Chapter I: Introduction

Chapter I: Introduction

G-protein coupled receptors (GPCRs) are heptahelical, membrane bound proteins involved in cellular communication and signaling. With the implication of G-protein coupled receptor (GPCR) in many diseases, (1, 2) the need to solve the high-resolution three-dimensional structure of this class of integral membrane proteins, to enable structure-based drug design, has become an important problem in structural biology. There is only one high resolution crystal structure for any GPCR that is of bovine rhodopsin (3) due to the difficulty in obtaining crystals for GPCRs. Availability of reliable three-dimensional structures for GPCRs will greatly aid subtype specific structure-based drug design that would minimize the cross-reactivity of drugs among other GPCRs. To partially solve the problem and to stimulate experiments to gain greater insight into GPCR structure and function, many research groups and pharma industry have utilized homology-modeling techniques based on the crystal structure of bovine rhodopsin (4). However a recent review by Baker, Sali and others (5) has shown that a homology model for a protein with ~80 amino acids and with a sequence identity of less than 30% to the template crystal structure is unreliable. The sequence identity of bovine rhodopsin to many GPCRs is less than 20%, meaning a homology-based approach is unlikely to provide a reliable structure to be used for making predictions. Despite these shortcomings, several groups have utilized a rhodopsin-based homology model with loose distance restraints obtained from mutation experiments to model the active site of many GPCR systems. Several other groups (6) have raised concern about the use of the bovine rhodopsin structure as a template for homology modeling since the bovine

rhodopsin crystal structure with 11-cis retinal, an inverse agonist, could be in an inactive conformation. Hence the homology models based on the bovine rhodopsin crystal structure, with optimization using experimental information, are likely to provide good correlation for antagonists (since the rhodopsin structure is co-crystallized with an inverse agonist), but unlikely to capture the correct nature of the agonist binding site (6). Moreover it is not clear if antagonists preferentially bind only to active or inactive conformation of the receptor. It must be noted that for many GPCRs, such as the dopamine receptor, subtype specific agonist therapies are high desirable and the current level of structure prediction technology does not allow for the rational design of such ligands.

To provide structural and ligand binding information on GPCRs, Vaidehi, Goddard and the Goddard group members have spent the past six years developing first principles computational techniques for predicting the three dimensional structure of GPCRs starting from the amino acid sequence (MembStruk), and for predicting binding site and binding energy of various ligands to GPCRs (HierDock). By "First principles method" we mean that no coordinates have been used from the high-resolution crystal structure of bovine rhodopsin. We also do not use any homology features to bovine rhodopsin. However we do use 3D structural information, specifically the starting relative orientation of the helical axes of the seven transmembrane(TM) helices that are obtained from the 7.5Å resolution electron density measurements of frog rhodopsin from Schertler's laboratory (7), and knowledge of the correspondence between TM sequences and specific helices in this structure. Using the low-resolution electron density map also implies using the correct initial positioning of the seven TM helices, which is 3D

structural information. This provides a starting point for the Monte Carlo optimization of the protein-lipid complex. Although some experimental data is used to start the process, we refer to MembStruk as first principles predictions to distinguish it from all other methods that use the atomistic 3D structure of bovine rhodopsin or bacteriorhodopsin, with the assumption that conserved residues would be in similar positions, plus assumptions of where the known agonists and antagonist bind to build an experimental based structure. The details of the MembStruk method are given in reference 11. In the next chapter the various steps of MembStruk as applied to dopamine receptors are explained.

The HierDock method to predict ligand binding sites has also been developed in the Goddard group over the last six years and tested for a number of co-crystals of gobular proteins (non-GPCR) in addition to bovine rhodopsin (8,9). The HierDock method for docking ligands into proteins does not utilize any experimental information to determine ligand binding sites and affinities.

The combination of MembStruk and HierDock methods has been used for prediction of structure and ligand binding sites of other GPCRs (8-11). We have validated our methods against the one crystal structure bovine rhodopsin and find good agreement, TM helices usually within one or two residues, RMSD in coordinates of main chain atoms in the structure of TM regions good to 2.84Å, binding site for retinal to 0.6Å. The RMSD in coordinates of the binding site of 11cis-retinal in the predicted structure of bovine rhodopsin is 2.92Å. About 75% of the residues in the binding site of 11cis-retinal in the crystal structure of rhodopsin are also in the predicted binding site (11). Hence

these methods are useful to predict the mutation candidates for ligand binding. The calculated binding energy of ligands using HierDock can differentiate two ligands with two or more orders of magnitude difference in binding affinity. The validation of these methods for GPCRs carried out over the last two years in the Goddard group includes all five dopamine receptors, all nine adrenergic receptors, the four histamine receptors, LPA_1 receptor, the G2A and TDAG8 choline-lipid receptors, the MrgC11 polypeptide receptor. As these validations have proceeded they have motivated improvements in the MembStruk and HierDock computational methods using more extensive energy optimization (minimization, Monte Carlo, and dynamics), more complete side chain optimizations, better charges and force fields based on quantum mechanics, and improved methods of including solvation, etc. Using these techniques we have reported the structure of 10 olfactory receptors (10), structure of bovine rhodopsin (9, 11), and other GPCRs (12).

Several members of the Goddard group have carried out the above-mentioned efforts. The continued development of the MembStruk and HierDock methods have been primarily carried out by senior members Vaidehi and Floriano, Graduate students Hall, Trabanino, and Kam. My role (with undergraduates Mr. Peter Freddolino and Mr. Maziyar Kalani) has been to validate these methods by applying them to a set of ligands binding to 22 GPCRs. My research studies have motivated many improvements in MembStruk and HierDock methods, although I have not personally changed the software code. These systems on which I have played a major role include:

Dopamine receptors: There are five subtypes of dopamine receptors, namely D1, D2, D3, D4, D5. Here the interest is in understanding the differential ligand interactions

between these closely related subtypes of dopamine receptors. I have conducted detailed studies of 11 experimentally well-studied agonists and antagonists to all five subtypes of the human dopamine receptors. The detailed results on the D2 studies are discussed in chapter 2, and the results of a blind prediction with Aventis Pharma on the D2 dopamine receptor is discussed in chapter 3. Appendices 3-6 present preliminary results on the D1, D3, D4 and D5 subtype of the human dopamine receptors. Appendix I is now published in the *Proceedings of the National Academy of Sciences* and involved significant contributions in the writing from Vaidehi and Goddard (my mentors), with some contributions from Mr. Peter Freddolino. Chapter 2 is an early version of a manuscript that is nearly ready to be submitted for publication.

Adrenergic receptors: There are nine subtypes of adrenergic receptors namely, α1, α2, α3, α1A, α1B, α1D, α2A, α2B, α2C. Here again the interest is in subtype specificity for endogenous ligands and pharmaceuticals. Also cross reactivity with dopamine receptors is an important shortcoming of the currently available therapies. Mr. Freddolino and I contributed equally to the structure and function study of these systems. We have conducted a detailed study of 11 experimentally well-studied ligands to the α2 adrenergic receptor. The results of the work on the α2 Adrenergic receptor is presented in Appendix II, which is now published in the *Proceedings of the National Academy of Sciences*. This work was written by Mr. Freddolino, and myself and involved significant contributions in the writing from Vaidehi and Goddard.

Histamine receptors: The four subtypes of histamine receptors H1, H2, H3, and H4 are also important pharmaceutical targets. Here again the interest is in subtype specificity for

endogenous ligand and pharmaceuticals. As is the case with many of the aminergic systems, cross reactivity with other GPCRs is a major shortcoming of the current pharmaceutical drugs for asthma or allergy. It is an undesirable side effect that a patient treated for asthma or heartburn suffers from schizophrenic symptoms or hallucinations. Therefore, there is great motivation to remove the receptor and subtype cross-reactivity for these systems. I studied the histamine GPCRs, with minor assistance from Mr. Maziyar Kalani and Peter Freddolino. We have considered detailed studies of 11 experimentally well-studied ligands to each of these systems. The results of these studies are currently being analyzed and reported.

Lipid receptors: ***G2A, TDAG8, OGR1, GPR4.*** In collaboration with Professor Owen Witte's group at UCLA, I have predicted the residues in the binding site of lipid lysophosphatidylcholine (LPC) in G2A GPCR for mutation studies. These predictions were performed with virtually no information on this system, except that LPC activates this receptor. The mutations that would affect the binding of the phosphate, choline and other polar groups in LPC have been tested experimentally and have positively shown that the predicted binding region is indeed correct. The mutations results for the binding site of the alkyl chain of the lipid are awaited. We are designing lipids that would bind to this receptor with better affinity than LPC. In this work Mr. Rene Trabanino and I predicted the structure of G2A receptor and I focused on the docking of these complex lipid ligands. This work is described in chapter 6 of this thesis.

Highlight of Results:

The structure and function prediction for the long isoform of the human D2

dopamine receptor is discussed in great detail in Chapter II. The predicted structure of D2 dopamine receptor has been validated with the experimental substituted cysteine accessibility method (SCAM) and site directed mutagenesis (13). We have also conducted computational alanine scanning mutations on the residues within 5-Å of the binding site of dopamine and haloperidol (an antagonist) to understand critical contacts for ligand stabilization. This chapter also discusses issues with homology modeling for this receptor. The agonist binding site in D2 dopamine receptor is located between TM3, 4, 5 and 6 whereas a class of lengthy antagonists have been found to be between helices TM3, 4, 5, 6 and also residues

The results of a blind study conducted in collaboration with Aventis Pharma, are discussed in Chapter III. The Aventis pharma team provided us with the two dimensional structures of 9 ligands (they did not specify agonists or antagonists) and we determined the binding site, binding affinities, and agonists or antagonist function of each ligand. We were 100% correct in the ordering of the high, medium and low affinity ligands with a correlation of 0.75 against experimental inhibition constants (which we were provided after we presented them with our data). During these studies, we also identified ways that Aventis could improve their ligands to increase their binding abilities. We successfully identified a possible stereoisomer for one of their ligands that would have improved binding to the D2 receptor. In similar studies we performed the same tasks for the D3 and α_{1A} receptors. From the predicted structures of these receptors we have a rationale for what causes a ligand to be D2 or α_{1A} specific, that can be directly tested out by experiments.

In Chapter IV I discuss the binding site of dopamine to the five subtypes of the

human dopamine receptors. In this chapter the focus is on a very quantitative, residue-by-residue analysis of contact points between the receptor and ligand. The residues that could be mutated by experimentalists to improve the binding affinity of dopamine to each receptor are identified (usually it is simpler to suggest mutations that would abolish binding). In this chapter we bring up residues that cause differential binding of dopamine to each receptor subtype and hypothesize critical contacts for receptor differentiation and how these points could be tested by experimentalists and utilized in subtype specific design of agonists useful for Parkinson's.

The binding site of classic antagonists to the five subtypes of the human dopamine receptors are discussed in Chapter V. The important residues that we believe cause differential binding of ligands to these receptors and suggest modifications that should improve the binding and selectivity of ligands to each receptor subtype are elucidated in this chapter.

The biggest challenge in my work was the G2A project. This protein is very different from the aminergic systems. G2A is known to bind selectively to choline lipids (LPC) and not to LPA and other non-choline lipids; there is no data available as to which residues are directly involved in binding. Moreover the lipid ligands are long and flexible and involve many rotatable bonds making it a challenge to the HierDock techniques. We predicted the protein structure in May 2003 and I predicted the binding site in June 2003. Based on these results, I provided the Witte group with ten predictions of mutation experiments that should validate the predicted site. Of these mutations, eight are predicted should decrease the binding while two suggested mutations should not affect

the binding affinity (one might increase it). Seven of these ten experiments have now been done and all seven worked as predicted. It is noteworthy that the seven mutations tested till now are the mutations that reduce binding to the phosphate head, or the choline group or other polar groups in the lipid. The mutations corresponding to the alkyl chain binding region remains to be tested. I have also predicted mutations that should convert G2A to an LPA receptor. The results of these studies are in preparation for publication.

The structure and function of the human D2 dopamine receptor and the binding site and binding affinities for pharmaceutical agonists and antagonists is discussed in Appendix I. In this study we gained insight into the difference between agonists vs. antagonist binding, critical contacts, we classified antipsychotics into two classes, and predicted ways of designing leads that would not cross-react with the chemokine receptors, some of which also have negatively charged amino acids in their third TM domains.

The details of structure and function of the human β2 adrenergic receptor and the binding site and binding affinities of pharmaceutical agonists and antagonists are presented in Appendix II. Similarly, in this study we gained insight into the differences between agonists and antagonists binding sites, and potential mechanisms of activation. Other members of the Goddard group are currently studying our hypotheses for the potential activation mechanism. The structure also allows for understanding of the mechanisms of cross reactivity with other aminergic systems.

The structure and function of the human D1, D3, D4 and D5 dopamine receptors and the differences are presented in Appendices III-VI.

References:

1. Pierce K.L., Premont R.T., Lefkowitz R.J. *Nature Reviews Molecular Cell Biology* 3, 639-650, (**2002**).

2. Schoneberg T., Schulz A., Gudermann T. *Rev. Physiol. Biochem. and Pharm.,* 144, 143-227 **(2002).**

3. Palczewski K., Kumasaka T., Hori T., Behnke C. A., Motoshima H., Fox B. A., Le Trong I., Teller D. C., Okada T., Stenkamp R. E. *Science* 289**,** 739–745 (**2000**).

4. (a) Teeter M.M., Froimowitz M., Stec B., DuRand C.J. *J Med Chem*. 37, 2874-2888 (**1994**). (b) Trumpp-Kallmeyer S., Hoflack J., Bruinvels A., Hibert M. *J Med Chem*. 35, 3448-3462 (**1992**). (c) Neve K.A., Cumbay M.G., Thompson K.R., Yang R., Buck D.C., Watts V.J., DuRand C.J., Teeter M.M. *Mol Pharmacol*. 60, 373-381 (**2001**). (d) Varady J., Wu X., Fang X., Min J., Hu Z., Levant B., Wang S. *J. Med. Chem*. 46, 4377-4392 (**2003**).

5. Baker D., Sali A. *Science* 294, 93-96 (**2001**).

6. For a review, please see: Furse K.E., Lybrand T.P. *J. Med. Chem.* 46, 4450-4462 (**2003**).

7. Schertler G.F.X., Villa C., Henderson R. *Eye* 12, 504-510 (**1998**).

8. (a) Datta D., Vaidehi N., Floriano W. B., Kim K. S., Prasadarao N. V., Goddard W. A., III *Proteins Struct. Funct. Genet.* 50**,** 213–221 (**2003**). (b) Datta D., Vaidehi N., Xu X., Goddard W. A., III *Proc. Natl. Acad. Sci. USA*

99**,** 2636–2641 (**2002**). (c) Zhang D., Vaidehi N., Goddard W. A., III, Danzer J. F., Debe D. *Proc. Natl. Acad. Sci. USA* 99**,** 6579–6584 (**2002**). (d) Wang P., Vaidehi N., Tirrell D. A., Goddard W. A., III *J. Am. Chem. Soc.* 124**,** 14442–14449 (**2002**). (e) Kekenes-Huskey P. M., Vaidehi N., Floriano W. B., Goddard W. A., III *J. Phys. Chem. B* 107**,** 11549–11557 (**2003**). (f) Floriano W. B., Vaidehi N., Zamanakos G., Goddard W. A., III *J. Med. Chem.* 47**,** 56–71 (**2004**).

9. Vaidehi N., Floriano W., Trabanino R., Hall S., Freddolino P., Choi E.J., Zamanakos G., Goddard W.A. *Proc. Natl. Acad. Sci. U.S.A.* 99, 12622-12627 (**2002**).

10. Floriano W.B., Vaidehi N., Singer M., Shepherd G. and Goddard III, W.A. *Proc. Natl. Acad. Sci. U.S.A*. 97, 10712-10716, (**2000)**.

11. Trabanino R.J., Hall S.E., Vaidehi N., Floriano W.B., Kam V., Goddard III, W.A. *BioPhys. J.* 86, 1904-1921 **(2003)**.

12. (a) Kalani M. Y., Vaidehi N., Goddard III, W.A. *Proc. Natl. Acad. Sci. U.S.A*. 101, 3815-3820 (**2004**). (b) Freddolino P. L., Kalani M. Y., Vaidehi N., Goddard III, W.A. *Proc. Natl. Acad. Sci. U.S.A*. 101, 2736-2741 (**2004**).

13. Shi L., Javitch J. *Annu Rev. Pharmacol. Toxicol*. 42, 437-67 (**2002**) *and references therein*.