Crystallography of Membrane Proteins, Major Targets in Drug Design

J. Wouters*

Institut de Recherches Microbiologiques JM Wiame, 1 Av E.Gryson, B-1070 Brussels, Belgium

Abstract: Protein crystallography has the potential to accelerate drug discovery greatly. High-resolution structures of membrane proteins of pharmaceutical interest open new perspectives in drug design. Recent structural data obtained for cyclooxygenases, monoamine oxidase, squalene cyclase, rhodopsine, porins, aquaporins, and ABC transporters are presented and briefly discussed.

Keywords: Membrane proteins crystallography.

INTRODUCTION

Together with progress in biology and bioinformatics, advances in crystallisation, diffraction data collection and processing have aided the development of high-throughput X-ray crystallography. The wealth of information made available through efforts in structural genomics and advances in computation has allowed structure-based drug design to emerge as a valuable tool in medicinal chemistry. In the past, combinatorial chemistry, coupled with high-throughput approaches, shifted attention away from the more structure-based methods. Large-scale determination of protein structures is reversing the drug discovery process by starting with the protein structure and using it to identify and design new ligands. It is the integration of structure-based methods, virtual screening, and combinatorial chemistry that will provide the basis for more efficient drug design in the future, significantly reducing the time of the design cycle and the cost per marketed drug.

Different aspects of the use of protein crystallography in drug design have been extensively reviewed by others (for a recent review see e.g. [1] and references therein) and will not be covered here. In the present paper recent advances in the field of **membrane protein crystallography** will be reviewed in connection with the great perspectives that those new structures offer in terms of drug design.

Obtaining large, well-ordered crystals is essential for a crystallographer to be able to analyse the three-dimensional molecular structure and active site of the proteins. The ability to design drugs using structure-based drug design is thus dependent upon the quality of the crystals that are available to examine. Crystals are grown from large quantities of the purified target enzyme and then by setting up a system to evaporate most of the water molecules that are mixed in with the enzyme. By removing the water, the enzyme is left in a hardened or crystallised form.

In this context, membrane proteins are a challenge for crystallographer as they prove very difficult to purify and crystallise. As a consequence, the number of membrane-spanning proteins structures solved to atomic resolution is very small. At the time of writing, over 18.800 crystal structures are available at the PDB [2] but only about 55 correspond to membrane proteins [3, 4]. Indeed, in recent years, the rate of structure determination of membrane proteins has steadily increased. But the total number of atomic structures of those membrane proteins, established over 20 years, is still smaller than the number of structures of soluble proteins determined in a single week !

A compilation of those membrane proteins of known structure, including crystallisation conditions and key references for the structure determinations is available on the net [5, 6]. This list is frequently updated. The crystallisation conditions used to obtain those crystals are also reported. In this context, one should underline the recent efforts made to develop techniques and reagents specifically dedicated to membrane protein crystallisation (design and purification of additives, detergents, use of cubic lipidic phases, crystallisation of proteins in complex with anti-bodies, ..). These aspects are not covered in the present work.

The membrane proteins listed on the net [5, 6] are divided into three categories:

(A) polytopic membrane proteins of the helical type from inner membranes of bacteria and mitochondria, and from eukaryotic membranes. 31 structures have currently been solved (among them 15 unrelated ones).

(B) ß-stranded membrane proteins from the outer membrane of gram negative bacteria and related membrane proteins. 20 structures are available, 1 four-stranded, 1 seven-stranded, 2 eight-stranded, 2 ten-stranded, 7 sixteen-stranded, 3 eighteen-stranded, 4 twenty two-stranded ß-barrel proteins (per monomer each).

(C) monotopic membrane proteins that are only inserted into the membrane, but do not cross it. 4 structures are available.

This is an extremely small number if it is considered that about one-third of all genes code for membrane proteins [7, 8]. In particular, in the human genome, about 10 000 genes are predicted to encode membrane proteins many of which (including G protein couple receptors, ion channels) are major drug targets. But only two crystal structures of human membrane proteins (human cyclooxygenase-2 and monoamine oxidase) are available.

*Address correspondence to this author at the IRMW, 1 Av E.Gryson, 1070 Brussels, Belgium; Tel:+32 2 527 36 34; Fax: +32 2 526 72 73; E-mail: jwouters@dbm.ulb.ac.be

1389-5575/03 $41.00+.00 © 2003 Bentham Science Publishers, Ltd.

Therefore, any new structure of membrane protein opens new perspectives in the understanding of their molecular mechanisms. Based on the three-dimensional (3D) structures of membrane proteins of direct pharmaceutical interest, structure-based drug design and virtual screening techniques can be applied in order to facilitate drug discovery. This is illustrated in the case of cyclooxygenases (nonsteroidal anti-inflammatory drug target) and monoamine oxidase (neurotransmitters metabolism). Both are monotopic membrane proteins of pharmaceutical interest. Determination of their 3D structures offered opportunities to direct design inhibitors. Other membrane proteins are not direct drug targets but allow precise modeling of major drug targets. This is illustrated by the structures of squalene cyclase and bovine rhodopsin that provide valuable templates for modeling of eukaryotic oxidosqualene cyclase (cholesterol biosynthesis) and many G-protein-coupled receptors (GPCRs) respectively. Other potentially interesting membrane proteins will be briefly presented : pore forming beta barrels (toxins and porins) and ABC transporters.

CYCLOOXYGENASES

Prostaglandins and glucocorticoids are potent mediators of inflammation. Non-steroidal anti-inflammatory drugs (NSAIDs) exert their effects by inhibition of prostaglandin production. The pharmacological target of NSAIDs is cyclooxygenase (COX, also known as PGH synthase), which catalyses the first committed step in arachidonic acid metabolism. Two isoforms of the membrane protein COX are commonly described: COX-1, which is constitutively expressed in most tissues, is responsible for the physiological production of prostaglandins; and COX-2, which is induced by cytokines, mitogens and endotoxins in inflammatory cells, is responsible for the elevated production of prostaglandins during inflammation. Recently a third distinct COX isozyme, COX-3, has been described [9]. Comparison of COX-3 activity with murine COX-1 and -2 demonstrates that this enzyme is selectively inhibited by analgesic/antipyretic drugs such as acetaminophen, phenacetin, antipyrine, and dipyrone, and is potently inhibited by some nonsteroidal antiinflammatory drugs. Thus, inhibition of COX-3 could represent a primary central mechanism by which these drugs decrease pain and possibly fever.

For a long time, NSAIDs available for clinical use inhibited both COX-1 and COX-2. They inhibit pro-inflammatory prostaglandins (PGs) derived from the activity of COX-2, as well as PGs in tissues like the stomach and kidney (*via* COX-1). New classes of compounds have been developed that have a high degree of selectivity for the inducible form of cyclooxygenase (COX-2) over the constitutive form (COX-1). Design and understanding of the molecular basis of selectivity of these compounds have been greatly facilitated by crystallographic studies (for a recent review see e.g. [10]).

The first 3D structure of prostaglandin H2 synthase-1 has been determined in 1994 (PDB accession number 1PRH [11]) at 3.5 Å resolution by X-ray crystallography. Shortly after this structure determination, it was demonstrated that aspirin exerts its anti-inflammatory effects through selective acetylation of serine 530 on prostaglandin H2 synthase (PGHS) by solving the 3.4 Å resolution X-ray crystal structure of PGHS isoform-1 inactivated by the potent aspirin analogue 2-bromoacetoxy-benzoic acid (PDB accesison number 1PTH [12]). Acetylation by this analogue abolishes cyclooxygenase activity by steric blockage of the active-site channel and not through a large conformational change. Two rotameric states of the acetyl-serine side chain are observed which block the channel to different extents, a result that may explain the dissimilar effects of aspirin on the two PGHS isoforms.

In 1996 the structures of unliganded murine COX-2 and complexes with flurbiprofen, indomethacin and SC-558, a selective COX-2 inhibitor, determined at 3.0 to 2.5 Å resolution have been reported (PDB accession numbers 3PGH and 1CX2 [13]). These structures explain the structural basis for the selective inhibition of COX-2 and demonstrate some of the conformational changes associated with time-dependent inhibition.

Both COXs are membrane-anchored proteins that exist as dimers and have remarkable structural similarity. They are bifunctional enzymes and comprise three independent folding units: an epidermal growth factor domain, a membrane-binding motif and an enzymatic domain. Two adjacent but spatially distinct active sites were found for haeme-dependent peroxidase and cyclooxygenase activities. The cyclooxygenase active site is created by a long, hydrophobic channel that is the site of non-steroidal anti-inflammatory drug binding. The conformation of the membrane-binding motif strongly suggests that the enzyme integrates into only one leaflet of the lipid bilayer and is thus a monotopic membrane protein.

While the tertiary structures of both COX isozymes are remarkably similar, COX-2 is characterised by a side pocket extension to the hydrophobic channel. Indeed a major difference between the COX-1 and COX-2 active sites is that the COX-1 active site is smaller than that of COX-2 (Fig. (**1**)). This is primarily due to valine to isoleucine substitutions at residues 434 and 523. The smaller Val-434 side chain in COX-2 enables the opening of a solvent accessible space, referred to as the side pocket, that increases the total volume of the cyclooxygenase active site. Indeed in the structure of the NSAID flurbiprofen (a non selective NSAID) bound to COX-2, side pocket residues of COX-1 differ from those of COX-2 [13]). The carboxylate of flurbiprofen forms an ion pair with the conserved Arg120 and a hydrogen bond with Tyr355. The side pocket of COX-2 is unoccupied in this complex. The position of flurbiprofen in the COX-2 complex is virtually identical to that observed with the COX-1 complex, consistent with the nonselective inhibition by this NSAID. In the structure of the COX-2-specific inhibitor SC-558 (1900-fold selectivity) bound to wild-type murine COX-2, the phenylsulfonamide moiety of SC-558 binds tightly in the COX-2 side pocket, forming several hydrogen bonds and van der Waals interactions. This appears to be one of the major determinants of inhibitor selectivity, although not the only factor. The actual position of Phe518 appears to be critical for the proper binding of SC-558 and other diarylheterocycle inhibitors. As the sidechain of Phe518 is located in the main channel of the active site, it could also dictate the isoform

selectivity of other classes of inhibitors that do not access the side pocket. It appears that the two conformations of Phe518 are influenced by sequence differences between COX-1 and COX-2 at residues Gly435 and Cys512 (COX-1 numbering), located in the second shell of residues around the active site.

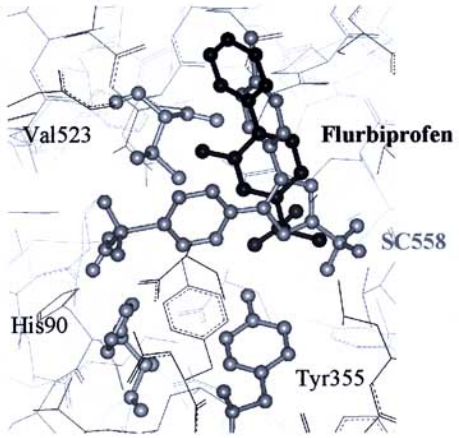

Fig. (1). Isoform-selective inhibitor binding. The COX-1 and COX-2 active sites are shown superimposed (COX-1, dark; COX-2, grey). Two inhibitors are seen: flurbiprofen (dark), a nonselective inhibitor, and SC-558 (grey), a COX-2–selective inhibitor. NSAIDs achieve COX inhibition by occupying the upper portion of the active site channel, preventing the fatty acid substrate from gaining access to an active site tyrosine. The COX-2–selective inhibitor projects into a side pocket that is not exploited by the nonselective inhibitor.

Mechanistical studies have also been possible thanks to the crystal structures. X-ray crystallographic studies indicate that the cyclooxygenase reaction occurs in a hydrophobic channel that extends from the membrane binding domain of the enzyme into the core of the globular domain. The structures of PGHS-1 with arachidonic acid (PDB accession number 1DIY [14]) bound in a chemically productive conformation, with dihomo-gamma-linolenic acid [15], with eicosapentaenoic and linoleic acids [16] have also been determined and compared.

The fatty acids adopt extended L-shaped conformations similar in all complexes. In this conformation, the 13proS hydrogen of arachidonic acid is well positioned for abstraction by tyrosine-385, the likely radical donor. A space also exists for oxygen addition on the antarafacial surface of the carbon in the 11-position (C-11). While this conformation allows endoperoxide formation between C-11 and C-9, it also implies that a subsequent conformational rearrangement must occur to allow formation of the C-8/C-12 bond and to position C-15 for attack by a second molecule of oxygen.

The arachidonic acid substrate gains thus access to the active site *via* a hydrophobic channel. This access is blocked by interpolation of an acetyl residue on Ser 530 (Ser 516 in COX-2 [12]). NSAIDs, by contrast, interact competitively with the active site.

Although the initial selective COX-2 inhibitors were discovered with the tools of classical biochemical pharmacology, structural studies reveal their localisation in the side pocket, where they interact with slow, tight-binding kinetics. The advent of COX-2 inhibitors and their introduction into clinical practice has rendered prostaglandin research fashionable (not to mention profitable) again. Based on the crystal structures of both forms of this membrane protein, new ligands have been proposed [17, 18].

MONOAMINE OXIDASES

Monoamine oxidase (MAO) is a key enzyme responsible for the degradation of serotonin, norepinephrine, dopamine, and phenylethylamine neurotransmitters [19, 20]. It is an outer membrane mitochondrial flavoenzyme existing in two isoforms, A and B, sharing about 70% sequence identity. Monoamine oxidase is a well-known target for antidepressant and neuroprotective drugs. The structure of the human type B enzyme has been determined by X-ray crystallography (PDB accession code 1GOS [21]). A key element towards the determination of the structure was the high level expression of human MAO B in *Pichia pastoris* in a fully-functional form and purification to homogeneity. This system provided sufficient amounts of enzyme to permit the solution of the structure to a better resolution than 3 Å by multicrystal averaging based on two crystal forms (P_1 and C_{222}) grown from different detergent crystallisation conditions. The structure was solved both as the free and as a covalent flavocyanine complex with pargyline. The structure of human MAO B provides a framework for probing the catalytic mechanism, understanding the differences between the B- and A-monoamine oxidase isoforms and designing specific inhibitors.

Together with COXs and squalene-hopene cyclase (see next section), MAO B structure is one of the first folds of monotopically-inserted membrane proteins to be solved. The enzyme binds to the membrane through a C-terminal transmembrane helix and apolar loops located at various positions in the sequence (Fig. (**2**)). The overall MAO B structure is dimeric but does not involve helix-helix interactions within the membrane inserted segments.

The substrate binding site is a flat cavity located in front of the flavin and lined by hydrophobic residues. A smaller 'entrance' cavity forms an entry way between the active site pocket and the protein surface. The opening of the entrance cavity is toward the membrane surface suggesting substrate access is from the surface of the membrane bilayer. The two-domain overall topology of MAO B is similar to the fold of *p*-hydroxybenzoate hydroxylase and polyamine oxidase, as predicted earlier by sequence analysis [22]. The electron density shows that pargyline, an analog of the clinically used MAO B inhibitor, deprenyl, binds covalently to the flavin N5 atom. The recognition site for the substrate amino group is an aromatic cage formed by Tyr 398 and Tyr 435.

Site-directed mutagenesis had already underlined the critical role of key amino acids in both MAO A and B. Based on the polyamine oxidase three-dimensional crystal structure, it was suggested that Lys-305, Trp-397 and Tyr-407 in MAO A (Lys-296, Trp-388, and Tyr-398 in MAO B)

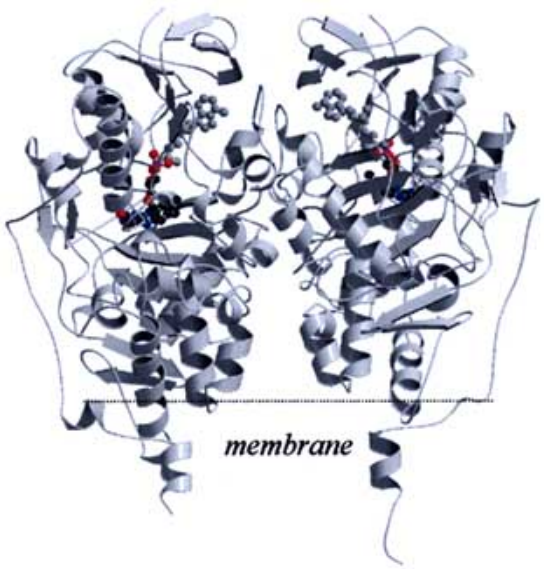

Fig. (2). Monoamine oxidase B adopts a dimeric structure with the C-terminal transmembrane helix and apolar loops anchoring the protein into the outer mitochondrial membrane [adapted from 21].

may be involved in the non-covalent binding to flavin (FAD). It was further proposed that Tyr-407 and Tyr-444 in MAO A (Tyr-398 and Tyr-435 in MAO B) may form an aromatic sandwich that stabilises the substrate binding.

The similar impact of analogous mutants in MAO A and MAO B suggests that these amino acids have the same function in both isoenzymes [23]. Three-dimensional modeling of MAO A using human MAO B as template suggests that the overall tertiary structure and the active sites of MAO A and B may be similar and provides a precise starting point for design of potent and selective inhibitors.

SQUALENE CYCLASE

Squalene-hopene cyclase (SHC) catalyses the conversion of squalene into pentacyclic compounds. It is the prokaryotic counterpart of the eukaryotic oxidosqualene cyclase (OSC) that catalyses the cyclisation step in cholesterol biosynthesis.

An elevated cholesterol level in the plasma is a major risk factor in the development of artherosclerotic vascular diseases. Currently, cholesterol biosynthesis is reduced by inhibiting the rate-limiting enzyme 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase by the so-called statins. These, however, can cause adverse effects due possibly to the suppression of ubiquinone, dolichol, and isopentenyl-tRNA farther along the mevalonate pathway [24]. Moreover, HMG-CoA reductase biosynthesis is upregulated upon inhibition [25].

OSC is an attractive target for novel anticholesteremic drugs because it acts downstream from important branching points in the cholesterol pathway and because the inhibition of OSC happens to avoid accumulation of steroidal intermediates. OSC inhibition raises the concentrations of 2,3,22,23-dioxidosqualene and oxysterols that down-regulate the HMG-CoA reductase activity *via* a negative-feedback loop. Potent, orally active inhibitors of human hepatic OSC are Ro48-8071 containing a benzophenone (BP) moiety (fig. (**3**)) [26], BIBX-79 [27], BIBB-515 [28] and quinuclidine inhibitors [29]. Because of clear sequence homology, SHC can serve as a template for OSC.

The crystal structure of squalene-hopene cyclase from *Alicyclobacillus acidocaldarius* was determined at 2.9 Å resolution (PDB accession code 2SQC [30]). The structure reveals a membrane protein with membrane-binding characteristics similar to those of prostaglandin-H2 synthase. The active site of the enzyme is located in a large central cavity that is of suitable size to bind squalene in its required conformation and that is bordered by aromatic residues. The structure supports a mechanism in which the acid starting the reaction by protonating a carbon-carbon double bond is an aspartate that is coupled to a histidine. Numerous surface alpha helices are connected by characteristic QW-motifs that tighten the protein structure, possibly for absorbing the reaction energy without structural damage.

The structure of SHC shows a dimeric monotopic integral membrane protein consisting of two β-barrel domains. A large central cavity is enclosed by loops and a short five-stranded sheet linking the barrel domains. Asp376 of the conserved motif DxD is located at the "top" of the cavity, where it initiates the cyclisation by protonating a terminal double bond. The cavity is connected by a nonpolar channel to a hydrophobic plateau structure, which is probably inserted into the nonpolar interior of the membrane. The channel is constricted by residues Phe166, Val174, Phe434 and Cys435. The carbocationic intermediates are stabilised by cation-pi interactions (for a recent review on these interactions see [31]) with tryptophans and phenylalanines, which provide an electron-rich but also non-nucleophilic environment

Recently the crystal structure of SHC in complex with Ro48-8071 (fig. (**3**)), a potent inhibitor of OSC and therefore of cholesterol biosynthesis has been established (PDB accession code 1GSZ [32]). Ro48-8071 is bound in the active-center cavity of SHC and extends into the channel that connects the cavity with the membrane. The binding site of Ro48-8071 is largely identical with the expected site of squalene. It differs from a previous model based on photoaffinity labeling. The knowledge of the inhibitor binding mode in SHC is likely to help develop more potent inhibitors for OSC.

G-PROTEIN-COUPLED RECEPTORS

The seven-transmembrane (7TM) receptors are the fourth largest superfamily in the human genome [33] with more than 600 predicted genes. Of the drugs used clinically in humans, 40% target 7TM receptors. 7TM receptors are ligand-activated transmembranar switches, almost all of which activate heterotrimeric GTP binding proteins (G-protein-coupled receptors (GPCRs)). Rhodopsin (structurally distinct from the microbial rhodopsins, which are light-

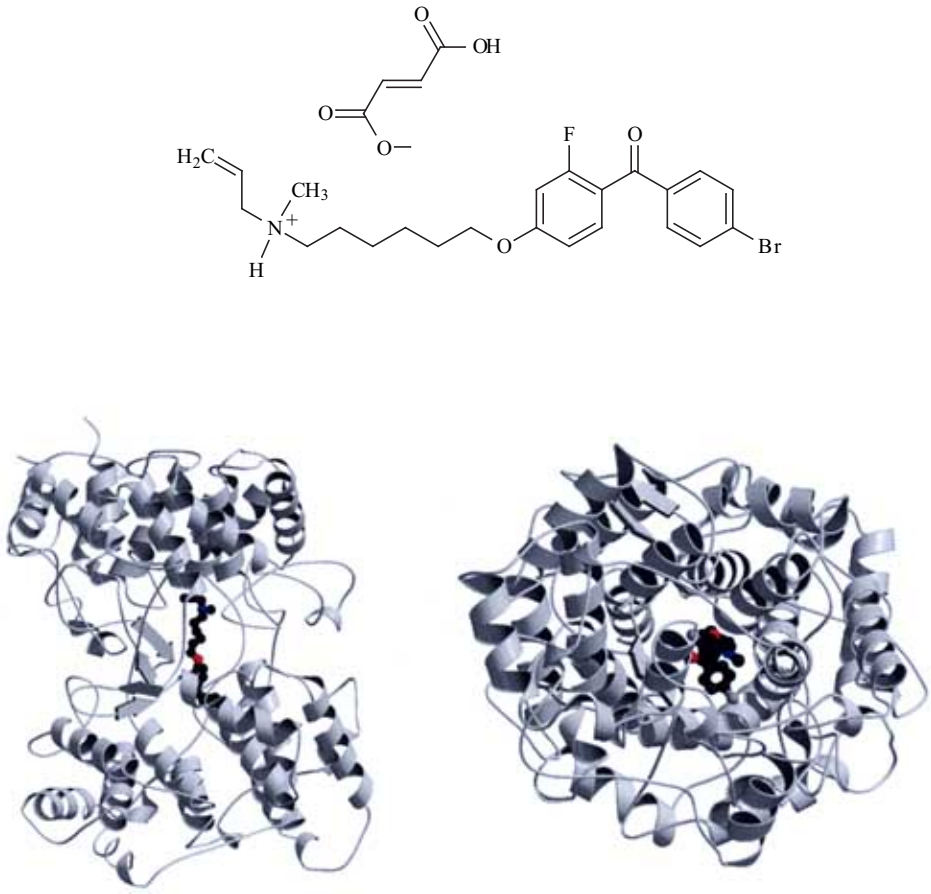

Fig. (3). Structure of inhibitor Ro48-8071 (top) and two perpendicular views of ribbon models (bottom) of squalene-hopene cyclase in complex with Ro48-8071 (shown in ball-and-stick). The protein is supposed to be oriented like in the left (bottom) orientation and plunge into the nonpolar membrane interior. The right (bottom) orientation is a perpendicular view from the membrane and shows the channel leading to the binding site occupied by the ligand.

activated pumps) is unique among the 7TM receptors because its ligand, 11-*cis*-retinal (RET), is covalently bound by formation of a Schiff's base linkage to the amino group of Lys296 in transmembrane helix VII. The recent determination of the structure of rhodopsin (Rho) by X-ray crystallography opens avenues to a deeper understanding of GPCR activation and transmembrane signaling [34-38]. It also provides an improved template for interpreting the huge body of structure activity, mutagenesis, and affinity labelling data available for related 7TM receptors of pharmaceutical interest (see for example [39] for a review on M1 muscarinic acetylcholine receptor modeling based on Rho crystal structure and [40]).

Indeed, Rho is the first GPCR for which a crystal structure has been reported (PDB accession numbers 1F88, 1HZX [39]; for an animation of the crystal structure of rhodopsin see [41]). To obtain crystals, bovine Rho was purified from rod outer segment membranes and crystallised from a detergent solution, nonyl-thiol-glucoside supplemented with the small amphiphile heptane 1,2,3-triol. The resolution of the crystallographic data is about 2.8 Å, but short segments of the cytoplasmic surface domain are not resolved. The structure represents the inactive form of Rho with its bound RET chromophore intact. A ribbon diagram of the Rho peptide backbone structure with the RET chromophore is presented in Fig. (**4**). The structure discussed in this review is that of the A chain in the crystal unit cell dimer.

Previous low-resolution electron density maps had roughly outlined the positions of the transmembrane helices [42], and the high-resolution structure is consistent with these maps. A similar arrangement is also observed in the crystal structure of light-driven ion pump bacteriorhodopsine [43-45]. The TM segments are tilted with respect to the presumed plane of the membrane bilayer. They are generally α-helical, but they contain significant kinks and irregularities. The helical segments form a compact bundle that contains the binding site for the RET chromophore. In addition to the seven-helix transmembrane bundle, the new structure shows, for the first time, the N-terminus and all three extracellular (EC) loops, plus (with two short gaps) the three intracellular (IC) loops and the C-terminal tail. The EC and IC structures offered several surprises, including a compact EC `lid', parts of which fold inwards to completely

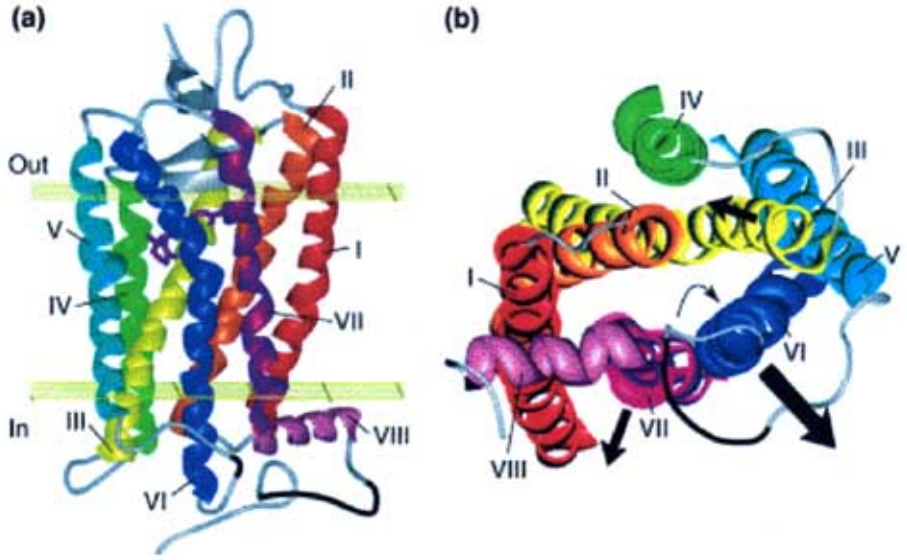

Fig. (4). (a) Crystal structure of rhodopsin. The seven transmembrane helices are shown with the following color coding: I, red; II, orange; III, yellow; IV, green; V, cyan; VI, dark blue; and VII, magenta. 11-*cis*-retinal is also shown in magenta. Helix VIII is shown in light pink, and all the nonhelical portions are shown in gray. Connections (black) have been drawn to roughly represent segments that are missing from the structure (i.e. four residues in intracellular loop 3 and six residues in the C-terminal tail). The yellow panels show the approximate boundaries of the hydrophobic core of the membrane.

enclose retinal, and an eighth helix (VIII) on the IC side, approximately parallel to the plane of the membrane and perpendicular to the seven-helix bundle.

In a recent structure report of Rho, resolution was extended to 2.6 Å and allowed identifiction of internal water molecules potentially essential for regulation of the activity of Rho [46]. The structure model further suggest that those water molecules, which are observed in the vicinity of highly conserved residues and in the retinal pocket, play an important role in all rhodopsin-like GPCRs.

The structure of rhodopsin's TM helix bundle was predicted, with considerable precision, in previous modeling studies. Models that now appear as the most accurate were guided by the low-resolution electron density of rhodopsin [42, 43] by homology with the structure of bacteriorhodopsin [45] and by patterns of conservation in the amino acid sequences of other GPCRs. This level of accuracy suggests that the TM bundle structure is indeed conserved among GPCRs. By contrast, the loops and termini are more divergent in amino acid sequence and probably in three-dimensional structure. Taking those limitations, methodological tools were developed to predict the structures, ligand binding sites and relative binding affinities of those membrane receptors involved in cell recognition and communication processes [47].

Despite the large number of models of GPCRs, our knowledge on how agonist binding to receptors results in G protein activation still remains unclear. The crystal structure of Rho in its ground state has represented a significant breakthrough in GPCR research. However, a better understanding of how the active conformation of rhodopsin is activated by agonists awaits the resolution of the active structure of the receptor. New structural approaches like those recently described by the group of Khorana [48-50] will be extremely useful in elucidating the architecture of the receptor-G protein interface. The results from systematic mutational analysis of different receptors represent another important step in determining the role of individual amino acids in GPCR function.

OTHER MEMBRANE PROTEINS OF POTENTIAL PHARMACEUTICAL INTEREST

Among other potential interesting membrane proteins of known 3D structure, one finds β-barrel membrane proteins (porins and toxins), ion channels (including aquaporins), ATP binding cassette transporters, ATPases, and respiratory proteins [3-5]. Some of them are briefly discussed here in connection with their therapeutical interest.

β-Barrel Membrane Proteins

Crystallographic studies of the past ten years have revealed that many outer membrane proteins and bacterial toxins are constructed on the β-barrel motif. Two structural classes can be identified. The first class, represented by the porins, includes monomeric or multimeric proteins where each β-barrel is formed from a single polypeptide. The second class features proteins where the beta-barrel is itself a multimeric assembly, to which each subunit contributes a

few beta-strands. In addition to structural investigations, much work has also been devoted to the functional aspects of these proteins, and to the relationships between structure and function. The structural and the functional properties of some of the best-studied examples of these various classes of proteins, namely the general-diffusion, specific and ligand-gated porins, multidrug efflux proteins and the staphylococcal toxin α-hemolysin have been reviewed recently [51].

Among those β-barrel membrane proteins, staphylococcal **α-hemolysin** is the prototype of a family of bacterial exotoxins with membrane-damaging function (for a review see *e.g.* [52]). These toxins are secreted in a soluble form that finally converts into a transmembrane pore by assembling an oligomeric β-barrel, with hydrophobic residues facing the lipids and hydrophilic residues facing the lumen of the channel. Besides α-hemolysin the family includes other single chain toxins forming homo-oligomers, e.g. β-toxin of *Clostridium perfringens*, hemolysin II and cytotoxin K of *Bacillus cereus*, but also the staphylococcal bi-component toxins, like γ-hemolysins and leucocidins, which are only active as the combination of two similar proteins which form hetero-oligomers. The molecular basis of membrane insertion has become clearer after the determination of the crystal structure of both the oligomeric pore and the soluble monomer (PDB accession numbers 7AHL, 1LKF [53, 54]).

Contained within the mushroom-shaped homo-oligomeric heptamer structure of *Staphylococcus aureus* the α-hemolysin pore (fig. (**5**)) [53] is a solvent-filled channel that runs along the sevenfold axis. The lytic transmembrane domain comprises the lower half of a 14-stranded antiparallel β barrel, to which each protomer contributes two β strands. The interior of the β barrel is primarily hydrophilic, and the exterior has a hydrophobic belt. The structure proves the heptameric subunit stoichiometry of the α-hemolysin oligomer, shows that a glycine-rich and solvent-exposed region of a water-soluble protein can self-assemble to form a transmembrane pore of defined structure, and provides insight into the principles of membrane interaction and transport activity of beta barrel pore-forming toxins. The crystal structure of the water soluble form of LukF (an homologue of alpha-hemolysin) [54] illustrates how a channel-forming toxin masks protein-protein and protein-membrane interfaces prior to cell binding and assembly, and together with the α-hemolysin heptamer structure, they define the end points on the pathway of toxin assembly.

Studies on this family of β-barrel pore-forming toxins are important for many aspects: (i) they are involved in serious pathologies of humans and farmed animals, (ii) they are a good model system to investigate protein-membrane interaction and (iii) they are the basic elements for the construction of nanopores with biotechnological applications in various fields (e.g. [55]).

The trimeric **TolC protein** from *E. coli*, involved in multidrug efflux and protein excretion is another example of single barrel assembled from multiple subunits. Recently the detergent solubilized TolC X-ray structure was solved (PDB

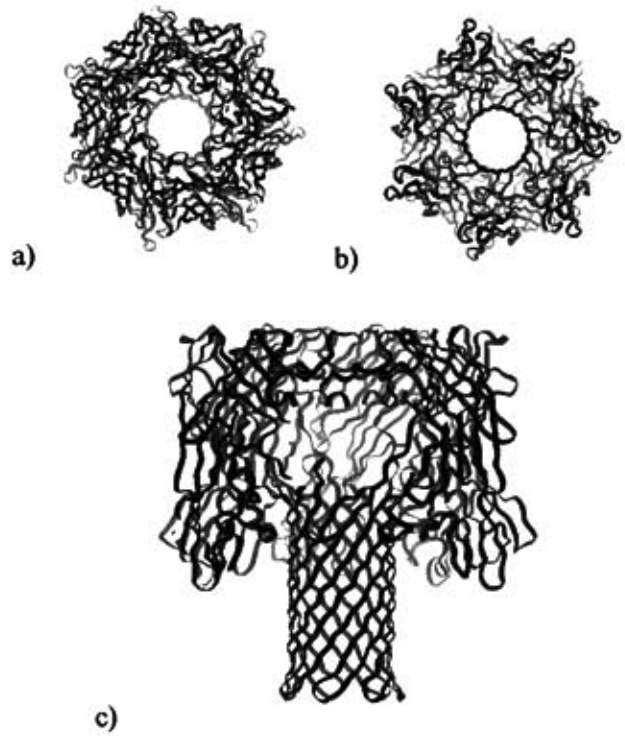

Fig. (5). Structure of the α-hemolysin oligomer. (a) Top (extracellular side), (b) bottom (intracellular side) and lateral (c) views.

accession code 1EK9 [56]) and disclosed a very peculiar architecture of a β-barrel extended into an α-helical barrel of coiled-coil helices. Each monomer contributes 4 β-strands and 4 α-helices, making the assembled 12-stranded β-barrel domain. The β-barrel is 40 Å in height, presumably spanning the outer membrane, and is indeed very similar to porin structures. The height of the α-helical domain is 100 Å, making this domain sufficiently long to span the entire periplasm. Thus the whole trimer ressembles a cannon, with a large solvent-filled internal cavity which would accomodate a variety of solutes. This structure can thus serve as a rational base to explain and understand protein export and drug efflux [51].

Ions and Other Channels

Several crystal structures of channels have recently been obtained. They include KcsA, a H^+ gated potassium channel of *Streptomyces lividans* at 3.2 Å (PDB accession code 1BL8 [57]) MthK, a Ca^{++} gated potassium channel of *Methanobacterium thermoautotrophicum* at 3.3 Å (PDB accession code 1LNQ [58]), the mechanosensitive channel MscL of *Mycobacterium tuberculosis* at 3.5 Å (PDB accession code 1MSL [59]), a chloride channel of *Salmonella typhimurium* at 3.0 Å (PDB accession code 1KPL [60]), aquaporin water channel (PDB accession code 1J4N [61]) and a glycerol facilitator channel (GlpF) of *E. coli* at 2.2 Å (PDB accession code 1FX8 [62].

Voltage-dependent K^+, Ca^{2+}, and Na^+ **channels** play vital roles in basic physiological processes, including management of the action potential, signal transduction, and secretion. They share the common function of passively transporting ions across cell membranes; thus, it would not be surprising if they should exhibit similarities of both structure and mechanism. Indeed, the principal pore-forming (α) subunits of each show either exact or approximate 4-fold symmetry and share a similar transmembrane topology, and all are gated by changes in membrane potential. Furthermore, all these channels possess an auxiliary polypeptide, designated the beta subunit, which plays an important role in their regulation. Despite considerable functional ressemblances and abilities to interact with structurally similar α subunits, however, there is considerable structural diversity among the β subunits. The similarities and differences in the structures and functions of the β subunits of the voltage-dependent K^+, Ca^{2+}, and Na^+ channels have been recently reviewed [63].

Existing drugs that modulate ion channels represent a key class of pharmaceutical agents across many therapeutic areas. For example, there is considerable potential for potassium channel drug discovery as they represent the largest and most diverse sub-group of ion channels and they play a central role in regulating the membrane potential of cells. Recent advances in genomics have greatly added to the number of these potential drug targets, and selecting a suitable potassium channel for drug discovery research is becoming a key step [64].

Determination of crystal structures of ion channels have also stimulated much interest in modeling, focusing especially on the question of structure-function relationships, and on how permeation models can be applied [65]. In this context, the conformational changes associated with gating transitions between closed and open states of channels have been reviewed [66], emphasizing the potential roles of helix-helix interactions in this process.

Among other transmembrane channels, **aquaporins** also deserve particular attention. These water channels facilitate the rapid transport of water across cell membranes in response to osmotic gradients. They are believed to be involved in many physiological processes that include renal water conservation, neuro-homeostasis, digestion, regulation of body temperature and reproduction. The atomic structure of AQP1 protein (PDB accession code 1J4N [61]) provides marked insight into several human disease states. Ten other aquaporin genes are also expressed in humans, and their structures are expected to be very similar. Each aquaporin is present in specific tissues where their permeabilities to water and small solutes may contribute to multiple physiological processes. So far, investigators have identified mutations in the water-selective human homologs AQP0, AQP1, and AQP2.

Aquaporins are suspected in numerous disorders involving fluid transport such as brain edema, cirrhosis, congestive heart failure, glaucoma, and pre-eclampsia (for a recent review : [67]). Multifunctional aquaglyceroporins that transport glycerol and other small molecules may have roles in energy metabolism and heavy metal transport, and at least one member of the aquaporin family may have a role in acid-base homeostasis as a kidney anion channel. Recent advances in determining the structures of aquaporins at the atomic level have revealed key mechanisms by which these channels maintain exquisite selectivity for substrates without sacrificing high rates of transport. Further investigation may reveal that structural signals lead to differences in several aspects of aquaporin function, such as the basis (and apparent requirement) for oligomerization, channel permeability properties, stability, and trafficking. Knowledge of protein structure may also provide insight into the spectrum of disease caused by distinct mutations in a single aquaporin gene, as in the case of mutations in AQP0 or AQP2. Finally, it is hoped that mechanistic and structural insights will lead to the development of new therapeutics through rational drug design. Considering the diverse expression patterns and functional properties of the aquaporins, the application of aquaporin structural biology holds promise for a wide range of clinical disorders.

ABC Transporters

The ABC transporters are ubiquitous membrane proteins that couple adenosine triphosphate (ATP) hydrolysis to the translocation of diverse substrates across cell membranes. Clinically relevant examples are associated with cystic fibrosis and with multidrug resistance of pathogenic bacteria and cancer cells. The crystal structure at 3.2 Å resolution of the *Escherichia coli* BtuCD protein, an ABC transporter mediating vitamin B12 uptake has been determined (PDB accession code 1L7V [68]). The two ATP-binding cassettes (BtuD) are in close contact with each other, as are the two membrane-spanning subunits (BtuC). This arrangement is distinct from that observed for the *E. coli* lipid flippase MsbA (PDB accession code 1JSQ [69])), another membrane

protein that was crystallised. The BtuC subunits provide 20 transmembrane helices grouped around a translocation pathway that is closed to the cytoplasm by a gate region whereas the dimer arrangement of the BtuD subunits resembles the ATP-bound form of the Rad50 DNA repair enzyme. A prominent cytoplasmic loop of BtuC forms the contact region with the ATP-binding cassette and appears to represent a conserved motif among the ABC transporters.

The nature of the mechanical changes that occur at each step of the chemical ATPase cycle are not fully understood. Crystal structures were determined of the MJ1267 ABC from *Methanococcus jannaschii* in Mg-ADP-bound and nucleotide-free forms [70]. Comparison of these structures reveals an induced-fit effect at the active site likely to be a consequence of nucleotide binding. These changes affect the region believed to mediate intercassette interaction in the ABC transporter complex.

Note added during revision process. Recently, the 2.8 angstrom structure of the integral membrane protein Fatty Acid Amide Hydrolase (FAAH) of rat has been solved (PDB accession code 1MT5 [71]). This enzyme degrades members of the endocannabinoid class of signaling lipids and terminates their activity. The structure of FAAH complexed with an arachidonyl inhibitor reveals how a set of discrete structural alterations allows this enzyme, in contrast to soluble hydrolases of the same family, to integrate into cell membranes and establish direct access to the bilayer from its active site.

REFERENCES

[1] Blundell, T.; Jhoti, H.; Abell, C. *Nature Reviews Drug Discovery*, **2002**, *1*, 45.

[2] Berman, H.M.; Battistuz, T.; Bhat, T. N.; Bluhm, W. F.; Bourne, P. E.; Burkhardt, K.; Feng, Z.; Gilliland, G. L.; Iype, L.; Jain, S.; Fagan, P.; Marvin, J.; Padilla, D.; Ravichandran, V.; Schneider, B.; Thanki, N.; Weissig, H.; Westbrook, J. D. ; Zardecki C. *Acta Cryst.,* **2002**, *D58*, 899. http://www.rcsb.org/pdb/

[3] Michel, H. Crystallization of membrane proteins in International Tables for Crystallography; Rossmann M.; Arnold E. Eds.; Kluwer Academic Publishers: London, **2001**, Vol. F, pp. 94-99.

[4] White, S.H.; Wimley, W.C. *Annu. Rev. Biophys. Biomol. Struct.,* **1999**, *28*, 319.

[5] http://blanco.biomol.uci.edu/membrane_proteins_xtal.html.

[6] http://www.mpibp-frankfurt.mpg.de/michel/public/memprotstruct.html

[7] Wallin, E.; von Heijne, G. *Protein Sci.,* **1998**, *7*, 1029

[8] Gerstei, M. *Proteins,* **1998**, *33*, 518.

[9] Chandrasekharan, N.V.; Dai, H.; Roos, K.L.; Evanson, N.K.; Tomsik, J.; Elton, T.S.; Simmons, D.L. *Proc. Natl. Acad. Sci. USA,* **2002,** *99*, 13926.

[10] Garavito, R.M.; Malkowski, M.G.; Dewitt, D.L. *Prostaglandins & other Lipid Mediators,* **2002,** *68*, 129.

[11] Picot, D; Loll, P.J; Garavito, R.M. *Nature,* **1994,** *367*, 243.

[12] Loll, P.J; Picot, D; Garavito, R.M. *Nat Struct Biol.,* **1995,** *2*, 605.

[13] Kurumbail, R.G; Stevens, A.M; Gierse, J.K; McDonald, J.J; Stegeman, R.A; Pak, J.Y; Gildehaus, D; Miyashiro, J.M; Penning, T.D; Seibert, K; Isakson, P.C; Stallings, W.C. *Nature,* **1996,** *384*, 644.

[14] Malkowski, M.G.; Ginell, S.L.; Smith, W.L.; Garavito, R.M. *Science,* **2000,** *289*, 1933.

[15] Thuresson, E.D.; Malkowski, M.G.; Lakkides, K.M.; Rieke, C.J.; Mulichak, A.M.; Ginell, S.L.; Garavito, R.M.; Smith, W.L. *J. Biol. Chem.,* **2001,** *276*, 10358.

[16] Malkowski, M.G.; Thuresson, E.D.; Lakkides, K.M.; Rieke, C.,J.; Micielli, R.; Smith, W.L.; Garavito, R.M. *J. Biol. Chem*., **2001,** *276*, 37547.

[17] Kalgutkar, A.S.; Crews, B.C.; Rowlinson, S.W.; Marnett, A.B.; Kozak, K.R.; Remmel, R.P.; Marnett, L.J. *Proc. Natl. Acad. Sci. USA,* **2000,** *97*, 925.

[18] Palomer, A.; Cabre, F.; Pascual; J.; Campos, J.; Trujillo, M.A.; Entrena, A.; Gallo, M.A.; Garcia, L.; Mauleon, D.; Espinosa, A. *J. Med. Chem.*, **2002,** *45*, 1402.

[19] Kalgutkar, A.S.; Castagnoli, N. Jr. *Med. Res. Rev.,* **1995**, *15*, 325.

[20] Wouters, J. *Curr. Med. Chem.,* **1998**, *5*, 137.

[21] Binda, C.; Newton-Vinson, P.; Hubalek, F.; Edmondson, D.E.; Mattevi, A. *Nat. Struct. Biol.,* **2002,** *9*, 22.

[22] Wouters, J.; Baudoux, G. *Proteins,* **1998**, *32*, 97.

[23] Geha, R.M.; Chen, K.; Wouters, J.; Ooms, F.; Shih, J.C. *J. Biol. Chem.,* **2002,** *277*, 17209.

[24] Grünler, J.; Ericsson, J.; Dallner, G. *Biochim. Biophys. Acta,* **1994,** *1212*, 259

[25] Qin, W.; Infante, J.; Wang, S.R.; Infante, R. *Biochim. Biophys. Acta,* **1992,** *1127*, 57.

[26] Morand, O.; Aebi, J.D.; Dehmlow, H.; Ji, YH.; Gains, N.; Lengsfeld, H.; Himber, J. *J. Lipid Res.,* **1997,** *38*, 373.

[27] Mark, M.; Müller, P.; Maier R.; Eisele, B. *J. Lipid Res.,* **1996,** *37*, 148

[28] Eisele, B.; Budzinski, R.; Müller, P.; Maier, R.; M. Mark, M. *J. Lipid Res.,* **1997,** *38*, 564.

[29] Brown G.R.; Hollinshead, D.M.; Stokes, E.S.; Clarke, D.S.; Eakin, M.A.; Foubister, A.J.; Glossop, S.C.; Griffiths, D.; Johnson, M.C.; McTaggart, F.; Mirrlees, D.J.; Smith, G.J.; Wood, R. *J. Med. Chem.,* **1999,** *42*, 1306.

[30] Wendt, K.U.; Poralla, K.; Schulz, G.E. *Science,* **1997,** *277,* 1811.

[31] Zacharias, N.; Dougherty, D. *Trends in Pharmaceutical Sciences,* **2002,** *23,* 281.

[32] Lenhart, A.; Weihofen, W.A.; Pleschke, A.E.; Schulz, G.E. *Chem. Biol.,* **2002,** *9*, 639.

[33] Teller, D.C.; Okada, T.; Behnke, C.A.; Palczewski, K.; Stenkamp, R.E. *Biochemistry*, **2001**, *40*, 7761; International Human genome Sequencing Consortium. *Nature,* **2001,** *409,* 860.

[34] Sakmar, T.P. *Curr. Opin. Cell Biol.,* **2002,** *14*, 189.

[35] Sakmar, T.P.; Menon, S.T.; Marin, E.P.; Awad, E.S. *Annu. Rev. Biophys. Biomol. Struct.,* **2002,** *31*, 443.

[36] Santosh, T.; Menon, S.T.; Sakmar, T. *Physiological Reviews,* **2001,** *81*, 1659.

[37] Elaine, C.; Bourne H. Trends in Pharmacological Sciences, **2001,** 22, 587.

[38] Palczewski, K.; Kumasaka, T.; Hori, T.; Behnke, C.A.; Motoshima, H.; Fox, B.A.; Le Trong, I.; Teller, D.C.; Okada, T.; Stenkamp, R.E; Yamamoto, M.; Miyano, M. *Science,* **2000,** *289*, 739

[39] Lu, Z.L.; Saldanha, J.; Hulme, H. *Trends in Pharmaceutical Sciences,* **2001**, *23* 140.

[40] Vaidehi, N.; Floriano, W.B.; Trabanino, R.; Hall, S.E.; Freddolino, P.; Choi, E.J.; Zamanakos, G.; Goddard, W.A. 3rd. *Proc. Natl. Acad. Sci. USA,* **2002**, *99*, 12622.

[41] http://archive.bmn.com/supp/tips/tips2211a.html.

[42] Schertler, G.; Hargrave, P.A. *Proc. Natl. Acad. Sci. USA,* **1995** *92*, 11578.

[43] Unger, V.M.; Hargrave, P.A.; Baldwin, J.M.; Schertler, G.F. *Nature,* **1997,** *389*, 203.

[44] Luecke, H.; Schobert, B.; Richter, H.T.; Cartailler, J.P.; Lanyi, J. *J. Mol. Biol.,* **1999,** *291*, 899.

[45] Pebay-Peyroula, E.; Rummel, G.; Rosenbusch, J.P.; Landau, E.M. *Science,* **1997,** *277*, 1676.
[46] Okada, T.; Fujiyoshi, Y.; Silow, M.; Navarro, J.; Landau, E.M.; Shichida, Y. *Proc. Natl. Acad. Sci. USA,* **2002**, *99*, 5982.
[47] Vaidehi, N.; Floriano, W.B.; Trabanino, R.; Hall, S.E.; Freddolino, P.; Choi, E.J.; Zamanakos, G.; Goddard, W.A. *Proc. Natl. Acad. Sci. USA,* **2002**, *99*, 12622.
[48] Cai, K.; Itoh, Y.; Khorana, H. G. *Proc. Natl. Acad. Sci. USA,* **2001**, *98*, 4877.
[49] Itoh, Y.; Cai, K.; Khorana, H. G. *Proc. Natl. Acad. Sci. USA,* **2001**, *98*, 4883.
[50] Loewen, M.C.; Klein-Seetharaman, J.; Getmanova, E.V.; Reeves, P.J.; Schwalbe, H.; Khorana, H.G. *Proc. Natl. Acad. Sci. USA,* **2001**, *98*, 4888.
[51] Delcour, A. J. Mol. Microbiol. Biotechnol., **2002,** 4, 1.
[52] Menestrina, G.; Serra, M.D.; Prevost, G. *Toxicon,* **2001,** *39*, 1661.
[53] Song, L.; Hobaugh, M.R.; Shustak, C.; Cheley, S.; Bayley, H.; Gouaux, J.E. *Science,* **1996,** *274*, 1859.
[54] Olson, R.; Nariya, H.; Yokota, K.; Kamio, Y.; Gouaux, E. *Nat. Struct. Biol.,* **1999,** *6*, 134.
[55] Cheley, S.; Gu, L.Q.; Bayley, H. *Chem. Biol.,* **2002,** *9*, 829
[56] Koronakis, V.; Sharff, A.; Koronakis, E.; Luisi, B.; Hughes, C. *Nature,* **2000,** *405*, 914.
[57] Doyle, D.A., Morais Cabral, J., Pfuetzner, R.A., Kuo, A., Gulbis, J.M., Cohen, S.L., Chait, B.T., MacKinnon, R. *Science,* **1998,** *280*, 69.
[58] Jiang, Y.; Lee, A.; Chen, J.; Cadene, M.; Chait, B.T.; MacKinnon, R. *Nature,* **2002**, *417*, 515.
[59] Chang, G.; Spencer, R.H.; Lee, A.T.; Barclay, M.T.; Rees, D.C. *Science,* **1998,** *282*, 2220.
[60] Dutzler, R. Campbell, E.B.; Cadene, M.; Chait, B.T.; MacKinnon, R. *Nature,* **2002,** *415*, 287.
[61] Sui, H.; Han, B.G.; Lee, J.K.; Walian, P.; Jap, B.K. *Nature,* **2001,** *414*, 872.
[62] Fu, D.; Libson, A.; Miercke, L.J.; Weitzman, C.; Nollert, P.; Krucinski, J.; Stroud, R.M. *Science,* **2000,** *290*, 481.
[63] Hanlon, M.R.; Wallace, B.A. *Biochemistry,* **2002,** *41*, 2886.
[64] Ford, J.W.; Stevens, E.B.; Treherne, J.M.; Packer, J.; Bushfield, M. *Prog. Drug Res.,* **2002,** *58,* 133.
[65] Chung, S.H.; Kuyucak, S. *Eur. Biophys. J.,* **2002,** *31*, 283.
[66] Spencer, R.H.; Rees, D.C. *Annu. Rev. Biophys. Biomol. Struct.,* **2002,** *31*, 207.
[67] Kozono, D.; Yasui, M.; King, L.S.; Agre, P. *J. Clin. Invest.,* **2002,** *109*, 1395.
[68] Locher, K.P.; Lee, A.T.; Rees, D.C. *Science,* **2002,** *296*, 1091.
[69] Chang, G.; Roth, C.B. *Science,* **2001,** *293*, 1793.
[70] Karpowich, N.; Martsinkevich, O.; Millen, L.; Yuan, Y.R.; Dai, P.L.; MacVey, K.; Thomas, P.J.; Hunt, J.F. *Structure*, **2001,** *9*, 571.
[71] Bracey, M.H.; Hanson, M.A.; Masuda, K.R.; Stevens, R.C.; Cravatt, B.F. *Science*, **2002**, *298*, 1793.

Copyright © 2003 EBSCO Publishing