Previously published in Biophysics Textbook Online (BTOL)

Chapter 1 Section 2

MEMBRANE PROTEIN STRUCTURE

Suzanne Scarlata

Overview

Membrane proteins reside in a highly asymmetric environment compared to their cytosolic counterparts which gives them unique properties. Rather than being in a homogeneous aqueous solvent, the "solvent" that surrounds membrane proteins consists of the aqueous region outside the membrane, the ionic region of the membrane surface and the oily region of the membrane interior. These differing dielectric environments impose certain structural features to membrane proteins that we will discuss. While the lipid environment also regulates the dynamic nature of the proteins, we will not discuss them here, but refer the reader to other sources (Yeagle, 1992; Gennis, 1989).

Along with the general features of the lipid properties mentioned above, cell membranes contain a variety of lipids and non-lipid components which "fine-tune" the physical properties of the particular membrane giving a range of environments that surround membrane proteins. Membrane composition varies from cell to cell and membrane composition also varies within cells. The outer membranes of cells isolate them from their environment and the proteins in them allow the cell to respond to external stimuli and regulate the entry of aqueous signals and nutrients. The internal membranes of cells compartmentalize organelles and the proteins in them allow these compartments to communicate with other regions of the cell. Since membranes are "two dimensional" surfaces, they will localize and concentrate cellular components that control the nature and directionality of cell signals. Although they play such critical roles in cell function, our knowledge of the structure of membrane proteins is limited due to the difficulty in solubilizing and crystallizing membrane proteins. Here, we will concentrate on the factors that govern membrane protein structure.

This section is divided into three parts. In section 1, we will review the key features of membranes that are necessary to understand membrane protein structure. In section 2 we will discuss the structural characteristics of membrane proteins. In section 3 we will review some examples of some membrane proteins whose structures are known.

2.1 Membrane Structure and Properties

2.1a *Natural membranes are two-dimensional fluids.* About ~25 years ago, biological membranes were described as being two dimensional fluids composed of two "leaflets" which contain mostly lipid molecules (Singer, 1972). This "fluid mosaic" model still best describes lipid membranes. In this model, the outer surfaces of the leaflet are composed primarily of the ionic and polar head groups and interact with the aqueous solution, whereas the inner portion of the membrane is composed of the hydrocarbon chains of the lipid (Figure 1). The chains are aligned parallel to each other, and under most biological conditions, have some flexibility that allows for rotation around the methylene groups. The nonpolar ends of the chains contact each other in the middle of the bilayer creating an oily barrier impermeable to

most aqueous soluble molecules but allow small, more non-polar agents such as ethanol, to pass. Integral membrane proteins and lipid molecules, which are confined to the bilayer, can diffuse laterally. *The most important feature of the fluid mosaic model is that it treats the membrane as a dynamic system in which the proteins and lipids can move and interact.* While this model has been refined in recent years to include nonhomogeneous distribution of components (Israelachvili, 1978) and the effect of cytoskeletal elements (Sackmann, 1995), it remains conceptually the best model to envision the environment of membrane proteins.

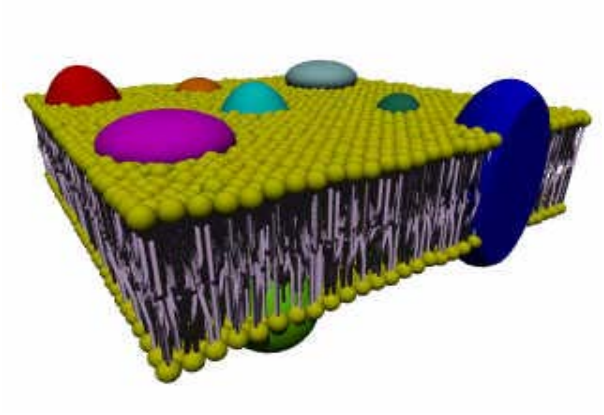

Figure1: Cartoon of a lipid membrane based on the Fluid Mosaic Model as proposed by Singer and Nicolson (1972). The key feature of the model is that both peripheral and integral proteins are free to diffuse (Figure by Paxton Provitera).

Just as the aqueous portions of cells have many components, so do membranes. Besides proteins, the major components of natural membranes are lipids and cholesterol, with a small amount of other materials such as fatty acids. The relative amounts of these species varies from membrane to membrane and even between membrane leaflets. The structures of lipids, cholesterol and fatty acids are shown in Figure 2. Because they impinge so much on membrane protein structure, we will briefly describe them. The student wishing a more comprehensive discussion of membrane components should see the following excellent references (Gennis, 1989; Petty, 1993; Yeagle, 1992).

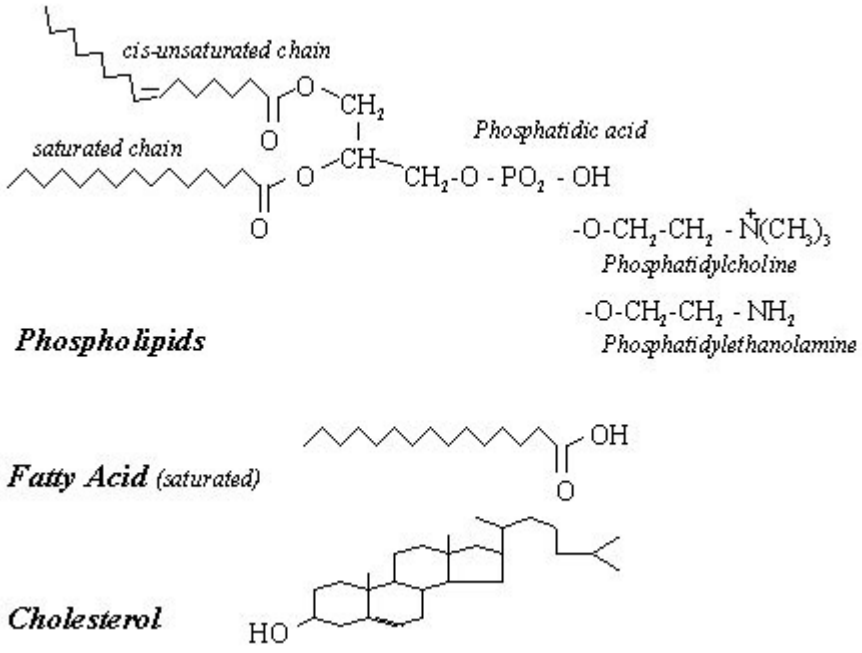

Figure 2: Structures of the major non-protein membrane components.

2.1b *Lipids* - Lipid molecules are characterized by two hydrocarbon chains and an ionic head group. Most lipids have a glycerol backbone with the hydrocarbon chains attached to the first two carbons and a phosphate group on the third (Figure 2). Other less common lipids are sphyingo-based or ceremide-based.

Phosphatidylcholine (PC) is one of the most biologically prevalent lipid head group and is commonly used in biophysical studies. PC is zwitterionic having a negative charge on the phosphate group and a positive charge on the amine. Phosphaditic acid (PA), phosphatidylglycerol (PG) and phosphatidylserine (PS) are common negatively charged lipid head groups. Some head groups (PE) are smaller and are less hydrated which causes stress on the bilayer surface since the hydrocarbon chains are more exposed. Some lipid head groups on the outer leaflet of plasma membranes are modified with large carbohydrate moieties that act as signatures allowing cells to recognize each other. *In terms of membrane protein structure, it is important to keep in mind that these head groups can participate in strong electrostatic and hydrogen-bonding interactions with the interfacial residues of membrane proteins, and their charge and hydration can directly influence the activity of peripheral membrane proteins.*

The hydrocarbon chains of lipids are usually 14-24 carbons long and, due to the mechanism of their synthesis, always have an even number of carbons. The chains tend to align parallel to each other stabilized by extended dispersion forces. Rotation can occur around the C-C bonds along the chains giving either a *trans* or *gauche* configuration (see Figure 3). The *trans* configuration allows the chains to come into close proximity of each other whereas the *gauche* configuration introduces kinks in the chain increasing the distances between the kinked chain and its neighbor. The closeness of the lipid chains or their "packing" dictates many of the physical properties of the bilayer such as the lateral movement of the lipid chains (see Yeagle, 1992), the permeability of the membrane to aqueous species and their gel to liquid crystal phase transition. In the gel phase, the lipid chains are usually well-aligned with little rotation around the C-C bonds, which are predominantly in the *trans* position (see Figure 2). This

reduced number of *gauche* isomers leads to better chain packing that in turn results in an increase in bilayer thickness and a reduction in the rate of lateral diffusion of membrane components. In contrast, the fluid phase is characterized by relatively rapid diffusion of membrane components and lower bilayer thickness due to the larger number of gauche isomers on the hydrocarbon chains.

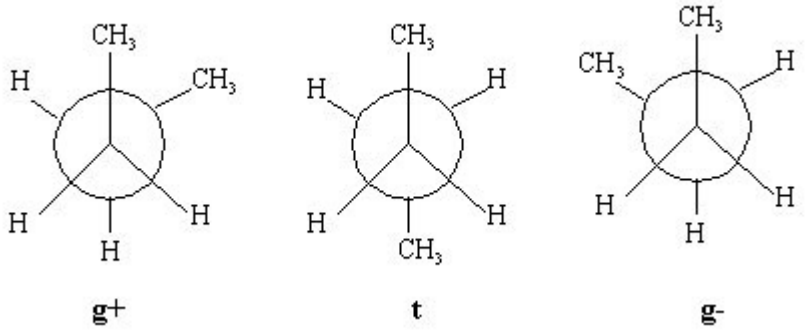

Figure 3: Newman projection diagram of the minimal energy *gauche* (g+ and g-) and *trans* conformations of butane.

Table 1 lists the Tm of different, commonly used lipids. These data show that longer chains have higher Tm values due to more extensive dispersion forces. Lipids with saturated or *trans*-unsaturated chains can pack more tightly as seen by higher Tm values. Lipids with *cis*-unsaturated chains, which are biologically more common, have lower Tm values because the *cis* configuration causes a substantial kink in the chain which increases the spacing between the lipid chains. Most double bonds occur further down the hydrocarbon chain keeping the center of the bilayer fluid. The usual number of double bonds found is 1-3. It should be kept in mind that the presence of a *cis* double bond shortens the length of the hydrocarbon chains and so bilayers composed of unsaturated chains will be thinner than those composed of saturated chains of the same length. The phase transition broadens as the number of components in the membrane increases and is barely detectable in many biological membranes.

TABLE 1

Gel to Liquid Crystal Phase Transition Temperature of Various Phospholipids (Silvius, 1982)

LIPID	Common Name	#Methylenes : # and position *cis* CH=CH	Tm (^{o}C)
dimyristoyl-phosphatidylcholine	DMPC	14:0 , 14:0	23
dipalmitoyl-phosphatidylcholine	DPPC	16:0 , 16:0	55

dioleoyl- phosphatidylcholine	DOPC	18:1(9) , 18:1 (9)	-20
dilinolenoyl- phosphatidylcholine	DLPC	18:2 (9,12), 18:2(9,12)	-53
1-palmitoyl-2-oleoyl- phosphatidylcholine	POPC	16:0 , 18:1 (9)	-2
dimyristoyl-phosphatidylserine	DMPS	14:0 , 14:0	35
dimyristoyl-phosphatidyl ethanolamine	DMPE	14:0 , 14:0	50
dimyristoyl-phosphatidic acid	DMPA	14:0 , 14:0	50

2.1c *Cholesterol*- Cholesterol is a fused ring structure with a single polar hydroxyl group which can interact with groups on or close to the membrane surface (Figure 2). Cholesterol has varying affects on membrane fluidity. In fluid phase, cholesterol tends to decrease the rotational freedom of the neighboring hydrocarbon chains and thus decreases the fluidity or "stiffens" the membrane. In gel phases, cholesterol acts as a contaminant which decreases the order of the well-packed lipid chains. Cholesterol can be found in high concentrations in some cell membranes (see Table 2).

2.1d *Fatty acids*- Fatty acids consist of a carboxylic group attached to a single14-24 CH_2 chain and are an important physiological end product of fat digestion (Figure 2). *A critical feature that distinguishes fatty acids from their corresponding two-chained lipid molecules is that they can freely partition into membranes, and can "flip-flop" or distribute evenly between the two leaflets of the membrane*. The latter mechanism occurs because the pK of carboxylic group is shifted from ~3 to ~ 7 when embedded in membranes giving the uncharged species some membrane solubility. Fatty acids and detergents, which are used in most methods to prepare purified membrane proteins, will act as carriers to membrane proteins to partition into the bilayer structure.

2.1e *Cells contain several different types of membranes which vary in composition*. As mentioned in section 2.1a, there is a substantial variability in the composition and physical properties of membranes in the cells. Listed in Table 2 is the composition of various membranes. Also consider that the lipid composition of the inner and outer leaflet of these membranes can also vary.

TABLE 2

Weight percent of membrane components in several types of membranes (adapted from Alberts et al., 1994)

Lipid	liver plasma membrane	erythrocyte plasma membrane	myelin	mitochondrian (inner and outer)	endoplasmic reticulum	*e. coli*
Cholesterol	17	23	22	3	6	0
POPE	7	18	15	35	17	70
POPS	4	7	9	2	5	trace
POPC	24	17	10	39	40	0
Sphingo-myelin	19	18	8	0	5	0
Glycolipids	7	3	28	trace	trace	0
Others	22	13	8	21	27	30

2.1f *The structure of membrane proteins is dictated by the length and magnitude of the dielectric gradient of the membrane bilayer thickness.*

Magnitude of the dielectric gradient. The feature that makes a lipid membrane a good barrier to aqueous species is the sharp dielectric gradient. The large change in dielectric constant between the ionic head groups to the oily interior is sharp (~80 to 2 Debeye) and occurs over a relatively short distance (see Fig. 4). This steep dielectric gradient makes it highly unfavorable to bury a charge (20 kCal/mol) or to leave an unsatisfied hydrogen bond (5 kCal/mol). In this way, the dielectric gradient will determine which amino acid side chains incorporate into the membrane

interior and which tend to remain in polar environments. Since the peptide backbone is composed of polar carbonyl and amino groups that hydrogen bond, burying a peptide backbone in the membrane interior will be energetically costly if these hydrogen bonds are not satisfied. As will be discussed, membrane proteins will adopt certain secondary and tertiary structures that can satisfy these hydrogen

bonds.

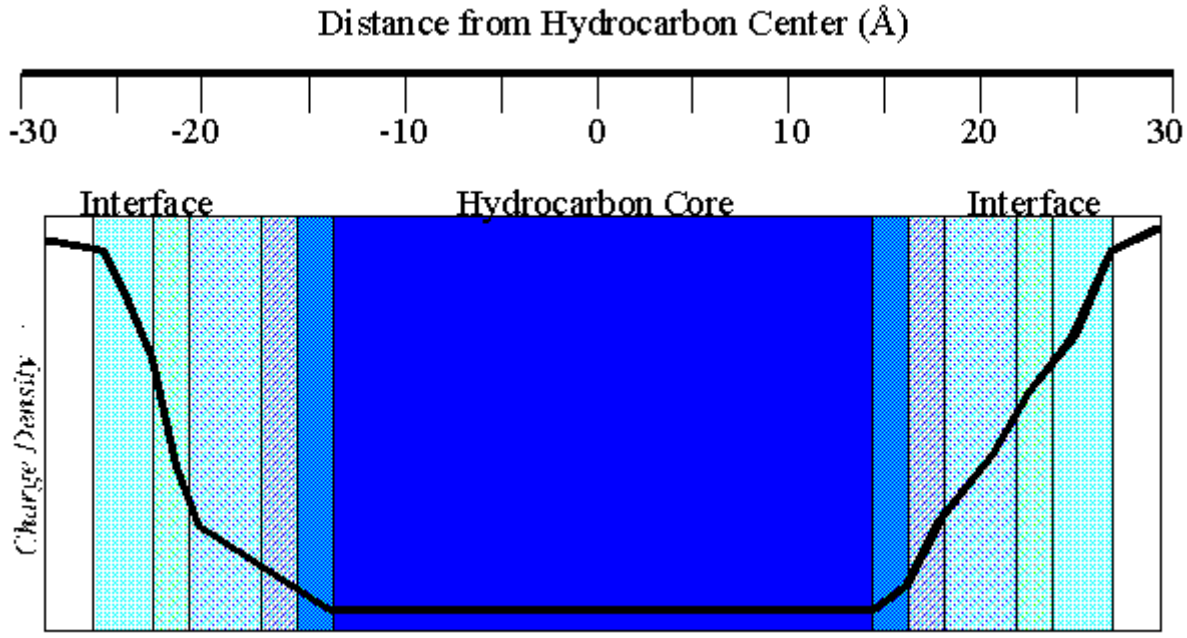

Figure 4: Redrawn with permission from Wimley and White *Biochim. Biophys. Acta* 1376: 339-352. (1998) showing the polarity gradient of a dioleoylphosphatidylcholine membrane (shades of blue) and the polarity profile (heavy line) obtained foom crystallographic data from the average charge density of the lipid groups. Further details of the definitions and methods used to calculate these parameters can be found in the original citation.

Bilayer Thickness. Bilayer thickness determines the length of the low dielectric well which in turn determines the regions of the protein that are interior, exterior and interfacial. Thus, the thickness of the bilayer may stabilize certain conformational states of the protein. The hydrophobic thickness of the membrane must closely match the hydrophobic length of the membrane protein. If the length is too long or too short, the protein may aggregate in an effort to minimize unfavorable interactions. Several models have been used to analyze the energetic cost of deforming a bilayer to accommodate a "mismatched" protein versus the hydrophobic interaction energy between the protein and lipid. Most often, the lipid molecules are treated as springs with a certain spring constant which allows the energy of deforming the lipid to be calculated (Mouritsen, 1984). Interestingly, it has been found that proteins will also deform to accommodate length mismatches (see Lee, 1998). These ideas are further discussed below.

Hydrocarbon Packing. The packing of the hydrocarbon chains may also stabilize certain membrane protein structures. Since the oily lipid chains find it energetically favorable to align with each other to maximize van der Waal interactions, they will prefer membrane components that do not greatly disrupt their interactions. In this way, membrane proteins can be thought of as contaminants. The end result is that proteins with a cylindrical shape will minimize the number of lipid chains that are disrupted by their presence and minimize the area of protein exposed to the bilayer.

2.2 Membrane Protein Structure

2.2a Membrane Protein Structure - basic principles

Figure 5 depicts the categories of membrane proteins. All membrane proteins have a specific upside-down or right-side-up orientation in the bilayer. Some proteins are anchored to the membrane by ionic interactions between residues with positively charged side chains and negatively charged lipid head groups since biological membranes tend to have a net negative charge. Other proteins are anchored by post-synthetic attachment of a hydrocarbon chain such as myristoyl, palmitoyl, farnesyl or gerenyl-gerenyl, or a lipid such as glycosylphosphatidylinositol (GPI) which confines them in regions close to their protein partners. Other proteins are anchored to the surface by ionic contacts. Since the structure of these types of membrane-anchored proteins follow the same rules as soluble, cytosolic proteins, they will only be mentioned briefly here.

The term monotopic or peripheral membrane protein refers to proteins that have a farily shallow penetration of the membrane surface. Many peripheral proteins can be released from the membrane by increasing the ionic strength of the solution. A second category of membrane proteins is integral or transmembrane bitopic or multitopic proteins. These proteins can only be released from the membrane by bilayer disruption with detergents. In the next section, we will discuss the primary, secondary and tertiary structure of bi- or multitopic membrane proteins. This discussion will be limited to the structural aspects of integral membrane proteins rather than the mechanisms through which these proteins insert, which generally involves a complex translocation process (see Simon, 1995). Before starting, it is important to keep in mind that the structure of a membrane protein will be a result of the interplay between three factors:

1- The need to match the dielectric properties of the side chains with the lipid.

2- The need to satisfy hydrogen bonds.

3- The need to align with the packing of the hydrocarbon chains.

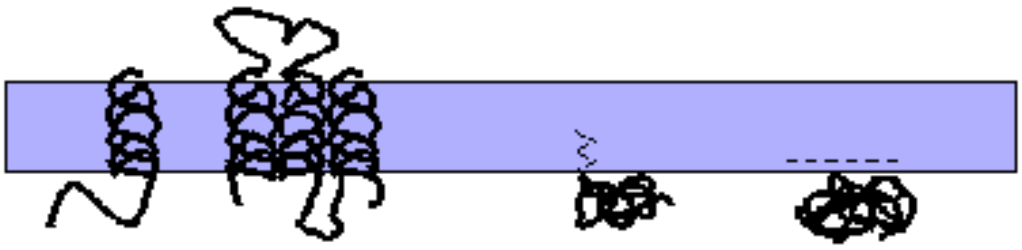

Figure 5: Cartoon of membrane proteins (see text).

Experimentally, we will concentrate on the structural information obtained by crystallographic and electron microscopy studies to narrow the scope of this chapter. However, other techniques, such as nuclear magnetic resonance, fluorescence resonance energy transfer, circular dichroism and protease

digestion have also provided useful and interesting structural information about membrane proteins.

2.2b Primary Structure of Transmembrane Proteins

Hydrophobicity Scales: Since the membrane interior is nonpolar, we would expect the surface residues of transmembrane proteins to be predominantly composed of nonpolar side chains. With this idea in mind, we can make predictions about the portions of the protein that may reside in the membrane interior. To aid in these predictions, scales have been formulated that give numerical values to the hydrophobic nature of amino acid side chains. However, this task is not as simple as it would first appear because although the nonpolar character of some side chains such as Leu is apparent, the hydrophobic nature of other side chains such as Ser and Gly are not. Several groups attacked the problem of assigning values of hydrophobicity to amino acids and many hydrophobicity scales have been generated (see Cornette, 1987). These scales are based on either solution studies, crystallographic data, or combinations of these two.

The solution-based hydrophobicity scales are generally based on free energy of transfer (ΔG_t) from aqueous solvent to a solvent that mimics the membrane, such as octanol. The molecule of interest (X) is added to a test tube containing an organic phase and the aqueous phase and its concentration in the two phases is measured using some standard technique. If the study is conducted at low amounts of X, then its concentration can be assumed to be equal to its chemical potential of X and can be written as

$$\Delta G_t = \mu^o_{H_2O} - \mu^o_{org} + RT \ln [X_{H_2O}] / [X_{org}]$$
$$\Delta G_t = \Delta G^o_t + RT \ln K_{eq}$$

and

$$\mu_{org} = \mu^o_{org} + RT \ln X_{org}$$
$$\mu_{H_2O} = \mu^o_{H_2O} + RT \ln X_{H_2O}$$

One problem with partitioning measurements is choosing the best solvent to mimic the membrane. A second problem is insolubility of some amino acid residues in one of the solvents. Early hydrophobicity scales have been improved by basing the scale on the difference between ΔG^o (transfer) of an amino acid with respect to a reference amino acid such as Gly in order to isolate the side chain contributions, and for residues in membrane-inserting or transmembrane peptides (see below).

Hydrophobicity values based on crystallographic methods have also been obtained. This approach surveys the solvent accessibility of amino acid residues in known crystallographic structures. Solvent accessibility is judged by rolling a water sphere of a specific radius, typically ~1.4 ’, and determining which residues can make van der Waals contact with the solvent.

Presently, both the solution-based and the crystallographic methods give similar values for most amino acids, but differ in residues whose side chains have both polar and non-polar character, such as Lys, Trp, Tyr, and Arg. To resolve these differences, recent hydrophobicity scales have evolved with improved methods to assign hydrophobicity values with the end goal of predicting membrane-spanning regions of proteins. The Kyte-Doolittle scale (Kyte, 1982) combines both crystallographic and solution-based

methods whereas the GES scale (Goldman, Engelman and Steitz, see Engelman, (1986) uses the solvent exposure of a residue on a 20 residue polyalanine helix. Another scale is based on the frequency that residues are found in membrane-spanning segments, e.g. (Rao, 1986). However, the interior of some ion channels and other transporters contain polar residues which may further blur features that distinguish the membrane spanning regions of unknown proteins, and also the interfacial region of some membrane proteins is not clearly defined. Recently, a scale has been formulated based on the free energy of transfer from water into a neutral, zwitterionic lipid membrane using tri- or penta-peptides where the amino acid of interest is located in the center of the peptide. To date, this is the only interfacial hydrophobicity scale that considers the energy contribution of all 20 amino acid side chains residues and peptide bonds. Table 3 lists the values of the free energy of transfer from water to POPC membrane interface and to n-octanol reported by Whimley and White (Wimely and White, 1996). The free energy value of transfer of the -CH_2-COHN glycyl unit, which has been added to n-octanol values, is 1.15 $\pm$ .11 kcal/mol and represents the free energy of burying a peptide bond (as discussed below). Thus, to avoid large energetic costs, interpeptide hydrogen bonds probably form before the insertion of the peptide into the membrane.

TABLE 3

Whole-Residue Hydrophobicity Scales from (White and Wimley, 1999)

Amino Acid	ΔG(interface) (kcal/mol)	ΔG(octanol) (kcal/mol)
Ala	0.17 $\pm$ 0.06	0.50 $\pm$ 0.12
Arg^+	0.81 $\pm$ 0.11	1.81 $\pm$ 0.13
Asn	0.42 $\pm$ 0.06	0.85 $\pm$ 0.12
Asp^-	1.23 $\pm$ 0.07	3.64 $\pm$ 0.17
Asp^0	-0.07 $\pm$ 0.11	0.43 $\pm$ 0.13
Cys	-0.24 $\pm$ 0.06	-0.02 $\pm$ 0.13
Gln	0.58 $\pm$ 0.08	0.77 $\pm$ 0.12
Glu^-	2.02 $\pm$ 0.11	3.63 $\pm$ 0.18
Glu^0	-0.01 $\pm$ 0.15	0.11 $\pm$ 0.12
Gly	0.01 $\pm$ 0.05	1.15 $\pm$ 0.11

His^+	0.96 ± 0.12	2.33 ± 0.11
His^0	0.17 ± 0.06	0.11 ± 0.06
Ile	-0.31 ± 0.06	-1.12 ± 0.11
Leu	-0.56 ± 0.04	-1.25 ± 0.11
Lys^+	0.99 ± 0.11	2.80 ± 0.11
Met	-0.23 ± 0.06	-0.67 ± 0.11
Phe	-1.13 ± 0.05	-1.71 ± 0.11
Pro	0.45 ± 0.12	0.14 ± 0.11
Ser	0.13 ± 0.08	0.46 ± 0.11
Thr	0.14 ± 0.06	0.25 ± 0.11
Trp	-1.85 ± 0.06	-2.09 ± 0.11
Tyr	-0.94 ± 0.06	-0.71 ± 0.11
Val	0.07 ± 0.05	-0.46 ± 0.11

2.2c Secondary Structure of Transmembrane Proteins

Since the membrane interior is essentially void of water, the only species that the atoms on the peptide backbone can hydrogen bond to are side chains or other backbone atoms. It is observed that membrane proteins prefer secondary structure elements where the peptide amide nitrogen and carbonyl oxygen hydrogen bond with each other in regular arrays in either α helices or β-sheets. The preference of hydrogen bonding of peptide backbone atoms to other backbone atoms can be understood when one considers that hydrogen bonds between the backbone and side chains in loop structures would limit the conformations available for the protein and thus be unfavorable in terms of configurational entropy. Also, random structures may leave unsatisfied hydrogen bonds and possibly disrupt the packing of the hydrocarbon chain.

Most transmembrane elements are α-helices. The two secondary structures that will satisfy all peptide backbone hydrogen bonds are β-sheets and α-helices. All bitopic membrane proteins whose structures have been solved cross the membrane as a helix. Most polytopic integral membrane proteins whose structures have been solved cross the membrane by helices. A few cross by β-sheets. Why the preference

for helices? One possibility may be that the length of a β-sheet may not accommodate small changes in bilayer thickness as easily as helices. Another possibility is that helices may insert individually whereas β-stands must align or zipper-up to form sheets before insertion which may be unfavorable under many circumstances.

Given that the hydrocarbon chains of lipids are 14-24 CH_2 in length, then a transmembrane α-helix must be ~18 residues long and a transmembrane β-sheet would only need to be ~7 residues long. These values have allowed investigators to predict the membrane spanning regions from the sequences of various membrane proteins. In this method, each residue of the protein is assigned a hydrophobicity value using one of the scales mentioned above. The protein is scanned starting from the N-terminus using a window of a particular number of residues, either 18 or 7. If the hydrophobicity of a particular window is high or above a base value, then that region may be a transmembrane region. This method is shown in Figure 6 where the GES or Kyte-Doolittle scales are used at 7 or 20 for bacteriorhodopsin, where it has been established to contain 7 transmembrane helices. Note that this method serves as a good test of hydrophobicity scales as well as a method that can predict transmembrane helices.

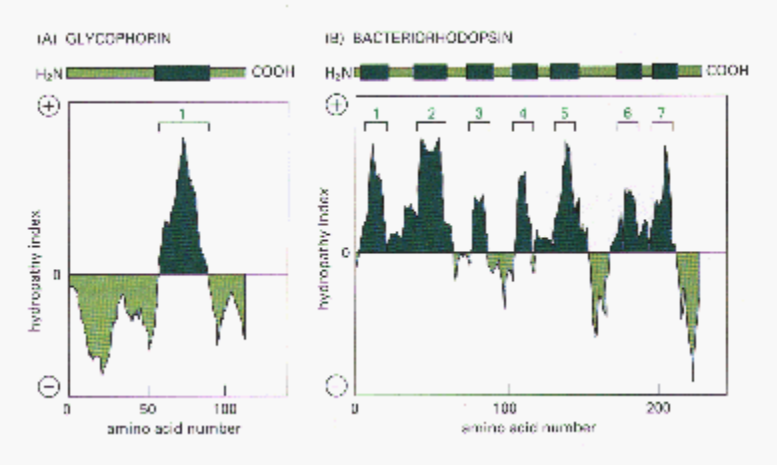

Figure 6: GES data for glycophorin and bacteriorhodopsin (see above) adapted from D. Eisenberg, Annu. Rev. Biochem. 53, 595-624 (1984).

Interestingly, many membrane-spanning regions contain Pro, a "helix-breaker" in the center. Pro may change the tilt of either end of the helix thereby adjusting the helical length or orientation. Also, Pro residues may help position the residues on its helix with neighboring helices or with a cofactor as seen in the structure of the photosynthetic reaction center (see below).

2.2d Tertiary Structure of Transmembrane Proteins

Folding of membrane proteins is expected to be similar to aqueous proteins in that the secondary structural elements first form and then these elements come together to the final tertiary structure. However, the point in the folding pathway where membrane insertion occurs is not completely clear. Very elegant structures of the transmembrane helical protein glycophorin (Hunt, 1997) suggest that folding of integral membrane proteins occurs by a two-step process. The first step involves insertion and formation of the helices and the second step involves association of the transmembrane helices. Since most membrane proteins are helical, this may be a general pathway. Alternately, folding of ú-sheet proteins most likely occurs through an alternate pathway involving formation of the sheet and then insertion into the membrane. As pointed out in a recent review of membrane protein folding by White & Wimley, (White and Wimley, 1999), key structural features that should be kept in mind are that the

tertiary structure of membrane proteins have similar packing as aqueous soluble proteins and that hydrogen bonding between the helices are rare and salt bridges are not observed. Thus, interactions between the structural elements occur by optimizing less energetic interactions between the secondary structural groups (i.e. by optimizing the packing between the helices), and can result in a reduced temperature sensitivity of these proteins.

How do transmembrane helices associate in the membrane interior? There are two mechanisms that the protein may use: First, the protein may arrange its nonpolar side chains in such a way as to maximize helix-helix packing. Second, the protein may use polar and hydrogen bonding side chains to stabilize interactions between the helices. Since the rise/residue of an α-helix is 3.6, then ~ every third residue should be polar for two interacting helices and two of three residues if multiple helices interact. to determine whether one or more sides of a helix interact, Eisenberg and coworkers formulated the *hydrophobic moment* (Rees, 1989). The hydrophobic moment ($\boldsymbol{\mu}_s$) is a vector of the sum of the hydrophobicity of the particular residues on the helix times the unit vector from the nucleus of the α C to the center of the side chain:

Thus, knowing the hydrophobic moment, it may be possible to predict which surfaces of the helix may interact with the membrane interior and which may interact with other helices.

$$\vec{\mu}_s = \sum_{n=1}^{n} H_n \vec{S}_n$$

Another interaction that promotes the association of helices is the *helical moment*. The helical moment is a relatively weak electrostatic dipole that results from the configuration of peptide bonds in an α-helix. Peptide bonds have a small dipole moment due to their resonance structure. When arranged in a helix, these small dipoles can sum together to give an overall dipole moment to the helix, and of course, the longer the helix the stronger the moment. The net result is that antiparallel configuration of helixes would be preferred over parallel. As we shall see later, transmembrane helices tend to align in an antiparallel configuration.

Even though some mechanisms such as the helical moment may play a role in the tertiary structure of membrane proteins, by far the factor that drives the association of helices in bilayers is the optimal packing of the helices. Most membrane-spanning helices are tilted ~21^{o} with respect to the bilayer plane and ~20^{o} to each other. This tilting appears to be the result of the "A knob-into-hole" packing arrangement of the helices causing the proteins to be arranged like left-handed coiled-coil proteins (Figure 7) (Walther et al., 1996).

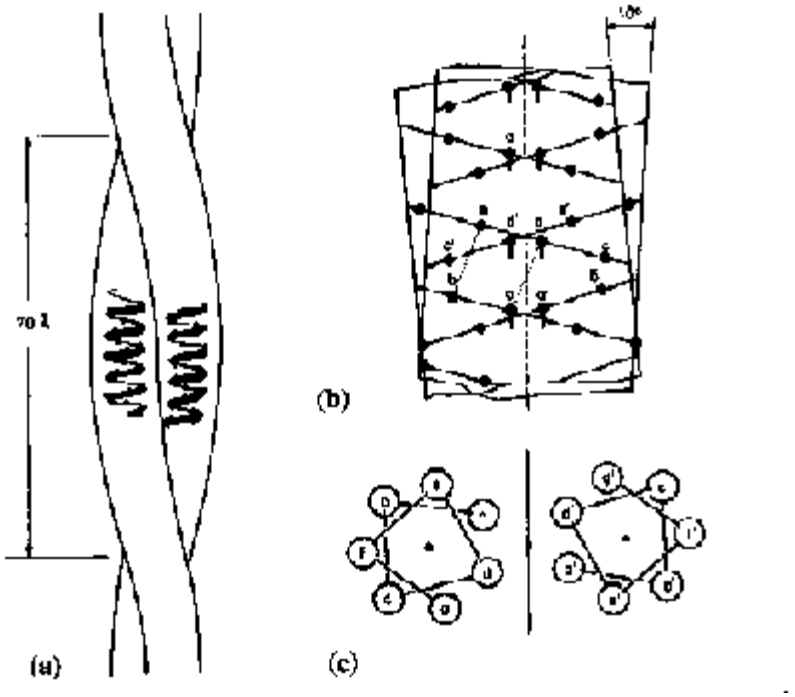

Figure 7: Depiction of coiled-coil helices taken from G.E. Schultz and R.H. Schirmer, Principles of Protein Structure, (1979) Springer Advanced texts in Chemistry, Charles R. Cantor, Editor. Springer-Verlag, New York, (printed with permission from Springer-Verlag).

2.2e An example of how interfacial contacts can stabilize tertiary structure: Case of gramicidin. Given that antiparallel, hydrophobic helical structures are the preferred membrane-spanning regions are there any other interactions between membranes and their integral proteins that can stabilize certain conformational states? One clear example where this is the case is the small antibiotic gramicidin. Gramicidin is a 15 residue peptide of alternating *l*- and *d*- amino acids. Although its primary function is thought to inhibit RNA synthesis, in membranes it forms channels or pores which are specific for monovalent cations. The N-terminus is formlyated and the C-terminus is acetylated and so the peptide is uncharged. Due to its availability and ability to be reconstituted in a variety of types of lipid membranes, gramicidin has become a key model system for the basic principles of interactions and properties of channels. Several years ago, it became apparent that this simple peptide can take on at least two major conformations (Figure 8) (Wallace, 1990). One is an intertwined helical dimer which functions as a pore for monovalent cations. The second structure is an N-N terminal dimer which functions as channel. The two structures can be distinguished by CD. For years, investigators found that certain lipid environments and certain reconstitution procedures would yield one conformation over the others. However, the reason for these preferences did not become apparent until it was realized that when the bilayer length matched the length of the peptide, the N-N terminal dimer would be preferred because this structure could be stabilized by hydrogen bonds from the indole protons of the Trp residues to the membrane head groups. If the length was mismatched, then either structure could form. This mismatch idea was initially observed by monitoring the rotational freedom of the 4 Trp residues that line the top of the gramicidin peptide. When the Trp residues participate in hydrogen bonds, the rotational motion is limited. Breaking these bonds, either by increasing the hydrostatic pressure, which increases bilayer length due to chain straightening, using longer chained lipids, or using smaller lipid head groups (PE instead of PC), allows the rotational motion of the Trp to increase (Scarlata, 1988; Scarlata, 1991; Scarlata and Gruner, 1997).

These interfacial hydrogen bonds, which were subsequently confirmed by nmr (Ketchum et al., 1993), stabilize the N-N terminal peptide structure by 2 kcals/mol per Trp.

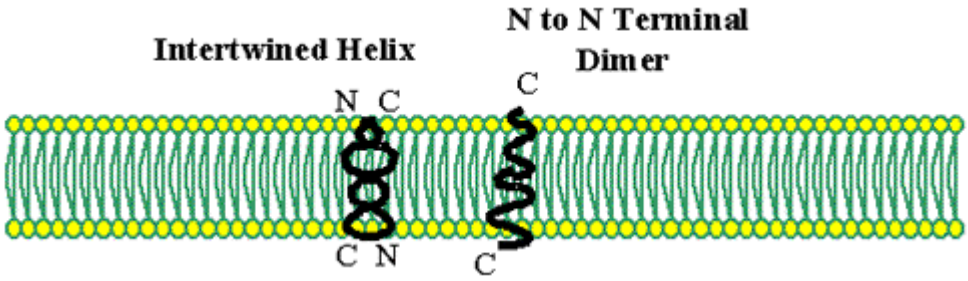

Figure 8: Cartoon of the two major conformational states of gramicidin.

2.2f Example of the stabilization of a functional conformation of a membrane protein by surface contacts: Case of GLUT1. If interfacial contacts can stabilize certain conformations of membrane proteins, could these interactions work to stabilize functional states of other proteins? Fluorescence-based studies were again used to investigate changes in interfacial position of the 12 interfacial Trp and Tyr of the hexose transporter, GLUT1 with increased lipid packing (Scarlata et al., 1995). In the ligand-bound active conformation, the interfacial residues are closer to the membrane surface and interact more strongly with the lipid head groups. As lipid packing was increased by raising the hydrostatic pressure, the interfacial residues moved closer to the surface and the liganded conformation was stabilized. Thus, bilayer length can stabilize certain states of transmembrane proteins.

2.2g General observations of the role of interfacial contacts and membrane protein structure. The above two sections discussed our work showing the importance of Trp-lipid contacts in stabilizing the structural and functional state of membrane proteins. Since those studies, very comprehensive investigations of the location of Trp and other residues have been carried out by White and colleagues, (e.g. see Wimely and White, 1996) which all indicate that Trp and Tyr are never buried in the hydrocarbon interior but remain interfacial and surface-exposed. Indeed, as more structures of membrane proteins become known, it is apparent that Trp and Try residues are located at interfacial positions. These residues can act as amphipathic anchors that can stabilize the protein structure by hydrogen bonding, but can be partially buried into the hydrocarbon interior or exposed to solvent with relatively low energy costs. Also, the orientation of the rings can be altered to best align with the lipid chains and to optimize hydrogen bond positioning with changes in bilayer thickness.

It is important to keep in mind that the contacts membrane proteins make with the lipid surface are very energetic as compared with the van der Waal interactions between the protein and the hydrocarbon chains. Also, these contacts may play a critical role in stabilizing conformational and functional states of membrane proteins. Since the energy of these contacts fall rapidly with distance ($1/r^3$), any minor changes in bilayer thickness may result in stabilization or destabilization of alternate protein conformations. However, the hydrophobic interactions between the transmembrane region of a protein and the lipid occur over a larger area, making the cumulative sum of these less energetic interactions significant. In this sense, the residues in the membrane interior stabilize the overall partitioning of the protein into the membrane, but the surface interactions can fine-tune the stabilization of particular protein conformations

2.2h Role of lipid packing on membrane protein conformation and function. The question of

whether lipid packing affected protein function was the focus of many studies several years ago. Generally, these studies were done by measuring the activity of the particular enzyme reconstituted in lipids of varying packing. The results of these studies were generally disappointing in that no clear picture emerged. However, in these studies could not control for the changes in bilayer length that accompanied changes in lipid packing. Based on the understanding that the surface contacts may be an important regulatory feature in membrane protein structure, it is likely that many of these changes can be traced back to changes caused by bilayer length and energetic interfacial interactions.

It is possible that movements of integral membrane proteins are sensitive to viscous damping by the lipid chains. Increasing the packing of the lipid in the fluid phase may be similar to moving in air versus moving in a swimming pool; although you may be able to move more slowly you can still move. Values estimated for the viscosity of lipid bilayers in the plane parallel to the lipid chains indicate that they may dampen motions, but still allow protein movement to occur since the intrinsic viscosity of the protein is expected to be much higher. Once the lipid enters the gel phase, it is as if the water in the swimming pool has turned to ice and backbone motions of protein are greatly inhibited. This extensive damping occurs either through the large increase in membrane viscosity, by changes in interfacial surface contacts or both.

2.2i Quaternary structures of membrane proteins. The same ideas discussed for the association of secondary structural elements also apply to protein-protein associations. Engelman and colleagues have investigated membrane protein oligomerization by measureing the dimerization of the single-transmembrane protein glycophorin as well as the association of bacteriorhodopsin helices in bilayers (see Hunt et al., 1997; Hunt, 1997). Their work shows the importance of packing interactions between the helices. However, aside from these model studies, our knowledge of the thermodynamics, kinetics and other physical properties that govern the oligomerization of integral membrane proteins is limited due to the lack of methods to monitor protein-protein interaction in the membrane. Although chemical cross-linking has been used in many systems, it is not clear whether the cross-links that form reflect specific complexes or are the result of lateral collisions between non-interacting species. Other methods include measuring the rotational volume of the protein complex with a long lived fluorescent or phosphorescent probe. While a criticism of this approach is that the protein may carry annular lipids in its rotation and give an erroneous molecular size, this does not appear to be the case (e.g.Roopnarine et al., 1993). A related approach is fluorescence homotransfer in which the oligomer is labeled with identical fluorophores and the number of transfer events is detected (e.g.Runnels and Scarlata, 1995). Another promising method is neutron diffraction techniques (see below). A problem with assessing the accuracy of these techniques is the lack of a basis set of membrane protein oligomers which one could use to validate these methods. Once this set is defined, then various techniques can be tested and better refined.

2.2j Structure of membrane-anchored and peripheral membrane proteins. As mentioned, many proteins are anchored to the membrane surface by a post-synthetic modification, such as myristyation, farnesylation or GPI-linkages. While their folding patterns will follow the same rules as aqueous-soluble proteins, some other factors should be kept in mind. The energy these modifications must contribute to keep their host protein bound to the membrane is ~0.2kcal/mol/ CH_2 following Tanford=s rule for hydrophobic interaction energy (see Peitzsch and McLaughlin, 1993). Some post-synthetic linkages, such as palmitoyl and farnesyl groups reversibly attach to Cys side chains thus promoting reversible membrane association. It is important to note that this energy is not sufficient to keep a protein bound to the membrane and these membrane-bound proteins must have other localization motifs, such as C2 and pleckstrin homology domains that target specific lipid head groups or clusters of basic residues that associate with negatively charged lipids. Keeping in mind that natural membranes have an abundance of negatively charged phospholipids, the most common is electrostatic interactions. A common theme of

peripheral proteins is a positively charged lobe either by virtue of its primary structure (eg RAS, HIV-MA) or by inclusion of a cation binding motif, such as the Ca^{2+}-binding C2 domains found in proteins such as protein kinase C (Nalefski and Falke, 1996).

Besides electrostatic charge interactions, many proteins bind to membranes through hydrophobic interactions by the formation of amphiphatic helices. An example of this is melittin, which forms an amphipathic tetramer whose hydrophobic regions can penetrate the membrane surface (Figure 9). Indeed, membrane insertion of amphipatic helices is a function of the nature of the membrane surface as well as the exposed hydrophobic area with highly curved small unilamellar vesicles being more penetrable than flatter large, unilamellar vesicles and PE lipids more penetrable than PC. Many toxins and smaller peptides use this mechanism for membrane insertion (see White and Wimley, 1999). The penetration of proteins into biological membranes may be promoted by the many non-uniformities of the membrane surface.

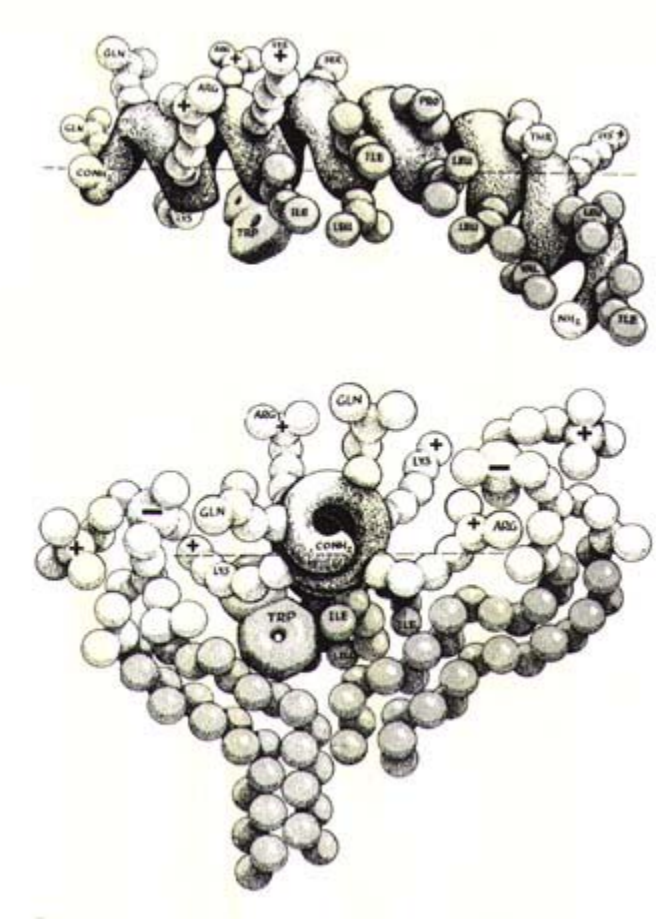

Figure 9: Reproduced with permission, from Terwilliger, Weissman and Eisenberg, *Biophys. J.* 37, 353-361 (1982).
(top) Structure of melittin showing the amphipathin nature of the helix where the dashed line depicts the membrane interface. (bottom) Model of the melittin-membrane complex.

2.3 Functional Families of Membrane Proteins

2.3a Families of transmembrane proteins. Many transmembrane proteins that are structurally related are also functionally related. For example, the EGF (epidermal growth factor receptor) and the insulin receptor fall into a family of growth factor receptors which have very large disufide-rich extracellular and a tyrosine kinase intracellular domains connected by a single-transmembrane helix (Figure 10). Most members of this family are monomers and binding of ligand induces dimerization and activation of the intracellular tyrosine kinase domain. Note that in the unliganded state the insulin receptor is a dimer and it is possible that in this case the binding of insulin changes the intersubunit orientation of the monomers allowing for activation.

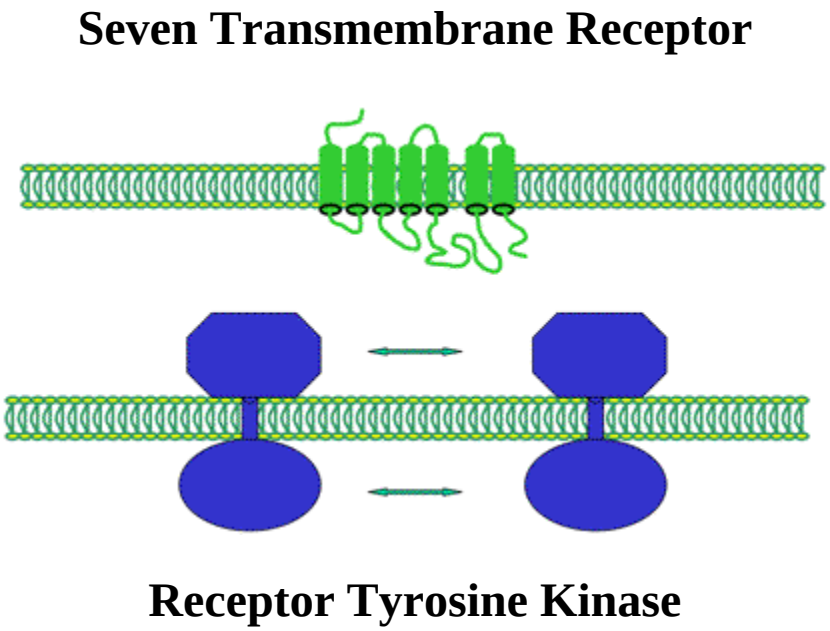

Figure 11: Cartoon of two families of transmembrane receptors.

Another important family of transmembrane proteins are the seven transmembrane family of G proteins (guanine nucleotide binding proteins) coupled receptors. These receptors are the most abundant class of receptors in mammalian cells and communicate an extremely diverse range of signals into the cell, from light (rhodopsin) to neurotransmitters (muscarinic or adrenergic receptors) to sex-related signals (oxytocin). Although their ligand activators are diverse, these receptors all couple to G proteins to transduce their signal. Structurally, they are similar in having seven transmembrane loops in a defined topology (see Figure 10). In contrast to the growth factor receptor family, these proteins have relatively small extramembrane loops.

Integral membrane proteins that transport species such as nutrients and ions must be able to shield their ligands from the surrounding hydrocarbon interior. Thus, these proteins are much larger than the signal transduction proteins mentioned above, and often contain several subunits. An example of this class is the 12 membrane spanning family belonging to transporters, such as GLUT1 and antibiotics. Other examples are given in the following section.

2.3b Examples of transmembrane protein structure. Now that the basic principles governing membrane protein structure have been laid, it would be interesting to view some examples of membrane proteins. A current listing of known membrane structures and their corresponding PBD codes is kept up by Dr. Steven White's laboratory at http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html . The

reader is encouraged to view many of these structures. To avoid being tedious, only a few will be discussed here.

Bacteriorhodopsin, photosynthetic complexes, respiratory complexes and the potassium channel can all be structurally classified as polytopic transmembrane helical proteins with small, structurally variable extramembrane regions that vary with the functional properties of the protein Figure 11. Rather than being a membrane-anchor, the orientation and alignment of the transmembrane helices are thought to mediate protein function.

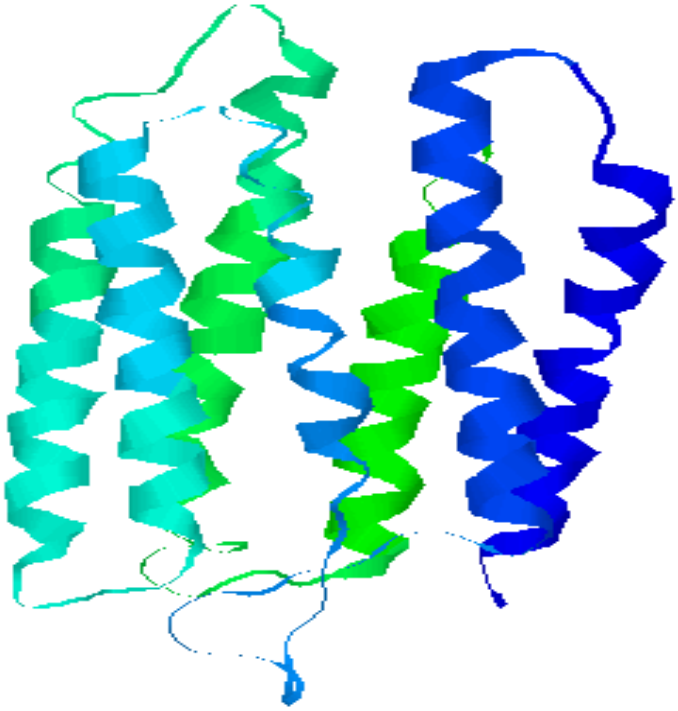

Figure 11: Structure of bacteriorhodopsin (PBD 1AP9).

Porins and related proteins form an antiparallel ú sheet hole in membranes which allow small molecules to pass. The porin from *R. capsulatus*, and PhoE and OmpF from *E. coli* have a 16 stranded antiparallel sheet while LamB and SrcY have 18 stranded antiparallel ú sheets (e.g. see Figure 12). The amino acid sequences of this class are predominantly polar with a hydrophobic belt of membrane-exposed residues.

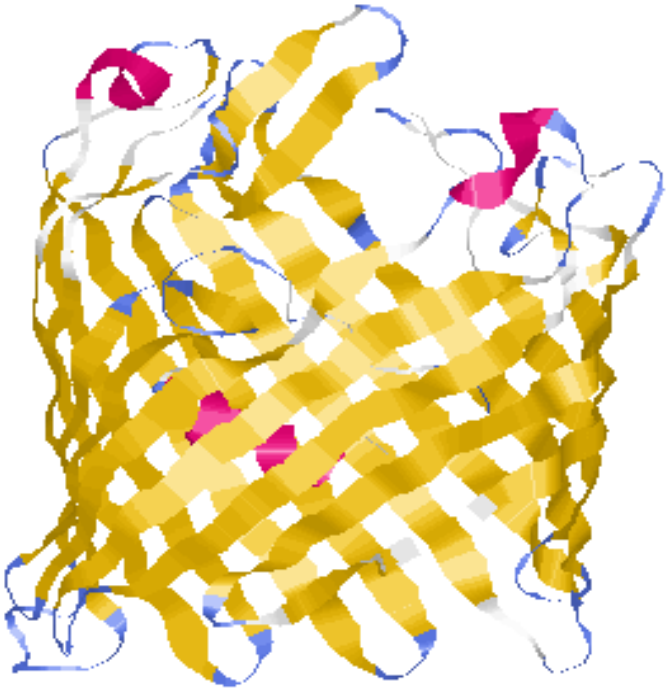

Figure 12: Example of a beta-structure membrane protein (PDB OPF).

Shown in Figure 13 is the structure of the heptomeric protein a-hemolysin. The top of `cap= of the protein is composed of 7 ú sandwiches between this domain and the transmembrane stem is a region rich in basic and aromatic residues. The transmembrane stem forms a channel with a 14-24’ depending on the side chain that protrude into the cylinder. Two hydrophobic bands consisting of Met and Lue residues are exposed to the hydrocarbon interior of the membrane with another band of aromotic close to the membrane interface.

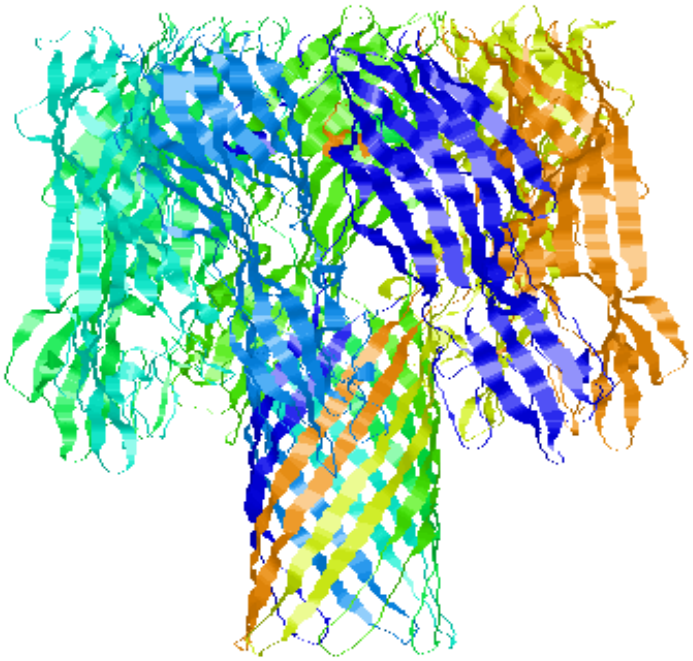

Figure 13: Structure of the heptapeptide alpha-hemolysin (PBD 7AHL).

These types of transmembrane proteins illustrate the various known membrane proteins topologies. It is certain that more topologies will be observed as more structures of membrane proteins are solved.

REFERENCES

Alberts, Bray, Lewis, Raff, Roberts, and Watson. 1994. Molecular Biology of the Cell. Garland Publishing, Inc., New York.

Cornette, J. L., Cease, K.B., Margalit, H., Spouge, J.L., Berzofsky,J.A. and DeLisi, C. 1987. Hydrophobicity scales and computational techniques for detecting amphipathic structures in proteins. *J.Molec.Biol.* 195:659-685.

Engelman, D. M., Steitz, T.A. and Goldman, A. 1986. Identifying nonpolar transmembrane helices in amino acid sequences of membrane proteins. *Ann. Rev. Biophys. Biophys.Chem.* 15:321-353.

Gennis, R. 1989. Biomembranes: Molecular Structure and Function. Springer-Verlag, New York.

Hunt, J., P. McCrea, G. Zaccai, and D. M. Engelman. 1997. Assessment of the aggregation state of integral membrane proteins in reconstituted phospholipid vesicles using small angle neutron scattering. *J.Molec. Biol.* 273:1004-1019.

Hunt, J. F., Earnest, T.N., Bousche, O., Kalshathi, K., Reilly, K., Horvath, C., Rothschild, K.J. and Engelman, D.M. 1997. A biophysical study of integral membrane protein folding. *Biochemistry.*

36:15156-15176.

Israelachvili, J. 1978. Refinements of the fluid-mosaic model of membrane structure. *Biochim.Biophys.Acta.* 469:221-225.

Ketchum, R., W. Hu, and T. Cross. 1993. High-resolution conformation of gramicidin A in a lipid bilayer by solid-state NMR. *Science.* 261:1457-1460.

Kyte, J. and Doolittle, R.F. 1982. A simple method for displaying the hydropathic character of a protein. *J. Molec. Biol.* 157:105-132.

Lee, A. G. 1998. How lipids interact with an intrinsic membrane protein: the case of the calcium pump. *Biol.Skr.Dan.Vid.Selsk.* 49.

Mouritsen, O. G. and Bloom, M. 1984. Mattress model of lipid protein interactions in membranes. *Biophys.J.* 36:141-153.

Nalefski, E., and J. J. Falke. 1996. The C2 domain calcium binding motif: Structural and functional diversity.

Peitzsch, R., and S. McLaughlin. 1993. Binding of acylated peptides and fatty acids to phospholipid vesicles: Relevance to myristolylated proteins. *Biochemistry.* 32:10436-10443.

Petty, H. 1993. Molecular Biology of Membranes. Plenum Press, Inc., New York.

Rao, J. K. M. and Argos, P. 1986. A conformational preference parameter to predict helices in integral membrane proteins. *Biochim.Biophys.Acta.* 869:197-214.

Rees, D. C., DeAntonio, L. and Eisenberg, D. 1989. Hydrophobic organization of membrane proteins. *Science.* 245:510-513.

Roopnarine, O., K. Hideg, and D. D. Thomas. 1993. Saturation transfer electron parametric resonance ofan indane-dione spin-label. Calibration with hemoglobin and application to myosin rotational dynamics. *Biophysical Journal.* 64:1896-907, 1993.

Runnels, L., and S. Scarlata. 1995. Theory and application of fluorescence homotransfer to melittin oligomerization. *Biophys.J.* 69:1569-1583.

Sackmann, E. 1995. Biological membranes. Architecture and function. *In* Handbook of Biological Physics. L. a. Sackmann, editor. Elsevier, Amsterdam. 1-63.

Scarlata, S. 1988. The effects of viscosity on gramicidin tryptophan rotational motion. *Biophys. J.* 54:1149-1157.

Scarlata, S. 1991. The effect of increased chain packing on gramicidin tryptophan-lipid interactions. *Biochemistry.* 30:9853-9859.

Scarlata, S., and S. Gruner. 1997. Role of phosphatidylethanolamine lipids in the stabilization of protein-lipid contacts. *Biophys.Chem.* 67:269-279.

Scarlata, S., H. McBath, and H. Haspel. 1995. Effect of lipid packing on the conformational states of purified GLUT-1 hexose transporter. *Biochemistry.* 34:7703-7711.

Silvius, J. R. 1982. Thermotropic Phase Transitions of Pure Lipids in Model Membranes and Their Modifications by Membrane Proteins. *In* Lipid-Protein Interactions. John Wiley & Sons, New York.

Simon, S. M. 1995. Protein-conducting channels for the translocation of proteins into and across membranes. *Cold Spring Harbor Sym. Quart. Biol.* 60:57-69.

Singer, S. and Nicolson, G.L. 1972. The fluid mosaic model of cell membranes. *Science.* 172:720-730.

Wallace, B. A. 1990. *Annu.Rev.Biophys.Biophys.Chem.* 19:127-157.

Walther, D., F. Eisenhaber, and P. Argos. 1996. Principles of helix-helix packing in proteins: The helical lattice superposition model. *J.Molec.Biol.* 255:536-553.

White, S., and W. Wimley. 1999. Membrane protein folding and stability: physical principles. *Annu.Rev.Biophys.Biomolec.Struc.* 28:319-365.

Wimely, W. C., and S. H. White. 1996. Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat.Struc.Biol.* 3:842-848.

Yeagle, P. 1992. The Structure of Biological Membranes. CRC Press, Boca Raton.

Last updated 10/20/99