5

The structure and function of membrane proteins

By 2004 the Protein Data Bank of structures contained over 22 000 submissions. Almost all of these files reflect the atomic coordinates of soluble proteins and less than 1 percent of all deposited coordinates belong to proteins found within membranes. It is clear from this small fraction that membrane proteins present special challenges towards structure determination.

Many of these challenges arise from the difficulty in isolating proteins embedded within lipid bilayers. For many decades membrane proteins have been represented as amorphous objects located in a sea of lipid with little consideration given to their organization. This is an unfortunate state of affairs because the majority of proteins encoded by genomes are located in membranes or associated with the membrane interface. Slowly, however, insight has been gained into the different types of membrane proteins, their domain structure and more recently the three dimensional structures at atomic levels of resolution. This chapter addresses the structural and functional properties of membrane proteins.

The molecular organization of membranes

Although this book is primarily about proteins it is foolish to ignore completely the properties of membranes since the lipid bilayers strongly influence protein structure and function. The term lipid refers to *any* biomolecule that is soluble in an organic solvent, and extends to quinones, cholesterol, steroids, fatty acids, triacylglycerols and non-polar vitamins such as vitamin K_3. All of these components are found in biological membranes. However, for most purposes the term lipid refers to the collection of molecules formed by linking a long-chain fatty acid to a glycerol-3-phosphate backbone; these are the major components of biological membranes.

Two long chain fatty acids are esterified to the C_1 and C_2 positions of the glycerol backbone whilst the phosphate group is frequently linked to another polar group. Generically all of these molecules are termed glycerophospholipids (Figure 5.1). Phosphatidic acid occurs in low concentrations within membranes and the addition of polar head groups such as ethanolamine, choline, serine, glycerol, and inositol forms phosphatidylethanolamine (PE), phosphatidylcholine (PC, also called lecithin), phosphatidylserine (PS), phosphatidylglycerol (PG) and phosphatidylinositol (PI), the key components of biological membranes (Table 5.1). Fatty acyl chains usually containing 16 or 18 carbon atoms occupy the R1 and R2 positions although chain length can range from 12 to 24 carbons.

Proteins: Structure and Function by David Whitford

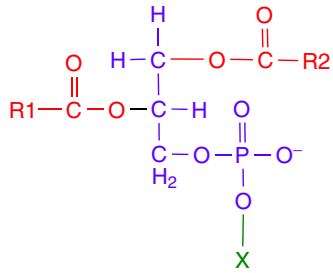


Figure 5.1 The structure of glycerophospholipids involves the esterification of two long-chain fatty acids to a glycerol-3-phosphate backbone with further additions of polar groups occurring at the phosphate group. When X=H the parent molecule phosphatidic acid is formed

In animal membranes the C_1 position is typically a saturated lipid of 16–18 carbons whilst the C_2 position often carries an unsaturated C_{16}–C_{20} fatty acid. Unsaturated acyl chains (Table 5.2) contain one or more double bonds usually found at carbons 5, 6, 8, 9, 11, 12, 14 and 15. Where more than one double bond occurs the sites are separated by three carbons so that, for example in linolenic acid the double bonds are at 9, 12 and 15, whilst in arachidonic acid the four double bonds are located at carbons 5, 8, 11 and 14. The double bonds are found in the *cis* state with *trans* isomers rarely found. With a large number of possible head groups coupled to a diverse number of fatty acyl chains (each potentially differing in their degree of saturation) the number of possible combinations of non-polar chain/polar head group is enormous.

Naming phospholipids follows the identity of the fatty acid residues; so a glycerophospholipid containing two 16 carbon saturated fatty acyl chains and a choline head group is commonly referred to as dipalmitoylphosphatidylcholine (DPPC). Although

Table 5.1 The common head groups attached to the glycerol-3-phosphate backbone of glycerophospholipids

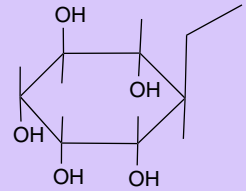
Donor	Head group	Phospholipid
Water	–H	Phosphatidic acid
Ethanolamine	$-CH_2-CH_2-NH_3^+$	Phosphatidylethanolamine
Choline	$-CH_2-CH_2-N^+(CH_3)_3$	Phosphatidylcholine
Serine	$-CH_2-CH(NH_3^+)-C(=O)-O^-$	Phosphatidylserine
Inositol		Phosphatidylinositol
Glycerol	$-CH_2-CH(OH)-CH_2-OH$	Phosphatidylglycerol

Table 5.2 Common acyl chains found in lipids of biological membranes

Name	Numeric representation	Name	Numeric representation
Lauric	12:0	Stearic	18:1
Myristic	14:0	Linoleic	18:2
Palmitic	16:0	Linolenic	18:3
Palmitoleic	16:1	Arachidic	20:0
Oleic	18:0	Arachidonic	20:4

For molecules such as linolenic acid the position of the three double bonds are denoted by the nomenclature $\Delta^{9,12,15}$ with carbon 1 starting from the carboxyl end. The IUPAC name for this fatty acid is 9,12,15 octadecenoic acid. A third alternative identifies the number of carbon atoms in the acyl chain, the number of double bonds and the geometry and carbon number associated with each bond. Thus, linolenic acid is 18:3 9-*cis*, 12-*cis*, 15-*cis*

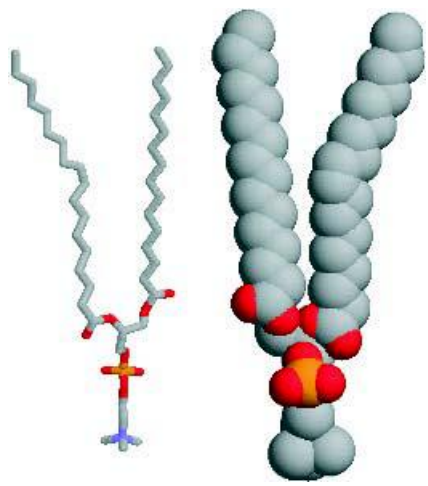


Figure 5.2 The diagram shows a wireframe and space filling model of two acyl chains, one 16 carbons in length, the other 18 carbons in length containing a *cis* double bond. A phosphatidylcholine head group is linked to a glycerol backbone. Protons have been omitted to increase clarity

glycerophospholipids are the most common lipids of biological membranes other variants occur such as sphingolipids, sulfolipids and galactolipids. In plant membranes sulfolipids and galactolipids occur alongside phospholipids, increasing the variety of lipids.

A phospholipid such as DPPC is described as amphipathic, containing polar and non-polar regions. Space filling models (Figure 5.2) show that introducing a single *cis* double bond produces a 'kink' in the acyl chain and an effective volume increase. This has pronounced effects on physical properties of lipids such as transition temperatures, fluidity and structures formed in solution.

When dispersed in aqueous solutions lipid molecules minimize contact between the hydrophobic acyl chain and water by forming structures where the polar head groups interact with aqueous solvents whilst the non-polar 'tails' are buried away from water. The size and shape of the lipid assembly is a complex combination of factors, including polar head group volume, charge, degree of unsaturation and length. Thus lipids with short acyl chains form micelles and pack efficiently into a spherical shape where the large volume of the hydrated polar head group exceeds that of the tail and favours micellar structure. Lipids with a single acyl chain, as well as detergents, adopt this shape. Longer acyl chains, particularly those with one or more double bonds, cannot pack efficiently into a spherical micelle and produce bilayers in most cases although alternative structures, such as inverted micelles where the polar head group is buried on the inside, can be formed by some of the common lipids. Additionally lipid aggregates are modulated by ionic strength or pH, but a detailed discussion of these states is beyond the scope of this text.

The ability of lipids or detergents to aggregate in solution is often expressed in terms of their critical micelle concentration (CMC). At concentrations above their CMC lipids are found as aggregated structures whilst below this concentration monomeric species exist in solution. The CMC of dodecylmaltoside (a single 12 carbon fatty acyl chain linked to a maltose group) is estimated at ~150 μM whilst the corresponding value for DPPC is 0.47 nM and means that this polar lipid is always found as an aggregated species. An isolated bilayer dispersed in a polar solvent is unlikely to be stable since at some point the ends of the bilayer will be exposed leading

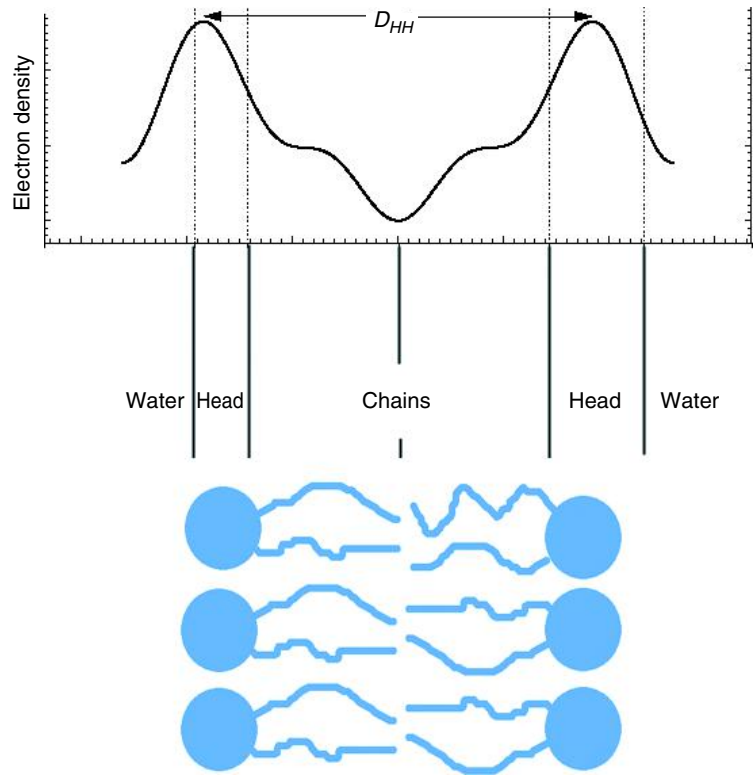


Figure 5.3 Electron density profile determined for a single bilayer. The electron dense regions correspond to the head group region whilst the non-polar domain are less electron dense. Such data established the basic dimensions of a bilayer (D_{HH}) at around 40 nm and the profile is shown alongside a diagram of a bilayer showing relative positions of head groups and acyl chains

to unfavourable interactions between non-polar and polar surfaces. As a result, extended bilayers in aqueous solution often form closed, sealed, solvent-filled structures known as liposomes or single lamellar vesicles that have proved to be good models for biological membranes.

Since integral membrane proteins are located in bilayers the properties and arrangement of lipids dictates the structure of any protein residing in this region. The structure of a lipid bilayer has been examined by neutron and X-ray diffraction to confirm a periodic arrangement for the distribution of electron density. The polar head groups are electron dense whilst the acyl chains represent comparatively 'diffuse' areas of electron density (Figure 5.3).

Fluidity and phase transitions are important properties of lipid bilayers

An important property of lipid bilayers is fluidity. Membrane fluidity depends on lipid composition, temperature, pH and ionic strength. At low temperatures the fluidity of a membrane is reduced with the acyl chains showing lower thermal mobility and packing into a 'frozen' gel state. At higher temperatures increased translational and rotational mobility results in a less ordered fluid state, sometimes called the liquid crystal (Figure 5.4). Transitions from the gel to fluid states are characterized by phase transition temperatures (Table 5.3). Short chains have lower transition temperatures when compared with longer lipids. The

transition temperature increases with chain length since hydrophobic interactions increase with the number of non-polar atoms. The transition temperature decreases with increasing unsaturation and is reflected by greater inherent disorder packing.

Biological membranes exhibit broad phase transition temperatures over a wide temperature range. All biological membranes exist in the fluid state and organisms have membranes whose lipid compositions exhibit transition temperatures significantly below their lowest body temperature. Thus, fish living in the depths of the ocean where temperatures may drop below −20 °C have membranes rich in unsaturated long-chain fatty acids such as linolenic and linoleic acid.

Real membranes are not homogeneous dispersions of lipids but contain proteins, often in considerable quantity. The lipid:protein ratio varies dramatically, with membranes such as the myelin sheath possessing comparatively little protein (Table 5.4). In contrast membranes concerned with energy transduction have larger amounts of protein and exhibit lower lipid:protein ratios.

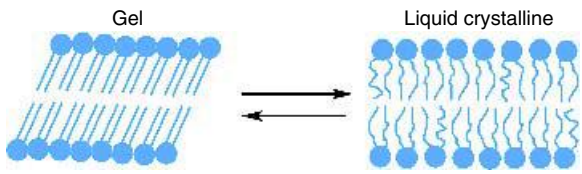


Figure 5.4 The gel to fluid (liquid crystalline) transition in lipid bilayers. In the gel phase the head groups are tightly packed together, the tails have a regular conformation and the bilayer width is thicker than the liquid crystalline state

Table 5.3 Phase transition temperatures of common diacylphospholipids

Diacylphospholipid	Transition temperature °C (T_m)
C22 phosphatidyl choline	75
C18 phosphatidyl choline	55
C16 phosphatidyl choline	41
C14 phosphatidyl choline	24
C18:1 phosphatidyl choline	−22
C16 phosphatidyl serine (low pH)	72
C16 phosphatidyl serine (high pH)	55
C16 phosphatidyl ethanolamine	60
C14 phosphatidyl ethanolamine	50

Phase transition temperatures are measured by differential scanning calorimetry (DSC) where heat is absorbed (ΔH increases) as the lipid goes through a phase transition

The fluid mosaic model

Any model of membrane organization must account for the properties of lipids, such as self-association into stable structures, mobility of lipids within monolayers, and the impermeability of cell membranes to polar molecules. In addition models must account for the presence of proteins both integral and peripheral and at a wide variety of ratios with the lipid component. In 1972, to account for these and many other observations, S.J. Singer and G. Nicolson proposed the fluid mosaic model; a model that has with only minor amendments stood the test of time. In this model proteins were originally viewed as floating like icebergs in a sea of lipid but in view of the known lipid:protein ratios found in some membranes this picture is unrealistic and representations should involve higher proportions of protein.

Table 5.4 Protein and lipid content of different membranes

Membrane	Protein	Lipid
Myelin sheath (Schwann cells)	21	79
Erythrocyte plasma membrane	49	43
Bovine retinal rod	51	49
Mitochondrial outer membrane	52	48
Mitochondrial inner membrane	76	24
Sarcoplasmic reticulum (muscle cells)	67	33
Chloroplast lamellae	70	30

The figures represent the percentage by weight. The figures do not always add up to 100% reflecting different contents of carbohydrate within each membrane (derived from Guidotti, G. *Ann. Rev. Biochem.* 1972, **41**, 731–752)

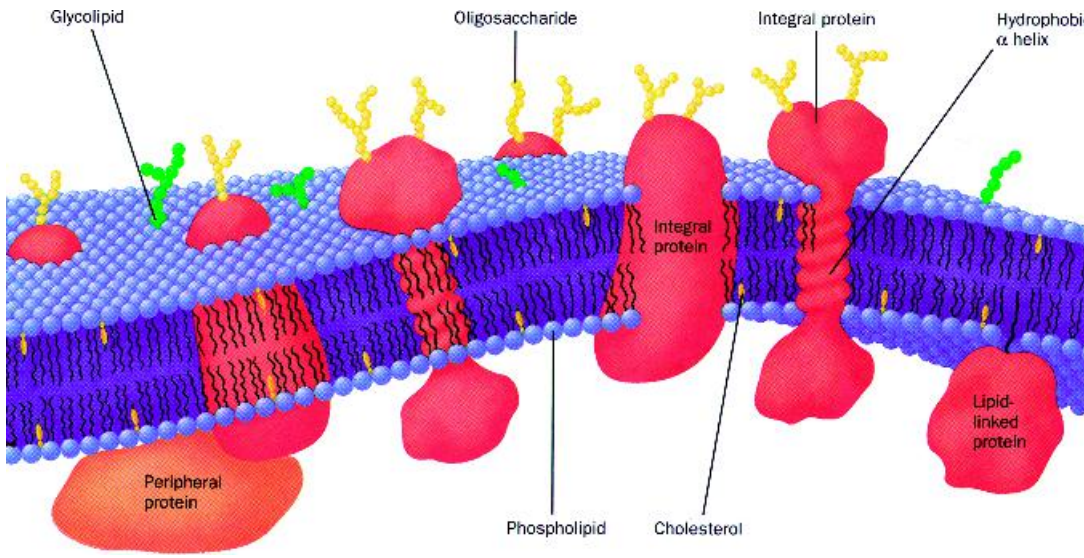


Figure 5.5 The organization of a typical cell membrane. The lipid in blue contains integral and peripheral proteins (red/orange) with cholesterol (yellow) found in animal membranes whilst polysaccharides are attached to the extra-cellular surface of proteins (reproduced with permission from Voet, D., Voet, J.G & Pratt, C.W., *Fundamentals of Biochemistry*, 1999 John Wiley & Sons, Ltd, Chichester)

In 1972 nothing was known about the structure of these proteins but the intervening period has elucidated details of organization within membranes. Interactions between protein and lipid alter the properties of both and in some systems there exists convincing data on the arrangement of proteins with respect to exterior and cytosolic compartments. However, a major problem is a comparative paucity of data on the secondary and tertiary structure of integral proteins that restricts understanding of their function in transport, cell–cell recognition, catalysis, etc. To address these problems it is necessary to examine the composition of membranes in more detail and one of the best model systems is the erythrocyte or red blood cell membrane. Figure 5.5 shows a typical cell membrane.

Membrane protein topology and function seen through organization of the erythrocyte membrane

The erythrocyte membrane is a popular system to study because blood can be obtained in significant quantities containing large amounts of cells (i.e membranes). The cell membrane is readily partitioned from haemoglobin and other cytosolic components by centrifugation. The resulting membranes are called 'ghosts' because they reseal to form colourless particles devoid of haemoglobin. Using ghost membranes identified two major groups of proteins. The first group of proteins shows weak association with membranes and is removed after mild washing treatments such as incubation with elevated concentrations of salt. The second class involved proteins that were only removed from the membrane by extreme procedures that included the use of detergents, organic solvents, chaotropic agents, lipases or mechanical fractionation. This second procedure involved the complete disruption of membrane environment with the result that many protein's lost structure and function. The terms *extrinsic* and *intrinsic* are sometimes applied to these two groups of membrane proteins. Other schemes of classification use the terms *non-integral* or peripheral and *integral* to describe each group. Unsurprisingly the properties of integral and non-integral proteins differ but further characterization of the erythrocyte membrane revealed a third type of

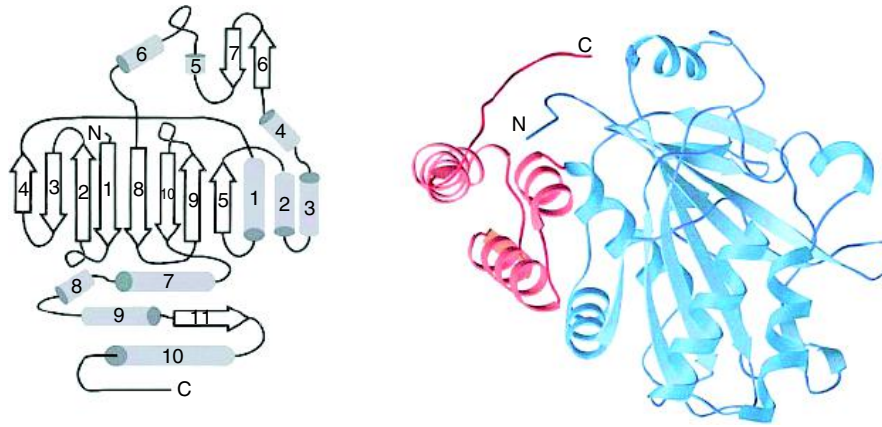


Figure 5.6 Structure of band 3 monomer. The secondary structure distribution (left) and a ribbon diagram of band 3 monomer (right). Regions shown in blue correspond to the peripheral protein binding domain and the red regions to the dimerization arm (reproduced with permission from Zhang, D., Kiyatkin, A., Bolin, J.T. & Low, P.S. *Blood* 2000, **96**, 2925–2933. The American Society of Hematology)

protein. This integral protein was attached firmly to the bilayer yet had substantial proportions of its sequence accessible to the aqueous phase.

An intrinsic protein of the erythrocyte membrane: band 3 protein

The name 'band 3' originally arose from sodium dodecyl sulfate–polyacrylamide gel electrophoresis studies of the protein composition of erythrocyte membranes where it was the subunit with the third heaviest molecular mass. It is the most abundant protein found in the erythrocyte membrane facilitating anion exchange and the transfer of bicarbonate (HCO_3^-) and chloride (Cl^-) ions. Band 3 protein is hydrophobic with many non-polar side chains and is released from the membrane only after extraction with detergents or organic solvents. As such it is typical of many membrane proteins and until recently there was little detail on the overall tertiary structure.

Band 3 consists of two independent domains (Figure 5.6). The N-terminal cytoplasmic domain links the membrane to the underlying spectrin–actin-based cytoskeleton using ankyrin and protein 4.1 as bridging proteins. The C-terminal domain mediates anion exchange across the erythrocyte membrane. As a result of its cytoplasmic location the structure of the N-terminal domain has been established by crystallography to a resolution of 2.6 Å but the membrane C-terminal domain remains structurally problematic. The cytoplasmic domain contains 11 β strands and 10 helical segments and belongs to the $\alpha + \beta$ fold class. Eight β-strands assemble into a central β-sheet of parallel and antiparallel strands whilst two of the remaining strands (β6 and β7, residues 176–185) form a β-hairpin that constitutes the core of the ankyrin binding site.

The approach of studying large extracellular regions belonging to membrane proteins is widely used to provide information since it represents in many instances the only avenue of structural analysis possible for these proteins.

Membrane proteins with globular domains

Integral membrane proteins are distinguished from soluble domains by their high content of hydrophobic residues. Residues such as leucine, isoleucine and valine often occur in blocks or segments of approximately 20–30 residues in length. Consequently, it is possible to identify membrane proteins by the distribution of residues with hydrophobic side chains throughout a primary sequence. Hydropathy plots reflect the preference of amino acid side chains for polar

and non-polar environments and as such numerical parameters are assigned to each residue (Table 5.5). The hydropathy values reflect measurements of the free energy of transfer of an amino acid from non-polar to polar solvents. By assigning numerical values to each residue along a sequence and then averaging these values over a pre-defined window size a hydropathy profile is established. For a small window size the precision is not great but for a window size of between 18 and 22 accurate prediction of transmembrane regions is possible. Increasingly sophisticated algorithms allow transmembrane region recognition and this approach is routinely performed on new, unknown, proteins.

Transmembrane regions occur as peaks in the hydropathy profile. An analysis of band 3 protein reveals a primary sequence containing many hydrophobic residues arranged in at least 13 transmembrane domains. This was therefore consistent with its role as an integral membrane protein. In contrast the 'integral' protein glycophorin also found in the erythrocyte has a low content of non-polar residues compatible with a single transmembrane helix (Figure 5.7). This was unexpected in view of its location and shows the dramatic variation of integral protein organization and secondary structure.

Table 5.5 Kyte and Doolittle scheme of ranking hydrophobicity of side chains

Amino acid	Parameter	Amino acid	Parameter
Ala	1.80	Leu	3.80
Arg	−4.50	Lys	−3.90
Asn	−3.50	Met	1.90
Asp	−3.50	Phe	2.80
Cys	2.50	Pro	−1.60
Gln	−3.50	Ser	−0.80
Glu	−3.50	Thr	−0.70
Gly	−0.40	Tyr	−0.90
His	−3.20	Trp	−1.30
Ile	4.50	Val	4.20

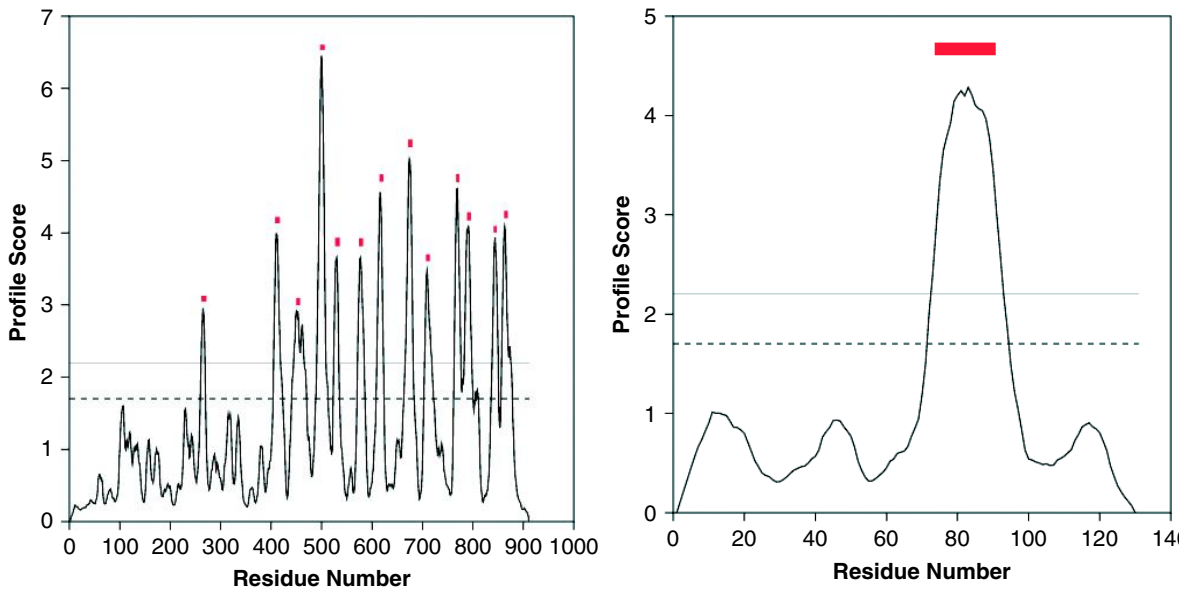


Figure 5.7 The putative transmembrane domains of band 3 and glycophorin. The red bars indicate the positions of the transmembrane domains

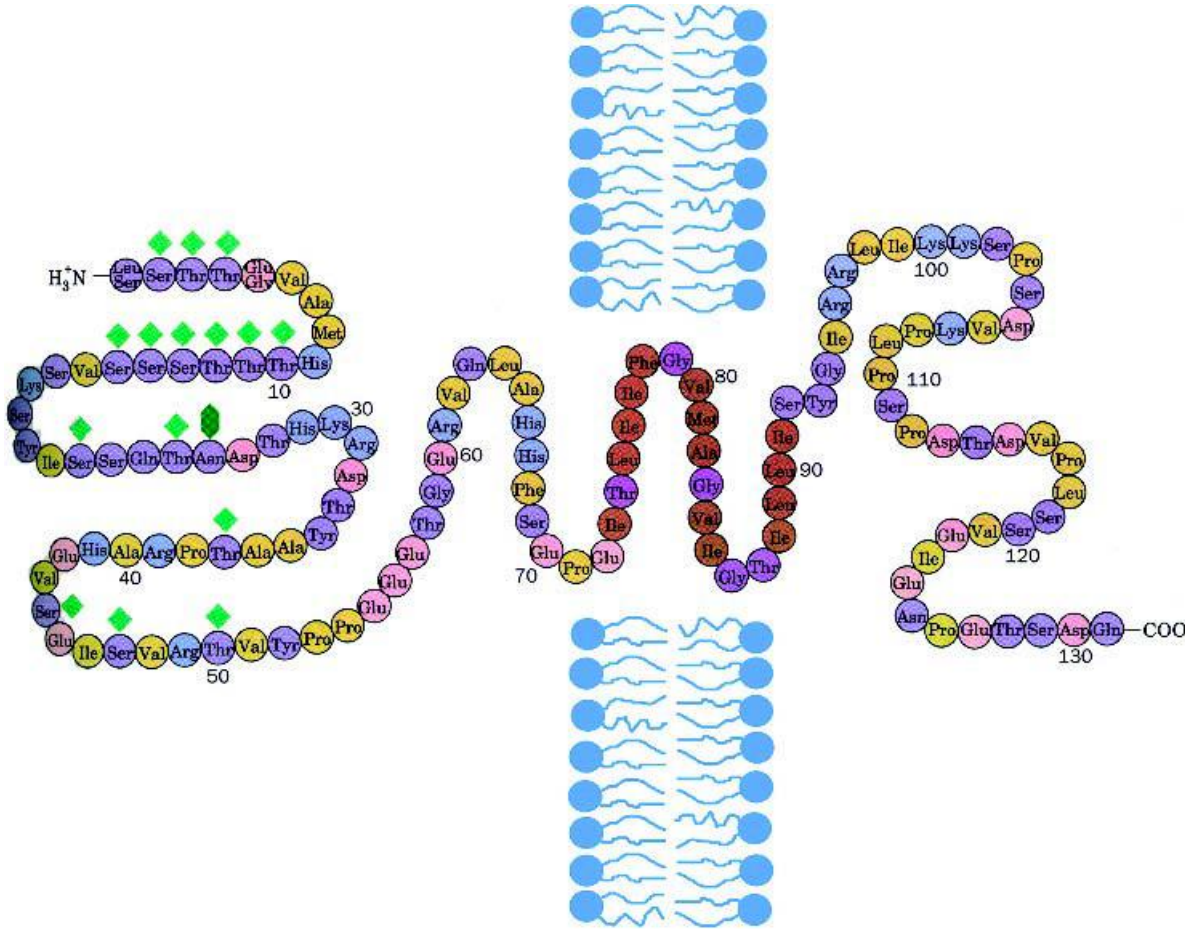


Figure 5.8 Sequence and postulated structure of glycophorin A. The external N terminal region has many polar residues and is extensively glycosylated. There is no glycosylation of the hydrophilic domain on the cytoplasmic side of the membrane. The non-polar region stretches for approximately 20 residues. A common genetic variant of glycophorin occurs at residues 1 and 5 with Ser and Gly replaced by Leu and Glu

Careful inspection of the hydropathy plot for glycophorin reveals that a single hydrophobic domain extends for approximately 20–25 residues and that less than 20 % of the protein's 131 residues are compatible with membrane locations. This observation was supported by amino acid sequencing of glycophorin, the first membrane protein to be sequenced, showing significant numbers of polar residues at the N and C terminals (Figure 5.8).

A combination of biochemical techniques established the arrangement of glycophorin – large hydrophilic domains protruded from either side of the erythrocyte membrane. One approach was to use membrane-impermeant agents that reacted with specific residues to yield fluorescent or radioactive labels that facilitated subsequent detection. Using this approach only the N-terminal domain was labelled and it was deduced that this region protruded out from the cell's surface. Rupturing the erythrocyte membrane labelled C-terminal domains but the transmembrane region was never labelled. Detailed studies established an N-terminal region of approximately 70 residues (1–72)

with the hydrophobic domain extending from 73–92 and the C-terminal region occupying the remaining fraction of the polypeptide from residues 92–131. *In vivo* the hydrophobic domain promotes dimerization and studies have shown that fusion of this protein sequence to 'foreign' proteins not only promotes membrane localization but also favours association into dimers. Dimerization is sequence dependent with mutagenesis of critical residues completely disrupting protein association. However, although considerable detail has been acquired on the organization of proteins within the erythrocyte membrane a description at an atomic level was largely lacking. In part this stemmed from the difficulty in isolating proteins in forms suitable for analysis by structural methods.

What was required was a different approach. A new approach came with two very different integral membrane proteins. The first was the purple protein, bacteriorhodopsin, found in the bacterium *Halobacterium halobium*. This membrane protein functions as a light-driven proton pump. The second system was a macromolecular photosynthetic complex from purple non-sulfur bacteria that ultimately converted light into chemical energy. The results from studies of these proteins would revolutionize our understanding of membrane protein structure.

Bacteriorhodopsin and the discovery of seven transmembrane helices

The study of the protein bacteriorhodopsin represents a landmark in membrane protein structure determination because for the first time it allowed an insight into the organization of secondary and tertiary structure, albeit at low resolution. Bacteriorhodopsin, a membrane protein of 247 residues (M_r~26 kDa), accumulates within the membranes of *H. halobium* – a halophilic (salt loving) bacterium that lives in extreme conditions and is restricted to unusual ecosystems. The cell membrane of *H. halobium* contains patches approximately 0.5 μm wide consisting predominantly of bacteriorhodopsin along with a smaller quantity of lipid. These patches contribute to a purple appearance of membranes and contain protein arranged in an ordered two-dimensional lattice whose function is to harness light energy for the production of ATP via light driven proton pumps. This

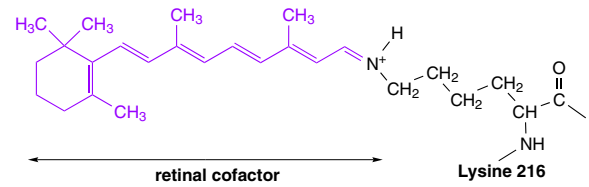


Figure 5.9 The retinal chromophore of bacteriorhodopsin linked to lysine 216 of bacteriorhodopsin. The retinal is responsible for the purple colour observed in the membranes of *H. halobium*

process is assisted by the presence of retinal co-factor linked covalently to lysine 216 in bacteriorhodopsin (Figure 5.9).

Bacteriorhodopsin is the simplest known proton pump and differs from those found in other systems in not requiring associated electron transfer systems. By dispersing purified purple membranes containing bacteriorhodopsin in lipid vesicles along with bovine heart ATP synthetase (the enzyme responsible for synthesizing ATP) Efraim Racker and Walter Stoeckenius were able to demonstrate that light-driven proton fluxes resulted in ATP synthesis. Not only was this a convincing demonstration of the light-driven pump of bacteriorhodopsin but it was also an important milestone in emphasizing the role of proton gradients in ATP formation.

Richard Henderson and Nigel Unwin used the two-dimensional patches of protein to determine the structure of bacteriorhodopsin using electron crystallography a technique (see Chapter 10) analogous to X-ray crystallography. By a process of reconstruction Henderson and Unwin were able to produce low-resolution structural images of bacteriorhodopsin in 1975 (Figure 5.10). This was the first time the structure of a membrane protein had been experimentally determined at anything approaching an atomic level of resolution.

The structure of bacteriorhodopsin revealed seven transmembrane helices and it was assumed that these regions were linked by short turns that projected into the aqueous phase on either side of the purple membrane. Shortly after initial structure determination the primary sequence of this hydrophobic protein was determined. This advance allowed a correlation

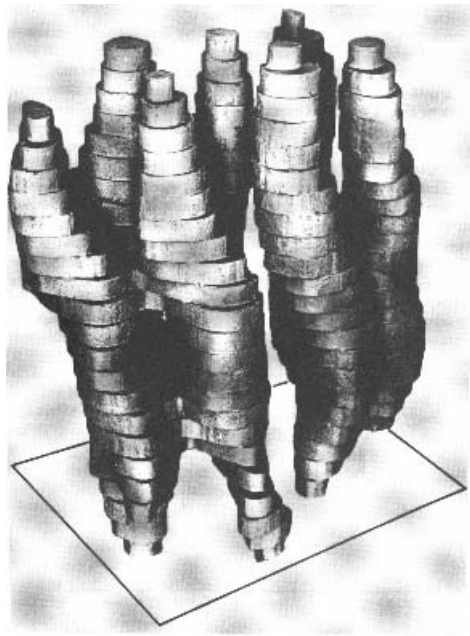


Figure 5.10 A model proposed for the structure of bacteriorhodopsin derived using electron microscopy (after Henderson, R. & Unwin, P.N.T *Nature* 1975, **257**, 28–32). The view shown lies parallel to the membrane with the helices running across the membrane. Although clearly 'primitive' and of low resolution this representation of seven transmembrane helices had a dramatic effect on perceptions of membrane protein structure

between structural and sequence data. Hydrophobicity plots of the amino acid sequence identified seven hydrophobic segments within the polypeptide chain and this agreed with the transmembrane helices seen in electron density maps. The primary sequence of bacteriorhodopsin was pictured as a block of seven transmembrane helices (Figure 5.11).

Subsequently, determination of the structure of bacteriorhodospin by X-ray crystallography to a resolution of 1.55 Å confirmed the organization seen by electron diffraction and showed that the seven transmembrane helices were linked by extra- and intracellular loops as originally proposed by Henderson and Unwin (Figure 5.12). Although not apparent from the early low-resolution structures a short antiparallel β strand is found in a loop between helices B and C. More significantly this structure allowed insight into the mechanism of signal transduction and ion pumping, the major biological roles of bacteriorhodopsin.

The light-driven pump is energized by light absorption by retinal and leads to the start of a photocycle where the conversion of retinal from an all-*trans* isomer to a 13-*cis* isomer is the trigger for subsequent reactions. Photobleaching of pigment leads to the vectorial release of a proton on the extracellular side of the membrane and a series of reactions characterized by spectral intermediates absorbing in different regions of the UV–visible spectrum.

The discovery that bacteriorhodopsin contained seven transmembrane helices arranged transversely across the membrane has assisted directly in understanding the structure and function of a diverse range of proteins. Of direct relevance is the presence of rhodopsin in visual cycle of the eye in vertebrates and invertebrates as well as a large family of proteins based on seven transmembrane helices.

Seven transmembrane helices and the discovery of G-protein coupled receptors

These receptors are embedded within membranes and are not easily isolated let alone studied by structural methods. However, cloning cDNA sequences for many putative receptors allowed their amino acid sequences to be defined with the recognition of similarity within this large class of receptors. The proteins were generally 400–480 residues in length with sequences containing seven conserved segments or blocks of hydrophobic residues. These seven regions of hydrophobic residues could only mean, in the light of the work on bacteriorhodopsin, transmembrane helices linked by cytoplasmic or extracellular loops. Amongst the first sequences to be recognized with this topology was human β_2 adrenergic receptor (Figure 5.13). Seven transmembrane helices approximately 24 residues in length were identified together with a glycosylated N-terminal domain projecting out from the extracellular surface and on the other side of the membrane a much larger C-terminal domain.

Prior to sequence analysis receptors such as the β_2 adrenergic receptors had been subjected to enormous experimental study. It was shown that receptors bind

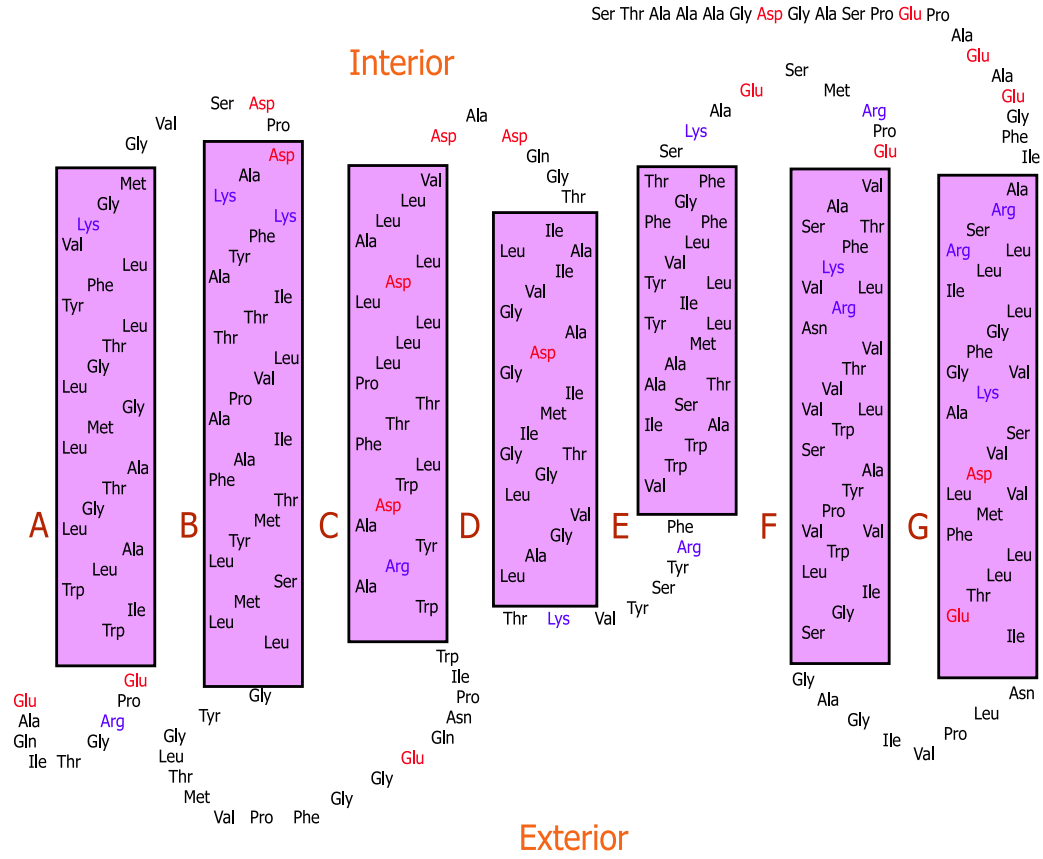


Figure 5.11 The primary sequence of bacteriorhodopsin superimposed on a seven transmembrane helix model. Charged side chains are shown in either red (negative) or blue (positive). Charged residues are usually found either in the loop regions or at the loop–helix interface. The presence of charged residues within a helix is rare and normally indicates a catalytic role for these side chains. This is seen in helix C and G of bacteriorhodopsin

ligands or hormones with kinetics resembling those exhibited by enzymes where the hormone (substrate) bound tightly with dissociation constants measured in the picomolar range (10^{-12} M). Binding studies also helped to establish the concept of agonist and antagonist. The use of structural analogues or agonists to the native hormone yielded a binding reaction and elicited a normal receptor response. In contrast antagonists compete for hormone binding site on the receptor inhibiting the normal biological reaction. One important hormone antagonist is propranolol – a member of a group of chemicals widely called β blockers – that competes with adrenaline (epinephrine) for binding sites on adrenergic receptors. These compounds have been developed as drugs to control pulse rate and blood pressure. Not only have useful drugs emerged from agonist and antagonist binding studies of adrenergic receptors but four major groups of adrenergic receptor have been recognized in the membranes of vertebrates (α_1, α_2 β_1 and β_2).

These receptors play a role in signal transduction and a major discovery in 1977 was the distinction between β adrenergic receptors and adenylate cyclase. Adenylate cyclase was shown in classic research on hormonal control of glycogen levels to be a membrane bound receptor for adrenaline. This led to the incorrect

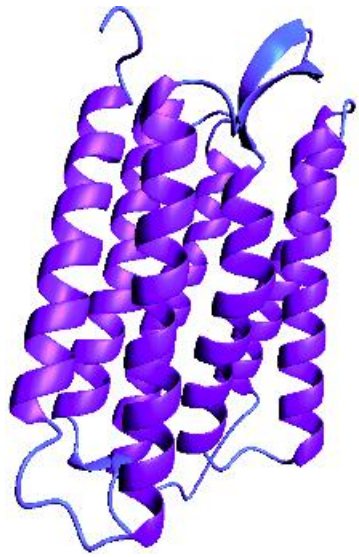


Figure 5.12 High resolution structure of bacteriorhodopsin (PDB:1C3W) showing seven transmembrane helices linked by extra and intracellular loops

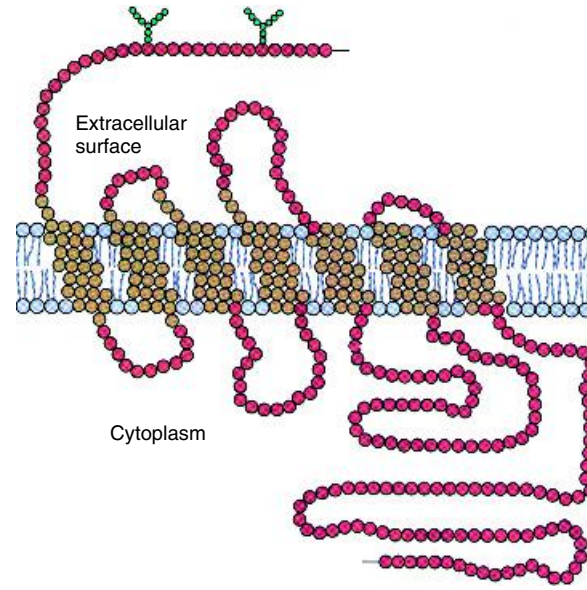


Figure 5.13 The amino acid sequence of the β_2 adrenergic receptor defines the organization of the protein into seven transmembrane helices. The adrenergic receptors are so called because they interact with the hormone adrenaline (after Dohlman *et al. Biochemistry* 1987, **26**, 2660–2666). The receptor interacts with cellular proteins in the C-terminal region and interaction is controlled by the reversible phosphorylation of Ser/Thr residues

idea that receptors were adenylate cyclase – a role shown later to be performed by G protein coupled receptors. Both adenylate cyclase and membrane bound receptors catalyse the activation or inhibition of cyclic AMP synthesis. A crucial discovery was the observation that transduction of the hormonal signal to adenylate cyclase involved not only a membrane bound receptor but a separate group of proteins called G proteins. G proteins are so called because they bind guanine nucleotides such as GTP/GDP. As a result the overall picture of signal transduction mechanisms changed to involve a hormone receptor, adenylate cyclase *and* a G protein. All of these proteins are associated with the membrane.

Studies of bacteriorhodopsin played a key role in understanding rhodopsin, the photoreceptive pigment at the heart of the visual cycle and responsible for the transduction of light energy in the eyes of vertebrates and invertebrates. Rhodopsin is the best characterized receptor and much of our current knowledge of signal transduction arose from studies of G proteins and G protein-coupled receptors in the visual cycle. The model of signal transduction receptors as seven transmembrane helices with extra and intracellular loops was in part possible because of the initial structural studies of bacteriorhodopsin – at first glance a completely unrelated protein.

Within the retina of the eye rhodopsin is the major protein of the disc membrane of rod cells and by analogy to the bacterial protein it contains a chromophore 11-*cis* retinal that binds to opsin (the protein). Rhodopsin absorbs strongly between wavelengths of 400 and 600 nm and absorption of light triggers the conversion of the bound chromophore to the all-*trans* form. Several conformational changes later metarhodopsin II is formed as an important intermediate in the multiple step process of visual transduction. The release of *trans*-retinal from the protein and the conversion of retinal back to the initial 11-*cis* state by the action of retinal isomerase signals the end of the cycle (Figure 5.14).

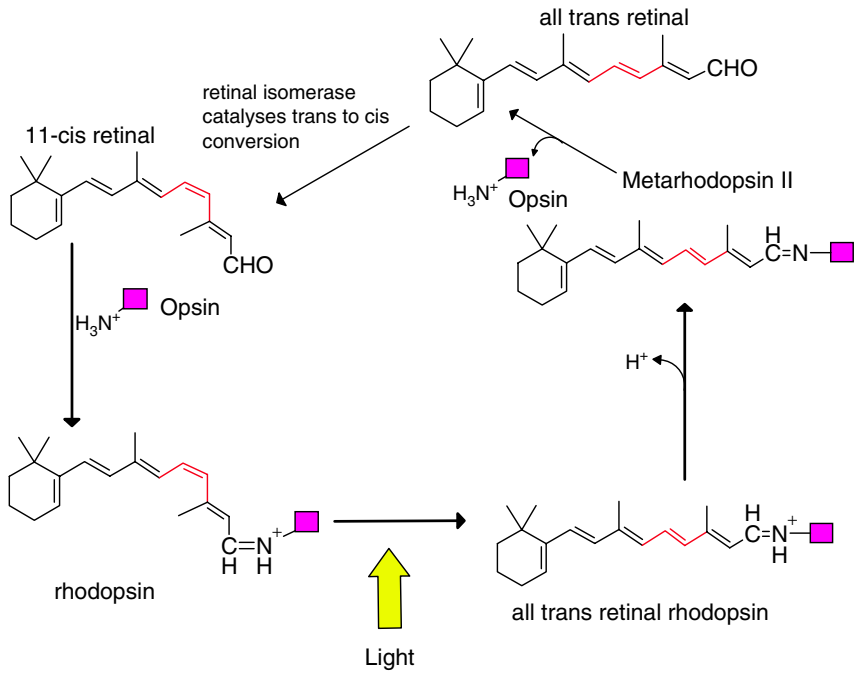


Figure 5.14 The basic visual cycle involving chemical changes in rhodopsin. The critical 11-*cis* bond is shown in red with retinal combining with opsin to form rhodopsin. Light converts the rhodopsin into the *trans*-retinal followed by the formation of metarhodopsin II. This intermediate reacts with transducin before dissociating after ~1 s into *trans*-retinal and opsin

Photochemically induced conformational changes in metarhodopsin II lead to interactions with a second protein transducin found in the disc membrane of retinal rods and cones. Transducin is a heterotrimeric guanine nucleotide binding protein (G protein) containing three subunits designated as α, β and γ of molecular weight 39 kDa, 36 kDa and 8 kDa. Photoexcited rhodopsin interacts with the α subunit of transducin promoting GTP binding in place of GDP and a switch from inactive to active states. The transducin α subunit is composed of two domains. One domain contains six β strands surrounded by six helices whilst the other domain is predominantly helical. Although in the visual cycle the extracellular stimulus and biochemical endpoint are different to hormone action the transmembrane events are remarkably similar.

Interaction between rhodopsin and G protein causes dissociation with the G_α subunit separating from the $G_{\beta\gamma}$ dimer (Figure 5.15). The transducin G_α–GTP complex activates a specific phosphodiesterase enzyme by displacing its inhibitory domain and thus triggering subsequent events in the signalling pathway. The phosphodiesterase cleaves the cyclic nucleotide guanosine 3′, 5′-monophosphate (cyclic GMP) in a process analogous to the production of cyclic AMP in hormone action and this in turn stimulates visual signals to the brain.

As a signalling pathway in humans the rhodopsin–transducin system contributed significantly to our understanding of other G proteins and G-protein coupled receptors. Subsequent work has suggested that all G proteins share a common structure of three subunits. The identification of human cDNA encoding 24 α proteins, 5 β proteins and 6 γ proteins testifies to the large number of G protein catalysed reactions and their importance within cell signalling pathways. The α subunits of G proteins are one group of an

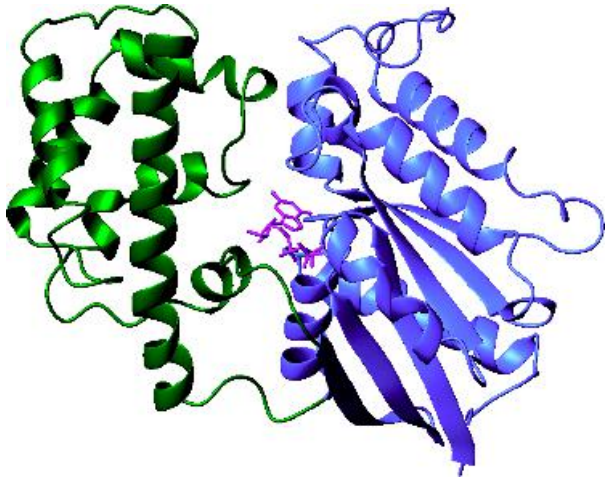


Figure 5.15 The structure of the G_α subunit of transducin. Two domains are shown in dark green (helical) and blue ($\alpha + \beta$) together with a GTP analogue (magenta) bound at the active site (PDB: 1TND)

important class of nucleotide binding proteins whose activity is enhanced by GTP binding and diminished by GDP binding. Other members with this group of G proteins include the elongation factors involved in protein synthesis and the products of oncogenes such as Ras.

Addressing hormone action is an entirely analogous process and is best visualized through inspection of the structure of the β_2 adrenergic receptor. Hormone binding at the extracellular surface promotes conformational change and the interaction of the C-terminal region with the G_α subunit. This stimulates GDP/GTP exchange and in its active state the G_α subunit binds to adenylate cyclase causing the formation of cAMP which functions downstream in protein phosphorylation reactions that stimulate or inhibit metabolic processes (Figure 5.16). G proteins are further sub-divided into two major classes: G_s and G_i. The former stimulate adenylate cyclase whilst the latter inhibit although both interact with a wide range of receptors and interact with other proteins besides adenylate cyclase.

Rhodopsin and β_2 adrenergic receptors are seven transmembrane receptor proteins that activate heterotrimeric GTP binding proteins and are grouped under the term G-protein coupled receptors (GPCRs). GPCRs are part of an even larger group of seven transmembrane receptors and represent the fourth largest superfamily found in the human genome. It is estimated that within the human genome there are over 600 genes encoding seven transmembrane helices receptors. Within this family there are three major structural subdivisions with the rhodopsin-like family representing the largest class. In almost all cases there is a common mechanism involving G protein activation and two types of intracellular or secondary messengers, calcium (Ca^{2+}) and cyclic AMP (cAMP). Approximately 70 % of GPCRs signal through the cAMP pathway and 30 % signal through the Ca^{2+} pathway.

Detailed structural data on GPCRs remains limited and this is unfortunate because many proteins within this family are important and lucrative pharmaceutical targets. Many drugs in clinical use in humans are directed at seven transmembrane receptors. Despite overwhelming sequence data and biophysical data the structure of a non-microbial rhodopsin has only recently been achieved. The crystal structure confirmed, perhaps belatedly, the eponymous seven transmembrane helical model but more importantly this structure provided a model for other GPCRs that allows insight into receptor–G protein complex formation and an understanding of the mechanism by which agonist-receptor binding leads to G protein activation.

The photosynthetic reaction centre of Rhodobacter viridis

Studies of bacteriorhodopsin defined a basic topology of helices traversing the membrane but did not reveal sufficient detail of helical arrangement with respect to each other and the membrane. Detailed analysis of membrane proteins required the use of X-ray crystallography, a technique whose resolution is often superior to electron crystallography. However, this experimental approach requires the formation of three-dimensional crystals of suitable size and this had proved extremely difficult for non-globular proteins. Bacteriorhodopsin forms a two-dimensional array of molecules naturally within the membrane but for almost all other membrane proteins the formation of crystals is a major hurdle to be overcome before acquisition of structural information.

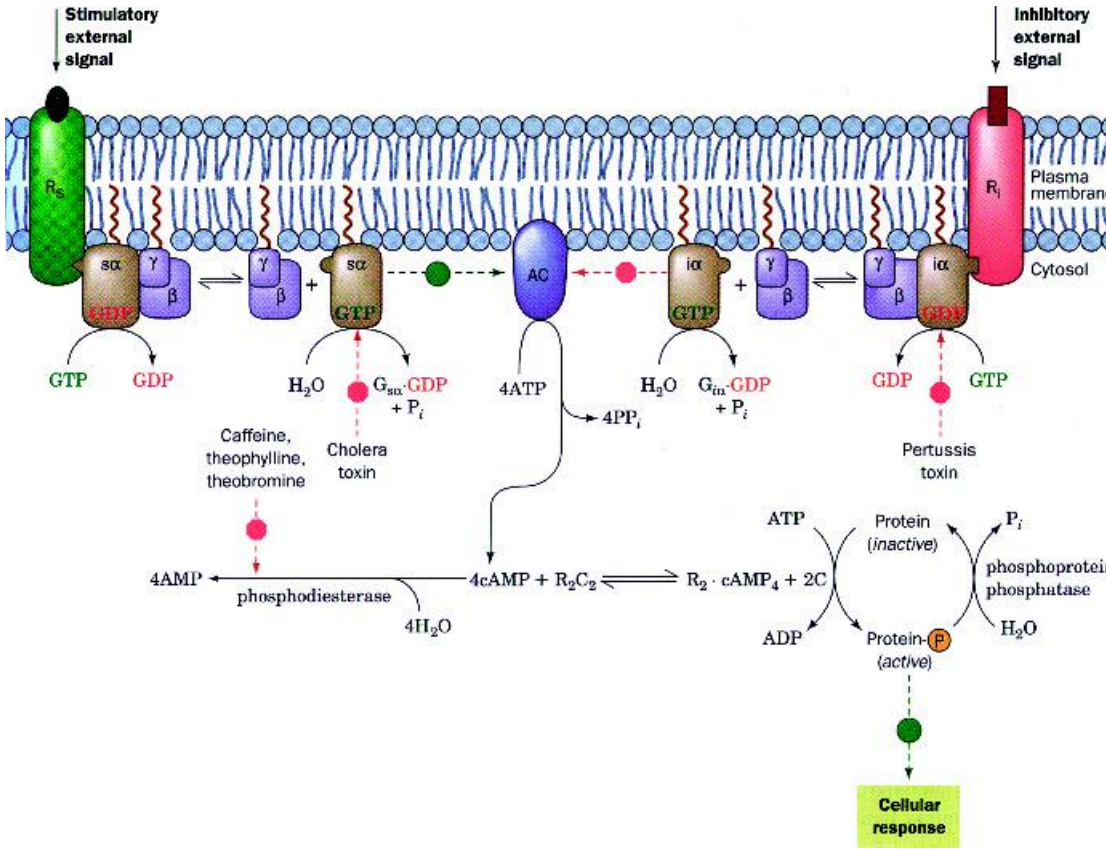


Figure 5.16 The adenylate cyclase signalling system. Binding of hormone to either receptor caused a stimulatory (R_s) or inhibitory (R_i) activity of adenylate cyclase. These receptors undergo conformational changes leading to binding of the G protein and the exchange of GDP/GTP in the Gα subunit. The action of the Gα subunit in stimulating or inhibiting adenylate cyclase continues until hydrolysis of GTP occurs. Adenylate cyclase forms cAMP which activates a tetrameric protein kinase by promoting the dissociation of regulatory and catalytic subunits (reproduced with permission from Voet, C, Voet, J.G & Pratt, C.W. *Fundamentals of Biochemistry* John Wiley & Sons, Ltd, Chichester, 1999).

Removing membrane proteins from their environment as a prelude to crystallization was frequently unsuccessful since the lipid bilayer plays a major role maintaining conformation. Moreover, exposure of hydrophobic residues that were normally in contact with non-polar regions of the bilayer to aqueous solutions resulted in many proteins unfolding with a concomitant loss of activity. One approach adopted in an attempt to improve methods of membrane protein crystallization was to substitute detergents that mimicked the structure and function of lipids. In many instances these detergents failed to preserve structure or function satisfactorily. The situation changed drastically in 1982 when Hartmut Michel prepared highly ordered crystals of the photosynthetic reaction centre from *Rhodobacter viridis* (originally *Rhodopseudomonas viridis*).

The reaction centre is a macromolecular complex located in heavily pigmented membranes of photosynthetic organisms (Figure 5.17). Within most

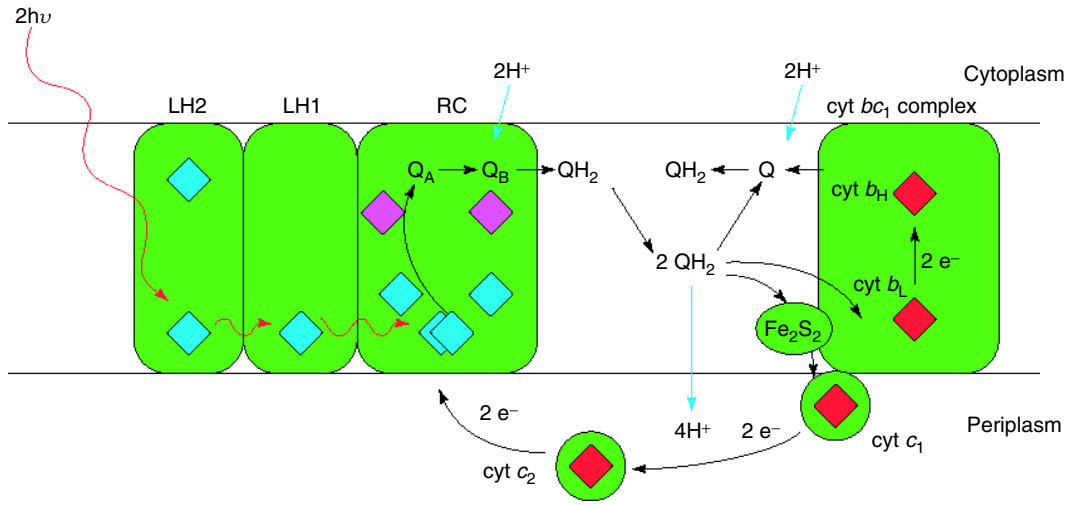


Figure 5.17 The organization of reaction centres and the photosynthetic electron transfer system in purple bacteria such as *R. sphaeroides*. In *R. viridis* a bound cytochrome is linked to the reaction centres but is absent in most species. Bacteriochlorophyll molecules are represented by blue diamonds, bacteriopheophytin molecules by purple diamonds and hemes by red diamonds. The red arrows represent excitation transfer from light harvesting chlorophyll complexes and the black and blue arrows correspond to electron and proton transfer, respectively (reproduced with permission Verméglio, A & Joliot, P. *Trends Microbiol.* 1999, **7**, 435–440. Elsevier).

bacterial membranes the centre is part of a cyclic electron transfer pathway using light to generate ATP production.

Absorption of a photon by light-harvesting complexes (LH1 and LH2) funnels energy to the reaction centre initiating primary charge separation where an electron is transferred from the excited primary donor, a bacteriochlorophyll dimer, to a quinone acceptor (Q_B) via other pigments. After a second turnover, the reduced quinone picks up two protons from the cytoplasmic space to form a quinol (QH_2) and is oxidized by cytochrome c_2, a reaction catalysed by the cytochrome bc_1 complex. This reaction releases protons to the periplasmic space for use in ATP synthesis and the cyclic electron transfer pathway is completed by the reduction of the photo-oxidized primary electron donor.

Fractionation studies of *R. viridis* membranes using detergents such as lauryl dimethylamine oxide (LDAO) yield a 'core' reaction centre complex containing four polypeptides lacking the light-harvesting chlorophyll and cytochrome bc_1 complexes normally associated with these centres. The subunits (H, M, and L) possessed all of the components necessary for primary photochemistry and were associated with a membrane bound cytochrome.

Analysis of the photosynthetic complex showed that its ability to perform primary photochemical reactions was due to the presence of chromophores that included four bacteriochlorophyll molecules, two molecules of a second related pigment called bacteriopheophytin, carotenoids, a non-heme iron centre, and two quinone molecules (ubiquinone and menaquinone). Using spectroscopic techniques such as circular dichroism, laser flash photolysis and electron spin resonance the order and time constants associated with electron transfer processes was established (Figure 5.18). In all cases electron transfer was characterized by extremely rapid reactions. The light harvesting complexes act as antennae and transfer energy to a pair of chlorophyll molecules (called P870, and indicating the wavelength associated with maximum absorbance) in the reaction centre located close to the periplasmic membrane surface. Absorption of a photon causes an excited state and at room temperature the photo-excited pair reduces

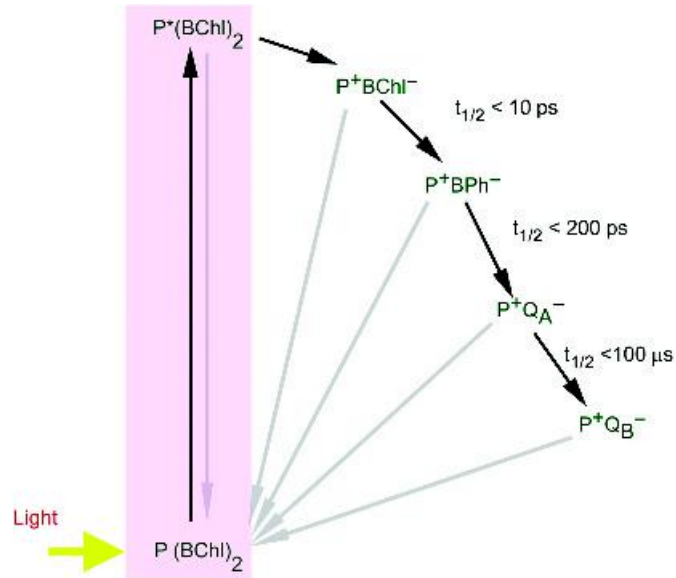


Figure 5.18 The sequence of donors and acceptors in the reaction centres of *R. viridis*. Q_A is menaquinone and Q_B is ubiquinone. Rapid back reactions occur at all stages and the primary photochemical reactions occur very rapidly (ps timescale) presenting numerous technical difficulties in accurate measurement

a pheophytin molecule within 3 ps. The initial primary charge separation generates a strong reductant and reduces the primary acceptor bacteriophaeophytin. Within a further 200 ps the electron is transferred from bacteriopheophytin to a quinone molecule. These reactions are characterized by high quantum efficiency and the formation of a stable charge separated state. Normally the quinone transfers an electron to a second quinone within 25 μs before further electron transfer to other acceptors located within membrane protein complexes. In isolated reaction centres electrons reaching the quinone recombine with a special pair of chlorophyll by reverse electron transfer.

The electron transfer pathway was defined by spectroscopic techniques but the arrangement of chromophores together with the topology of protein subunits remained unknown until Michel obtained ordered crystals of reaction centres.

Crystallization requires the ordered precipitation of proteins. Membrane proteins have both hydrophilic and hydrophobic segments due to their topology, with the result that the proteins are rarely soluble in either aqueous buffers or organic media. By partitioning the membrane protein in detergent (LDAO) micelles it was possible to purify subunits but it proved very difficult to crystallize reaction centres from LDAO. As a general rule macromolecular protein complexes will only aggregate into ordered crystalline arrays if their exposed polar surfaces can approach each other closely. The use of detergents, particularly if they formed micelles of large size, tended to limit the close approach of polar surfaces preventing crystallization. The major advance of Michel was to introduce small amphiphilic molecules such as heptane 1,2,3 triol into the detergent mixture. Although the precise mechanism of action of these amphiphiles is unclear they are thought to displace large detergent molecules from the crystal lattice, thereby assisting aggregation. In addition they do not form micelles themselves but lower the effective concentration of detergent by becoming incorporated within micelles and limiting protein denaturation. Lastly, amphiphiles interact with protein but do not prevent close proximity of solvent-exposed polar regions.

The structure of the reaction centre revealed the topology of structural elements within the H,M,L

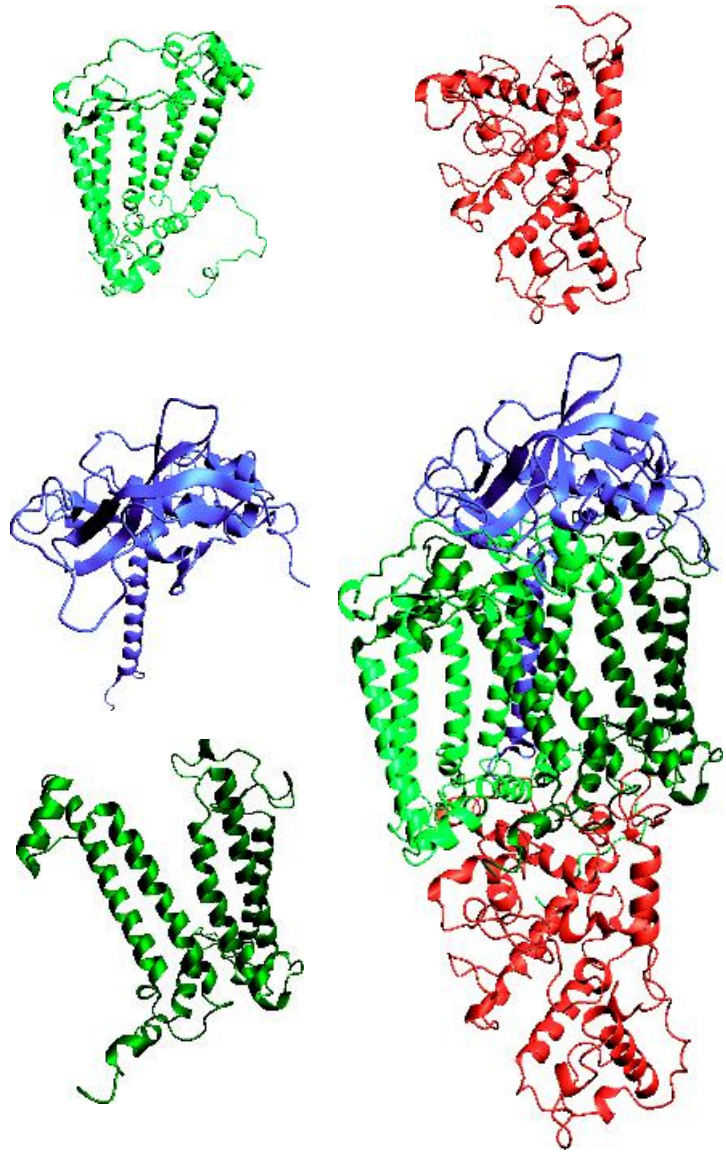


Figure 5.19 Distribution of secondary structure in each subunit of the photosynthetic reaction centre together with the whole complex. The cytochrome subunit is shown in orange, the L subunit in light green, the M subunit in dark green and the H subunit in blue (PDB:1PRC)

and cytochrome subunits and importantly the position of the pigments (Figure 5.19). The structure offered the possibility of understanding how reaction centres achieved picosecond charge separation following light absorption.

The structure of the bacterial reaction centre

One surprise from the crystal structure of the four-subunit complex was the H subunit. Sequencing the

gene for subunit H revealed that it was the smallest subunit in terms of number of residues. Its migration on SDS gels was anomalous when compared to the L and M subunits leading to slightly inaccurate terminology. Careful inspection of Figure 5.19 reveals that the H subunit is confined to the membrane by a single transmembrane α helix (shown in blue) whilst most of the protein is found as a cytoplasmic domain. The physiological role of the H subunit is unclear since it lacked all pigments. The structure of subunit H is divided into three distinct segments; a membrane-spanning region, a surface region from residues 41 to 106, and a large globular domain formed by the remaining residues. The membrane-spanning region of the H subunit is a single regular α helix from residues 12 to 35 with some distortion near the cytoplasmic membrane surface. Following the single transmembrane helix are ∼65 residues forming a surface region interacting with the L and M subunits and based around a single α helix and 2 β strands. In the cytoplasmic domain of subunit H an extended system of antiparallel and parallel β strands are found together with a single α helix from residues 232 to 248.

The L and M subunits each contain five transmembrane helices (Table 5.6) but more significantly each subunit is arranged with an approximate two-fold symmetry that allows the superposition of atoms from one subunit with those on the other. The superposition is surprising and occurs with low levels of sequence identity (∼26 percent). In each subunit the five helices are designated A–E; three (A, B and D) are extremely regular and traverse the membrane as straight helices whilst C and E are distorted. Helix E in both subunits shows a kink caused by a proline residue that disrupts regular patterns of hydrogen bonding. Although the helices show remarkably regular torsion angles and hydrogen bonding patterns they are inclined at angles of ∼11° with respect to the membrane. Elements of β strand were absent and the results confirmed those obtained with bacteriorhodopsin. In the reaction centre between 22 and 28 residues were required to cross the membrane each helix linked by a short loop.

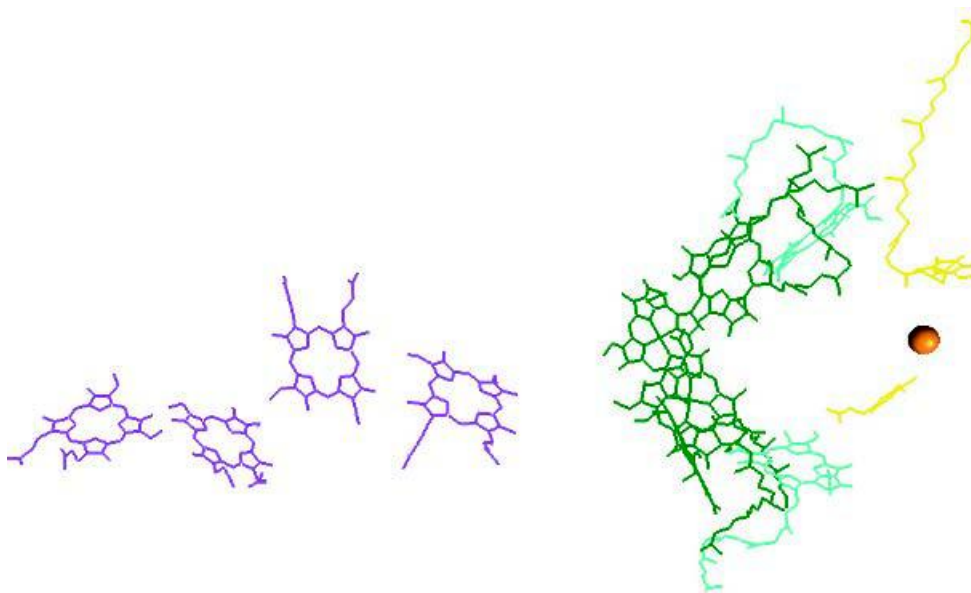


Figure 5.20 The arrangement of heme, bacteriochlorophyll, bacteriopheophytin, quinone and non heme iron in the hydrophobic region between L and M subunits. Electron transfer proceeds from the hemes shown on the left through to the quinones shown on the right

Table 5.6 The transmembrane helices found in the L/M subunits of the reaction centre

Helical segments in subunits L and M		
Region and number of residues in brackets		
Helix	Subunit L	Subunit M
A	L33–L53 (21)	M52–M76 (25)
B	L84–L111 (28)	M111–M137 (27)
C	L116–L139 (24)	M143–M166 (24)
D	L171–L198 (28)	M198–M223 (26)
E	L226–L249 (24)	M260–M284 (25)

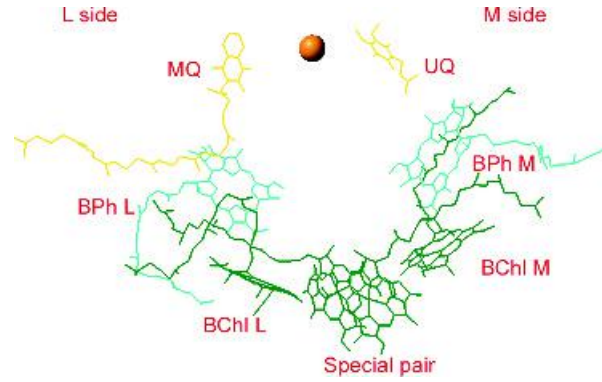


Figure 5.21 Arrangement of chromophores in the bacterial reaction centre. The cytochrome redox co-factors have been removed

For the L and M subunits the five transmembrane helices defined the bilayer region of the membrane. Surprisingly, direct interactions between subunits are limited by the presence of chromophores between the subunits and most of the subunit contacts were restricted to periplasmic surfaces close to the H subunit.

The cytochrome subunit contained 336 residues with a total of nine helices and four heme groups covalently linked to the polypeptide via two thioether bridges formed with cysteine residues. The fifth and six ligands to the hemes were methionine and histidine for three of the centres whilst the fourth showed bis-histidine ligation. The heme groups have different redox potentials and serve to shuttle electrons to the special pair chlorophyll after excitation. Although predominantly in the aqueous phase the cytochrome remains membrane bound by a covalently linked diglyceride attached to the N-terminal Cys residue.

It is, however, the arrangement of pigments that provides the most intriguing picture and sheds most light on biological function. The active components ($4 \times$ BChl, $2 \times$ BPh, $2 \times$ Q + Fe) are confined to a region between the L and M subunits in a hydrophobic core with each chromophore bound via specific side chain ligands. The special pair of chlorophyll molecules is located on the cytoplasmic side of the membrane close to one of the heme groups of the cytochrome. Stripping the L and M polypeptides reveals the two-fold symmetry of the chromophores (see Figures 5.20 & 5.21). The structural picture agrees well with the order of electron transfer established earlier using kinetic techniques as

$$B(Chl)_2 \longrightarrow BPh \longrightarrow Q_A \longrightarrow Q_B \longrightarrow \text{other acceptors}$$

The additional chlorophylls found in the reaction centre do not participate directly in the electron transfer process and were called ‘voyeur’ chlorophylls.

The co-factors between the L and M subunits are arranged as two arms diverging from the special pair of chlorophylls. A single ‘voyeur’ chlorophyll together with a pheophytin molecule lies on each side and they lead separately to menaquinone and ubiquinone molecules with a single non-heme iron centre located between these acceptors. It was therefore a considerable surprise to observe that despite the high degree of inherent symmetry in the arrangement of co-factors the left hand branch (L side) was the more favoured electron transfer route by a ratio of 10:1 over the right hand or M route.

The underlying reasons for asymmetry in electron transfer are not clear but rates of electron transfer are very dependent on free energy differences between excited and charge separated states as well as the distance between co-factors. Both of these parameters are influenced by the immediate protein surroundings of each co-factor and studies suggest that rates of electron transfer are critically coupled to the properties of the bacteriochlorophyll monomer (the voyeur) that

lies between the donor special pair and bacteriopheophytin acceptor.

The X-ray crystal structure of the *R. viridis* reaction centre was determined in 1985 and was followed by the structure of the *R. sphaeroides* reaction centre that confirmed the structural design and pointed to fold conservation. Research started with crystallization of the bacterial photosynthetic resulted in one of the outstanding achievements of structural biology and Hartmut Michel, Johan Diesenhofer, and Robert Huber were rewarded with the Nobel Prize for Chemistry in 1988. Over the last two decades this membrane protein has been extensively studied as a model system for understanding the energy transduction, the structure and assembly of integral membrane proteins, and the factors that govern the rate and efficiency of biological electron transfer.

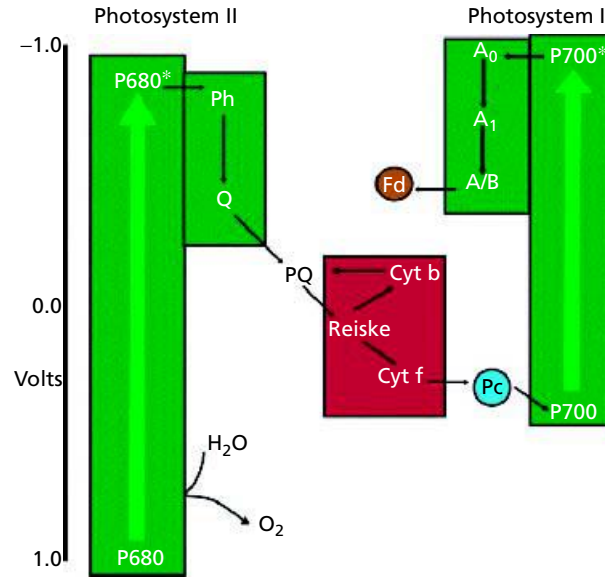


Figure 5.22 The oxygenic photosynthetic system found in cyanobacteria and eukaryotic algae and higher plants. Two photosystems acting in series generate reduced NADPH and ATP via the absorption of light. Photosystem II precedes photosystem I

Oxygenic photosynthesis

The photosynthetic apparatus of most bacteria is simpler than that found in algae and higher plants. Here two membrane-bound photosystems (PSI and PSII) act in series splitting water into molecular oxygen and generating ATP/NADPH via photosynthetic electron transfer (Figure 5.22). The enzymatic splitting of water is a unique reaction in nature since it is the only enzyme to use water as a principal substrate. Oxygen is one of the products and the evolution of this reaction represents one of the most critical events in the development of life. It generated an atmosphere rich in oxygen that supported aerobic organisms. The geological record suggests the appearance of oxygen in the atmosphere fundamentally changed the biosphere.

The water splitting reaction is carried out by PSII and involves Mn ions along with specific proteins. Within PSII are two polypeptides called D1 and D2 that show a low level of homology with the L and M subunits of anoxygenic bacteria. The structure of PSII from the thermophilic cyanobacterium (sometimes called blue–green algae) *Synechococcus elongatus* confirmed structural homology with bacterial reaction centres. PSII is a large membrane-bound complex comprising of at least 17 protein subunits, 13 redox-active co-factors and possibly as many as 30 accessory chlorophyll pigments. Of these subunits 14 are located within the photosynthetic membrane and include most significantly the reaction centre proteins D1 (*PsbA*) and D2 (*PsbD*). The observation that D1 and D2 each contained five transmembrane helices arranged in two interlocked semicircles and related by the pseudo twofold symmetry axis emphasized the staggering similarity to the organization of the L and M subunits (Figure 5.23). Moreover the level of structural homology between D1/D2 and L/M subunits suggested that all photosystems evolved from common ancestral complexes.

Photosystem I

Oxygenic photosynthesis requires two photosystems acting in series and the structure of the cyanobacterial PSI reaction centre further confirmed the design pattern of heterodimers (Figure 5.24). Two subunits, PsaA

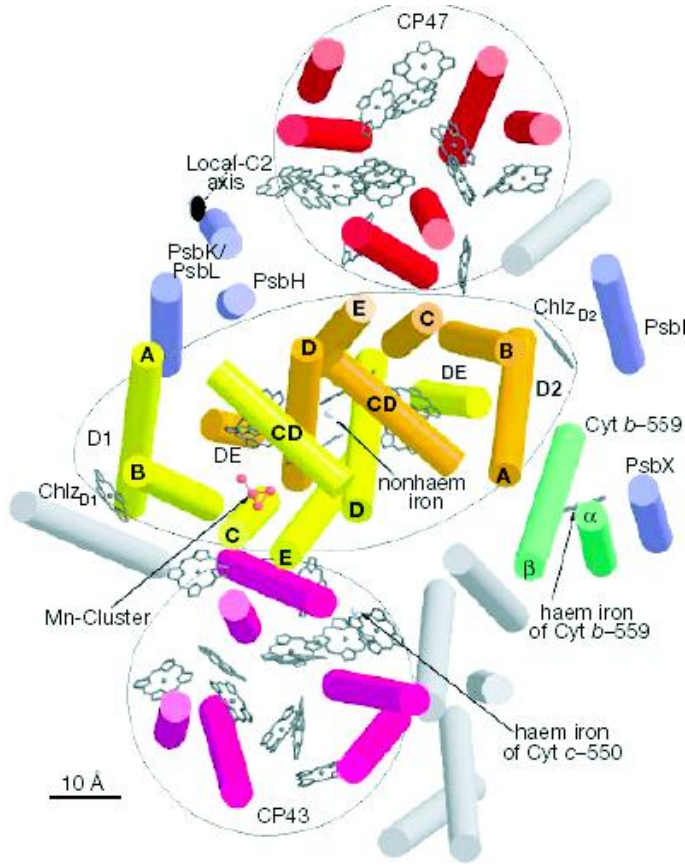


Figure 5.23 Structure of PSII with assignment of protein subunits and co-factors. Arrangement of transmembrane α helices and co-factors in PSII. Chl a and hemes are indicated by grey wireframe representations. The direction of view is from the lumenal side, perpendicular to the membrane plane. The α helices of D1, D2 are shown in yellow and enclosed by an ellipse. Antennae chlorophylls in subunits CP43 and CP47 are enclosed by circles. Unassigned α helices are shown in grey (reproduced with permission from Zouni, A, *et al. Nature* 2001 **409** 739–743. Macmillan)

and PsaB make up a reaction centre that contains eleven transmembrane helices. By now this sounds very familiar, and five helices from each subunit traverse the membrane and together with a third protein, PsaC, contain all of the co-factors participating in the primary photochemistry. A further six subunits contribute transmembrane helices to the reaction centre complex whilst two subunits (PsaD and E) bind closely to the complex on the stromal side of the membrane.

Support for the hypothesis that reaction centres evolved from a common ancestor were strengthened further by the observation that the five C-terminal helices of PsaA and PsaB in PSI showed homology to the D1/D2 and L/M heterodimers described previously. Reaction centres are based around a special pair of chlorophyll molecules that transfer electrons via intermediary monomeric states of chlorophyll and quinone molecules to secondary acceptors that are quinones in type II reaction centres and iron–sulfur proteins in type I reaction centres. Figure 5.25 summarizes the organization of reaction centre complexes.

Oxygenic photosynthesis is the principal energy converter on Earth converting light energy into biomass and generating molecular oxygen. As a result of the

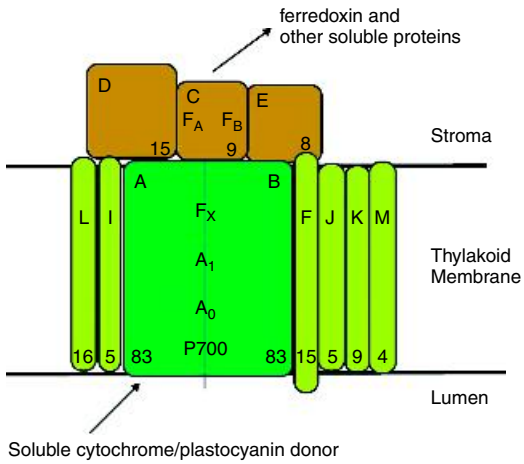


Figure 5.24 The organization of PSI showing subunit composition and co-factors involved in electron transfer

success of membrane protein structure determination we are beginning to understand the molecular mechanisms underlying catalysis in complexes vital for continued life on this planet.

Membrane proteins based on transmembrane β barrels

So far membrane proteins have been characterized by the exclusive use of the α helix to traverse membranes. However, some proteins form stable membrane protein structure based on β strands. Porins are channel-forming proteins found in the outer membrane of Gram-negative bacteria such as *Escherichia coli* where they facilitate the entry of small polar molecules into the periplasmic space. Porins are abundant proteins within this membrane with estimates of ~100 000

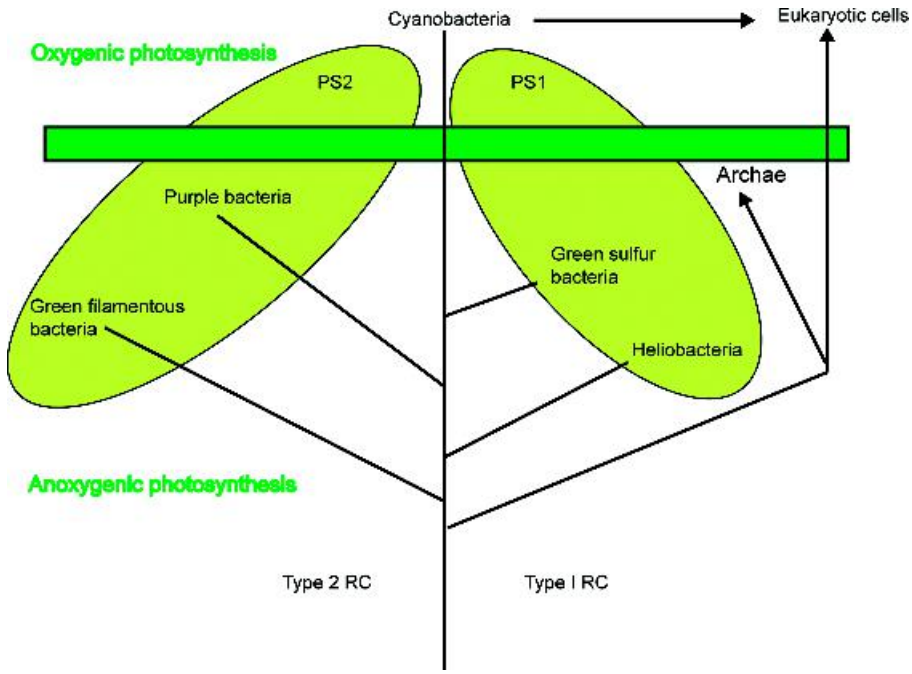


Figure 5.25 A summary of the distribution of photosynthetic reaction centre complexes in pro- and eukaryotes. Oxygenic photosynthesis is based around two reaction centre complexes, PSI and PSII, while anoxygenic photosynthesis uses only one reaction centre. This reaction centre can be either type II or type I. Type-II reaction centres are found in purple non-sulfur bacteria, green filamentous bacteria and oxygenic PSII and type-I reaction centres are typical of green sulfur bacteria, heliobacteria and oxygenic PSI

copies of this protein per cell. Porins have a size limit of less than 600 for transport by passive diffusion through pores. Although most porins form non-specific channels a few display substrate-specificity such as maltoporin.

The OmpF porin from the outer membrane *of E. coli* contains 16 β strands arranged in a barrel-like fold (Figure 5.26). The significance of this topology is that it overcomes the inherent instability of individual strands by allowing inter-strand hydrogen bonding. OmpF appears to be a typical *E. coli* porin and will serve to demonstrate structural properties related to biological function. Although shown as a monomer the functional unit for porins is more frequently a trimeric complex of three polypeptide barrels packed closely together within the bilayer.

The individual OmpF monomer contains 362 residues within a subunit of 39.3 kDa where the length of the monomer (5.5 nm) is sufficient to span a lipid bilayer and to extrude on each side into the aqueous phase. In the barrel hydrophobic groups, particularly aromatic side chains, occupy regions directly in contact with the lipid bilayer (Figure 5.27). As a consequence porins have a 'banded' appearance when the distribution of hydrophobic residues is plotted over the surface of the molecule. The barrel fold creates a cavity on the inside of this protein that is lined with polar groups. These groups located away from the hydrophobic bilayer form hydrogen bonds with water found in this cavity.

The cavity itself is not of uniform diameter – the pore range is 1.1 nm at its widest but it is constricted near the centre with the diameter narrowing to ~0.7 nm. In this cavity are located side chains that influence the overall pore properties. For example, OmpF shows a weak cation selectivity whilst highly homologous porins such as PhoE phosphoporin (63 percent identity) shows a specificity towards anions. The molecular basis for this charge discrimination has become clearer since the structures of both porins have been determined. A region inside the barrel contributes to a constriction of the channel and in most porins this region is a short α helix. In OmpF the short helix extends from residues 104 to 112 and is followed by a loop region that, instead of forming the hairpin

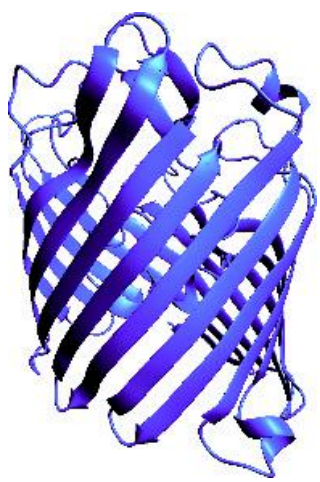


Figure 5.26 The arrangement of strands in OmpF from the outer membrane of *E. coli*. OmpF stands for outer membrane protein f. The strands make an angle of between 35 and 50° with the barrel axis with the barrel often described as having a rough end, marked by the presence of extended loop structures and a smooth end where shorter loop regions define the channel

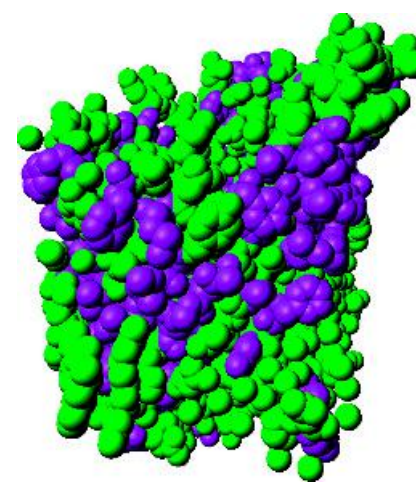


Figure 5.27 Distribution of hydrophobic residues around the barrel of OmpF porin. The space filling model of porin shows the asymmetric distribution of hydrophobic side chains (Gly, Ala, Val, Ile, Leu, Phe, Met, Trp and Pro) in a relatively narrow band. This band delineates the contact region with the lipid bilayer

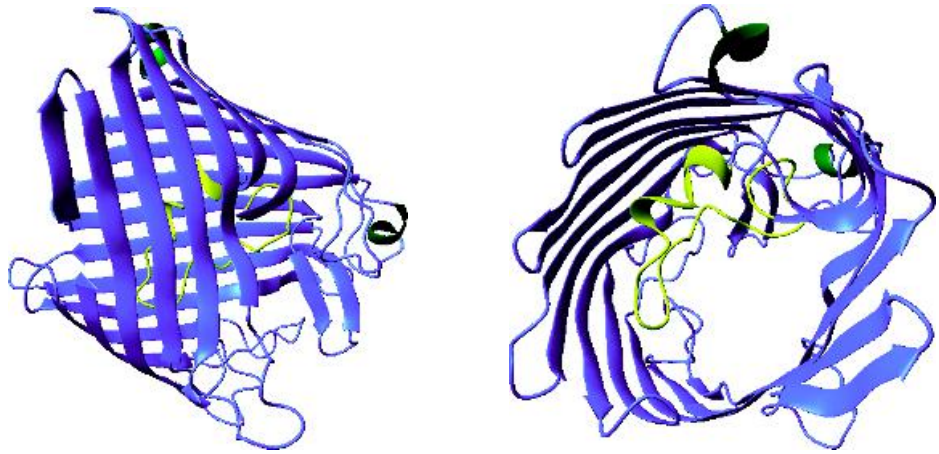


Figure 5.28 Two views of the OmpF monomer showing the L3 loop region (yellow) located approximately half way through the barrel and forming a constriction site

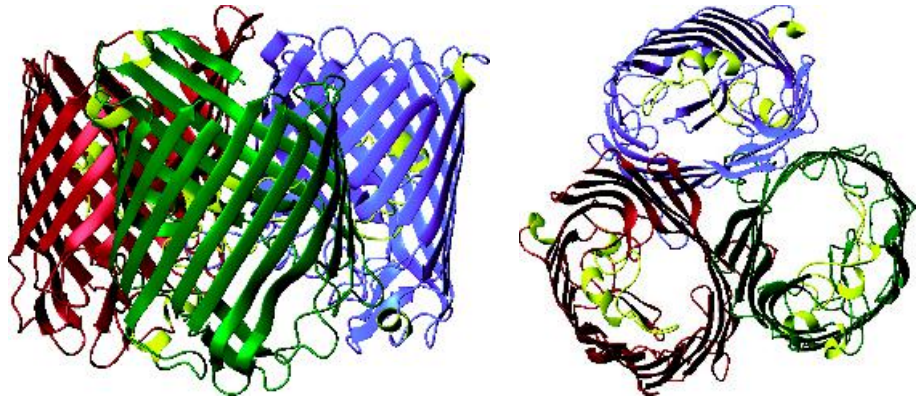


Figure 5.29 The biological unit of OmpF (and other porins) is a trimer of barrels with interaction occurring via the hydrophobic surfaces of each monomer. The three monomers are shown in a side view (viewed from a position within the membrane) and a top view

structure between two strands, projects into the channel (Figure 5.28). In the vicinity of this constriction but on the opposite side of the barrel are charge side chains that influence ion selectivity. The organization and properties of OmpF is shared by the PhoE porin but also by evolutionary more distant porins from *Rhodobacter* and *Pseudomonas* species.

Porins are stable proteins because the 16-stranded antiparallel β-barrel is closed by a salt bridge formed between the N- and C-terminal residues (Ala and Phe, respectively) whilst the biological unit packs together as a trimeric structure enhanced in stability by hydrophobic interactions between the sides of the barrel (Figure 5.29).

Another important example of a membrane protein based on a β barrel is α-haemolysin, a channel-forming toxin, exported by the bacterium *Staphylococcus aureus*. Unusually this protein is initially found as a water soluble monomer (M_r~33 kDa) but assembles in a multistep process to form a functional heptamer on

Figure 5.30 The structure of α-haemolysin showing the complex arrangement of β strands. The heptamer is mushroom shaped as seen from the above figure with the transmembrane domain residing in the 'stalk' portion of the mushroom and each protomer contributing two strands of the 14-stranded β barrel. This arrangement is not obvious from the above structure but becomes clearer when each monomer is shown in a different colour and from an altered perspective (Figure 5.31)

the membranes of red blood cells. Once assembled and fastened to the red blood cell membrane α-haemolysin causes haemolysis by creating a pore or channel in the membrane of diameter ~2 nm. If left untreated infection will result in severe illness.

The monomeric subunit associates with membranes creating a homoheptameric structure of ~230 kDa that forms an active transmembrane pore that allows the bilayer to become permeable to ions, water and small solutes. The three-dimensional structure of α-haemolysin (Figures 5.30 and 5.31) shows a structure containing a solvent-filled channel of length 10 nm that runs along the seven-fold symmetry axis of the protein. As with porins the diameter of this channel varies between 1.4 and 4.6 nm.

Porins and the α-haemolysins make up a large class of membrane proteins based on β barrels. Previously thought to be just an obscure avenue of folding followed by a few membrane proteins it is now clear that this class is increasing in their number and distribution. Porins are found in a range of bacterial membranes as well as the mitochondrial and chloroplast membranes. The OmpF structure is based on 16 β strands and porins from other membranes

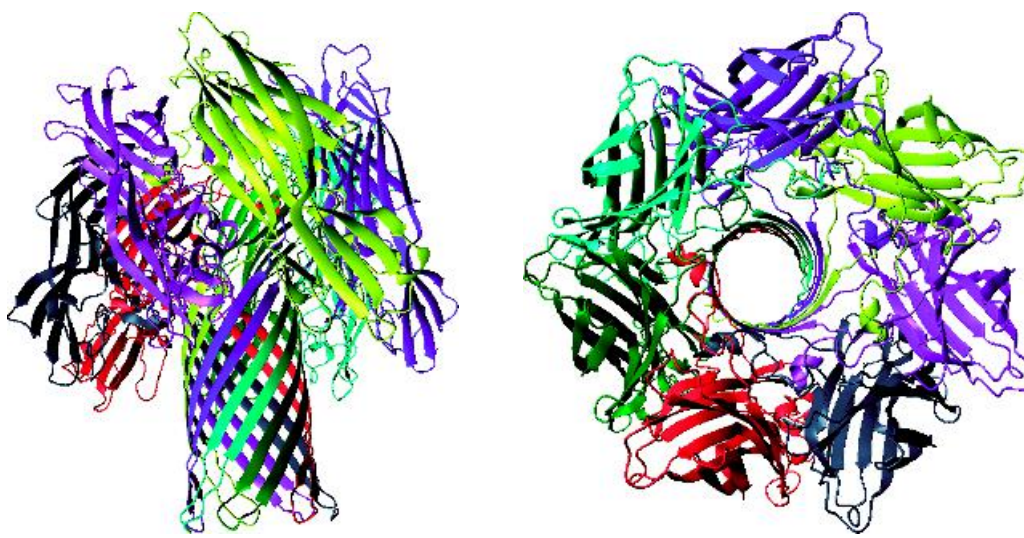


Figure 5.31 The structure of α-haemolysin showing the contribution of two strands from each protomer to the formation of a 14-stranded β barrel representing the major transmembrane channel. The channel or pore is clearly visible from a top view looking down on the α-haemolysin (PDB: 7AHL)

contain between 8 and 22 strands (in all cases the number of strands is even) sharing many structural similarities.

Respiratory complexes

Structure determination of bacterial reaction centres provided enormous impetus to solving new membrane protein structures. In the next decade success was repeated for membrane complexes involved in respiration.

Respiration is the major process by which aerobic organisms derive energy and involves the transfer of reducing equivalents through series of electron carriers resulting in the reduction of dioxygen to water. In eukaryotes this process is confined to the mitochondrion, an organelle with its own genome, found in large numbers in cells of metabolically active tissues such as flight muscle or cardiac muscle. The mitochondrion is approximately 1–2 μm in length although it may sometimes exceed 10 μm, with a diameter of ~0.5–1 μm and contains a double membrane. The outer membrane, a semipermeable membrane, contains porins analogous to those described in previous sections. However, the inner mitochondrial membrane is involved in energy transduction with protein complexes transferring electrons in steps coupled to the generation of a proton gradient. The enzyme ATP synthetase uses the proton gradient to make ATP from ADP in oxidative phosphorylation.

Purification of mitochondria coupled with fractionation of the inner membrane and the use of specific inhibitors of respiration established an order of electron transfer proteins from the oxidation of reduced substrates to the formation of water. The inner mitochondrial membranes called cristae are highly invaginated with characteristic appearances in electron micrographs and are the location for almost all energy transduction. This process is dominated by four macromolecular complexes that catalyse the oxidation of substrates such as reduced nicotinamide adenine dinucleotide (NADH) or reduced flavin adenine dinucleotide ($FADH_2$) through the action of metalloproteins such as cytochromes and iron sulfur proteins (Figure 5.32).

Some complexes of mitochondrial respiratory chains are amenable to structural analysis and although structures of complex I (NADH-ubiquinone oxidoreductase; see Figure 5.33) and complex II (succinate dehydrogenase) are not available, advances in defining complex III (ubiquinol-cytochrome c oxidoreductase) and the final complex of the respiratory chain, complex IV or cytochrome oxidase, have assisted understanding of catalytic mechanisms.

Complex III, the ubiquinol-cytochrome c oxidoreductase

The cytochrome bc family developed early in evolution and universally participates in energy transduction. All but the most primitive members of this family of protein complexes contain a b type cytochrome, a c type cytochrome[1] and an iron sulfur protein (ISP). In addition there are often proteins lacking redox groups with non-catalytic roles controlling complex assembly. In all cases the complex oxidizes quinols and transfers electrons to soluble acceptors such as cytochrome c.

Two-dimensional crystals of low resolution (~20 Å) defined the complex as a dimer with ‘core’ proteins lacking co-factors projecting out into solution. Much later the structure of the extrinsic domains of the iron–sulfur protein and the bound cytochrome were derived by crystallography but the main body of the complex, retained within the confines of the lipid bilayer, remained structurally ‘absent’. At the beginning of the 1990s several groups obtained diffraction quality crystals of the cytochrome bc_1 complex from bovine and chicken mitochondria. An extensive structure for the cytochrome bc_1 complex from bovine heart mitochondria was produced by S. Iwata and co-workers in 1998 at a resolution of 3 Å after several groups had produced structures for the majority of the complex. The structures showed similar organization although small structural differences were noticed and appeared to be of functional importance (see below).

Isolated cytochrome bc_1 complexes from eukaryotic organisms contain 10 (occasionally 11) subunits

[1]In plants and algae the c type cytochrome is sometimes called cytochrome f but in all other respects it has analogous structure and function. The f refers to the latin for leaf, *frons*

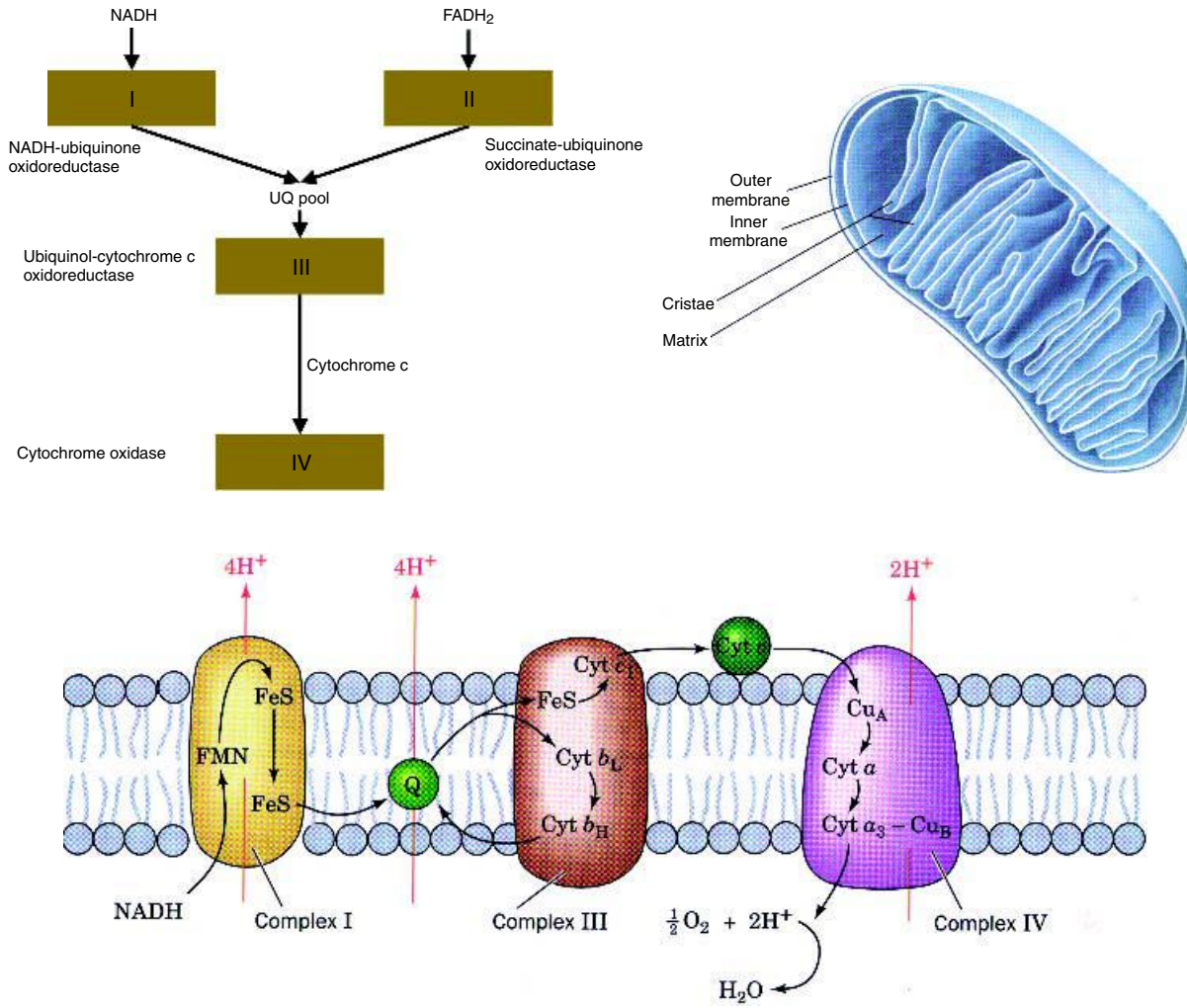


Figure 5.32 The organization of the respiratory chain complexes within the inner mitochondrial membrane. The traditional view of mitochondria (centre) was derived from electron microscopy of fixed tissues

including a b type cytochrome with two heme centres, an iron–sulfur protein (called the Reiske protein after its discoverer) and a mono heme c type cytochrome. DNA sequences for all subunits are known and considerable homology exists between the three metazoan branches represented by bovine, yeast and potato. Detailed UV-visible absorbance studies combined with potentiometric titrations established that the two hemes of cytochrome b had slightly different absorbance maxima in the reduced state (α bands at 566 and 562 nm) and different equilibrium redox potentials (E_m~120 and −20 mV respectively). This led to each heme of cytochrome b being identified as b_H or b_L for high and low potential.

Further understanding of electron transfer in the bc_1 complex came from detailed spectroscopic investigations coupled with the use of inhibitors that block electron (and proton) transfer at specific sites by binding to different protein subunits. One commonly used inhibitor is an anti-fungal compound called

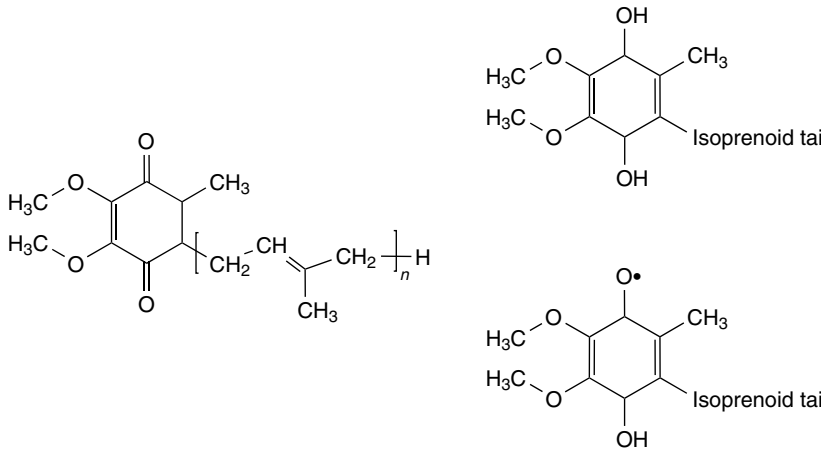


Figure 5.33 The structure of ubiquinone, sometimes called coenzyme Q_{10}, contains an isoprenoid tail linked to a substituted benzoquinone ring. The isoprenoid tail contains a series of isoprenyl units (CH_2–CH=C(CH_3)–CH_2). Various numbers of these units can occur but 10 is the most common leading to the designation UQ_{10} or Q_{10}. The molecules form semiqinone and quinol states with the pattern of reduction complicated by protonation leading to species such as anionic semiquinones

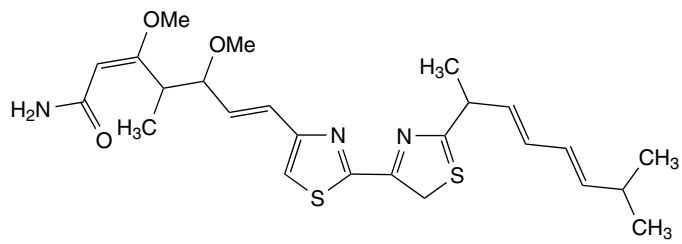


Figure 5.34 The mitochondrial cytochrome bc_1 complex inhibitor myxathiazol

myxathiazol (Figure 5.34) that binds to the Reiske iron–sulfur protein blocking quinol oxidation. It was observed to block reduction of the low potential b heme. Other inhibitors such as stigmatellin, 5-*n*-undecyl-6-hydroxy-4,7-dioxobenzothiazol (UHDBT) and β-methoxyacrylate (MOA)–stilbene showed similar inhibitory profiles. Another group of inhibitors bound differently to the bc_1 complex allowing reduction of the iron-sulfur centre and cytochrome b but prevented their oxidation. These inhibitors included antimycin A and led to the idea of two distinct inhibitor-binding sites within the cytochrome bc_1 complex.

Inhibition of electron transfer by antimycin A yielded unexpected patterns of reduction and oxidation; the normal reduction of cytochrome c_1 by ubiquinol was prevented but was accompanied by an increased reduction of cytochrome b. This observation indicated a branch point in the electron transfer chain where electrons were shuttled to cytochromes c_1 *and* b. The electron transfer pathway (Figure 5.35) from ubiquinol to cytochrome c was rationalized by the Q cycle. Whilst cytochrome c_1 is oxidized by the soluble electron carrier (cytochrome c) the only possible acceptor of electrons from cytochrome b (b_H) is ubiquinone.

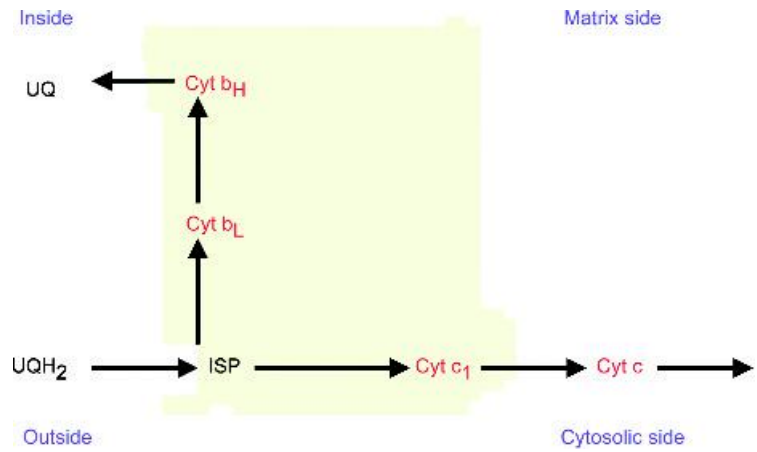


Figure 5.35 Outline of electron transfer reactions in cytochrome bc_1 complex. UQ/UQH_2 is ubiquinone/ubiquinol; ISP, Reiske iron sulfur protein; Cyt c_1, cytochrome c_1; b_L and b_H low and high potential hemes of cytochrome b

The Q cycle accounts for the fact that ubiquinol is a $2e/2H^+$ carrier whilst the bc_1 complex transfers electrons as single entities. Quinones have roles as oxidants *and* reductants. Quinol is oxidized at a Q_o site within the complex and one electron is transferred to the high-potential portion of the bc_1 complex whilst the other electron is diverted to the low-potential part (cytochrome b) of the pathway. The high-potential chain, consisting of the Reiske iron–sulfur protein, cytochrome c_1 and the natural acceptor of the inter-membrane space (cytochrome c) transfers the first electron from quinol through to the final complex of the respiratory chain, cytochrome oxidase. The low-potential portion consists of two cytochrome b hemes and forms a separate pathway for transferring electrons across the membrane from the Q_o site to a Q_i site. The rate-limiting step is quinol (QH_2) oxidation at the Q_o site and it is generally supposed that an intermediate semiquinone is formed at this site. In order to provide two electrons at the Q_i site for reduction of quinone, the Q_o site oxidizes two molecules of quinol in successive reactions. The first electron at the Q_i site generates a semiquinone that is further reduced to quinol by the second electron from heme b_H. The overall reaction generates four protons in the inter-membrane space, sometimes called the P-side, and the formation of quinol uses up two protons from the mitochondrial matrix. The whole reaction is summarized by the scheme

$$QH_2 + 2\text{cyt c}(Fe^{3+}) + 2H^+ \longrightarrow Q + 2\text{cyt c}(Fe^{2+}) + 4H^+$$

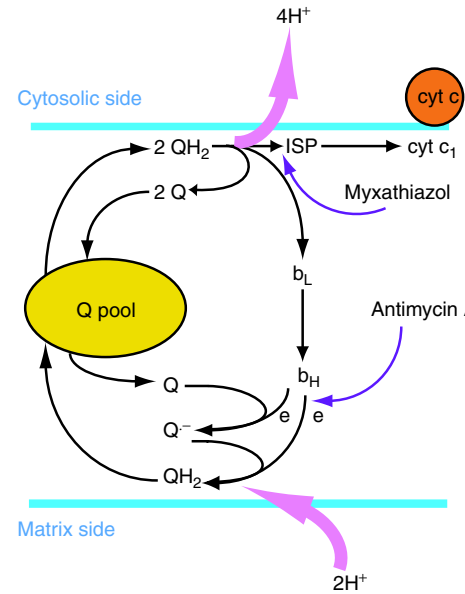


Figure 5.36 The Q cycle and the generation of a vectorial proton gradient

but does not reflect the subtlety of this reaction pathway that allows the bc_1 complex to generate a proton gradient via electron flow through the complex.

The Q cycle (Figure 5.36) became an established concept long before the structure of the cytochrome bc_1 complex was determined. However, this scheme places stringent constraints on the topology of redox co-factors. With the determination of the structure of the cytochrome bc_1 complex came the opportunity to critically assess the Q cycle and to explore the potential mechanism of charge transfer.

One of the first impressions gained from the structure of the bc_1 complex is the large number of extrinsic domains. The Reiske protein is anchored to the membrane by two N terminal helices. The ligands are the side chains of cysteines 139 and 158 and, unlike most iron–sulfur proteins, two histidines (His 141 and 161). The nitrogen containing ligands lead to a higher redox potential for the iron (~100 mV) than is normal in this class of proteins. The iron sulfur centre located in a layer of β sheets is solvent accessible and probably accounts for interaction with quinones and inhibitors such as stigmatellin.

Cytochrome c_1 is also anchored to the bilayer by a single transmembrane helix although in this case it is located at the C terminus. The transmembrane helix is short (20 residues) and the majority of the protein projects ~4 nm into the inter-membrane space. This domain contains a single heme centre and presents a binding surface for cytochrome c containing large numbers of negatively charged residues.

On the matrix side of the complex are two unfortunately named 'core' proteins. These proteins surprisingly project outwards into the aqueous phase and are certainly not located at the core of the complex. Sequence analysis has ascertained that these core subunits show significant homology to a family of heterodimeric Zn proteases and current views of their function envisages a role for these domains in processing matrix proteins by specific peptidase activity. It remains unclear why these proteins remain affiliated to the cytochrome bc_1 complex (Figure 5.37).

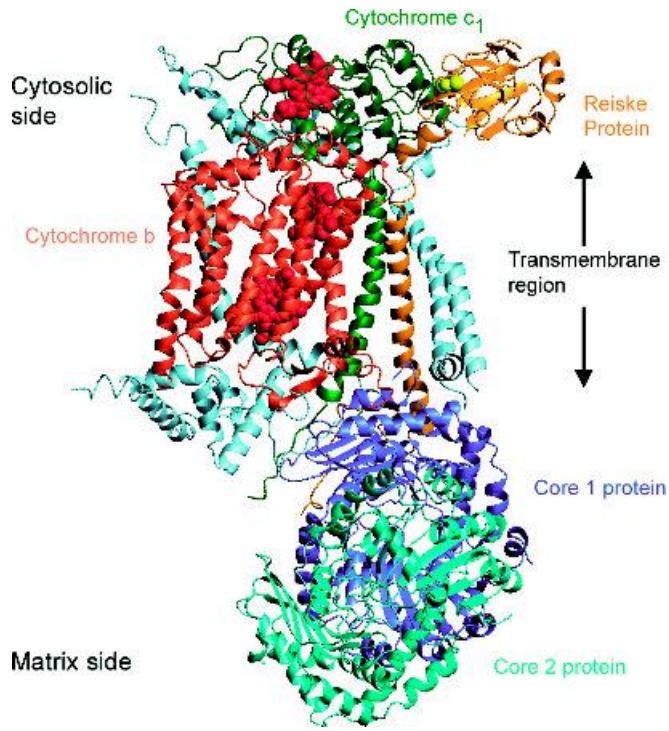


Figure 5.37 The structure of monomeric bovine cytochrome bc_1 complex (PDB:1BE3)

The major transmembrane protein is cytochrome b and the protein has two hemes arranged to facilitate transmembrane electron transfer exactly as predicted by Q cycle schemes. Each heme is ligated by two histidine residues and these side chains arise from just two (out of a total of eight) transmembrane helices found in cytochrome b. The two hemes are separated by 20.7 Å (Fe–Fe centre distances) with the closest approach reflected by an edge–edge distance of 7.9 Å – a distance compatible with rapid inter-heme electron transfer. Cytochrome b is exclusively α helical and lacks secondary structure based on β strands. The helices of cytochrome b delineate the membrane-spanning region and are tilted with respect to the plane of the membrane. The remaining subunits of the bc_1 complex (see Table 5.7) are small with masses below 10 kDa with undefined roles.

Although all of the structures produced from different sources are in broad agreement it was noticed that significant differences existed in the position of the Reiske iron–sulfur protein with respect to the other redox co-factors. This different position reflected crystallization conditions such as the presence of inhibitors and led to a position for the iron–sulfur protein that varied from a location close to the surface of cytochrome b to a site close to cytochrome c_1. In the presence of the inhibitor stigmatellin the iron–sulfur protein shifted towards a position proximal to cytochrome b and suggested intrinsic mobility. The concept of motion suggested that the Reiske protein might act as a 'gate' for electron transfer modulating the flow of electrons through the complex (Figure 5.38).

Table 5.7 The conserved subunits of bovine mitochondrial cytochrome bc_1

Subunit	Identity	No of residues	Mr
1	Core 1	446	53604
2	Core 2	439	46524
3	Cyt b	379	42734
4	Cyt c_1	241	27287
5	Reiske	196	21609
6	Su6	111	13477
7	QP-Cb	82	9720
8	Hinge	78	9125
9	Su10	62	7197
10	Su 11	56	6520

The colour shown in the first column identifies the relevant subunit in the structure shown in Figure 5.37.

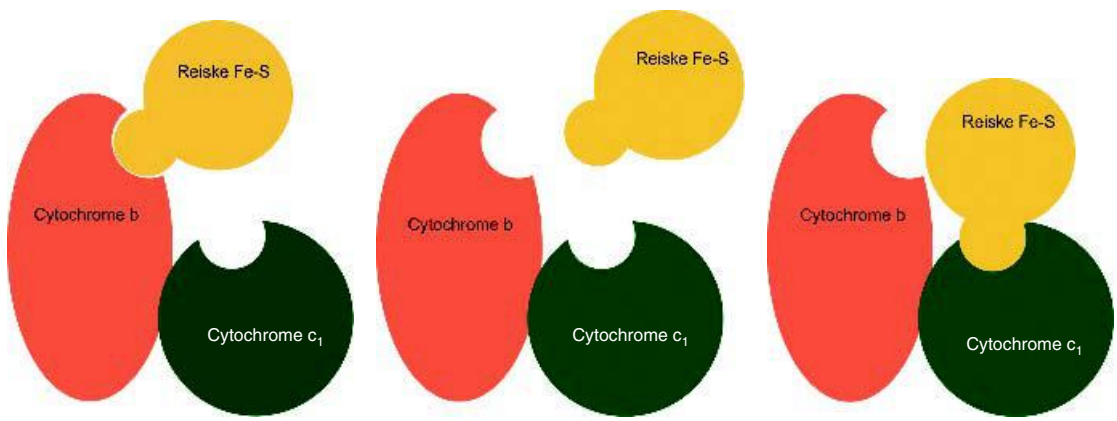


Figure 5.38 The Reiske iron–sulfur protein acts as a redox gate. The schematic diagrams show the Reiske Fe-S centre associated with the cytochrome b, in an intermediate state and associated with cytochrome c_1. Motion at the Reiske Fe-S centre is believed to control electron and proton transfer at the cytochrome bc_1 complex

Complex IV or cytochrome oxidase

Cytochrome c oxidase is the final complex of the respiratory chain catalyzing dioxygen reduction to water. For many years it has occupied biochemists thoughts as they attempted to unravel the secrets of its mechanism. Isolation of cytochrome oxidase has two heme groups (designated a and a_3) together with 2 Cu centres (called Cu_A and Cu_B). Techniques such as electron spin resonance and magnetic circular dichroism alongside conventional spectrophotometric methods showed that heme *a* is a low spin iron centre, whilst heme a_3 was a high spin state compatible with a five-coordinate centre and a sixth ligand provided by molecular oxygen. The catalytic function of cytochrome oxidase is summarized as

$$4\ \text{Cyt c}(Fe^{2+}) + O_2 + 8H^+ \longrightarrow 4\ \text{Cyt c}(Fe^{3+}) + 2H_2O + 4H^+$$

where oxidation of ferrous cytochrome c leads to four-electron reduction of dioxygen and the generation of a proton electrochemical gradient across the inner membrane. Armed with knowledge of the overall reaction spectroscopic studies established an order for electron transfer of

$$\text{Cyt c} \longrightarrow Cu_A \longrightarrow \text{heme a} \longrightarrow \text{heme } a_3\text{–}Cu_B$$

Cu_B and heme a_3 were located close together modulating their respective magnetic properties and this cluster participated directly in the reduction of oxygen with inhibitors such as cyanide binding tightly to heme a_3–Cu_B and abolishing respiration. Transient intermediates are detected in the reduction of oxygen and included the formation of dioxygen adjunct (Fe^{2+}–O_2), the ferryl oxide (Fe^{4+}–O) and the hydroxide (Fe^{3+}–OH). Other techniques established histidine side chains as the ligands to hemes a and a_3–Cu_B on subunit I and Cu_A as a binuclear centre.

The catalytic cycle occurs as discrete steps involving electron transfer from cytochrome c via Cu_A, heme a and finally to the heme a_3–Cu_B site binuclear complex (Figure 5.39). The initial starting point for catalysis involves a fully oxidized complex with ferric (Fe^{3+}) and cupric centres (Cu^{2+}).

Complex IV, with oxidized heme a_3–Cu_B centres, receives two electrons in discrete (1e) steps from the preceding donors. Both centres are reduced and bind two protons from the mitochondrial matrix. In this state oxygen binds to the binuclear centre and is believed to bridge the Fe and Cu centres. Reduction of dioxygen forms a stable peroxy intermediate that remains bound (Fe^{3+}– O^-=O^-–Cu^{2+}) and further single electron (from cytochrome c) and proton transfers lead to the formation of hydroxyl and then ferryl (Fe^{4+}) states. Rearrangement at the catalytic centre during the last two stages completes the cycle forming two molecules of water for every oxygen molecule reduced.

In the absence of a structure for cytochrome oxidase the scheme proposed would need to consider the route and mechanism of proton uptake for the reduction of oxygen from the matrix side of the enzyme to the catalytic site, the mechanism and route of substrate (O_2) binding to the active site and finally the route for product removal from the catalytic site. A further difficulty in understanding the catalytic mechanism of cytochrome oxidase lies with the observation that in addition to the protons used directly in the reduction of water it appears from measurements of H^+/e ratios that up to four further protons are pumped across the membrane to the inter-membrane space. It seems likely that conformational changes facilitating electron and proton transfer lie at the heart of oxidase functional activity.

Cytochrome oxidase from denitrifying bacteria such as *Paracoccus denitrificans* has a simpler subunit composition with only three subunits: all of the redox components are found on the two heaviest polypeptides (subunits I and II). In eukaryotic cytochrome oxidase at least 13 subunits are present with subunits I–III, comparable to those observed in prokaryotic oxidases, encoded by the mitochondrial genome. The crystallization of the complete 13 subunit complex of bovine cytochrome oxidase by Yoshikawa and colleagues in 1995 resolved or suggested answers to several outstanding questions concerning the possible electron transfer route through the complex, a mechanism and pathway of proton pumping and the organization of the different subunits.

The mitochondrially coded subunits are the largest and most hydrophobic subunits and contain four redox centres. Surprisingly the oxidase binds two additional

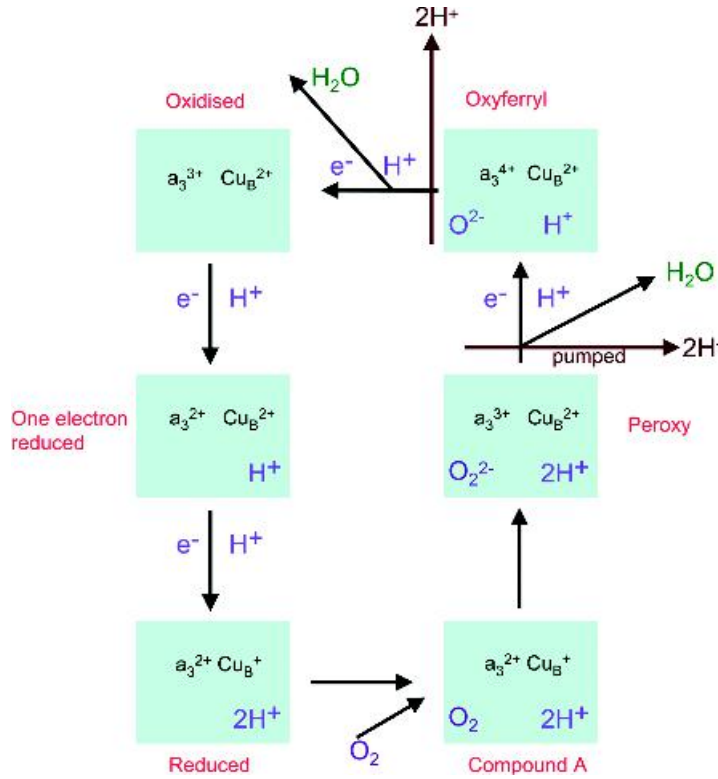


Figure 5.39 A possible scheme for catalytic cycles of cytochrome oxidase. The redox state of the binuclear heme a_3–Cu_B site together with reduction of oxygen, proton binding and protons pumped across membrane is highlighted

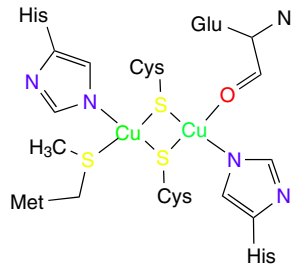


Figure 5.40 The ligands of the Cu_A centre of subunit II of bovine mitochondrial cytochrome oxidase

divalent metal ions namely magnesium and zinc. The zinc is ligated to a nuclear coded subunit (called subunit Vb) whilst the magnesium is held by subunit II. The Cu_A centre as expected was a binuclear cluster of Cu atoms and was contained within subunit II on a cytosolic projecting domain. The Cu_A centre lies ~0.8 nm above the surface of the membrane and is separated from the Fe atoms of hemes a and a_3 by distances of ~1.9 and 2.2 nm respectively. The two Cu atoms are joined in a bridge formed by two sulfur atoms from the side chains of Cys residues in a geometry reminiscent of that found in [2Fe–2S] clusters of iron–sulfur proteins (Figure 5.40). Besides the two sulfur ligands provided by cysteine side chains at positions 196 and 200 of subunit II additional ligands originated from side chains of His161 and His204, Met207 together with the backbone oxygen from Glu198. Glu 198 is also involved in binding the divalent Mg^{2+} ion. In this case the ligand originates from the side chain oxygen atom as opposed to backbone groups with the cation located on a pathway between the Cu_A and heme a_3 centres.

Heme a is coordinated by the side chains of His61 and His378 of subunit I. The iron is a low spin ferric centre lying in the plane of the heme with two imidazole ligands – a geometry entirely consistent with the spin properties. The heme plane is positioned perpendicularly to the membrane surface with the hydroxyethylfarnesyl sidechain of the heme pointing towards the membrane surface of the matrix side in a fully extended conformation. The remaining two redox centres heme a_3 and Cu_B make up the oxygen reduction site and lie very close to heme a. Hemes a and a_3 are within 0.4 nm of each other and this short distance is believed to facilitate rapid electron transfer.

Of some considerable surprise to most observers was the location in the oxidase structure of the fifth ligand to heme a_3. The side chain of histidine 376 provides the fifth ligand to this heme group and lies extremely close to one of the ligands to heme a (Figure 5.41). In the oxidase structure the Cu_B site is close to heme a_3 (~0.47 nm) and has three imidazole ligands (His240, 290 and 291) together with a strong possibility that a fourth unidentified ligand to the metal ion exists. It is generally believed that trigonal coordination is unfavourable and a more likely scenario will involve a tetrahedral geometry around the copper.

A hydrogen-bond network exists between the Cu_A and heme a and includes His204, one of the ligands to Cu_A, a peptide bond between Arg438 and Arg439 and the propionate group of heme a. Although direct electron transfer between these centres has not been observed and they are widely separated the arrangement of bonds appears conducive for facile electron transfer. In addition, a network connects Cu_A and heme a_3 via the magnesium centre and may represent an effective electron transfer path directly between these centres. Direct electron transfer between Cu_A and heme a_3 has never been detected with kinetic measurements, suggesting that electron transfer from Cu_A to heme a is at least two orders of magnitude greater than the rate from Cu_A to heme a_3. From the arrangement of the intervening residues in subunit I the structural reasons underlying this quicker electron transfer from $Cu_A \longrightarrow$ heme a than $Cu_A \longrightarrow$ heme a_3 are not obvious.

Before crystallization the composition and arrangement of subunits within eukaryotic oxidases was unclear. This has changed dramatically with the crystallization of oxidases from bovine heart mitochondria and *P. denitrificans* so that there is now precise information on the arrangement of subunits (Table 5.8). In the crystal structure of bovine cytochrome oxidase the enzyme is a homodimer with each monomer containing 13 polypeptides and six co-factors. The assembly of the 13 subunits into a complex confirmed that these polypeptides are intrinsic constituents of the active enzyme and not co-purified contaminants. The crystal structure confirmed the presence of Zn and Mg atoms as integral components albeit of unknown function.

In the mammalian enzyme the 13 subunits of cytochrome oxidase contain 10 transmembrane proteins that contribute a total of 28 membrane-spanning helices. The three subunits encoded by mitochondrial genes form a core to the monomeric complex with the 10 nuclear-coded subunits surrounding these subunits and seven contributing a single transmembrane helix. The remaining three nuclear coded subunits (subunits Va, Vb, VIb) do not contain transmembrane regions.

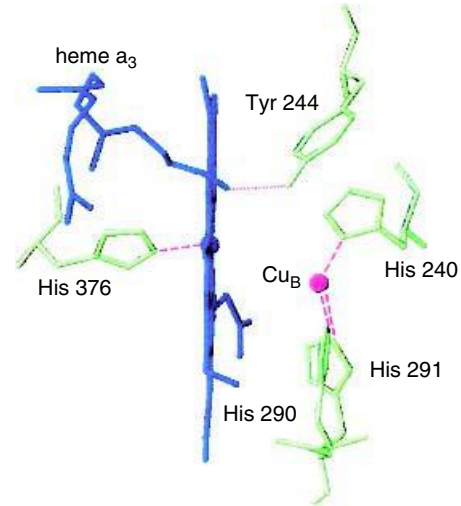


Figure 5.41 The heme a_3–Cu_B redox centre of cytochrome oxidase. The dotted line denotes a hydrogen bond and broken lines denote coordination bonds. Heme a_3 is shown in blue, Cu_B in pink and amino acid residues in green. (Reproduced with permission from Yoshikawa, S. *Curr. Opin. Struct. Biol.* 1997, **7**, 574–579)

Table 5.8 The polypeptide composition of eukaryotic and prokaryotic oxidases

Subunit	Prokaryotic	Eukaryotic
I	554, 620112	514, 57032
II	279, 29667	227, 26021
III	273, 30655	261, 29919
IV	*49, 5370	147, 19577
Va	Absent	109, 12436
Vb		98, 10670
VIa		97, 10800
VIb		85, 10025
VIc		73, 8479
VIIa		80, 9062
VIIb		88, 10026
VIIc		63, 7331
VIII		70, 7639

**Paracoccus denitrificans* has a fourth subunit that is unique to prokaryotes with no obvious homology to that found in eukaryotes. Three of the subunits of the eukaryotic enzyme, VIa, VIIa, and VIII, have two isoforms. Of the two isoforms one is expressed in heart and skeletal muscle and the other is expressed in the remaining tissues. The masses and length of the peptides includes signal sequences for nuclear-coded proteins.

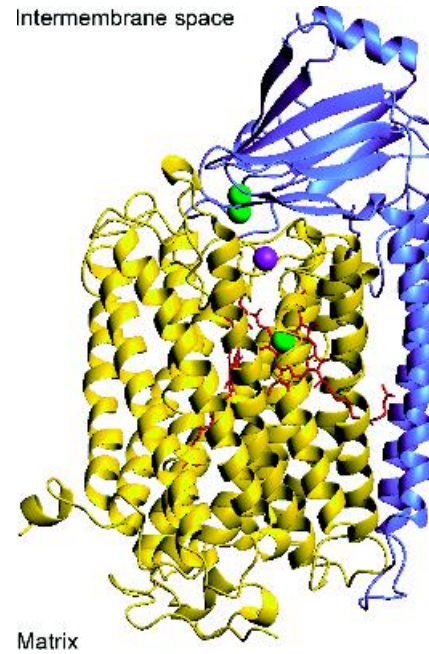


Figure 5.42 The monomeric 'core' of bovine cytochrome oxidase showing subunits I and II. Subunit I is shown in gold whilst subunit II is shown in blue. The limited number of transmembrane helices in subunit II is clear from the diagram in contrast to subunit I. The Cu atoms are shown in green, Mg in purple and the hemes in red (PDB: 1OCC)

In addition to the three metal centre (heme a, heme a_3 and Cu_B) subunit I has 12 transmembrane α helices and this was consistent with structural predictions based on the amino acid sequence. Subunit I is confined to the bilayer region with few loops extending out into the soluble phase. The subunit is the largest found in cytochrome oxidase containing just over 510 residues (M_r ~57 kDa) (Figure 5.42).

Subunits II and III both associate with transmembrane regions of subunit I without forming any direct contact with each other. From the distribution of charged residues on exposed surfaces of subunit II an 'acidic' patch faces the inter-membrane space and defines a binding site for cytochrome c. Unsurprisingly in view of its role as the point of entry of electrons into the oxidase subunit II has just two transmembrane helices interacting with subunit I, but has a much larger polar domain projecting out into the inter-membrane space. This domain consists of a 10-stranded β barrel with some structural similarities to class I copper proteins, although it is a binuclear Cu centre. The magnesium ion is located at the interface between subunits I and II.

Subunit III lacks redox co-factors but remains a significant part of the whole enzyme complex. It has seven α helices and is devoid of large extramembrane projections. The seven helices are arranged to form a large V-shaped crevice that opens towards the cytosolic side where it interacts with subunit VIb (Figure 5.43).

Subunits Va and Vb (Figure 5.44) are found on the matrix side and both are predominantly located in the soluble phase. Subunit Va is extensively helical with subunit Vb possessing both helical- and strand-rich domains and binding a single Zn ion within a series

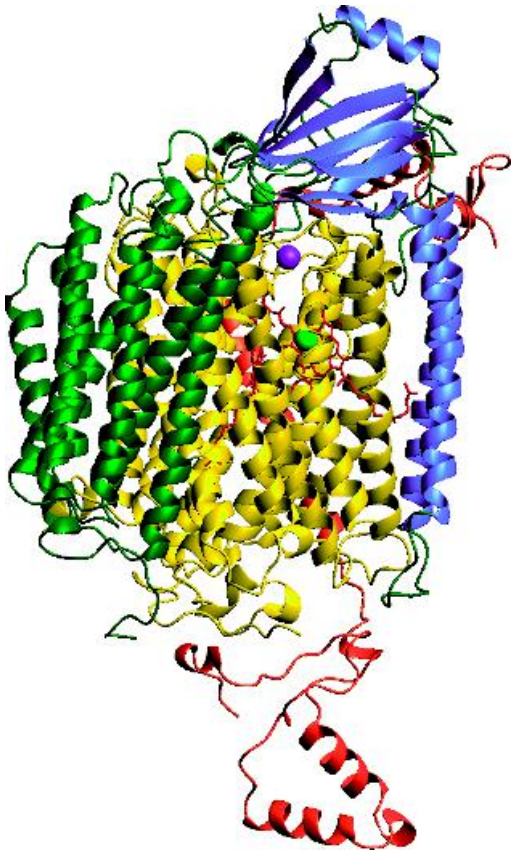


Figure 5.43 To the monomeric core of subunits I and II the remaining mt DNA coded subunit (III) has been added along with subunit IV. Subunit IV (red) has a single helix crossing the membrane with a helix rich domain projecting into the matrix phase. Subunit III (green) does not make contact with subunit II and consists almost exclusively of transmembrane helices

of strands linked by turns that form a small barrel-like structure.

Subunit VIa contributes a single transmembrane helix and is located primarily towards the cytosolic side of the inner mitochondrial membrane. Subunit VIb is located on the same side of the membrane but does not contribute transmembrane helices to the oxidase structure. It is affiliated primarily with subunit III. Subunit VIc contributes a single transmembrane helix and interacts predominantly with subunit II (Figure 5.45). Subunits VIIa, VIIb and VIIc, together with subunit VIII (Figure 5.46), are small subunits with less than 90 residues in each of the polypeptide chains. Their biological function remains unclear. Unlike complex III a far greater proportion of cytochrome oxidase is retained with the boundaries imposed by the lipid bilayer.

The structural model for cytochrome oxidase provides insight into the mechanism of proton pumping, a vital part of the overall catalytic cycle of cytochrome oxidase. Proton pumping results in the vectorial movement of protons across membranes and

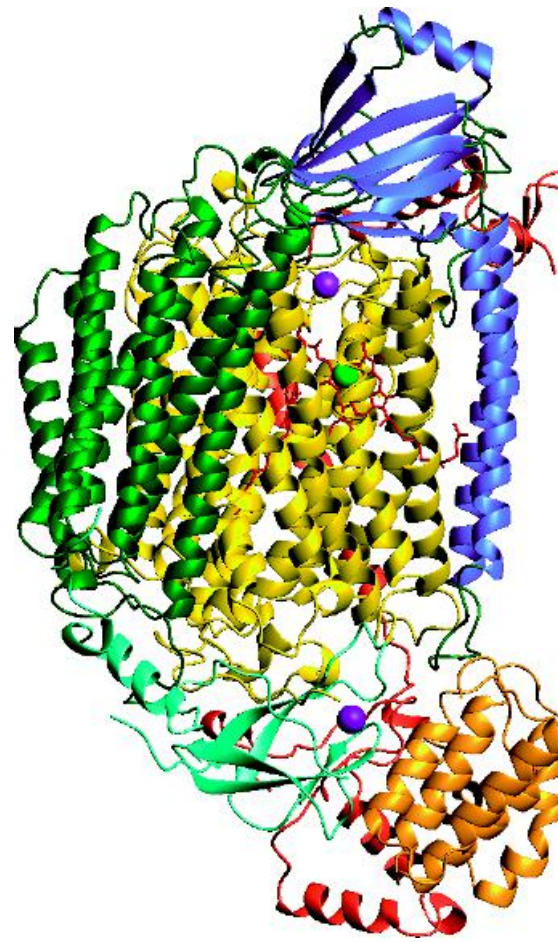


Figure 5.44 The structure of cytochrome oxidase monomer (PDB: 1OCC). Subunits Va (orange) and Vb (aquamarine) have been added. Subunit Vb binds Zn (shown in purple) found in the crystal structure

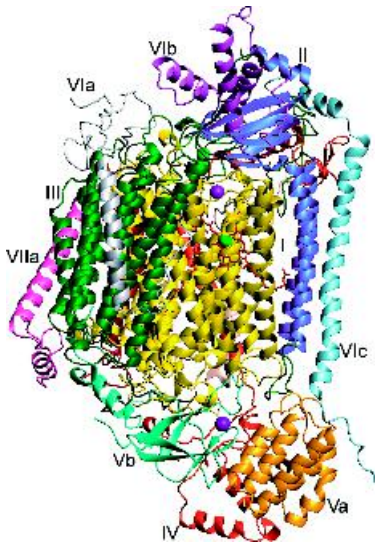


Figure 5.45 The three-dimensional structure of cytochrome oxidase monomer (PDB: 2OCC). The seven smaller subunits VIa (grey) VIb (purple), VIc (light blue), VIIa (pink), VIIb (beige), VIIc (brown) and VIII (black) have been added to the structure shown in Figure 5.44. Subunits VIIb, VIIc, and VIII are largely hidden by the larger subunits (I and III) in this view

in bacteriorhodopsin occurred via strategically placed Asp and Glu residues facilitating proton movement from one side of the membrane to another. This picture, together with that provided by ion channels such as porins, suggests a similar mechanism in cytochrome oxidase. The exact mechanism of proton pumping is hotly debated but it was of considerable interest to observe that the bovine crystal structure contained two possible proton pathways that were shared by simpler bacterial enzymes (Figure 5.47).

The oxygen binding site is located in the hydrophobic region of the protein approximately (35 Å) from the M side. Pathways were identified from hydrogen bonds between side chains, the presence of internal cavities containing water molecules, and structures that could form hydrogen bonds with small possible conformational changes to side chains. Two pathways called the D- and K-channels and named after residues Asp91 and Lys319 of bovine subunit I were identified. (In Paracoccus subunit I these residues are Asp124 and Lys354.) Despite significant advances the proton pumping mechanism of cytochrome oxidase remains largely unresolved. This arises mainly from difficulties associated with resolving proton pumping events during dioxygen

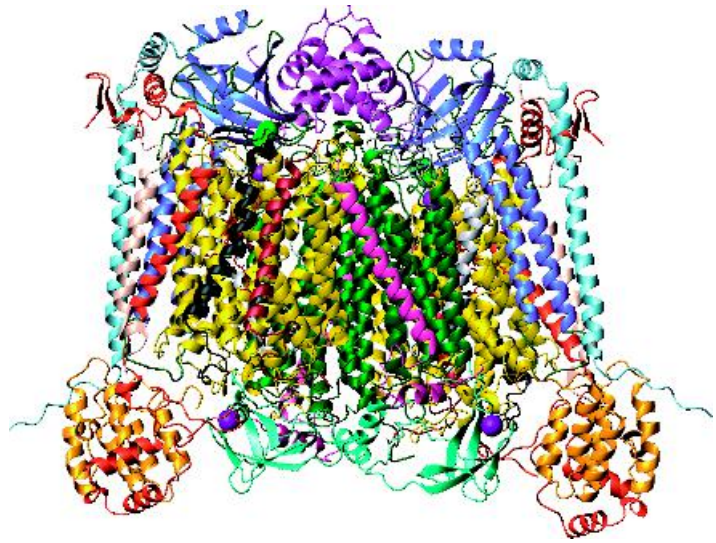


Figure 5.46 The homodimer of bovine cytochrome oxidase using previous colour schemes. This view shows subunits VIIb, VIIc, and VIII in more detail and is from the opposite side to that of the preceding figures

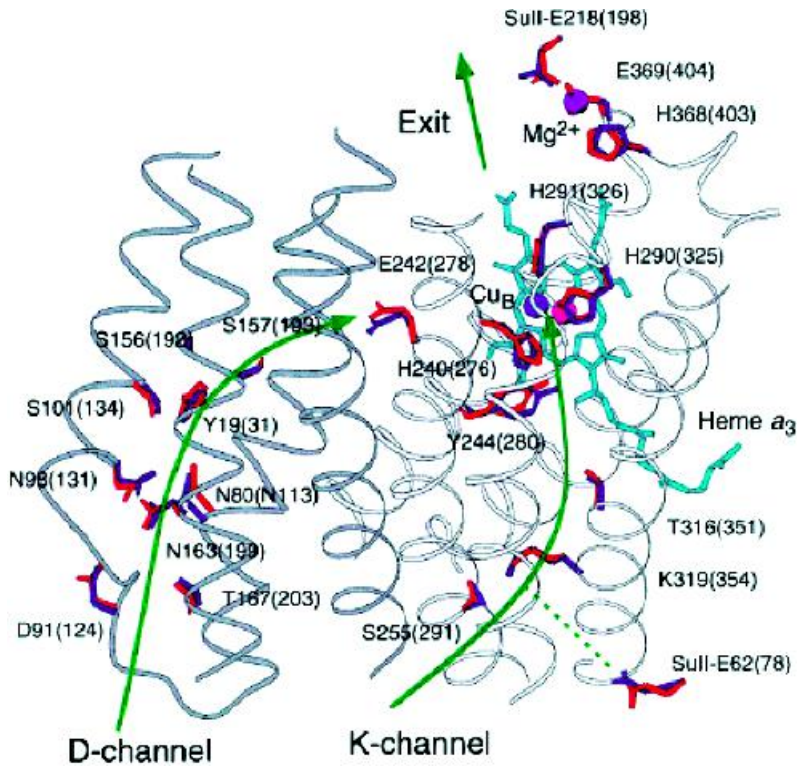


Figure 5.47 The proton pathways in bovine and *Paracoccus denitrificans* cytochrome oxidase. Residues implicated in proton translocation in bovine cytochrome c oxidase are shown in red and in blue for P. denitrificans oxidase. The structures of helices, heme a_3 and metal centres such as the Mg^{2+} binding site are based on PDB: 1ARI (reproduced with permission from Abramson, S *et al. Biochim. Biophys. Acta* 2001, **1544**, 1–9. Elsevier)

reduction and it is currently unclear whether pumping precedes oxygen reduction and what structural rearrangements in the reaction cycle occur. In bacteriorhodopsin intermediate structures were detected and associated with proton movements by combining genetic, spectroscopic and structural studies, and it is likely that such methods will resolve the mechanism of proton transfer in cytochrome oxidase in the near future.

From the study of different membrane proteins a common mode of proton (or cation) transfer involves specifically located aspartate and glutamate residues to shuttle protons across membranes. Both bacteriorhodopsin and cytochrome oxidase provide compelling evidence in their structures for this method of proton conduction across membranes.

The structure of ATP synthetase

It has been estimated that even the most inactive of humans metabolizes kg quantities of ATP in a normal day. This results from repeated phosphorylation reactions with on average each molecule of ADP/ATP being 'turned over' about 1000 times per day. The enzyme responsible for the synthesis of ATP is the F_1F_o ATPase, also called the ATP synthase or synthetase. Dissection of the enzyme complex and genetic studies show the synthase to contain eight different subunits in *E. coli* whilst a greater number of subunits are found in mammalian enzymes. Despite variations in subunit number similarity exists in their ratios and primary sequences with most enzymes having a combined mass between 550 and 650 kDa. With a

central role in energy conservation and a presence in bacterial membranes, the thylakoid membranes of plant and algae chloroplasts and the mitochondrial inner membrane of plants and animals the enzyme is one of the most abundant complexes found in nature.

The link between electron transfer and proton transfer was the basis of the chemiosmotic theory proposed in its most accessible form by Peter Mitchell at the beginning of the 1960s. One strong line of evidence linking electron and proton transfer was the demonstration that addition of chemicals known as 'uncouplers' dissipates proton gradients and halts ATP synthesis. Many uncouplers were weak, membrane permeable, acids and a major objective involved elucidating the link between transmembrane proton gradients and the synthesis of ATP from ADP and inorganic phosphate.

Fractionation and reconstitution studies showed that ATP synthases contained two distinct functional units each with different biological and chemical properties. In electron micrographs this arrangement was characterized by the appearance of the enzyme as a large globular projection extending on stalks from the surface of cristae into the mitochondrial matrix (Figure 5.48). A hydrophobic, membrane-binding, domain called F_o interacted with a larger globular domain peripherally associated with the membrane. F_o stands for 'factor oligomycin' referring to the binding of this antibiotic to the hydrophobic portion whilst F_1 is simply "factor one". F_o is pronounced 'ef-oh'. The peripheral protein, F_1, was removed from membranes by washing with solutions of low ionic strength or those containing chelating agents. This offered a convenient method of purification. Although not strictly a membrane protein it is described in this section because its physiological function relies intimately on its association with the F_o portion of ATP synthetase, a typical membrane protein. In an isolated state the F_1 unit hydrolysed ATP but could not catalyse ATP synthesis. For this reason the F_1 domain is often described as an ATPase reflecting this discovery. A picture of ATP synthesis evolved where F_o formed a 'channel' that allowed a flux of protons and their use by F_1 in ATP synthesis.

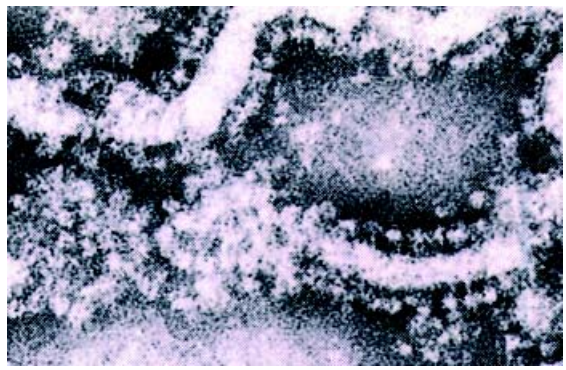


Figure 5.48 An electron micrograph of cristae from isolated mitochondria showing the F_1 units projecting like 'lollipops' from the membrane and into the matrix (reproduced with permission from Voet, D., Voet, J.G & Pratt, C.W. *Fundamentals of Biochemistry*. John Wiley & Sons, Chichester, 1999)

In *E. coli* F_o is extremely hydrophobic and consists of three subunits designated a, b and c (Table 5.9). In *E. coli* a ratio of a_1, b_2, c_{9-14} exists with the a and c subunits making contact and appearing to form a proton channel. Within both subunits are residues essential for proton translocation and amongst those to have been identified from genetic and chemical modification studies are Arg210, His245, Glu196, and Glu219 of the *a* subunit and Asp61 of the *c* subunit. Since all of these residues possess side chains that bind protons the concept of a proton channel formed by these groups is not unrealistic.

Table 5.9 The composition of ATP synthetases

Subunits in F_1 region	Eukaryotic (bovine) $\alpha_3\beta_3\gamma\delta\varepsilon$
α	509 residues, M_r~55,164
β	480, 51,595
γ	272, 30,141
δ	190, 20,967
ε	146, 15,652
Subunits in F_o region	Prokaryotic (*E. coli*) a,b_2,c_{9-14}
a	271, 30,285
b	156, 17,202
c	79, 8,264

The much larger F_1 unit contains five different subunits ($\alpha-\varepsilon$) occurring with a stoichiometry $\alpha_3\beta_3\gamma_1\delta_1\varepsilon_1$. The total mass of the F_1F_o ATPase from bovine mitochondria is ~450 kDa with the F_1 unit having a mass of ~370 kDa and arranged as a spheroid of dimensions 8×10 nm wide supported on a stem or stalk 3 nm in length. Chemical modification studies identified catalytic sites on the β subunits with affinity labelling and mutagenesis studies highlighting the importance of Lys155 and Thr156 in the *E. coli* β subunit sequence ^{149}Gly-Gly-Ala-Gly-Val-Gly-Lys-Thr-Ala^{157}.

A scheme for ATP synthesis envisaged three discrete stages:

1. Translocation of protons by the F_o
2. Catalysis of ATP synthesis *via* the formation of a phosphoanhydride bond between ADP and Pi by F_1
3. A coupling between synthesis of ATP and the controlled dissipation of the proton gradient *via* the coordinated action and interaction of F_1 and F_o.

The mechanism of ATP synthesis

A mechanism for ATP synthesis was proposed by Paul Boyer before structural data was available and has like all good theories stood the test of time and proved to be compatible with structural models derived subsequently for F_1/F_o ATPases. Boyer envisaged the F_1 ATPase as three connected, interacting, sites each binding ADP and inorganic phosphate. The model embodied known profiles for the kinetics of catalysed reactions and the binding of ATP and ADP by isolated F_1. Each catalytic site was assumed to exist in a different conformational state although inhibition of ATP synthesis by modification of residues in the β subunit suggested that these sites interacted cooperatively.

Rate-limiting steps were substrate binding or product release (i.e ATP release or ADP binding) but not the formation of ATP. Careful examination of the ATP hydrolysis reaction suggested three binding sites for ATP each exhibiting a different K_m. When ATP is added in sub-stoichiometric amounts so that only one of the three catalytic sites is occupied substrate binding is very tight (low K_d) and ATP hydrolysis occurs slowly. Addition of excess ATP leads to binding at all three sites but with lowered affinity at the second and third sites. The K_d for binding ATP at the first site is <1 nM whereas values for sites 2 and 3 are ~1 μM and ~30 μM. Upon occupancy of the third site, the rate of overall ATP hydrolysis increases by a factor of $\sim 10^4 - 10^5$.

As long as the results from hydrolysis of ATP can be related directly to that of the synthetic reaction and isolated F_1 behaves comparably to the complete F_1F_o complex then these results point towards key features of the catalytic mechanism. Isolated F_1 exhibits negative cooperativity with respect to substrate binding where ATP binding at the site with the lowest K_m makes further binding of ATP molecules difficult. This occurs because the K_m for other sites is effectively increased. In contrast the enzyme displays positive cooperativity with respect to catalytic activity with the rate increasing massively in a fully bound state. To explain these unusual properties, Boyer proposed the 'binding change' mechanism (Figure 5.49).

A central feature of this hypothesis are three catalytic sites formed by the three α/β subunit pairs each with a different conformation at any one time. One site is open and ready for substrate binding (ATP or ADP + P_i) whilst the second and third sites are partly open and closed, respectively. Substrate binding results in the closure of the open site and produces a cooperative conformational change in the other two sites; the closed one becomes partly open and the partly open one becomes fully open. In this manner the sites alternate between conformational states. Obvious questions from this model arise. How are conformational changes required in the binding change mechanism propagated? How are these events linked to proton translocation and the F_o part of the ATPase?

From a consideration of the organization of F_o G.B. Cox independently suggested a rotary motor involved the c subunit ring turning relative to the a and b subunits. This idea was adopted and incorporated as a major refinement of the binding change mechanism with the rotary switch causing a sequential change in binding. At the time (1984) the concept of enzymes using rotary motion to catalyse reactions was unusual. The cooperative kinetics of the enzyme (negative for ATP binding and positive for catalysis) supported the

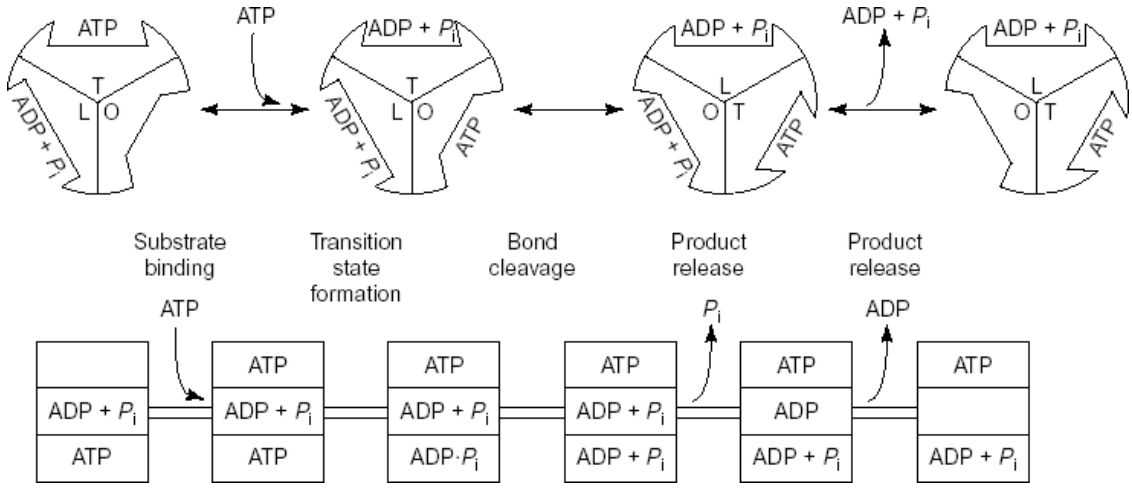


Figure 5.49 The binding change hypothesis and catalytic site activity at the β subunit. Each catalytic site cycles through three states: T, L and O. ATP binds to the O (open and empty) site converting it into a T (tight and ATP-occupied) site. After bond cleavage at this site the T site is converted into the L (loose and ADP-occupied) site where the products escape to recover the O state. The concerted switching of states in each of the sites results in the hydrolysis of one ATP molecule. Sub-steps in the hydrolysis of one ATP molecule based on kinetic and inhibitor studies are shown for a three-site mechanism. Thus, ATP binding in the open site (top) leads to transition state formation and then bond cleavage in a closed site (bottom), followed by release from a partly open site as it opens fully and releases ADP (middle) (reproduced with permission from *Trends Biochem. Sci* 2002, **27**, 154–160. Elsevier)

binding-change mechanism. However, the idea of a rotary switch was so novel that there were few serious attempts to test it until over a decade later when in 1995 the first structures for bovine F_1 ATPase produced by John Walker and co-workers appeared to support Boyer's model.

The structure of the F_1 ATPase

As a peripheral membrane protein the water-soluble F_1 unit is easily purified and the structure showed an alternate arrangement of three α and three β subunits based around a central stalk formed by two α helices running in opposite directions derived from the γ subunit. These two helices form a coiled-coil at the centre of F_1 and the remaining part of the γ subunit protrudes from the $(\alpha\beta)_3$ assembly and interacts with polar regions of the c subunit of F_o (Figure 5.50).

The α and β subunits exhibit sequence identity of ~20 percent and show similar overall conformations

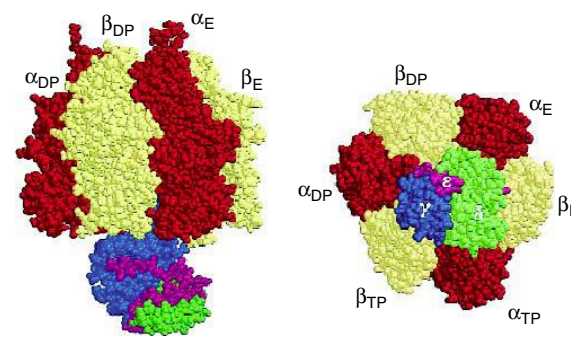


Figure 5.50 A space filling model of bovine F1-ATPase from a side view (left) and from the membrane (right) showing α subunits (red) and β subunits (yellow) and the attached central stalk containing the γ, δ and ε subunits (blue, green and magenta, respectively). The β_{TP}, β_{DP} and β_E subunits had AMP-PNP, ADP and no nucleotide bound, respectively. The α subunits are named according to the catalytic interface to which they contribute

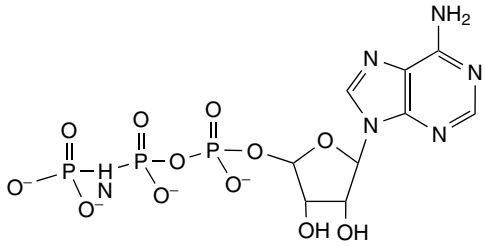


Figure 5.51 The structure of AMP-PNP – a non-hydrolysable derivative of ATP

arranged around the central stalk made up of the γ, δ and ε subunits. F_1 was co-crystallized with ADP and AMP-PNP (5′-adenylyl-imidophosphate, a non-hydrolysed analogue of ATP; Figure 5.51). Only five of the six nucleotide binding sites were occupied by substrate. Each of the α subunits bound AMP-PNP and exhibited identical conformation.

In contrast one β subunit bound AMP-PNP (β_{TP}), another bound ADP (β_{DP}) and the third subunit remained empty (β_E). The structure provided compelling evidence for three different modes of nucleotide binding as expected from the Boyer model. Although both α and β subunits have nucleotide binding sites only those on the β subunits are catalytically important.

Each α and β subunit contains three distinct domains (Figure 5.52). A β barrel located at the N-terminal is followed by a central nucleotide-binding domain containing a Rossmann fold. The β barrels form a crown that sits on top of the central nucleotide-binding domain. At the C terminal is a helix rich domain. The conformation of the β subunit changes with nucleotide-binding forming states associated with high, medium and low nucleotide affinities. These three states are most probably related to release of product (ATP), binding of substrates (ADP and inorganic phosphate), and formation of ATP, respectively. In an isolated state the structure of the β subunit has the open form with nucleotide binding causing transition to closed forms. On the β subunit the principal residues involved in ADP binding have been determined by examining the crystal structure of this subunit in the presence of ADP-fluoroaluminate (ADP-AlF_3) (Figure 5.53). Positively charged side chains of arginine and lysine have a major role in binding ADP-fluoroaluminate whilst the side

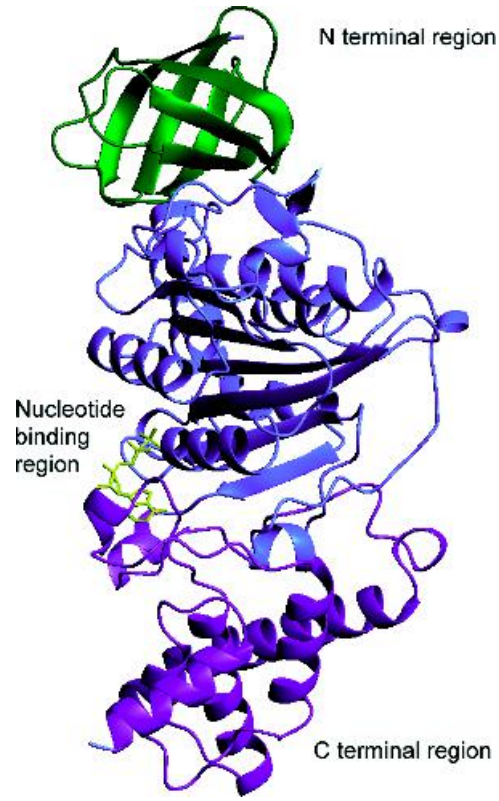


Figure 5.52 An individual β subunit showing the three distinct domains together with the bound nucleotide – in this case ADP. The ADP is shown in yellow, the barrel structure in dark green, the nucleotide binding region in blue and the helix rich region in violet (PDB:1BMF)

chain of Glu188 binds a water molecule and the amide of Gly159 binds the second phosphate group of ADP. Arginine 373 is part of the α subunit and contributes to the binding site for ADP. ADP-fluoroaluminate mimics the transition state complex with fluorine atoms occupying similar conformations to the extra phosphate group of ATP but acting as a potent inhibitor of ATP synthesis. Significantly, the structure adopted by the β subunit in the presence of ADP-AlF_3 is very similar to the catalytically active conformation for this subunit (β_{TP}). These results further suggest that once the substrate is bound only small conformational changes are required to promote synthesis of ATP.

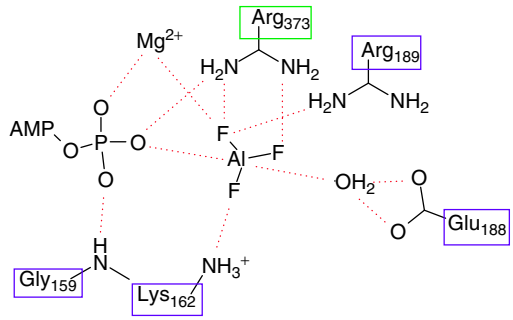


Figure 5.53 Schematic representation of the nucleotide-binding site of the β subunit of the AlF_3–F_1 complex, showing the coordination of the aluminofluoride group. Possible hydrogen-bond interactions are shown by dotted lines

The stoichiometry of the bovine mitochondrial F_1 ATPase means that the single γ, δ and ε subunits lack symmetry, and in the case of the γ subunit the intrinsic asymmetry causes a different interaction with each of the three catalytic β subunits of F_1. This interaction endows each β subunit with different nucleotide affinity and arises from rotation of the central stalk. Rotation of the γ subunit driven by proton flux through the F_o region drives ATP synthesis by cycling through structural states corresponding to the tight, open and loose conformations.

Initial structures of F_1 ATPase showed disorder for the protruding end of the γ subunit from the $(\alpha\beta)_3$ hexamer. Examining enzyme inhibited with dicyclohexylcarbodiimide (DCCD), a known stoichiometric inhibitor of ATP synthesis that reacts specifically with Glu 199 of the β subunit, resulted in improved order and confirmed the orientation of the central pair of helices of the γ subunit, the structures of the δ and ε subunits and their interaction with the F_o region at the end of the central stalk. The protruding part of the central stalk is composed of approximately half of the γ subunit and the entire δ and ε subunits. The overall length of the central stalk from the C-terminus of the γ subunit to the foot of the protruding region is ~11.4 nm. It extends ~4.7 nm from the $(\alpha\beta)_3$ domain thereby accounting for 'lollipop' appearance in electron micrographs (Figure 5.48). The central stalk is the key rotary element in the catalytic mechanism. The γ subunit interacts directly with the ring of c subunits located in the membrane of F_o and in the absence of covalent bonding between c subunits and the stalk region this interaction must be sufficiently robust to permit rotation. In the γ subunit, three carboxyl groups (γAsp 194, γAsp 195 and γAsp 197) are exposed on the lower face of the foot suggesting that they may interact with basic residues found in loop regions of the c-ring.

It is established that all species contain a, b and c subunits within the F_o part of the enzyme although there is currently no clear agreement on the total number of subunits. In yeast ATP synthetase at least thirteen subunits are found in the purified enzyme. As well as the familiar α–ε subunits of F_1 and the a, b and c subunits of F_o additional subunits called f and d are found in F_o whilst OSCP, ATP8 and h occur in F_1. Using the yeast enzyme the F_1 region was co-crystallized with a ring of 10 c subunits and although the loss of c subunits during preparation cannot be eliminated definitively the lack of symmetry relations between the hexameric collection of α/β subunits and the 10 c subunits was surprising. The structure of the c subunit monomer was established independently by NMR spectroscopy (Figure 5.55) and showed two helical regions that were assumed to be transmembrane segments linked by an extramembranous loop region. These loop regions make contact with γ and δ subunits and represent sites of interaction (Figure 5.56). Only the a and c subunits are essential for proton translocation (Figure 5.57) with Arg210, His245, Glu196, and Glu219 (a subunit) and Asp61 of the c subunit implicated as essential residues.

Further refinement of the ATP synthetase structure has shown a second, peripheral, stalk connecting the F_1 and F_o domains in the enzyme (Figure 5.58). The second stalk visible using cryoelectron microscopy of detergent solubilized F_1F_o from *E. coli*, chloroplast and mitochondrial ATPases is located around the outside of the globular F_1 domain and may act as a stator to counter the tendency of the $(\alpha\beta)_3$ domain to follow the rotation of the central stalk.

With the structural model derived from X-ray crystallography in place for the F_1 subunit a number of elegant experiments established that rotation *is* the link between proton flux and ATP synthesis. The

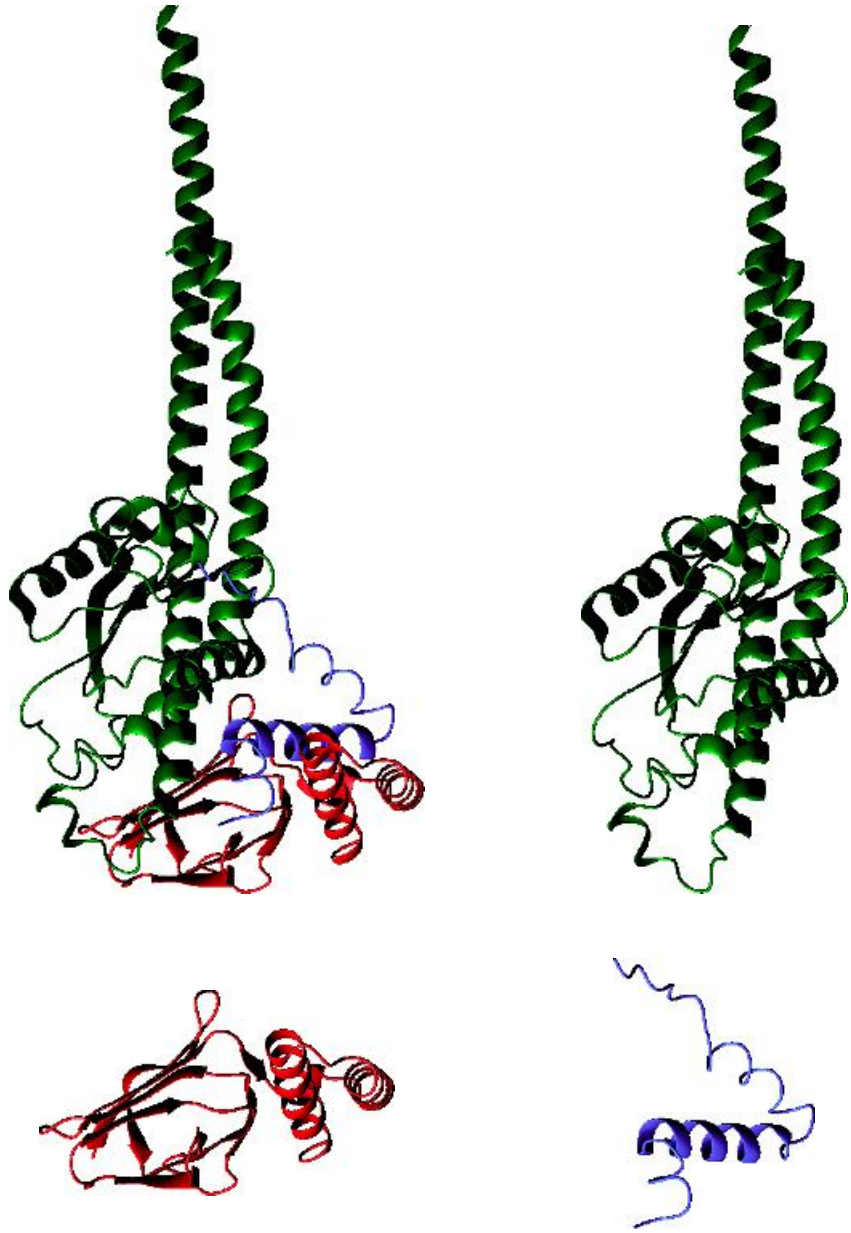


Figure 5.54 The structures of the central stalk assembly with the individual subunits shown in isolation alongside. The γ subunit contains the long α helical coiled coil region that extends through the $(\alpha\beta)_3$ assembly (top right). The bovine ε subunit is very small ($\sim$50 residues in length) with one major helix together with a poorly organized helix (bottom right). The δ subunit contains an extensive β sheet formed from seven β strands together with a distinct helix turn helix region (bottom left)

idea of rotary switches is now no longer heretical but is accepted as another solution by proteins to the mechanistic problem of harnessing energy from proton gradients.

Covalently linking the γ subunit of F_1 to the C terminal domain of the β subunit blocked activity whilst cross-linking to the α subunit was without effect. Complementary studies introducing cysteine residues into the structure of β and γ subunits with the formation of thiol cross links inhibited activity but reducing these bonds followed by further cross-linking resulted in the modification of a different β subunit. This could only occur if rotation occurred within ATPases. Finally, covalently joining the γ, ε and c subunit ring did not inhibit ATP synthesis and suggested that the subunits act in unison as a rotating unit. Direct observation of rotation in the synthetase was obtained using actin filaments attached to the γ subunit. By following the fluorescence of covalently linked fluorophores in a confocal microscope it was shown that ATP hydrolysis led to a 360° rotation of the filament in three

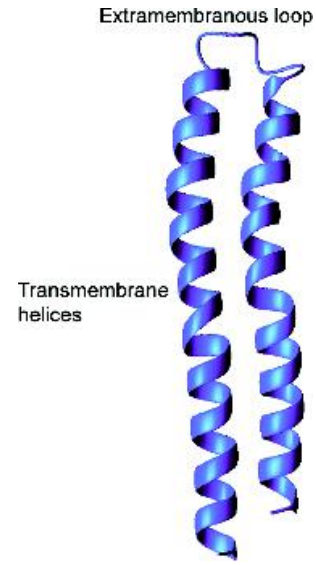


Figure 5.55 The structure of the c ring monomer determined using NMR spectroscopy is a simple HTH structure (PDB:1A91)

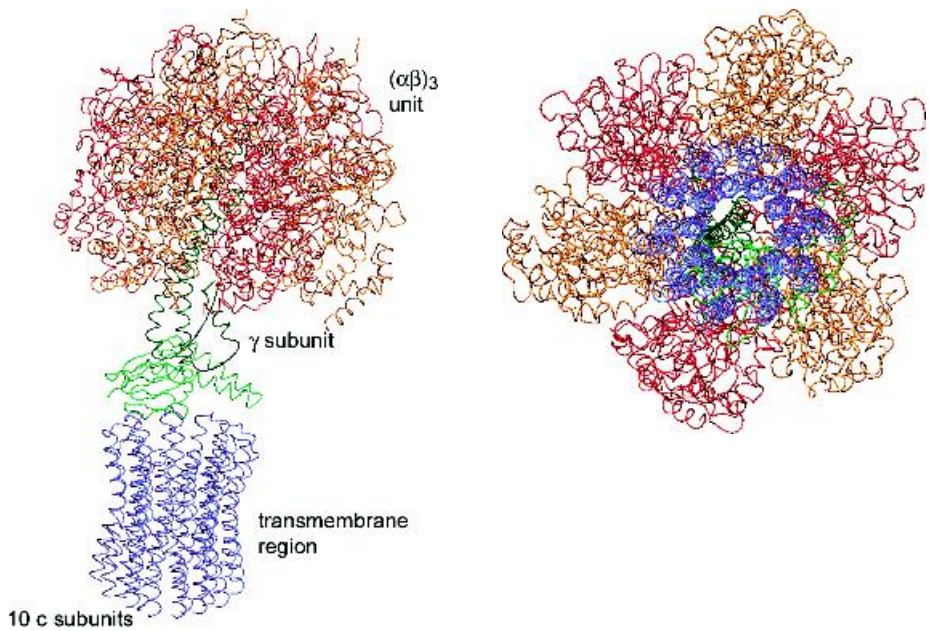


Figure 5.56 The arrangement of the c ring with the F_1 catalytic domain in yeast ATP synthetase. The c ring of 10 monomers is shown in blue, the $(\alpha\beta)_3$ unit is shown in red and orange and the γ subunit in dark green. A view of the F_1 F_o ATPase from the inter-membrane space looking through the membrane region is shown on the right

uni-directional steps of 120°. The complex showed rotation rates of ~130 Hz entirely consistent with a maximal rate of hydrolysis of ATP of ~300 s^{-1}.

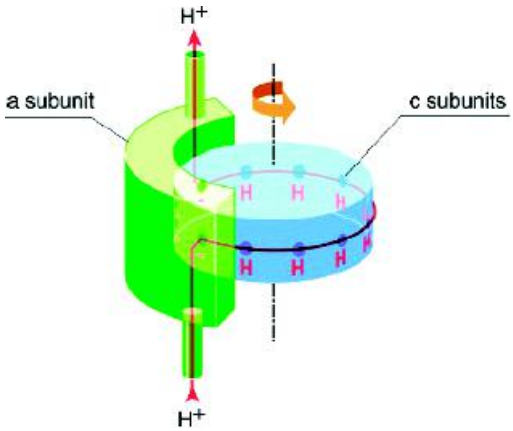


Figure 5.57 Hypothetical model for the generation of rotation by proton transfer through the F_o domain of ATP synthase. The central cylinder (blue) consists of c subunits; the external part (green) corresponds to a single a subunit. The red line indicates the proton path

ATP, the universal carrier of cell energy, is manufactured from ADP and phosphate by the enzyme ATP synthase (or F_1F_o ATPase) using the free energy of an electrochemical gradient of protons. This proton gradient is generated by the combined efforts of numerous respiratory or photosynthetic protein complexes whose catalytic properties include rapid electron transfer coupled to movement of protons across membranes.

ATPase family

The F_1F_o ATPases comprise a huge family of enzymes with members found in the bacterial cytoplasmic membrane, the inner membrane of mitochondria and the thylakoid membrane of chloroplasts. These enzymes are usually classified as proton-transporting F type ATPases to distinguish them from enzymes that hydrolyse ATP as part of their normal catalytic function. Further classes can be recognized in these ATPases through the use of alternative ions to the proton such as Na^+ or Li^+ ions. Minor changes to the a and c subunit sequences alter the coupling ion specificity although the primary mechanism remains very similar. In all cases the ATPases are based around a large globular catalytic domain that projects from the membrane surface but

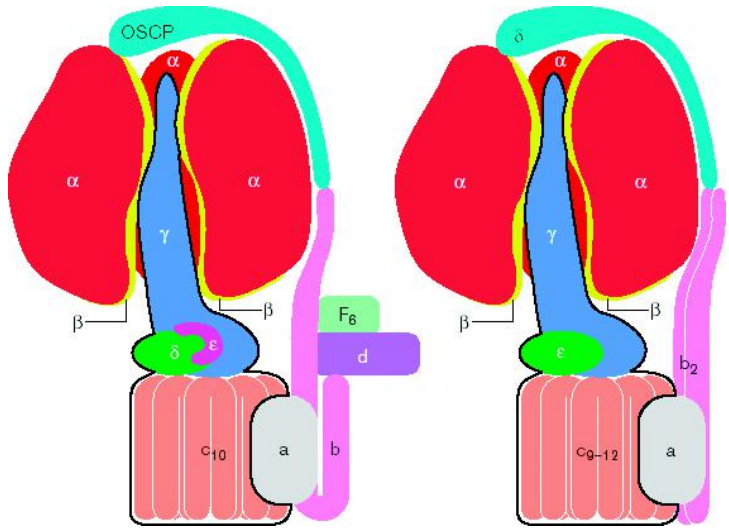


Figure 5.58 Organization of ATP synthases. The bovine F_1 – ATPase (left) and the ATP synthase of *E. coli* (right). (Reproduced with permission from Stock, D. *et al. Curr. Opin. Struct. Biol.* 2000, **10**, 672–679. Elsevier)

interacts with a hydrophobic complex that controls ion translocation via rotary motion.

In all cells there are membrane proteins whose function relies on the hydrolysis of ATP and the coupling of this energy to the movement of ions such as Na^+, K^+ or Ca^{2+} across membranes. Transport proteins represent a major category of membrane proteins and a number of different mechanisms exist for transport. Transport across membranes is divided into mediated and non-mediated processes. Non-mediated transport involves the diffusion of ions or small molecules down a concentration gradient. Transport is closely correlated with solute solubility in the lipid phase with steroids, anaesthetics, narcotics and oxygen diffusing down a concentration gradient and across membranes. Mediated transport is divided into two categories. Passive-mediated transport also called facilitated diffusion relies on the transfer of a specific molecule from high concentration to low concentration. Transport proteins involved in this mechanism of transport have already been described in this chapter and are exemplified by the β barrel porin proteins. Unfortunately these proteins cannot transport substrates against concentration gradients; this requires active transport and the hydrolysis of ATP.

Na^+–K^+ ATPase is one member of a class of membrane bound enzymes that transfer cations with ATP hydrolysis providing the driving force. The Na^+–K^+ ATPases, a highly conserved group of integral proteins, are widely expressed in all eukaryotic cells and are very active with a measure of their importance seen in the estimate that ~25 percent of all cytoplasmic ATP is hydrolysed by sodium pumps in resting human cells. In nerve cells, a much higher percentage of ATP is consumed (>70 percent) to fuel sodium pumps critical to neuronal function.

The Na^+–K^+ ATPase is a transmembrane protein containing two types of subunit. An α subunit of ~110 kDa contains the ion-binding site and is responsible for catalytic function whilst a β subunit, glycosylated on its exterior surface, is of less clear function. Sequence analysis suggests that the α subunit contains eight transmembrane helices together with two cytoplasmic domains. In contrast the β subunit (M_r~55 kDa) has a single transmembrane helix and a much larger domain projecting out into the extracellular environment where glycosylation accounts for ~20 kDa of the total mass of the protein. The β subunit is required for activity and it is thought to play a role in membrane localization and activation of the α subunit. *In vivo* the Na^+–K^+ ATPase functions as a tetramer of composition $\alpha_2\beta_2$.

The Na^+–K^+ ATPase pumps sodium out of the cell and imports potassium with the hydrolysis of ATP at the intracellular surface of the complex. The overall stoichiometry of the reaction results in the movement of 3 Na^+ ions outward and 2 K^+ ions inwards and is represented by the equation

$$3Na^+_{(in)} + 2K^+_{(out)} + ATP + H_2O \leftrightharpoons 3Na^+_{(out)} + 2K^+_{(in)} + ADP + P_I$$

The ATPase is described as an antiport where there is the simultaneous movement of two different molecules in opposite directions, and leads to the net movement of positive charge to the outside of the cell. A uniport involves the transport of a single molecule whilst a symport simultaneously transports two different molecules in the same direction (Figure 5.59).

In the presence of Na^+ ions ATP phosphorylates a specific aspartate residue on the α subunit triggering conformational change. In the presence of high levels of K^+ the dephosphorylation reaction is enhanced and suggests two distinct conformational states for the Na^+–K^+ ATPase. These states, designated E_1 and E_2, differing in conformation show altered affinity and specificity. On the inside of the cell the Na^+–K^+ ATPase (in the E_1 state) binds three sodium ions at specific sites followed by ATP association to yield an E_1–ATP–$3Na^+$ complex. ATP hydrolysis occurs to form an aspartyl phosphate intermediate $E_1 \sim P$–$3Na^+$ that undergoes a conformational change to form the E_2 state ($E_2 \sim P$–$3Na^+$) and the release of bound Na outside the cell. The presence of a 'high energy' phosphoryl intermediate (denoted by ~P) has led to this class of membrane proteins being called P type ATPases. As an E_2–P state the enzyme binds two K^+ ions from the outside to form the E_2–P–$2K^+$ complex. Hydrolysis of the phosphate group to reform the original aspartate side chain causes conformation change with the result that $2K^+$ ions are released on the inside of the cell. The ATPase ion pump, upon release of the potassium ions, forms the E_1 state capable

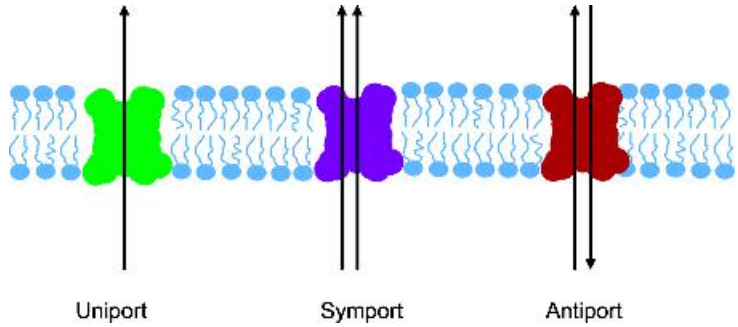


Figure 5.59 Diagrammatic representation of uniport, symport and antiport systems used in processes of active transport across membranes

of binding three Na^+ ions and restarting the cycle again.

Study of the function of Na^+–K^+ ATPases has been aided by the use of cardiac glycosides, natural products that increase the intensity of muscular contraction in impaired heart muscle. Abnormalities have been identified in the function of Na^+–K^+ ATPases and are involved in several pathologic states including heart disease and hypertension. Several types of heart failure are associated with significant reductions in the myocardial concentration of Na^+–K^+ ATPase or impaired activity. Similarly excessive renal reabsorption of sodium due to oversecretion of the hormone aldosterone is associated with forms of hypertension. Oubain is a naturally occurring steroid and is still prescribed as a cardiac drug. Its action is to block the efflux of Na^+ ions leading to an increase in their intracellular concentration (Figure 5.60). The resulting increase in Na^+ concentration causes a stimulation of secondary active transport systems in cardiac muscle such as the Na^+–Ca^{2+} antiport system. This pump removes Na^+ ions from the cell and transports Ca^{2+} to the inside. Transient increases in cytosolic Ca^{2+} ion concentration often triggers many intracellular responses one of which is muscular contraction. The Ca^{2+} ions stimulate cardiac muscle contraction producing a larger than normal response to aid cardiovascular output. Although of major

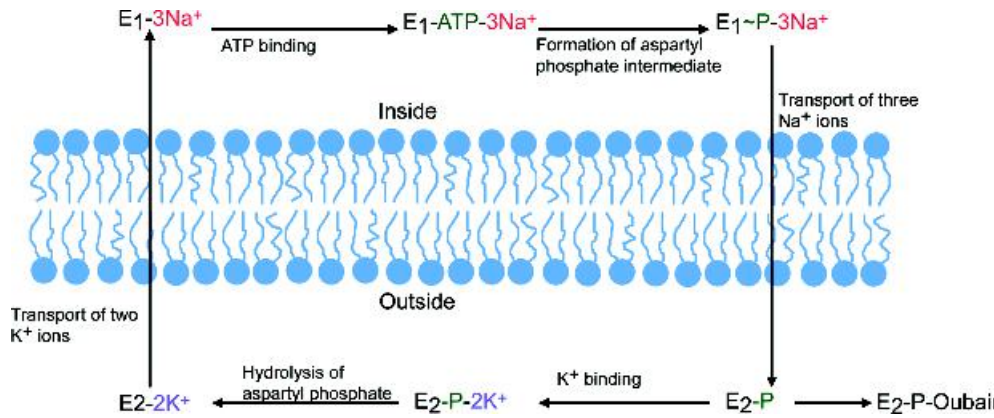


Figure 5.60 The active transport of Na^+ and K^+ ions by the Na^+–K^+ ATPase. Oubain (pronounced 'wabane') is a cardiac glycoside promoting increased intracellular Na^+ concentration

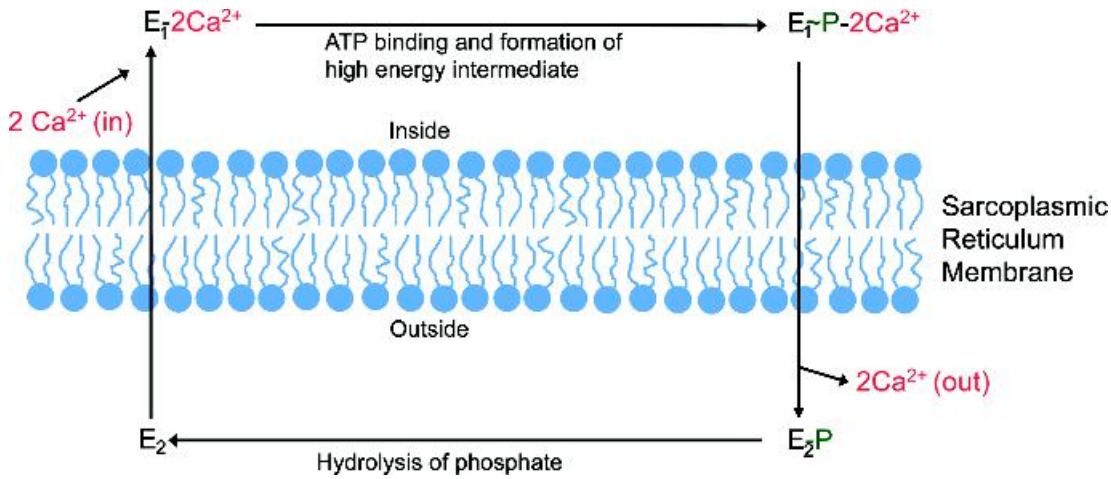


Figure 5.61 Scheme for active transport of Ca^{2+} by SR Ca ATPase. The inside refers to the cytosol and the outside represents the lumen of the sarcoplasmic reticulum of skeletal muscle cells. Similar Ca pumps exist in the plasma membranes and ER

importance to the physiology of all cells detailed studies of the Na^+–K^+ ATPase protein and ultimately the development of improved drugs is hampered by the absence of a well-defined structure.

Fortunately, another prominent member of the P type ATPases, the Ca^{2+}–ATPase from the sarcoplasmic reticulum (SR) of muscle cells, has been crystallized to reveal some aspects of the organization of transport proteins. SR Ca^{2+} ATPase is amongst the simplest proteins in this large family consisting of a single polypeptide chain of ~110 kDa. Its biological function is to transport $2Ca^{2+}$ ions against a concentration gradient for every molecule of ATP hydrolysed (Figure 5.61). Within the cytosol the concentration of Ca^{2+} is ~0.1 μM – about 10000 times lower than the extracellular or lumenal concentration (~1.0–1.5 mM). When muscle contracts large amounts of Ca^{2+} stored in the SR are released into the cytosol of muscle cells. In order to relax again after contraction the Ca concentration must be reduced by pumping ions back into the SR against a concentration gradient.

As an ion pump the Ca^{2+} ATPase works comparatively slowly with a turnover number of ~60 Ca^{2+} ions per second. Pumps such as the Na^+–K^+ ATPase transfer ions at rates of 10^6 s^{-1} and in order to overcome this rate deficiency for Ca^{2+} transfer the cell compensates by accumulating large amounts of enzyme within the SR membrane. The accumulation of enzyme in this membrane leads to ~60 percent of the total protein being Ca^{2+} ATPase.

The structure of the SR Ca^{2+} ATPase reveals two Ca^{2+} ions bound in a transmembrane domain of 10 α helical segments. In addition to the transmembrane region three cytosolic regions were identified as phosphorylation (P), nucleotide binding (N) and actuator (A) domains. This last domain transmits conformational change through the protein leading to movement of Ca across membranes. The organization of these domains can be seen in the beautiful structure of Ca bound SR-ATPase where ten transmembrane α helices (M1–M10) segregate into two blocks (M1–M6 and M7–M10), a structural organization consistent with the lack of M7–M10 in simpler bacterial P-type ATPases (Figure 5.62).

The lengths and inclination of the ten helices differ substantially; helices M2 and M5 are long and straight whilst others are unwound (M4 and M6) or kinked (M10) in the middle of the membrane. The Ca ions are located side by side surrounded by four transmembrane helices (M4, M5, M6 and M8), with the ligands to Ca^{2+} provided at site I by side-chain oxygen atoms

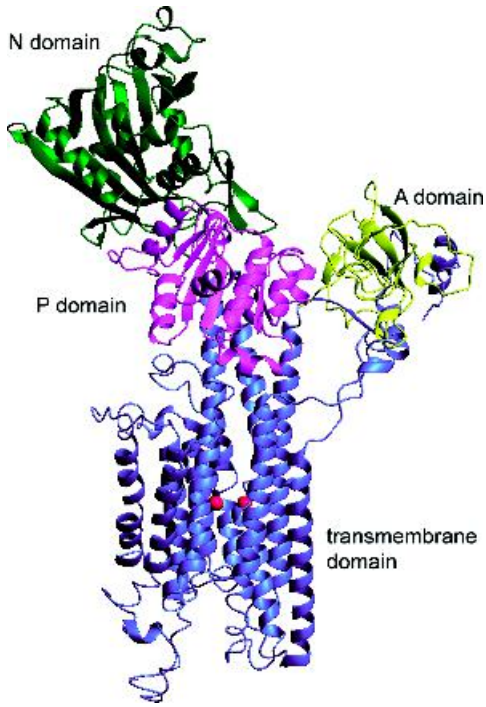


Figure 5.62 The structure of SR Ca-ATPase determined to a resolution of A. (PDB: 1EUL). The four distinct domains are shown in different colours; transmembrane domain (blue), Actuator or A domain (yellow), phosphorylation or P domain (magenta) and the nucleotide binding or N domain (green). Two bound Ca ions are shown in red. In this view the lumen of the SR is at the bottom whilst the cytosol is found at the top containing the A, P and N domains

derived from Asn768, Glu771 (M5), Thr799, Asp800 (M6) and Glu908 (M8). Site II is formed almost entirely by helix M4 with backbone oxygen atoms of Val304, Ala305 and Ile307 together with the side-chain oxygens of Asn796, Asp800 (M6) and Glu309 (M4) providing the key ligands. The coordinating ligands provided by residues 304, 305, 307 and 309 occur in residues located either sequentially or very close together in the primary sequence. Such an arrangement is not compatible with regular helical conformations and is only possible as a result of local unfolding of the M4 helix in this region. In this context a proline residue at position 308 almost certainly contributes to unfolding whilst the sequence PEGL is a key motif in all ATPases. Although the position of the Ca^{2+} is clearly seen within the transmembrane region the structure of the SR-CaATPase does not offer any clues as to the route of ion conduction. There were no obvious vestibules in the bound state, as seen in porins, and conformational changes facilitate transport of Ca ions across the membrane by coupling events in the transmembrane domain with those in the cytoplasmic domains. This involves Ca binding and release, nucleotide binding as well as Asp351 phosphorylation and dephosphorylation.

Summary

Following the determination of the structure of the bacterial photosynthetic reaction centre there has been a progressive increase in the number of refined structures for membrane bound.

Compared with the large number of structures existing for globular domains in the protein databank it is clear that the number of highly resolved (<5 Å) structures available for membrane proteins is much smaller. This will change rapidly as more membrane proteins become amenable to methods of structure determination.

Peripheral or extrinsic proteins are associated with the membrane but are generally released by mild disruptive treatment. In contrast, integral membrane proteins remain firmly embedded within the hydrophobic bilayer and removal from this environment frequently results in a loss of structure and function.

Crystallization of membrane proteins requires controlled and ordered association of subunits via the interaction of exposed polar groups on the surfaces of proteins. The process is assisted by the introduction of amphiphiles that balance conflicting solvent requirements of membrane proteins.

The vast majority of proteins rely on the folding of the primary sequence into transmembrane helical structures. These helices have comparable geometry to those occurring in soluble proteins exhibiting similar bond lengths and torsion angles.

The major differences with soluble proteins lie in the relative distribution of hydrophobic amino acid residues.

Table 5.10 Some of the membrane protein structures deposited in PDB; Most structures were determined using X-ray crystallography

Protein	PDB	Resolution (Å)	Function
Rhodopsin family and G-protein coupled receptors			
Bacteriorhodopsin (EM)	2BRD	3.5	Light driven proton pump
(X-ray)	1C3W	1.55	
Halorhodopsin	1E12	1.8	
Rhodopsin	1F88	2.8	Visual cycle transducing light energy into chemical signals
Sensory rhodopsin II	1H68	2.1	
Photosynthetic and light harvesting complexes			
R. viridis reaction centre	1PRC	2.3	Primary charge separation initiated by a photon of light
R. sphaeroides reaction centre	1PSS	3.0	
Light harvesting chlorophyll complexes from *R. acidophila*	1 KZU	2.5	Gathering light energy and focussing photons to the photosynthetic reaction centre
Photosystem I from *S. elongatus*	1JB0	2.5	Primary photochemistry
Photosystem II from *S. elongatus*	1FE1	3.8	Oxygenic photosynthesis
Bacterial and mitochondrial respiratory complexes			
Cytochrome oxidase (aa_3) from *P. denitrificans*	1AR1	2.8	Catalyses terminal step of aerobic respiration
Cytochrome oxidase (ba_3) from *T. thermophilus*	1EHK	2.4	
Cytochrome oxidase from bovine heart mitochondria	1OCC	2.8	
Bovine heart mitochondria bc_1 complex	1BGY	2.8	Electron transfer complex that oxidizes ubiquinonols and reduces cytochrome c
Chicken heart cytochrome bc_1 complex	1BCC	3.2	
S. cerevisiae cytochrome bc_1 complex	1EZV	2.3	

(*continued*)

Table 5.10 (*continued*)

Protein	PDB	Resolution (Å)	Function
Other energy transducing membrane proteins			
ATP synthase (F_1 c10) *S. cerevisiae*	1E79	3.9	Membrane portion forms proton channel
Fumarate reductase complex from *Wolinella succinogenes*	*1QLA*	2.2	
β *barrel membrane proteins of porin family*			
Porin from *R. capsulatus*	2POR	1.8	Semi-selective pore proteins
Porin from *R. blastica*	1PRN	2.0	
OmpF from *E. coli.*	2OMF	2.4	
PhoE from *E. coli.*	1PHO	3.0	
Maltoporin from *S. typhimurium*	2MPR	2.4	
Maltoporin from *E. coli*	1MAL	3.1	
OmpA from *E. coli.* (NMR)	1G90		
(X-ray)	1BXW	2.5	
Ion and other channel proteins			
Calcium ATPase from SR of rabbit	1EUL	2.6	Controls Ca flux in sarcoplasmic reticulum
Toxins			
α-haemolysin from *S. aureus*	7AHL	1.9	Channel forming toxin
LukF from *S. aureus*	3LKF	1.9	Homologous in structure and function to α-haemolysin

Membrane proteins such as toxins or pore proteins are based on β strands assembling into compact, barrel structures that assume great stability via inter-strand hydrogen bonding.

Sequencing proteins assists in the definition of transmembrane domains *via* identification of hydrophobic side chains. Proteins containing high proportions of residues with non-polar side chains such as Leu, Val, Ile, Phe, Met and Trp in blocks of ~20–30 residues are probably membrane bound.

The structures of bacteriorhodopsin revealed seven transmembrane helices of ~23 residues in length correlating with seven 'blocks' in the primary sequence composed predominantly of residues with non-polar side chains.

The seven transmembrane helix structure is a common architecture in membrane proteins with G-protein coupled receptors sharing these motifs.

Methods developed for crystallization of the bacterial reaction centre from *R. viridis* proved applicable to other classes of membrane proteins including respiratory complexes found in mitochondrial membranes and in aerobic bacteria. These complexes generate proton gradients vital for the production of ATP and represent the 'power stations' providing energy for all cellular processes.

The structure of Ca ATPase reveals the coordinated action of cytosolic and membrane located domains in Ca transport across the membrane and points the way to understanding the movement of charged ions across impermeant lipid bilayers.

ATP is produced by ATP synthase a large enzyme composed of a peripheral F_1 complex together with a smaller membrane bound region (F_o). These domains form a rotary motor where the flux of protons through F_o drives rotation of c subunits via transmission to a central stalk (γ, δ and ε subunits) that mediates conformational changes at the catalytic centres in $(\alpha\beta)_3$ subunits promoting ATP formation.

Problems

1. Having determined the nucleotide sequence of a putative membrane protein describe how you would use this data to obtain further information on this protein.
2. List the post-translational modifications added to proteins to retain them within or close to the lipid bilayer. Give an example protein for each modification.
3. Find the membrane-spanning region in the following sequence: . . . . . .GELHPDDRSKITKPSESIITTIDS NPSWWTNWLIPAISALFVALIYHLYTSEN. . . . .
4. How many residues are required to span a lipid bilayer if they are all form within a regular α helix? How many turns of a regular α helix are required to cross the membrane once?
5. Using the crystal structure for the bacterial reaction centre estimate the average length of the transmembrane helices. Are there any differences and if so how do you account for them?
6. Describe the advantages and disadvantages of using the erythrocyte membrane as a model for studying membrane protein structure and organization.
7. Use protein databases to find membrane protein structures determined at resolution below 0.3 nm and those between 0.3 and 0.5 nm.
8. Gramicidin is a small polypeptide of 15 residues that adopts a helical conformation. This molecule has been observed to act as an ionophore allowing K^+ ions to cross the membrane. How might this occur?