atom approximations have been used in MD simulations of biological membranes. The aliphatic tails of lipids are represented by a few pseudo-atoms by gathering 2-4 methylene groups into each pseudo-atom.

The parameterisation of these very coarse-grained models must be done empirically, by matching the behaviour of the model to appropriate experimental data or all-atom simulations. Ideally, these parameters should account for both enthalpic and entropic contributions to free energy in an implicit way. When coarse-graining is done at higher levels, the accuracy of the dynamic description may be less reliable. But very coarse-grained models have been used successfully to examine a wide range of questions in structural biology.

Examples of applications of coarse-graining in biophysics:

- Protein folding studies are often carried out using a single (or a few) pseudo-atoms per amino acid;
- DNA supercoiling has been investigated using 1-3 pseudo-atoms per basepair, and at even lower resolution;
- Packaging of double-helical DNA into bacteriophage has been investigated with models where one pseudo-atom represents one turn (about 10 basepairs) of the double helix;
- RNA structure in the ribosome and other large systems has been modelled with one pseudo-atom per nucleotide.

The simplest form of coarse-graining is the "united atom" (sometimes called "extended atom") and was used in most early MD simulations of proteins, lipids and nucleic acids. For example, instead of treating all four atoms of a CH_3 methyl group explicitly (or all three atoms of CH_2 methylene group), one represents the whole group with a single pseudo-atom. This pseudo-atom must, of course, be properly parameterised so that its van der Waals interactions with other groups have the proper distance-dependence.

Similar considerations apply to the bonds, angles, and torsions in which the pseudo-atom participates. In this kind of united atom representation, one typically eliminates all explicit hydrogen atoms except those that have the capability to participate in hydrogen bonds (polar hydrogens). An example of this is the Charmm 19 force-field.

The polar hydrogens are usually retained in the model, because proper treatment of hydrogen bonds requires a reasonably accurate description of the directionality and the electrostatic interactions between the donor and acceptor groups. A hydroxyl group, for example, can be both a hydrogen bond donor and a hydrogen bond acceptor, and it would be impossible to treat this with a single OH pseudo-atom.

Note that about half the atoms in a protein or nucleic acid are non-polar hydrogens, so the use of united atoms can provide a substantial savings in computer time.

Examples of Applications

Molecular dynamics is used in many fields of science.

- First macromolecular MD simulation published (1977, Size: 500 atoms, Simulation Time: 9.2 ps = 0.0092 ns, Programme: CHARMM precursor) Protein: Bovine Pancreatic Trypsine Inhibitor. This is one of the best studied proteins in terms of folding and kinetics. Its simulation published in Nature magazine paved the way for understanding protein motion as essential in function and not just accessory.

The following two biophysical examples are not run-of-the-mill MD simulations.

They illustrate almost heroic efforts to produce simulations of a system of very large size (a complete virus) and very long simulation times (500 microseconds):

- MD simulation of the complete satellite tobacco mosaic virus (STMV) (2006, Size: 1 million atoms, Simulation time: 50 ns, programme: NAMD) This virus is a small, icosahedral plant virus which worsens the symptoms of infection by Tobacco Mosaic Virus (TMV). Molecular dynamics simulations were used to probe the mechanisms of viral assembly. The entire STMV particle consists of 60 identical copies of a single protein that make up the viral capsid (coating), and a 1063 nucleotide single stranded RNA genome. One key finding is that the capsid is very unstable when there is no RNA inside. The simulation would take a single 2006 desktop computer around 35 years to complete. It was thus done in many processors in parallel with continuous communication between them.
- Folding Simulations of the Villin Headpiece in All-Atom Detail (2006, Size: 20,000 atoms; Simulation time: 500 μs = 500,000 ns, Programme: folding@home) This simulation was run in 200,000 CPU's of participating personal computers around the world. These computers had the folding@home programme installed, a large-scale distributed computing effort coordinated by Vijay Pande at Stanford University. The kinetic properties of the Villin Headpiece protein were probed by using many

independent, short trajectories run by CPU's without continuous real-time communication. One technique employed was the Pfold value analysis, which measures the probability of folding before unfolding of a specific starting conformation. Pfold gives information about transition state structures and an ordering of conformations along the folding pathway. Each trajectory in a Pfold calculation can be relatively short, but many independent trajectories are needed.

Molecular Dynamics Algorithms

Integrators

- Verlet integration
- Constraint algorithms (for constrained systems)
- Symplectic integrator

Short-range Interaction Algorithms

- Cell lists
- Verlet list
- Bonded interactions

Long-range Interaction Algorithms

- Ewald summation
- Particle Mesh Ewald (PME)
- Particle-Particle-Particle-Mesh (P3M)
- Reaction Field Method

Parallelisation Strategies

- Domain decomposition method (Distribution of system data for parallel computing)
- Molecular Dynamics - Parallel Algorithms

Major Software for MD Simulations

- ABINIT (DFT)
- AMBER (classical)
- CASTEP (ab initio)

- CPMD (DFT)
- CHARMM (classical)
- ESPResSo (classical, coarse-grained, parallel, extensible)
- Fireball (ab initio)
- DL_POLY (classical)
- GROMACS (classical)
- GROMOS (classical)
- LAMMPS (classical, large-scale with spatial-decomposition of simulation domain for parallelism)
- NAMD
- PWscf
- SIESTA (DFT)
- VASP (Ab-initio)
- TINKER (classical)
- YASARA (classical)
- ORAC (classical)

Related Software

- VMD - MD simulation trajectories can be loaded and visualised.
- PyMol - Molecular Modelling software written in python.
- Sirius - Molecular modelling, analysis and visualisation of MD trajectories.
- Esra - Lightweight molecular modelling and analysis library (Java/Jython/Mathematica).

Cell Membrane Mechanism

Biomembrane

A biological membrane or biomembrane is an enclosing or separating tissue which acts as a barrier within or around a cell. It is, almost invariably, a lipid bilayer, composed of a double layer of lipid-class molecules, specifically phospholipids, with occasional proteins intertwined, some of which function as channels.

Function

Such membranes typically define enclosed spaces or compartments in which cells may maintain a chemical or biochemical environment that differs from the outside. For example, the membrane around peroxisomes shields the rest of the cell from peroxides, and the plasma membrane separates a cell from its surrounding medium. Most organelles are defined by such membranes, and are called membrane-bound organelles.

Probably the most important feature of a biomembrane is that it is a selectively permeable structure. This means that the size, charge and other chemical properties of the atoms and molecules attempting to cross it will determine whether they succeed to do so. Selective permeability is essential for effective separation of a

cell or organelle from its surroundings. Biological membranes also have certain mechanical or elastic properties.

If a particle is too large or otherwise unable to cross the membrane by itself, but is still needed by a cell, it could either go through one of the protein channels, or be taken in by means of endocytosis.

Types of Biological Membranes

- Cell membrane
- Mucous membrane
- S-layer

Uses of Cell Adhesion

Cellular adhesion is the binding of a cell to another cell or to a surface or matrix. Cellular adhesion is regulated by specific adhesion molecules that interact with molecules on the opposing cell or surface. Such adhesion molecules are also termed "receptors" and the molecules they recognise are termed "ligands" (and sometimes "counterreceptors").

Since cells are not often found in isolation, rather they tend to stick to other cells or non-cellular components of their environment, a fundamental question is: what makes cells sticky? Cell adhesion generally involves protein molecules at the surface of cells, so the study of cell adhesion involves cell adhesion proteins and the molecules that they bind to.

Cytoskeletal Interactions

For a cell adhesion protein the intracellular domain binds to protein components of the cell's cytoskeleton. This allows for very tight adhesion. Without attachment to the cytoskeleton, a cell adhesion protein that is tightly bound to a ligand would be in danger of being hydrolysed by extracellular hydrolytic enzymes. This will rip out the adhesion protein from the fragile cell membrane. For example, cadherin cell adhesion proteins are typically coupled to the cytoskeleton by way of special linking proteins called "catenins".

Adhesion in Prokaryotes

Prokaryotes have adhesion molecules usually termed "adhesins". Adhesins may occur on pili (fimbriae), flagellae, or the cell surface. Adhesion of bacteria is the first step in colonisation and regulates tropism (tissue- or cell-specific interactions).

Adhesion in Virii

Viruses also have adhesion molecules required for viral binding to host cells. For example, influenza virus has a haemagglutinin on its surface that is required for recognition of the sugar sialic acid on host cell surface molecules. HIV has an adhesion molecule termed gp120 that binds to its ligand CD4, which is expressed on lymphocytes.

Adhesion in Eukaryotes

Eukaryotic protozoans also express multiple adhesion molecules. An example of a pathogenic protozoan is the malarial parasite (*Plasmodium falciparum*), which uses one adhesion molecule called the circumsporozoite protein to bind to liver cells, and another adhesion molecule called the merozoite surface protein to bind red blood cells. In human cells, which have many different types of adhesion molecules, the major classes are named integrins, Ig superfamily members, cadherins, and selectins. Each of these adhesion molecules has a different function and recognises different ligands. Defects in cell adhesion are usually attributable to defects in expression of adhesion molecules.

Human Genetic Diseases

There are human genetic diseases caused by inability to express a specific adhesion molecule. An example is leukocyte adhesion deficiency-I (LAD-I), where patients do not express the β -2-integrin subunit precursor. This integrin is required for leukocytes to adhere to the blood vessel wall during inflammation in order to fight infection. The leukocytes from LAD-I patients fail to adhere and patients exhibit serious episodes of infection that can be life threatening.

Efflux System

Active efflux is a mechanism responsible for extrusion of toxic substances and antibiotics outside the cell. Its importance lies in

its contribution to bacterial anti-microbial resistance. Efflux systems function via an energy-dependent mechanism (Active transport) to pump out unwanted toxic substances through specific efflux pumps.

Some efflux systems are drug-specific while others may accommodate multiple drugs and thus contribute significantly to bacterial multi-drug resistance (MDR).

Efflux in Bacteria

Bacterial Efflux Pumps: Efflux pumps are proteinaceous transporters localised in the cytoplasmic membrane of all kind of cells. They are active transporters meaning that they require a source of chemical energy to perform their function.

Some are primary active transporters utilising Adenosine triphosphate hydrolysis as a source of energy, while others are secondary active transporters (uniporters, symporters or anti-porters) in which transport is coupled to an electrochemical potential difference created by pumping out hydrogen or sodium ions outside the cell.

Bacterial efflux transporters are classified into five major superfamilies, based on the amino acid sequence and the energy source used to export their substrates:

1. The major facilitator superfamily (MFS);
2. The ATP-binding cassette superfamily (ABC);
3. The small multidrug resistance family (SMR);
4. The resistance-nodulation-cell division superfamily (RND); and
5. The multidrug and toxic compound extrusion family (MATE).

Of these only the ABC superfamily are primary transporters, the rest being secondary transporters utilising proton or sodium gradient as a source of energy. While MFS dominates in Gram positive bacteria , the RND family is unique to Gram-negatives.

Function

Although antibiotics are the most clinically important substrates of efflux systems, it is probable that most efflux pumps have other natural physiological functions.

Examples include:

- The E.coli AcrAB efflux system which has a physiologic role of pumping out bile acids and fatty acids to lower their toxicity.
- The MFS family Ptr pump in Streptomyces pristinaespiralis appears to be an auto-immunity pump for this organism when it turns on production of pristinamycins I and II.
- The AcrAB–TolC system in E.coli is suspected to have a role in the transport of the calcium-channel components in the E. coli membrane.
- The MtrCDE system plays a protective role by providing resistance to faecal lipids in rectal isolates of Neisseria gonorrhoeae.
- The AcrAB efflux system of Erwinia amylovora is important for this organism's virulence, plant (host) colonisation and resistance to plant toxins.

The ability of efflux systems to recognise a large number of compounds other than their natural substrates is probably because substrate recognition is based on physicochemical properties, such as hydrophobicity, aromaticity and ionisable character rather than on defined chemical properties, as in classical enzyme-substrate or ligand-receptor recognition. Because most antibiotics are amphiphilic molecules—possessing both hydrophilic and hydrophobic characters, they are easily recognised by many efflux pumps.

Impact on Anti-microbial Resistance

The impact of efflux mechanisms on anti-microbial resistance is large, this is usually attributed to the following:

- The genetic elements encoding efflux pumps may be encoded on chromosomes and/or plasmids, thus contributing to both intrinsic (natural) and acquired resistance respectively. As an intrinsic mechanism of resistance, efflux pump genes can survive a hostile environment (for example in the presence of antibiotics) which allows for the selection of mutants that over-express

these genes. Being located on transpoable genetic elements as plasmids or transposons is also advantageous for the microorganisms as it allows for the easy spread of efflux genes between distant species.

- Antibiotics can act as inducers and regulators of the expression of some efflux pumps.
- Expression of several efflux pumps in a given bacterial species may lead to a broad spectrum of resistance when considering the shared substrates of some multi-drug efflux pumps, where one efflux pump may confer resistance to a wide range of anti-microbials.

Efflux in Eukaryotes

Efflux Inhibitors: Several trials are currently being conducted to develop drugs that can be co-administered with antibiotics to act as inhibitors for the efflux-mediated extrusion of antibiotics. None of the efflux inhibitors tested is yet in clinical use. However, some of them are used to determine the efflux prevalence in clinical isolates. Its shown that Verapamil can inhibit P-glycoprotein mediated efflux which can increase oral absorption of some compounds.

Flexibility of Cell Membranes

A cell membrane defines a boundary between the living cell and its environment. It consists of lipids, proteins, carbohydrates, etc. Lipids and proteins are dominant components of membranes. One of the principal types of lipids in membranes is phospholipid. A phospholipid molecule has a polar hydrophilic head group and two hydrophobic hydrocarbon tails. When a quantity of lipid molecules disperse in water, they will assemble themselves into a bilayer in which the hydrophilic heads shield the hydrophobic tails from the water surroundings because of the hydrophobic forces.

The widely accepted model for cell membranes is the fluid mosaic model proposed by Singer and Nicolson in 1972 [Science 175 (1972) 720]. In this model, the cell membrane is considered as a lipid bilayer where the lipid molecules can move freely in the membrane surface like fluid, while the proteins are embedded in the lipid bilayer.

Some proteins are called integral membrane proteins because they traverse entirely in the lipid bilayer and play the role of information and matter communications between the interior of the cell and its outer environment. The others are called peripheral membrane proteins because they are partially embedded in the bilayer and accomplish the other biological functions. Beneath the lipid membrane, the membrane skeleton, a network of proteins, links with the proteins in the lipid membrane.

Elasticity of Lipid Vesicles

The first step to study the elasticity of cell membranes is to study lipid bilayers. Usually, the thickness of the lipid bilayer is much smaller than the scale of the whole lipid bilayer. It is reasonable to describe the lipid bilayer by a mathematical surface. In 1973, Helfrich [Z. Naturforsch. C 28 (1973) 693] recognised that the lipid bilayer was just like a nematic liquid crystal at room temperature, and then proposed the curvature energy per unit area of the bilayer

$$f_c = \frac{k_c}{2}(2H + c_0)^2 + \bar{k}K, \tag{1}$$

where k_c, $\bar{k}$ are bending rigidities. c_0 is called the spontaneous curvature of the membrane. H and K are the mean and Gaussian curvature of the membrane surface, respectively. The free energy of a closed bilayer under the osmotic pressure $-p$ (the outer pressure minus the inner one) as:

$$F_H = \int (f_c + \lambda)dA + \Delta p \int dV, \tag{2}$$

where dA and dV are the area element of the membrane and the volume element enclosed by the closed bilayer, respectively. λ is the surface tension of the bilayer. By taking the first order variation of above free energy, Ou-Yang and Helfrich derived an equation to describe the equilibrium shape of the bilayer as:

$$\Delta p - 2\lambda H + k_c(2H + c_0)(2H^2 - c_0 H - 2K) + k_c\nabla^2(2H) = 0 \tag{3}$$

They also obtained that the threshold pressure for the instability of a spherical bilayer was

$$\Delta p_c \propto k_c R^3 \tag{4}$$

where R being the radius of the spherical bilayer.

Using the shape equation (3) of closed vesilces, Ou-Yang predicted that there was a lipid torus with the ratio of two generated radii being exactly $\sqrt{2}$. His prediction was soon confirmed by the experiment. Additionally, researchers obtained an analytical solution to (3) which explained the classical problem—the biconcave discoidal shape of normal red cells.

Elasticity of Open Lipid Membranes

The opening-up process of lipid bilayers by talin was observed by Saitoh *et al.* arose the interest of studying the equilibrium shape equation and boundary conditions of lipid bilayers with free exposed edges. Capovilla *et al.*, Tu and Ou-Yang carefully studied this problem. The free energy of a lipid membrane with an edge C is written as

$$F_o = \int \left(f_c + \lambda \right) dA + \gamma \oint_c ds \tag{5}$$

where ds and γ represent the arclength element and the line tension of the edge, respectively. The first order variation gives the shape equation and boundary conditions of the lipid membrane:

$$k_c(2H + c_0)(2H^2 - c_0 H - 2K) - 2\lambda H + k_c \nabla^2 (2H) = 0 \tag{6}$$

$$[k_c(2H + c_0) + \bar{k} K_n]_c = 0 \tag{7}$$

$$\left[-2k_c \frac{\partial H}{\partial e_2} + \gamma k_n = \bar{k} \frac{d\tau_g}{ds} \right]_c = 0 \tag{8}$$

$$\left[\frac{k_c}{2}(2H + c_0)^2 + \bar{k} K + \lambda + \gamma k_g \right]\Big|_c = 0 \tag{9}$$

where k_n, k_g, and τ_g are normal curvature, geodesic curvature, and geodesic torsion of the boundary curve, respectively. e_2 is the unit vector perpendicular to the tangent vector of the curve and the normal vector of the membrane.

Elasticity of Cell Membranes

A cell membrane is simplified as lipid bilayer plus membrane skeleton. The skeleton is a cross-linking protein network and

joints to the bilayer at some points. Assume that each proteins in the membrane skeleton have similar length which is much smaller than the whole size of the cell membrane, and that the membrane is locally 2-dimensional uniform and homogenous. Thus the free energy density can be expressed as the invariant form of $2H$, K, $tr\varepsilon$ and $det\varepsilon$:

$$f_{cm} = f(2\mathrm{H},\ \mathrm{K},\ tr\varepsilon,\ det\varepsilon) \tag{10}$$

where ε is the in-plane strain of the membrane skeleton. Under the assumption of small deformations, and invariant between $tr\varepsilon$ and $-tr\varepsilon$, (10) can be expanded up to second order terms as:

$$f_m = \frac{k_c}{2}(2H + c_0)^2 + \bar{k}K + \lambda + \frac{k_d}{2}(\mathrm{tr}\varepsilon)^2 - 2\mu(\det\varepsilon) \tag{11}$$

where k_d and μ are two elastic constants. In fact, the first two terms in (11) are the bending energy of the cell membrane which contributes mainly from the lipid bilayer. The last two terms come from the entropic elasticity of the membrane skeleton.

Role of Gram-negative Bacteria

Gram-negative bacteria are those that do not retain crystal violet dye in the Gram staining protocol. Gram-positive bacteria will retain the dark blue dye after an alcohol wash. In a Gram stain test, a counterstain (commonly Safranin) is added after the crystal violet, colouring all Gram-negative bacteria a red or pink colour. The test itself is useful in classifying two distinctly different types of bacteria based on structural differences in their cell walls.

Many species of Gram-negative bacteria are pathogenic, meaning they can cause disease in a host organism. This pathogenic capability is usually associated with certain components of Gram-negative cell walls, in particular the lipopolysaccharide (also known as LPS or endotoxin) layer. The LPS is the trigger which the body's innate immune response receptors sense to begin a cytokine reaction. It is toxic to the host. It is this response which begins the inflammation cycle in tissues and blood vessels.

Characteristics

The following characteristics are displayed by Gram-negative bacteria:

1. Cell walls only contain a few layers of peptidoglycan (which is present in much higher levels in Gram-positive bacteria).
2. Cells are surrounded by an outer membrane containing lipopolysaccharide (which consists of Lipid A, core polysaccharide, and O-polysaccharide) outside the peptidoglycan layer.
3. Porins exist in the outer membrane, which act like pores for particular molecules.
4. There is a space between the layers of peptidoglycan and the secondary cell membrane called the periplasmic space.
5. The S-layer is directly attached to the outer membrane, rather than the peptidoglycan.
6. If present, flagella have four supporting rings instead of two.
7. No teichoic acids or lipoteichoic acids are present.
8. Lipoproteins are attached to the polysaccharide backbone whereas in Gram-positive bacteria no lipoproteins are present.
9. Most do not sporulate (*Coxiella burnetti* forms spore-like structures).

Example Species

The proteobacteria are a major group of Gram-negative bacteria, including Escherichia coli, Salmonella, and other Enterobacteriaceae, Pseudomonas, Moraxella, Helicobacter, *Stenotrophomonas*, Bdellovibrio, acetic acid bacteria, Legionella and many others.

Other notable groups of Gram-negative bacteria include the cyanobacteria, spirochaetes, green sulphur and green non-sulphur bacteria. Crenarchaeota: Unique because most bacteria have gram-positive molecules in their capsules, it has gram-negative.

Medically relevant Gram-negative cocci include three organisms, which cause a sexually transmitted disease (Neisseria gonorrhoeae), a meningitis (Neisseria meningitidis), and respiratory symptoms (Moraxella catarrhalis).

Medically relevant Gram-negative bacilli include a multitude of species. Some of them primarily cause respiratory problems (Hemophilus influenzae, Klebsiella pneumoniae, Legionella pneumophila, Pseudomonas aeruginosa), primarily urinary problems (Escherichia coli, Proteus mirabilis, Enterobacter *cloacae*, Serratia marcescens), and primarily gastrointestinal problems (Helicobacter pylori, Salmonella enteritidis, Salmonella typhi).

Nosocomial gram negative bacteria include Acinetobacter baumanii, which cause bacteremia, secondary meningitis, and ventilator-associated pneumonia in intensive care units of hospital establishments.

Medical Treatment

One of the several unique characteristics of Gram-negative bacteria is the outer membrane. This outer membrane is responsible for protecting the bacteria from several antibiotics, dyes, and detergents which would normally damage the inner membrane or cell wall (peptidoglycan). The outer membrane provides these bacteria with resistance to lysozyme and penicillin. Fortunately, alternative medicinal treatments such as lysozyme with EDTA, and the antibiotic ampicillin have been developed to combat the protective outer membrane of some pathogenic Gram-negative organisms.

Role of Gram-positive Bacteria

Gram-positive bacteria are those that retain a crystal violet dye during the Gram stain process. Gram-positive bacteria appear blue or violet under a microscope, while Gram-negative bacteria appear red or pink. The Gram classification system is empirical, and largely based on differences in cell wall structure. The purpose of Gram staining is to visually differentiate groups of bacteria, primarily for identification.

Characteristics

The following characteristics are generally present in a Gram-positive bacterium:

1. A very thick cell wall (peptidoglycan).
2. If a flagellum is present, it contains two rings for support as opposed to four in Gram-negative bacteria because Gram-positive bacteria have only one membrane layer.

3. Teichoic acids and lipoteichoic acids are present, which serve to act as chelating agents, and also for certain types of adherence.

History of Gram-positive Bacteria

In the original bacterial phyla, the Gram-positive forms made up the phylum Firmicutes, a name now used for the largest group. It includes many well-known genera such as Bacillus, Listeria, Staphylococcus, Streptococcus, Enterococcus, and Clostridium. It has also been expanded to include the Mollicutes, bacteria like Mycoplasma that lack cell walls and so cannot be stained by Gram, but are derived from such forms.

The actinobacteria are another major group of Gram-positive bacteria; they and the Firmicutes are referred to as the high and low G+C groups based on the guanosine and cytosine content of their DNA.

If the second membrane is a derived condition, the two may have been basal among the bacteria; otherwise they are probably a relatively recent monophyletic group. They have been considered as possible ancestors for the archaeans and eukaryotes, both because they are unusual in lacking the second membrane and because of various biochemical similarities such as the presence of sterols.

The Deinococcus-Thermus bacteria also have Gram-positive stains, although they are structurally similar to Gram-negative bacteria.

Both Gram-positive and Gram-negative bacteria may have a membrane called an S-layer. In Gram-negative bacteria, the S-layer is directly attached to the outer membrane. In Gram-positive bacteria, the S-layer is attached to the peptidoglycan layer. Unique to Gram-positive bacteria is the presence of teichoic acids in the cell wall. Some particular teichoic acids, lipoteichoic acids, have a lipid component and can assist in anchoring peptidoglycan, as the lipid component is embedded in the membrane.

Definition of Biological Thermodynamics

In thermodynamics, biological thermodynamics (Greek: *bios* = life and *logikos* = reason + Greek: *thermos* = heat and *dynamics*

= power) or bioenergetics is the study of energy transformation in the biological sciences.

More definitively, biological thermodynamics may be defined as the quantitative study of the energy transductions that occur in and between living organisms, structures, and cells and of the nature and function of the chemical processes underlying these transductions. Biological thermodynamics may address the question of whether the benefit associated with any particular phenotypic trait is worth the energy investment it requires.

History

German-British medical doctor and biochemist Hans Krebs' 1957 book *Energy Transformations in Living Matter* (written with Hans Kornberg) was the first major publication on the thermodynamics of biochemical reactions.

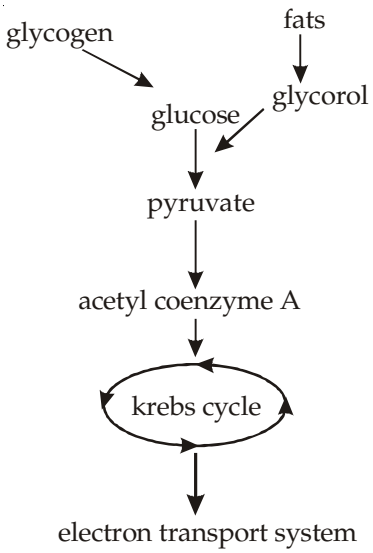
In addition, the appendix contained the first-ever published thermodynamic tables, written by K. Burton, to contain equilibrium constants and Gibbs free energy of formations for chemical species, able to calculate biochemical reactions that had not yet occurred.

Phenomenon of Bioenergetics

Growth, development and metabolism are some of the central phenomena in the study of biological organisms. Living cells and organisms must perform work to stay alive, to grow, and to reproduce themselves.

The energy concept is useful to explain such biological processes. The ability to harness energy from a variety of metabolic pathways and channelise it into activities of organism is a fundamental property of all living organisms. Sustenance of life is critically dependent on energy transformations; living organisms survive because of exchange of energy within and without.

In a living organism chemical bonds are constantly broken and made to make the exchange and transformation of energy possible. These chemical bonds are most often bonds in carbohydrates, including sugars.



Other chemical bonds include bonds in chemical compounds that are important for metabolism, for example, those in a molecule of ATP or fats and oils. These molecules, along with oxygen, are common stores of concentrated energy for the biological processes. One can therefore assert that transformation of energy from a more to a less concentrated form is the driving force of all biological processes or chemical processes that are responsible for the life of a biological organism. Molecular biology and biochemistry are in fact scientific studies concerning the making and breaking of chemical bonds in the molecules found in biological organisms.

As with other natural phenomena, the exchange of metabolic energy in biological organisms is often accompanied with an increase in the entropy of the universe.

Energy transformations in biology are primarily due to the chemical synthesis and decompositions that are brought about by the energy absorbed by organisms from sunlight through insolation and photosynthesis. The total energy captured by photosynthesis in green plants from the solar radiation is about $2 \quad 10^{23}$ joules of energy per year. Annual energy captured by photosynthesis in green plants is about 4 per cent of the total sunlight energy which reaches Earth.

The energy transformations in biological communities surrounding hydrothermal vents are exceptions. They oxidise sulphur, obtaining their energy via chemosynthesis rather than photosynthesis. The oxygen used to do this is photosynthetically derived, but the sulphur in the thermodynamically unstable, non-oxidised state exists due to geothermal energy.

Food that is ingested by an organism contains energy. Not all metabolisable energy is available for the production of ATP. Some energy is utilised during the metabolic processes associated with digestion, absorption and intermediary metabolism of food and can be measured as heat production; this is referred to as dietary-induced thermogenesis (DIT), or thermic effect of food, and varies with the type of food ingested. The predator-prey relationships, food chains, are in effect energy transformations within ecosystems.

Central Point of Thermodynamics

The field of biological thermodynamics is focused on thermodynamic applications of the principles of chemical thermodynamics in biology and biochemistry. Principles covered include the first law of thermodynamics, the second law of thermodynamics, Gibbs free energy, statistical thermodynamics, reaction kinetics, and on hypotheses of the origin of life.

Presently, biological thermodynamics concerns itself with the study of internal biochemical dynamics as: ATP hydrolysis, protein stability, DNA binding, membrane diffusion, enzyme kinetics, and other such essential energy controlled pathways. Thermodynamically, the amount of energy capable of doing work during a chemical reaction is measured quantitatively by the change in the Gibbs free energy. The physical biologist Alfred Lotka attempted to unify the change in the Gibbs free energy with evolutionary theory.

System of Cell Membrane

The cell membrane (also called the plasma membrane, plasmalemma or phospholipid bilayer) is a semi-permeable lipid bilayer common to all living cells. It contains a variety of biological molecules, primarily proteins and lipids, which are involved in

a vast array of cellular processes, and also serves as the attachment point for both the intracellular cytoskeleton and, if present, the cell wall.

Robert Hooke was the first to name the parts of cells, including the plasma membrane.

The cell membrane surrounds the cytoplasm of a cell and physically separates the intracellular components from the extracellular environment, thereby serving a function similar to that of skin.

The barrier is selectively permeable and able to regulate what enters and exits the cell, thus facilitating the transport of materials needed for survival.

The movement of substances across the membrane can be either *passive*, occurring without the input of cellular energy, or *active*, requiring the cell to expend energy in moving it.

Specific proteins embedded in the cell membrane can act as molecular signals which allow cells to communicate with each other. Protein receptors are found ubiquitously and function to receive signals from both the environment and other cells. These signals are transduced into a form which the cell can use to directly effect a response.

Other proteins on the surface of the cell membrane serve as "markers" which identify a cell to other cells. The interaction of these markers with their respective receptors forms the basis of cell-cell interaction in the immune system.

Structure

Lipid Bilayer: The cell membrane consists of a thin layer of amphipathic lipids which spontaneously arrange so that the hydrophobic "tail" regions are shielded from the surrounding polar fluid, causing the more hydrophilic "head" regions to associate with the cytosolic and extracellular faces of the resulting bilayer.

This forms a continuous, spherical lipid bilayer containing the cellular components approximately 7 nm thick which is barely discernible with a transmission electron microscope.

The arrangement of hydrophilic and hydrophobic heads of the lipid bilayer prevents hydrophilic solutes from passively diffusing across the band of hydrophobic tail groups, allowing the cell to control the movement of these substances via transmembrane protein complexes such as pores and gates.

Integral Membrane Proteins

The cell membrane contains many integral membrane proteins which pepper the entire surface. These structures, which can be visualised by electron microscopy or fluorescence microscopy, can be found on the inside of the membrane, the outside, or through-and-through. They include synapses, desmosomes, clathrin-coated pits, caveolaes, and different structures involved in cell adhesion.

Membrane Skeleton

The cytoskeleton is found underlying the cell membrane in the cytoplasm and provides a scaffolding for membrane proteins to anchor to, as well as forming organelles which extend from the cell.

Anchoring proteins restricts them to a particular cell surface—for example, the apical surface of epithelial cells that line the vertebrate gut — and limits how far they may diffuse within the bilayer.

The cytoskeleton is able to form appendage-like organelles, such as cilia, which are covered by the cell membrane and project from the surface of the cell. The apical surfaces of the aforementioned epithelial cells are dense with finger-like projections, called microvilli, which increase cell surface area and thereby increase the absorption rate of nutrients.

Definition of Energetics

Energetics is the scientific study of energy flows under transformation. Because energy flows at all scales, from the quantum level, to the biosphere and cosmos, energetics is therefore a very broad discipline, encompassing for example thermodynamics, chemistry, biological energetics, biochemistry and ecological energetics.

Where each branch of energetics begins and ends is a topic of constant debate. For example, Lehninger contended that when the science of thermodynamics deals with energy exchanges of all types, it can be called energetics.

Aims

In general, energetics is concerned with seeking principles that accurately describe the useful and non-useful tendencies of energy flows under transformation. 'Principles' are understood here as phenomena which behave like historical invariants under multiple observations.

When some critical number of people have observed such invariance, such a principle is usually then given the status of a 'fundamental law' of science. Like in all science, whether or not a theorem or principle is considered a fundamental law appears to depend on how many people agree to such a proposition. The ultimate aim of energetics therefore is the description of fundamental laws. Philosophers of science have held that the fundamental laws of thermodynamics can be treated as the laws of energetics.

History

Energetics has a controversial history. Some authors maintain that the origins of energetics can be found in the work of the ancient Greeks, but that the mathematical formalisation began with the work of Leibniz. Lt. Col. Richard de Villamil (1928) said that Rankine formulated the Science of Energetics in his paper *Outlines of the Science of Energetics* published in the Proceedings of the Philosophical Society of Glasgow in 1855. W. Ostwald and E. Mach subsequently developed the study and in the late 1800s energetics was understood to be incompatible with the atomic view of the atom forwarded by Boltzmann's gas theory. Proof of the atom settled the dispute but not without significant damage.

In the 1920s, Lotka then attempted to build on Boltzmann's views through a mathematical synthesis of energetics with biological evolutionary theory. Lotka proposed that the selective principle of evolution was one which favoured the maximum

useful energy flow transformation. This view subsequently influenced the further development of ecological energetics, especially the work of Howard T. Odum.

De Villamil attempted to clarify the scope of energetics with respects to other branches of physics by contriving a system that divides mechanics into two branches; energetics (the science of energy) and "pure", "abstract" or "rigid" dynamics (the science of momentum). According to Villamil energetics can be mathematically characterised by scalar equations, and rigid dynamics by vectorial equations. In this division the dimensions for dynamics are *space*, time and mass, and for energetics, *length*, time and mass. This division is made according to fundamental presuppositions about the properties of bodies which can be expressed according to how one answers to following two questions:

1. Are particles rigidly fixed to together?
2. Is there any machinery for stopping moving bodies?

In Villamil's classification system, dynamics says yes to 1 and no to 2, whereas energetics says no to 1 and yes to 2. Therefore, Villamil's in system, dynamics assumes that particles are rigidly fixed together and cannot vibrate, and consequently must all be at zero temperature. The conservation of momentum is a consequence of this view, however it is considered valid only in logic and not to be a true representation of the facts. In contrast energetics does not assume that particles are rigidly fixed together, particles are therefore free to vibrate, and consequently can be at non-zero temperatures.

Principles of Energetics

As a general statement of energy flows under transformation, the principles of energetics include the first four laws of thermodynamics which seek a rigorous description. However the precise place of the laws of thermodynamics within the principles of energetics is a topic currently under debate.

If the ecologist Howard T. Odum was right, then the principles of energetics take into consideration a hierarchical ordering of

energy forms, which aims to account for the concept of energy quality, and the evolution of the universe. Albert Lehninger called these hierarchical orderings the "successive stages in the flow of energy through the biological macrocosm"

Odum proposed 3 further energetic principles and one corollary that take energy hierarchy into account. The first four principles of energetics are related to the same numbered laws of thermodynamics, and are expanded upon in that article. The final four principles are taken from the ecological energetics of H.T. Odum.

- *Zeroth Principle of Energetics:* If two thermodynamic systems A and B are in thermal equilibrium, and B and C are also in thermal equilibrium, then A and C are in thermal equilibrium.
- *First Principle of Energetics:* The increase in the internal energy of a system is equal to the amount of energy added to the system by heating, minus the amount lost in the form of work done by the system on its surroundings.
- *Second Principle of Energetics:* The total entropy of any isolated thermodynamic system tends to increase over time, approaching a maximum value.
- *Third Principle of Energetics:* As a system approaches absolute zero of temperature all processes cease and the entropy of the system approaches a minimum value or zero for the case of a perfect crystalline substance.
- *Fourth Principle of Energetics:* There seem to be two opinions on the fourth principle of energetics:
  - — The Onsager reciprocal relations are sometimes called the fourth law of thermodynamics. As the fourth law of thermodynamics Onsager reciprocal relations would constitute the fourth principle of energetics.
  - — In the field of ecological energetics H.T. Odum considered maximum power, the fourth principle of energetics. Odum also proposed the Maximum empower principle as a corollary of the maximum

power principle, and considered it to describe the propensities of evolutionary self-organisation.

- *Fifth Principle of Energetics:* The energy quality factor increases hierarchically. From studies of ecological food chains, Odum proposed that energy transformations form a hierarchical series measured by Transformity increase.

  Flows of energy develop hierarchical webs in which inflowing energies interact and are transformed by work processes into energy forms of higher quality that feedback amplifier actions, helping to maximise the power of the system."
- *Sixth Principle of Energetics:* Material cycles have hierarchical patterns measured by the emergy/mass ratio that determines its zone and pulse frequency in the energy hierarchy. M.T. Brown and V. Buranakarn write, "Generally, emergy per mass is a good indicator of recycle-ability, where materials with high emergy per mass are more recyclable".

Ecological Energetics

Ecological energetics is the quantitative study of the flow of energy through ecological systems. It aims to uncover the principles which describe the propensity of such energy flows through the trophic, or 'energy availing' levels of ecological networks. In systems ecology the principles of ecosystem energy flows or "ecosystem laws" (i.e. principles of ecological energetics) are considered formally analogous to the principles of energetics.

History

Ecological energetics appears to have grown out of the Age of enlightenment and the concerns of the physiocrats. It began in the works of Podolinksy in the late 1800s, and subsequently was developed by the soviet ecologist Stanchinskii, the Austro-American Lotka, and American limnologists, Lindeman and Hutchenson. It underwent substantial development by H.T.Odum and was applied by systems ecologists, and radiation ecologists.

Definition of Gravitational Biology

Gravitational Biology is the study of the effects gravity has on living organisms. Throughout the history of the Earth life has evolved to survive changing conditions, such as changes in the climate and habitat. The only constant factor in evolution since life first began on Earth is the force of gravity. As a consequence, all biological processes are accustomed to the ever-present force of gravity and even small variations in this force can have significant impact on the health and function of organisms.

Gravity and Life on Earth

The force of gravity on the surface of the Earth, normally denoted *g*, has remained constant in both direction and magnitude since the formation of the planet. As a result, both plant and animal life have evolved to rely upon and cope with it in various ways.

Plant use of Gravity

Plant tropisms are directional movements of a plant with respect to a directional stimulus. One such tropism is Gravitropism, or the growth or movement of a plant with respect to gravity. Plant roots grow towards the pull of gravity and away from sunlight, and shoots and stems grow against the pull of gravity and towards sunlight.

Animal Struggles with Gravity

Gravity has had an effect on the development of animal life since the first single-celled organism. The size of single biological cells is inversely proportional.

That is, in stronger gravitational fields the size of cells decreases, and in weaker gravitational fields the size of cells increases. Gravity is thus the limiting factor in the growth of individual cells.

Cells which were naturally larger than gravity alone would allow for had to develop means to protect against internal sedimentation. Several of these methods are based upon protoplasmic motion, thin and elongated shape of the cell body,

increased cytoplasmic viscosity, and a reduced range of specific gravity of cell components relative to the ground-plasma.

The effects of gravity on many-celled organisms is even more drastic. During the period when animals first evolved to survive on land some method of directed locomotion and thus a form of inner skeleton or outer skeleton would have been required to cope with the increase in the force of gravity due to the weakened upward force of buoyancy. Prior to this point, most lifeforms were small and had a worm- or jellyfish-like appearance, and without this evolutionary step would not have been able to maintain their form or move on land.

In larger terrestrial vertebrates gravitational forces influence musculoskeletal systems, fluid distribution, and hydrodynamics of the circulation.

Gravity and Life Elsewhere

Every day the realisation of space habitation becomes closer, and even today space stations exist and are home to long-term, though not yet permanent residents.

Because of this there is a growing scientific interest in how changes in the gravitational field influence different aspects of the physiology of living organisms, especially mammals since these results can normally be closely related to the expected effects on humans. All current research in this field can be classified into two groups.

The first group consists of the experiments that involve gravitational fields of less than one *g*, termed *hypogravity*. All space travel is done in hypogravity, and effective gravitational fields on any space station without Artificial gravity are on the order of hypogravity, and therefore the understanding of the effects of hypogravity on the human body is necessary for prolonged space travel and colonisation.

The second group consist of those involving gravitational fields of more than one *g*, termed *hypergravity*. For brief periods, furing take-off and landing of space craft astronauts are under the influence of hypergravity. Understanding the effects of

hypergravity are also necessary if colonisation of planets larger than the Earth is ever to take place.

Experiments

Recent experiments have proven that alterations in metabolism, immune cell function, cell division, and cell attachment all occur in the hypogravity of space. For example, after a matter of days in microgravity ($< 10^{-3}$ *g*), human immune cells were unable to differentiate into mature cells.

One of the large implications of this is that if certain cells cannot differentiate in space, organisms may not be able to reproduce successfully after exposure to zero gravity.

Scientists believe that the stress associated with space flight is responsible for the inability of some cells to differentiate. These stresses can alter metabolic activities and can disturb the chemical processes in living organisms. A specific example would be that of bone cell growth.

Microgravity impedes the development of bone cells. Bone cells must attach themselves to something shortly after development and will die if they can not. Without the downward pull of a gravitational force on these bone cells, they float around randomly and eventually die off. This suggests that the direction of gravity may give the cells clues as to where to attach themselves.

Construction and Fluid Mosaic Model

According to the fluid mosaic model of S. Jonathan Singer and Nicholson, the biological membranes can be considered as a two-dimensional liquid where all lipid and protein molecules diffuse more or less freely. This picture may be valid in the space scale of 10 nm.

However, the plasma membranes contain different structures or domains that can be classified as

(a) Protein-protein complexes;

(b) Lipid rafts,

(c) Pickets and fences formed by the actin-based cytoskeleton; and

(d) Large stable structures, such as synapses or desmosomes.

Composition

Lipids: The cell membrane consists of three classes of amphipathic lipids: phospholipids, glycolipids, and steroids. The relative composition of each depends upon the type of cell, but in the majority of cases phospholipids are the most abundant.

The fatty acid chains in phospholipids and glycolipids usually contain an even number of carbon atoms, typically between 14 and 24. The 16- and 18-carbon fatty acids are the most common. Fatty acids may be saturated or unsaturated, with the configuration of the double bonds nearly always *cis*. The length and the degree of unsaturation of fatty acids chains have a profound effect on membranes fluidity.

The entire membrane is held together via non-covalent interaction of hydrophobic tails, however the structure is quite fluid and not fixed rigidly in place. Phospholipid molecules in the cell membrane are "fluid" in the sense that they are free to diffuse and exhibit rapid lateral diffusion along the layer they are present in.

However, movement of phospholipid molecules between layers is not energetically favourable and does not occur to an appreciable extent. Lipid rafts and caveolae are examples of cholesterol-enriched microdomains in the cell membrane.

In animal cells cholesterol is normally found dispersed in varying degrees throughout cell membranes, where it confers a stiffening and strengthening effect on the membrane. It resides in the irregular spaces between the hydrophobic tails of the membrane lipids.

Cell Membrane Proteins

The cell membrane plays host to a large amount of protein which is responsible for its various activities. The amount of protein differs between species and according to function, however the typical amount in a cell membrane is 50 per cent.

These proteins are undoubtedly important to a cell: approximately a third of the genes in yeast code specifically for them, and this number is even higher in multicellular organisms.

Three groups of membrane proteins can be identified:

Type	***Description***	***Examples***
Integral proteins	Located spanning the membrane and thus have a hydrophilic cytosolic domain which interacts with internal molecules, a hydrophobic membrane-spanning domain which anchors it within the cell membrane, and a hydrophilic extracellular domain which interacts with external molecules. The hydrophobic domain consists of one, multiple, or a combination of α-helices and β sheet protein motifs. Also known as *transmembrane proteins*.	Ion channels, proton pumps, G protein-coupled receptor
Lipid anchored proteins	Located covalently bound to single or multiple lipid molecules, which hydrophobically insert into the cell membrane and anchor the protein. The protein itself is not in contact with the membrane.	G proteins
Peripheral proteins	They are attached to integral membrane proteins, or associated with peripheral regions of the lipid bilayer. These proteins tend to have only temporary interactions with biological membranes, and once reacted the molecule dissociates to carry on its work in the cytoplasm.	Some enzymes, some hormones

The cell membrane, being exposed to the outside environment, is an important site of cell communication. As such, a large variety of protein receptors and identification proteins, such as antigens, are present on the surface of the membrane.

Main Categories: Integral membrane proteins are permanently attached to the membrane. They can be defined as those proteins which require a detergent (such as SDS or Triton X-100) or some other apolar solvent to be displaced. They can be classified according to their relationship with the bilayer:

- Transmembrane proteins span the entire membrane. The transmembrane regions of the proteins are either beta-barrels or alpha-helical. The alpha-helical domains are present in all types of biological membranes including outer membranes. The beta-barrels were found only in

outer membranes of Gram-negative bacteria, cell wall of Gram-positive bacteria, and outer membranes of mitochondria and chloroplasts.

- Integral monotopic proteins are permanently attached to the membrane from only one side.

Peripheral membrane proteins are temporarily attached either to the lipid bilayer or to integral proteins by a combination of hydrophobic, electrostatic, and other non-covalent interactions. Peripheral proteins dissociate following treatment with a polar reagent, such as a solution with an elevated pH or high salt concentrations.

Integral and peripheral proteins may be post-translationally modified, with added fatty acid or prenyl chains, or GPI (glycosylphosphatidylinositol), which may be anchored in the lipid bilayer.

Classification of membrane proteins to integral and peripheral does not include some polypeptide toxins, such as colicin A or alpha-hemolysin, and certain proteins involved in apoptosis. These proteins are water-soluble but can aggregate and associate irreversibly with the lipid bilayer and form alpha-helical or beta-barrel transmembrane channels.

An alternative classification is to divide all membrane proteins to integral and *amphitropic*. The *amphitropic* are proteins that can exist in two alternative states: a water-soluble and a lipid bilayer-bound, whereas *integral* proteins can be found only in the membrane-bound state. The amphitropic protein category includes water-soluble channel-forming polypeptide toxins, which associate irreversibly with membranes, but excludes peripheral proteins that interact with other membrane proteins rather than with lipid bilayer.

Functions: In animal cells the cell membrane alone establishes a separation between interior and environment, whereas in fungi, bacteria, and plants an additional cell wall forms the outermost boundary. However, the cell wall plays mostly a mechanical support role rather than a role as a selective boundary. One of the key roles of the membrane is to maintain the cell potential. The functions of the cell membrane include, but are not limited to:

- Controlling what goes in and out of the cell.
- Anchoring of the cytoskeleton to provide shape to the cell.
- Attaching to the extracellular matrix to help group cells together in the formation of tissues.
- Transportation of particles by way of ion pumps, ion channels, and carrier proteins.
- Containing receptors that allow chemical messages to pass between cells and systems.
- Participation in enzyme activity important in such things as metabolism and immunity.

New material is incorporated into the membrane, or deleted from it, by a variety of mechanisms:

- Fusion of intracellular vesicles with the membrane not only excretes the contents of the vesicle, but also incorporates the vesicle membrane's components into the cell membrane. The membrane may form blebs that pinch off to become vesicles.
- If a membrane is continuous with a tubular structure made of membrane material, then material from the tube can be drawn into the membrane continuously.
- Although the concentration of membrane components in the aqueous phase is low (stable membrane components have low solubility in water), exchange of molecules with this small reservoir is possible.

In all cases, the mechanical tension in the membrane has an effect on the rate of exchange. In some cells, usually having a smooth shape, the membrane tension and area are interrelated by elastic and dynamical mechanical properties, and the time-dependent interrelation is sometimes called homeostasis, *area regulation* or *tension regulation*.

Peptides Bound by Membrane

There are also numerous membrane-associated peptides, some of which are non-ribosomal peptides. They can form transmembrane channels (for example, gramicidins and

peptaibols), travel across the membrane as ionophores (valinomycin and others), or associate with lipid bilayer surface, as daptomycin and other lipopeptides. These peptides are usually secreted. So, they probably should be classified as *amphitropic*, although some of them are poorly soluble in water and associate with membrane irreversibly.

This page intentionally left blank

Matters in Biology

Functions of Vitamins

A vitamin is a compound that cannot be synthesised by a given organism but is nonetheless vital to its survival or health (for example coenzymes). These compounds must be absorbed, or eaten, but typically only in trace quantities. When originally discovered by a Polish doctor, he believed them to all be basic. He therefore named them vital amines. The l was dropped to form the word vitamines.

Types of Biomolecules

A23187: A23187 is a mobile ion-carrier that forms stable complexes with divalent cations (ions with a charge of +2). A23187 is also known as Calcimycin, Calcium Ionophore, biotic A23187 and Calcium Ionophore A23187. It is produced at fermentation of Streptomyces Chartreusensis.

Actions and Uses

A23187 has antibiotic properties against gram positive bacteria and fungi. It also acts as a divalent cation ionophore, allowing

these ions to cross cell membranes, which are usually impermeable to them. A23187 is most selective for Mn^{2+}, somewhat less selective for Ca^{2+} and Mg^{2+}, much less selective for Sr^{2+}, and even less selective for Ba^{2+}.

The ionophore is used in laboratories to increase intracellular Ca^{2+} levels in intact cells. It also uncouples oxidative phosphorylation, the process cells use to synthesise Adenosine triphosphate which they use for energy. In addition, A23187 inhibits mitochondrial ATPase activity. A23187 also induces apoptosis in some cells (e.g. mouse lymphoma cell line, or S49, and Jurkat cells) and prevents it in others (e.g. cells dependent on interleukin 3 that have had the factor withdrawn).

Inex Pharmaceuticals Corporation (Canada) reported an innovative application of A23187. Inex used A23187 as a molecular tool in order to make artificial liposomes loaded with anti-cancer drugs such as Topotecan.

Commercial Availability

Commercially, A23187 is available as free acid, Ca^{2+} salt, and 4-brominated analog.

Component of Abamectin

Abamectin is a mixture of avermectins containing more than 80 per cent avermectin B1a and less than 20 per cent avermectin B1b. These two components, B1a and B1b have very similar biological and toxicological properties.

The avermectins are insecticidal or anthelmintic compounds derived from the soil bacterium Streptomyces avermitilis. Abamectin is a natural fermentation product of this bacterium. Abamectin is used to control insect and mite pests of a range of agronomic, fruit, vegetable and ornamental crops, and it is used by homeowners for control of fire ants. Doses of 50 to 200 µg/kg of ivermectin, a similar member of the avermectin family of compounds, is widely used to treat humans in the World Health Organisation onchocerciasis (river blindness) programme.

Abamectin is also known as Avermectin B1 and MK-936. Trade names include Abba, Affirm, Agri-Mek, Avid, Dynamec, Vertimec and Zephyr.

Significance of Abietic Acid

Abietic acid (also known as *abietinic acid* or *sylvic acid*), a resin acid, is the primary irritant in pine, isolated from resin (via isomerisation). It is soluble in alcohols, acetone, and ethers. It is used in lacquers, varnishes, and soaps, and for the analysis of resins and the preparation of metal resinates. It is listed in the Toxic Substances Control Act inventory.

Its ester is called an *abietate*.

Abietic acid is a weak contact allergen, however compounds resulting of its oxidation by air elicit stronger response. It is the primary irritant in pine wood and resin.

Significance of Acetic Acid

Acetic acid, also known as ethanoic acid, is an organic chemical compound with the formula CH_3COOH best recognised for giving vinegar its sour taste and pungent smell. Pure, water-free acetic acid (*glacial acetic acid*) is a colourless liquid that attracts water from the environment (hygroscopy), and freezes below 16.7 C (62 F) to a colourless crystalline solid. Acetic acid is corrosive, and its vapour causes irritation to the eyes, a dry and burning nose, sore throat and congestion to the lungs, however, it is considered a weak acid due to its limited ability to dissociate in aqueous solutions.

Acetic acid is one of the simplest carboxylic acids (the second-simplest, next to formic acid). It is an important chemical reagent and industrial chemical that is used in the production of polyethylene terephthalate mainly used in soft drink bottles; cellulose acetate, mainly for photographic film; and polyvinyl acetate for wood glue, as well as many synthetic fibres and fabrics. In households diluted acetic acid is often used in descaling agents. In the food industry acetic acid is used under the food additive code E260 as an acidity regulator.

The global demand of acetic acid is around 6.5 million tonnes per year (Mt/a), of which approximately 1.5 Mt/a is met by recycling; the remainder is manufactured from petrochemical feedstocks or from biological sources.

Nomenclature

The trivial name *acetic acid* is the most commonly used and officially preferred name by the IUPAC. This name derives from *acetum*, the Latin word for vinegar. The synonym *ethanoic acid* is a systematic name that is sometimes used in introductions to chemical nomenclature.

Glacial acetic acid is a trivial name for water-free acetic acid. Similar to the German name *Eisessig* (literally, ice-vinegar), the name comes from the ice-like crystals that form slightly below room temperature at 16.7 C (about 62 F). The most common and official abbreviation for acetic acid is AcOH or HOAc where *Ac* stands for the acetyl group $CH_3^-C(=O)^-$;. In the context of acid-base reactions the abbreviation HAc is often used where *Ac* instead stands for the acetate anion (CH_3COO^-), although this use is regarded by many as misleading. In either case, the *Ac* is not to be confused with the abbreviation for the chemical element actinium.

Acetic acid has the empirical formula CH_2O and the molecular formula $C_2H_4O_2$. The latter is often written as CH_3-COOH, CH_3COOH, or CH_3CO_2H to better reflect its structure. The ion resulting from loss of H+ from acetic acid is the *acetate* anion. The name *acetate* can also refer to a salt containing this anion, or an ester of acetic acid.

History

Vinegar is as old as civilization itself, perhaps older. Acetic acid-producing bacteria are present throughout the world, and any culture practising the brewing of beer or wine inevitably discovered vinegar as the natural result of these alcoholic beverages being exposed to air.

The use of acetic acid in chemistry extends into antiquity. In the 3rd century BC, the Greek philosopher Theophrastos described how vinegar acted on metals to produce pigments useful in art, including *white lead* (lead carbonate) and verdigris, a green mixture of copper salts including copper(II) acetate. Ancient Romans boiled soured wine in lead pots to produce a highly sweet syrup called *sapa*.

Sapa was rich in lead acetate, a sweet substance also called *sugar of lead* or *sugar of Saturn*, which contributed to lead poisoning among the Roman aristocracy. The 8th century Persian alchemist

Jabir Ibn Hayyan (Geber) concentrated acetic acid from vinegar through distillation.

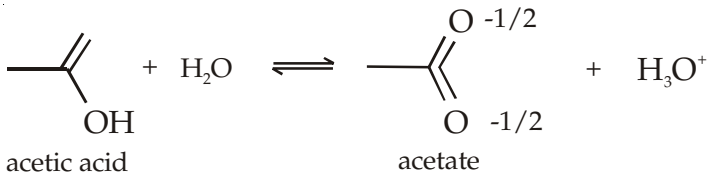
In the Renaissance, glacial acetic acid was prepared through the dry distillation of metal acetates. The 16th-century German alchemist Andreas Libavius described such a procedure, and he compared the glacial acetic acid produced by this means to vinegar. The presence of water in vinegar has such a profound effect on acetic acid's properties that for centuries many chemists believed that glacial acetic acid and the acid found in vinegar were two different substances. The French chemist Pierre Adet proved them to be identical.

In 1847, the German chemist Hermann Kolbe synthesised acetic acid from inorganic materials for the first time. This reaction sequence consisted of chlorination of carbon disulphide to carbon tetrachloride, followed by pyrolysis to tetrachloroethylene and aqueous chlorination to trichloroacetic acid, and concluded with electrolytic reduction to acetic acid.

By 1910, most glacial acetic acid was obtained from the "pyroligneous liquor" from distillation of wood. The acetic acid was isolated from this by treatment with milk of lime, and the resultant calcium acetate was then acidified with sulphuric acid to recover acetic acid. At this time Germany was producing 10,000 tons of glacial acetic acid, around 30 per cent of which was used for the manufacture of indigo dye.

Chemical Properties

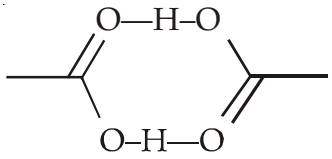
Acidity: The hydrogen (H) atom in the carboxyl group (COOH) in carboxylic acids such as acetic acid can be given off as an H^+ ion (proton), giving them their acidic character. Acetic acid is a weak, effectively monoprotic acid in aqueous solution, with a pKa value of 4.8. Its conjugate base is acetate (CH_3COO^-). A 1.0 M solution (about the concentration of domestic vinegar) has a pH of 2.4, indicating that merely 0.4 per cent of the acetic acid molecules are dissociated.



acetic acid acetate

Cyclic Dimer

Cyclic dimer of acetic acid; dashed lines represent hydrogen bonds.



The crystal structure of acetic acid shows that the molecules pair up into dimers connected by hydrogen bonds. The dimers can also be detected in the vapour at 120 C. They also occur in the liquid phase in dilute solutions in non-hydrogen-bonding solvents, and to some extent in pure acetic acid, but are disrupted by hydrogen-bonding solvents. The dissociation enthalpy of the dimer is estimated at 65.0–66.0 kJ/mol, and the dissociation entropy at 154–157 J mol^{-1} K^{-1}. This dimerisation behaviour is shared by other lower carboxylic acids.

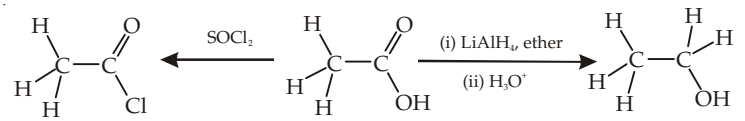
Solvent: Liquid acetic acid is a hydrophilic (polar) protic solvent, similar to ethanol and water. With a moderate dielectric constant of 6.2, it can dissolve not only polar compounds such as inorganic salts and sugars, but also non-polar compounds such as oils and elements such as sulphur and iodine. It readily mixes with many other polar and non-polar solvents such as water, chloroform, and hexane. This dissolving property and miscibility of acetic acid makes it a widely used industrial chemical.

Chemical Reactions

Acetic acid is corrosive to many metals including iron, magnesium, and zinc, forming hydrogen gas and metal salts called acetates. Aluminium, when exposed to oxygen, forms a thin layer of aluminium oxide on its surface which is relatively resistant, so that aluminium tanks can be used to transport acetic acid. Metal acetates can also be prepared from acetic acid and an appropriate base, as in the popular "baking soda + vinegar" reaction. With the notable exception of chromium(II) acetate, almost all acetates are soluble in water.

$$Mg(s) + 2\ CH_3COOH(aq) \rightarrow (CH_3COO)_2Mg(aq) + H_2(g)$$

$$NaHCO_3(s) + CH_3COOH(aq) \rightarrow CH_3COONa(aq) + CO_2(g) + H_2O(l)$$



Acetic acid undergoes the typical chemical reactions of a carboxylic acid, such producing ethanoic acid when reacting with alkalis, producing a metal ethanoate when reacted with a metal, and producing a metal ethanoate, water and carbon dioxide when reacting with carbonates and hydrogencarbonates. Most notable of all its reactions is the formation of ethanol by reduction, and formation of derivatives such as acetyl chloride via nucleophilic acyl substitution. Other substitution derivatives include acetic anhydride; this anhydride is produced by loss of water from two molecules of acetic acid. Esters of acetic acid can likewise be formed via Fischer esterification, and amides can also be formed. When heated above 440 C, acetic acid decomposes to produce carbon dioxide and methane, or to produce ketene and water.

Detection

Acetic acid can be detected by its characteristic smell. A colour reaction for salts of acetic acid is iron(III) chloride solution, which results in a deeply red colour that disappears after acidification. Acetates when heated with arsenic trioxide form cacodyl oxide, which can be detected by its malodourous vapours.

Biochemistry

The acetyl group, derived from acetic acid, is fundamental to the biochemistry of virtually all forms of life. When bound to coenzyme A it is central to the metabolism of carbohydrates and fats. However, the concentration of free acetic acid in cells is kept at a low level to avoid disrupting the control of the pH of the cell contents. Unlike some longer-chain carboxylic acids (the fatty acids), acetic acid does not occur in natural triglycerides. However, the artificial triglyceride triacetin (glycerin triacetate) is a common food additive, and is found in cosmetics and topical medicines.

Acetic acid is produced and excreted by certain bacteria, notably the Acetobacter genus and Clostridium acetobutylicum. These bacteria are found universally in foodstuffs, water, and soil, and acetic acid is produced naturally as fruits and some other foods spoil. Acetic acid is also a component of the vaginal lubrication of humans and other primates, where it appears to serve as a mild anti-bacterial agent.

Production: Acetic acid is produced both synthetically and by bacterial fermentation. Today, the biological route accounts for only about 10 per cent of world production, but it remains important for vinegar production, as many of the world food purity laws stipulate that vinegar used in foods must be of biological origin.

About 75 per cent of acetic acid made for use in the chemical industry is made by methanol carbonylation, explained below. Alternative methods account for the rest.

Total worldwide production of virgin acetic acid is estimated at 5 Mt/a (million tonnes per year), approximately half of which is produced in the United States. European production stands at approximately 1 Mt/a and is declining, and 0.7 Mt/a is produced in Japan.

Another 1.5 Mt are recycled each year, bringing the total world market to 6.5 Mt/a. The two biggest producers of virgin acetic acid are Celanese and BP Chemicals. Other major producers include Millennium Chemicals, Sterling Chemicals, Samsung, Eastman, and Svensk Etanolkemi.

Methanol Carbonylation: Most virgin acetic acid is produced by methanol carbonylation. In this process, methanol and carbon monoxide react to produce acetic acid according to the chemical equation:

$$CH_3OH + CO \rightarrow CH_3COOH$$

The process involves iodomethane as an intermediate, and occurs in three steps. A catalyst, usually a metal complex, is needed for the carbonylation (step 2).

(1) $CH_3OH + HI \rightarrow CH_3I + H_2O$

(2) $CH_3I + CO \rightarrow CH_3COI$

(3) $CH_3COI + H_2O \rightarrow CH_3COOH + HI$

By altering the process conditions, acetic anhydride may also be produced on the same plant. Because both methanol and carbon monoxide are commodity raw materials, methanol carbonylation long appeared to be an attractive method for acetic acid production. Henry Drefyus at British Celanese developed a methanol carbonylation pilot plant as early as 1925. However, a lack of practical materials that could contain the corrosive reaction mixture at the high pressures needed (200 atm or more) discouraged commercialisation of these routes for some time. The first commercial methanol carbonylation process, which used a cobalt catalyst, was developed by German chemical company BASF in 1963. In 1968, a rhodium-based catalyst (*cis* $^{-}[Rh(CO)_2I_2]^{-}$) was discovered that could operate efficiently at lower pressure with almost no by-products.

The first plant using this catalyst was built by US chemical company Monsanto in 1970, and rhodium-catalysed methanol carbonylation became the dominant method of acetic acid production. In the late 1990s, the chemicals company BP Chemicals commercialised the Cativa catalyst ($[Ir(CO)_2I_2]^{-}$), which is promoted by ruthenium. This iridium-catalysed process is greener and more efficient and has largely supplanted the Monsanto process, often in the same production plants.

Acetaldehyde Oxidation

Prior to the commercialisation of the Monsanto process, most acetic acid was produced by oxidation of acetaldehyde. This remains the second most important manufacturing method, although it is uncompetitive with methanol carbonylation. The acetaldehyde may be produced via oxidation of butane or light naphtha, or by hydration of ethylene.

When butane or light naphtha is heated with air in the presence of various metal ions, including those of manganese, cobalt and chromium; peroxides form and then decompose to produce acetic acid according to the chemical equation

$$2\ C_4H_{10} + 5\ O_2 \rightarrow 4\ CH_3COOH + 2\ H_2O$$

Typically, the reaction is run at a combination of temperature and pressure designed to be as hot as possible while still keeping

the butane a liquid. Typical reaction conditions are 150 C and 55 atm. Several side products may also form, including butanone, ethyl acetate, formic acid, and propionic acid. These side products are also commercially valuable, and the reaction conditions may be altered to produce more of them if this is economically useful. However, the separation of acetic acid from these by-products adds to the cost of the process.

Under similar conditions and using similar catalysts as are used for butane oxidation, acetaldehyde can be oxidised by the oxygen in air to produce acetic acid

$$2\ CH_3CHO + O_2 \rightarrow 2\ CH_3COOH$$

Using modern catalysts, this reaction can have an acetic acid yield greater than 95 per cent. The major side products are ethyl acetate, formic acid, and formaldehyde, all of which have lower boiling points than acetic acid and are readily separated by distillation.

Fermentation

Oxidative Fermentation: For most of human history, acetic acid, in the form of vinegar, has been made by bacteria of the genus Acetobacter. Given sufficient oxygen, these bacteria can produce vinegar from a variety of alcoholic foodstuffs. Commonly used feeds include apple cider, wine, and fermented grain, malt, rice, or potato mashes. The overall chemical reaction facilitated by these bacteria is:

$$C_2H_5OH + O_2 \rightarrow CH_3COOH + H_2O$$

A dilute alcohol solution inoculated with *Acetobacter* and kept in a warm, airy place will become vinegar over the course of a few months. Industrial vinegar-making methods accelerate this process by improving the supply of oxygen to the bacteria.

The first batches of vinegar produced by fermentation probably followed errors in the winemaking process. If must is fermented at too high a temperature, acetobacter will overwhelm the yeast naturally occurring on the grapes. As the demand for vinegar for culinary, medical, and sanitary purposes increased, vintners quickly learned to use other organic materials to produce vinegar

in the hot summer months before the grapes were ripe and ready for processing into wine. This method was slow, however, and not always successful, as the vintners did not understand the process.

One of the first modern commercial processes was the "fast method" or "German method", first practised in Germany in 1823. In this process, fermentation takes place in a tower packed with wood shavings or charcoal. The alcohol-containing feed is trickled into the top of the tower, and fresh air supplied from the bottom by either natural or forced convection. The improved air supply in this process cut the time to prepare vinegar from months to weeks.

Most vinegar today is made in submerged tank culture, first described in 1949 by Otto Hromatka and Heinrich Ebner. In this method, alcohol is fermented to vinegar in a continuously stirred tank, and oxygen is supplied by bubbling air through the solution. Using this method, vinegar of 15 per cent acetic acid can be prepared in only 2–3 days.

Anaerobic Fermentation

Some species of anaerobic bacteria, including several members of the genus Clostridium, can convert sugars to acetic acid directly, without using ethanol as an intermediate. The overall chemical reaction conducted by these bacteria may be represented as:

$$C_6H_{12}O_6 \rightarrow 3\ CH_3COOH$$

More interestingly from the point of view of an industrial chemist, many of these acetogenic bacteria can produce acetic acid from one-carbon compounds, including methanol, carbon monoxide, or a mixture of carbon dioxide and hydrogen:

$$2\ CO_2 + 4\ H_2 \rightarrow CH_3COOH + 2\ H_2O$$

This ability of *Clostridium* to utilise sugars directly, or to produce acetic acid from less costly inputs, means that these bacteria could potentially produce acetic acid more efficiently than ethanol-oxidisers like *Acetobacter*. However, *Clostridium* bacteria are less acid-tolerant than *Acetobacter*. Even the most acid-tolerant *Clostridium* strains can produce vinegar of only a few

per cent acetic acid, compared to some *Acetobacter* strains that can produce vinegar of up to 20 per cent acetic acid. At present, it remains more cost-effective to produce vinegar using *Acetobacter* than to produce it using *Clostridium* and then concentrating it. As a result, although acetogenic bacteria have been known since 1940, their industrial use remains confined to a few niche applications.

Applications

Acetic acid is a chemical reagent for the production of many chemical compounds. The largest single use of acetic acid is in the production of vinyl acetate monomer, closely followed by acetic anhydride and ester production. The volume of acetic acid used in vinegar is comparatively small.

Vinyl Acetate Monomer: The major use of acetic acid is for the production of vinyl acetate monomer (VAM). This application consumes approximately 40 to 45 per cent of the world's production of acetic acid. The reaction is of ethylene and acetic acid with oxygen over a palladium catalyst.

$$2\ H_3C\text{-}COOH + 2\ C_2H_4 + O_2 \rightarrow 2\ H_3C\text{-}CO\text{-}O\text{-}CH{=}CH_2 + 2\ H_2O$$

Vinyl acetate can be polymerised to polyvinyl acetate or to other polymers, which are applied in paints and adhesives.

Acetic Anhydride: The condensation product of two molecules of acetic acid is acetic anhydride. The worldwide production of acetic anhydride is a major application, and uses approximately 25 to 30 per cent of the global production of acetic acid. Acetic anhydride may be produced directly by methanol carbonylation bypassing the acid, and Cativa plants can be adapted for anhydride production.

acetic acid acetic acid acetic anhydride

Acetic anhydride is a strong acetylation agent. As such, its major application is for cellulose acetate, a synthetic textile also used for photographic film. Acetic anhydride is also a reagent for the production of aspirin, heroin, and other compounds.

Vinegar: In the form of vinegar, acetic acid solutions (typically 5 to 18 per cent acetic acid, with the percentage usually calculated by mass) are used directly as a condiment, and also in the pickling of vegetables and other foodstuffs. Table vinegar tends to be more diluted (5 to 8 per cent acetic acid), while commercial food pickling generally employs more concentrated solutions. The amount of acetic acid used as vinegar on a worldwide scale is not large, but historically, this is by far the oldest and most well-known application.

Use as Solvent: Glacial acetic acid is an excellent polar protic solvent. It is frequently used as a solvent for recrystallisation to purify organic compounds. Pure molten acetic acid is used as a solvent in the production of terephthalic acid (TPA), the raw material for polyethylene terephthalate (PET).

Although currently accounting for 5–10 per cent of acetic acid use worldwide, this specific application is expected to grow significantly in the next decade, as PET production increases.

Acetic acid is often used as a solvent for reactions involving carbocations, such as Friedel-Crafts alkylation. For example, one stage in the commercial manufacture of synthetic camphor involves a Wagner-Meerwein rearrangement of camphene to isobornyl acetate; here acetic acid acts both as a solvent and as a nucleophile to trap the rearranged carbocation. Acetic acid is the solvent of choice when reducing an aryl nitro-group to an aniline using palladium-on-carbon.

Glacial acetic acid is used in analytical chemistry for the estimation of weakly alkaline substances such as organic amides. Glacial acetic acid is a much weaker base than water, so the amide behaves as a strong base in this medium. It then can be titrated using a solution in glacial acetic acid of a very strong acid, such as perchloric acid.

Other Applications

Dilute solutions of acetic acids are also used for their mild acidity. Examples in the household environment include the use in a stop bath during the development of photographic films, and in descaling agents to remove limescale from taps and kettles. The acidity is also used for treating the sting of the box jellyfish by

disabling the stinging cells of the jellyfish, preventing serious injury or death if applied immediately, and for treating outer ear infections in people in preparations such as Vosol. Equivalently, acetic acid is used as a spray-on preservative for livestock silage, to discourage bacterial and fungal growth. Glacial acetic acid is also used as a wart and verruca remover. Several organic or inorganic salts are produced from acetic acid, including:

- Sodium acetate, used in the textile industry and as a food preservative.
- Copper(II) acetate, used as a pigment and a fungicide.
- Aluminium acetate and iron(II) acetate—used as mordants for dyes.
- Palladium(II) acetate, used as a catalyst for organic coupling reactions such as the Heck reaction.

Substituted acetic acids produced include:

- Monochloroacetic acid (MCA), dichloroacetic acid (considered a by-product), and trichloroacetic acid. MCA is used in the manufacture of indigo dye.
- Bromoacetic acid, which is esterified to produce the reagent ethyl bromoacetate.
- Trifluoroacetic acid, which is a common reagent in organic synthesis.

Amounts of acetic acid used in these other applications together (apart from TPA) account for another 5–10 per cent of acetic acid use worldwide. These applications are, however, not expected to grow as much as TPA production.

Safety: Concentrated acetic acid is corrosive and must therefore be handled with appropriate care, since it can cause skin burns, permanent eye damage, and irritation to the mucous membranes. These burns or blisters may not appear until several hours after exposure. Latex gloves offer no protection, so specially resistant gloves, such as those made of nitrile rubber, should be worn when handling the compound. Concentrated acetic acid can be ignited with some difficulty in the laboratory. It becomes a flammable risk if the ambient temperature exceeds 39 C (102 F), and can form explosive mixtures with air above this temperature (explosive limits: 5.4–16 per cent).

Solutions at more than 25 per cent acetic acid are handled in a fume hood because of the pungent, corrosive vapour. Dilute acetic acid, in the form of vinegar, is harmless. However, ingestion of stronger solutions is dangerous to human and animal life. It can cause severe damage to the digestive system, and a potentially lethal change in the acidity of the blood.

Due to incompatibilities, it is recommened to keep acetic acid away from chromic acid, ethylene glycol, nitric acid, perchloric acid, permanganates, peroxides and hydroxyls.

Acetylcholine (Ach)

The chemical compound acetylcholine, often abbreviated as ACh, was the first neurotransmitter to be identified. It is a chemical transmitter in both the peripheral nervous system (PNS) and central nervous system (CNS) in many organisms including humans. Acetylcholine is the neurotransmitter in all autonomic ganglia. In layman's terms, it is a chemical that allows neurons to communicate with each other within humans and other organisms.

Chemistry

Acetylcholine is an ester of acetic acid and choline with chemical formula $CH_3COOCH_2CH_2N^+(CH_3)_3$. This structure is reflected in the systematic name, *2-acetoxy-N,N,N-trimethylethanaminium*.

Acetylcholine (ACh) was first identified in 1914 by Henry Hallett Dale for its actions on heart tissue. It was confirmed as a neurotransmitter by Otto Loewi who initially gave it the name vagusstoff because it was released from the vagus nerve. Both received the 1936 Nobel Prize in Physiology or Medicine for their work.

Later work showed that when acetylcholine binds to acetylcholine receptors on skeletal muscle fibres, it opens ligand gated sodium channels in the membrane. Sodium ions then enter the muscle cell, stimulating muscle contraction.

Acetylcholine, while inducing contraction of skeletal muscles, instead induces decreased contraction in cardiac muscle fibres. This distinction is attributed to differences in receptor structure

between skeletal and cardiac fibres. Acetylcholine is also used in the brain, where it tends to cause excitatory actions. The glands that receive impulses from the parasympathetic part of the autonomic nervous system are also stimulated in the same way.

Synthesis and Degradation

Acetylcholine is synthesised in certain neurons by the enzyme choline acetyltransferase from the compounds choline and acetyl-CoA. Organic mercurial compounds have a high affinity for sulphydryl groups, which causes dysfunction of the enzyme choline acetyl transferase. This inhibition may lead to acetylcholine deficiency, and can have consequences on motor function.

Normally, the enzyme acetylcholinesterase converts acetylcholine into the inactive metabolites choline and acetate. This enzyme is abundant in the synaptic cleft, and its role in rapidly clearing free acetylcholine from the synapse is essential for proper muscle function.

The devastating effects of organophosphate-containing nerve agents (e.g. Sarin gas) are due to their irreversible inactivation of this enzyme. The resulting accumulation of acetylcholine causes continuous stimulation of the muscles, glands and central nervous system; victims commonly die of suffocation as they cannot contract their diaphragm. Other organophosphates and some carbamates are effective insecticides because they inhibit acetylcholinasterase in insects. On the other hand, since a shortage of acetylcholine in the brain has been associated with Alzheimer's disease, some drugs that inhibit acetylcholinesterase are used in the treatment of that disease. A recent study has shown that THC is one such drug, effective at reducing the formation of characteristic neurofibrillary tangles and amyloid beta plaques.

Release Sites

- Acetylcholine is released in the autonomic nervous system:
  - — Pre- and post-ganglionic parasympathetic neurons.
  - — Preganglionic sympathetic neurons (and also post-ganglionic sudomotor neurons, i.e., the ones that control sweating).

Botulin acts by suppressing the release of acetylcholine; where the venom from a black widow spider has the reverse effect.

- All pre-ganglionic autonomic fibres including:
  - — All pre-ganglionic sympathetic fibres.
  - — All pre-ganglionic parasympathetic fibres.
  - — Preganglionic sympathetic fibres to suprarenal medulla, the modified sympathetic ganglion. On stimulation by acetylcholine, it releases adrenaline and noradrenaline.
- All post-ganglionic parasympathetic fibres.
- Some post-ganglionic sympathetic fibres.
  - — Secretory fibres to sweat glands.
  - — Vasodilator fibres to blood vessels of skeletal muscles.

Pharmacology

There are two main classes of acetylcholine receptor (AChR), nicotinic acetylcholine receptors (nAChR) and muscarinic acetylcholine receptors (mAChR). They are named for the ligands used to discover the receptors.

Nicotinic AChRs are ionotropic receptors permeable to sodium, potassium, and chloride ions. They are stimulated by nicotine and acetylcholine and blocked by curare. Most peripheral AChRs are nicotinic, such as those on the heart and blood vessels or at the neuromuscular junction. They are also found in wide distribution through the brain, but in relatively low numbers.

Muscarinic receptors are metabotropic and affect neurons over a longer time frame. They are stimulated by muscarine and acetylcholine, and blocked by atropine. Muscarinic receptors are found in both the central nervous system and the peripheral nervous system, in heart, lungs, upper GI tract and sweat glands.

Extracts from the plant included this compound, and its action on muscarinic AChRs that increased pupil size was used for attractiveness in many European cultures in the past. Now, ACh is sometimes used during cataract surgery to produce rapid constriction of the pupil. It must be administered intraocularly because corneal cholinesterase metabolises topically administered

ACh before it can diffuse into the eye. It is sold by the trade name Miochol-E (CIBA Vision). Similar drugs are used to induce mydriasis (dilation of the pupil) in cardiopulmonary resuscitation and many other situations.

The disease myasthenia gravis, characterised by muscle weakness and fatigue, occurs when the body inappropriately produces antibodies against acetylcholine receptors, and thus inhibits proper acetylcholine signal transmission. Over time the motor end plate is destroyed.

Drugs that competitively inhibit acetylcholinesterase (e.g., neo-stigmine or physostigmine) are effective in treating this disorder.

They allow endogenously released acetylcholine more time to interact with its respective receptor before being inactivated by acetylcholinesterase in the gap junction.

Blocking, hindering or mimicking the action of acetylcholine has many uses in medicine. Cholinesterase inhibitors, an example of enzyme inhibitors, increase the action of acetylcholine by delaying its degradation; some have been used as nerve agents (Sarin and VX nerve gas) or pesticides (organophosphates and the carbamates).

Clinically they are used to reverse the action of muscle relaxants, to treat myasthenia gravis and in Alzheimer's disease (rivastigmine, which increases cholinergic activity in the brain).

ACh Receptor Agonists

Direct Acting:

- Acetylcholine
- Bethanechol
- Carbachol
- Cevimeline
- Pilocarpine
- Suberylcholine

Indirect Acting (Reversible): Reversibly inhibit the enzyme acetylcholinesterase (which breaks down acetylcholine), thereby increasing acetylcholine levels:

- Ambenomium
- Donepezil
- Edrophonium
- Galantamine
- Neostigmine
- Physostigmine
- Pyridostigmine
- Rivastigmine
- Tacrine
- Carbamate Insecticides (Aldicarb)

Indirect Acting (Irreversible): Semi-permanently inhibit the enzyme acetylcholinesterase:

- Echothiophate
- Isoflurophate
- Organophosphate Insecticides (Malathion, Parathion, Azinphos Methyl, Chlorpyrifos, among others)

Reactivation of Acetylcholine Esterase:

- Pralidoxime

ACh Receptor Antagonists

Antimuscarinic Agents:

- Atropine
- Ipratropium
- Scopolamine
- Tiotropium

Ganglionic Blockers:

- Mecamylamine
- Hexamethonium
- Nicotine (in high doses)
- Trimethaphan

Neuromuscular Blockers:

- Atracurium
- Cisatracurium
- Doxacurium
- Metocurine
- Mivacurium
- Pancuronium
- Rocuronium
- Succinylcholine
- Tubovurarine
- Vecuronium

Others/Uncategorised/Unknown:

- Surugatoxin
- Organophosphates block the breakdown of acetylcholine. Tetrachlorvinphos and other organophosphates operate by blocking acetylcholinesterase, which is an enzyme that breaks down acetylcholine.

Neuromodulatory Effects

In the central nervous system, ACh has a variety of effects as a neuromodulator.

Given its prominent role in learning, ACh is naturally involved with synaptic plasticity. It has been shown to enhance the amplitude of synaptic potentials following long-term potentiation in many regions, including the dentate gyrus, CA1, piriform cortex, and neo-cortex.

This effect most likely occurs either through enhancing currents through NMDA receptors or indirectly by suppressing adaptation. The suppression of adaptation has been shown in brain slices of regions CA1, cingulate cortex, and piriform cortex as well as *in vivo* in cat somatosensory and motor cortex by decreasing the conductance of voltage-dependent M currents and Ca^{2+}-dependent K^{+} currents.

Acetylcholine also has other effects on excitability of neurons. Its presence causes a slow depolarisation by blocking a tonically

active K^+ current, which increases neuronal excitability. Paradoxically, it increases spiking activity in inhibitory interneurons while decreasing strength of synaptic transmission from those cells.

This decrease in synaptic transmission also occurs selectively at some excitatory cells: for instance, it has an effect on intrinsic and associational fibres in layer Ib of piriform cortex, but has no effect on afferent fibres in layer Ia. Similar laminar selectivity has been shown in dentate gyrus and region CA1 of the hippocampus. One theory to explain this paradox interprets Acetylcholine neuromodulation in the neo-cortex as modulating the estimate of expected uncertainty, acting counter to Norepinephrine (NE) signals for unexpected uncertainty. Both would then decrease synaptic transition strength, but ACh would then be needed to counter the effects of NE in learning a signal understood to be noisy.

Actin and its Genetics

Actin is a globular structural, 42-47 kDa protein found in many eukaryotic cells, with concentrations of over 100 μM. It is also one of the most highly conserved proteins, differing by no more than 5 per cent in species as diverse as algae and humans. It is the monomeric subunit of microfilaments, one of the three major components of the cytoskeleton.

Genetics

Actin is one of the most highly conserved proteins, with 80.2 per cent sequence conservation at the gene level between Homo sapiens and Saccharomyces cerevisiae (a species of yeast), and 95 per cent conservation of the primary structure of the protein product.

Although most yeasts have only a single actin gene, higher eukaryotes generally express several isoforms of actin encoded by a family of related genes. Mammals have at least six actins, which are divided into three classes (alpha, beta and gamma) according to their isoelectric point.

Alpha actins are generally found in muscle, whereas beta and gamma isoforms are prominent in non-muscle cells. Although there are small differences in sequence and properties between the isoforms, all actins assemble into microfilaments and are essentially identical in the majority of tests performed in vitro.

The typical actin gene has an approximately 100 nucleotide 5' UTR, a 1200 nucleotide translated region, and a 200 nucleotide 3' UTR. The majority of actin genes are interrupted by introns, with up to 6 introns in any of 19 well-characterised locations. The high conservation of the family makes actin the favoured model for studies comparing the introns-early and introns-late models of intron evolution.

All non-spherical prokaryotes appear to possess genes such as MreB which encode homologues of actin; these genes are required for the cell's shape to be maintained. The plasmid-derived gene ParM encodes an actin-like protein whose polymerised form is dynamically unstable, and appears to partition the plasmid DNA into the daughter cells during cell division by a mechanism analogous to that employed by microtubules in eukaryotic mitosis.

Functions

Actin has two main functions in cells. In all cells it forms the thinnest part of the cytoskeleton, which allows motility, while in muscle cells it also forms the contractile apparatus.

Microfilaments

Individual subunits of actin are known as globular actin (G-actin). G-actin subunits assemble into long filamentous polymers called F-actin. Two parallel F-actin strands twist around each other in a helical formation, giving rise to microfilaments of the cytoskeleton. Microfilaments measure approximately 7 nm in diameter with a loop of the helix repeating every 37nm.

Polarity

The polarity of an actin filament can be determined by decorating the microfilament with myosin "S1" fragments, creating barbed (+) and pointed (-) ends on the filament. An S1 fragment is composed of the head and neck domains of myosin II.

Muscle Contraction

In muscle, actin forms thin filaments, which together with the motor protein myosin (which forms thick filaments), are arranged into actomyosin myofibrils. These provide the mechanism of contraction. Using the hydrolysis of ATP for energy, myosin heads extend to bind with actin filaments.

In contractile bundles, the actin-bundling protein alpha-actinin separates each thin filament by ~35 nm. This increase in distance allows thick filaments to fit in between and interact, enabling deformation or contraction. In deformation, one end of myosin is bound to the plasma membrane while the other end "walks" towards the plus end of the actin filament.

This pulls the membrane into a different shape relative to the cell cortex. For contraction, the myosin molecule is usually bound to two separate filaments and both ends simultaneously "walk" towards their filament's plus end, sliding the actin filaments closer to each other. This results in the shortening, or contraction, of the actin bundle (but not the filament). This mechanism is responsible for muscle contraction and cytokinesis, the division of one cell into two.

History

Actin was first observed experimentally in 1887 by W.D. Halliburton, who extracted a protein from muscle which 'coagulated' preparations of myosin, and which he dubbed "myosin-ferment" . However, Halliburton was unable to further characterise his findings and the discovery of actin is generally credited instead to Bruno F. Straub, a young biochemist working in Albert Szent-Gyorgyi's laboratory at the Institute of Medical Chemistry at the University of Szeged, Hungary.

In 1942, Straub developed a novel technique for extracting muscle protein that allowed him to isolate substantial amounts of relatively pure actin. Straub's method is essentially the same as that used in laboratories today.

Szent-Gyorgyi had previously described the more viscous form of myosin produced by slow muscle extractions as 'activated' myosin, and since Straub's protein produced the activating effect, it was dubbed 'actin'.

The hostilities of World War II meant that Szent-Gyorgyi and Straub were unable to publish the work in Western scientific journals; it became well-known in the West only in 1945, when it was published as a supplement to the Acta Physiologica Scandinavica.

Straub continued to work on actin and in 1950 reported that actin contains bound ATP and that, during polymerisation of the protein into microfilaments, the nucleotide is hydrolysed to ADP and inorganic phosphate (which remain bound in the microfilament). Straub suggested that the transformation of ATP-bound actin to ADP-bound actin played a role in muscular contraction. In fact this is only true in smooth muscle, and was not experimentally supported until 2001.

The crystal structure of G-actin was solved in 1990 by Kabsch and colleagues. In the same year a model for F-actin was proposed by Holmes and colleagues. The model was derived by fitting a helix of G-actin structures according to low-resolution fibre diffraction data from the filament. Several models of the filament have been proposed since. However there is still no high-resolution x-ray structure of F-actin.

The Listeria bacteria uses the cellular machinery to move around inside the host cell: it induces directed polymerisation of actin by the ActA transmembrane protein, thus pushing the bacterial cell around.

Cellular Constituent

Biological matter or biological material refers to the unique, highly organised substances of which cellular life is composed of, for instance membranes, proteins, and nucleic acids. They may also be called cellular components.

Most biological matter has the characteristics of soft matter, being governed by relatively small energies.

All known life is made of biological matter. To be differentiated from other theoretical or fictional life forms, such life may be called carbon-based, *cellular*, *organic*, *biological*, or even simply *living*—as some definitions of life exclude alternative biochemistry.

Variations of Tissue

Biological tissue is a collection of interconnected cells that perform a similar function within an organism.

The study of tissue is known as histology, or, in connection with disease, histopathology.

The classical tools for studying the tissues are the wax block, the tissue stain, and the optical microscope, though developments in electron microscopy, immunofluorescence, and frozen sections have all added to the sum of knowledge in the last couple of decades.

With these tools, the classical appearances of the tissues can be examined in health and disease, enabling considerable refinement of clinical diagnosis and prognosis.

There are four basic types of tissue in the body of all animals, including the human body and lower multicellular organisms such as insects. These compose all the organs, structures and other contents.

- *Epithelium:* Tissues composed of layers of cells that cover organ surfaces such as surface of the skin and inner lining of digestive tract: the tissues that serve for protection, secretion, and absorption.
- *Connective Tissue:* As the name suggests, connective tissue holds everything together. Some people consider blood a connective tissue. It should be noted that blood is not 'considered' connective tissue, it is connective tissue (i.e. cells of connective tissue (blood) separated by an inorganic material (plasma). Plasma is the extracellular matrix that includes everything but the red/white blood cells. These tissues contain extensive extracellular matrix.
- *Muscle Tissue:* Muscle cells contain contractile filaments that move past each other and change the size of the cell. Muscle tissue also is separated into three distinct categories: visceral or smooth muscle, which is found in the inner linings of organs; skeletal muscle, which is found attached to bone in order for mobility to take place; and cardiac muscle which is found in the heart.

- *Nervous Tissue:* Cells forming the brain, spinal cord and peripheral nervous system.

Plant Tissues

Examples of tissue in other multicellular organisms are vascular tissue in plants, such as xylem and phloem. Plant tissues are categorised broadly into three tissue systems: the epidermis, the ground tissue, and the vascular tissue. Together they are often referred to as biomass.

- *Epidermis:* Cells forming the outer surface of the leaves and of the young plant body.
- *Vascular Tissue:* The primary components of vascular tissue are the xylem and phloem. These transport fluid and nutrients internally.
- *Ground Tissue:* Ground tissue is less differentiated than other tissues. Ground tissue manufactures nutrients by photosynthesis and stores reserve nutrients.

Types of Biomolecules: A diverse range of biomolecules exist, including:

- Small molecules:
  - — Lipid, Phospholipid, Glycolipid, Sterol
  - — Vitamin
  - — Hormone, Neurotransmitter
  - — Carbohydrate, Sugar
  - — Disaccharide
- Monomers:
  - — Amino acid
  - — Nucleotide
  - — Phosphate
  - — Monosaccharide
- Polymers:
  - — Peptide, Oligopeptide, Polypeptide, Protein
  - — Nucleic acid, i.e. DNA, RNA
  - — Oligosaccharide, Polysaccharide
- Macromolecules:

— Prion

Nucleosides and Nucleotides

Nucleosides are molecules formed by attaching a nucleobase to a ribose ring. Examples of these include cytidine, uridine, adenosine, guanosine, thymidine and inosine. Nucleosides can be phosphorylated by specific kinases in the cell, producing nucleotides, which are the molecular building blocks of DNA (deoxyribonucleic acid) and RNA (ribonucleic acid).

Types of Saccharides

Monosaccharides are carbohydrates in the form of simple sugars. Examples of monosaccharides are the hexoses glucose, fructose, and galactose and pentoses, ribose, and deoxyribose.

Disaccharides are formed from two monosaccharides joined together. Examples of disaccharides include sucrose, maltose, and lactose. Monosaccharides and disaccharides are sweet, water soluble, and crystalline.

Polysaccharides are polymerised monosaccharides, complex, unsweet carbohydrates. Examples are starch, cellulose, and glycogen. They are generally large and often have a complex, branched, connectivity. They are insoluble in water and do not form crystals. Shorter polysaccharides, with 2-15 monomers, are sometimes known as oligosaccharides.

Mechanism of Actinomycin

Actinomycin is any of a class of polypeptide antibiotics isolated from soil bacteria of the genus Streptomyces.

Mechanism

Actinomycin-D is primarily used as an investigative tool in cell biology to inhibit transcription. It does this by binding DNA at the transcription initiation complex and preventing elongation by RNA polymerase.

As it can bind DNA duplexes, it can also interfere with DNA replication, although other chemicals such as hydroxyurea are better suited for use in the laboratory as inhibitors of DNA synthesis.

Clinical Use

As Chemotherapy: Actinomycin-D is marketed under the trade name Dactinomycin. Actinomycin-D is one of the older chemotherapy drugs which has been used in therapy for many years.

It is a clear, yellow liquid which is administered intravenously and most commonly used in treatment of a variety of cancers, including:

- Gestational trophoblastic neo-plasia
- Wilms' tumour
- Rhabdomyosarcoma

As an Antibiotic

It was the first antibiotic shown to have anti-cancer activity, but is not normally used as such, as it is highly toxic, causing damage to genetic material.

It was the first antibiotic ever isolated by Selman Waksman.

Research Use

Actinomycin-D and its fluorescent derivative, 7-amino-actinomycin D, are used as stains in microscopy and flow cytometry applications. The affinity of these stains compounds for GC-rich regions of DNA strands makes them excellent markers for DNA.

Uses and Effects of Adenosine

Adenosine is a nucleoside comprised of adenine attached to a ribose (ribofuranose) moiety via a β-N_9-glycosidic bond.

Adenosine plays an important role in biochemical processes, such as energy transfer—as adenosine triphosphate (ATP) and adenosine diphosphate (ADP)—as well as in signal transduction as cyclic adenosine monophosphate, cAMP. It is also an inhibitory neurotransmitter, believed to play a role in promoting sleep and suppressing arousal, with levels increasing with each hour an organism is awake.

Pharmacological Effects

Adenosine is an endogenous purine nucleoside that modulates many physiologic processes. Cellular signalling by adenosine occurs through four known adenosine receptor subtypes (A1, A2A, A2B, and A3), all of which are seven transmembrane spanning G-protein coupled receptors. These four receptor subtypes are further classified based on their ability to either stimulate or inhibit adenylate cyclase activity.

The A2A and A2B receptors couple to G ά s and mediate the stimulation of adenylate cyclase, while the A1 and A3 adenosine receptors couple to G ά i which inhibits adenylate cyclase activity. Additionally, A1 receptors couple to G ά o, which has been reported to mediate adenosine inhibition of Ca^{2+} conductance, whereas A2B and A3 receptors also couple to G ά q and stimulate phospholipase activity Extracellular adenosine concentrations from normal cells are approximately 300 nM; however, in response to cellular damage (e.g. in inflammatory or ischemic tissue), these concentrations are quickly elevated (600-1,200 nm). Thus, in regards to stress or injury, the function of adenosine is primarily that of cytoprotection preventing tissue damage during instances of hypoxia, ischemia, and seizure activity. Activation of A2A receptors produces a constellation of responses that in general can be classified as anti-inflammatory.

Anti-inflammatory Properties

Adenosine is a potent anti-inflammatory agent, acting at its four G-protein coupled receptors. Topical treatment of adenosine to foot wounds in diabetes mellitus has been shown in lab animals to drastically increase tissue repair and reconstruction. Topical administration of adenosine for use in wound healing deficiencies and diabetes mellitus in humans is currently under clinical investigation.

Action on the Heart

When administered intravenously, adenosine causes transient heart block in the AV node. It also causes endothelial dependent

relaxation of smooth muscle as is found inside the artery walls. This causes dilatation of the "normal" segments of arteries where the endothelium is not separated from the tunica media by atherosclerotic plaque. This feature allows physicians to use adenosine to test for blockages in the coronary arteries, by exaggerating the difference between the normal and abnormal segments.

In individuals suspected of suffering from a supraventricular tachycardia (SVT), adenosine is used to help identify the rhythm. Certain SVTs can be successfully terminated with adenosine. This includes any re-entrant arrhythmias that require the AV node for the re-entry (e.g., AV re-entrant tachycardia (AVRT), AV nodal re-entrant tachycardia (AVNRT). In addition, atrial tachycardia can sometimes be terminated with adenosine.

Adenosine has an indirect effect on atrial tissue causing a shortening of the refractory period. When administered via a central lumen catheter, adenosine has been shown to initiate atrial fibrillation because of its effect on atrial tissue. In individuals with accessory pathways, the onset of atrial fibrillation can lead to a life threatening ventricular fibrillation.

Fast rhythms of the heart that are confined to the atria (e.g., atrial fibrillation, atrial flutter) or ventricles (e.g., monomorphic ventricular tachycardia) and do not involve the AV node as part of the re-entrant circuit are not typically converted by adenosine, however the ventricular response rate will be temporarily slowed.

Because of the effects of adenosine on AV node-dependent SVTs, adenosine is considered a class V anti-arrhythmic agent. When adenosine is used to cardiovert an abnormal rhythm, it is normal for the heart to enter asystole for a very brief period. While the adenosine is necessary to save to patient, the event of the heart stopping for several seconds is very disconcerting to a normally conscious patient.

This effect of temporary arrest is often overlooked and not mentioned, except in professional medical literature. The pharmacological effects of adenosine are blunted in individuals who are taking methylxanthines (e.g., caffeine (found in coffee) and theophylline (found predominantly in tea)).

Caffeine's stimulatory effects are primarily (although not entirely) credited to its inhibition of adenosine by binding to the same receptors. By nature of caffeine's purine structure it binds to some of the same receptors as adenosine, effectively blocking adenosine receptors in the central nervous system. This reduction in adenosine activity leads to increased activity of the neurotransmitter dopamine.

Dosage

When given for the evaluation or treatment of an SVT, the initial dose is 6 mg, given as a fast IV/Intraosseous IO push. Due to adenosine's extremely short half-life, start the IV line as proximal to the heart as possible, such as the antecubital fossa. If this has no effect (e.g., no evidence of transient AV block), a 12mg dose can be given 1-2 minutes after the first dose. If the 12mg dose has no effect, a second 12mg dose can be administered 1-2 minutes after the previous dose.

Some clinicians may prefer to administer a higher dose (typically 18 mg), rather than repeat a dose that apparently had no effect.

When given to dilate the arteries, such as in a "stress test", the dosage is typically 0.14 mg/kg/min, administered for 4 or 6 minutes, depending on the protocol.

Consider *increasing* the recommended dose in patients on theophylline since methylxanthines prevent binding of adenosine at receptor sites.

Consider *decreasing* the dose in patients on dipyridamole (Persantine) and diazepam (Valium) because adenosine potentiates the effects of these drugs.

Consider *decreasing* the recommended dose in half in patients who are presenting Congestive Heart Failure, Myocardial Infarction, shock, hypoxia, and/or hepatic or renal insufficiency.

Consider *decreasing* the recommended dose in half for elderly patients.

Drug Interactions

Beta blockers and dopamine may precipitate toxicity in the patient.

Contraindications

Poison/Drug induced tachycardia, Asthma (relative contraindication), 2nd or 3rd degree heart block, Atrial fibrillation, atrial flutter, Ventricular tachycardia, Sick sinus syndrome, Stokes-Adams Attack, Wolf-Parkinson-White syndrome, bradycardia with Pre-mature Ventricular Contractions (PVCs).

- WPW- adenosine may be administered if equipment for cardioversion is immediately available as a backup.
- A-flutter W/rvr - when it first presents with SVT

Side Effects

Many individuals experience facial flushing, lightheadedness, diaphoresis, or nausea after administration of adenosine. These symptoms are transitory, usually lasting less than one minute.

Metabolism

When adenosine enters the circulation, it is broken down by adenosine deaminase, which is present in red cells and the vessel wall.

Dipyridamole, an inhibitor of adenosine deaminase, allows adenosine to accumulate in the blood stream. This causes an increase in coronary vasodilatation.

Role of Lipids

Lipids are chiefly fatty acid esters, and are the basic building blocks of biological membranes. Another biological role is energy storage (e.g., triglycerides). Most lipids consist of a polar or hydrophilic head (typically glycerol) and one to three non-polar or hydrophobic fatty acid tails, and therefore they are amphiphilic. Fatty acids consist of unbranched chains of carbon atoms that are connected by single bonds alone (saturated fatty acids) or by both single and double bonds (unsaturated fatty acids). The chains are usually 14-24 carbon groups long, but it is always an even number.

For lipids present in biological membranes, the hydrophilic head is from one of three classes:

- Glycolipids, whose heads contain an oligosaccharide with 1-15 saccharide residues.
- Phospholipids, whose heads contain a positively charged group that is linked to the tail by a negatively charged phosphate group.
- Sterols, whose heads contain a planar steroid ring, for example, cholesterol.

Other lipids include prostaglandins and leukotrienes which are both 20-carbon fatty acyl units synthesised from arachidonic acid.

Activity of Hormones

Hormones are produced in the endocrine glands, where they are secreted into the bloodstream. They perform a wide range of roles in the various organs including the regulation of metabolic pathways and the regulation of membrane transport processes.

Hormones may be grouped into three structural classes:

- The steroids are one class of such hormones. They perform a variety of functions, but they are all made from cholesterol.
- Simple amines or amino acids.
- Peptides or proteins.

Significance of Amino Acids

Amino acids are molecules that contain both amino and carboxylic acid functional groups. (In biochemistry, the term amino acid is used when referring to those amino acids in which the amino and carboxylate functionalities are attached to the same carbon, plus proline which is not actually an amino acid).

Amino acids are the building blocks of long polymer chains. With 2-10 amino acids such chains are called peptides, with 10-100 they are often called polypeptides, and longer chains are known as proteins. These protein structures have many structural and functional roles in organisms.

There are twenty amino acids that are encoded by the standard genetic code, but there are more than 500 natural amino acids.

When amino acids other than the set of twenty are observed in proteins, this is usually the result of modification after translation (protein synthesis). Only two amino acids other than the standard twenty are known to be incorporated into proteins during translation, in certain organisms:

- Selenocysteine is incorporated into some proteins at a UGA codon, which is normally a stop codon.
- Pyrrolysine is incorporated into some proteins at a UAG codon. For instance, in some methanogens in enzymes that are used to produce methane.

Besides those used in protein synthesis, other biologically important amino acids include carnitine (used in lipid transport within a cell), ornithine, GABA and taurine.

Structure of Protein

The particular series of amino acids that form a protein is known as that protein's primary structure. Proteins have several, well-classified, elements of local structure and these are termed secondary structure. The overall 3D structure of a protein is termed its tertiary structure. Proteins often aggregate into macromolecular structures, or quaternary structure.