NATURE|Vol 459|21 May 2009|doi:10.1038/nature08142

Biophysical dissection of membrane proteins

Stephen H. White[1]

The first atomic-resolution structure of a membrane protein was solved in 1985. Twenty-four years and more than 180 unique structures later, what have we have learned? An examination of the atomic details of several diverse membrane proteins reveals some remarkable biophysical features and suggests that we can expect to achieve much more in the decades to come.

Two events define the beginning of the modern era of membrane-protein biophysics: the determination of the three-dimensional structure of bacteriorhodopsin[1] at low resolution by Richard Henderson and Nigel Unwin in 1975, and the atomic-resolution structure of the *Rhodopseudomonas viridis* photosynthetic reaction centre[2] by Johann Deisenhofer and Hartmut Michel in 1985. Although spectroscopic studies[3,4] of isolated membranes had suggested in the mid-1960s that the proteins of plasma membranes are largely α-helical, Henderson and Unwin's 1975 structure[1] established unequivocally the existence of transmembrane α-helices in a membrane protein. Deisenhofer and Michel's structure confirmed the presence of transmembrane α-helices and allowed an atomic-level interpretation of biophysical data for the first time. As important as the structure was on its own, the more important accomplishment may simply have been to demonstrate that membrane proteins could be crystallized, thereby opening the way for atomic-resolution X-ray crystallography.

As other groups picked up the challenge, progress accelerated (Fig. 1a) such that now, 24 years later, we have high-resolution structures for more than 180 unique proteins (Fig. 1b). What have we learned about membrane-protein biophysics as a result? In this Overview, I extract some general features of membrane-protein biophysics from the five reviews that follow in this Insight, which dissect the operation of several membrane proteins at the atomic level. The clearest feature is that interactions with the lipid bilayer are important, but no general principles about these have yet emerged. Symmetry, which arises from gene duplications, is seen to be a building block of transporter function. In addition, we find that water can penetrate deeply into membrane proteins, despite the non-polar character of the bilayer. As expected from their biosynthetic pathways, all plasma-membrane proteins are α-helical bundles, but there are wide variations in geometry, related in part to water penetration and interactions with the bilayer. Seven-helix bundles, which are hallmarks of G-protein-coupled receptors, seem ideally minimalist and are extremely versatile signalling platforms.

Transporter structural motifs

Sodium-coupled transporters, which are reviewed by Eric Gouaux (see page 347), are an important class of membrane protein, exhibiting pseudo-two-fold symmetry and deep water penetration. These so-called secondary transporters are 'couplers' because they couple the energetic 'uphill' movement of one solute to the 'downhill' movement of another solute. For sodium-coupled transporters, the energy gained from the movement of sodium ions down an electrochemical gradient is used to concentrate substrates, such as aspartate[5] and leucine[6], on one side of the membrane. These transporters, which take advantage of the two-fold structural symmetry, work as 'rocker switches' to provide alternating access to the two sides of the membrane[7]. As Gouaux points out, a broad range of transporters share these features despite having low sequence identity and differing numbers of transmembrane helices. For example, the vSGLT sodium/galactose symporter[8] has 14 transmembrane helices and the Mhp1 benzyl-hydantoin transporter[9] has 12 transmembrane helices, but the transporter function is carried out by pseudo-two-fold-related 5+5 transmembrane repeats.

Signalling and seven-helix bundles

Another class of couplers are the G-protein-coupled receptors (GPCRs). These seven-helix membrane proteins receive an optical or chemical signal on the extracellular membrane surface and initiate G-protein-based signalling cascades in the cytoplasm. GPCRs, which are categorized into six subclasses, form the largest single class of eukaryotic membrane proteins and are the targets of many of the drugs being developed by pharmaceutical companies[10]. Because of their pharmaceutical importance, there have been extensive efforts — dating back to 1992 at least[11] — to build molecular models of ligand-activated GPCRs to guide drug discovery. The first models were based on bacteriorhodopsin, which has seven transmembrane helices and a covalently linked retinal whose photoexcitation drives proton transport. But when the structure of bovine rhodopsin, a true GPCR, was determined[12], modellers turned their attentions to that, with little success[13]. Finally, in 2007, the first structures of ligand-activated GPCRs became available, starting with human β_2 adrenergic receptor co-crystallized with Fab5 antibodies[14]. On page 356, Daniel Rosenbaum, Søren Rasmussen and Brian Kobilka review the structural and mechanistic insights into GPCR function gained from these structures, which include higher-resolution structures of β_2 adrenergic receptor[15,16], turkey β_1 adrenergic receptor[17], human A_{2A} adenosine receptor[18], opsin[19] and squid rhodopsin[20].

A recent mini-review provides a close structural comparison of GPCR structures[21]. The lesson that emerges is that the seven-transmembrane-helix bundle is extremely adaptable, and perhaps also minimalist. Why minimalist? Bacteriorhodopsin can be refolded in lipid vesicles from separated helices (A, B and CDEFG)[22]. Using retinal absorbance as a measure of folding, neither A+CDEFG nor B+CDEFG gives native spectra, whereas A+B+CDEFG does. This implies that seven is the minimum number of helices needed to maintain the native environment of retinal. It may be that seven helices provide ample space for a wide range of ligands through relatively minor helix distortions and reorientations, without the need to increase or decrease the number of transmembrane helices.

[1]Department of Physiology and Biophysics, and Center for Biomembrane Systems, University of California, Irvine, California 92697, USA.

Helix hairpins

A physicist friend once opined that the concept sketches, such as rocker switches, drawn by biologists must be hopelessly naive. Surely the real mechanisms must be far more subtle, drawing on sophisticated physics not yet revealed? Well, membrane protein machines really do function largely according to the principles embodied in the sketches (which, of course, summarize years of painstaking work, a fact overlooked by my friend). This is certainly true of F_OF_1 ATPase (or simply ATP synthase). Paul Boyer introduced in the 1980s the concept of rotary catalysis[23] by F_1 ATPase; John Walker and colleagues[24] provided the nearly complete structural details of rotary catalysis by solving the structure of the ATP synthase at atomic resolution.

ATP synthase generates ATP from proton electrochemical gradients in mitochondria. The ATP-producing F_1 domain is a rotor with a shaft that passes through the membrane-embedded F_O sector, which is composed of ten transmembrane two-helix c-subunits. The multiple two-helix-hairpin motif is apparently ideal for flexibly stabilizing large complexes with circular symmetry, such as light-harvesting complexes[25,26]. ATP synthases of different types can have up to 14 c-subunit hairpins[27]. Although the exact biochemical reason for variations in c-subunit number is not understood, structurally it seems to be an efficient way to expand a boundary of a membrane protein. Need a larger boundary? Just add a few more hairpins. Protons that pass through F_O on their way down the electrochemical gradient drive the F_1 rotor to produce ATP via the mechanical energy produced. Remarkably, the rotation of the rotor can be observed and quantified by attaching a fluorescent actin filament to the γ-subunit shaft[28]. Wolfgang Junge and his colleagues Hendrik Sielaff and Siegfried Engelbrecht describe the structure and operation of this remarkable machine in their review (see page 364). Junge provides excellent movies of the synthase in action as Supplementary Information. More movies and structural information can be found at www.biologie.uni-osnabrueck.de/biophysik/junge/picsmovs.html and www.mrc-dunn.cam.ac.uk/research/atp_synthase.

Bilayer distortion

Intramembrane-cleaving proteases (iCLIPs), first identified by Michael Brown, Joseph Goldstein and colleagues[29,30], irreversibly cleave transmembrane helices to release tethered signalling domains intra- or extracellularly. For example, site-2 protease (S2P) cleaves the cytoplasmic domain of the mammalian sterol regulatory element-binding protein (SREBP). The liberated domain then travels to the nucleus to activate genes involved in cholesterol and fatty-acid biosynthesis. An iCLIP of current interest is the catalytic subunit of γ-secretase, presenilin, which liberates the amyloid-β peptide. S2P and presenilin are examples of metallo- and aspartyl proteases, respectively. A third class of iCLIP comprises the rhomboid serine proteases, which in *Drosophila* liberate the epidermal growth factor domain by cleaving the single-transmembrane Spitz protein.

The conceptual difficulty with iCLIPs is that they must carry out hydrolytic reactions within the hydrophobic membrane interior. The mystery of how they achieve this is emerging from the recently published structures of S2P[31] and rhomboid proteases from the bacteria *Escherichia coli*[32–34] and *Haemophilus influenzae*[35], as discussed by Elinor Erez, Deborah Fass and Eitan Bibi in their comprehensive but succinct review of intramembrane proteases (see page 371). As with sodium-coupled transporters, there is deep penetration of water into the heart of the rhomboid, allowing peptide-bond hydrolysis by a catalytic dyad composed of serine and histidine. An unusual (perhaps even unique) feature of the six-helix rhomboids is the rather large L1 helical loop between helix 1 and helix 2 that protrudes parallel to the membrane plane with deep penetration into the bilayer interface. The exact function of this loop is uncertain, but it may be involved in the regulation of proteolysis. Helix 5 sits at the membrane entrance to the catalytic site, but is on the opposite side of the protein to L1. Molecular dynamic simulations[36] suggest that there is dynamic coupling between L1 and helix 5, consistent with a regulatory role for L1. The simulations also show that the irregular shape of the protein causes significant bilayer deformation in the vicinity of the active site, and that hydrogen-bond interactions with bilayer lipids affect protein orientation and dynamics. Distortion of the lipids in the vicinity of the active site may be necessary to admit both water and transmembrane-helix substrates.

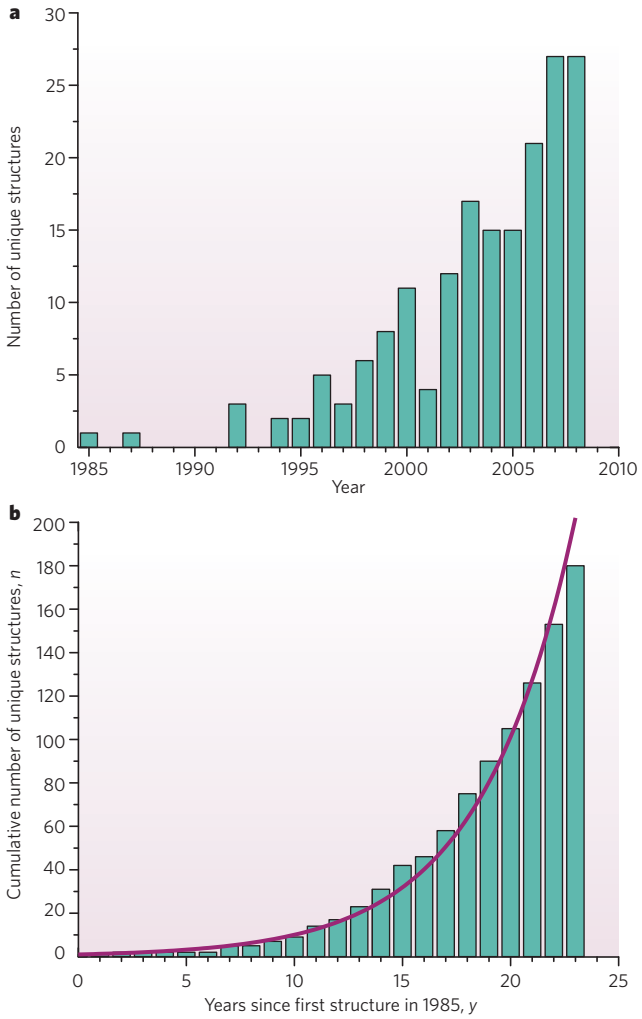

Figure 1 | Progress in determining membrane protein structures. Only unique structures are included in the statistics. Proteins of the same type from different species are included, but structures of mutagenized versions of proteins are excluded, as are proteins that differ only in terms of substrate bound or physiological state. **a**, The number of structures reported each year since 1985. **b**, The bars represent the cumulative number (n) of structures plotted against the number of years (y) since the first structure was reported. The solid curve is the best fit to the equation $n = \exp(ay)$, where $a = 0.23$; the reduced χ^2 of the fit is 0.6. Data are from a curated database of membrane proteins of known structure at http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html.

Restoring the bilayer to membrane proteins

The function of rhomboid proteases requires careful consideration of the effect of the lipid membrane itself, within which all membrane proteins must exist and function. The difficulty with membrane-protein crystallographic structures is that we never see the surrounding membrane, except to the extent that a few lipids or solubilizing detergent molecules may form part of a crystal's unit cell. A powerful approach to understanding a membrane protein in context is to restore the membrane and its lipids computationally, which can be done either by atomic-level molecular dynamics simulations[37] or by continuum mechanics. Rob Phillips, Tristan Ursell, Paul Wiggins and Pierre Sens describe the use of the latter approach (see page 379), which is especially useful for understanding the interplay between tension-gated ion channels and the membrane. It is also useful for describing the

interactions between proteins in the crowded membrane environment[38]. A fundamental question for the future is how to move seamlessly between continuum and atomic-level simulations.

Structural themes

Some unexpected structural themes have emerged from the X-ray structures of membrane proteins discussed here. The bacteriorhodopsin structure, finally solved to atomic resolution by electron crystallography[39] in 1996 and by X-rays[40] in 1997, revealed the transmembrane-helix bundle as the fundamental structural motif of plasma-membrane proteins (bacterial, mitochondrial and chloroplast outer-membrane proteins generally have a β-barrel motif[41,42]). The seven-helix bundle characteristic of bacteriorhodopsin is a remarkably versatile motif, as is apparent from G-protein-coupled receptors. Proteins involved in proton and electron transport form very tight 'waterproof' bundles, whereas transporters such as the *E. coli* lactose permease[43] and the mitochondrial ADP/ATP carrier[44] have large, water-filled cavities that extend almost completely across the membrane. Sodium-coupled transporters share this feature, albeit to a lesser extent. Also discovered in the first structure of an ion channel, the KcsA potassium channel[45], these water-filled cavities complicate the prediction of membrane-protein topology by hydropathy analysis.

Another theme to emerge is an internal structural repeat and inversion of the first half of the protein to form a helix bundle with a pseudo-two-fold symmetry about an axis parallel to the membrane plane. This pattern apparently arises from ancient gene duplications, as described in detail in a recent review[46]. In a few cases, such as the *E. coli* EmrE multidrug transporter[47], the two-fold symmetry may arise from two dual-topology monomers inserted with opposite topologies[48] to form the functional homodimer. Finally, from studies of voltage-sensor domains[49] that control voltage-dependent ion channels and enzymes, and which can act as ion channels themselves, we now know that charged amino acids, especially arginine, can be buried directly in the lipid bilayer as a result of salt bridges and lipid polar-group interactions.

The biophysical motifs discussed here are highlights from a rich menu of structures. The 180 unique structures available at the end of 2008 have revolutionized our understanding of membrane proteins. What does the future hold? What new motifs and folds will appear? James Bowie and colleagues[50] have estimated that about 1,700 membrane-protein structures are needed to account for each structural family. At the present pace and level of technology, that will take about 30 years. So stay tuned for the remarkable progress of the past 24 years to continue. ■

1. Henderson, R. & Unwin, P. N. T. Three-dimensional model of purple membrane obtained by electron microscopy. *Nature* **257,** 28–32 (1975).
2. Deisenhofer, J., Epp, O., Miki, K., Huber, R. & Michel, H. Structure of the protein subunits in the photosynthetic reaction centre of *Rhodopseudomonas viridis* at 3Å resolution. *Nature* **318,** 618–624 (1985).
3. Maddy, A. H. & Malcolm, B. R. Protein conformations in the plasma membrane. *Science* **150,** 1616–1618 (1965).
4. Lenard, J. & Singer, S. J. Protein conformation in cell membrane preparations as studied by optical rotatory dispersion and circular dichroism. *Proc. Natl Acad. Sci. USA* **56,** 1828–1835 (1966).
5. Yernool, D., Boudker, O., Jin, Y. & Gouaux, E. Structure of a glutamate transporter homologue from *Pyrococcus horikoshii*. *Nature* **431,** 811–818 (2004).
6. Yamashita, A., Singh, S. K., Kawate, T., Jin, Y. & Gouaux, E. Crystal structure of a bacterial homologue of Na^+/Cl^--dependent neurotransmitter transporter. *Nature* **437,** 203–205 (2005).
7. Mitchell, P. A general theory of membrane transport from studies of bacteria. *Nature* **180,** 134–136 (1957).
8. Faham, S. *et al.* The crystal structure of a sodium galactose transporter reveals mechanistic insights into Na^+/Sugar symport. *Science* **321,** 810–814 (2008).
9. Weyand, S. *et al.* Structure and molecular mechanism of a nucleobase-cation-symport-1 family transporter. *Science* **322,** 709–713 (2008).
10. Lagerström, M. C. & Schiöth, H. B. Structural diversity of G-protein-coupled receptors and significance for drug discovery. *Nature Rev. Drug Discov.* **7,** 339–357 (2008).
11. Pardo, L., Ballesteros, J. A., Osman, R. & Weinstein, H. On the use of the transmembrane domain of bacteriorhodopsin as a template for modeling the three-dimensional structure of guanine nucleotide-binding regulatory protein-coupled receptors. *Proc. Natl Acad. Sci. USA* **89,** 4009–4012 (1992).
12. Palczewski, K. *et al.* Crystal structure of rhodopsin: A G protein-coupled receptor. *Science* **289,** 739–745 (2000).
13. Oliveira, L., Hulsen, D., Hulsik, D. J., Paiva, A. C. M. & Vriend, G. Heavier-than-air flying machines are impossible. *FEBS Lett.* **564,** 269–273 (2004).
14. Rasmussen, S. G. F. *et al.* Crystal structure of the human β_2 adrenergic G-protein-coupled receptor. *Nature* **450,** 383–388 (2007).
15. Rosenbaum, D. M. *et al.* GPCR engineering yields high-resolution structural insights into β_2-adrenergic receptor function. *Science* **318,** 1266–1273 (2007).
16. Cherezov, V. *et al.* High-resolution crystal structure of an engineered human β_2-adrenergic G protein-coupled receptor. *Science* **318,** 1258–1265 (2007).
17. Warne, T. *et al.* Structure of a β-adrenergic G-protein-coupled receptor. *Nature* **454,** 486–492 (2008).
18. Jaakola, V.-P. *et al.* The 2.6 angstrom crystal structure of a human A_{2A} adenosine receptor bound to an antagonist. *Science* **322,** 1211–1217 (2008).
19. Park, J. H., Scheerer, P., Hofmann, K. P., Choe, H.-W. & Ernst, O. P. Crystal structure of the ligand-free G-protein-coupled receptor opsin. *Nature* **454,** 183–188 (2008).
20. Murakami, M. & Kouyama, T. Crystal structure of squid rhodopsin. *Nature* **453,** 363–368 (2008).
21. Hanson, M. A. & Stevens, R. C. Discovery of new GPCR biology: One receptor structure at a time. *Structure* **17,** 8–14 (2009).
22. Kahn, T. W. & Engelman, D. M. Bacteriorhodopsin can be refolded from two independently stable transmembrane helices and the complementary five-helix fragment. *Biochemistry* **31,** 6144–6151 (1992).
23. Boyer, P. D. The ATP synthase: A splendid molecular machine. *Annu. Rev. Biochem.* **66,** 717–749 (1997).
24. Stock, D., Leslie, A. G. W. & Walker, J. E. Molecular architecture of the rotary motor in ATP synthase. *Science* **286,** 1700–1705 (1999).
25. Koepke, J., Hu, X. C., Muenke, C., Schulten, K. & Michel, H. The crystal structure of the light-harvesting complex II (B800- 850) from *Rhodospirillum molischianum*. *Structure* **4,** 581–597 (1996).
26. McDermott, G. *et al.* Crystal structure of an integral membrane light-harvesting complex from photosynthetic bacteria. *Nature* **374,** 517–521 (1995).
27. Weber, J. & Senior, A. E. ATP synthesis driven by proton transport in F_1F_O-ATP synthase. *FEBS Lett.* **545,** 61–70 (2003).
28. Noji, H., Yasuda, R., Yoshida, M. & Kinosita, K. Jr Direct observation of the rotation of F_1-ATPase. *Nature* **386,** 299–302 (1997).
29. Sakai, J. *et al.* Sterol-regulated release of SREBP-2 from cell membranes requires two sequential cleavages, one within a transmembrane segment. *Cell* **85,** 1037–1046 (1996).
30. Rawson, R. B. *et al.* Complementation cloning of S2P, a gene encoding a putative metalloprotease required for intramembrane cleavage of SREBPs. *Mol. Cell* **1,** 47–57 (1997).
31. Feng, L. *et al.* Structure of a site-2 protease family intramembrane metalloprotease. *Science* **318,** 1608–1612 (2007).
32. Wang, Y., Zhang, Y. & Ha, Y. Crystal structure of a rhomboid family intramembrane protease. *Nature* **444,** 179–183 (2006).
33. Wu, Z. *et al.* Structural analysis of a rhomboid family intramembrane protease reveals a gating mechanism for substrate entry. *Nature Struct. Mol. Biol.* **13,** 1084–1091 (2006).
34. Ben-Shem, A., Fass, D. & Bibi, E. Structural basis for intramembrane proteolysis by rhomboid serine proteases. *Proc. Natl Acad. Sci. USA* **104,** 462–466 (2007).
35. Lemieux, M. J., Fischer, S. J., Cherney, M. M., Bateman, K. S. & James, M. N. G. The crystal structure of the rhomboid peptidase from *Haemophilus influenzae* provides insight into intramembrane proteolysis. *Proc. Natl Acad. Sci. USA* **104,** 750–754 (2007).
36. Bondar, A.-N., del Val, C. & White, S. H. Rhomboid protease dynamics and lipid interactions. *Structure* **17,** 395–405 (2009).
37. Lindahl, E. & Sansom, M. S. P. Membrane proteins: molecular dynamics simulations. *Curr. Opin. Struct. Biol.* **18,** 425–431 (2008).
38. Engelman, D. M. Membranes are more mosaic than fluid. *Nature* **438,** 578–580 (2005).
39. Grigorieff, N., Ceska, T. A., Downing, K. H., Baldwin, J. M. & Henderson, R. Electron-crystallographic refinement of the structure of bacteriorhodopsin. *J. Mol. Biol.* **259,** 393–421 (1996).
40. Pebay-Peyroula, E., Rummel, G., Rosenbusch, J. P. & Landau, E. M. X-ray structure of bacteriorhodopsin at 2.5 Å from microcrystals grown in lipidic cubic phases. *Science* **277,** 1676–1681 (1997).
41. Buchanan, S. K. β-Barrel proteins from bacterial outer membranes: Structure, function and refolding. *Curr. Opin. Struct. Biol.* **9,** 455–461 (1999).
42. Schulz, G. E. β-Barrel membrane proteins. *Curr. Opin. Struct. Biol.* **10,** 443–447 (2000).
43. Abramson, J. *et al.* Structure and mechanism of the lactose permease of *Escherichia coli*. *Science* **301,** 610–615 (2003).
44. Pebay-Peyroula, E. *et al.* Structure of mitochondrial ADP/ATP carrier in complex with carboxyatractyloside. *Nature* **426,** 39–44 (2003).
45. Doyle, D. A. *et al.* The structure of the potassium channel: Molecular basis of K^+ conduction and selectivity. *Science* **280,** 69–77 (1998).
46. von Heijne, G. Membrane-protein topology. *Nature Rev. Mol. Cell Biol.* **7,** 909–918 (2006).
47. Schuldiner, S. EmrE, a model for studying evolution and mechanism of ion-coupled transporters. *Biochim. Biophys. Acta* **1794,** 748–762 (2009).
48. Rapp, M., Seppälä, S., Granseth, E. & von Heijne, G. Emulating membrane protein evolution by rational design. *Science* **315,** 1282–1284 (2007).
49. Swartz, K. J. Sensing voltage across lipid membranes. *Nature* **456,** 891–897 (2008).
50. Oberai, A., Ihm, Y., Kim, S. & Bowie, J. U. A limited universe of membrane protein families and folds. *Protein Sci.* **15,** 1723–1734 (2006).

Acknowledgements This work was supported in part by grants from the National Institute of General Medical Science and the National Institute of Neurological Disorders and Stroke.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The author declares no competing financial interests. Correspondence should be addressed to the author (stephen.white@uci.edu).