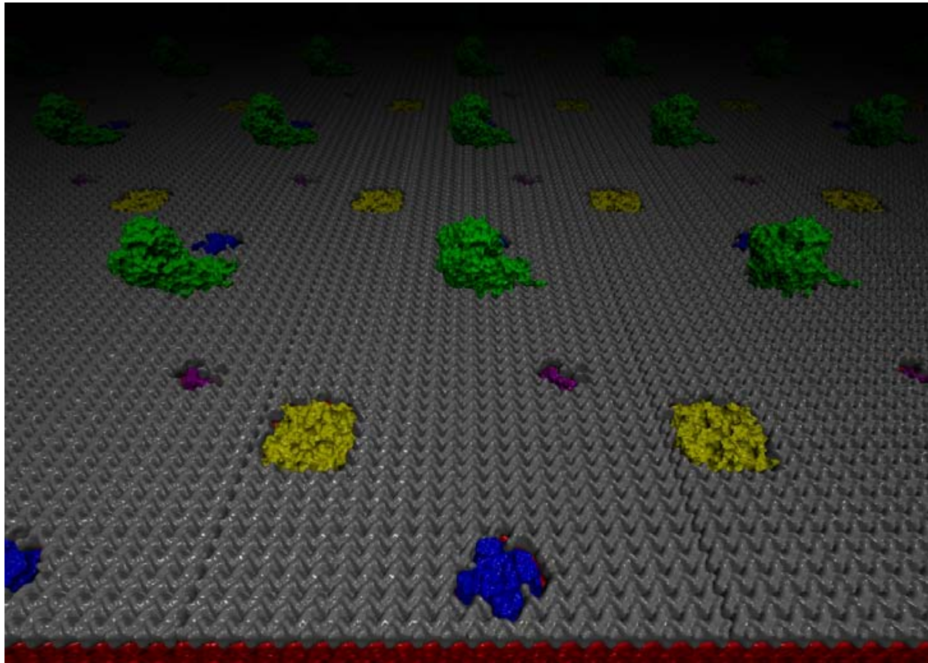
Case Study: Membranes

Peter Freddolino and Amy Shih
April 13, 2006



1 Introduction to Lipids and Membranes

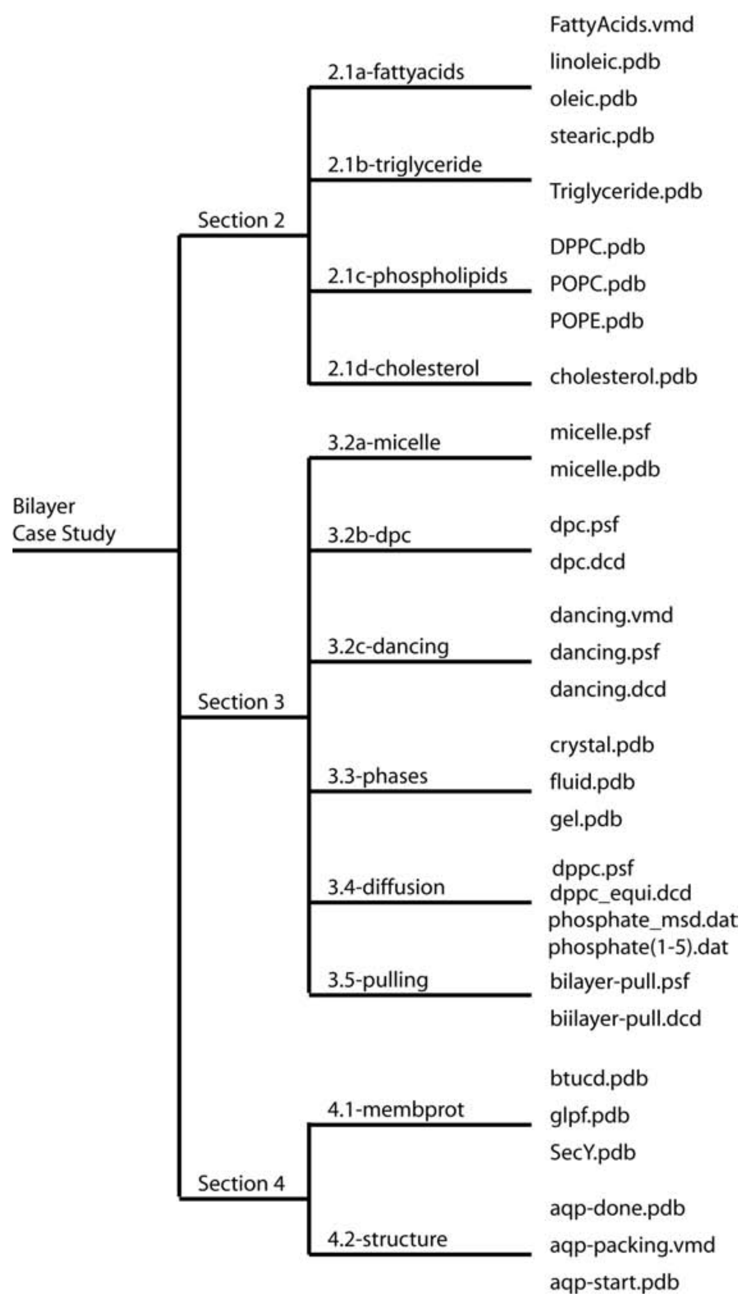
Membranes are essential to cellular organisms. They are like fortresses in that they provide a barrier between the inside and outside with guarded drawbridges in the form of proteins that regulate the influx and efflux of material. Unlike the rigid walls of a fortress, membranes are fluid and are able to bend and move. The bricks forming a membrane, called lipids, freely move. Plasma membranes enclose and define the boundaries of a cell, maintaining a barrier between the interior of a cell, the cytosol, and the extracellular environment. In eukaryotic cells, membranes also surround specialized organelles such as the mitochondria, nucleus, and other membrane bound

organelles. Some cells can also have large membrane bound compartments called vacoules which serve a variety of functions including capturing food, sequestering toxic material, surrounding and eliminianting cellualr debris and maintianing fluid balance (turgor). Vesicles are relatively smaller membrane bound compartments which can store, transport or digest cellular products and waste. For example, cell organelles like the endoplasmic reticulum contain membranes seperating the cytoplasm from their lumen. In the case of the rough endoplasmic reticulum, ribosomes attach to the surface releasing newly synthesized proteins, through membrane channels, translocons, either into the lumen or into the membrane itself, in the case of membrane proteins. Another example are lysosomes are membrane-bound digestive vesicles found inside cells. They contain lysozymes (digestive enzymes) that can break down macromolecules allowing cells to feed. The membrane of lysosomes are impervious to lysozymes allowing cells to break down their food without being digested by their own digestive enzymes [12]. Although biological membranes have diverse functions and compositions, they are all structurally similar in that they are comprised of a thin, 5-10 nm, layer of lipids and proteins held together by primarily non-covalent interactions.

1.1 Membrane Function

The function of membranes is diverse. Membranes form continuous sheets enclosing a defined inner compartment. This compartmentalization allows for specialized activities with regulated interaction with the surrounding environment. Although membranes form a barrier allowing for compartmentalization, they do not form an impenetrable barrier. Membranes are selectively pervious in that they allow certain molecules through but exclude other molecules. A pure lipid bilayer is slightly permeable to water, but mostly impermeable to water-soluble molecules such as sodium, calcium, and potassium. It is the responsibility of membrane proteins to regulate the transport of such molecules. The distribution and types of membrane proteins found within a particular membrane varies, thus allowing for the customization of each interior space to the specific requirements of the cell or organelle. Imagine in our fortress analogy that drawbridges are manned by guards who can tell the difference between friend and foe. Thus friendly material can be allowed entrance, while adverse materials can be kept at bay.

In this case study, we will use the following files as shown below.



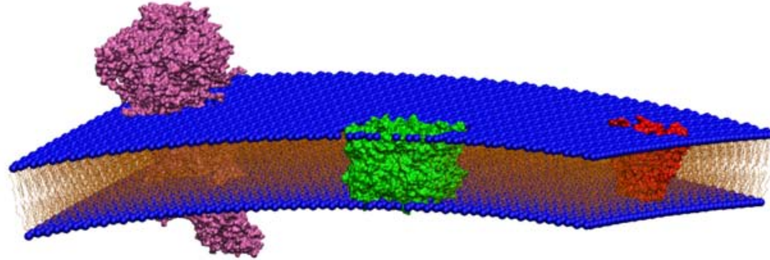


Figure 1: Fluid mosaic model with membrane proteins shown in surface representation embedded in the lipid bilayer. The hydrophobic interior of the lipid bilayer is shown in orange with the hydrophillic head groups shown in blue.

Transport of solutes across the plasma membrane allows cells to accumulate high cytoplasmic concentrations of substances such as amino acids and sugars that are important to the metabolism and maintenance of the cell. The ionic composition within a cell usually differs from that of the surrounding environment. Cells generally maintain a high cytosolic concentration of potassium ions compared to the exterior environment and a low concentration of cytosolic sodium ions. This creates an ionic gradient, which allows the passive diffusion of ions from high to low concentration when membrane protein channels are open. The movement of ions through channels creates an electric potential difference where the inside of the cell is roughly -70 mV more negative then the exterior. Thus the plasma membrane is also called a capacitor. Proteins within membranes, furthermore play a role in intercellular communication and signal transduction [10].

1.2 Membrane Structure and Composition

In 1972 a membrane model, the fluid mosaic-model, was proposed by Singer and Nicolson and has developed into the fundamental structural model for membranes (Figure 1) [10]. In this model, lipid bilayers form the structural backbone of the membrane with a heterogeneous collection of proteins. Integral membrane proteins span the entire lipid bilayer while peripheral membrane proteins loosely associate with the surface of a membrane. The lipids in the fluid-mosaic model diffuse laterally, but rarely flip-flop from an inside-out to an outside-in orientation. The fluidity of the membrane is de-

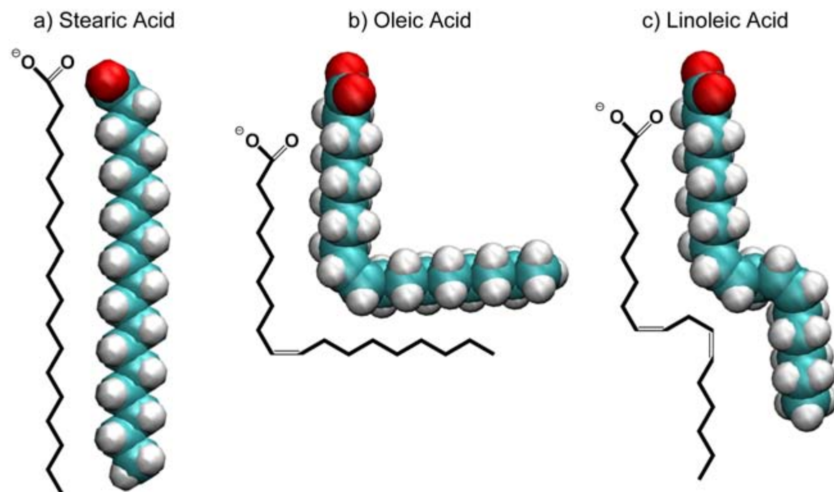


Figure 2: Fatty acid chemical structure and space filling model. a) Saturated stearic acid, b) monounsaturated oleic acid with a single cis-double bond, and c) polyunsaturated linoleic acid with two cis-double bonds. (View these fatty acids by loading the files `linoleic.pdb`, `oleic.pdb`, and `stearic.pdb` or by loading the VMD saved state `FattyAcids.vmd`.)

termined by the composition of lipids and proteins. A biological membrane is comprised of roughly half lipids and half proteins. A wide variety of lipids are found in biological membranes, but the most prevalent are phospholipids, sphingolipids, and glycolipids.

2 Lipids

Lipids are molecules which are insoluble in water, but soluble in organic solvents. We are most familiar with lipids in the form of fats and oils; hormones such as testosterone, are also lipids. Membranes are primarily composed of amphiphilic lipids, such as phospholipids. These lipids have a hydrophobic, “water fearing”, tail group and a hydrophilic, “water loving”, head group. The lipids in membranes form a bilayer structure with the hydrophobic tail groups interacting with each other and the polar hydrophilic head groups interacting with the aqueous interior and exterior environments.

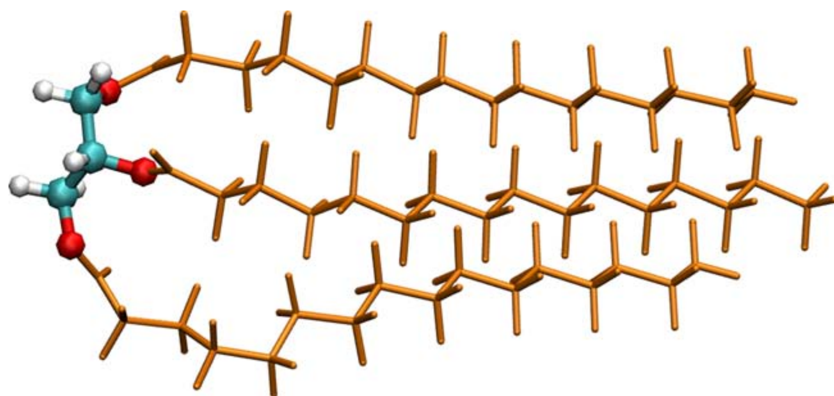


Figure 3: Triglycerides are triesters of glycerol with three fatty acid chains. The hydrophobic fatty acid tail groups are shown in orange and the glycerol atoms are shown colored according to their atom type, with carbons in cyan, oxygens in red, and hydrogens in white. (View the structure of a triglyceride using `Triglyceride.pdb`).

2.1 Types of Lipids

Fatty acids are the building blocks of lipids. They are composed of a long hydrocarbon chain with a carboxcylic acid head group. Although free fatty acids do not occur in nature, they are the hydrocarbon chains that make up the tail groups of most lipids. There are usually between 14 and 20 carbons within a fatty acid hydrocarbon chain; carbons are usually even in number. Fats on food labels are often described as being saturated, unsaturated, and polyunsaturated. These descriptions are used to describe the type of fatty acid hydrocarbon chains. A saturated fatty acid contains no double bonds. These produce long straight chain hydrocarbons, whereas unsaturated fatty acids contain double bonds and polyunsaturated two or more double bonds. The presence of a double bond causes a kink or bend in the hydrocarbon chain (Figure 2).

Most naturally occurring fats and oils in plants and animals consists of mixtures of triglycerides. They are triesters of glycerol with three fatty acid chains connected (Figure 3). Simple triglycerides have three identical fatty acids whereas the more common mixed triglycerides have 2 or 3 types of fatty acids. Fats, a solid at room temperature, and oils, a liquid at room temperature, are made from mixtures of simple and mixed triglycerides. Although triglycerides are not present in membranes and bilayers, they are the

primary way in which fats are stored for energy needs of an organism.

The most prevalent type of lipid found in biological membranes are phospholipids. Phospholipids have a glycerol phosphate where the glycerol is esterified in the C1 and C2 positions to two fatty acids, and the phosphoryl group to a variable head group. In biological membranes the head groups are often derived from polar alcohols such as choline, serine, and ethanolamine. The C1 position is often esterified to saturated fatty acids containing 16 or 18 carbons. In contrast, the C2 position often has unsaturated fatty acids of between 16 to 20 carbons. Phospholipids are all amphiphilic in nature with polar hydrophillic head groups and non-polar hydrophobic tails (Figure 4). Commonly found types of synthetic phospholipids in laboratory experiments include dipalmitoyl-glycero-phosphocholine (DPPC), 1-palmitoyl-2-oleoyl-glycerol-phosphocholne (POPC), and 1-palmitoyl-2-oleoyl-glycero-phosphoethanolamine (POPE). We will use VMD to look at the structures of these lipids (Figure 5). DPPC has unsaturated fatty acid tail groups, while POPC and POPE have one saturated and one unsaturated tail. DPPC and POPC both have a phosphocholine head group, while POPE has a smaller phosphoethanolamine head. The composition of the head and tail groups of lipids affects the overall structure, including packing density and surface area per lipid, of resulting lipid aggregates.

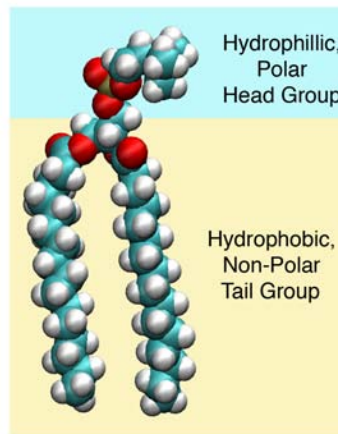


Figure 4: Phospholipids are amphiphilic with a hydrophillic head group and two hydrophobic tail groups. Phosphlipids have a cylindrical shape and easily form lipid bilayers.

Another type of lipid found in biological membranes are sphingolipids which instead of glycerol in phospholipids, have a long-chain amino alcohol sphingosine (Figure 7) with a fatty acid bonded to the C2 position in the sphingosine forming a ceramide. Addition of a polar head group to the C1 position forms a sphingolipid. There are three types of sphingolipids dependent upon the type of polar head group attached. Sphingomyleins are ceramides with either a phosphocholine or phosphoethanoline head group. They can and often are classified as a phospholipid due to the presence of the phosphate group. Glycolipids have sugar head groups. They occur primarily

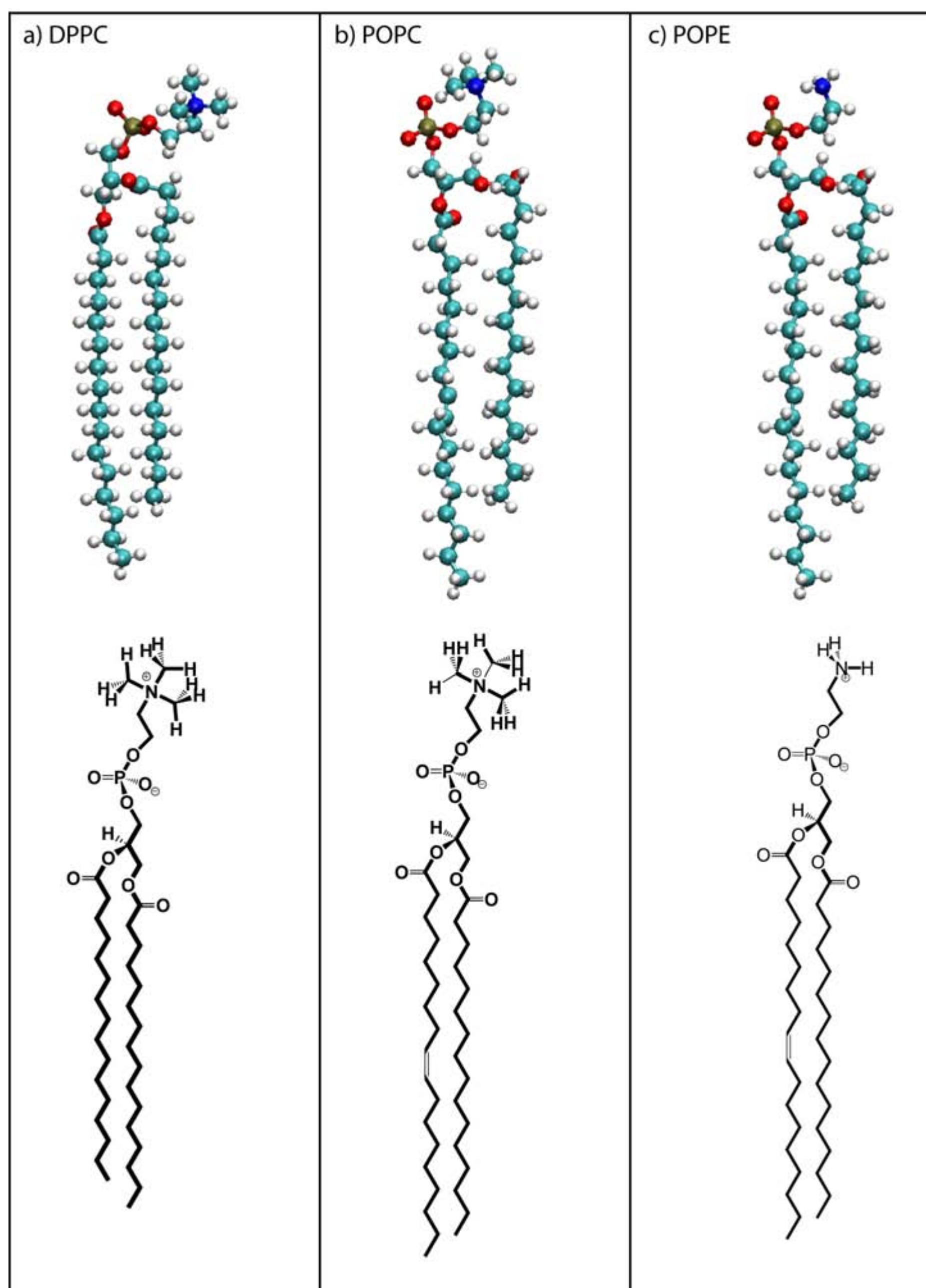


Figure 5: Chemical Structure and molecular view of three common phospholipids. a) DPPC, b) POPC, and c) POPE. (View the differences in these lipids by loading the following files into VMD: `DPPC.pdb`, `POPC.pdb`, and `POPE.pdb`.)

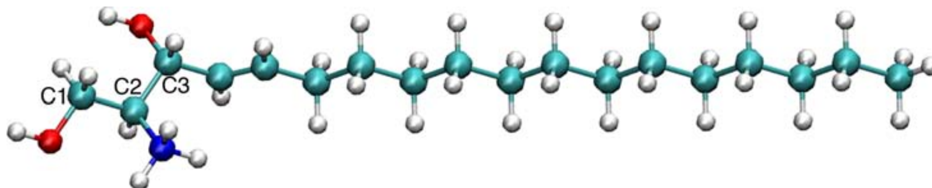


Figure 7: Sphingosine is the base unit in which a fatty acid attaches at the C2 position and a polar head group attaches to the C1 position forming a sphingolipid.

on the surface of the plasma membrane. Gangliosides are more complex sphingolipids, in that they have olgiosacharides as head groups with at least one sialic acid residue. This produces a negatively charged lipid at neutral pH. Sphingolipids often serve as receptors or biological recognition sites. The inability to break down sphingolipids can lead to fatal diseases [8, 1]. For example, a fatal inherited diseases, called Tay-Sachs, results fromthe lack of an enzyme (hexosamindase A) to break down glanglisoide GM2. Accumulation of glanglioside GM2 within the brain and nerve cells results in blindness, paralysis and eventually death [12].

The final type of lipids are steroids. which have tetracyclic cyclopentaphenanthrene skeletons. Steroids include the hormones, testosterone and estrogen, which regulate sexual development; and carbohydrate metabolism, as well as cholesterol. We will focus our attention on cholesterol, which can be found in the plasma membranes of animal cells. Cholesterol is important in a wide range of functions including fatty acid transport and metabolism and in the production of hormones, bile salts and vitamins. High levels of cholesterol in the blood can cause an increase in the risk for heart disease. Cholesterol is an amphiphilic molecule with a hydroxyl polar head group and a hydrophobic

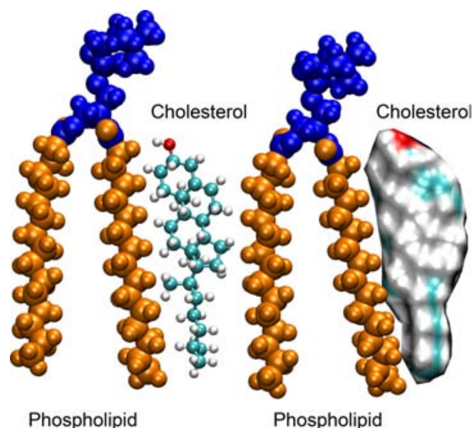


Figure 6: Cholesterol is a steroid which can be found in the plasma membrane of animals. It wedges itself in beween lipid molecules in a bilayer holding the first several lipid hydrocarbons rigid, decreasing the fluidity of a bilayer. (View the structure of cholesterol by loading the following file into VMD: `cholesterol.pdb`.)

hydrocarbon body. The ringed steroid region is rigid and when cholesterol is inserted into a lipid bilayer it interacts with the first several carbons on phospholipids holding them immobile (Figure 6). Membranes containing cholesterol are thus less flexible and less suseptible to small water-soluble molecule permeation. Cholesterol is only found in the plasma membranes of animals. Those organisms where cholesterol is not found, such as bacteria, generally require a cell wall to hold them rigid. Cholesterol concentration in plasma membrane is a balancing act. Too little cholesterol and the plasma membrane becomes too fluid, but too much cholesterol and the cells become too rigid.

Exercise 1: Energy Storage. The human body stores energy in the form of triglycerides. Fatty acid oxidation is the process in which triglycerides are broken down and converted into energy. The process yields approximately 9 kcal/g of energy, almost twice as much energy as obtained from the breakdown of proteins or carbohydrates. For an average 150 lb human male with 20% body fat, calculate the total amount of energy reserves. If the energy requirement for this human male were 2200 kcal per day, how many days would they survive if their sole source of energy came from their stored fat?

3 Lipid Bilayers

The lipid bilayer of membranes often consists of a particular ratio of various lipid types and proteins. This ratio varies among organisms and specific functions of membranes. Human mylein sheath has 30% by weight proteins, 30% phospholipids, 19% cholesterol, and trace amounts of other lipids. However, the inner cellular membrane of *E. coli* has 75% proteins, 25% phospholipids, and no cholesterol. As with fortresses, there can be a wide array of bricks and drawbridges which make up the structure, tailoring it for specific functions and environments.

Due to the hydrophobic effect, lipids form aggregates in aqueous environments. They can form micelles, bilayers, and liposomes depending on the type of lipids used. Biological membranes are composites of various types of lipids, however due to the slow lateral diffusion of lipids within a bilayer (diffusion rate of lipids is on the order of 10^{-8} cm^2/s [8]) most molecular dynamics simulations embed membrane proteins in bilayers of a single lipid type.

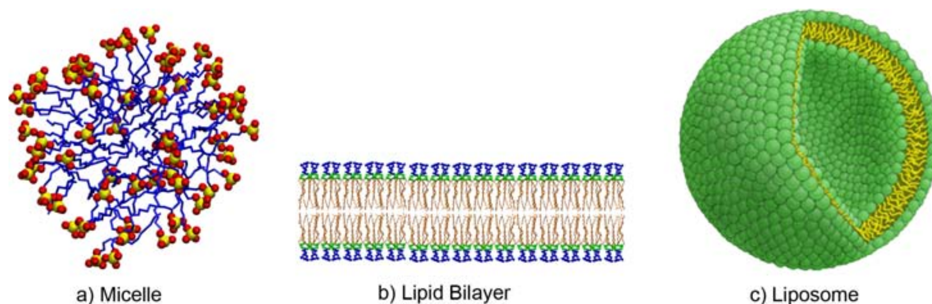


Figure 8: Types of lipid aggregates. a) Sodium dodecyl sulfate (SDS) is a common detergent with a large hydrophillic head and a small hydrophobic tail. When present in concentrations above its critical micellular concentration, SDS forms micelles in solution. The hydrophobic tails of the SDS are buried in the interior of the micelle and are shielded from the aqueous environment by their hydrophilic head groups. b) Lipid bilayers are planar and form when one lipid monolayer orients its hydrophobic tails towards the hydrophobic tails of a second monolayer, forming a hydrophobic interior. c) Liposomes also assume a bilayer structure, but instead of assuming a planar shape, the bilayer bends back onto itself (in order to stabilize the edges of a lipid bilayer) forming an aqueous cavity in the middle. (Prepare the image of a micelle using the files `micelle.psf` and `micelle.pdb`. Look at the structure of an individual SDS molecule by selecting a single residue. Additionally, load the files `dpc.psf` and `dpc.dcd` to view the aggregation of a DPC micelle from randomly placed lipids.)

3.1 Energetics of Lipid Aggregate Formation

When placed in aqueous solution lipids form into lipid aggregates. This spontaneous assembly is due to the hydrophobic effect. Lipids are amphiphilic in nature, and thus have a water-loving head group and a water-fearing tail group. The tail groups do not want to interact with water and, thus, group together in order to exclude water. This leaves the hydrophillic head groups to interact with the polar water resulting in a phase separation. This separation occurs due to the strong need for water to hydrogen bond with other polar molecules, especially other water molecules. The hydrophobic fatty acid tail groups of lipids do not offer hydrogen bonding partners for water and are thus excluded from being near water.

3.2 Lipid Aggregates

Various types of lipid aggregates form depending on to the specific proportion and structure of the lipids. Lipids with larger head groups (cross sectional

area) than that of their hydrophobic fatty acid side chain tend to form micelles in solution. Micelles are spherical structures in which the hydrophobic lipid tails are buried in the interior core and shielded from water (Figure 8a). The polar head groups are at the surface interacting with the aqueous environment. Micelles require a minimum concentration of lipids, such as the detergent sodium dodecyl sulfate, in order to form. This minimum concentration is called the critical micelle concentration.

Another type of lipid aggregate is the bilayer, in which a sheet is formed with two lipid monolayers. The hydrophobic tails of one lipid monolayer are oriented towards the hydrophobic tails of the second monolayer (Figure 8b). This type of lipid aggregate forms when the lipid head group and tail groups have roughly similar cross sectional areas. However this type of aggregate leaves exposed hydrophobic tail groups at the edges of the bilayer. To further stabilize the structure, lipid bilayers fold back onto themselves and spontaneously form into liposomes forming an aqueous cavity inside the aggregate (Figure 8c).

Exercise 2: Movement of Lipids. In order to examine the fluidity and movement of lipids within a bilayer, as described by the fluid mosaic model. We can use VMD to view the results of a 2 ns NAMD simulation of a DPPC bilayer simulation (Figure 9). Load the file `dancing.psf` and then loading the trajectory `dancing.dcd` into the same molecule. Select several lipids to view (i.e. `segid L200 L225 L250 L275`) and play the trajectory to see the movements of these lipids. Do the lipids all have a single structure or many different structures? Describe the lateral movement of each lipid compared with the movements of just the tail and head groups.

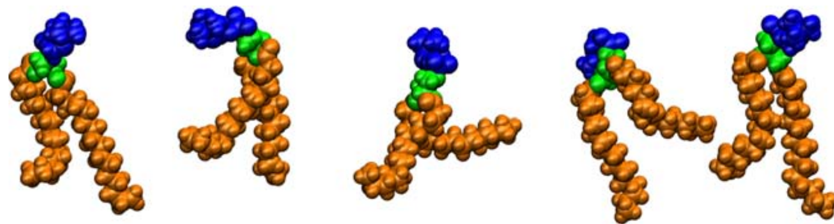


Figure 9: Five "average" lipids from a DPPC bilayer molecular dynamics simulation using NAMD. Notice that the lipids have assumed a wide range of shapes and orientations. (View the dynamics of these five DPPC lipids by loading the VMD saved state `dancing.vmd`.)

3.3 Lipid Phases

Even within the realm of lipid bilayers, there are a number of different 'phases' that the lipids can take, analogous to the classical solid-liquid-gas progression of most substances. In the case of lipids, the three most important states are the liquid (L_α), gel (L'_β), and crystal states. Depending on the lipid in question, other states, such as hexagonally packed tubes, may occur [17].

Equilibrated structures of the three most common phases of lipids are shown in Figure 10. The crystal state of the lipid, as its name implies, shows a perfectly ordered arrangement that is repeated through space. In contrast, the gel phase shows a bit more disorder in the tail region and a number of different head group conformations, and the fluid phase shows a liquid-like disorder throughout the entire lipid. Note that even in the L_α phase the bilayer cannot be properly termed a 'liquid' because it is still held in its planar shape by the hydrophobic effect (and the head groups are still kept on the outside of the bilayer). However, the orientations of individual lipids are essentially random within this constraint. A phase diagram of the lipid DLPC is shown in Figure 11, illustrating the liquid-gel transition and the differences in packing between these two states. The liquid state is by far the most physiologically relevant, and most biological membranes will closely resemble this state, although their exact physical properties may be altered by the inclusion of proteins and cholesterol. The gel state has, however, been implicated in the formation of blood clots, and in lipid rafts that commonly form around some classes of membrane proteins.

Lipid rafts are specialized structures that comprise a small but very important part of animal cell membranes. These regions of the membrane contain increased concentrations of sphingolipids and cholesterol, which causes a gel phase to become more favorable. As a result, these regions are much less mobile, to the point that they survive the destruction of the rest of the bilayer by detergents. Physiologically, lipid rafts are known to be involved in cholesterol transport, and are also the sole location of several important signaling receptors.

Blood clotting also involves a transition of platelet membranes from the liquid to gel phase, and it has recently been shown that the early stages of this process are similar to lipid raft formation. Upon injury, tissue damage causes the release of thrombin, which is the beginning of a long protein cascade that forms a plug to stop the bleeding; this plug is created when

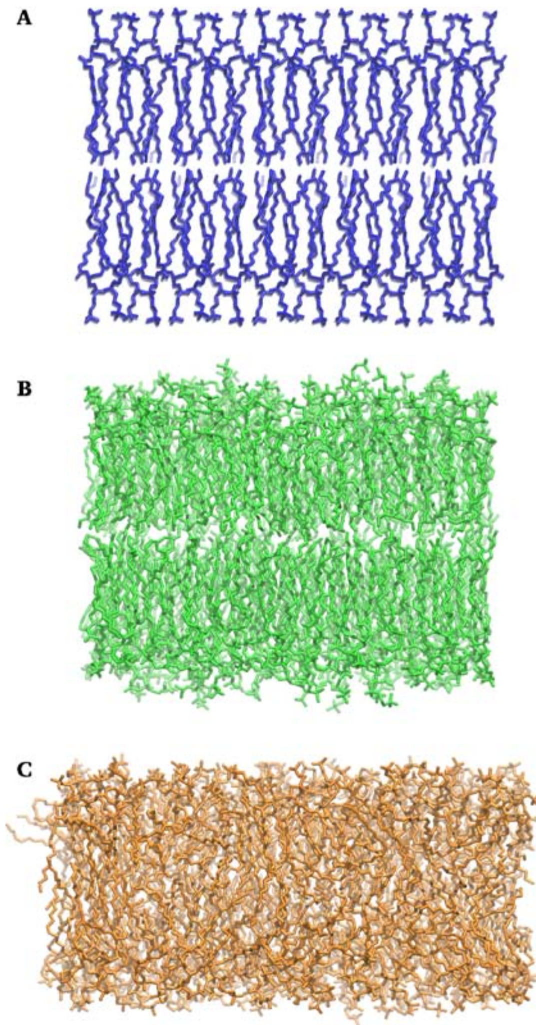


Figure 10: Structures of the crystal (A), gel (B), and liquid (C) phases of a DPPC bilayer, calculated through long term MD equilibration ([9]). PDB structures are in the files `crystal.pdb`, `gel.pdb`, and `fluid.pdb` from the `3.3-phases` subdirectory.

the membranes small blood cells called platelets transition into the gel phase (for more information on this transition, see [5]). Cold-induced transition to the gel state of isolated platelets is a major factor in the difficulty of storing platelets for transfusion, and active research is underway to suppress this.

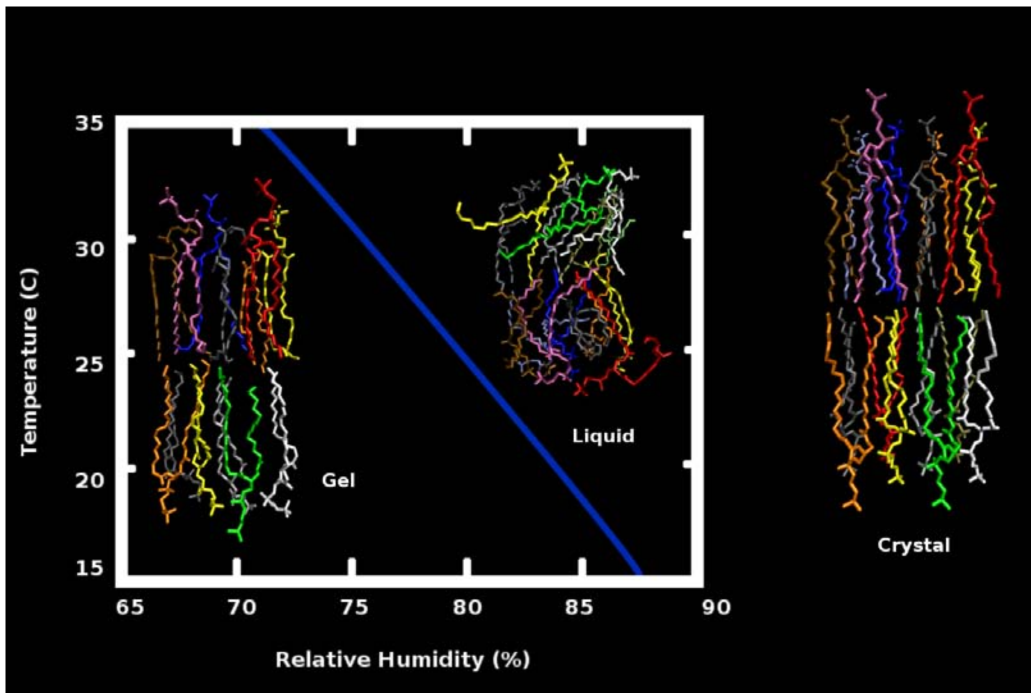


Figure 11: Phase diagram of the lipid DPPC, showing the transition between the gel and liquid phases. A diagram of the crystal phase is shown at right for comparison. Structures were obtained from [9].

Exercise 3: Physical Properties of Lipid Phases. Load the structures of the three lipid phases illustrated in Figure 10 into VMD from the files `crystal.pdb`, `gel.pdb`, and `fluid.pdb` in the `3.3-phases` subdirectory. Each structure contains a total of 200 DPPC lipids. Use VMD to calculate the height of the bilayer, surface area per lipid, and density of the bilayer for each of these phases. You should use the command `measure minmax $sel` to obtain the approximate dimensions of the bilayer in each case (be sure that `$sel` contains only lipids); observe that the bilayer is oriented in the x-z plane, with its height along the y axis. Explain possible reasons for the differences in these characteristics that you see in these states, based on the packing of lipids in each phase. Compare the values you obtained with experimental values [11].

3.4 Lipid Diffusion Through and Across the Bilayer

Biological membranes are not rigid structures, and diffusion both in and across the bilayer is possible for small, apolar molecules. Within the plane

of the membrane, most lipids have diffusion coefficients on the order of 10^{-8} cm^2/s. Simple diffusion in two dimensions is governed by the equation:

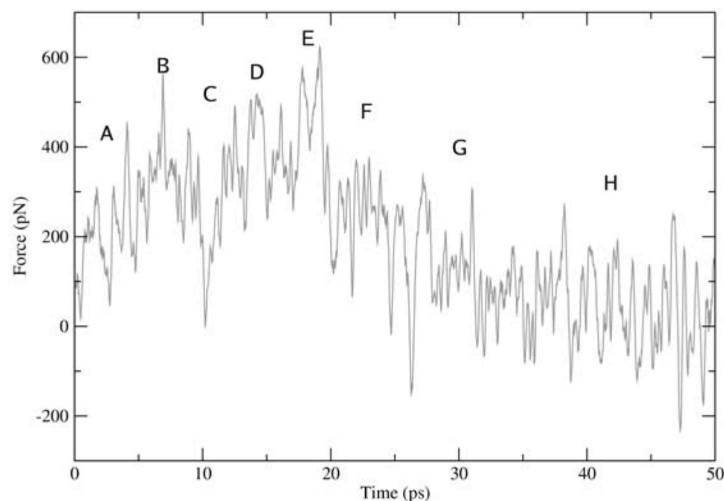
$$\langle r^2 \rangle = 4Dt \tag{1}$$

In this equation, $\langle r^2 \rangle$ is the expected value of the square of the distance a given particle has travelled from its starting point. This means that it would be expected to take on the order of 10 ms for a lipid to diffuse 100 nm in the plane of the membrane (see Appendix I for a detailed derivation of the diffusion law and the calculations involved for lipid diffusion). This is not an unreasonably long time, and indeed lipid diffusion and mixing are very rapid on biological time scales. In the world of simulations, however, this is an incredibly lengthy process; a 'long' simulation of large biological systems will generally run for a few tens of nanoseconds, and on such time scales one would expect lipids to diffuse a distance on the order of a few angstroms.

Exercise 4: Lipid Diffusion in the Bilayer. The results of a molecular dynamics simulation of a lipid bilayer are contained in the files `dppc.psf` and `dppc_equi.dcd` from the `3.4-diffusion`. The mean square deviation of phosphorus atoms as a function of time has been extracted from this simulation, and can be found in the file `phosphate_msd.dat` as a set of time/value pairs, with time in picoseconds. Use this data and equation 1 to calculate the observed diffusion coefficient of lipids in the bilayer in these simulations. Is this result reasonable? Observe the data from this file in graph form; how long from the start of the simulation does it take for the system to begin observing the diffusion law? The deviation of five selected phosphorus atoms from their starting position as a function of time has been placed in the files `phosphate#.dat`, with # an integer between 1 and 5. Do any of these individual atoms come close to obeying the diffusion law?

In addition to diffusing within the bilayer, it is also conceivable for lipids to flip *across* it. In a pure bilayer, however, such an event is almost inconsiderably rare; bilayers require around a few weeks to equilibrate in this way in the absence of proteins. However, lipid flipping is known to be very important for some signaling processes, as well as in properly regulating the leaflet composition of the bilayer. In living cells specialized enzymes called *flippases* are known to facilitate this process, bringing lipid flipping into a much more reasonable time scale. The flipping of lipids and other hydrophobic molecules contained in the bilayer is important for processes including proper lipid composition, blood clotting, and resistance of bacteria and cancer cells to a number of drugs [14, 7].

Exercise 5: Analysis of Steered Molecular Dynamics. The graph below shows the results of an SMD run where a single lipid was pulled out of the bilayer at a velocity of 0.0001 Å/fs. The y axis shows the force applied on the lipid to maintain this velocity; positive forces are in the direction pulling the lipid away from the rest of the bilayer. Major points of interest in the simulation are labeled A-H on the chart. Load the simulation into VMD using the files `bilayer-pull.pdb` and `bilayer-pull.dcd` from the `3.5-pulling` directory, and use the information from this trajectory to identify the events that lead to the lettered force changes. Be sure to pay attention to both hydrophillic and hydrophobic effects. Note that the snapshots in the dcd are taken every 500 fs.



The results from such experiments allow for analysis of the important factors in membrane integrity, and illustrate that polar contacts, van der Waals interactions, and the hydrophobic effect all play roles. If this simulation were repeated many times, the results could be used to estimate the free energy of partition of the lipid between water and the bilayer interior. Nonetheless, even the results of a single pulling experiment can be useful in qualitatively understanding the interactions between lipids in the bilayer.

3.5 Lipid-Lipid Interactions

As discussed in Section 3.1, the primary impetus for lipid bilayer assembly is the hydrophobic effect, which drives the tails of lipids to aggregate and escape an aqueous medium. Aside from the hydrophobic effect, there are a number of other interactions which can take place between neighboring lipids to stabilize the bilayer. In addition to the hydrophobic effect, the long hydrophobic tails of lipids are attracted through van der Waals interactions, which although weak can have a significant impact when there are large areas of surface contact between molecules. In the hydrophillic head region, most biological lipids have a negatively charged phosphate group attached to a second positively charged or polar group. These regions of the lipids interact favorably with water, but interactions between positive and negative groups on the heads of neighboring lipids can also have a powerful influence on lipid structure; these interactions have an impact on the surface area of each lipid and cause the heads of most lipids to bend rather than staying perpendicular to the membrane surface.

4 The Bilayer and Embedded Proteins

The ability of the membrane bilayer to block transit of large molecules into and out of the cell is precisely what makes it useful for separating and protecting the cell from its environment. If all cells were surrounded only by a pure lipid bilayer they would be unable to obtain food, maintain an osmotic balance with the environment, or communicate with other cells. Living cells accomplish all these tasks using *transmembrane (TM) proteins*: a class of proteins which span the entirety of the lipid bilayer. Such proteins comprise an average of 50% of the mass of biological membranes, and are responsible for such tasks as selectively allowing certain substances through the membrane, pumping important ions and small molecules into or out of the cell, and activating intracellular signaling pathways in response to extracellular signals.

4.1 Membrane Protein Structure and Lipid Packing

Most soluble proteins follow a well-studied formula in their structure, consisting of a hydrophobic core that lends stability to the folded state, and a

mostly hydrophillic outer surface to ensure that the protein is water soluble. Obviously, this configuration is completely inappropriate for membrane bound proteins, since a large fraction of their surface is in contact with the hydrophobic membrane interior. Instead, membrane proteins will tend to either look like inverted globular proteins, with a hydrophobic exterior and hydrophillic interior, or be completely hydrophobic. This can be a bit misleading though; membrane proteins don't always form a globular hydrophillic core in the same way that soluble proteins have a hydrophobic core. Proteins requiring a transmembrane domain solely for structural reasons (such as anchoring a complex to the membrane) often do not have a hydrophillic portion to their transmembrane domain at all, although obviously some transmembrane proteins, such as ion channels, require a hydrophillic channel for their function. For the purposes of this section we will consider three exemplary transmembrane proteins, each of which requires a hydrophillic central region to function: BtuCD, a bacterial protein which transports vitamin B12 across the membrane; GlpF, a water and glycerol channel; and SecY, a pore which regulates the secretion of proteins. The hydrophobic and hydrophillic patterns along the exterior of these proteins can be seen in Figure 12.

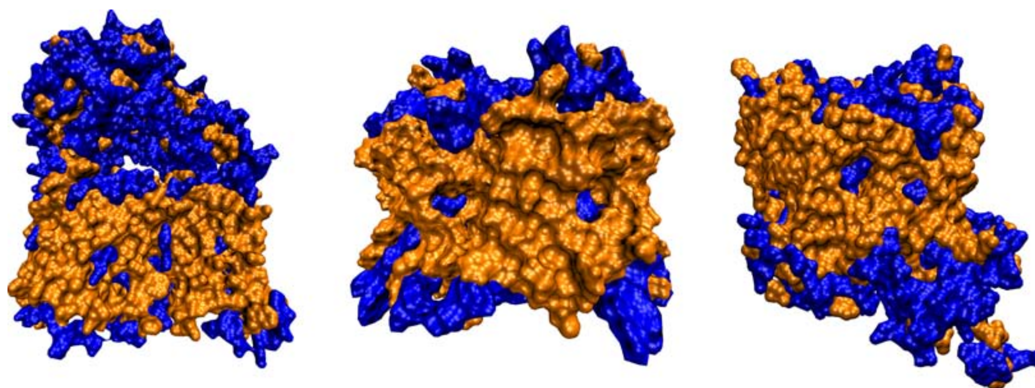


Figure 12: Patterns of hydrophobic and hydrophillic amino acids in three membrane bound proteins: BtuCD (top), GlpF, and SecY. In each case, hydrophobic amino acids are shown in orange and hydrophillic amino acids in blue. Note that BtuCD is a complex consisting of a membrane bound (bottom) and soluble (top) domain; the other two proteins also have a few soluble loops at the top and bottom where the protein is exposed to solution. Structures for the proteins shown in Figures 12, 13, and 14 can be found in the files `btucd.pdb`, `glpf.pdb`, and `SecY.pdb` in the `4.1-membprot` directory; you can use these to take a closer look at the patterns of hydrophobicity in these proteins.

The site of the bilayer-aqueous interface is a special case in membrane-protein interactions because of the presence of charged lipid head groups. Certain amino acids that are capable of forming favorable interactions with the head groups are found with extremely high frequencies at this interface, to the point that many computer programs use them to predict the end of membrane spanning regions of proteins. The residues most commonly found at these positions are tryptophan, which interacts favorably with the phospholipid head groups; and arginine, which is thought to 'snorkel' so that its aliphatic chain remains in the hydrophobic part of the bilayer but its head group is positioned to interact with the lipid phosphate. The locations of tryptophan, tyrosine, and arginine residues in the protein BtuCD are illustrated in Figure 13; note the presence of bands of these amino acids at the membrane-solution interface.

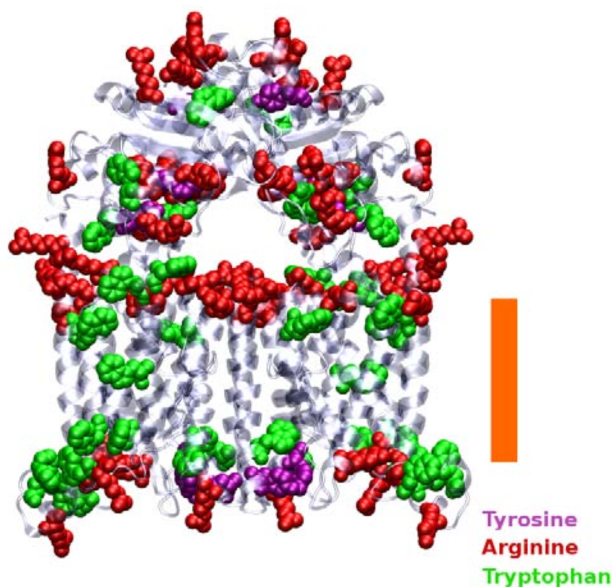


Figure 13: Side view of BtuCD with selected residues highlighted. The orange stripe along the right side of the figure indicates the approximate extent of the membrane.

Viewing the GlpF structure (Fig. 14) illustrates another important principle of transmembrane protein structure. The diagram is colored by secondary structure, emphasizing that the transmembrane region is composed entirely of α-helices. This is not at all unusual for a transmembrane protein. The

only common alternative to a heavily helical transmembrane region is the β-barrel: a tube composed of a large array of β sheets folded into a cylinder. The need for an ordered secondary structure in the membrane environment stems from the fact that every amino acid has at least two polar groups on its main chain. Thus, placing the protein main chain in a hydrophobic environment would normally carry an energetic penalty. In secondary structure elements, though, all main chain polar groups except those on the edges are involved in hydrogen bonds, and as a result these structures will be much more stable in a lipid environment than in extended conformations.

Transmembrane proteins containing a water-filled cavity or channel often contain *amphipathic* α-helices: helices which are hydrophillic on one side and hydrophobic on the other, as seen in the inset of Figure 14. Such an arrangement is favorable because the helices must be simultaneously interacting with water and lipid on opposite sides. This pattern is distinctive enough that it can actually be used in searching sequences for transmembrane proteins; regions of alternating hydrophobic and hydrophillic amino acids with a period of 3 or 4 very often indicates the presence of an amphipathic α-helix.

4.2 Importance of the Bilayer in Membrane Protein Structure

The presence of lipophilic amino acids on the outer surface makes this part of the protein soluble in lipid and allows the bilayer to pack closely around the protein. Since one of the main functions of transmembrane proteins is to provide *selective* transport across the membrane, a close seal here is absolutely essential.

In addition to the need for a tight seal between membrane and protein, the interaction between proteins and lipids can have an effect on both membrane protein function and the ability of proteins to be incorporated into the membrane. Because the membrane must deform to accomidate the inclusion of a protein, the ease of protein inclusion, as well as the rate of aggregation of membrane proteins, is highly dependent on the strength of lipid-lipid interactions and the ability of the membrane to curve in resonse to variable sized inclusions [13]. In addition, membranes will exert a tension on the sides of a protein which can enable or prevent function [6, 4]. This is especially important for so-called mechanosensitive proteins, which are channels that open in response to changing pressures within the cell by measuring the tension in

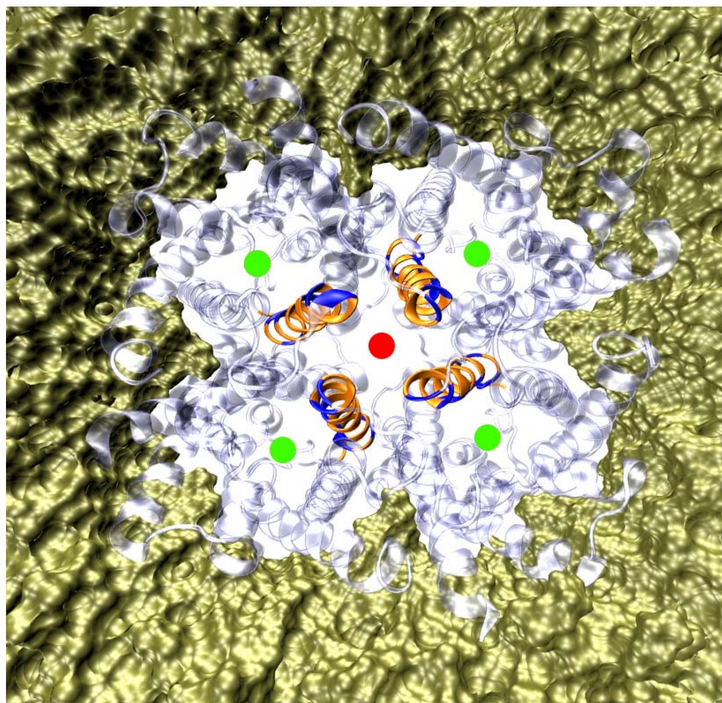


Figure 14: Diagram of the α-helices of GlpF viewed from the outside of the cell looking in, with the protein in cartoon representation. One transmembrane helix from each GlpF monomer has been highlighted, with hydrophobic amino acids colored orange and hydrophillics colored blue. To aid in understanding the organization, the four water-filled pores of the protein are highlighted with green dots, and a red dot indicates the central 'pore' between the monomers.

the membrane [16].

To see an example of lipid packing around a membrane protein, open the VMD savestate `aqp-packing.vmd` from the `4.2-structure` directory. The structures shown here are snapshots of the aquaporin crystal structure in a POPC bilayer; in the initial snapshot lipids are left in their normal structure, with a simple hole cut out from the bilayer to make room for the protein, whereas in the final state, they have been allowed to adjust to the structure of the protein. In the final snapshot, one can observe that the lipids have packed around the protein, changing from a fairly flat interior surface to one that has adapted to the outer surface of the protein to allow for enhanced interaction. This is best seen by looking at a single TM helix

of the protein (for example, the region between residues 210 and 235) and the lipid immediately adjacent to this region. Observe how the lipid actually forms indentations for hydrophobic side chains to fit.

Exercise 6: Energetics of Lipid Packing. The solvent accesible surface area (sasa) of a selection in VMD can be calculated with the `measure sasa` command. The structures for aquaporin that you have been given contain only protein and lipid, which means that the sasa calculated for this structure is the area interacting with water. To calculate the area accessible to water for a given subset of this structure, use the command `measure sasa 1.4 $all -restrict $selection`, where $all is an atom selection for all atoms in the system, and $selection is the atoms you actually want the sasa of. What is the difference in the water exposed surface area of *hydrophobic side chains* between the pre- and post-equilibration structures of aquaporin given here? Assuming that the energy of partition of a hydrophobic side chain from lipid to water is 0.04 $\frac{\mathrm{kJ}}{\mathrm{mol}\cdot\mathrm{\AA}^2}$, what is the difference in energy due solely to the solvation of hydrophobic side chains between these two structures? Is this enough energy to be significant at room temperature? Why?

Appendix I: Diffusion

To determine how far a lipid molecule undergoing free diffusion is likely to travel in a given period of time, we can model it as a single particle freely diffusing in two dimensions. For the sake of simplicity, we assume that it moves in discrete steps, once every τ units of time, and thus will take its jth step at $t = t_0 + j\tau$. Note that this will essentially duplicate actual physical circumstances in the limit $\tau \to 0$. We consider each of these steps to be of a fixed length L, in a random direction.

For this freely diffusing particle, we wish to calculate $\langle x^2 \rangle$, the expected magnitude of the square of the distance that the particle has travelled from the starting point, as a function of t. This can be determined by first noting that $\vec{R}_N$, the position vector of the particle with respect to the origin after N steps, is equal to $(\vec{R}_{N-1} + \vec{L})$; where $\vec{L}$ is a random step of length L. We seek $\langle x^2 \rangle$ after N steps, which is equivalent to $\vec{R}_N \cdot \vec{R}_N$. Combining these identities,

$$\begin{aligned} \langle x^2 \rangle &= \vec{R}_N \cdot \vec{R}_N = (\vec{R}_{N-1} + \vec{L})^2 \\ &= R_{N-1}^2 + 2\vec{R}_{N-1} \cdot L + L^2 \end{aligned}$$

Considered over many trials, $\langle \vec{L} \rangle = 0$, since the particle is equally likely to travel in any direction. This in turn means that over many trials $\langle R_{N-1} \cdot L \rangle = 0$, and thus, $\langle x^2 \rangle = R_{N-1}^2 + L^2$. Note that for N=1, $\vec{R}_{N-1} = 0$, since the particle starts at the origin, so $\langle R_1^2 \rangle = L^2$. For step 2, then,

$$\langle R_2^2 \rangle = \langle R_1^2 \rangle + L^2 = L^2 + L^2 = 2L^2$$

Following this pattern inductively, we find that after N steps,

$$\langle x^2 \rangle = \langle R_N^2 \rangle = NL^2$$

Since we have defined each step as lasting τ units of time, $N = t/\tau$, and thus the mean square deviation can be defined as $\langle x^2 \rangle = NL^2 = \frac{tL^2}{\tau}$.

In practice, the constant $\frac{L^2}{N}$ is replaced by $2nD$, where D is defined as the diffusion coefficient of the molecule (which can be empirically determined) and n is the number of dimensions that the motion occurs in. The diffusion coefficient is useful not only in mean squared deviation calculation, but also in much more general diffusive properties. The inclusion of this factor allows a variety of different processes for a given system to be connected. A more detailed discussion of the physical principles involved in diffusion within the bilayer, as well as a description for how the two-dimensional diffusion constants of lipids may be measured, is given in [3].

For the case of the lipid bilayer, $n = 2$, since diffusion occurs in two dimensions, and the biologically relevant diffusion coefficient D has been measured to be on the order of 10^{-8} cm^2/s. For one to calculate, for example, the expected distance for one lipid molecule to travel in 10 ms (as was done in the text), the calculations are as follows:

$$\begin{aligned}\langle x^2 \rangle = 4Dt &= 4 * 10^{-8}\text{cm}^2/\text{s} * 10\text{ms} \\ &= 4 * 10^{-8}\text{cm}^2/\text{s} * 10^{-2}\text{s} \\ &= 4 * 10^{-10}\text{cm}^2\end{aligned}$$

Taking the square root of this to obtain the root mean square distance travelled, one obtains $\sqrt{4 * 10^{-10}\text{cm}^2} = 2 * 10^{-5}$ cm = 200 nm.

References

[1] B. Alberts, D. Bray, J. Lewis, M. Raff, K. Roberts, and J. D. Watson. *Molecular Biology of the Cell.* Garland Publishing, third edition, 1994.

[2] G. S. Ayton and G. A. Voth. Mesoscopic lateral diffusion in lipid bilayers. *Biophysical Journal*, 87:3299–3311, 2004.

[3] A. Brünger, R. Peters, and K. Schulten. Continuous fluorescence microphotolysis to observe lateral diffusion in membranes: Theoretical methods and applications. *J. Chem. Phys.*, 82:2147–2160, 1985.

[4] R. S. Cantor. Lipid composition and the lateral pressure profile in bilayers. *Biophys. J.*, 76:2625–2639, 1999.

[5] J. Craft, K. Epps, S. Brancato, and A. Serfis. Incorporation of blood clotting factor x into phospholipid model membranes; fluorescence microscopy imaging and subphase effects surrounding the lipid phase transition region. *J. Colloid and Interface Science*, 268(1):181–187, 2003.

[6] N. Dan and S. A. Safran. Effect of lipid characteristics on the structure of transmembrane proteins. *Biophys. J.*, 75:1410–1414, 1998.

[7] W. T. Doerrler and C. R. H. Raetz. Atpase activity of the msba lipid flippase of *Escherichia coli*. *J. Biol. Chem.*, 39:36697–36705, 2002.

[8] R. B. Gennis. *Biomembranes: Molecular Structure and Function*. Springer, first edition, 1989.

[9] H. Heller, M. Schaeffer, and K. Schulten. Molecular dynamics simulation of a bilayer of 200 lipids in the gel and in the liquid-crystal phases. *J. Phys. Chem.*, 97:8343–8360, 1993.

[10] G. Karp. *Cell and Molecular Biology: Concepts and Experiments*. Wiley, third edition, 2002.

[11] J. F. Nagle and S. Tristram-Nagle. Structure of lipid bilayers. *Biochim. Biophys. Acta*, 1469:159–195, 2000.

[12] D. L. Nelson and M. M. Cox. *Lenhinger Principles of Biochemistry*. Worth Publishers, third edition, 2000.

[13] V. Pata and N. Dan. The effect of chain length on protein solubilization in polymer-based vesicles (polymersomes). *Biophys. J.*, 85:2111–2118, 2003.

[14] Y. Romsicki and F. J. Sharom. Phospholipid flippase activity of the reconstituted p-glycoprotein multidrug transporter. *Biochemistry*, 40:6937–6947, 2001.

[15] D. Voet and J. G. Voet. *Biochemistry*. Wiley, second edition, 1995.

[16] P. Wiggins and R. Phillips. Membrane-protein interactions in mechanosensitive channels. *Biophys. J.*, 88:880–902, 2005.

[17] P. Yager, J. Chappell, and D. D. Archibald. When lipid bilayers won't form liposomes: Tubules, helices, and cochleate cylinders. In B. P. Gaber and K. R. K. Easwaran, editors, *Biomembrane Structure and Function: The State of the Art*, pages 001–019, 1992.