

## 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  $\eta$  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.<sup>12</sup> For several years, only one GPCR, bovine rhodopsin, was available for this kind of modeling.<sup>28-36</sup> 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.<sup>65-67</sup> The publication of several other GPCR crystal structures, including the human  $\beta 2$  adrenergic receptor ( $\beta 2$ ),<sup>37-40</sup> the closely related turkey  $\beta 1$  adrenergic receptor ( $\beta 1$ ),<sup>41</sup> and the human adenosine  $A_{2A}$  receptor<sup>42</sup> revealed close structural similarity among the structures. This, as well as advances in ligand-steered homology modeling,<sup>122</sup> 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.<sup>123</sup> 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  $\beta 2$  and turkey  $\beta 1$  GPCRs with the other adrenergic receptors range from 24%–60% overall and up to 85% in the TM regions.

	$\beta 2$ Human		$\beta 1$ Turkey	
	All%	TM%	All%	TM%
$\beta 1$	41	68	60	85
$\beta 3$	42	59	39	63
$\alpha 1a$	28	41	26	44
$\alpha 1b$	30	44	27	44
$\alpha 1d$	30	43	27	43
$\alpha 2a$	26	39	27	49
$\alpha 2b$	24	38	24	46
$\alpha 2c$	24	38	24	44
$\beta 2$	–	–	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  $\beta 1$  receptor and the human  $\beta 2$  receptor, which is closer by sequence identity than that between human  $\beta 2$  and human  $\beta 1$ . 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  $\beta 2$  with respect to  $\eta$ ,  $\theta$ , and  $\phi$  angles. Although there are large differences, the most consistently similar structures are the ones with the highest sequence identity, such as  $\beta 1$  and  $\beta 2$  with 52% TM sequence identity with respect to  $\beta 2$ . 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  $\beta 2$  and

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

	$\eta$				$\theta$				$\phi$			
	$\beta 1$	A2a	rhod.	opsin	$\beta 1$	A2a	rhod.	opsin	$\beta 1$	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  $\beta 2$  receptors are from the same species,  $\beta 1$  (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  $\beta$  receptors should be modeled the most easily, while the more distantly related  $\alpha 2$  receptors may require more refinement. In this work, both the  $\beta 1$  turkey and  $\beta 2$  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<sup>92</sup> with CHARMM22<sup>93</sup> charges. Side chain placement was determined with SCREAM<sup>90</sup> 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  $\beta 2$  (2RH1) and  $\beta 1$  (2VT4) crystal structures, referred to as template structures. The  $\beta 2$  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., " $\alpha 1a$ - $\beta 1$ " is the human  $\alpha 1a$  homology model built

on the turkey  $\beta 1$  template.

*Prediction of transmembrane regions:* Preliminary TM helices were obtained via MAFFT<sup>95-97</sup> alignment to the template receptor. These alignments were compared to the predicted helical regions from secondary structure prediction methods Porter,<sup>124</sup> APSSP2,<sup>119</sup> and PSIPred.<sup>125</sup> 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  $\beta 1$  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  $\eta$  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.



Model,<sup>126</sup> and Mulliken charge calculation with Jaguar (B3LYP with 6-31G\*\* basis set).<sup>127</sup> For antagonists with multiple nitrogens, pKa values were calculated with the pKa module in Jaguar.

Each receptor was validated using both  $\beta 1$  and  $\beta 2$  models, docked individually with canonical subtype selective antagonists. The binding sites were predicted using ScanBindSite.pl.<sup>80</sup> 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<sup>3.32</sup> 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<sup>70</sup> 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 antagonists were chosen and docked for each subtype. For the  $\alpha 1$  subtype we chose WB4101 and prazosin. WB4101 has been extensively studied<sup>128,129</sup> as an  $\alpha 1$  selective antagonist and a particularly potent  $\alpha$  blocker, while prazosin was one of the first commercially developed  $\alpha$  blocker. Yohimbine is a traditional  $\alpha 2$  antagonist,<sup>130</sup> 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.<sup>131,132</sup> For  $\beta 1$  and  $\beta 3$ , the  $\beta 1$  selective antagonist (–) RO 363 was docked. The residues implicated in binding<sup>133</sup> 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

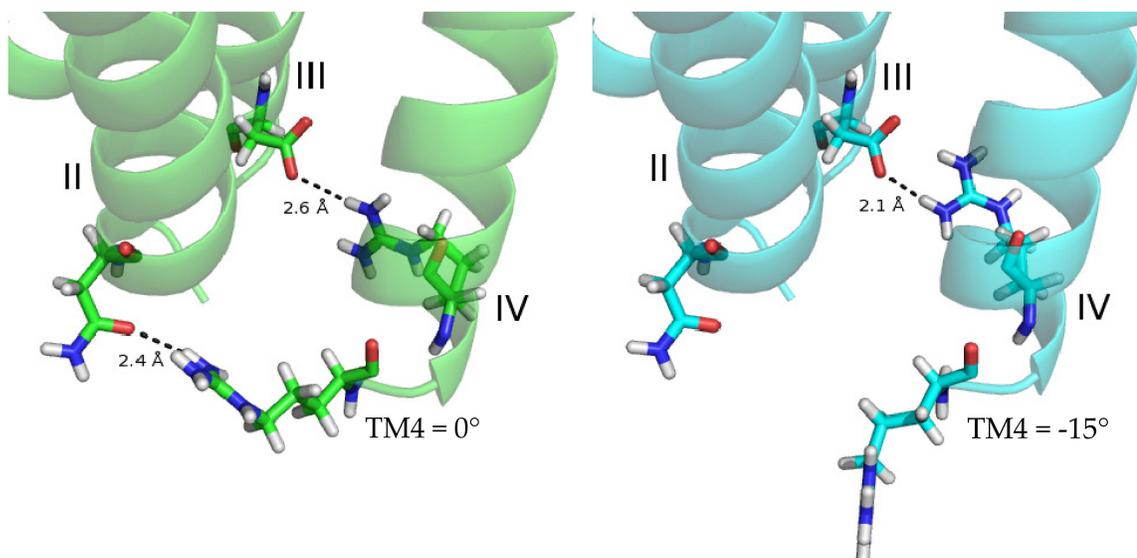
While the GenSembler method was developed using 30° increments for sampling, later studies in the Goddard group showed that BiHelix will resolve differences in helix  $\eta$  rotation as fine as 10°. This analysis exponentially increases the time necessary for full 360° sampling, so the 30° scan remains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

	$\beta 2$ Template								$\beta 1$ Template							
	H1	H2	H3	H4	H5	H6	H7	Energy	H1	H2	H3	H4	H5	H6	H7	Energy
$\beta 1$	0	0	0	345	0	0	0	-178.9	0	0	0	345	0	0	0	-202.1
$\beta 3$	0	0	0	0	0	0	0	-183.1	0	0	0	345	0	0	0	-186.3
$\alpha 1a$	0	0	0	345	0	15	0	254.2	0	0	0	345	0	0	0	209.8
$\alpha 1b$	0	0	0	345	0	0	345	246.5	0	0	0	345	0	60	315	193.8
$\alpha 1d$	0	0	0	105	270	0	90	63.2	0	0	0	90	0	75	0	138.0
$\alpha 2a$	0	0	0	345	0	0	0	153.8	0	0	0	345	0	0	0	117.3
$\alpha 2b$	180	15	0	345	0	0	105	193.1	0	0	0	345	270	105	0	119.0
$\alpha 2c$	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  $\eta$  rotations analogous to the template crystal structure as the lowest in minimized energy after building the full bundles. Those least related to the template structures,  $\alpha 2b$ ,  $\alpha 2c$ , and  $\alpha 1d$ , 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<sup>4.41</sup> and Arg<sup>4.40</sup> are vertically positioned to interact with Asp<sup>3.49</sup> and Asn<sup>2.40</sup>, respectively, if the  $\eta$  rotation allows it. In the 0° position, both residues on TM4 may interact with their



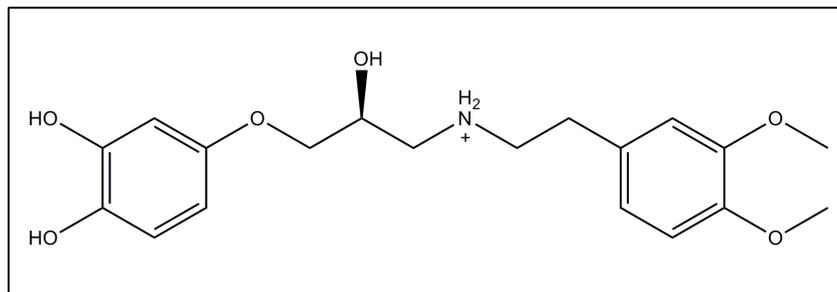
**Figure 4.1:** BiHelix/CombiHelix results indicate a counterclockwise rotation of TM4 is favored over the crystallographic orientation for most receptor-template pairs. The interactions responsible for this preference are illustrated below for the  $\beta_1$ - $\beta_1$  case.

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  $\beta 2$  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  $\beta 1$  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  $\beta 1$  template and the  $\beta 1$  human and  $\beta 3$  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,  $\beta 1$ - $\beta 2$  and  $\beta 3$ - $\beta 2$ , have TM identities of 68% and 59% and overall identities of 41% and 42%. The overall sequence identity implies the  $\beta 3$ - $\beta 2$  structure should be slightly better than the  $\beta 1$ - $\beta 2$  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  $\beta$  receptors, the predicted binding site should indicate which residues are responsible for this ligand's selectivity.

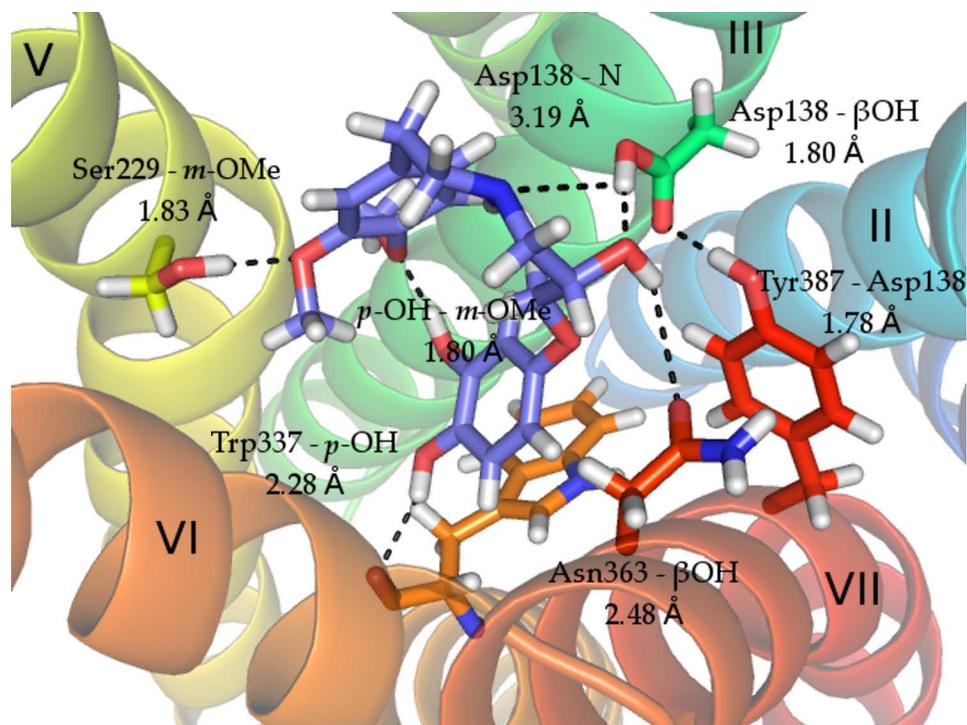
rotations, the energy of the best bundle increases, indicating a less favorable match between template and target.

### 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.*<sup>133</sup> 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<sup>2.63</sup> or Phe359<sup>7.35</sup> to alanine only have a small effect on binding, but combined they decrease antagonist affinity 25-fold. Leu110<sup>2.56</sup> mutated to alanine results in a seven-fold decrease in affinity, but that change is augmented when paired with the Phe359<sup>7.35</sup> 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.

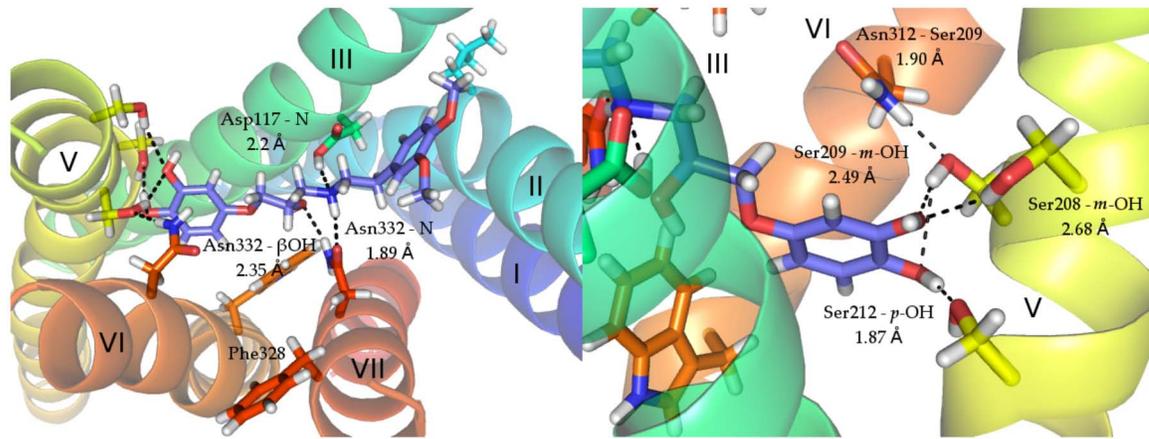
The BiHelix / CombiHelix for both  $\beta_1$  human and  $\beta_3$  in both templates resulted in TM bundles similar to the crystal templates. The only variation occurred in TM4, and for  $\beta_3$ - $\beta_2$  all helices were found in the crystal orientations. Neither template-target pair featured 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<sup>7.35</sup>. One pose, however, positioned one side of the ligand where it might interact with Leu110<sup>2.56</sup> 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.



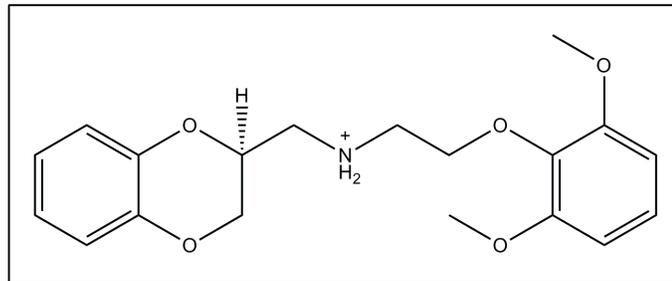
**Figure 4.3:** The best docked conformation of (–) RO-363 (in blue) to the  $\beta_1$ - $\beta_2$  homology model was a folded conformation, with stabilizing interactions from the adrenergic agonist pharmacophore as well as an internal hydrogen bond.

For both cases, the  $\beta_2$  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  $\beta_1$  and  $\beta_3$  human receptors.

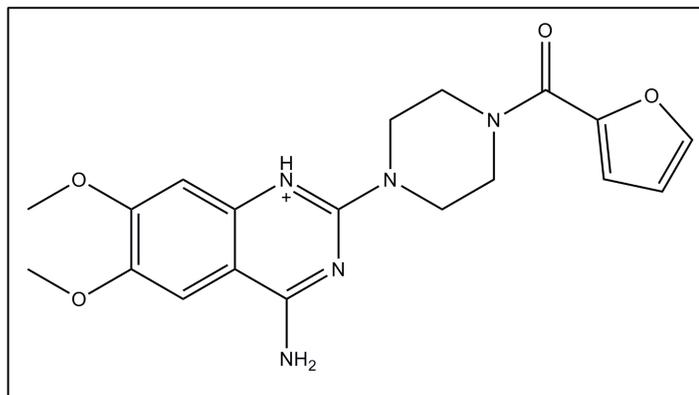
Two sets of studies explored antagonist binding to  $\alpha_1$  receptors: one focusing primarily on WB4101 (Figure 4.5) and prazosin (Figure 4.6) binding to  $\alpha_1a$ ,<sup>129</sup> the other performing extensive SAR on WB4101 and comparing the differences among  $\alpha_1a$ ,  $\alpha_1b$ , and a serotonin receptor.<sup>128</sup> Waugh, *et al.* only studied  $\alpha_1a$ , 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.



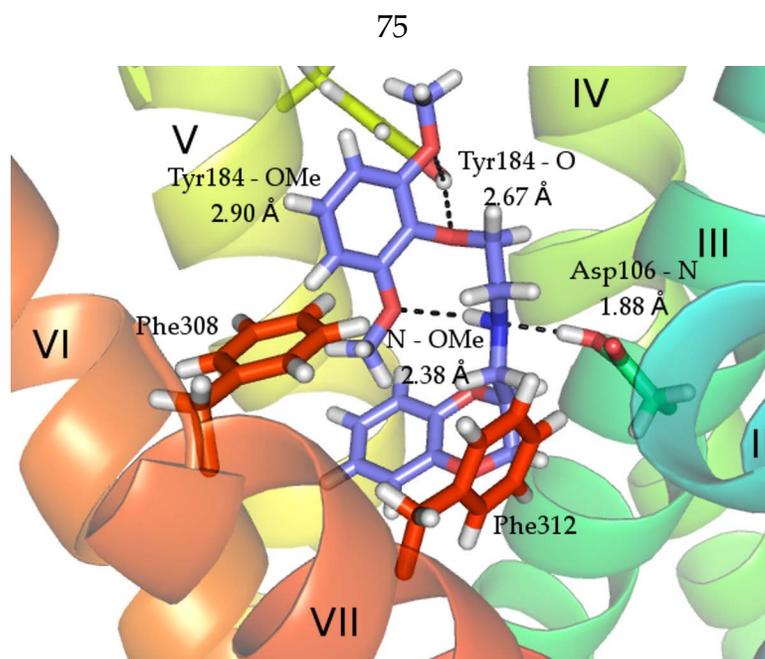
**Figure 4.5:** WB4101 has been studied extensively as an  $\alpha_1$  selective antagonist.



**Figure 4.6:** Prazosin is a selective  $\alpha_1$  antagonist, and is prescribed as an antihypertensive drug.

residues mutated are conserved throughout the  $\alpha_1$  subtype: Phe<sup>7.39</sup> and Phe<sup>7.35</sup>, both at the top of TM7. WB4101 is  $\alpha_1$  selective, but binds more strongly to  $\alpha_{1a}$  than to  $\alpha_{1b}$  or  $\alpha_{1d}$ . 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<sup>7.35</sup>, is more important for prazosin binding than for WB4104, and that the opposite holds for Phe<sup>7.39</sup>. The predicted binding sites should reflect this difference.

The low-energy rotations for  $\alpha_{1a}$ - $\beta_1$  are the most similar to the template rotations of all the  $\alpha_1$  homology results, with only the consistent TM4 anticlockwise rotation deviating from the strictly analogous structure. This structure binds prazosin well, and while the key Phe308<sup>7.35</sup> residue does appear in the cavity analysis (Table 4.5) as a stabilizing residue, Phe312<sup>7.39</sup> 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<sup>3.32</sup> interacts with the protonated amine (1.88 Å); Tyr184<sup>5.36</sup> binds to the O-methyl and ether groups simultaneously



**Figure 4.7:** The  $\alpha 1a$ - $\beta 1$  homology model docked WB4101 in a curled conformation stabilized primarily through interactions with Asp106<sup>3,32</sup>, 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<sup>7,35</sup> and Phe312<sup>7,39</sup> 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.  $\alpha 1a$ - $\beta 2$ , 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  $\alpha 1a$  to the  $\beta$  templates, this structure requires further refinement before the binding site can be verified.

$\alpha 1b$  favors a slight anticlockwise rotation of TM7, the conserved rotation of TM4, and in the  $\beta 1$  template, an additional clockwise rotation of TM6. This  $\alpha 1b$ - $\beta 1$  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  $\alpha 1a$ - $\beta 1$  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	<b>Phe312</b>	<b>-7.386</b>	<b>-1.411</b>	<b>-8.797</b>
Met292	-2.039	0.025	-2.014	Asp106	-2.373	-1.299	-3.672
<b>Phe308</b>	<b>-1.582</b>	<b>-0.062</b>	<b>-1.644</b>	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	<b>Phe308</b>	<b>-1.049</b>	<b>-0.180</b>	<b>-1.229</b>
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
<b>Phe312</b>	<b>-0.142</b>	<b>0.224</b>	<b>0.082</b>	Leu80	-0.173	0.063	-0.110

of the  $\alpha 1$  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  $\alpha 1a$  receptor. These structures require more specialized refinement before they can be fully validated.

With comparable sequence identity to the other  $\alpha 1$  receptors,  $\alpha 1d$  favors rotations similarly divergent from the template rotations. Both templates showed TM4 rotations, but not the same magnitude as the other structures:  $\alpha 1d\text{-}\beta 2$  prefers a  $105^\circ$  clockwise rotation of TM4, and  $\alpha 1d\text{-}\beta 1$  favors a  $90^\circ$  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  $\eta$  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.

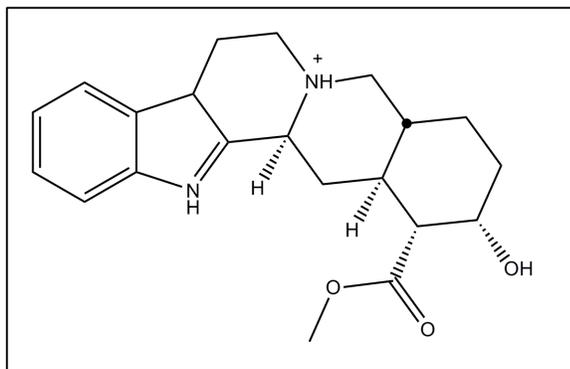
Even with this shift in helix  $\eta$  rotations, it was possible to obtain plausible docked structures for WB4101 and prazosin for the structure built on the  $\beta 2$  template. The WB4101 binding site is shifted towards the TM1-2-7 pocket, with Ser153<sup>2,61</sup> forming a hydrogen bond with one of the dioxane oxygens, the larger phenyl group with O-methyl substituents buried deep in the binding pocket, and Phe384<sup>7,35</sup>

**Table 4.6:** The top ten residues in the cavity analysis for  $\alpha 1d$ - $\beta 2$  include the key residues implicated in mutation studies, Phe384<sup>7.35</sup> and Phe388<sup>7.39</sup> 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<sup>7.39</sup> both creating  $\pi$  stacking interactions with the other side of the ligand. The key residue for WB4101 (Phe388<sup>7.39</sup>) 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<sup>7.39</sup>. 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<sup>7.34</sup> than with Phe388<sup>7.39</sup>, also verifying the integrity of this binding site. That these structures were obtained for the  $\beta 2$  template structure, for which there is slightly higher sequence identity, rather than for the  $\beta 1$  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  $\alpha 2$  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  $\alpha_2$  adrenergic inverse agonist.

*et al.* recently characterized the binding affinity of several agonists and antagonists with  $\alpha_2a$  and mutant receptors containing changes to TMs 2, 3, and 5.<sup>131</sup> The most striking effect came from the mutation of Asp113<sup>3.32</sup> 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<sup>5.42</sup> and Ser204<sup>5.46</sup>. 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<sup>3.49</sup> and Asp79<sup>2.50</sup>, are deeply buried in the TM core. The highly conserved Asp<sup>3.49</sup> 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<sup>2.50</sup> is positioned to interact with Asn<sup>1.50</sup> and Asn<sup>7.49</sup> 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<sup>7.39</sup>, implying the upper section of TM7 is more likely to

directly interact with the antagonist.<sup>132</sup>

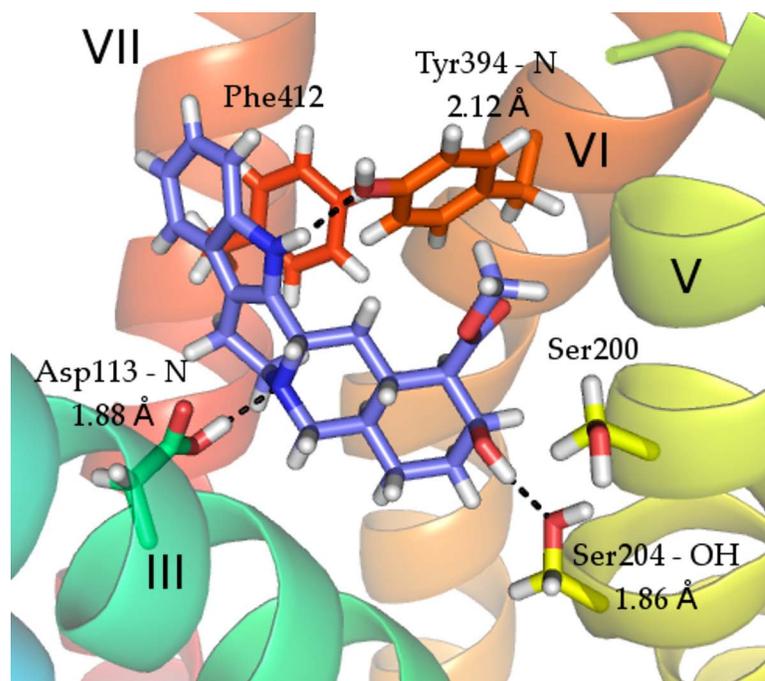
Both the  $\beta 1$  and  $\beta 2$  templates yielded identical helix rotations for the  $\alpha 2a$  structure, so the differences in binding can be attributed to changes in the helix orientations due to the template. The best  $\alpha 2a$ - $\beta 1$  docked structure by local cavity interaction (Figure 4.9) showed close interactions with Asp113<sup>3.32</sup> (1.88 Å to the protonated amine) and Ser204<sup>5.46</sup> (1.86 Å to the hydroxide). Ser200<sup>5.42</sup> 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<sup>7.39</sup> forms a  $\pi$  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  $\alpha 2a$ - $\beta 2$  structure, which is consistent with the increased sequence identity between target and template:  $\alpha 2a$  and  $\beta 1$  turkey share 46% TM sequence identity, while  $\alpha 2a$  and  $\beta 2$  only share 39%.

This structure also shows Tyr394<sup>6.55</sup> interacting directly with the indole nitrogen on the ligand (2.12 Å). This residue, analogous to Asn<sup>6.55</sup> in the  $\beta$  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  $\alpha 2b$  rotations shift TMs 1, 2, 4, and 7 for the  $\beta 2$  template and TMs 4, 5, and 6 for the  $\beta 1$  template. While the  $\alpha 2b$ - $\beta 2$  structure does form good protein-ligand interactions with Asp92<sup>3.32</sup> (1.81 Å to the protonated amine) and an addi-

**Table 4.7:** The key residue implicated by mutation studies, Phe412<sup>7,39</sup>, appears in the cavity analysis for  $\alpha2a$ - $\beta1$  docked with yohimbine. However, Asp113<sup>3,32</sup> 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
<b>Phe412</b>	<b>-4.819</b>	<b>0.614</b>	<b>-4.206</b>
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  $\alpha1$  and  $\beta$  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  $\alpha2a$  docked structure is positioned close to Phe412<sup>7,39</sup> implicated in mutation studies.

tional unpredicted interaction with Glu73<sup>2.65</sup> (1.75 Å to the hydroxide), Phe412<sup>7.39</sup> 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 responsible for a conserved interaction, but considering the lower sequence identity between  $\alpha 2b$  and the available crystal structures, it is more likely that these rotations are not correct for  $\alpha 2b$ . Both  $\alpha 2b-\beta 1$  and  $\alpha 2b-\beta 2$  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  $\eta$  rotation analysis may show a better binding site for this system.

The final rotations favored for  $\alpha 2c$  are closer to the template crystal structures than those for  $\alpha 1b$ : the recurring 15° anticlockwise rotation of TM4 appears for the  $\alpha 2c-\beta 2$  structure, and TM6 is rotated 45° clockwise for  $\alpha 2c-\beta 1$ . The primary interaction in the binding cavity for the  $\alpha 2c-\beta 1$  structure involves Phe423<sup>7.39</sup>, a cation- $\pi$  interaction with the protonated amine, but this functional group on the ligand is expected to interact with the necessary Asp131<sup>3.32</sup> 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<sup>5.42</sup> or Ser218<sup>5.46</sup>, 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  $\alpha 2c-\beta 2$  structure did not support a plausible

**Table 4.8:** Both key residues involved in yohimbine binding, Phe423<sup>7,39</sup> and Asp131<sup>3,32</sup>, appear in the cavity analysis for  $\alpha 2c$ - $\beta 1$  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  $\beta 2$  template may be too different from the native  $\alpha 2c$  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  $\eta$  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  $\theta$  and  $\phi$  to the  $\eta$  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  $\alpha 1a$  antagonist interaction with the conserved Tyr184<sup>5,36</sup> 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<sub>2A</sub> 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  $\beta 2$ ) with those that have less (like the  $\alpha 1$  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  $\beta 2$  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  $\beta 2$  and  $\beta 1$  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.