Adrenergic Receptors: Model Systems for Investigation of GPCR Structure and Function

Thesis by

Heather L. Wiencko

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2009

(Defended May 18, 2009)

Acknowledgements

"I may not have gone where I intended to go, but I think I have ended up where I needed to be."

– *Douglas Adams*

I started my Caltech journey with catalysis, but ended up doing computational chemistry with biochemistry applications. I was fortunate to make this massive leap into Bill Goddard's group, where Bill supported, taught, and challenged me for five years. I am grateful for the chance to make the switch, and for all I've learned from Bill over the years. He defies the stereotype of the highly specialized academic, and demonstrates that breadth as well as depth of knowledge is both possible and desirable. Along the way I was lucky to work with the Biogroup, from Nagarajan Vaidehi who started working with me as soon as I arrived in the group, to Ravinder Abrol who continued to observe, discuss, critique, and explain patiently. My colleagues in the Biogroup Pete, Caglar, Bartosz, Soo-Kyung, Jenelle, Adam, and our second years Andrea and Caitlin all contributed in some way to the work described in this thesis, and I've enjoyed working with all of them. The rest of the Goddard group also deserves thanks, especially John Keith, who pointed out over coffee one afternoon that I should really consider looking into the GPCR

work being done in the group.

I would be remiss if I didn't thank my undergraduate adviser, Tim Warren, without whom I would not have made it to Caltech, much less all the way to a PhD. Even though I'm not currently using all the inorganic chemistry he taught me, his ideas about science and teaching, as well as his unflagging encouragement remain the foundation of much of my work and ambitions.

The greater Caltech community gave me a reason to step out of the lab for reasons other than sleep and lunch. Desiree LaVertu and the Caltech Glee Club were a constant source of music, friendship, fun, and solace from my very first term here to the very last. It's hard to imagine what my Caltech experience would have been like without music, and on the days when giving up grad school and joining the circus seemed like a good idea the thing that kept me tethered here was knowing I had rehearsal that I needed to get to. The Caltech knitters drew me out of the BI sub-basement once a week to stitch and converse and think about things other than science, however briefly. The Chemistry department staff provided support and assistance, and a student couldn't ask for a better department to belong to.

Few people can boast an international network of friends and family, and I have enjoyed the best of many possible worlds and timezones. Through the most difficult times out here, Billy Healy served as the belt with which I held up my sanity. All the friends I made here are too numerous to name, but the entire Fellowship of the Flat and their d20 hangers-on made my Pasadena exile enjoyable. Katie Brenner, Heidi Privett, Michelle Robbins, Ariele Hanek, Mike Adams, Greg Ver Steeg,

and all the rest of them are great friends and offered invaluable support. Katie Saliba, Warren Stramiello, and Chris Borths allowed me to indulge my gaming habit every week, and I cherish the many evenings we spent over a table of dice and books. Farther afield, Riz Kassim dragged me out into LA and made me enjoy a city that struck me as weird and foreign when I first arrived. It's still weird and foreign, but my friends helped it seem like home.

My sister Christy let me show up in her house with almost no notice to write my proposals, and was the first person over the course of my PhD to figure out that I really didn't want to talk about when I would be finished. The Nortons and Laceys became my home away from home, and the entire gang back in Dublin have assured me over the years that I'd be able to go back and still have a life. I am astonishingly lucky to have the support of so many people here and abroad, and although I can't name everyone I am certainly grateful.

Finally, I must thank my husband, Ger. Perhaps the most important lesson I learned in the past seven years came from him, that together two people can indeed be greater than the sum of their parts.

Abstract

Membrane proteins mediate intercellular communication through a wide variety of modes, resulting in changes in the membrane and within the cell itself. One superfamily of integral membrane proteins, G-protein coupled receptors (GPCRs), are responsible for a vast diversity of processes including three of the five senses: sight, smell, and taste. GPCRs comprise 4% of the human genome but are disproportionately represented as pharmaceutical targets: over 50% of the best selling drugs target some member of this superfamily, mitigating the effects of diseases ranging from hypertension to schizophrenia. These receptors exist in equilibrium between their active and inactive states, and either of these states may be stabilized by the binding of an extracellular stimulus that may be either a small molecule or a peptide. The active state of the receptor triggers a response from the associated G-protein, which then controls the release of a second messenger within the cell that initiates other downstream processes. The ubiquity of GPCRs in key biological processes makes them both an attractive target for drug development and a challenge for selective drug design. Their conformational flexibility and membrane environment pose challenges for direct structural characterization, and to date only five of the more than 1,000 known GPCRs have been characterized by

high-resolution crystallography.

The nine adrenergic GPCRs mediate the stress response throughout the body, and are implicated in diseases including hypertension and asthma. While they are among the best studied subtypes of GPCRs, much remains to be learned about selectivity and activation. The first section of this work describes the *ab initio* structure prediction of the turkey β1 receptor and validation using a series of stabilizing mutations. This work preceded the currently available turkey β1 structure, but shows good agreement, especially in the binding site. It validates the latest methods developed for GPCR structure prediction, emphasizes the role of a neutral charge scheme in energy determination, and explores a structure validation strategy based on stabilizing mutations rather than ligand docking. The next section uses the experimental crystal structure as a starting point for nanosecond timescale molecular dynamics, exploring the roles of ligand binding in helix movement that contribute to the transition to an active state. These simulations reveal the early steps in receptor activation, beginning with tilting motions of transmembrane helices 5 and 6 and movement of transmembrane helix 1 closer into the protein core. The last section also uses newly available crystal structures as a starting point, and builds homology models of the human adrenergic receptors for which there are not yet crystal structures. The receptors most closely related to the target structures show the best results, while the less related ones will require further refinement. The best structures provide insight into the binding site of subtype selective antagonists, and can serve as the foundation for future studies.

The central idea of this thesis is that theory and experiment can and must work in concert, with the findings from one propelling advances in the other in the mutual pursuit of knowledge. The methods developed in the course of this work are applied to systems with a great deal of experimental knowledge, but may be applied to those that have been less thoroughly characterized. Over the course of these explorations, new subtleties in adrenergic structure have been illuminated, and may drive further exploration into selective binding and the activation mechanism of these and other receptors.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

G-protein coupled receptors (GPCRs) make up 4% of the human genome, comprising one of the largest and most studied superfamilies of membrane proteins.[1–4] These proteins integrate in the cell membrane and mediate communication with the cell by responding to an extracellular ligand and activating a guanine nucleotide-binding protein (G-protein) at the intracellular side of the membrane.[5–7] The ligand may be a hormone, neurotransmitter, drug, odorant, or covalently bound molecule that responds to light. Once in an active state, the associated G-protein initiates a second messenger cascade, a signaling mechanism first characterized in 1958.[8] The processes mediated by GPCRs range from sensory recognition to immune responses to system regulation, and are not yet fully understood.[9,10] Their ubiquity and ligand-based mode of activation make them attractive drug targets: more than 50% of the current top selling drugs target GPCRs.[11–15] Of the 266 human proteins targeted by approved drugs, 26% of them are Class A GPCRs.[16] Small molecule drug development still focuses heavily on this superfamily for new advances in medicine.

With accurate binding site models of important GPCRs, it may be possible to

design new potent and selective drugs as well as improve the selectivity of existing drugs. Early work[17] in GPCR study confirmed the seven transmembrane (TM) helix structure that is conserved throughout the superfamily. Ligands bind primarily to the TM bundle rather than to the flexible extracellular loops,[18] and the conserved structure continues to drive understanding of GPCR structure and binding sites.

GPCRs are divided into three classes, each sharing a series of conserved residues and interactions. Class A, or rhodopsin-like receptors feature the conserved seven-TM structure, bind small ligands inside this TM core, and are the primary GPCR targets for drug development.[14,16,19] Conserved residues include a cluster of charged residues in TMs 1, 2, and 7 that interact via interhelical salt bridges and are important for stabilization of the receptor[20] and ion regulation,[21] a WXPFF motif in TM6 responsible for TM6 shape and implicated in both ligand binding[22] and receptor activation,[23,24] an NPXXY motif at the intracellular side of TM7, and a three-residue D- or ERY motif at the intracellular side of TM3 that can interact with TMs 5 or 6 to stabilize the inactive receptor.[25,26] In general, the universally conserved receptors contribute to stabilization of either an active or inactive state, while residues only conserved within a subfamily are responsible for subtype selective binding.

These highly conserved residues form the basis of the Ballesteros-Weinstein numbering system,[27] used throughout this work as a way of describing residues in a particular GPCR so that comparisons with related receptors can be made easily. This numbering scheme names a conserved residue in each TM as X.50, where

X is the TM helix number, and residues around it are numbered in sequence. A residue of interest will be denoted with its three-letter code and its number in the sequence, with the Ballesteros-Weinstein designation in superscript. When referring to a conserved residue without reference to a particular receptor, only the three-letter code and Ballesteros-Weinstein designation are used.

The increasing availability of GPCR X-ray crystal structures has dramatically advanced understanding of GPCR structure and ligand binding sites. The first structure available was bovine rhodopsin, first published in 2001.[28–36] It verified previously determined information about general structure, the role of highly conserved residues, and the ligand binding site, but it also provided a starting point for further structure predictions through homology modeling. The conserved 7-TM structure allowed this distantly related receptor to provide insight into human receptors in different systems, but the low sequence identity between bovine rhodopsin and many receptors of interest cast doubt on the validity of the resulting models. This changed again in 2007 with the publication of the human β2 adrenergic receptor (β2),[37–40] then in 2008 with the turkey β1 (β1) adrenergic receptor[41] and human adenosine A_{2a}.[42] An engineered mutant of another GPCR, BLT1, has been reported[43] and may result in yet another crystal structure. These structures revitalized the homology model as a valid avenue of structure prediction, both because of the similarity between the two related structures and the new availability of crystal templates expected to be more similar to receptors of interest.

Many important GPCRs are constitutively active, or self-activating, adding an-

other avenue of exploration.[44] GPCRs exist in equilibrium between active and inactive states, and ligand binding may stabilize one state or preserve the equilibrium by preventing other ligands from binding.[45–47] In general, ligands that stabilize an active or inactive state are known as agonists or inverse agonists; many known antagonists are now understood to be allosteric antagonists that simply block the binding site but do not affect the activation state.[48] As a result of this dynamic behavior, ligand studies alone do not completely illuminate the mechanism of activation. As structure identification has progressed, due to advances both in methods and availability of X-ray structures, activation has become the new frontier of GPCR research.

The bovine rhodopsin crystal structure contains a salt bridge between TMs 3 and 6,[28,31,49] connecting the highly conserved ERY motif at the intracellular end of TM3 with a conserved asparagine at the intracellular end of TM6. These residues are conserved throughout the family, and although this interaction is not observed in subsequent crystal structures there is evidence that the interaction is conserved in the native protein.[25] When considering the absence of this interaction in the available adrenergic crystal structures, it is important to consider the necessary modifications for crystallization may have disrupted this interaction. Experimental[25,50] and theoretical evidence show that this salt bridge creates an "ionic lock," stabilizing TMs 3 and 6 in their inactive positions. This lock may break during activation, but MD simulations imply the lock breaks late in the activation process, after some receptor reorganization takes place. One available crystal structure shows

the conformation of an active receptor.[32] This structure shows the ionic lock broken, but also that the intracellular end of TM5 is extended beyond the membrane into the cytosol. The changes observed between the active and inactive conformations of rhodopsin are useful as a starting point for further study into activation, but without an understanding of the intermediate steps the complete mechanism remains unclear.

1.1 Adrenergic Receptors

Adrenergic receptors are a family of nine GPCRs that mediate the stress response to endogenous catecholamine agonists epinephrine and norepinephrine (Figure 1.1). Much of the seminal work in GPCR research, including two of the five currently available crystal structures, has targeted this family of receptors. They are found throughout the sympathetic nervous system, and regulate functions such as blood pressure and heart rate at rest and in response to stress. The β receptors' role in regulating blood pressure drew initial attention to the family as drugs were developed to fight hypertension, then as the α subtypes were discovered and characterized it was possible to create drugs that targeted a specific part of the sympathetic nervous system.

β adrenergic receptors are perhaps the best studied subtype, as they include targets for widely prescribed drugs for asthma and blood pressure. β1 antagonists are standard anti-hypertensive and cardiovascular disease treatments,,[51] including propranolol, the first selective β blocking drug,.[52] β2 agonists including

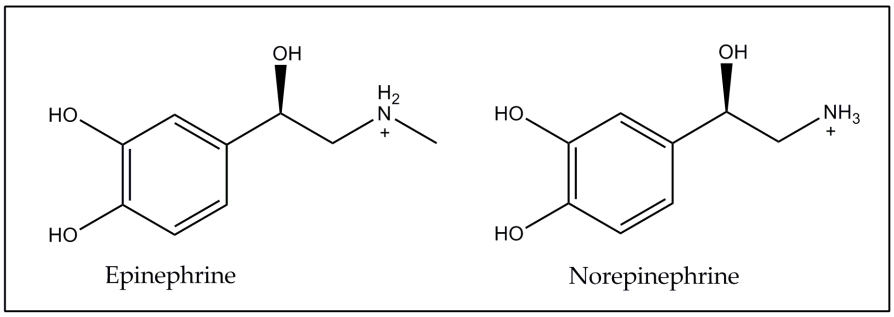

Figure 1.1: Adrenergic receptors respond to the endogenous catecholamine agonists epinephrine and norepinephrine.

salbutamol and formoterol are widely prescribed for asthma.[53] Cross-reaction of these drugs can have adverse effects for patients who suffer from both asthma and hypertension, so the utility of selective drugs for this subtype is immediate and widespread.

These receptors stimulate the production of cAMP through the activation of adenylyl cyclase and are found in the heart, lungs, and adipose tissue. Much of the structure-activity relationship work done for the adrenergic family has been on β2,[54–57] and currently the only human adrenergic crystal structure available is for β2.[37–40] As many binding site residues are conserved, this structure offers an excellent foundation for homology modeling, as discussed in Chapter 4.

The α1 receptors are located in vascular smooth muscle, the digestive tract, the liver, and central nervous system. The three subtypes are designated α1a, α1b, and α1d (the α1c designation was later discovered to be identical to α1a).[58] They stimulate production of inositol triphosphate and Ca^{++} [59] and regulate vasoconstriction in a variety of systems. α-blockers are sometimes prescribed as alternative antihypertensives instead of the traditional β-blockers.

Unlike the other two subtypes, the α2 receptors are inhibitory, decreasing the intracellular concentration of second messenger and decreasing the release of norepinephrine by inhibiting adenylyl cyclase.[60,61] These receptors are involved in blood pressure,[62] development, the startle reflex, and locomotion, among other processes. They are implicated in cognitive functions and working memory.[63] The canonical α2 antagonist, yohimbine, is an aphrodesiac, but also shows effects similar to α1 and β agonists, reflecting its role in the adrenergic feedback loop.[48,64]

All nine adrenergic receptors respond to the endogenous ligands epinephrine and norepinephrine, but their effects differ greatly depending on their associated G-protein and distribution throughout the body. In addition to the Class A GPCR conserved motifs, the adrenergic receptors share a series of conserved residues responsible for binding the endogenous agonists. A TM3 aspartic acid recognizes the protonated amine in nearly all adrenergic ligands, both agonists and antagonists. TM5 contains two absolutely conserved serine residues and a third serine conserved through all adrenergic receptors except α1a (and shifted one position in α2a and α2c). The Class A WXPFF motif participates in agonist binding.[22] TM7 contains a conserved tyrosine residue that may interact with the protonated amine on a ligand to form a ligand-mediated salt bridge between TMs 3 and 7 (some examples shown in Chapter 4). These conserved residues create an adrenergic pharmacophore for the family, and subtype selectivity depends on the binding site residues around these conserved motifs.

1.2 Computational Advances in GPCR Research

Theoretical predictions and models offer atom-level understanding and fine control of systems that are otherwise difficult to explore scientifically. Through structure prediction and molecular dynamics (MD) simulations, theory can illuminate details of both static structures as well as dynamic systems, providing insight important to drug development and understanding of GPCR activation. Homology models are often used to visualize results of experimental data, as well as provide a starting point for such studies. *Ab intio* methods that do not rely directly on an X-ray crystal structure are used to visualize systems that cannot be accurately modeled using existing structural data. Finally, molecular dynamics provides a way of monitoring the motions and activation mechanisms of GPCR systems that "snapshot" structural data cannot yet achieve.

1.2.1 Homology Modeling

The conserved seven-TM helix structure of GPCRs offers a useful starting point in structure prediction for GPCRs with no direct structural data. Once the crystal structure of bovine rhodopsin was available, direct homology modeling was widely used to create models of GPCR structures of interest as well as binding site models for drug development and screening for off-target effects. Rhodopsin's low sequence identity to human receptors and receptor families of interest complicated the structure prediction process, however, and homology models based on rhodopsin often needed heavy modifications.[65–67] The publication of β2 and

adenosine A_{2A} structures offered human template receptors to work with and turkey β1 provided a comparison to a well-studied, highly targeted family of receptors. While ligand-based homology modeling can work well,[12] it does not provide a useful method for systems that have yet to be extensively studied, or that have no known endogenous ligand. Chapter 4 explores the utility of simple homology modeling in generating structures for the entire human adrenergic family based on the two available, highly similar adrenergic X-ray crystal structures.

1.2.2 *Ab initio* Structure Prediction

For systems with less mutation and SAR data, and those that are only distantly related to available crystal structures, methods must be developed to predict their structures without direct use of the high-resolution crystal structures homology models rely upon. The PREDICT method[68] is one such method starting from the amino acid sequence of the receptor of interest, and Fanelli *et al.* developed a method for the prediction of the α1b receptor.[69]

The conserved structure of all GPCRs that inspired the homology modeling approach may also be a starting point for *ab initio* structure prediction. By considering the hydrophobicity of a region of amino acids as well as their tendency towards αhelical structure, it is possible to determine what parts of a GPCR sequence will form the TM helices and which will form the loops using only the amino acid sequence and thermodynamic information. Developed in the Goddard group, the Membstruk[70] and later MembSCREAM[71] protocols focus on optimization of these

predicted TM bundles, as most residues involved in ligand binding are located in the TM core.

1.2.3 Molecular Dynamics

Molecular dynamics (MD) provide a method to observe a GPCR interacting with its environment. While X-ray crystal structures offer high-resolution snapshots of the GPCR in an environment optimized for crystallization rather than native behavior, force field calculations can approximate protein structure and motion in a more native environment. MD also offers the opportunity to model sections of and interactions in the protein unavailable via crystallization.

This approach has been applied extensively to bovine rhodopsin to determine the regions and interactions important to activation.[72–76] These involve a variety of approaches, including constrained dynamics to test specific interactions, observation of the isomerization of retinal from *cis* to *trans* initiating the transition to the activated state, and they observe a range of changes including changes to the orientations of TMs 5 and 6. α1b has been studied with molecular dynamics based on a predicted structure,[26,69] and recently the crystal structure of β2 has been studied for 600 ns.[77] The potential for insight into activation from MD is profound, but in order to sample the entire millisecond timescale of GPCR activation it is necessary to incorporate some information from experimental studies of activation.

1.3 Subject of this Thesis

This thesis explores each of the major theoretical efforts in GPCR research: *ab initio* structure prediction, nanosecond (ns) timescale dynamics to elucidate activation mechanism, and homology modeling to obtain an entire family of structures based on closely related crystal structures. It explores two methods of validation for predicted structures: docking to verify the integrity of the ligand binding site and verification that residues implicated in binding are involved in the predicted structure, and comparison to stabilizing mutation data. I focus on the adrenergic family, utilizing the wealth of experimental data available to develop and test robust methods, leading to a greater understanding of the adrenergic receptors as well as methods that may be applied with confidence to other systems with less experimental data.

The thesis first addresses β1 turkey structure prediction directly from the amino acid sequence without assistance from crystal data, and validates the structure using stabilizing mutation data, before the crystal structure was available. This served as a test case for many of the recent developments in *ab initio* structure prediction. The next section starts with the β1 crystal data to initiate 10 ns of molecular dynamics on the apo protein, followed by ligand binding and a subsequent 10 ns of equilibration in the presence of ligand. These simulations also include a full intracellular loop, a region of the protein not currently available in the β1 crystal structure. Finally, the thesis turns towards homology models of the entire human adrenergic family, exploring the utility of available crystal structures for obtaining

3D models of highly similar receptors. Several of these models are validated with docked antagonists.

The central idea of this thesis is that theory and experiment can and must work in concert, with the findings from one propelling advances in the other in the mutual pursuit of knowledge. The methods developed in the course of this work are applied to systems with a great deal of experimental knowledge, but may be applied to those that have been less thoroughly characterized. Over the course of these explorations, new subtleties in adrenergic structure have been illuminated, and may drive further exploration into selective binding and the activation mechanism of these and other receptors.

Chapter 2

β1 Structure Prediction and Validation

2.1 Overview

Over several years, GPCR prediction methods in the Goddard lab have evolved to keep pace with the changing field of GPCR structure. Despite some success in the field with homology modeling, initial models based on existing crystal structures require some effort before they can be trusted to accurately reflect nuances of the desired system. (Chapter 4 details some of these efforts.) Over several years, Goddard *et al.* have developed,[70,78,79] applied,[53,80–88] and refined[71] techniques to determine the structure of GPCRs starting from the amino acid sequence and an initial template for arranging predicted TM regions in space.

The MembStruk method and its refinement in the GenSemble method begin by predicting the TM regions from the amino acid sequence, refining TM helices, then determining the rotations of the predicted helices in the membrane. TMPred[70,78] used a variation of the Eisenberg scale to determine the hydrophobicity of a position in an amino acid sequence and the hydrophobic moment of the helix. The

GenSemble method builds on the MembStruk method. PredicTM, like TMPred, uses an alignment of related sequences to determine the hydrophobicity along a sequence of interest, and uses moving window averages to reduce noise. It removes any gap penalty in the hydrophobic calculation by disregarding gaps in the target sequence or related sequences. It also uses the octanol scale,[89] a thermodynamic measurement, to determine hydrophobicity. This removes the need to adjust the hydrophobicity baseline in determining the starting and stopping point of helices. Both methods apply capping rules to extend predicted TM helices to charged or helix breaker residues. These hydrophobicites are used to calculate the hydrophobic center of the helices, which determines the vertical position of helices in the membrane.

Both methods build canonical helices and optimize them individually using short-term molecular dynamics in vacuum, with some changes in helix preparation and final structure choice, but the methods differ in the initial placement of these helices relative to each other. MembStruk aligns the optimized helices to the electron density map of the low-resolution frog rhodopsin structure.[17] It uses the hydrophobic center to position each helix along its axis, and hydrophobic moment to determine the helices' initial η rotation. GenSemble, developed as new GPCR crystal structures became more readily available, uses information either from fully optimized structure predictions (such as the prostaglandin receptor[84]) or from available X-ray structure data. Initial η values are less important because of refinements in the η rotation determination, so this initial rotation is determined

simply based on the conserved X.50 residue for the template and target structures. (TM3 uses 3.32, as the conserved 3.50 arginine is usually outside the predicted hydrophobic helix.)

MembStruk and GenSemble differ the most in their η rotation prediction procedures. While MembStruk rotates through 360° in 5° increments and optimizes each rotation with minimization only, GenSemble is built on the BiHelix and CombiHelix methods which optimize helix rotations with SCREAM.[90] This allows for rotation sampling in up to 30° increments. The BiHelix method isolates two helices at a time and samples every rotation of each helix (144 cases per pair), then combines the minimized pairwise energies to determine which TM bundles have the lowest energies for the system. The best bundles are built with CombiHelix and then optimized with SCREAM and short force-field minimization to determine the low-lying rotation combinations. The complete 360° sampling ensures that starting conformation is irrelevant to the method, and this has been verified with all available crystal structures.

Once predicted, a structure must be validated. Where ligand binding and mutation data are available, ligand binding site prediction is the method of choice. For some receptors, however, there is little ligand data. In addition, docking may introduce errors based on the binding site prediction method, and with methods still in development this technique may slow the validation of otherwise finalized structures. In the case of this β1 structure prediction, the timing necessitated swift prediction and validation, and there was a set of stabilizing mutation data avail-

able for validation.[91] Using SCREAM and minimization to mutate the predicted receptor structure to the mutants described for the impending crystal system, it is possible to validate the predicted β1 structure without the use of ligand binding predictions.

2.2 Methods

General Methods: All calculations were carried out using the DREIDING force field[92] with charges from CHARMM22.[93] Side chain placement was determined with SCREAM[90] Unless otherwise noted, all simulations were performed in the gas phase with a dielectric of 2.5.

2.2.1 Structure Prediction of the Turkey β1 Adrenergic Receptor

Prediction of transmembrane regions: We predicted the TM regions using PredicTM. First, we used NCBI BLAST[94] to obtain a set of 1,100 protein sequences homologous to the target receptor, including all adrenergic receptors from a variety of species. These sequences ranged from 7% to 56% sequence identity with turkey β1 (12 – 88% in TM regions). Next, we obtained the pairwise multiple-sequence alignment of these sequences using MAFFT,[95–97] modifying the result to compress gaps in the target sequence. We then used this alignment to determine the TM regions using a moving-window average hydrophobicity analysis based on the octanol scale.[89] The resulting hydrophobicity profile shows seven clear hydrophobic regions corresponding to seven TM helices (Figure 2.1). The final PredicTM he-

lices were determined by comparing the N- and C- termini of each helix to known "helix breaker" residues, typically glycine, proline, aspartic acid, glutamic acid, arginine, and lysine, and extending each TM region to include the nearest charged residue. The N-termini of TMs 1 and 3 were extended an additional three and four residues, respectively, upon comparison with the PredicTM-generated TM predictions of four related receptors: human β1 (P08588), turkey β4c (P43141), frog β1 (O42574), and human β2 (P07550). Table 2.1 shows the final TM predictions. Three sets of hydrophobic centers were calculated and carried forward to the next steps: "area" centers based on the area of the hydrophobicity peak, "rawmid" centers chosen as the geometric center of the helix, and crystal centers based on the 2RH1 β2 crystal structure.

Initial Helix Optimization (OptHelix) and Helix Bundle Assembly: After determining the sequence of each TM region, we optimized each TM individually using the OptHelix method.[71] Each TM is built individually as a polyalanine α-helix with Gly and Pro residues in locations corresponding to the target TM region. One extra residue was added to each end of each helix to stabilize the ends during this state. We minimized this helix with a conjugate gradient minimization, added Ser and Thr residues in appropriate locations using SCREAM, then performed short-term (2 ns) dynamics on each individual helix as described in Abrol, *et al.* The final helices were taken from the snapshot with an RMSD closest to that of the average structure during the last 1.5 ns of simulation. We removed the extra ala-

nine residues from the termini and arranged the finished helices in a seven-helix bundle using the orientations, distances, and tilt angles from the β2 human crystal structure (2RH1).[39] Although the alignment to the 1GZM bovine rhodopsin crystal template was evaluated, the resulting bundles had higher energies than those aligned to the β2 crystal, indicating it was a less favorable orientation. Only the results based on the β2 template are reported here in detail.

Helix Rotation Optimization (BiHelix/CombiHelix): The BiHelix/CombiHelix method determines the low-energy helix bundles for a system by considering combinations of helix rotations. This method is independent of the starting rotation. Twelve isolated pairs of helices are generated: 1–2, 1–7, 2–3, 2–4, 2–7, 3–4, 3–5, 3–6, 3–7, 4–5, 5–6, and 6–7. In each pair, each helix is rotated through 360° 30° increments, resulting in 144 combinations per pair. We optimize the sidechain orientations for each combination with SCREAM, then combine the pairwise energies using a mean field analysis to obtain the energy of each possible combination of rotations ($12^7 = 35$ million possibilities). We build the best 1,000 of these combinations, minimize for ten steps, then rank each of the top 1,000 bundles by energy.

For this work, the initial bundle was modified to allow polar interactions in the TM core to optimize fully. The ends of the TMs, those residues added at the end of the TM prediction process based on capping and consensus decisions, were mutated to alanine except for glycine and proline where appropriate. In the rest of the TM region, large, nonpolar residues (phenylalanine, isoleucine, leucine, va-

line, tryptophan, and tyrosine) were mutated to alanine. The BiHelix analysis was performed on this modified bundle. The best 1,000 structures were built with CombiHelix then mutated back to wild type residues. These structures were ranked by interhelical energies, determined by calculating the energy of each isolated helix, then subtracting it from the total calculated energy of the bundle.

2.2.2 Structure Mutation Calculations

We used SCREAM to replace sidechains in each mutant receptor and to simultaneously optimize residues within 5 Å of the mutated residue. This inclusion of surrounding residues is necessary to allow the receptor to adjust to the change in a reasonable way; for a mutation such as Val to Ala, surrounding residues may adjust to fill the gap, and for a mutation like Ala to Leu the opposite should be true. This 5 Å cavity was then minimized for ten steps to resolve steric conflicts, and single point energies were determined using the DREIDING force field. After this optimization, we converted the residues to neutral forms (to reduce possible bias due to long-range coulomb interactions) and determined both the total bundle energy and the interhelical interaction energy. A lower energy should correspond to a higher T_m, as both indicate a more stable structure.

2.3 Results and Discussion

Using PredicTM and the octanol scale, we obtained the seven TM regions based on the hydrophobicity of the β1 sequence and related sequences, shown in Fig-

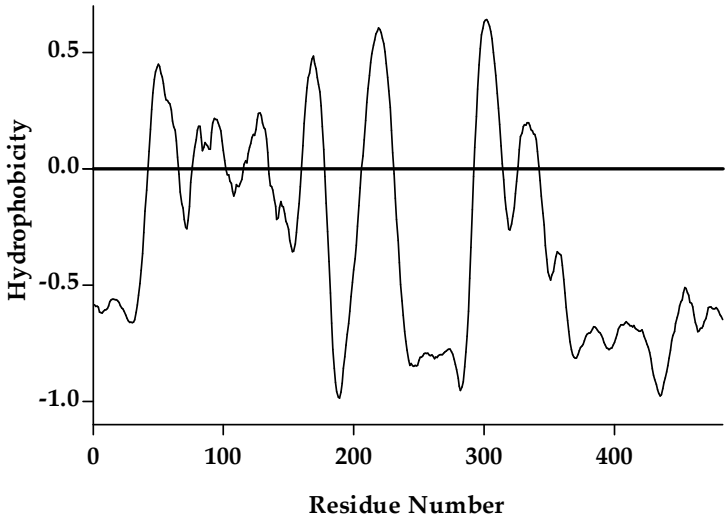

Figure 2.1: The hydrophobicity plot for $\beta 1$ turkey shows seven peaks representing to seven highly hydrophobic regions in the $\beta 1$ amino acid sequence. These regions correspond to the seven TM helices. These data were compared to similar plots for the human adrenergic receptors, and helices were lengthened to agree with the "consensus" predictions for the adrenergic family.

ure 2.1. These raw regions were lengthened based on comparison to the human adrenergic receptors; the differences between the raw predictions and the final predictions based on consensus are shown in Table 2.1. In general, the crystallographic helices are longer than the predicted helices, especially for TM3, since this prediction method relies solely on hydrophobicity rather than another indicator of helical character. In Chapter 4, this prediction method is supplemented with secondary structure predictions that consider helical character calculations as well as hydrophobicity. This consideration does produce helices closer in length to the crystallographic helices, but for the purposes of determining the TM core, the hydrophobicity and consensus analyses are sufficient to produce useful TM bundles.

After generating and optimizing the best 1,000 structures from the 35 million possibilities evaluated by BiHelix, several possibilities exist for choosing the best structure. Although ranking by minimized energy works consistently for BiHelix for crystallographic helices and 3D orientations, it is more difficult to encourage plausible combinations of rotations to appear within the top 0.05% of structures it is currently feasible to optimize. Selectively "alanizing" the input structure temporarily removes large nonpolar residues that might, as a result of the approximate template alignment, interfere with the formation of stabilizing polar contacts. It also eliminates the influence of polar or charged residues at the ends of the TM helices. In the native protein, these end residues can interact with the loops or charged lipid head groups. For the TM bundle calculations, however, these residues may distort the bundle rotation calculations and populate the top

Table 2.1: While the raw helices from PredicTM (β1 raw) are often truncated due to conserved polar and charged helices at the termini, comparison with other, related receptors provide plausible helices for structure prediction (β1 final). The 2VT4 β1 crystal sequence with α helical character is included for comparison, even though this data was not available at the time of these predictions. The raw midpoint hydrophobic centers are indicated in bold.

	β1 raw	AGMSLLMALV**V**LLIVAGNVLVIAAI
TM1	β1 final	QQWEAGMSLLMALV**V**LLIVAGNVLVIAAI
	β1 crystal	WEAGMSLLMALVVLLIVAGNVLVIAAIGST
	β1 raw	TNLFITSLACADL**V**MGLLVVPFGATLV
TM2	β1 final	TNLFITSLACADL**V**MGLLVVPFGATLVVR
	β1 crystal	TLTNLFITSLACADLVVGLLVVPFGATLVVRG
	β1 raw	CWTSLDVLCV**T**ASIETLCV
TM3	β1 final	SFLCECWTSLDVLCV**T**ASIETLCVIAID
	β1 crystal	GSFLCECWTSLDVLCVTASIETLCVIAIDRYLAI
	β1 raw	VIICTVWA**I**SALVSFLPI
TM4	β1 final	KVIICTVWA**I**SALVSFLPIMMH
	β1 crystal	TRARAKVIICTVWAISALVSFLPIMM
	β1 raw	YAIASSIISFY**I**PLLIMIFVYLRV
TM5	β1 final	RAYAIASSIISFY**I**PLLIMIFVYLRVYR
	β1 crystal	NRAYAIASSIISFYIPLLIMIFVALRAYREAKE
	β1 raw	GIIMGVFTL**C**WLPFFLVNIVNV
TM6	β1 final	KTLGIIMGVFTL**C**WLPFFLVNIVNV
	β1 crystal	REHKALKTLGIIMGVFTLCWLPFFLVNIVNVFN
	β1 raw	VFFNWLGY**A**NSAFNPII
TM7	β1 final	VFFNWLGY**A**NSAFNPIIYC
	β1 crystal	PDWLFVAFNWLGYANSAMNPIIYC

Table 2.2: When a selectively alanized structure is analyzed with BiHelix then the resulting bundles built and ranked by interhelical energies, many plausible structures appear in the top 100. The top 20 for turkey β1 aligned to the human β1 template using the raw midpoint centers are shown here.

H1	H2	H3	H4	H5	H6	H7	kcal/mol
30	240	60	210	30	30	270	-397.5
0	240	60	210	30	60	210	-383.9
270	30	60	210	60	0	330	-379.9
0	30	60	210	30	30	270	-379.8
30	240	60	210	30	30	120	-379.0
180	30	60	210	60	0	330	-378.5
0	30	60	120	60	0	330	-376.6
120	30	60	210	60	0	330	-374.7
30	330	60	210	30	30	270	-373.6
240	30	60	210	60	0	330	-372.5
0	300	90	120	60	0	60	-369.3
30	240	60	210	60	0	120	-367.9
0	30	60	210	30	30	330	-367.7
0	240	60	210	30	30	120	-366.9
0	30	60	210	30	30	0	-365.5
0	300	90	120	60	0	30	-365.4
90	30	60	210	60	0	330	-364.4
0	30	90	120	60	0	30	-363.8
0	30	60	210	60	0	330	-363.2
0	240	60	210	30	30	60	-362.9

1,000 structures with false positives. After these polar contacts are optimized by SCREAM and minimization, the larger nonpolar residues are restored and optimized. These full structures are then ranked by interhelical energy, which includes not only the polar interactions but any nonpolar interactions that may stabilize or destabilize the bundle. The best 20 of these structures are shown, ranked by interhelical energy, in Table 2.2.

The best structure by energy is based on the β2 2RH1 crystal structure template, using the raw midpoint hydrophobic centers (the geometric center of the hydrophobic helix prediction, before capping and TM extension to consensus se-

quences) to vertically orient the helices. This structure places most conserved residues in the appropriate positions for ligand binding and bundle stabilization. The TM1-TM2 salt bridge between Asn59$^{1.50}$ and Asp87$^{2.50}$ is closer to the protein-lipid interface than in related crystal structures, but the interaction itself (2.2 Å) is preserved. TMs 5, 6, and 7 are oriented such that Asp121$^{3.32}$, Ser212$^{5.43}$, Ser215$^{5.46}$, Trp303$^{6.48}$, Phe306$^{6.51}$, Asn310$^{6.55}$, Asn329$^{7.39}$, and Tyr333$^{7.43}$ are all accessible to the expected ligand binding pocket. This structure has an RMSD from the 2VT4 crystal structure of 3.7 Å and is close enough to a native conformation that it could be validated using the mutation data produced for the system. Figure 2.2 shows this structure with sidechains for binding site residues.

The experiments measured T_m for a series of mutations, defined as the temperature at which the protein begins to denature. After identifying point mutations that increased T_m for the mutant receptor, the group created and refined combinatorial sets of the best point mutations to eventually find the most stable mutant receptor that still preserved wild type binding and activation profiles. The final structure with the highest T_m contains six point mutations, but two of these (R68S and A282L) are expected to fall in the intracellular loops. As our methods focus primarily on the TM regions, any sets of mutations we consider must differ from one another without considering mutations sites in the loops. In order to ensure direct comparison between experimental and calculated energies, the mutations compared should also include the same loop residues.

With this in mind, we found three sets of mutations that fall within our pre-

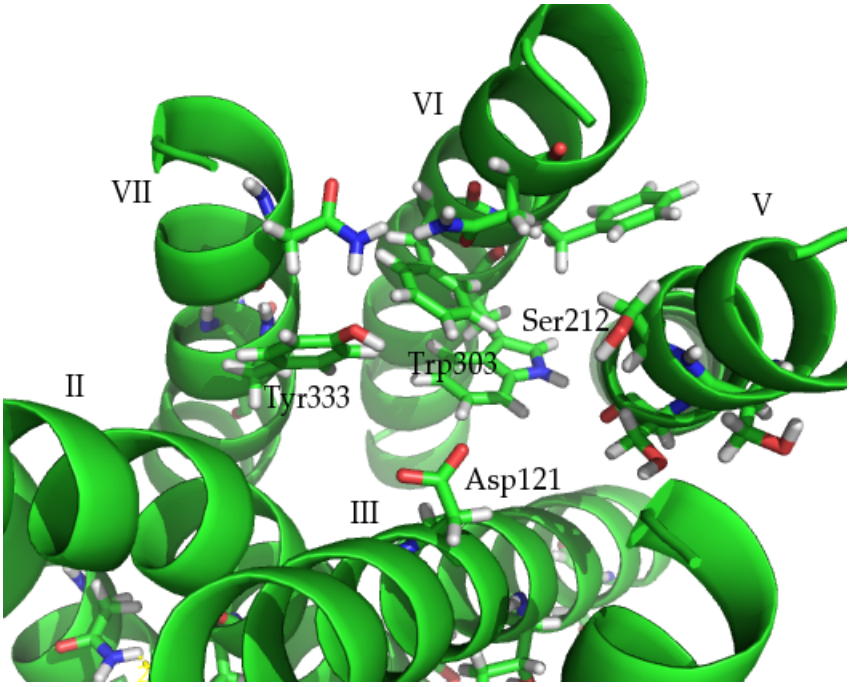

Figure 2.2: The top structure by interhelical energy shows most key residues for agonist binding oriented towards the binding pocket. Asp121$^{3.32}$ is particularly important for recognizing a protonated amine that is common to both agonists and antagonists. The conserved Trp303$^{6.48}$ and Phe306$^{6.51}$ are also oriented favorably. Ser215$^{5.46}$, part of the conserved adrenergic motif responsible for interaction with agonist catechol hydroxides, is also turned towards the binding pocket.

Table 2.3: Mutations probed include changes to TMs 2, 3, 5, and 7 in varying combinations. Set 1 includes m23, the combination of mutations eventually crystallized and reported as PDB 2VT1. T_m is listed relative to wild type.

	Label	T_m	2.53	2.61	3.40	5.58	5.61	7.32	7.37	7.44	7.48
Set 1	m6-11	7.4								A334L	
	m6-10	15.7				Y227A				A334L	
	m23	21.1	M90V			Y227A			F327A		F338M
Set 2	m4-6	3.3		G98A				D322A			
	m7-6	8.3	M90V		I129V						
	m7-5	13.5	M90V			Y227A					
	m7-7	13.5	M90V			Y227A					F338M
Set 3	m10-4	15.2	M90V				V230A			A334L	
	m10-8	15.6	M90V				V230A		F327A	A334L	
	m22	15.7	M90V			Y227A			F327A	A334L	
	m19	17.3	M90V			Y227A	V230A		F327A		F338M
	m18	17.9	M90V			Y227A			F327A		F338M

dicted TM bundle and affect T_m differently (summarized in Table 2.3). Set 1 includes mutations in TMs 2, 5, and 7: m6-11 with only A334$^{7.44}$L, m6-10 which adds Y227$^{5.58}$A, and m23 that adds M90$^{2.53}$V and replaces A334$^{7.44}$L with F327$^{7.37}$A and F338$^{7.48}$M. (This set includes the final combination of mutations that was characterized in the β1 crystal structure, m23.) Each of these experimental combinations also contains R68S (IC1) and A282L (IC3), residues not considered in the calculated energies. Set 2 begins with two pairs of different mutations: m4-6 with G98$^{2.61}$A and D332$^{7.32}$A, and m7-6 with M90$^{2.53}$V and I129$^{3.40}$V. Mutation combination m7-5 replaces I129$^{3.40}$V with Y227$^{5.58}$A, and m7-7 adds F338$^{7.48}$M to that combination. Set 3 contains the most combinations of mutations: five combinations of M90$^{2.53}$V, Y227$^{5.58}$A, V230$^{5.61}$A, F327$^{7.37}$A, A334$^{7.44}$L, and F338$^{7.48}$M. Each combination in set 3 includes R68S (IC1) that is not included in energy calculations.

The T_m and SCREAM energy data correlate significantly for sets 1 and 2, with

R^2 values ranging from 0.82 to 0.96. Set 2, shown in Figure 2.3, shows significant correlation for all four energy evaluations, while set 1 (Figure 2.4) correlates significantly for all evaluations except charged interhelical energies. For both sets, neutral evaluations show the strongest correlation and the closest fitting trendline. Neutral energies have been shown in the Goddard group to provide more reliable predictions, reduce noise in sampling, and increase resolution between structures that are expected to be closer in energy. Even in a dielectric, long-range coulomb interactions may play a role in force field energy calculations, a role that is not reflected in the reality of a protein solvated in water and lipid. Using a neutral residue scheme, the noise of long-range coulomb interactions is eliminated, resulting in more physical energies.

Despite the increased quality of energies calculated with a neutral residue scheme, the T_m values for set 3 are too close for a force field calculation to tell apart. The first three combinations shown in Table 2.3 span only 0.5, and the second two only 0.3, with the two groups only 1.6 apart, which is within experimental error for several of the cases. Even if precise, these differences correspond to an energy change of less than 0.01 kcal/mol which is not a reliable energy difference to probe with force field calculations. The combinations of mutations in sets 1 and 2 span 8 kcal/mol, which is a small change but reasonable to expect with the neutral residue paradigm. So while set 3 does not show agreement between the calculated mutation data and experimental T_m, the agreement with the wider energy ranges of sets 1 and 2 as well as the helix rotations placing known binding site residues in

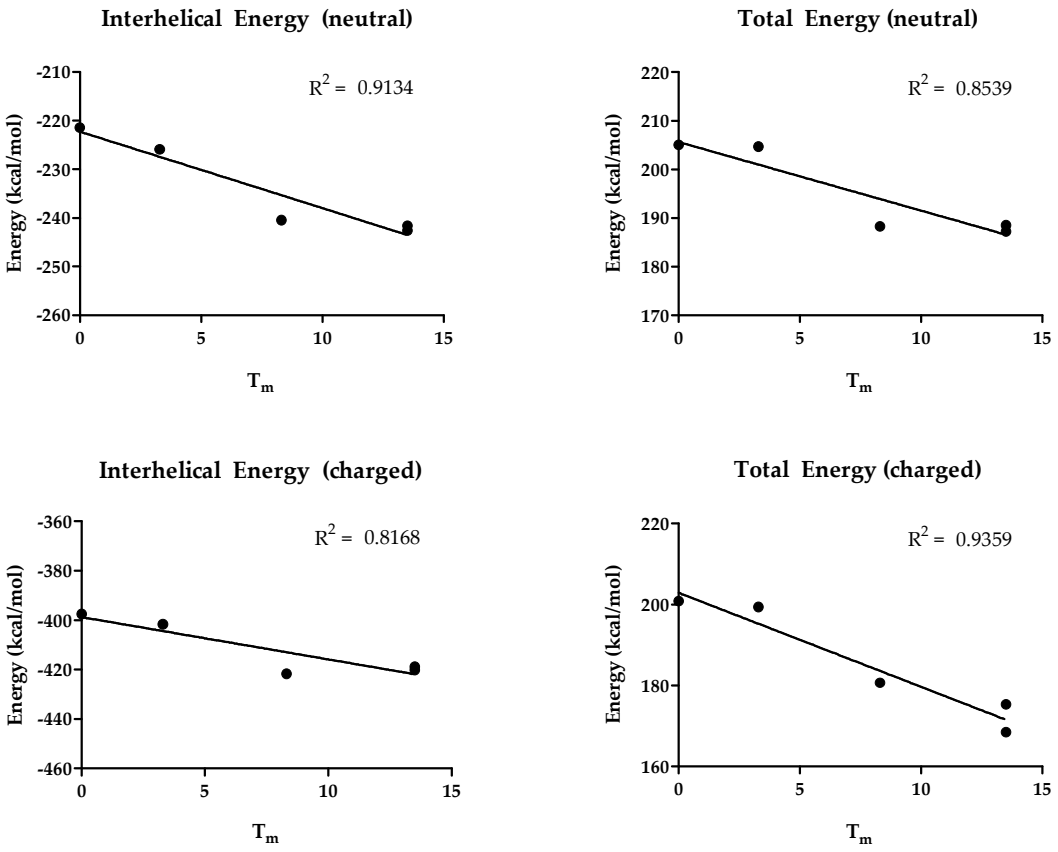

Figure 2.3: Set 2 shows the best overall correlation between calculation mutation energies and experimental T_m values, as all four energy evaluations correlate significantly with experiment. Note that the negative slope is expected, as an increase in T_m indicates the protein does not denature as easily and is therefore more stable, and a lower force field energy indicates a more stable structure.

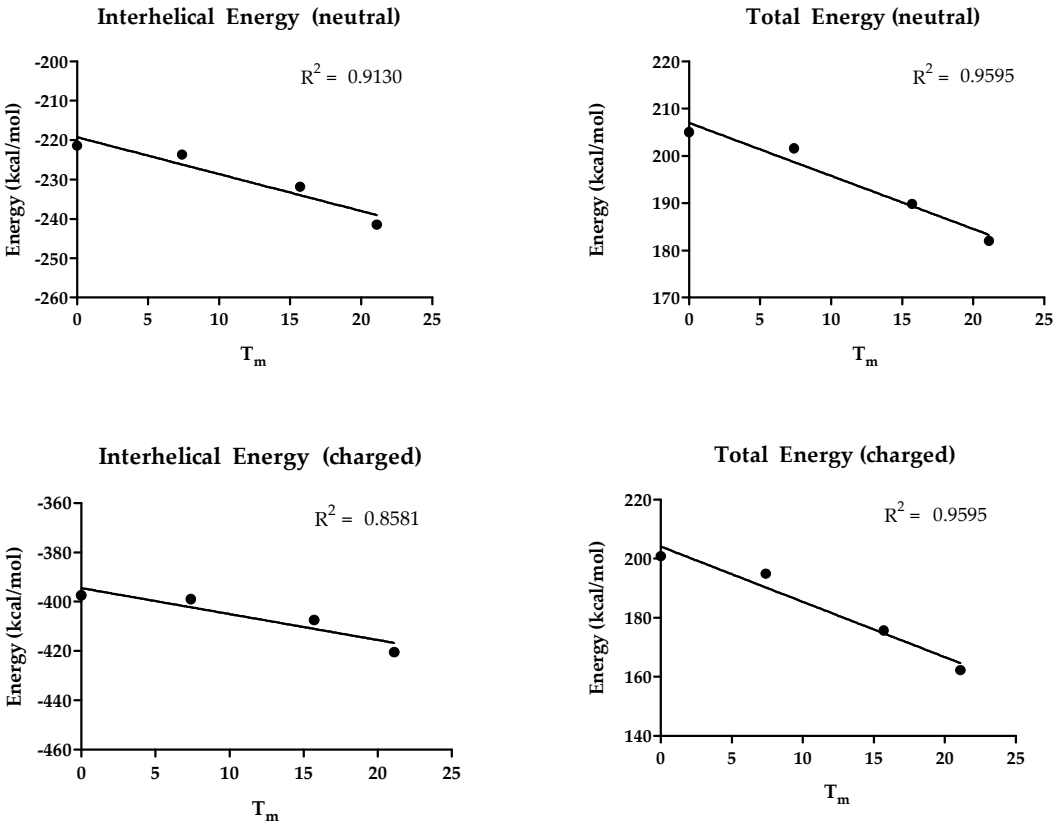

Figure 2.4: Set 1 shows significant correlation between calculated and experimental changes to turkey $\beta 1$ for both interhelical energy evaluations and the neutral total energy calculation. Using a neutral residue scheme allows for more accurate energy values, and should be considered a more reliable guide than the charged residue scheme.

expected locations near the agonist binding pocket give confidence that the predicted structure is indeed a valid conformation of the receptor.

2.4 Conclusion

Using the refined GenSemble method and stabilizing mutations reported in advance of the β1 crystal structure publication, a stable conformation for the turkey β1 adrenergic receptor was determined and validated. Although the resulting structure differed slightly from the eventual crystal data (CRMSD: 3.72 Å), the good agreement with the stabilizing mutations implies this structure is a valid alternate conformation for the system.

The validation study illustrates the utility of using neutral energies for comparison of structures. This approach has already enjoyed success in ligand docking studies (such as those discussed in Chapter 4 to validate homology models of the entire family of human adrenergic receptors) and is used here to confirm the agreement of this predicted structure with experimental data. It provides a sensible alternative to other methods of dampening spurious long-range coulomb interactions such as using a distance-dependent dielectric, and allows for confident calculation of smaller energy differences than a charged residue scheme.

Because the final predicted structure does have TMs 3, 5, 6, and 7 rotated such that important ligand binding residues are rotated inward towards the binding pocket, ligand docking studies may be performed. A docked β1-cyanopindolol complex with predicted intracellular and extracellular loops may also be equili-

brated in explicit water and lipid, similar to the procedures in Chapter 3. The structure may move towards the "native" structure observed in the crystal, or it may equilibrate to a different but similarly stable inactive conformation. This optimized structure may be a desirable starting point for further studies into GPCR activation, as it does not rely directly on a crystal structure for its initial conformation.

Chapter 3

Molecular Dynamics Simulations of Unbound, Agonist-Bound, and Inverse-Agonist-Bound Turkey β1

3.1 Overview

As structural data for GPCRs become more readily available, the primary focus of GPCR research has shifted towards the mechanism of activation. It is believed that GPCRs exist in equilibrium between active and inactive states, even in the absence of ligand, and that ligand binding occurs through a series of incremental steps that eventually stabilize one form of the receptor.[57,98] Drug development for GPCRs focuses by necessity on the "snapshot" perspective provided by X-ray crystal structures and predicted binding site structures, but the variety of conformations accessible to a target GPCR or family of targets complicates this design process. While the majority of GPCR-targeting drugs are antagonists or inverse agonists, a deeper understanding of the activation mechanism and its effect on the binding site can assist the development of agonist drugs to turn on rather than

shut down a specific receptor. In studying this problem with molecular dynamics (MD), we consider the movements of conserved regions thought to be important to activation or stabilization.

The role of the highly conserved D/ERY motif at the intracellular side of TM3 has been the subject of debate since the release of the β adrenergic and human adenosine receptors. In the bovine rhodopsin crystal structure, this triplet interacts with a conserved glutamate at the intracellular side of TM6, and is believed to function as an "ionic lock" that partially stabilizes an inactive form of the receptor. The β2, β1, and adenosine A_{2a} crystal structures lack this interaction, which may imply the ionic lock is not a conserved interaction. However, the long, flexible IC3 connecting TM6 to TM7 is treated differently in each crystal structure depending on what strategy was used to successfully crystallize the protein. It may be difficult, then to draw conclusions about the base of TM6 as its behavior and interactions may be governed by the presence (or absence) of IC3 and the associated G-protein. Fluorescence studies of human β2 do show evidence for an ionic lock interaction that is disrupted upon activation by all but the weakest agonists, even though this interaction is not observed in the crystal structure.[25] For this debate, MD can offer an opportunity to explore the stability of these interactions in the presence of a complete IC3 loop and bound ligands as well as explicit solvent.

Class A GPCRs also share a conserved WXPFF motif in the center of TM6 that affects both ligand binding and receptor activation. The orientation of the Trp in this motif has been implicated as an early step in activation, as it shifts from

a "vertical" orientation with the aromatic sidechain parallel to the TM helix to a "horizontal" orientation perpendicular to TM6.[99] Yao, *et al.* illustrated the role of this rotation for β2 occuring for full agonists and strong partial agonists.[25] In all available crystal structures, Trp$^{6.48}$ is observed in the vertical, inactive orientation. A recent computational study of bovine rhodopsin also showed motion of this tryptophan "toggle switch" as well as involvement of the ionic lock and TM helix motions.[100]

While motions of TM6 seem highly likely to play a role in activation, TM5 may also play a role that has only recently been explored. The crystal structure of active rhodopsin[32] as well as the squid opsin structure[101] show the TM5 helix elongated past the lipid membrane, an important difference between the active and inactive structures. Secondary structure predictions of several GPCRs, including the adrenergic receptors, show the possibility of this helix being extended. A virtual ligand screening study found that shifting TM5 of the human β2 structure inwards towards the binding pocket improved predicted agonist binding, another possible link between TM5 movement and activation.[102] Mutations of TM5 in rhodopsin affect activation by disrupting possible interactions with TM3.[103–105] The predicted helical region of turkey β1 contains several charged residues that can interact with either TM3 or TM6 to form an alternative ionic lock, but the precise nature of this helix's contribution to activation remains unknown.

Some GPCRs are naturally constitutively active, showing second messenger activity even in the absence of ligand.[44] This complicates crystallization but allows

for easy study of ligand effects and GPCR deorphanization.[106] Systems with less constitutive activity have been engineered for study purposes, providing further insight into which parts of the TM bundle play a significant role in activation. Mutations in TMs 3 and 7 of the AT1 receptor create a constitutively active mutant.[107] Active mutants of β2,[108] histamine H2,[109] and α1b[26] receptors result from mutaton of the highly conserved DRY motif at the intracellular end of TM3. Mutations in TM6 cause constitutive activation for the rat μ opiod receptor,[110] the α1b adrenergic receptor,[111] and bovine rhodopsin.[112] One site in IC3 close to the beginning of TM6 also produces a constitutively active form of α1b.[113] Additionally, Rasmussen *et al.* discovered activating mutations in the DRY motif of β2 also caused a counterclockwise rotation of TM6.[108] It is clear that TM6 plays a role in the transition between active and inactive states, but TMs 3 and 7 likely also play a role.

The challenges posed by the dynamic behavior of GPCRs have led some researchers to explore stabilizing mutations in an effort to shut down GPCR activation and create a less dynamic, more easily crystallized receptor. The best stabilizing mutations for turkey β1 described in Chapter 2 occur in TMs 2, 5, and 7 as well as two intracellular loops (IC1 joining TMs 1 and 2, and IC3 joining TMs 5 and 6). Joining the intracellular ends of TMs 3 and 6 results in an inactive form of rhodopsin.[114] Recently, a stabilized variation of the leukotriene B_4 receptor BLT1 was engineered by including a metal binding site to connect the intracellular ends of TMs 3 and 6.[43] The common helices for both activating and stabilizing mutations are nearly always TMs 3 and 6, especially the conserved D/ERY motif in

TM3. These changes usually do not affect ligand binding dramatically, except that the stabilizing mutations may decrease agonist affinity in accord with the ternary complex model.

To test the role of ligand effects in these possible conformation changes in turkey β1, MD with a variety of ligands can be performed and the resulting motions analyzed. As a receptor with little constitutive activity, β1 offers an opportunity to observe the transition to an active state. Although inactive, the β1 crystal structure does not contain the putative ionic lock. This interaction as well as the missing intracellular loop, can be built in for MD studies. Through 10 ns MD of the apo, inverse agonist bound, and agonist bound structures, we can observe the motions of TM helices that begin the process of GPCR activation.

3.2 Methods

NVT Langevin dynamics were carried out using NAMD[115] and the CHARMM force field.[93] The complete systems include about 38,000 atoms each, including water. The SHAKE algorithm was used to prevent drift from fast modes.[116] The bath temperature was set to 310K, and a Particle Mesh Ewald (PME) grid was used to calculate long-range interactions. For both minimization and equilibration, femtosecond timesteps were used. Before analysis, all frames were aligned to the last frame to minimize noise from structural drift.

The 2VT4 crystal structure[41] served as the starting point for this study. The crystal ligand was removed, as were the crystal ions and other accompanying

molecules with the exception of the cyanopindolol ligand and the seven crystal waters. This structure was minimized, then the IC3 loop modeled with PLOP.[117,118] This loop was built in 20-residue sections, beginning with the 10 residues at the intracellular ends of TMs 5 and 6, then the 10 residues attached to those, and so on until the entire loop was built and connected. The last seven residues of IC3 are found to be helical by the APSSP2 secondary structure prediction method,[119] so the helical constraints available in PLOP were used to build these residues as an α helix. In order to test the role of Arg139$^{3.50}$ and Glu285$^{6.30}$ in stabilizing the receptor through an ionic lock, this interaction was built using SCREAM.[90] This modified receptor was minimized for 1,000 steps, ions were added, then it was placed into lipid (palmitoyloleoyl phosphatidylcholine) and water (TIP3) and the lipid/water minimized for 5,000 steps and equilibrated for 0.5 ns. This entire system was minimized for 5,000 steps, then equilibrated for 10 ns. Figure 3.1 shows the energy fluctuations for the system fall after 6,000 steps.

The initially optimized receptor-solvent system became the foundation for ligand dynamics, both inverse agonist and agonist. The protein after initial optimization, but before the 10 ns equilibration, was modified to accommodate either the inverse agonist cyanopindolol or the agonist isoproterenol. For cyanopindolol, the crystal orientation of the ligand was merged with the optimized apo protein. The binding site, then the entire protein, was minimized for 5,000 steps. This system was then equilibrated for 10 ns. We predicted the agonist orientation using GenMSCDock in a procedure similar to that described in Chapter 4. Rather than

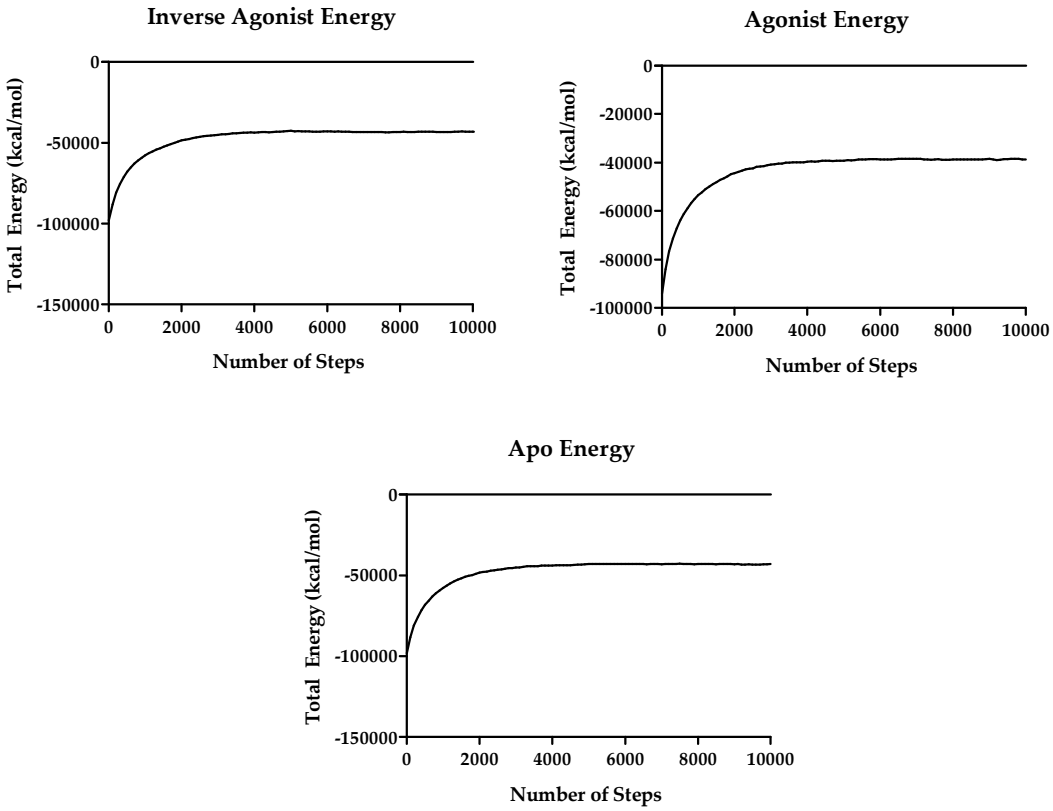

Figure 3.1: All three simulations reached a stable energy within the first 10 ps of the simulation, with the isoproterenol bound structure higher in energy, on average, than the apo or cyanopindolol bound structures.

DarwinDock, MSCDock based on USCF Dock 4[120] generated the docked poses using diversity and voronoi clustering, but the final optimization of the structures was similar. While both epinephrine and norepinephrine were docked, the best structure by energy that agreed with general adrenergic mutation data was a norepinephrine structure. This structure was converted to isoproterenol, a potent β selective agonist. Isoproterenol binds β1 more tightly than do the endogenous agonists, and activates the receptor more quickly.[121] Once built, this ligand was merged with the equilibrated apo protein and initial positions for the crystallographic waters, then optimized similarly to the cyanopindolol case. In both cases, energy fluctuations subsided within the first 10 ps of the simulation, shown in Figure 3.1.

3.3 Results and Discussion

Overall, the ligand bound structures shifted more than the apo protein. The RMSD over 10 ns changed less for unbound structure than for the ligand bound structures, as the ligand bound structures moved towards either an active or an inactive conformation (Figure 3.2). The ionic lock also does not move significantly over the course of the simulation, with a heavy-atom distance between Arg139$^{3.50}$ and Glu285$^{6.30}$ stable around 4 Å. This interaction does break in the last 2 ns of the apo simulation, as Arg275$^{6.20}$ begins to interact with Glu285$^{6.30}$ (illustrated in Figure 3.3). The lock does remain stable in the ligand-bound simulations, however, indicating that the lock interaction is relatively stable and that the competing residues

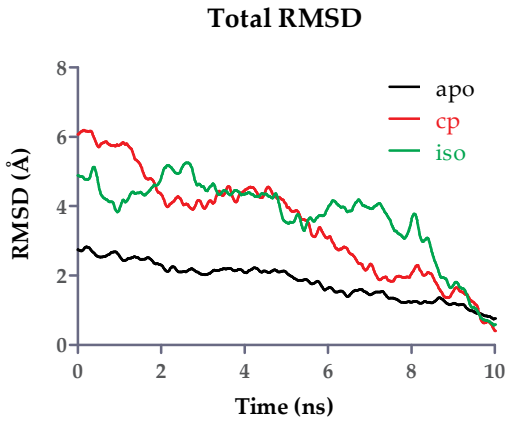

Figure 3.2: Total protein root-mean squared deviation is shown for all three MD simulations: apo protein in black, cyanopindolol (cp) bound in red, and isoproterenol (iso) bound in green. The apo protein moves the least, while the ligand bound forms move slightly more as they begin motion towards either a completely active or inactive conformation.

on IC3 are less likely to form a stable interaction due to the flexibility of the loop when ligand is present. The ionic lock distances are shown in Figure 3.4.

The TM6 toggle, Trp303$^{6.48}$, also does not shift significantly during the simulation (Figure 3.5). In β2 simulations in the Goddard group, this residue does move, while the ionic lock remains stable over the 50 ns of that simulation. This system differs from β2 in that turkey β1 has less constitutive activity than β2; this may indicate that the inactive state is considerably more stable for turkey β1 than for β2. The hallmarks of activation may not occur until later in the activation process for this receptor.

The beginning of the receptor-wide conformational change as illustrated here occurs with small movements of the TM helices themselves shifting towards or away from each other depending on the ligand involved. For the apo protein, the helices move relatively little. After about 7 ns of isoproterenol-bound dynamics,

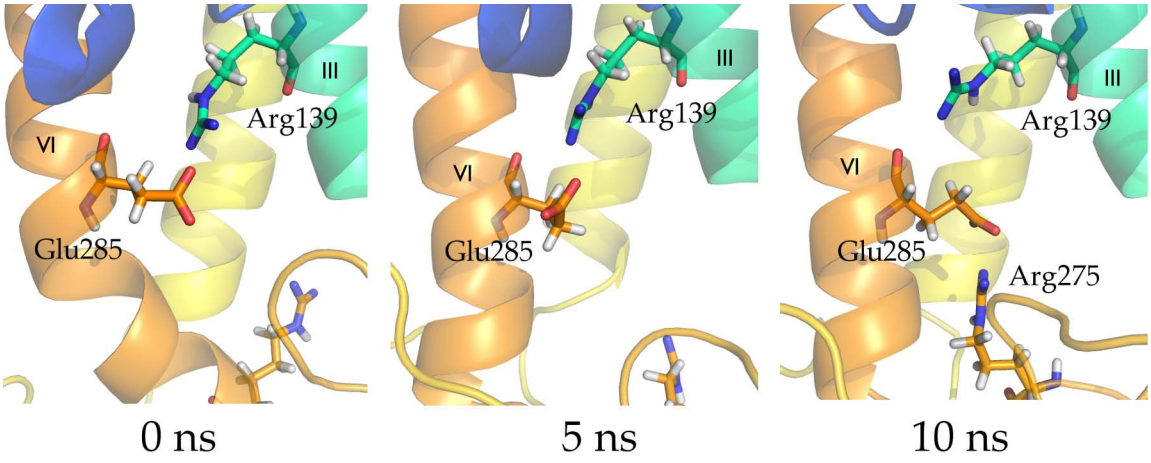

Figure 3.3: Although stable in the ligand bound simulations, the TM3-6 ionic lock breaks during 10 ns MD of the apo protein. In the last 2 ns, Arg275$^{6.20}$ in the third intracellular loop shifts towards TM6 and interacts with Glu285$^{6.30}$

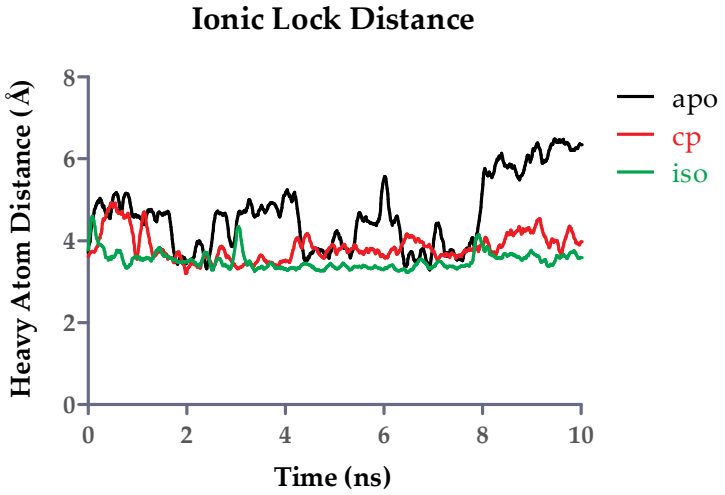

Figure 3.4: The ionic lock distance does not change significantly during the course of ligand-bound MD. At the end of the apo protein simulation, a residue from IC3 shifts to interact with Glu285$^{6.30}$, disrupting the ionic lock. Further studies will probe the stability of this interaction with alternate conformations of the receptor.

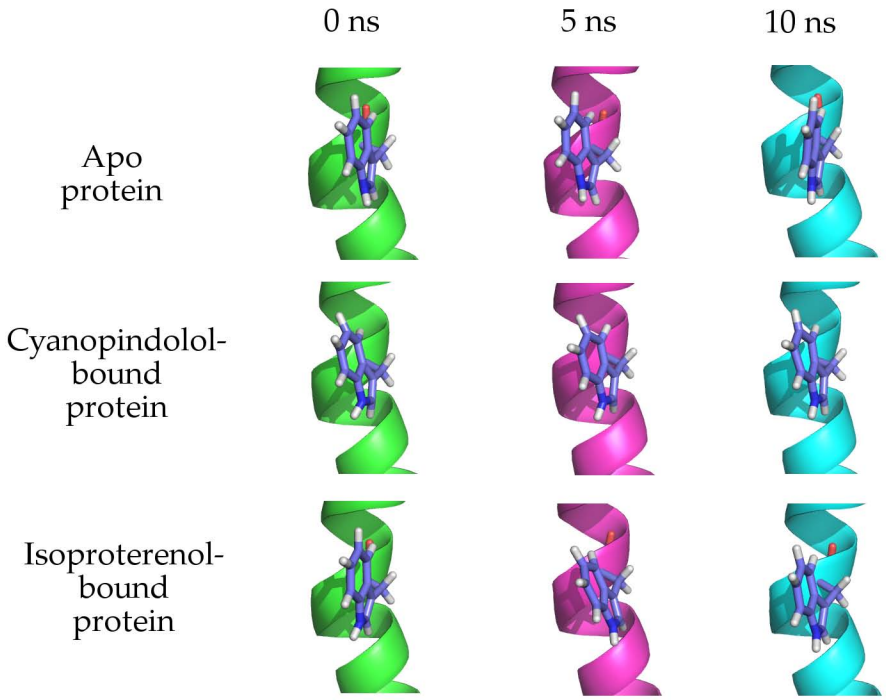

Figure 3.5: Although the residue fluctuates slightly, the 90°switch from the vertical rotamer of Trp303$^{6.48}$ shown here and the horizontal rotamer thought to signify an activated state is not observed during the course of this study.

however, the total RMSDs of TMs 3, 5, and 6 begin to shift (Figure 3.6). Measuring the distance between the C and N termini of the TM helices gives a clearer sense of the helix motions involved. Figure 3.7 shows the distances between the tops of TMs 5 and 6, the bottoms of TMs 3 and 6, and the tops of TMs 1 and 7. The tops of TMs 5 and 6 move away from each other slightly in the apo and agonist bound forms, but towards each other as the inverse agonist binds. This is accompanied by a smaller inward shift by the intracelluar ends of TMs 5 and 6. If the intracellular end of TM5 converts to a longer α helix extending into the cytosol, this inward movement may signify an alternate ionic lock between TMs 5 and 6 replacing the interaction between TMs 3 and 6.

As TM5 moves towards TM6, TM3 also shifts with respect to TM6. The distance between the intracellular ends of TMs 3 and 6, shown in Figure 3.7, indicates a trend towards breaking the ionic lock for the activated case (in green) and stabilizing it for the inactive case (in red). The distance fluctuates for the apo case, perhaps indicating the region is flexible. It may also be sensitive to the presence of G protein, which is not considered in this study. As the ionic lock only breaks in the last 2 ns of the apo simuation and not at all in the ligand-bound cases, this helix motion is likely the very beginning of the motion that eventually breaks the ionic lock and continues towards activation rather than the lock releasing the helices to move. These motions can also be seen in the changes to θ (defined as helix tilt angle, illustrated in Figure 3.9) over the course of the simulation. While the apo protein holds θ for the TM5 generally constant, TM5 in the isoproterenol-bound

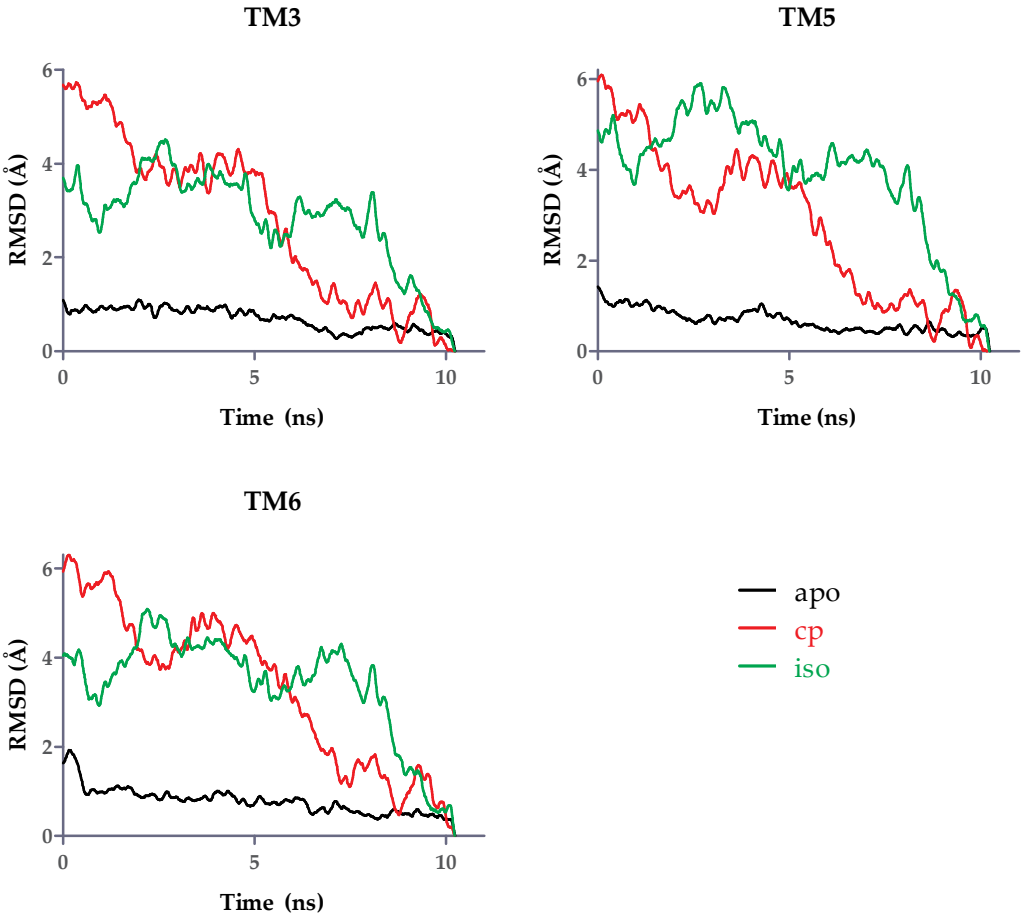

Figure 3.6: While the apo protein remains relatively still over the 10 ns of equilibration, the isoproterenol-bound system begins to diverge after 7 ns.

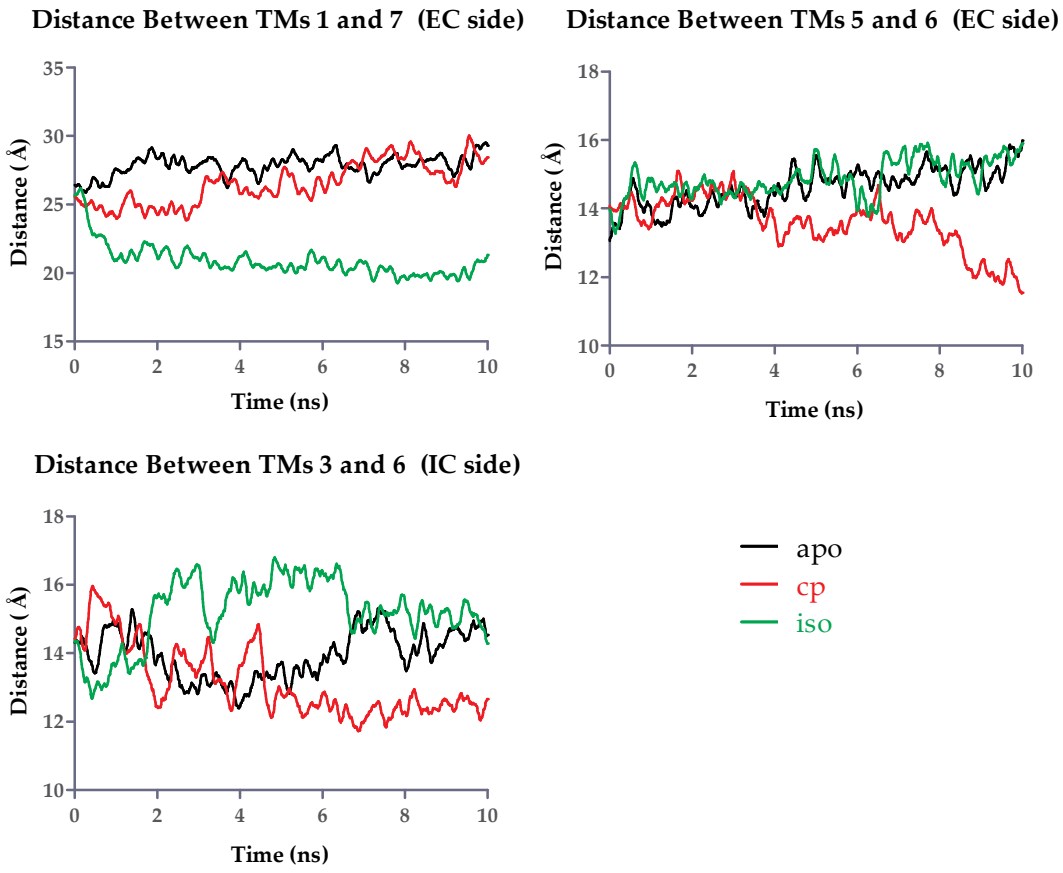

Figure 3.7: Distances between Cα of the first or last residues of TMs 3, 5, 6, 1, and 7. These show the degree to which the ends of the helices are moving towards or away from one another, depending on the ligand bound. Inverse agonist binding causes the intracellular ends of TMs 3 and 6 and the extracellular ends of TMs 5 and 6 to move closer, while agonist binding causes TM1 to move closer to TM7.

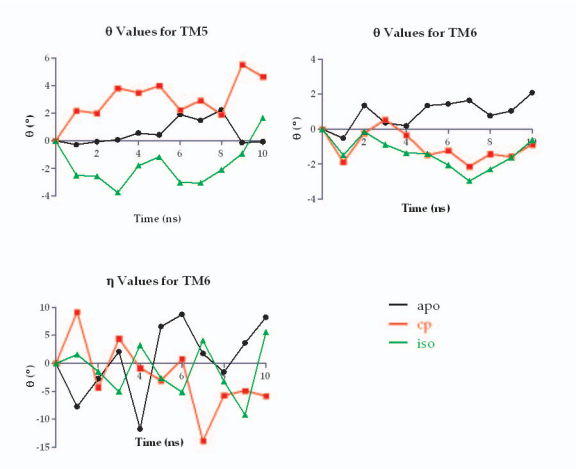

Figure 3.8: The changes in η for TM5 compared to TM6 reveal the closing of the distance between the intracellular ends of TMs 5 and 6 is due more to helix tilting for TM5 than any concerted motion by TM6. With ligand bound, (green and red) the rotation fluctuates but does not move decisively in one direction when the ligand is bound.

structure tilts "back," moving the intracellular side in towards the TM bundle, a motion that is mirrored in the cyanopindolol-bound structure. TM6, on the other hand, is flexible but there is no strong trend towards tilting in one direction or another. While η rotation of TM6 has been implicated in receptor activation (shown in Figure 3.8), the overall helix rotation fluctuates and does not show a dramatic, concerted change. These results suggest the motion of surrounding helices initiate the transition, allowing TM6 to rotate later during activation.

An additional helix motion is observed over the course of this simulation: the intracellular end of TM1 moves slightly towards TM7 and remains close in the agonist bound case. For the apo and inverse agonist cases, the distances remain constant for most of the simulation and increase a small amount. The areas of contact between the two TMs are primarily hydrophobic, as the conserved Asn339$^{7.49}$ is not positioned to interact with Asn59$^{1.50}$. This interaction could be tested ex-

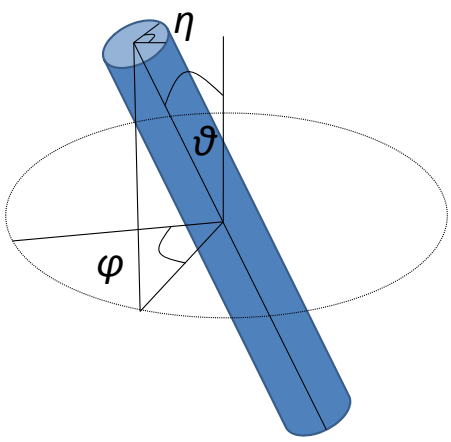

Figure 3.9: Decomposition of helix rotations produces three rotational modes: η, the rotation of the helix about the helical axis; θ, the helical tilt; and ϕ, the sweep angle of the helix.

perimentally by mutating some of these hydrophobic residues to polar ones to encourage TM1 to remain close to TM7 and facilitate activation of the system. Val62$^{1.53}$ is positioned close to Tyr343$^{7.53}$, the conserved tyrosine in the NPXXY motif. If Val62$^{1.53}$ is mutated to a polar or negatively charged residue, it might strengthen the interaction between the two helices. In addition, as Tyr$^{7.53}$ is conserved throughout Class A GPCRs, this may be a general interaction useful for a variety of systems.

The β1 crystal structure contains seven water molecules in the receptor structure, illustrated in Figure 3.10. Three of these remain stable during the isoproterenol simulation: one in the EC3 loop halfway between the extracellular ends of TMs 6 and 7, one in the EC1 loop close to TM2, and one associated with the NPXXY motif in TM7. Two of these remain stable during the apo protein simulation, the ones closest to TM7. Only the water associated with the NPXXY motif stays relatively fixed throughout the cyanopindolol simulation. It mediates the interaction

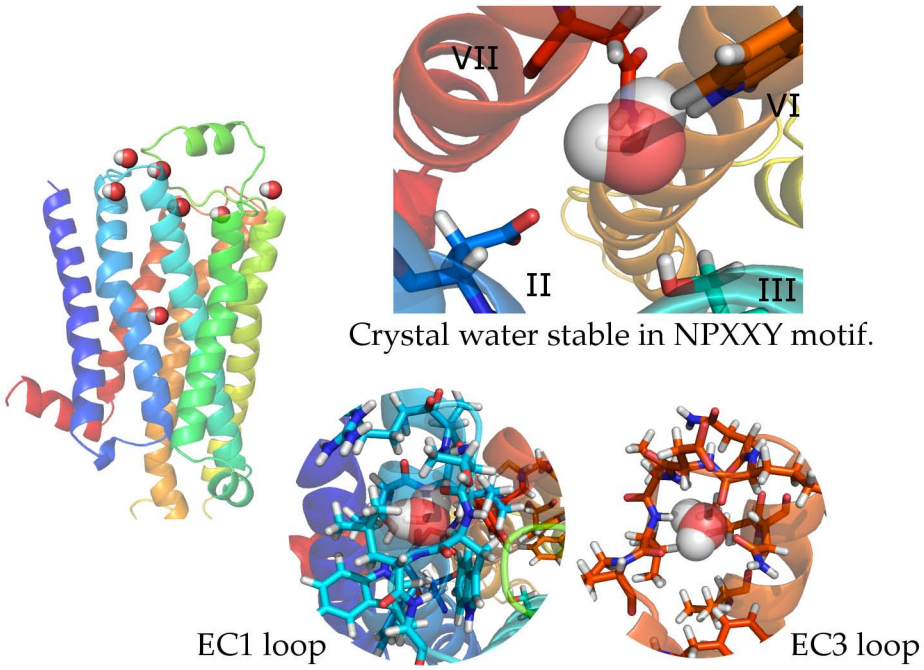

Figure 3.10: The crystal water stable in the NPXXY motif is shown in the upper right, with the conserved Asn335$^{7.49}$ and Asp87$^{2.50}$ holding it in place. The apo protein simulation shows the water associated with the EC3 loop, shown in the bottom right, is also stable. The EC1 loop water in the center bottom remains associated with the loop through the isoproterenol-bound simulation only. All seven crystal waters are illustrated at left.

between the conserved residues Asp87$^{2.50}$ and Asn335$^{7.49}$, a region associated with ion regulation in α2a[21] and universally conserved in Class A GPCRs. It is possible this is an ion binding site and the crystal water for this case is an ion responsible for stabilizing the TM2-7 interaction.

3.4 Conclusion

MD analysis of the turkey β1 crystal structure reveals changes towards an equilibrated apo protein and helical shifts during the first 10 ns of full activation or deactivation. TMs 1, 3, and 5 move dramatically differently based on whether the

inverse agonist or full agonist is bound, but the hallmarks of activation involving TM6 do not begin during the 10 ns studied. TMs 1, 3, and 5 must therefore move in order to allow TM6 to rotate and change its conformation for full receptor activation.

The motion of TM1 towards TM7 upon agonist binding indicates a strengthened interaction between these helices that may contribute to constitutive activation. A mutation to TM1 to tether the helix to a conserved polar residue on TM7 may be useful in creating constitutively active mutants of orphan receptors for the purposes of identifying a native ligand. It may also stabilize an active state of the receptor enough to allow crystallization and X-ray characterization of an active form.

This study covered the very beginning steps of β1 activation, and highlighed the movements of TMs 3, 6, and 5 that initiate the transition of the receptor from the inactive to the active state. The agonist equilibrated structure provides another starting point for further investigation into activation. With modifications to this equilibrated structure such as elongating the TM5 α-helix, creating an alternate salt bridge between TM5 and 6, and rotating Trp303$^{6.48}$ to the active conformation, another MD simulation may illuminate which of these changes is stable in an active conformation and more likely to contribute to receptor activation.

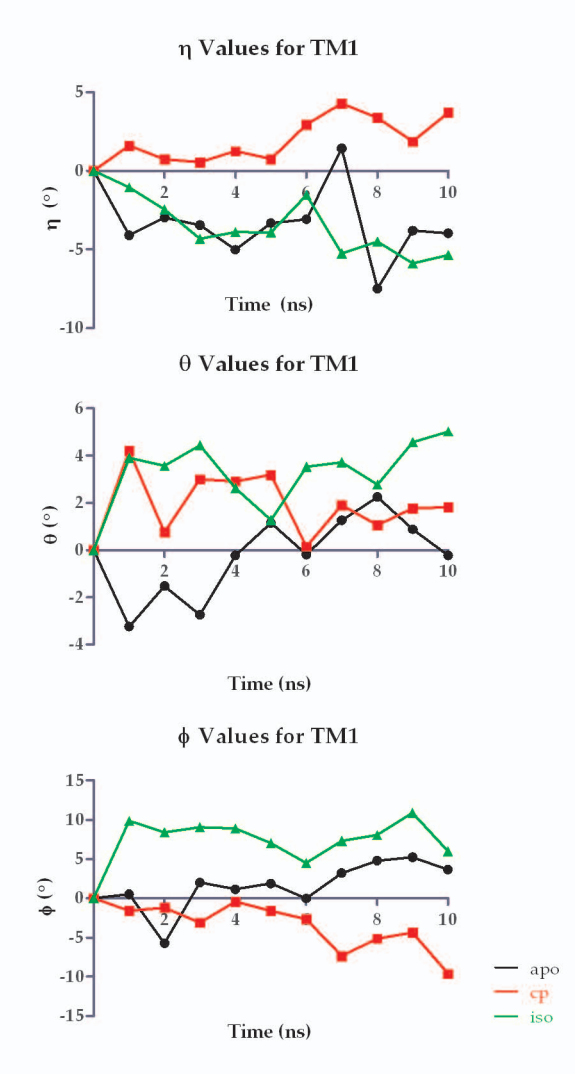

Figure 3.11: Changes to η, θ, and ϕ for TM 1 over the course of 10 ns full-solvent dynamics.

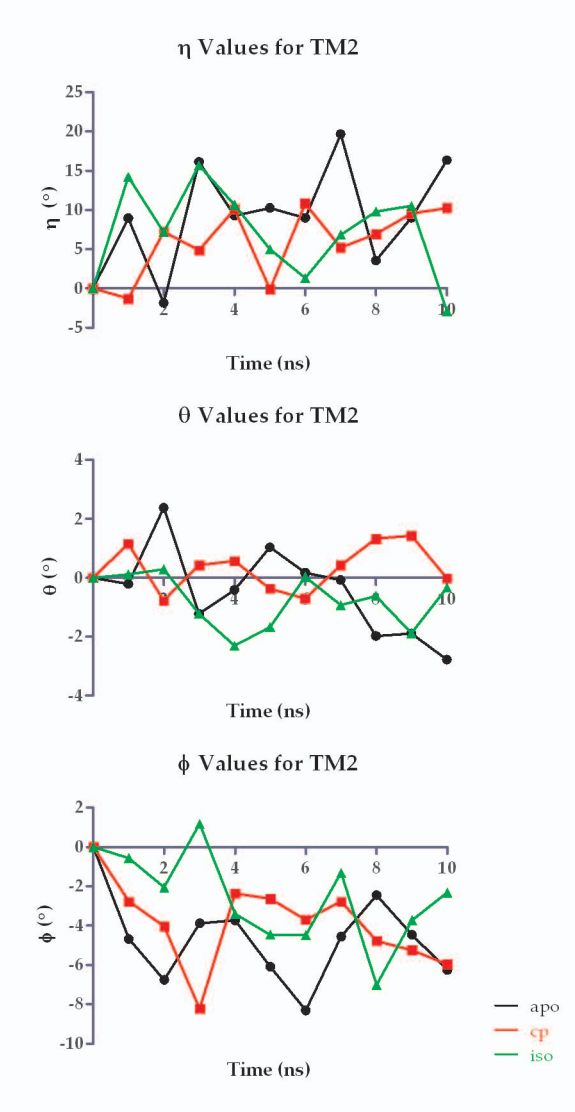

Figure 3.12: Changes to η, θ, and ϕ for TM 2 over the course of 10 ns full-solvent dynamics.

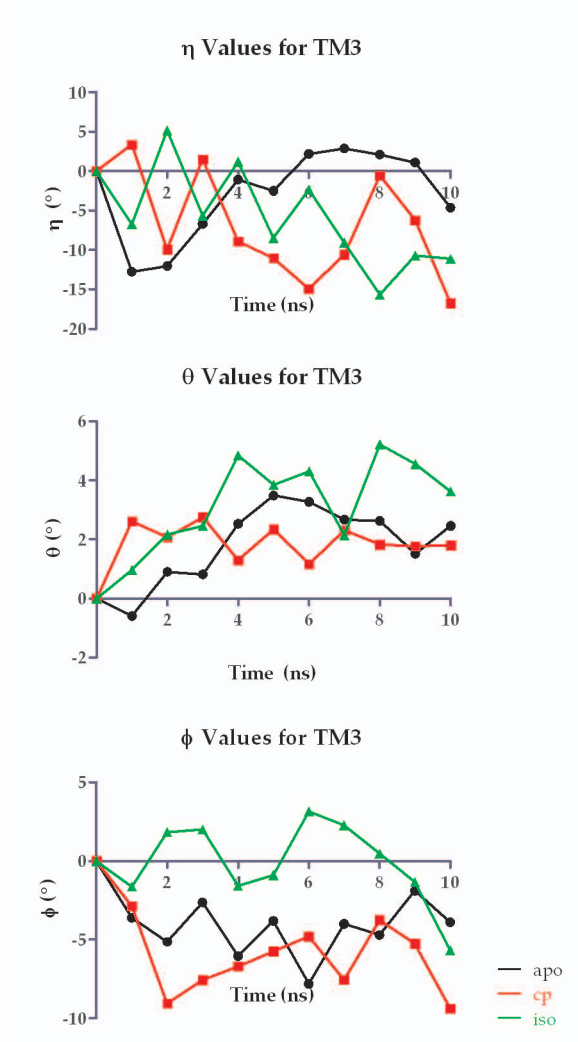

Figure 3.13: Changes to η, θ, and ϕ for TM 3 over the course of 10 ns full-solvent dynamics.

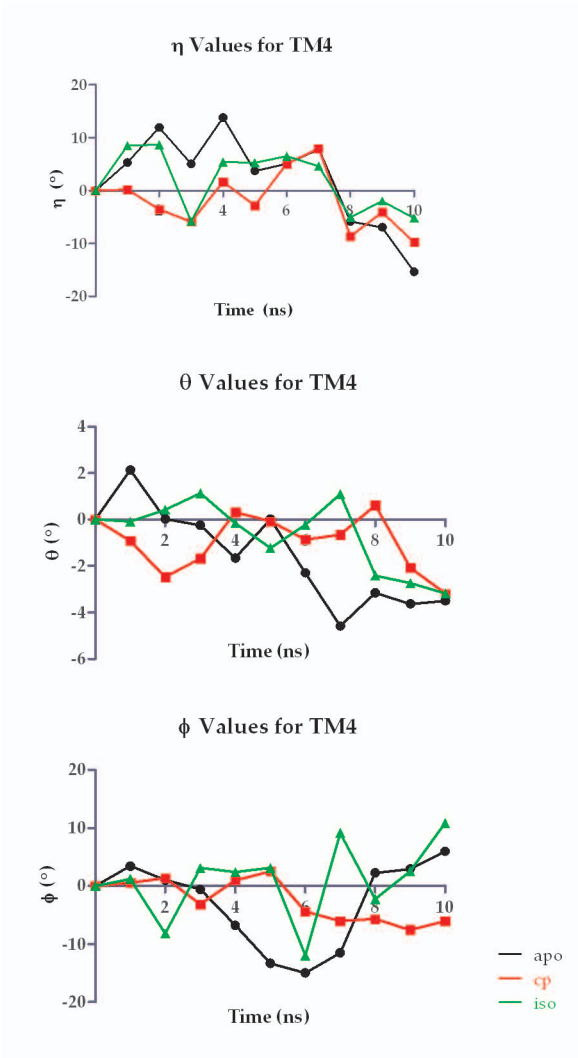

Figure 3.14: Changes to η, θ, and ϕ for TM 4 over the course of 10 ns full-solvent dynamics.

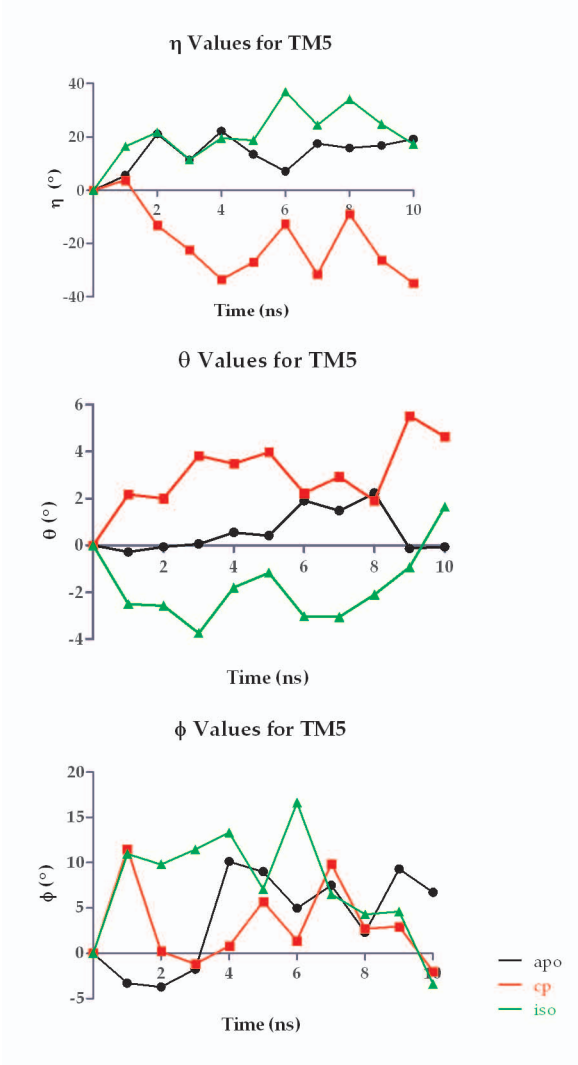

Figure 3.15: Changes to η, θ, and ϕ for TM 5 over the course of 10 ns full-solvent dynamics.

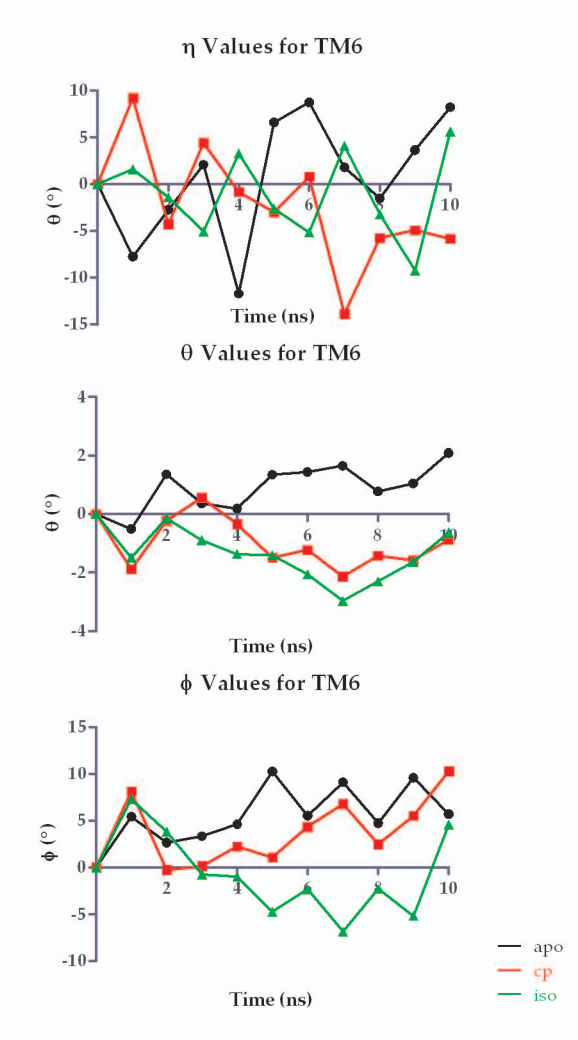

Figure 3.16: Changes to η, θ, and ϕ for TM 6 over the course of 10 ns full-solvent dynamics.

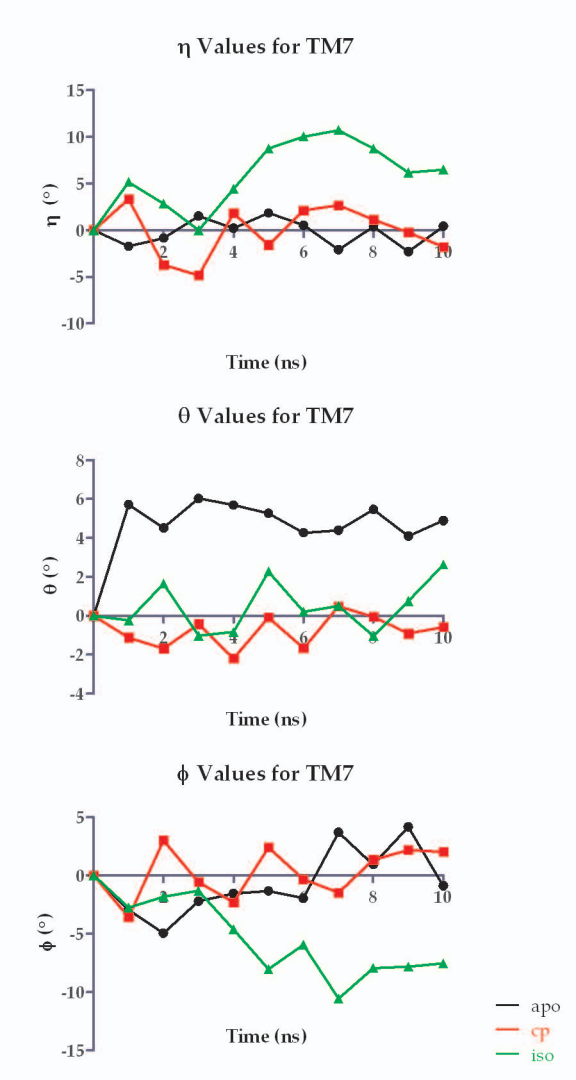

Figure 3.17: Changes to η, θ, and ϕ for TM 7 over the course of 10 ns full-solvent dynamics.

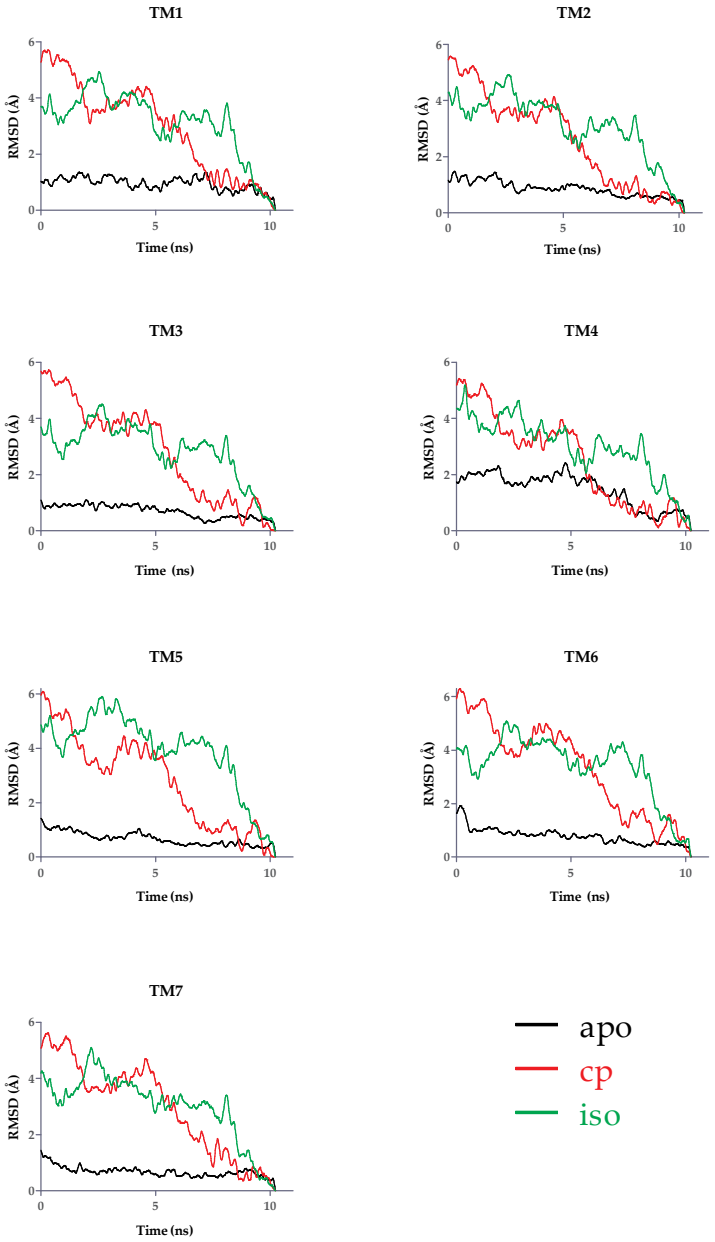

Figure 3.18: TM helix RMSD for apo (black), isoproterenol-bound (green), and cyanopindolol-bound (red) turkey β1 shown for each helix individually.

Chapter 4

Homology Models of Human Adrenergic Receptors

The increased availability of high-resolution GPCR crystal structures has enabled deeper investigation into homology modeling as a method for GPCR structure prediction. Most homology models currently available were built based on bovine rhodopsin, but low sequence identity between rhodopsin and most GPCR targets of interest for drug development casts some doubt on the utility of these models. The most recent structures for two adrenergic receptors offer the opportunity to create high-quality homology models for the entire human adrenergic GPCR family. Using the homology model approach combined with the BiHelix / CombiHelix method for helix η rotation determination, structures for the eight remaining human adrenergic receptors were built. These structures were validated with ligand docking studies, and the predicted binding sites offer some insight into subtype selectivity in the family.

4.1 Overview

Homology modeling, a method to model a GPCR structure based on an existing experimental structure, is widely used to determine overall structure, binding sites, and subtype selectivity for GPCRs.[12] For several years, only one GPCR, bovine rhodopsin, was available for this kind of modeling.[28–36] Most receptors of interest are human GPCRs from other subtypes however, with low sequence identity to bovine rhodopsin, complicating the process of building a homology model.[65–67] The publication of several other GPCR crystal structures, including the human β2 adrenergic receptor (β2),[37–40] the closely related turkey β1 adrenergic receptor (β1),[41] and the human adenosine A_{2A} receptor[42] revealed close structural similarity among the structures. This, as well as advances in ligand-steered homology modeling,[122] have inspired more confidence in the homology approach. Differences between individual receptors can arise, however, and they may be significant enough to mislead conclusions during binding site studies or enquiry into the mechanism of activation.[123] For receptors with no closely related crystal data, *ab initio* structure prediction methods such as those described in Chapter 2 are in continuing development. For receptors with high similarity to existing crystal structures, however, homology modeling remains a viable option for quick and accurate determination of useful receptor data, especially for binding site studies. The recent publication of two adrenergic GPCRs creates a unique opportunity for high-quality homology modeling of the other adrenergic receptors.

Table 4.1 shows the sequence identity of the adrenergic receptors as well as

Table 4.1: Identities between the human β2 and turkey β1 GPCRs with the other adrenergic receptors range from 24%–60% overall and up to 85% in the TM regions.

	β2 Human		**β1 Turkey**	
	All%	TM%	All%	TM%
β1	41	68	60	85
β3	42	59	39	63
α1a	28	41	26	44
α1b	30	44	27	44
α1d	30	43	27	43
α2a	26	39	27	49
α2b	24	38	24	46
α2c	24	38	24	44
β2	–	–	42	70

the identity of the TM regions used for the homology models. The nine adrenergic receptors are closely related to one another, even between different species, as shown in the relationship between the turkey β1 receptor and the human β2 receptor, which is closer by sequence identity than that between human β2 and human β1. This close similarity implies that a homology approach for closely related receptors can be successful.

Along with relatively high sequence identity, the adrenergic receptors share similar TM bundle arrangements. Table 4.2 shows the differences between available crystal structures relative to β2 with respect to η, θ, and ϕ angles. Although there are large differences, the most consistently similar structures are the ones with the highest sequence identity, such as β1 and β2 with 52% TM sequence identity with respect to β2. The next most similar structures are the two rhodopsin examples, squid and bovine, with 27% identity. The adenosine receptor is less related to either of these two subtypes, with 25% identity with respect to β2 and

Table 4.2: Rotation comparison of available GPCR crystal structures relative to human adrenergic β2 (PDB: 2RH1), all expressed in degrees relative to β2. Structures are turkey β1 (PDB: 2VT4), human adenosine A2a (PDB: 3EML), bovine rhodopsin (PDB: 1GZM), and squid opsin (PDB: 2Z73).

	η				θ				ϕ			
	β1	A2a	rhod.	opsin	β1	A2a	rhod.	opsin	β1	A2a	rhod.	opsin
tm1	-3.85	-7.44	-17.19	-11.92	-6.27	-6.28	-2.53	-5.87	-4.18	-24.00	-35.92	-30.12
tm2	4.97	15.80	-9.57	-16.64	-6.50	-8.86	-2.00	-0.89	-8.91	-8.85	-11.36	-6.56
tm3	-0.56	10.01	-11.2	16.35	-4.24	-6.58	-2.34	-1.85	4.17	-4.26	-23.97	-10.44
tm4	-4.32	6.18	1.72	17.98	2.18	4.86	-0.61	-0.54	-35.17	-7.45	-9.18	-3.00
tm5	-1.73	-3.32	-43.45	-37.16	0.85	0.37	-0.45	0.51	12.06	-13.01	-22.54	2.98
tm6	2.64	-3.98	-9.01	21.44	4.35	-9.19	-1.04	-3.81	-4.94	8.91	-1.07	-44.33
tm7	-4.14	-5.65	-7.02	-5.29	6.61	-2.60	6.65	2.96	-12.12	-10.22	-0.51	9.67

18% identity with respect to bovine rhodopsin. Even though the adenosine and β2 receptors are from the same species, β1 (turkey) is more closely related, in both amino acid sequence and helix orientation. The similarity in helix orientation is even preserved despite different crystallization strategies.

This similarity indicates that homology modeling is a viable option for determination of all nine adrenergic GPCR structures. A simple procedure for homology modeling can only provide a snapshot of the receptor, however, and as sequence identities decrease through the family the possibility of that snapshot being inaccurate increases. The methods developed for refinement of *ab initio* structure prediction can not only provide insight into other possible conformations of a known crystal structure and very closely related receptors, it can also predict the native conformations of more distantly related structures. The rotation sampling methods also lend insight into which helices of each template structure are rigid or flexible; which helices are locked into a conformation with strong conserved interactions, and which are allowed to rotate or shift within the structure.

In general, structures with higher identity between the target and template GPCRs are expected to have better results with rapid, coarse homology modeling. The closely related β receptors should be modeled the most easily, while the more distantly related α2 receptors may require more refinement. In this work, both the β1 turkey and β2 human structures are tested as templates for the eight human adrenergic receptors that lack X-ray crystal structures, and the resulting models are validated with a novel docking protocol. Those structures with the highest target-template identity enjoyed the most success, while those less related will require further refinement before full validation.

4.2 Methods

General Methods: All calculations were carried out using the DREIDING force field[92] with CHARMM22[93] charges. Side chain placement was determined with SCREAM[90] Unless otherwise noted, all simulations were performed in the gas phase with a dielectric of 2.5.

4.2.1 Building the Homology Models

Each receptor was built based on both the β2 (2RH1) and β1 (2VT4) crystal structures, referred to as template structures. The β2 human structure was not built, as it was used as a validation case for the homology method previously in the Goddard group. Throughout the descriptions, structures are labeled according to their name and their template, i.e., "α1a-β1" is the human α1a homology model built

on the turkey β1 template.

Prediction of transmembrane regions: Preliminary TM helices were obtained via MAFFT[95–97] alignment to the template receptor. These alignments were compared to the predicted helical regions from secondary structure prediction methods Porter,[124] APSSP2,[119] and PSIPred.[125] The final helix used for the model included the shorter helix between the secondary structure predictions and the crystal alignment; in the cases where no method predicted a helix but the crystal helix continued, the crystal helix was truncated for the model. The chosen helix for β1 human, as an example, is shown in Table 4.3. As a result, the TM lengths for a given receptor may be different for different templates.

Simplified Helix Optimization: After creating the TM helices, we minimized the helices in vacuum. This is a different procedure from the OptHelix method discussed in Section 2.2. OptHelix, while offering a useful starting point for structure predictions that rely on no other structural data other than a template for helix orientations, does not reproduce crystallographic helices with enough accuracy to determine crystallographic η rotations. For a homology model, helix shapes provided from the crystal data are a suitable starting point for minimal receptor-specific optimization. Each receptor's TM bundle was optimized by isolating each helix and minimizing for 100 steps.

Table 4.3: To determine the termini for homology helices, we considered results from secondary structure prediction methods Porter, APSSP2, and PsiPred. Starting with the crystal helix from the chosen template, homology helices were truncated if all three secondary structure predictions indicated a position would not be an αhelix. TM1 for β1 human and α1a are shown here as an example. The template crystal helix is indicated in red, and the final homology TM based on the combination of crystal alignment and secondary structure prediction is indicated in red and bold.

β1 xtal	GAELLSQQWEAGMSLLMALVVLLIVAGNVLVIAAIGSTQRL
β1 human	SPEPLSQQWTAGMGLLMALIVLLIVAGNVLVIVAIAKTPRL
Porter	ccccccHHHHHHHHHHHHHHHHHHHHHHHHHHHHcccc
APSSP2	cccccHHHHHHHHHHHHHHHHHHHHHHHHHHHHcccc
PsiPred	ccccccHHHHHHHHHHHHHHHHHHHHHHHEEEEEEEcccc
β1 homology	SPEPLSQQ**WTAGMGLLMALIVLLIVAGNVLVIVAIAK**TPRL
α1a	PPAPVNISKAILLGVILGGLILFGVLGNILVILSVACHRHL
Porter	ccccHHHHHHHHHHHHHHHHHHHHHHHHHHHHcccccc
APSSP2	cccccHHHHHHHHHHHHHHHHHHHHHHHHHHHHcccccc
PsiPred	cccccHHHHHHHHHHHHHHHHHHHHHHHHEEEEEEEcccc
α1a homology	PPAPVNIS**KAILLGVILGGLILFGVLGNILVILSVA**CHRHL

Helix Rotation Optimization: The BiHelix and CombiHelix methods determine the low-energy helix η rotations, and are described in detail in Section 2.2. For each receptor, the possible bundles are ranked by minimized energy.

4.2.2 Validation With Docking

A selection of structures was validated using docking and mutation studies. Each receptor validation included a group of ligands with consistent experimental data, with binding affinities sufficiently different (preferably 50-fold or more) to reliably separate using force field energies.

Ligands were prepared by building in Maestro, conformation search with Macro-

Model,[126] and mulliken charge calculation with Jaguar (B3LYP with 6-31G** basis set).[127] For antagonists with multiple nitrogens, pKa values were calculated with the pKa module in Jaguar.

Each receptor was validated using both β1 and β2 models, docked individually with canonical subtype selective antagonists. The binding sites were predicted using ScanBindSite.pl.[80] The receptors were modified to replace bulky nonpolar residues (tryptophan, tyrosine, phenylalanine, valine, isoleucine, and leucine) with alanine, allowing polar interactions to dominate the binding site selection and initial ligand enrichment steps.

For all models, the binding sites were chosen using the endogenous agonist epinephrine. In the adrenergic receptors, the conserved $Asp^{3.32}$ serves as the key anchor point for both agonists and antagonists bearing a protonated nitrogen. As many antagonists function by blocking the agonist binding site as well as stabilizing the receptor's inactive conformation, using the endogenous agonist for binding site determination is both plausible from a theoretical standpoint and practical to execute using current methods. ScanBindSite.pl uses a coarse implementation of HierDock[70] to place a ligand in a variety of small box regions over the entire empty space of the receptor. These placements are ranked by energy and ligand buried surface, and the regions corresponding to the best average energy at 80% buried surface were selected for further analysis.

Once the binding site was selected, two diverse ligand conformations were docked to the site using the recently developed DarwinDock, a Monte Carlo dock-

ing procedure. The protein was converted back to the wild type with SCREAM, then the ligands and charged residues were neutralized. This neutral complex was minimized twice, first the binding site, then the entire complex. The final docked poses were selected based on the energy from this final minimization.

In order to probe the binding sites of each receptor, subtype selective anatagonists were chosen and docked for each subtype. For the $\alpha 1$ subtype we chose WB4101 and prazosin. WB4101 has been extensively studied[128,129] as an $\alpha 1$ selective antagonist and a particularly potent α blocker, while prasosin was one of the first commercially developed α blocker. Yohimbine is a traditional $\alpha 2$ antagonist,[130] and its rigid structure allowed for excellent docking with less sampling than the other antagonists with more rotational degrees of freedom. In addition, the endogenous agonists epinephrine and norepinephrine were docked to the $\alpha 2$ receptors to compare with results from ligand binding studies.[131,132] For $\beta 1$ and $\beta 3$, the $\beta 1$ selective antagonist $(-)$ RO 363 was docked. The residues implicated in binding[133] are conserved between $\beta 1$ and $\beta 3$, and a high quality docking pose can illuminate the residues that give rise to this ligand's selectivity.

4.3 Results and Discussion

While the GenSemble method was developed using 30° increments for sampling, later studies in the Goddard group showed that BiHelix will resolve differences in helix η rotation as fine as 10°. This analysis exponentially increases the time necessary for full 360° sampling, so the 30° scan remains the method of choice

Table 4.4: The best TM bundles by minimized energy (kcal/mol) are shown for each template-receptor pair. Bundles are represented as combinations of η rotations in degrees relative to the template crystal structure. TMs 1, 2, and 3 are largely static, while TM4 shows some flexibility and a preference for a 15°counterclockwise rotation from the crystallographic orientation.

	β2 Template								β1 Template							
	H1	H2	H3	H4	H5	H6	H7	Energy	H1	H2	H3	H4	H5	H6	H7	Energy
β1	0	0	0	345	0	0	0	-178.9	0	0	0	345	0	0	0	-202.1
β3	0	0	0	0	0	0	0	-183.1	0	0	0	345	0	0	0	-186.3
α1a	0	0	0	345	0	15	0	254.2	0	0	0	345	0	0	0	209.8
α1b	0	0	0	345	0	0	345	246.5	0	0	0	345	0	60	315	193.8
α1d	0	0	0	105	270	0	90	63.2	0	0	0	90	0	75	0	138.0
α2a	0	0	0	345	0	0	0	153.8	0	0	0	345	0	0	0	117.3
α2b	180	15	0	345	0	0	105	193.1	0	0	0	345	270	105	0	119.0
α2c	0	0	0	345	0	0	0	263.5	0	0	0	0	0	45	0	220.8

when scanning the entire range of possible rotations. Once a bundle has been chosen, finer sampling in 15° increments within a 120° range is performed. The final structures are chosen from this analysis, also based on minimized energy, and the results are shown in Table 4.4. Most receptors in this study showed η rotations analagous to the template crystal structure as the lowest in minimized energy after building the full bundles. Those least related to the template strutures, α2b, α2c, and α1d, had alternate rotations for TMs 4, 5, 6, and 7 in varying combinations depending on the receptor and template. These alternatives were the starting point for the finer rotational analysis, and the final structures are reported in Table 4.4 relative to the initial crystal-derived structure.

The final structures after the fine rotational analysis show a clear preference for the 15° counterclockwise rotation of TM4. Conserved residues Arg$^{4.41}$ and Arg$^{4.40}$ are vertically positioned to interact with Asp$^{3.49}$ and Asn$^{2.40}$, respectively, if the η rotation allows it. In the 0° position, both residues on TM4 may interact with their

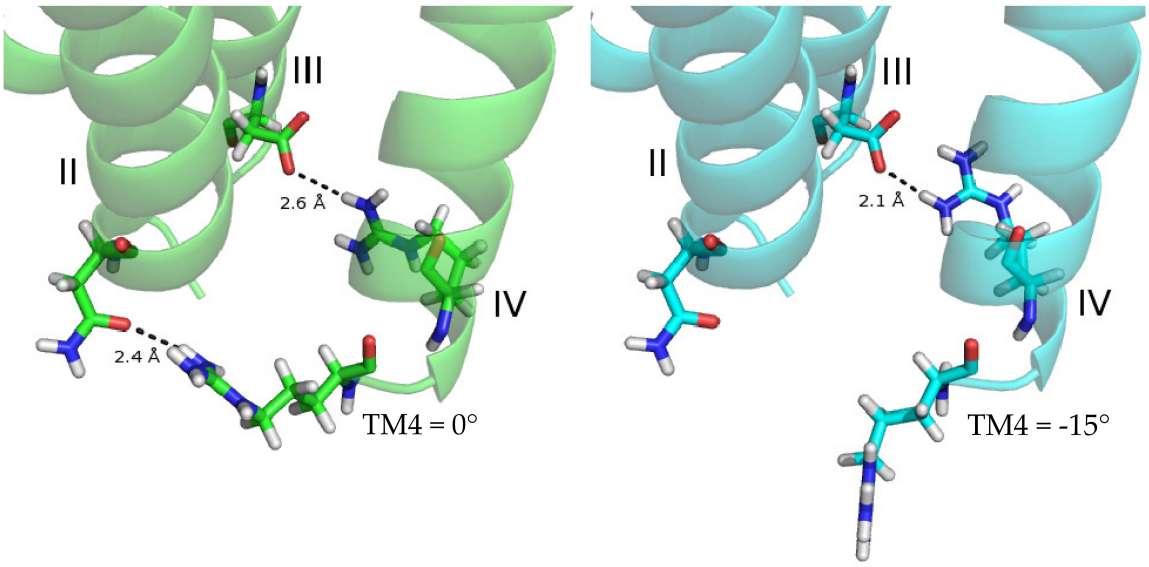

Figure 4.1: BiHelix/CombiHelix results indicate a counterclockwise rotation of TM4 is favored over the crystallographic orientation for most receptor-template pairs. The interactions responsible for this preference are illustrated below for the β1-β1 case.

partners in TMS 2 and 3, but only loosely. After the counterclockwise rotation the interaction between TMs 3 and 4 is strengthened, the distance between the two residues decreasing to 2.1 Å from the crystal orientation. This change is illustrated in Figure 4.1. Optimizing this interaction has a greater effect on the overall protein structure than compromising between the two, and this is reflected in the final helix orientations for these models. This change is difficult to validate, as ligand binding data for the adrenergic receptors confirms roles for all helices except TM4. The discussion of helix motion in Chapter 3 indicates that TM4 may simply be more dynamic than the other TM helices, and less important in ligand binding.

These charged residues in the helix termini may interact with loops or lipids rather than other helices in the native protein. In Chapter 2, a modified BiHelix method removed charged residues at the ends to avoid spurious interaction en-

ergies. That case, however, involved helix shapes determined by OptHelix and alignment to an average β2 crystal structure. In this homology model, the template choice and helix shapes are matched, and expected to interact more like the native protein. In previous studies, BiHelix/CombiHelix performed on the β1 crystal structure with truncated helices, the energy difference between the crystallographic rotations and the next most stable structure was significantly smaller than with full helices. This indicates that the intrahelical interactions in the helix ends are important for TM bundle stabilization, and care should be taken when deciding to ignore them. In the "blind" prediction case, it was appropriate to do so; in this homology model case, it is not.

In *ab initio* structure prediction, total bundle energy governs the choice of helix alignment template. As described in Chapter 2, several plausible templates are built, then after an ensemble of TM bundles are built a final template decision is made based on which template yields the lowest energy. In these homology models, the final bundle energies reflect a good match between the β1 template and the β1 human and β3 structures. With TM sequence identities of 85% and 63%, respectively, this indicates that sequence identity predicts the quality of a proposed homology model. The next best structures, β1-β2 and β3-β2, have TM identities of 68% and 59% and overall identities of 41% and 42%. The overall sequence identity implies the β3-β2 structure should be slightly better than the β1-β2 structure, and this is seen in the relative energies. As the target sequence deviates from the template sequence, and as the best rotations deviate from the initial crystallographic

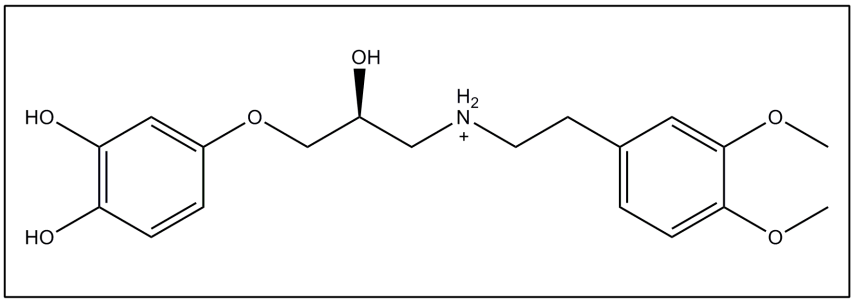

Figure 4.2: $(-)$RO-363 is a selective $\beta 1$ antagonist. While residues implicated in $\beta 1$ binding are conserved in the β receptors, the predicted binding site should indicate which residues are responsible for this ligand's selectivity.

rotations, the energy of the best bundle increases, indicating a less favorable match between template and target.

4.3.1 Validation with Docking and Mutation Studies

For swift validation of the adrenergic binding sites, we chose both agonists and antagonists based on studies with mutation or SAR data available for subtype selective ligands. The template crystal structures were crystallized with inverse agonists, so the resulting structures are more likely to resemble the inactive forms of the target receptors. The canonical "blocker" ligands should bind well to these inactive forms. However, as mentioned in Chapter 2, use of a new docking protocol meant to predict binding sites and poses without a knowledge-based judgement call introduces uncertainty into the validation process. Many structures built for this study could be validated with antagonist docking, but for those that could not be, further efforts should concentrate first on refinement of the docking procedure before revisiting the structure prediction.

$(-)$ RO-363 (Figure 4.2) is a $\beta 1$ selective antagonist with recent, detailed mutation data from studies carried out by Sugimoto, *et al.*[133] Half of the molecule resembles epinephrine, while the other half contains O-methyl groups that mimic some $\alpha 1$ antagonists. The sites studied for human $\beta 1$ antagonist binding are found at the top of TMs 2 and 7, and single, double, and triple mutations are considered. Alone, the mutation of Thr117$^{2.63}$ or Phe359$^{7.35}$ to alanine only have a small effect on binding, but combined they decrease antagonist affinity 25-fold. Leu110$^{2.56}$ mutated to alanine results in a seven-fold decrease in affinity, but that change is augmented when paired with the Phe359$^{7.35}$ mutation. The triple mutation also produces a 25-fold decrease in affinity. Although RO-363 is a $\beta 1$ selective antagonist, the residues studied are conserved in $\beta 3$. A good binding site for both receptors can indicate what part of the binding site is important for this subtype selectivity.

The BiHelix / CombiHelix for both $\beta 1$ human and $\beta 3$ in both templates resulted in TM bundles similar to the crystal templates. The only variation occurred in TM4, and for $\beta 3$-$\beta 2$ all helices were found in the crystal orientations. Neither template-target pair featured TM2 rotated such that the residues tested by Sugimoto, *et al.* were accessible to the binding site, and both binding pockets were too deep in the receptor to interact directly with Phe$^{7.35}$. One pose, however, positioned one side of the ligand where it might interact with Leu110$^{2.56}$ with an alternate rotation of TM2, and appropriate rotations do appear in the low-energy TM bundles from CombiHelix. Both poses featured strong interactions in the binding pocket with canonical adrenergic binding site residues, as shown in Figures 4.3 and 4.4.

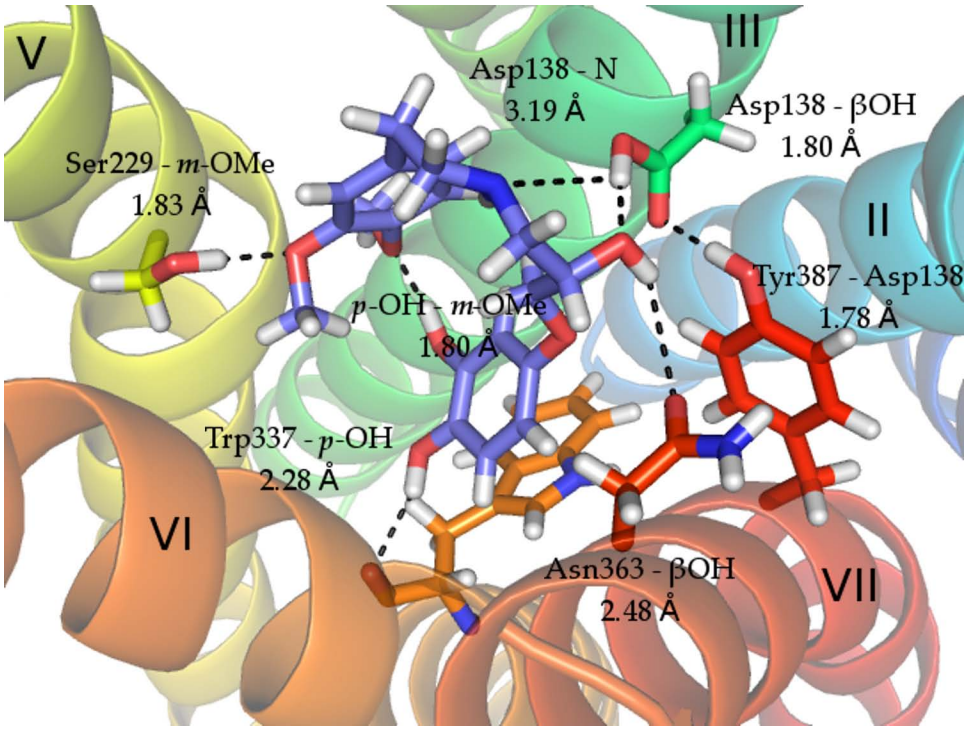

Figure 4.3: The best docked conformation of $(-)$ RO-363 (in blue) to the β1-β2 homology model was a folded conformation, with stabilizing interactions from the adrenergic agonist pharmacophore as well as an internal hydrogen bond.

For both cases, the β2 template provided the best docked poses, possibly due to all seven TM helices found in their crystallographic orientations. Although these structures do not directly support some of the mutation data, the strong polar interactions with canonical adrenergic binding site residues such as the TM5 serines and TM3 aspartic acid imply these poses and structures are plausible forms of the β1 and β3 human receptors.

Two sets of studies explored antagonist binding to α1 receptors: one focusing primarily on WB4101 (Figure 4.5) and prazosin (Figure 4.6) binding to α1a,[129] the other performing extensive SAR on WB4101 and comparing the differences among α1a, α1b, and a serotonin receptor.[128] Waugh, *et al.* only studied α1a, but the

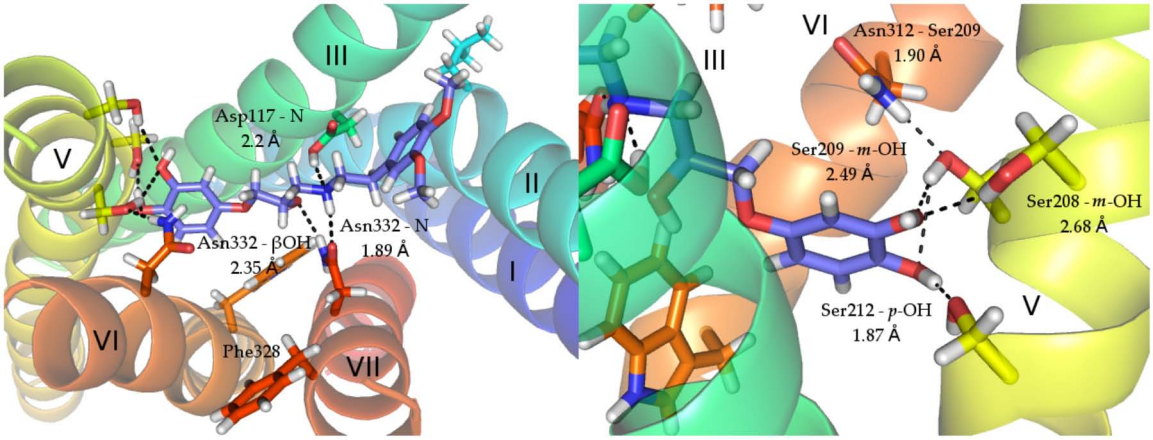

Figure 4.4: The $\beta 3$-$\beta 2$ model binds a linear conformation of $(-)$ RO-363, spanning the entire TM core from the antagonist binding pocket near TMs 2 and 7 to the agonist binding residues on TM5. The TM5 hydrogen bonding network is shown separately on the right.

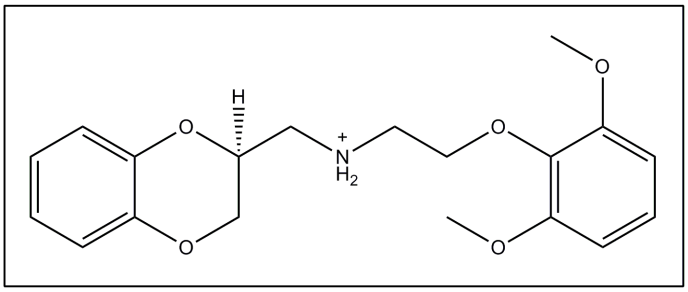

Figure 4.5: WB4101 has been studied extensively as an $\alpha 1$ selective antagonist.

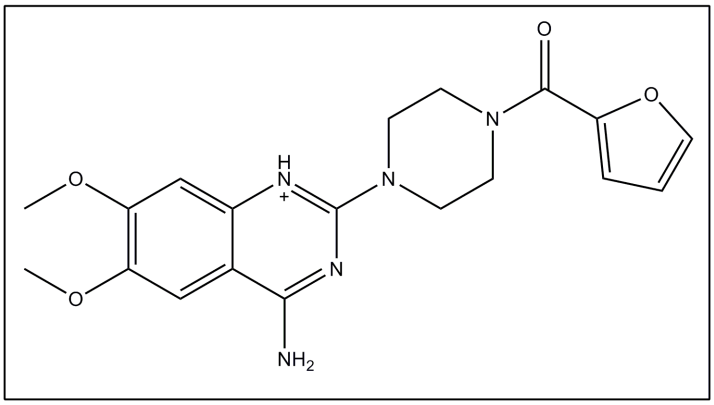

Figure 4.6: Prazosin is a selective α1 antagonist, and is prescribed as an antihypertensive drug.

residues mutated are conserved throughout the α1 subtype: Phe$^{7.39}$ and Phe$^{7.35}$, both at the top of TM7. WB4101 is α1 selective, but binds more strongly to α1a than to α1b or α1d. A good binding site for this ligand, then, can indicate which nonconserved residues are important for subtype selectivity. This study found that the upper phenylalanine, Phe$^{7.35}$, is more important for prazosin binding than for WB4104, and that the opposite holds for Phe$^{7.39}$. The predicted binding sites should reflect this difference.

The low-energy rotations for α1a-β1 are the most similar to the template rotations of all the α1 homology results, with only the consistent TM4 anticlockwise rotation deviating from the strictly analagous structure. This structure binds prazosin well, and while the key Phe308$^{7.35}$ residue does appear in the cavity analysis (Table 4.5) as a stabilizing residue, Phe312$^{7.39}$ is the most important residue in the best binding pose. By contrast, the WB4101 binding site does show several residues interacting with the ligand: Asp106$^{3.32}$ interacts with the protonated amine (1.88 Å); Tyr184$^{5.36}$ binds to the O-methyl and ether groups simultaneously

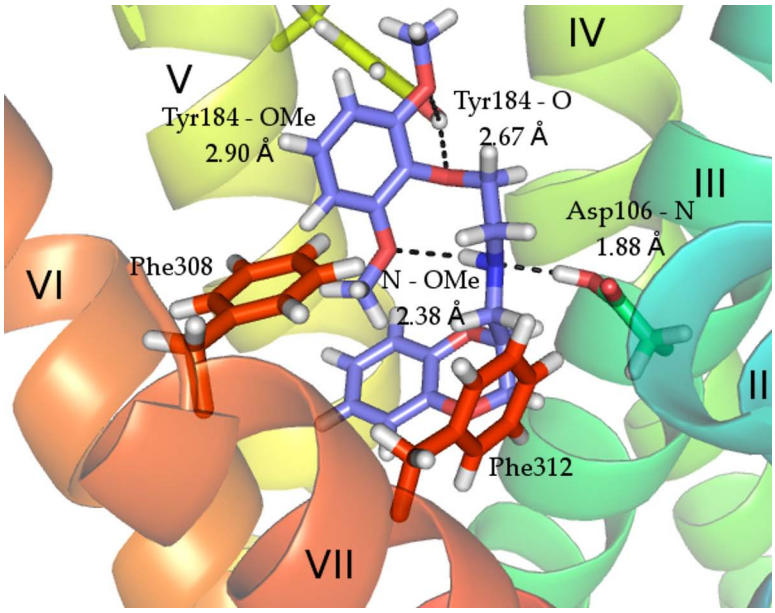

Figure 4.7: The α1a-β1 homology model docked WB4101 in a curled conformation stabilized primarily through interactions with Asp106$^{3.32}$, TM5, and TM7 phenylalanine residues implicated in antagonist binding through mutation studies.

(2.90 and 2.67 Å); and an internal hydrogen bond between the amine and the remaining O-methyl group stabilizes another polar interaction (2.38 Å). These interactions are illustrated in Figure 4.7. Both Phe308$^{7.35}$ and Phe312$^{7.39}$ are present in the binding site and contribute to ligand binding, but as with prazosin the relative contributions are the reverse of what is expected. α1a-β2, which favors rotation of TM6 15° clockwise in addition to the conserved TM4 rotation, could not be docked with a plausible pose. Despite the higher sequence identity for α1a to the β templates, this structure requires further refinement before the binding site can be verified.

α1b favors a slight anticlockwise rotation of TM7, the conserved rotation of TM4, and in the β1 template, an additional clockwise rotation of TM6. This α1b-β1 structure shows the greatest deviation from the crystallographic orientations

Table 4.5: While most of the key residues implicated by mutation studies appear in the cavity analysis for α1a-β1 docked with WB-4101 and prazosin, they are not in the order of priority suggested by the experiments. Further refinement of these docked structures may yield more accurate results. Energies are reported in kcal/mol.

WB-4101				**Prazosin**			
Residue	VdW	Coulomb	Total	Residue	VdW	Coulomb	Total
Tyr184	-2.357	-1.109	-3.465	Phe312	-7.386	-1.411	-8.797
Met292	-2.039	0.025	-2.014	Asp106	-2.373	-1.299	-3.672
Phe308	-1.582	-0.062	-1.644	Trp285	-2.952	-0.437	-3.390
Phe288	-1.152	-0.178	-1.330	Lys309	-1.820	-0.836	-2.656
Ala103	-1.187	0.020	-1.167	Trp102	-2.053	-0.431	-2.484
Tyr316	-0.688	-0.290	-0.978	Cys110	-2.672	0.254	-2.418
Trp102	-0.597	-0.363	-0.960	Trp313	-0.837	-0.997	-1.834
Val185	-0.994	0.103	-0.890	Phe308	-1.049	-0.180	-1.229
Phe289	-1.101	0.266	-0.835	Val107	-0.908	-0.008	-0.916
Leu162	-0.852	0.027	-0.825	Tyr316	-1.646	0.745	-0.901
Phe193	-0.805	0.248	-0.557	Leu75	-1.076	0.195	-0.881
Trp285	0.234	-0.684	-0.450	Ala103	-0.433	-0.332	-0.765
Cys110	-1.315	0.868	-0.447	Ser83	-0.385	-0.232	-0.617
Glu181	-0.153	-0.213	-0.367	Phe281	-0.152	-0.386	-0.538
Thr111	-0.255	-0.054	-0.309	Ser319	-0.160	-0.341	-0.500
Ala189	-0.560	0.269	-0.291	Met292	-0.595	0.122	-0.473
Ser188	-0.358	0.151	-0.207	Phe193	-0.404	0.041	-0.363
Val79	-0.138	0.088	-0.050	Phe289	-0.595	0.290	-0.304
Ser192	-0.331	0.376	0.046	Val79	-0.794	0.513	-0.282
Phe312	-0.142	0.224	0.082	Leu80	-0.173	0.063	-0.110

of the α1 receptors modeled here. Despite some docked structures that involved loose interactions with residues implicated by homology in antagonist binding, none of the favored docked structures of either prazosin or WB4101 supported the mutation data for antagonist binding to the related α1a receptor. These structures require more specialized refinement before they can be fully validated.

With comparable sequence identity to the other α1 receptors, α1d favors rotations similarly divergent from the template rotations. Both templates showed TM4 rotations, but not the same magnitude as the other structures: α1d-β2 prefers a 105° clockwise rotation of TM4, and α1d-β1 favors a 90° clockwise rotation. This may be due to a nonconserved glutamic acid at the intracellular end of TM4; the effect of the charged residue at the ends of TMs can be dramatic, and this may have contributed to the final η residue determination. This may be tested by mutating these charged residues in the TM caps to alanine and repeating the BiHelix analysis. The distinction between alanization in this case but not for the TM4 interactions described above is that this case concerns a non-conserved residue. A highly conserved residue is likely to mediate a significant interaction, while it is more possible in this case that the residue is creating noise.

Even with this shift in helix η rotations, it was possible to obtain plausible docked structures for WB4101 and prazosin for the structure built on the β2 template. The WB4101 binding site is shifted towards the TM1-2-7 pocket, with Ser153$^{2.61}$ forming a hydrogen bond with one of the dioxane oxygens, the larger phenyl group with O-methyl substituents buried deep in the binding pocket, and Phe384$^{7.35}$

Table 4.6: The top ten residues in the cavity analysis for α1d-β2 include the key residues implicated in mutation studies, Phe384$^{7.35}$ and Phe388$^{7.39}$ in the correct order of importance for both WB-4101 and prazosin. Energies are reported in kcal/mol.

WB-4101				**Prazosin**			
Residue	VdW	Coulomb	Total	Residue	VdW	Coulomb	Total
Phe388	-5.125	-0.983	-6.108	Glu157	4.205	-4.366	-5.979
Val149	-2.776	-0.154	-2.931	Ser153	-2.241	-1.166	-3.407
Phe384	-1.855	-0.715	-2.570	Asp176	-1.749	-1.231	-2.981
Leu390	-2.337	0.322	-2.016	Phe384	-2.800	-0.040	-2.840
Met156	-1.810	0.237	-1.573	Phe388	-2.281	-0.547	-2.828
Asp176	0.241	-1.735	-1.494	Leu150	-2.355	-0.405	-2.761
Glu157	-1.093	-0.142	-1.236	Met156	-3.121	0.637	-2.484
Tyr392	-0.749	-0.133	-0.882	Trp361	-1.895	-0.210	-2.105
Ala173	-0.350	-0.518	-0.868	Val149	-1.895	0.161	-1.734
Trp172	-0.349	-0.486	-0.835	Gly391	-2.264	0.814	-1.450

and Phe388$^{7.39}$ both creating π stacking interactions with the other side of the ligand. The key residue for WB4101 (Phe388$^{7.39}$) lies 3 Å away from the ligand, and the less important residue is 3.5 Å away. The cavity analysis for this pose shows a more than twofold contribution to the binding energy for Phe388$^{7.39}$. This structure and binding pose do agree with these mutation data, according to the cavity analysis (Table 4.6). The prazosin docked structure shows a stronger interaction between the ligand and Phe384$^{7.34}$ than with Phe388$^{7.39}$, also verifying the integrity of this binding site. That these structures were obtained for the β2 template structure, for which there is slightly higher sequence identity, rather than for the β1 template supports the assertion that a template-target pair with higher sequence identity is more likely to produce a high quality structure.

Yohimbine (Figure 4.8) is a potent α2 inverse agonist, often used as a benchmark ligand against which other ligand binding affinities are measured. Wang,

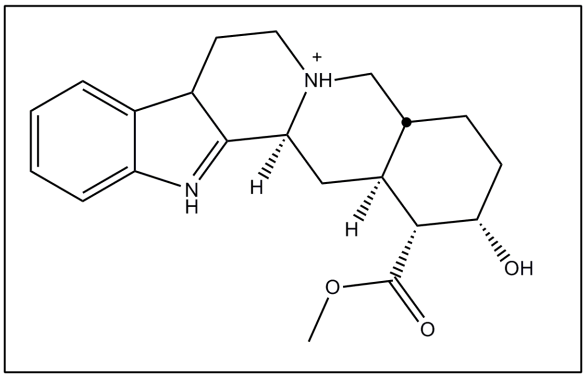

Figure 4.8: Yohimbine, originally discovered for its aphrodesiac properties, is a potent and selective α2 adrenergic inverse agonist.

et al. recently characterized the binding affinity of several agonists and antagonists with α2a and mutant receptors containing changes to TMs 2, 3, and 5.[131] The most striking effect came from the mutation of Asp113$^{3.32}$ to asparagine, which completely eliminated yohimbine binding but allowed a small amount of agonist mediated activation. The next most important residues are in TM5, Ser200$^{5.42}$ and Ser204$^{5.46}$. These had a greater effect on agonist binding, where they are expected to form crucial interactions with the catechol hydroxides, but mutation to Alanine resulted in a three- to four-fold decrease in yohimbine affinity. The least important residues, Asp130$^{3.49}$ and Asp79$^{2.50}$, are deeply buried in the TM core. The highly conserved Asp$^{3.49}$ is expected to be involved in the TM3-TM6 ionic lock controlling activation, and is too far away from other important residues for direct ligand interaction. Similarly, Asp$^{2.50}$ is positioned to interact with Asn$^{1.50}$ and Asn$^{7.49}$ to form the stabilizing polar network seen in the bovine rhodopsin crystal structure. An additional study shows a dramatic, 300-fold decrease in yohimbine binding upon mutation of Phe412$^{7.39}$, implying the upper section of TM7 is more likely to

directly interact with the antagonist.[132]

Both the β1 and β2 templates yielded identical helix rotations for the α2a structure, so the differences in binding can be attributed to changes in the helix orientations due to the template. The best α2a-β1 docked structure by local cavity interaction (Figure 4.9) showed close interactions with Asp113$^{3.32}$ (1.88 Å to the protonated amine) and Ser204$^{5.46}$ (1.86 Å to the hydroxide). Ser200$^{5.42}$ was not observed interacting with the ligand, but the residue is positioned such that it may interact with the ester group in a dynamic environment. Most importantly, Phe412$^{7.39}$ forms a π stacking interaction with the indole rings and appears in the cavity analysis with a favorable -4.206 kcal/mol stabilization energy (Table 4.7). This docking pose, even before further refinement with molecular dynamics, agrees well with the published mutation data. This agreement is better than the corresponding best docked α2a-β2 structure, which is consistent with the increased sequence identity between target and template: α2a and β1 turkey share 46% TM sequence identity, while α2a and β2 only share 39%.

This structure also shows Tyr394$^{6.55}$ interacting directly with the indole nitrogen on the ligand (2.12 Å). This residue, analagous to Asn$^{6.55}$ in the β receptors shown to be important for stereoselectivity, is a good candidate for further mutation studies and may be important for adrenergic ligand selectivity.

The favored α2b rotations shift TMs 1, 2, 4, and 7 for the β2 template and TMs 4, 5, and 6 for the β1 template. While the α2b-β2 structure does form good protein-ligand interactions with Asp92$^{3.32}$ (1.81 Å to the protonated amine) and an addi-

Table 4.7: The key residue implicated by mutation studies, Phe412$^{7.39}$, appears in the cavity analysis for α2a-β1 docked with yohimbine. However, Asp113$^{3.32}$ with a favorable Coulomb energy of -2.423 kcal/mol does not show a favorable energy due to Van der Waals repulsion after cavity optimization places the protonated amine very close to that residue (1.88 Å). Energies are reported in kcal/mol.

Residue	VdW	Coulomb	Total
Tyr394	-1.381	-0.834	-4.591
Ser204	2.415	-2.415	-4.323
Phe412	-4.819	0.614	-4.206
Phe390	-3.122	0.153	-2.970
Leu110	-1.555	-0.456	-2.012
Cys201	-1.843	0.090	-1.753
Ser200	-1.371	-0.312	-1.683
Phe205	-1.853	0.183	-1.670
Lyn409	-1.408	-0.021	-1.429
Val197	-1.164	-0.113	-1.278

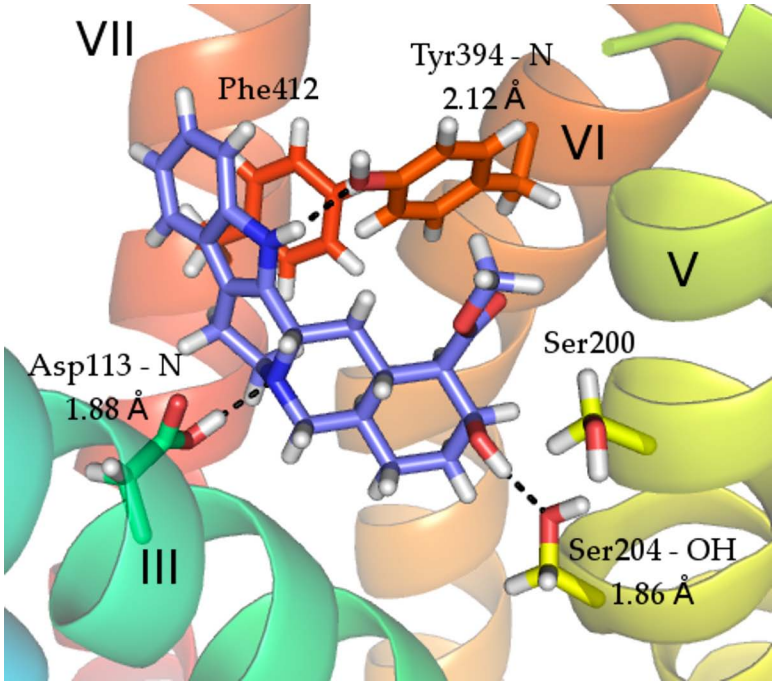

Figure 4.9: Yohimbine has less conformational flexibility than the antagonists docked for the α1 and β receptors, so only one primary conformation was important for docking. The docked ligand interacts with conserved residues on TMs 3 and 5, and in the α2a docked structure is positioned close to Phe412$^{7.39}$ implicated in mutation studies.

tional unpredicted interaction with Glu73$^{2.65}$ (1.75 Å to the hydroxide), Phe412$^{7.39}$ is rotated too far towards the lipid to allow the ligand to interact with the implicated TM5 serines. This may indicate that this conserved phenylalanine is not reponsible for a conserved interaction, but considering the lower sequence identity between α2b and the available crystal structures, it is more likely that these rotations are not correct for α2b. Both α2b-β1 and α2b-β2 have small rotations in TMs 5 and 7, respectively, that shift the important binding residues closer to the binding site, and it is possible that a broader docking strategy involving a diverse ensemble of low-energy TM bundles from the η rotation analysis may show a better binding site for this system.

The final rotations favored for α2c are closer to the template crystal structures than those for α1b: the recurring 15° anticlockwise rotation of TM4 appears for the α2c-β2 structure, and TM6 is rotated 45° clockwise for α2c-β1. The primary interaction in the binding cavity for the α2c-β1 structure involves Phe423$^{7.39}$, a cation-π interaction with the protonated amine, but this functional group on the ligand is expected to interact with the necessary Asp131$^{3.32}$ instead. While this aspartate does appear in the cavity analysis (Table 4.8), it contributes only a small amount to the overall binding energy. There is no polar interaction with Ser214$^{5.42}$ or Ser218$^{5.46}$, though the O-methyl groups on the ligand are accessible to alternative rotamers of both residues. This structure could be refined to reflect the mutation data more clearly, both with selective SCREAM rotamer scans and with annealing dynamics. In addition, the α2c-β2 structure did not support a plausible

Table 4.8: Both key residues involved in yohimbine binding, Phe423$^{7.39}$ and Asp131$^{3.32}$, appear in the cavity analysis for α2c-β1 docked to yohimbine, but key TM5 serine residues are conspicuously absent. Like many of the coarsely docked structures, this binding site should improve with further refinement with binding site annealing or selective sidechain optimization with SCREAM. Energies are reported in kcal/mol.

Residue	VdW	Coulomb	Total
Phe423	-5.370	-0.001	-5.371
Tyr402	-3.336	-0.198	-3.571
Val132	-2.965	0.353	-2.612
Phe219	-2.977	0.407	-2.570
Trp395	-2.036	-0.394	-2.431
Leu128	-1.174	-0.733	-1.907
Cys215	-1.476	-0.319	-1.795
Cys135	-2.396	0.670	-1.726
Tyr427	-2.192	0.561	-1.631
Ile211	-0.914	-0.611	-1.525
Asp131	-0.448	-0.920	-1.368

yohimbine docked pose, indicating that the overall orientation of the β2 template may be too different from the native α2c structure for a simple homology approach to model.

4.4 Conclusion

In building homology models for the human adrenergic receptors based on available crystal data, sequence identity between the target receptor and the template was a good but not perfect indicator of the model's ultimate success. As this work focused primarily on obtaining swift, coarse results for all eight uncrystallized human adrenergic receptors, many options remain for further structure optimization. The OptHelix method for helix optimization provides another set of bundles for each receptor-pair, and may result in more native-like helices. Replacement of

charged residues in the TM caps as in Chapter 2 may reveal different low-energy η rotations, especially for cases with lower sequence identity between target and template where the orientation of the TM bundle is less likely to resemble the native structure. The SuperBiHelix method, which adds θ and ϕ to the η rotational sampling of BiHelix, can fine-tune the template orientation of these low-identity cases as well.

The subtype selective antagonist docking suggest a series of mutation studies that may be performed both *in silico* and in the lab. In particular, the α1a antagonist interaction with the conserved Tyr184$^{5.36}$ points towards a possible important interaction for allosteric antagonists, allowing a ligand to engage TM5 without disrupting residues closer to the intracellular side where activation takes place.

For the structures with plausible helix rotations but unsatisfactory docking, further investigation into docking strategies will yield more plausible results. The newly developed DarwinDock features a panoply of variations for both pose prediction and final complex optimization that may be tuned to obtain a reasonable structure, and alternate methods for choosing a binding site may be employed. Recent studies of the adenosine A_{2A} receptor in the Goddard group as well as mutation studies for a variety of systems have shown that the extracellular loops play a role in ligand binding, so complete validation of these systems will include fully modeled loops.

Finally, with full predicted structures for all human adrenergic receptors, detailed activation studies may be performed with full-solvent molecular dynamics.

The comparison of systems with constitutive activity (like β2) with those that have less (like the α1 receptors) may illuminate important processes for activation as well as interactions that stabilize an inactive state. The structures predicted and optimized here are antagonist-stabilized inactive structures, so attention should be paid to interhelical interactions that anchor the TM helices into place, preventing activation or ligand replacement by a diffusible agonist.

From *ab initio* structure prediction to nanosecond timescale MD through to a complete set of 3D structures for the human adrenergic receptors, theory and experiment influence and benefit from one another. The full structure of the turkey β2 receptor was predicted using the amino acid sequence and validated with stabilizing mutation data from experiment. The methods used for that prediction can now be used with more confidence to predict the structures of systems with less experimental data, and in turn provide further leads for experiment. The crystal structure itself served as an experimental starting point for MD studies of the activating protein, and offered insight into helix movements that initiate activation as well as possible mutations to stabilize an active receptor. Finally, both the β2 and β1 crystal structures were built upon to create quick and useful models of the closely related adrenergic subfamily of receptors. While some of the models require refinement, they do give some idea of residue positioning, binding site similarities and differences among the different subtypes, and potential mutation studies. With increasing confidence in theoretical methods ensured by validation, theory can enable avenues of scientific exploration otherwise unavailable.

References

1. Fredriksson, R., and Schioth, H. B. (2005) "The repertoire of G-protein-coupled receptors in fully sequenced genomes." *Molecular Pharmacology 67*, 1414–1425.

2. Foord, S. M., Bonner, T. I., Neubig, R. R., Rosser, E. M., Pin, J. P., Davenport, A. P., Spedding, M., and Harmar, A. J. (2005) "International Union of Pharmacology. XLVI. G protein-coupled receptor list." *Pharmacological Reviews 57*, 279–288.

3. Schöneberg, T., Schulz, A., Biebermann, H., Hermsdorf, T., Römpler, H., and Sangkuhl, K. (2004) "Mutant G-protein-coupled receptors as a cause of human diseases." *Pharmacology and Therapeutics 104*, 173–206.

4. Josefsson, L. G. (1999) "Evidence for kinship between diverse G-protein coupled receptors." *Gene 239*, 333–340.

5. Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J. *Molecular Biology of the Cell*, 3rd ed.; Garland Publishing: New York, 1994.

6. Gutkind, J. S. (1998) "The Pathways Connecting G Protein-coupled Recep-

tors to the Nucleus through Divergent Mitogen-activated Protein Kinase Cascades." *Journal of Biological Chemistry 273*, 1839–1842.

7. Altenbach, C., Kusnetzow, A. K., Ernst, O. P., Hofmann, K. P., and Hubbell, W. L. (2008) "High-resolution distance mapping in rhodopsin reveals the pattern of helix movement due to activation." *Proceedings of the National Academy of Sciences of the United States of America 105*, 7439–7444.

8. Sutherland, E. W., and Rall, T. W. (1958) "Fractionisation and characterisation of a cyclic adenine ribonucletide formed by tissue peptides." *Journal of Biological Chemistry 232*, 1077–1091.

9. Urban, J. D., Clarke, W. P., von Zastrow, M., Nichols, D. E., Kobilka, B., Weinstein, H., Javitch, J. A., Roth, B. L., Christopoulos, A., Sexton, P. M., Miller, K. J., Spedding, M., and Mailman, R. B. (2007) "Functional Selectivity and Classical Concepts of Quantitative Pharmacology." *Journal of Pharmacology and Experimental Theraputics 320*, 1–13.

10. Mailman, R. B. (2007) "GPCR functional selectivity has therapeutic impact." *Trends in Pharmacological Sciences 28*, 390–396.

11. Siehler, S. (2008) "Cell-based assays in GPCR drug discovery." *Biotechnology Journal 3*, 471–483.

12. Evers, A., and Klabunde, T. (2005) "Structure-based Drug Discovery Using GPCR Homology Modeling: Successful Virtual Screening for Antagonists of the α1A Adrenergic Receptor." *Journal of Medicinal Chemistry 48*, 1088–1097.

13. Becker, O. M., Marantz, Y., Shacham, S., Inbal, B., Heifetz, A., Kalid, O., Bar-Haim, S., Warshaviak, D., Fichman, M., and Noiman, S. (2004) "G protein-coupled receptors: In silico drug discovery in 3D." *Proceedings of the National Academy of Sciences of the United States of America 101*, 11304–11309.

14. Hopkins, A. L., and Groom, C. R. (2002) "The druggable genome." *Nature Reviews: Drug Discovery 1*, 727–730.

15. Klabunde, T., and Hessler, G. (2002) "Drug Design Strategies for Targeting G-Protein-Coupled Receptors." *ChemBioChem 3*, 928–944.

16. Overington, J. P., Al-Lazikani, B., and Hopkins, A. L. (2006) "How many drug targets are there?." *Nature Reviews: Drug Discovery 5*, 993–996.

17. Unger, V. M., Hargrave, P. A., Baldwin, J. M., and Schertler, G. F. X. (1997) "Arrangement of rhodopsin transmembrane α-helices." *Nature 389*, 203–205.

18. Strader, C. D., Fong, T. M., Graziano, M. P., and Tota, M. R. (1995) "The family of G-protein-coupled receptors." *FASEB Journal 9*, 745–754.

19. Kobilka, B. (1992) "Adrenergic Receptors as Models for G Protein-Coupled Receptors." *Annual Review of Neuroscience 15*, 87–114.

20. Sealfon, S. C., Chi, L., Ebersole, B. J., Rodic, V., Zhang, D., Ballesteros, J. A., and Weinstein, H. (1995) "Related Contribution of Specific Helix 2 and 7 Residues to Conformational Activation of the Serotonin 5-HT Receptor." *Journal of Biological Chemistry 270*, 16683–16688.

21. Ceresa, B. P., and Limbird, L. E. (1994) "Mutation of an aspartate residue highly conserved among G-protein- coupled receptors results in nonreciprocal disruption of α2- adrenergic receptor-G-protein interactions. A negative charge at amino acid residue 79 forecasts α2A-adrenergic receptor sensitivity to allosteric modulation by monovalent cations and fully effective receptor/G-protein coupling." *Journal of Biological Chemistry 269*, 29557–29564.

22. Piascik, M. T., and Perez, D. M. (2001) "α1-Adrenergic Receptors: New Insights and Directions." *Journal of Pharmacology and Experimental Theraputics 298*, 403–410.

23. Shi, L., Liapakis, G., Xu, R., Guarnieri, F., Ballesteros, J. A., and Javitch, J. A. (2002) "β2 Adrenergic Receptor Activation: Modulation of the Proline Kink in Transmembrane 6 by a Rotamer Toggle Switch." *Journal of Biological Chemistry 277*, 40989–40996.

24. Vilardaga, J.-P. (2006) "Switching modes for G protein-coupled receptor activation." *Nature Chemical Biology 2*, 395–396.

25. Yao, X., Parnot, C., Deupi, X., Ratnala, V. R. P., Swaminath, G., Farrens, D., and Kobilka, B. (2006) "Coupling ligand structure to specific conformational switches in the α2-adrenoceptor." *Nature Chemical Biology 2*, 417–422.

26. Scheer, A., Fanelli, F., Costa, T., De Benedetti, P. G., and Cotecchia, S. (1996) "Constitutively active mutants of the α1B-adrenergic receptor: role of highly

conserved polar amino acids in receptor activation." *The EMBO Journal 15*, 3566.

27. Ballesteros, J. A., Weinstein, H., and Stuart, C. S. "Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors." In *"Methods in Neurosciences*; Academic Press, 1995; Vol. 25, pp 366–428.

28. Paleczewski, K., Kumasaka, T., Hori, T., Behnke, C. A., Motoshima, H., Fox, B. A., Le Trong, I., Teller, D. C., Okada, T., Stenkamp, R. E., Yamamoto, M., and Miyano, M. (2000) "Crystal Structure of Rhodopsin: a G Protein-Coupled Receptor." *Science 289*, 739–45.

29. Li, J., Edwards, P. C., Burghammer, M., Villa, C., and Schertler, G. F. X. (2004) "Structure of Bovine Rhodopsin in a Trigonal Crystal Form." *Journal of Molecular Biology 343*, 1409–1438.

30. Okada, T., Fujiyoshi, Y., Silow, M., Navarro, J., Landau, E. M., and Shichida, Y. (2002) "Functional role of internal water molecules in rhodopsin revealed by x-ray crystallography." *Proceedings of the National Academy of Sciences of the United States of America 99*, 5982–5987.

31. Okada, T., Sugihara, M., Bondar, A.-N., Elstner, M., Entel, P., and Buss, V. (2004) "The Retinal Conformation and its Environment in Rhodopsin in Light of a New 2.2Å Crystal Structure." *Journal of Molecular Biology 342*, 571–583.

32. Salom, D., Lodowski, D. T., Stenkamp, R. E., Trong, I. L., Golczak, M., Jastrzebska, B., Harris, T., Ballesteros, J. A., and Palczewski, K. (2006) "Crystal structure of a photoactivated deprotonated intermediate of rhodopsin." *Proceedings of the National Academy of Sciences of the United States of America 103*, 16123–16128.

33. Standfuss, J., Xie, G., Edwards, P. C., Burghammer, M., Oprian, D. D., and Schertler, G. F. X. (2007) "Crystal Structure of a Thermally Stable Rhodopsin Mutant." *Journal of Molecular Biology 372*, 1179–1188.

34. Nakamichi, H., Buss, V., and Okada, T. (2007) "Photoisomerization Mechanism of Rhodopsin and 9-cis-Rhodopsin Revealed by X-ray Crystallography." *Biophysical Journal 92*, L106–L108.

35. Park, J. H., Scheerer, P., Hofmann, K. P., Choe, H.-W., and Ernst, O. P. (2008) "Crystal structure of the ligand-free G-protein-coupled receptor opsin." *Nature 454*, 183–187.

36. Scheerer, P., Park, J. H., Hildebrand, P. W., Kim, Y. J., Krausz, N., Choe, H.-W., Hofmann, K. P., and Ernst, O. P. (2008) "Crystal structure of opsin in its G-protein-interacting conformation." *Nature 455*, 497–502.

37. Cherezov, V., Rosenbaum, D. M., Hanson, M. A., Rasmussen, S. G. F., Thian, F. S., Kobilka, T. S., Choi, H.-J., Kuhn, P., Weis, W. I., Kobilka, B. K., and Stevens, R. C. (2007) "High-Resolution Crystal Structure of an Engineered Human β2-Adrenergic G Protein Coupled Receptor." *Science 318*, 1258–1265.

38. Rosenbaum, D. M., Cherezov, V., Hanson, M. A., Rasmussen, S. G. F., Thian, F. S., Kobilka, T. S., Choi, H.-J., Yao, X.-J., Weis, W. I., Stevens, R. C., and Kobilka, B. K. (2007) "GPCR Engineering Yields High-Resolution Structural Insights into β2-Adrenergic Receptor Function." *Science 318*, 1266–1273.

39. Rasmussen, S. G. F., Choi, H.-J., Rosenbaum, D. M., Kobilka, T. S., Thian, F. S., Edwards, P. C., Burghammer, M., Ratnala, V. R. P., Sanishvili, R., Fischetti, R. F., Schertler, G. F. X., Weis, W. I., and Kobilka, B. K. (2007) "Crystal structure of the human β2 adrenergic G-protein-coupled receptor." *Nature 450*, 383–387.

40. Hanson, M. A., Cherezov, V., Griffith, M. T., Roth, C. B., Jaakola, V.-P., Chien, E. Y. T., Velasquez, J., Kuhn, P., and Stevens, R. C. (2008) "A Specific Cholesterol Binding Site Is Established by the 2.8 Å Structure of the Human β2-Adrenergic Receptor." *Structure 16*, 897–905.

41. Warne, T., Serrano-Vega, M. J., Baker, J. G., Moukhametzianov, R., Edwards, P. C., Henderson, R., Leslie, A. G. W., Tate, C. G., and Schertler, G. F. X. (2008) "Structure of a β1-adrenergic G-protein-coupled receptor." *Nature 454*, 486–491.

42. Jaakola, V.-P., Griffith, M. T., Hanson, M. A., Cherezov, V., Chien, E. Y. T., Lane, J. R., Ijzerman, A. P., and Stevens, R. C. (2008) "The 2.6 Å Crystal Structure of a Human A_{2A} Adenosine Receptor Bound to an Antagonist." *Science 322*, 1211–1217.

43. Martin, A., Damian, M., Laguerre, M., Parello, J., Pucci, B., Serre, L., Mary, S., Marie, J., and Baneres, J. L. (2009) "Engineering a G protein-coupled receptor for structural studies: Stabilization of the BLT1 receptor ground state." *Protein Science 18*, 727–734.

44. Leurs, R., Smit, M. J., Alewijnse, A. E., and Timmerman, H. (1998) "Agonist-independent regulation of constitutively active G-protein-coupled receptors." *Trends in Biochemical Sciences 23*, 418 – 422.

45. Samama, P., Cotecchia, S., Costa, T., and Lefkowitz, R. J. (1993) "A mutation-induced activated state of the β2-adrenergic receptor. Extending the ternary complex model." *Journal of Biological Chemistry 268*, 4625–4636.

46. Lefkowitz, R. J., Cotecchia, S., Samama, P., and Costa, T. (1993) "Constitutive activity of receptors coupled to guanine nucleotide regulatory proteins." *Trends in Pharmacological Sciences 14*, 303–307.

47. Costa, T., and Herz, A. (1989) "Antagonists with negative intrinsic activity at delta opioid receptors coupled to GTP-binding proteins." *Proceedings of the National Academy of Sciences of the United States of America 86*, 7321–7325.

48. Wade, S. M., Lan, K.-L., Moore, D. J., and Neubig, R. R. (2001) "Inverse Agonist Activity at the α2A-Adrenergic Receptor." *Molecular Pharmacology 59*, 532–542.

49. Vogel, R., Mahalingam, M., Lüdeke, S., Huber, T., Siebert, F., and Sakmar, T. P. (2008) "Functional Role of the "Ionic Lock."–An Interhelical Hydrogen-Bond

Network in Family A Heptahelical Receptors." *Journal of Molecular Biology 380*, 648–655.

50. Ballesteros, J. A., Jensen, A. D., Liapakis, G., Rasmussen, S. G. F., Shi, L., Gether, U., and Javitch, J. A. (2001) "Activation of the β2-Adrenergic Receptor Involves Disruption of an Ionic Lock between the Cytoplasmic Ends of Transmembrane Segments 3 and 6." *Journal of Biological Chemistry 276*, 29171–29177.

51. Lopez-Sendo, J. et al. (2004) "Expert consensus document on beta-adrenergic receptor blockers: The Task Force on β-Blockers of the European Society of Cardiology." *European Heart Journal 25*, 1341–1362.

52. Prichard, B. N. C., and Gillam, P. M. C. (1964) "Use of propranolol (Inderal) in treatment of hypertension." *British Medical Journal 2*, 725.

53. Freddolino, P. L., Kalani, M. Y. S., Vaidehi, N., Floriano, W. B., Hall, S. E., Trabanino, R. J., Kam, V. W. T., and Goddard, W. A. (2004) "Predicted 3D structure for the human β2 adrenergic receptor and its binding site for agonists and antagonists." *Proceedings of the National Academy of Sciences of the United States of America 101*, 2736–2741.

54. Ruffolo, R. R., Bondinell, W., and Hieble, J. P. (1995) "α- and β-Adrenoceptors: From the Gene to the Clinic. 2. Structure-Activity Relationships and Therapeutic Applications." *Journal of Medicinal Chemistry 38*, 3681–3716.

55. Liapakis, G., Ballesteros, J. A., Papachristou, S., Chan, W. C., Chen, X., and

Javitch, J. A. (2000) "The Forgotten Serine." *Journal of Biological Chemistry 275*, 37779–37788.

56. Strader, C. D., Candelore, M. R., Hill, W. S., Sigal, I. S., and Dixon, R. A. (1989) "Identification of two serine residues involved in agonist activation of the β-adrenergic receptor." *Journal of Biological Chemistry 264*, 13572–13578.

57. Gether, U., and Kobilka, B. K. (1998) "G Protein-coupled Receptors. II. Mechanism of Agonist Activation." *Journal of Biological Chemistry 273*, 17979–17982.

58. Hieble, J. P., Bylund, D. B., Clarke, D. E., Eikenburg, D. C., Langer, S. Z., Lefkowitz, R. J., Minneman, K. P., and Ruffolo, J., R. R. (1995) "International Union of Pharmacology. X. Recommendation for nomenclature of α1-adrenoceptors: consensus update." *Pharmacological Reviews 47*, 267–270.

59. Minneman, K. P. (1988) "α1-adrenergic receptor subtypes, inositol phosphates, and sources of cell Ca^{2+}." *Pharmacological Reviews 40*, 87–119.

60. Kable, J. W., Murrin, L. C., and Bylund, D. B. (2000) "In Vivo Gene Modification Elucidates Subtype-Specific Functions of α2-Adrenergic Receptors." *Journal of Pharmacology and Experimental Theraputics 293*, 1–7.

61. Philipp, M., Brede, M., and Hein, L. (2002) "Physiological significance of α2-adrenergic receptor subtype diversity: one receptor is not enough." *The American Journal of Physiology - Regulatory, Integrative and Comparative Physiology 283*, R287–295.

62. MacMillan, L. B., Hein, L., Smith, M. S., Piascik, M. T., and Limbird, L. E. (1996) "Central Hypotensive Effects of the α2a-Adrenergic Receptor Subtype." *Science 273*, 801–803.

63. Franowicz, J. S., Kessler, L. E., Borja, C. M. D., Kobilka, B. K., Limbird, L. E., and Arnsten, A. F. T. (2002) "Mutation of the α2A-adrenoceptor impairs working memory performance and annuls cognitive enhancement by guanfacine." *Journal of Neuroscience 22*, 8771.

64. Southwick, S. M., MorganIii, C. A., Charney, D. S., and High, J. R. (1999) "Yohimbine use in a natural setting: effects on post traumatic stress disorder." *Biological Psychiatry 46*, 442–444.

65. Chung, S. Y., and Subbiah, S. (1996) "A structural explanation for the twilight zone of protein sequence homology." *Structure 4*, 1123–1127.

66. Bissantz, C., Bernard, P., Hibert, M., and Rognan, D. (2003) "Protein-based virtual screening of chemical databases. II. Are homology models of g-protein coupled receptors suitable targets?." *Proteins: Structure, Function, and Genetics 50*, 5–25.

67. Patny, A., Desai, P. V., and Avery, M. A. (2006) "Homology Modeling of G-Protein-Coupled Receptors and Implications in Drug Design." *Current Medicinal Chemistry 13*, 1667–1691.

68. Shacham, S., Marantz, Y., Bar-Haim, S., Kalid, O., Warshaviak, D., Avisar, N., Inbal, B., Heifetz, A., Fichman, M., and Topf, M. (2004) "PREDICT modeling

and in-silico screening for G-protein coupled receptors." *Proteins: Structure, Function, and Bioinformatics 57*, 51–86.

69. Fanelli, F., Menziani, C., Scheer, A., Cotecchia, S., and De Benedetti, P. G. (1998) "*Ab Initio* Modeling and Molecular Dynamics Simulation of the α1b-Adrenergic Receptor Activation." *Methods 14*, 302–317.

70. Vaidehi, N., Floriano, W. B., Trabanino, R., Hall, S. E., Freddolino, P., Choi, E. J., Zamanakos, G., and Goddard, W. A. (2002) "Prediction of structure and function of G protein-coupled receptors." *Proceedings of the National Academy of Sciences of the United States of America 99*, 12622–12627.

71. Abrol, R., Goddard, W. A., Griffith, A. R., and Kam, V. W. T. "Methods for Predicting Three-dimensional Structures for αHelical Membrane Proteins and their Use in Design of Selective Ligands." Patent 20090037118, 2009.

72. Niv, M. Y., Skrabanek, L., Filizola, M., and Weinstein, H. (2006) "Modeling activated states of GPCRs: the rhodopsin template." *Journal of Computer-Aided Molecular Design 20*, 437–448.

73. Gouldson, P. R., Kidley, N. J., Bywater, R. P., Psaroudakis, G., Brooks, H. D., Diaz, C., Shire, D., and Reynolds, C. A. (2004) "Toward the active conformations of rhodopsin and the β2-adrenergic receptor." *Proteins: Structure, Function, and Bioinformatics 56*, 67–84.

74. Saam, J., Tajkhorshid, E., Hayashi, S., and Schulten, K. (2002) "Molecular

Dynamics Investigation of Primary Photoinduced Eventsin the Activation of Rhodopsin." *Biophysical Journal 83*, 3097–3112.

75. Crozier, P. S., Stevens, M. J., and Woolf, T. B. (2007) "How a small change in retinal leads to G-protein activation: initial events suggested by molecular dynamics calculations." *Proteins: Structure, Function, and Bioinformatics 66*, 559–574.

76. Isin, B., Schulten, K., Tajkhorshid, E., and Bahar, I. (2008) "Mechanism of signal propagation upon retinal isomerization: insights from molecular dynamics simulations of rhodopsin restrained by normal modes." *Biophysical Journal 95*, 789–803.

77. Huber, T., Menon, S., and Sakmar, T. P. (2008) "Structural Basis for Ligand Binding and Specificity in Adrenergic Receptors: Implications for GPCR-Targeted Drug Discovery." *Biochemistry 47*, 11013–11023.

78. Trabanino, R. J., Hall, S. E., Vaidehi, N., Floriano, W. B., Kam, V. W. T., and Goddard, W. A. (2004) "First Principles Predictions of the Structure and Function of G-Protein-Coupled Receptors: Validation for Bovine Rhodopsin." *Biophysical Journal 86*, 1904–1921.

79. Goddard, W. A., and Abrol, R. (2007) "3-Dimensional Structures of G Protein-Coupled Receptors and Binding Sites of Agonists and Antagonists." *Journal of Nutrition 137*, 1528S–1538.

80. Floriano, W. B., Vaidehi, N., and Goddard, W. (2004) "Making Sense of Olfaction through Predictions of the 3-D Structure and Function of Olfactory Receptors." *Chemical Senses 29*, 269–290.

81. Hall, S. E., Floriano, W. B., Vaidehi, N., and Goddard, W. A. (2004) "Predicted 3-D Structures for Mouse I7 and Rat I7 Olfactory Receptors and Comparison of Predicted Odor Recognition Profiles with Experiment." *Chemical Senses 29*, 595–616.

82. Kalani, M. Y. S., Vaidehi, N., Hall, S. E., Trabanino, R. J., Freddolino, P. L., Kalani, M. A., Floriano, W. B., Kam, V. W. T., and Goddard, W. A. (2004) "The predicted 3D structure of the human D2 dopamine receptor and the binding site and binding affinities for agonists and antagonists." *Proceedings of the National Academy of Sciences of the United States of America 101*, 3815–3820.

83. Vaidehi, N., Schlyer, S., Trabanino, R. J., Floriano, W. B., Abrol, R., Sharma, S., Kochanny, M., Koovakat, S., Dunning, L., Liang, M., Fox, J. M., de Mendonca, F. L., Pease, J. E., Goddard, I., William A., and Horuk, R. (2006) "Predictions of CCR1 Chemokine Receptor Structure and BX 471 Antagonist Binding Followed by Experimental Validation." *Journal of Biological Chemistry 281*, 27613–27620.

84. Li, Y., Zhu, F., Vaidehi, N., Goddard, W., Sheinerman, F., Reiling, S., Morize, I., Mu, L., Harris, K., Ardati, A. et al. (2007) "Prediction of the 3D Structure and

dynamics of Human DP G-Protein Coupled Receptor bound to an agonist and an antagonist." *Journal of the American Chemical Society 129*, 10720.

85. Heo, J., Han, S.-K., Wendel, N. V. J., Kekenes-Huskey, P., and William A. Goddard, I. (2007) "Prediction of the 3D Structure of FMRF-amide Neuropeptides Bound to the Mouse MrgC11 GPCR and Experimental Validation." *ChemBioChem 8*, 1527–1539.

86. Heo, J., Vaidehi, N., Wendel, J., and Goddard, W. A. (2007) "Prediction of the 3-D structure of rat MrgA G protein-coupled receptor and identification of its binding site." *Journal of Molecular Graphics and Modelling 26*, 800–812.

87. Heo, J., Ja, W. W., Benzer, S., and Goddard, W. A. (2008) "The Predicted Binding Site and Dynamics of Peptide Inhibitors to the Methuselah GPCR from Drosophila melanogaster." *Biochemistry 47*, 12740–12749.

88. Bray, J. K., and Goddard, W. A. (2008) "The structure of human serotonin 2c G-protein-coupled receptor bound to agonists and antagonists." *Journal of Molecular Graphics and Modelling 27*, 66–81.

89. White, S. H., and Wimley, W. C. (1998) "Hydrophobic interactions of peptides with membrane interfaces." *Biochimica et Biophysica Acta (BBA) - Reviews on Biomembranes 1376*, 339–352.

90. Tak Kam, V. W., and Goddard, W. A. (2008) "Flat-Bottom Strategy for Improved Accuracy in Protein Side-Chain Placements." *Journal of Chemical Theory and Computation 4*, 2160–2169.

91. Serrano-Vega, M. J., Magnani, F., Shibata, Y., and Tate, C. G. (2008) "Conformational thermostabilization of the β1-adrenergic receptor in a detergent-resistant form." *Proceedings of the National Academy of Sciences of the United States of America 105*, 877–882.

92. Mayo, S. L., Olafson, B. D., and Goddard, W. A. (1990) "DREIDING: a generic force field for molecular simulations." *The Journal of Physical Chemistry 94*, 8897–8909.

93. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) "CHARMM: A program for macromolecular energy, minimization, and dynamics calculations." *Journal of Computational Chemistry 4*, 187–217.

94. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.." *Nucleic Acids Research 25*, 3389–402.

95. Katoh, K., Misawa, K., Kuma, K.-i., and Miyata, T. (2002) "MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform." *Nucleic Acids Research 30*, 3059–3066.

96. Katoh, K., Kuma, K.-i., Toh, H., and Miyata, T. (2005) "MAFFT version 5: improvement in accuracy of multiple sequence alignment." *Nucleic Acids Research 33*, 511–518.

97. Katoh, K., and Toh, H. (2008) "Recent developments in the MAFFT multiple sequence alignment program." *Briefings in Bioinformatics 9*, 286–298.

98. Swaminath, G., Xiang, Y., Lee, T. W., Steenhuis, J., Parnot, C., and Kobilka, B. K. (2004) "Sequential Binding of Agonists to the β2 Adrenoceptor: Kinetic Evidence for Intermediate Conformational States." *Journal of Biological Chemistry 279*, 686–691.

99. Schwartz, T. W., Frimurer, T. M., Holst, B., Rosenkilde, M. M., and Elling, C. E. (2006) "Molecular mechanism of 7TM receptor activation — a global toggle switch model." *Annual Review of Pharmacology and Toxicology 46*, 481–519.

100. Bhattacharya, S., Hall, S. E., and Vaidehi, N. (2008) "Agonist-induced conformational changes in bovine rhodopsin: insight into activation of G-protein-coupled receptors." *Journal of Molecular Biology 382*, 539–555.

101. Murakami, M., and Kouyama, T. (2008) "Crystal structure of squid rhodopsin." *Nature 453*, 363–367.

102. Reynolds, K. A., Katritch, V., and Abagyan, R. (2009) "Identifying conformational changes of the β2 adrenoceptor that enable accurate prediction of ligand/receptor interactions and screening for GPCR modulators." *Journal of Computer-Aided Molecular Design 23*, 273–288.

103. Beck, M., Sakmar, T. P., and Siebert, F. (1998) "Spectroscopic Evidence for Interaction between Transmembrane Helices 3 and 5 in Rhodopsin." *Biochemistry 37*, 7630–7639.

104. Weitz, C. J., and Nathans, J. (1993) "Rhodopsin activation: effects of the metarhodopsin I-metarhodopsin II equilibrium of neutralization or introduction of charged amino acids within putative transmembrane segments." *Biochemistry 32*, 14176–14182.

105. Weitz, C. J., and Nathans, J. (1992) "Histidine residues regulate the transition of photoexcited rhodopsin to its active conformation, metarhodopsin II." *Neuron 8*, 465–472.

106. Chalmers, D. T., and Behan, D. P. (2002) "The use of constitutively active GPCRs in drug discovery and functional genomics." *Nature Reviews: Drug Discovery 1*, 599–608.

107. Balmforth, A. J., Lee, A. J., Warburton, P., Donnelly, D., and Ball, S. G. (1997) "The Conformational Change Responsible for AT1 Receptor Activation Is Dependent upon Two Juxtaposed Asparagine Residues on Transmembrane Helices III and VII." *Journal of Biological Chemistry 272*, 4245–4251.

108. Rasmussen, S., Jensen, A., Liapakis, G., Ghanouni, P., Javitch, J., and Gether, U. (1999) "Mutation of a highly conserved aspartic acid in the β2 adrenergic receptor: constitutive activation, structural instability, and conformational rearrangement of transmembrane segment 6." *Molecular Pharmacology 56*, 175–184.

109. Alewijnse, A. E., Timmerman, H., Jacobs, E. H., Smit, M. J., Roovers, E., Cotecchia, S., and Leurs, R. (2000) "The Effect of Mutations in the DRY Motif on the

Constitutive Activity and Structural Instability of the Histamine H2 Receptor." *Molecular Pharmacology 57*, 890–898.

110. Huang, P., Li, J., Chen, C., Visiers, I., Weinstein, H., and Liu-Chen, L. (2001) "Functional Role of a Conserved Motif in TM6 of the Rat μ Opioid Receptor: Constitutively Active and Inactive Receptors Result from Substitutions of Thr6.34 (279) with Lys and Asp." *Biochemistry 40*, 13501–13509.

111. Chen, S., Lin, F., Xu, M., and Graham, R. M. (2002) "Phe303 in TMVI of the α1B-Adrenergic Receptor Is a Key Residue Coupling TM Helical Movements to G-protein Activation." *Biochemistry 41*, 588–596.

112. Han, M., Smith, S. O., and Sakmar, T. P. (1998) "Constitutive activation of opsin by mutation of methionine 257 on transmembrane helix 6." *Biochemistry 37*, 8253–8261.

113. Kjelsberg, M. A., Cotecchia, S., Ostrowski, J., Caron, M. G., and Lefkowitz, R. J. (1992) "Constitutive activation of the α1b-adrenergic receptor by all amino acid substitutions at a single site. Evidence for a region which constrains receptor activation." *Journal of Biological Chemistry 267*, 1430–1433.

114. Farrens, D. L., Altenbach, C., Yang, K., Hubbell, W. L., and Khorana, H. G. (1996) "Requirement of Rigid-Body Motion of Transmembrane Helices for Light Activation of Rhodopsin." *Science 274*, 768–770.

115. Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E.,

Chipot, C., Skeel, R. D., Kalé, L., and Schulten, K. (2005) "Scalable molecular dynamics with NAMD." *Journal of Computational Chemistry 26*, 1781–1802.

116. Timothy R. Forester, W. S. (1998) "SHAKE, rattle, and roll: Efficient constraint algorithms for linked rigid bodies." *Journal of Computational Chemistry 19*, 102–111.

117. Jacobson, M. P., Pincus, D. L., Rapp, C. S., Day, T. J. F., Honig, B., Shaw, D. E., and Friesner, R. A. (2004) "A hierarchical approach to all-atom protein loop prediction." *Proteins: Structure, Function, and Bioinformatics 55*, 351–367.

118. Zhu, K., Pincus, D. L., Zhao, S., and Friesner, R. A. (2006) "Long loop prediction using the protein local optimization program." *Proteins: Structure, Function, and Bioinformatics 65*, 438–452.

119. Raghava, G. P. S. (2002) "APSSP2: A combination method for protein secondary structure prediction based on neural network and example based learning." *CASP 5*, A–132.

120. Ewing, T. J. A., Makino, S., Skillman, A. G., and Kuntz, I. D. (2001) "DOCK 4.0: search strategies for automated molecular docking of flexible molecule databases." *Journal of Computer-Aided Molecular Design 15*, 411–428.

121. Nakanishi, J., Takarada, T., Yunoki, S., Kikuchi, Y., and Maeda, M. (2006) "FRET-based monitoring of conformational change of the α2 adrenergic receptor in living cells." *Biochemical and Biophysical Research Communications 343*, 1191–1196.

122. Cavasotto, C. N., Orry, A. J. W., Murgolo, N. J., Czarniecki, M. F., Kocsi, S. A., Hawes, B. E., O'Neill, K. A., Hine, H., Burton, M. S., Voigt, J. H., Abagyan, R. A., Bayne, M. L., and Monsma, F. J. (2008) "Discovery of Novel Chemotypes to a G-Protein-Coupled Receptor through Ligand-Steered Homology Modeling and Structure-Based Virtual Screening." *Journal of Medicinal Chemistry 51*, 581–588.

123. Mobarec, J. C., Sanchez, R., and Filizola, M. (2009) "On Template Selection for Homology Modeling of G-Protein Coupled Receptors." *Biophysical Journal 96*, 653a.

124. Pollastri, G., Vullo, A., Frasconi, P., and Baldi, P. (2006) "Modular DAG-RNN architectures for assembling coarse protein structures." *Journal of Computational Biology 13*, 631–650.

125. Jones, D. T. (1999) "Protein secondary structure prediction based on position-specific scoring matrices." *Journal of Molecular Biology 292*, 195–202.

126. Mohamadi, F., Richards, N. G. J., Guida, W. C., Liskamp, R., Lipton, M., Caufield, C., Chang, G., Hendrickson, T., and Still, W. C. (1990) "MacroModel-an integrated software system for modeling organic and bioorganic molecules using molecular mechanics." *Journal of Computational Chemistry 11*, 440.

127. *Jaguar*, version 7.5, Schrödinger, LLC, New York, NY, 2008..

128. Pallavicini, M., Fumagalli, L., Gobbi, M., Bolchi, C., Colleoni, S., Moroni, B., Pedretti, A., Rusconi, C., Vistoli, G., and Valoti, E. (2006) "QSAR study for

a novel series of ortho disubstituted phenoxy analogues of α1-adrenoceptor antagonist WB4101." *European Journal of Medicinal Chemistry 41*, 1025–1040.

129. Waugh, D. J. J., Gaivin, R. J., Zuscik, M. J., Gonzalez-Cabrera, P., Ross, S. A., Yun, J., and Perez, D. M. (2001) "Phe-308 and Phe-312 in Transmembrane Domain 7 Are Major Sites of α1-Adrenergic Receptor Antagonist Binding; Imidazoline Agonists Bind Like Antagonists." *Journal of Biological Chemistry 276*, 25366–25371.

130. Starke, K., Borowski, E., and Endo, T. (1975) "Preferential blockade of presynaptic alpha-adrenoceptors by yohimbine." *European Journal of Pharmacology 34*, 385.

131. Wang, C. D., Buck, M. A., and Fraser, C. M. (1991) "Site-directed mutagenesis of α2A-adrenergic receptors: identification of amino acids involved in ligand binding and receptor activation by agonists." *Molecular Pharmacology 40*, 168–179.

132. Suryanarayana, S., Daunt, D. A., Von Zastrow, M., and Kobilka, B. K. (1991) "A point mutation in the seventh hydrophobic domain of the α2 adrenergic receptor increases its affinity for a family of beta receptor antagonists." *Journal of Biological Chemistry 266*, 15488–15492.

133. Sugimoto, Y., Fujisawa, R., Tanimura, R., Lattion, A. L., Cotecchia, S., Tsujimoto, G., Nagao, T., and Kurose, H. (2002) "β1-Selective Agonist (-)-1-(3, 4-Dimethoxyphenetylamino)-3-(3, 4-dihydroxy)-2-propanol [(-)-RO363] Differ-

entially Interacts with Key Amino Acids Responsible for β1-Selective Binding in Resting and Active States." *Journal of Pharmacology and Experimental Therapeutics 301*, 51–58.