

Chapter 3. Development of Computational Methods

3.1. Introduction

This chapter presents the computational methods employed for the results presented in Chapters 4 and 5. These methods are presented generally, and the later chapters describe specific deviations from this procedure as necessary.

Briefly, the goal of the thesis is to understand the physical and chemical binding properties of lysophospholipids, sphingolipids, and associated ligands with G-protein coupled receptors (GPCRs). Because x-ray crystallographic structures are not available for these structures, we build each protein virtually and the computational techniques determine the tertiary structure, as well as the binding between the GPCRs and their ligands. Therefore, each GPCR must be constructed in the computer before ligands may be docked and explored. Sections 3.2 and 3.3 describe the procedure for building the proteins using TM2ndS to assign the helical regions and MembStruk to determine the helical packing. The outlined procedure follows the development of the LPA₂ structure.

Once the protein is successfully constructed, we consider potential binding sites and perform molecular-dynamics simulations on the interaction of ligands with these binding sites. Sections 3.4 and 3.5 outline the procedures for isolating the binding site of the endogenous ligand as well as identifying the binding modes of other agonists. The procedure for building the explicit lipid and water solvent systems for long time-scale molecular dynamics is provided in Section 3.6

3.2. Force Field3s and Charges

We modified the generic DREIDING force field¹ to include parameters for the phosphate group, and used it for minimizations of the proteins and the ligands. Protein atoms are assigned CHARMM22² charges, while ligands are assigned HF/6-31G** Mulliken charges calculated by Jaguar v6.5. As previously mentioned, deviations from this procedure are presented, as necessary, in later chapters.

3.3. Prediction of the Transmembrane Barrel

TM2ndS was developed by the Goddard group to quickly identify the amino acid sequence segments that belong to the seven transmembrane regions. Trabanino and others detail the technique of predicting the sequence range of the seven transmembrane helices,³⁻⁵ but a brief overview of the method, known as TM2ndS, follows here.

An NCBI BLAST⁶ search provides a library of sequences to use as a basis of comparison for the hydrophobicity of the human LPA₁ (primary accession number: Q92633) sequence. As shown in Appendix A.1, we selected a series of GPCR sequences with wide-ranging homologies to LPA₁. Using default parameters, CLUSTAL W⁷ compares these sequences by creating a set of pair-wise alignments, shown in Appendix A.2. The alignment developed from the sequence of LPA₁ was used for the transmembrane predictions of LPA₂.

It is expected that the transmembrane helices are less polar than the loop regions, which are exposed to an aqueous environment. In order to isolate the amino acid segments that constitute the transmembrane helices, a measure of relative hydrophobicity

is determined for each residue in LPA₂ by comparing the Eisenberg hydrophobicity of the residue to an average of the hydrophobicities of each sequence in the alignment created from the BLAST search. A window size of 12 residues was used for this analysis. The Eisenberg scale provides a normalized thermodynamic measure of the hydrophobicity for each amino acid residue where higher hydrophobicity yields a more positive Eisenberg value.⁸ Table 3.1 lists the original and normalized Eisenberg value for each amino acid residue. Previously published capping rules, based on known trends of helix breaking, determine the ends of each helix.⁴ Generally prolines, charged residues, and/or two consecutive glycine residues are found at the termini of helices.

Following the TM2ndS generation of the seven helices, the generalized MembStruk method^{5,9,10} determines the correct packing of the helices. Although the subject of recent publications, this research required significant improvements to the method which have not been described previously. These improvements are described in detail below and will become the subject of future communications.

There are three major components to the MembStruk method:

- 1) Construction and relaxation of each transmembrane (TM) helix (Section 3.3.1)
- 2) Packing and translation of the helices relative to one another (Section 3.3.2)
- 3) Optimization of the helical rotations (Section 3.3.3)

3.3.1. Construction and Relaxation of the Transmembrane Helices

TM2ndS defines and constructs the helices as canonical alpha helices with extended side chains. Structure minimization optimizes each helix using conjugate gradients, and 50 ps of NEIMO molecular dynamics (MD) at 300 K.¹¹ Added counter-ions prevent unnatural salt bridges between charged amino acids. Natural kinks in the helices develop upon relaxation at prolines and, occasionally, glycines.

3.3.2. Packing and Translation of the Helices Relative to One Another

A low-resolution (7 Å) crystal structure of frog rhodopsin¹² serves as a template for packing the seven helices into the “TM bundle” and orienting the tilts of each helix. This template has a topology that can be described as six helices packed in a circular shape that encompasses helix 3 (Figure 3.1). The mid-plane of the lipid bilayer defines the Cartesian $z = 0$ plane. The x - and y -coordinates of each main chain atom are defined by the template structure, while the z -coordinate is determined by the position of an atom relative to the “center” of its respective helix. We defined the center of a helix to be the geometric center adjusted to the peak of hydrophobicity, according to the normalized Eisenberg¹³ hydrophobicity scale given in Table 3.1.

Physically, one would expect the most hydrophobic region of a helix to be submerged in the most hydrophobic region of the cellular bilayer. To provide the most physical model of the protein in the cellular membrane, the helices are oriented so that their hydrophobic centers align on one plane, $z=0$, the middle of the cellular bilayer.

3.3.3. Optimization of the Helical Rotations

Identification of optimal helix packing is a highly coupled problem involving a combinatorial analysis of helix rotations. In order to simplify the problem and minimize computational expense, we approach the problem iteratively. The goal is to identify which face of each helix is solvated by the lipid membrane and which sector is pointing into the barrel.

Initially, the orientation for each helix has the net hydrophobic moment of the middle of the helix pointed outward from the TM barrel. This initial structure undergoes side chain optimization of all residues in the barrel using the SCREAM program, with the associated default parameters developed in our group.

Because of the highly conserved hydrogen-bonding network between helices 1, 2, and 7, the rotation procedure begins with helix 2. We rotate the helix through 360° in 30° increments to find the rotational orientation that maximizes the thermodynamic benefit from solvating the helix in the membrane. This is quantified at each increment by a solvent-exposed surface area penalty given by Equation 3.1. The total barrel penalty, P_{Tot} , is calculated using a hydrophobic scale derived from averaging the Wimley-White¹⁴ and White Biological scales¹⁵, given in Table 3.2. The solvent-accessible surface area (SASA) of each residue is determined, and we calculate the residue penalty, P_{res} , for exposure to the lipid membrane for all residues with a SASA of 45% or larger.

$$\begin{aligned}
 P_{res} &= SASA_{res} \times HP_{res} \\
 P_{Tot} &= \sum P_{res}
 \end{aligned}
 \tag{3.1}$$

Finally, visual inspection of plots of P_{Tot} versus rotational angle determines the limited angle range chosen to proceed to the next optimization.

Those 30°-increment structures that lay in the thermodynamically favored sector undergo side chain optimization using SCREAM on just the polar residues to optimize the hydrogen-bonding network between the helices. For this analysis, Arg, Lys, Asp, Asn, Glu, Gln, Ser, His, Tyr, and Thr are polar residues, but any polar residue that is at either end of a helix is excluded from optimization. During this side chain optimization, multiple energy terms are summed to give a total SCREAM energy, E_{screeam} , shown in Equation 3.2. The individual energy terms included the internal energy, E_{int} , of each “SCREAMed” rotamer in an empty lattice; the interaction energies— van der Waals (V), Hydrogen bonding (H) and Coulombic (C)—between each rotamer and its own backbone (sBB); other backbone atoms (oBB); fixed side chain residues (fSc); and atoms in other SCREAMed residues (ScSc). All energy terms have a scale factor of one, except for the van der Waals fixed side chain, V_{fSc} , term. Because the non-polar residues were fixed, the interactions between the polar side chains and these fixed side chains was not optimal. This term, therefore, contained a significant amount of noise. Only 20% of the V-fSc term is included in E_{screeam} . This factor was derived semi-empirically and did not change the rankings of E_{screeam} for most helices. The helix rotation with the lowest E_{screeam} was fixed in that orientation for optimization of the rest of the helices.

$$E_{\text{screeam}} = \text{Int} + V_{\text{sBB}} + H_{\text{sBB}} + C_{\text{sBB}} + V_{\text{oBB}} + H_{\text{oBB}} + C_{\text{oBB}} - 0.8(V_{\text{fSc}}) + H_{\text{fSc}} + C_{\text{fSc}} + V_{\text{ScSc}} + H_{\text{ScSc}} + C_{\text{ScSc}} \quad (3.2)$$

After fixing the rotational orientation of helix 2, all side chains are re-optimized. Each helix, in the order 7, 1, 3, 4, 5, and 6, is subjected to the rotation/penalty calculation/side chain optimization procedure. Once all seven helices are fixed into an improved position, another iteration of rotation analysis is performed. The thermodynamic penalty is no

longer considered during the second iteration, and polar residue side chain optimization is performed at every 30° increment in order to prevent the elimination of any low-energy structures. All side chains are optimized at the new $\rho=0$, where ρ is the rotational angle of the helix, prior to rotating the helix. Again, the helices were rotated in serial in the order 2, 7, 1, 3, 4, 5, 6, and the energy term E_{rot} characterizes the interactions for the structures. Subsequent iterations can be performed to insure convergence but are not necessary for LPA₂. After all helical rotations converge, the TM barrel is minimized for 150 steps.

At this stage, we consider the *apo* transmembrane region of the GPCR structurally optimized and use it to analyze potential binding pockets and the binding interactions with ligands.

3.4. Determination of Binding Modes

Identification of ligand binding sites in a protein generally consists of three steps: isolation of the potential binding pocket(s), docking conformations of the ligand(s) into the binding pocket, and scoring of the conformations. To maximize the internal surface area of each protein, the procedure requires mutation of bulky residues (TYR, TRP, VAL, LEU, ILE, PHE) in the transmembrane region to alanines prior to isolation of the putative binding regions. The positively charged residues at the intracellular termini of the helices also mutate to ALA to minimize results biased toward biologically unlikely structures. This “alaninized” version of the TM barrel is used for docking. When scoring the docked protein-ligand complexes, the protein mutates back to the wild-type structures.

3.4.1. Isolation of Potential Binding Pockets

The Putative Active Sites with Spheres (PASS)¹⁶ method was used to identify potential binding sites within the transmembrane region. Under the assumption that for binding a ligand must, at a minimum, sterically fit into a cavity within a protein, the PASS algorithm uses the geometric topology of the protein to identify putative binding pockets.

The algorithm covers all exposed surfaces of the protein with spheres. Clusters of these spheres define void spaces within the protein that could house an arbitrary ligand. To be selected as a member of a potential binding site, each sphere must be tangentially in contact with the van der Waals radii of three atoms, have a buried surface area above a cut-off threshold, and not be within 1 Å of a more buried sphere. All spheres that remain after these three filters have been applied are considered to be inside a potential binding site.

Within these clumps of probe spheres, the PASS algorithm identifies “active site points (ASPs), points at the center of putative binding sites. This selection process defines which clusters of spheres are independent from another, and therefore, is the final determination of each putative binding site. These ASPs are based solely on the geometry of the protein, so that an arbitrary ligand, regardless of its hydrophobic or hydrophilic properties, could potentially bind in the cavity.

3.4.2. Docking of Ligand Conformations

Upon isolating a set of putative binding sites within a protein, the endogenous ligands are docked into each sphere set using the anchor search method in DOCK 4.0.¹⁷ Spacing of 0.25 Å and a dielectric constant of 2.5 are used for building the grids. ModMSCDock, a program developed in the Goddard group, which integrates our methods with DOCK 4.0, was used for running all docking calculations.

The *cis* isomer of LPA 18:1 was docked into each of the sphere sets generated by PASS. The phospholipid ligands were docked with a formal charge of -1, and the methylene groups in the alkyl chains are in the staggered conformation.

3.4.3. Preliminary Scoring of Docked Ligands

After an ensemble of *cis* ligand conformations has been docked into the sphere region, the collection is sorted by buried surface area. Given the hypothesis that binding of ligands in the TM region dominates agonism of GPCRs, we assume ligands are docked better when they have a greater buried surface area. The Connolly surface area quantifies this, and the ensemble for each sphere set is pared down to a maximum of 30 conformations that have at least 70% buried surface area. These 30 (or fewer) conformations are minimized inside the protein to a root-mean-square deviation (RMSD) of 0.1 (kcal/mol)/Å or a maximum of 100 steps, and Delphi^{18,19} calculates the binding energy using Poisson-Boltzmann solvation.

Each of the 30 minimized ligand conformations from a PASS sphere set is matched into the TM barrel of the wild-type protein structure, and all residues within

5.0 Å of the docked ligand were SCREAMed to optimize their side chains. Then the ligand and all residues within the 5.0 Å binding pocket of the ligand are minimized to 0.5 (kcal/mol)/Å RMSD force. Following the minimization of the binding pocket region, the entire protein-ligand complex is minimized to an RMSD force of 0.2 (kcal/mol)/Å.

Total energy of the ligand-protein complex allows quantitative ranking of the minimized complexes. High total energy eliminates some complexes as candidate structures, and the ensemble that is carried on to the next level of structure completion must satisfy certain requirements. The complex must be within 10 kcal/mol of the lowest energy complex; have a root-mean-squared deviation greater than 0.6 Å from all other members of the ensemble; or, if not within 10 kcal/mol of the lowest energy complex, be one of the five lowest energy complexes and still satisfy the diversity criterion. These remaining structures are further refined, scored, and possibly eliminated by the procedure outlined in section 3.4.4. In addition, any structure with the phosphate head bound to ARG or LYS residues at the intracellular end of the bundle are discarded.

3.4.4. Completion of the Protein Structure

For all remaining candidate complexes, a continuous configurational biased (ccb) Monte Carlo algorithm adds both extracellular and intracellular loops to the ligand-docked protein.²⁰ For simplicity, alanines replace proline residues during loop building. Once the loop growth is complete, we mutate the relevant alanines to proline residues. After the addition of an extracellular loop and an intracellular loop, the two loops are minimized to an RMSD of 0.3 (kcal/mol)/Å while the rest of the protein remains fixed. Only one Monte-Carlo sample is created for each loop on each protein structure. After

scoring the loop-containing complexes, each of the top ligand conformations are matched into each of the top protein structures with loops. This insures that each element of the looped protein/ligand matrix is computed, to maximize the ensemble of potential complex structures.

3.4.5. Calculation of Binding Energies and Interaction Energies

Upon identification of the lowest energy looped-protein/ligand complex, one cycle of annealing dynamics (up to a temperature of 600 K) relaxes the entire binding pocket. Then, we convert ARG, LYS, HIS, ASP, and GLU to their neutral counterparts to insure that Coulombic interactions are not overly weighted and that salt bridges between residues can be interrupted with favorable ligand-protein interactions. The QM charges for the ligand are recalculated after fully protonating the phosphate head. The entire complex is minimized to an RMSD of 0.1 (kcal/mol)/Å. The binding energy is calculated using Equation 3.3. Calculations of the ligand energy in solvent utilize the Analytical Volume Generalized Born (AVGB) method for solvation and water as the solvent. The binding energies are calculated out of a sense of tradition, not because we obtain any information from them. As discussed in Chapter 2, the low solubility of lipids in water necessitates that other proteins transport these ligands through the body. To get a true binding energy, one should use the binding energy of LPA and albumin (or S1P and high density lipoprotein) for $E_{solvent}$. Unfortunately, our survey of the literature did not yield that number. Others have published approximate binding affinities of LPA for albumin, though. These values are in the low nanomolar range, similar to the EC₅₀ values seen for LPA and its G protein-coupled receptors.²¹

$$B.E. = E_{complex} - E_{solvent} \quad (3.3)$$

In addition to calculating the binding energy, we also directly calculate the interaction energy of the ligand with all residues within 5.0 Å of the ligand. These binding and cavity energies are not used for determining the best complex for the endogenous ligand but are useful for comparison of the binding of different ligands (described below) with the experimental data from the literature. Generally, the cavity energy is more positive than the binding energy and correlates with experimental results better than the binding energy. Use of the cavity energy does have some limitations, though, especially when comparing ligands of different sizes. As the cavity energy is the sum of all of the interactions between the ligand and the binding pocket, larger ligands will inherently have stronger cavity energies.

3.5. Binding of Other Agonists

All other agonists are docked into the protein as it exists after the annealing dynamics described in Section 3.4.5. In order to insure that the protein remains in the active state we have determined through binding of the endogenous ligand, the other ligands are built through perturbation of the endogenous ligand. The 3D-Sketcher in Cerius2 modifies LPA *cis*-18:1 into other members of the LPA lipid family and into members of a series of *N*-acyl ethanolamide phosphatidic acid (NAEPA) derivatives that show a range of activity in LPA₂. Tables 3.3 and 3.4 provide structures of all agonists examined in this study. This family of NAEPA derivatives was chosen because activation by these ligands is, to some degree, stereoselective. After building the new ligands from LPA *cis*-18:1,

the ligand and protein structures are merged together and all side chains within 5.0 Å of the ligand are optimized using SCREAM. The binding pocket is minimized for 100 steps, or to an RMSD of 0.3 (kcal/mol)/Å. All charged residues are converted to their neutral counterparts, and the entire complex is minimized to an RMSD of 0.1 (kcal/mol)/Å. The methods described in Section 3.4.5 are used to calculate binding and interaction energies.

3.6. Molecular Dynamics Studies

Full explicit-solvent, long time-scale molecular dynamics are run on the endogenous ligand/LPA₂ complex to study global structural changes in the protein upon relaxation, as well as localized changes in the binding pocket. Both explicit lipid and water are included in these simulations. The complex has charged residues and the same structure seen after the annealing dynamics described in Section 3.4.5.

To build the lipid bilayer and the water boxes above and below the membrane, the `solvate` plugin in VMD 1.8.5 was used.²² Palmitoyloleoylphosphatidylcholine (POPC) is the lipid used for building these membranes. Sodium or chloride ions are used to make the system charge neutral.

After the explicit-solvent system is built, all minimization and dynamics calculations are run with NAMD 2.6.²³ The water and lipid molecules are minimized for 1000 steps while keeping the protein and ligand fixed. The solvent is then relaxed for 100 ps, again with the protein-ligand complex fixed. The entire system is minimized for another 1000 steps, and then the system relaxes for one nanosecond using molecular dynamics.

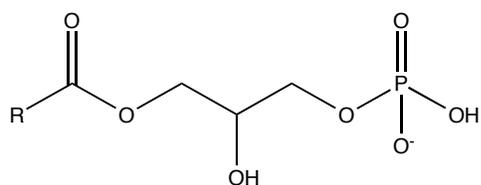
3.7. Figures and Tables

| Residue | Original Eisenberg | Normalized Eisenberg |
|---------|--------------------|----------------------|
| ALA | 0.250 | 0.620 |
| GLY | 0.160 | 0.480 |
| CYS | 0.040 | 0.290 |
| ILE | 0.730 | 1.380 |
| LEU | 0.530 | 1.060 |
| MET | 0.260 | 0.640 |
| VAL | 0.540 | 1.080 |
| PHE | 0.610 | 1.190 |
| TRP | 0.370 | 0.810 |
| TYR | 0.020 | 0.260 |
| ASN | -0.640 | -0.780 |
| GLN | -0.690 | -0.850 |
| THR | -0.180 | -0.050 |
| PRO | -0.070 | 0.120 |
| SER | -0.260 | -0.180 |
| ARG+ | -1.760 | -2.530 |
| ASP- | -0.720 | -0.900 |
| GLU- | -0.620 | -0.740 |
| HIS0 | -0.400 | -0.400 |
| HIS+ | - | -0.400 |
| LYS+ | -1.100 | -1.500 |

Table 3.1 The original and optimized Eisenberg hydrophobicity value for each amino acid residue. Larger positive numbers indicate increasing hydrophobicity.

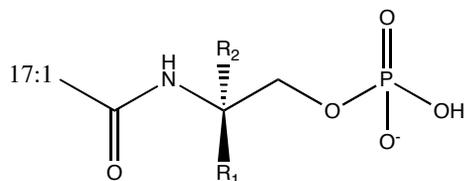
| | aWW | Biological | (aWW+Bio)/2 |
|------|-------|------------|-------------|
| ALA | 0.50 | 0.11 | 0.31 |
| GLY | 1.15 | 0.74 | 0.95 |
| CYS | -0.02 | -0.13 | -0.08 |
| ILE | -1.12 | -0.60 | -0.86 |
| LEU | -1.25 | -0.55 | -0.90 |
| MET | -0.67 | -0.10 | -0.39 |
| VAL | -0.46 | -0.31 | -0.39 |
| PHE | -1.71 | -0.32 | -1.02 |
| TRP | -2.09 | 0.30 | -0.90 |
| TYR | -0.71 | 0.68 | -0.02 |
| ASN | 0.85 | 2.05 | 1.45 |
| GLN | 0.77 | 2.36 | 1.57 |
| THR | 0.25 | 0.52 | 0.39 |
| PRO | 0.14 | 2.23 | 1.19 |
| SER | 0.46 | 0.84 | 0.65 |
| ARG+ | 1.81 | 2.58 | 2.20 |
| ASP- | 3.64 | 3.49 | 3.57 |
| GLU- | 3.63 | 2.68 | 3.16 |
| HIS0 | 0.11 | - | - |
| HIS+ | 2.33 | 2.06 | 2.20 |
| LYS+ | 2.80 | 2.71 | 2.76 |

Table 3.2 An average of the Wimley-White^{14,24} and White Biological²⁵ scales are used to determine the thermodynamic benefit to having a certain face of a helix solvated by the lipid membrane.



| Ligand | Details of Saturation |
|---------------|--|
| LPA 18:1 | <i>Cis</i> at Δ^9 |
| LPA 18:0 | Saturated |
| LPA 16:0 | Saturated |
| LPA 20:0 | Saturated |
| LPA 18:2 | <i>Cis</i> at Δ^9 and Δ^{12} |
| LPA 18:3 | <i>Cis</i> at Δ^9 , Δ^{12} and Δ^{15} |

Table 3.3 A description of the members of the LPA family docked into LPA₂. All of these ligands are built as the *R*-stereoisomer.



| Functional Group | R ₁ functionalized, R ₂ = H | R ₂ functionalized, R ₁ =H |
|-------------------|---|--|
| Methylene Hydroxy | VPC31143 | VPC31144 |
| Carbomethyl | VPC31139 | VPC31180 |
| Methylene Amino | VPC12178 | VPC12048 |
| Methyl | VPC12086 | VPC12101 |
| Ethyl | VPC12109 | VPC12115 |

Table 3.4 A description of the NAIPA derivatives used to explore the stereoselectivity of our model for LPA₂.

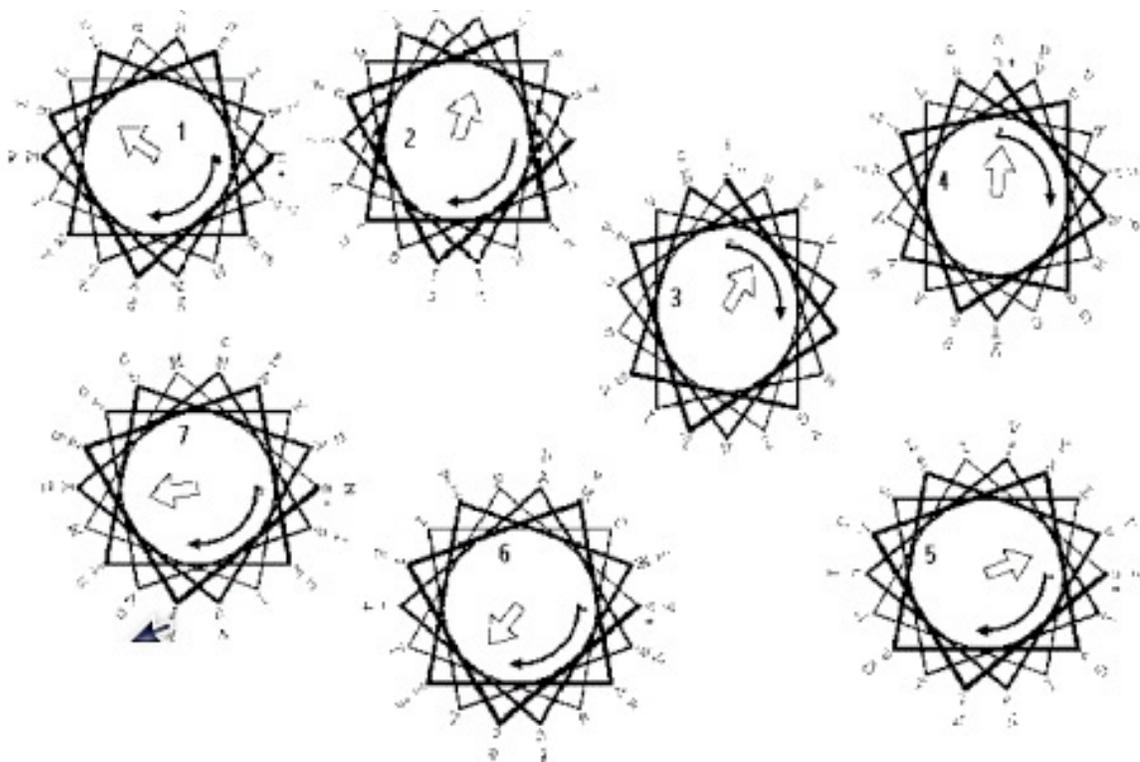


Figure 3.1 A sketch of the topology of the low-resolution frog rhodopsin crystal structure. This template is used for arranging the centers of helices relative to one another.

3.8. References

- (1) Mayo, S. L. O., B.D.; Goddard, W.A., III *Journal of Physical Chemistry* **1990**, *94*, 8897-8909.
- (2) MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *Journal of Physical Chemistry B* **1998**, *102*, 3586-3616.
- (3) Trabanino, R. Prediction of structure, function, and spectroscopic properties of G-protein-coupled receptors: methods and applications, Ph.D. Thesis, California Institute of Technology, 2004.
- (4) Trabanino, R. J.; Hall, S. E.; Vaidehi, N.; Floriano, W. B.; Kam, V. W. T.; Goddard, W. A. *Biophysical Journal* **2004**, *86*, 1904-1921.
- (5) Vaidehi, N.; Floriano, W. B.; Trabanino, R.; Hall, S. E.; Freddolino, P.; Choi, E. J.; Zamanakos, G.; Goddard, W. A. *Proceedings of the National Academy of Sciences of the United States of America* **2002**, *99*, 12622-12627.
- (6) Gasteiger, E.; Gattiker, A.; Hoogland, C.; Ivanyi, I.; Appel, R. D.; Bairoch, A. *Nucleic Acids Research* **2003**, *31*, 3784-3788.
- (7) Thompson, J. D.; Higgins, D. G.; Gibson, T. J. *Nucleic Acids Research* **1994**, *22*, 4673-4680.
- (8) Eisenberg, D.; Weiss, R. M.; Terwilliger, T. C.; Wilcox, W. *Faraday Symposia of the Chemical Society* **1982**, 109-120.
- (9) Hall, S. E. Development of a structure prediction method for G-protein coupled receptors, Ph.D. Thesis, California Institute of Technology, 2005.
- (10) Hall, S. E.; Floriano, W. B.; Vaidehi, N.; Goddard, W. A. *Chemical Senses* **2004**, *29*, 595-616.
- (11) Vaidehi, N.; Jain, A.; Goddard, W. A. *Journal of Physical Chemistry* **1996**, *100*, 10508-10517.
- (12) Schertler, G. F. X. *Eye* **1998**, *12*, 504-510.
- (13) Eisenberg, D.; Schwarz, E.; Komaromy, M.; Wall, R. *Journal of Molecular Biology* **1984**, *179*, 125-142.
- (14) Wimley, W. C.; White, S. H. *Nature Structural and Molecular Biology* **1996**, *3*, 842-848.
- (15) Hessa, T.; Kim, H.; Bihlmaier, K.; Lundin, C.; Boekel, J.; Andersson, H.; Nilsson, I.; White, S. H.; von Heijne, G. *Nature* **2005**, *433*, 377-381.

- (16) Brady, G. P., Jr.; Stouten, P.F. *Journal of Computationally-Aided Molecular Design* **2000**, *14*, 383-401.
- (17) Ewing, T. J., S. Makino, A. G. Skillman, and I. D. Kuntz *Journal of Computational Aided Molecular Design* **2001**, *15*, 411-428.
- (18) Rocchia, W.; Sridharan, S.; Nicholls, A.; Alexov, E.; Chiabrera, A.; Honig, B. *Journal of Computational Chemistry* **2002**, *23*, 128-137.
- (19) Rocchia, W.; Alexov, E.; Honig, B. *Journal of Physical Chemistry B* **2001**, *105*, 6507-6514.
- (20) Sadanobu, J. G., W.A. *Journal of Chemical Physics* **1997**, *106*, 6722-6729.
- (21) Thumser, A. E. A., Voysey, Joanne E., and Wilton, David C. *Biochemical Journal* **1994**, *301*, 801-806.
- (22) Humphrey, W., Dalke, A. and Schulten, K. *Journal of Molecular Graphics* **1996**, *14*, 33-38.
- (23) Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kale, L., Schulten, K. *Journal of Computational Chemistry* **2005**, *26*, 1781-1802.
- (24) Wimley, W. C.; Creamer, T. P.; White, S. H. *Biochemistry* **1996**, *35*, 5109-5124.
- (25) White, S. H.; Wimley, W. C. *Annual Review of Biophysics and Biomolecular Structure* **1999**, *28*, 319-365.