Chapter 5

Prediction of Structure and Antagonist Binding Site:

Rat CCR1

Abstract

This chapter summarizes preliminary work on structural prediction of another rodent chemokine receptor, rat CCR1 (rCCR1). Using MembStruk v4.0, we aim to develop a structural model that can lend insight into the differential binding of BX471 to human, mouse, and rat CCR1. Earlier studies in our group (Chapter 2, Chapter 4) focused on function prediction as a means of structural validation, an approach which is not yet possible for rCCR1 due to lack of experimental binding affinity data. Thus, the scope of this study is discussion of structure prediction and analysis of the BX471 binding cavity.

Our results implicate two residues conserved in the rodent structures, Phe172 (TM4) and Phe260 (TM6), in ligand recognition and stabilization by pi-pi stacking interactions. The conformation of the docked BX471 antagonist does not allow these energetically favorable interactions in mCCR1structure, suggesting comparatively stronger binding to rCCR1. Additionally, we find that Tyr113 (TM3), the key contributor in hCCR1 binding cavities, is quite repulsive in the rCCR1 binding cavity, The calculated binding energy of -28.38 kcal/mol indicates that BX471 binds best to hCCR1 (-39.25 kcal/mol), then rCCR1 (-28.38 kcal/mol), and mCCR1 (-21.81 kcal/mol). Comparison to the experimental trend in binding affinities yields exact agreement with our data.

Computational Methods

Structure Prediction

The computational methods for structure and ligand binding predictions are the same as those described for hCCR1 in Chapter 2 (MembStruk v4.0)[1-3]. Where appropriate, the changes in procedure for rCCR1 are discussed below.

I. TM domain length prediction

The multiple sequence alignment profile for rCCR1 was generated by the set of sequences (same for hCCR1 and mCCR1) shown in Appendix A. The primary sequence of rCCR1 is shown in Figure 5-1. Illustrated in Figure 5-2 is a hydrophobicity profile calculated using TM2ndS. The TM domains (Figure 5-3) are determined from the baseline in this hydrophobicity plot. Both the hydrophobic centers and the TM regions are identical to what we calculated for mCCR1, as expected for such similar receptors. With respect to hCCR1, there are 34 mutations in the TM regions (same as mCCR1). The distribution is also very similar, with a maximum difference of one residue change in each helix: TM1 (7), TM2 (4), TM3 (2), TM4 (5), TM5 (5), TM6 (8), and TM7 (3). These mutations are anywhere from 3-9 residues away from either side of the hydrophobic center.

II. TM bundle assembly and optimization of helical translation and rotation

The 7.5Å electron density map of frog rhodopsin[4] is used to position each canonical α-helix in the bundle. The helices are then translated to fit the plane of maximum hydrophobicity, and rotations optimized via net hydrophobic moment reorientation for TMs 1-2, and 4-7. The RotMin energy minimization procedure is used for TM3, the only TM not directly interacting with the membrane. An energy scan profile is obtained after 360° rotation (5° increments) for the main chain, while side chains are reassigned and the potential energy minimized. This profile shows two minima, at -55° and 180° for TM3 (Figure 5-4). Both rotational configurations were visually inspected, and the latter orientation appeared to make more interhelical contacts in the bundle (discussed in the next section).

Results and Discussion

Structural Features of rCCR1: Interhelical Interactions

I. Location of Conserved Residues and Comparison to BR, hCCR1, and mCCR1

We find fewer polar interactions in the predicted rCCR1 structure compared to both hCCR1 and mCCR1. The typical water mediated salt bridge between Lys94 (TM2) and Glu287 (TM7) is much weaker in this structure (8.61 Å) compared to either hCCR1 or mCCR1 (Figure 5-7). As a result, we investigated the possibility of rotating TM7 by 90°, but doing so would orient many of the hydrophobic residues (Ile, Leu, Val) inside the bundle, not facing the lipid membrane. Thus, this rotation was not performed. The bidentate salt bridge between Glu120 (TM3) and His293 (TM7) is very similar in magnitude to the mCCR1 model (Figure 5-8). As noted in previous chapters, the corollary to this interaction in BR is between Glu122 (TM3) and His211 (TM5). Another hydrogen bond is seen between the very highly conserved residues (across many GPCRs) Asp80 (TM2) and Asn297 (TM7), at a distance of 3.13 Å (Figure 5-9). Although this interaction is seen in the human and mouse structures, it is much stronger here and not mediated by water, contrary to the observation in BR[5].

Figure 5-10 shows Ser79 (TM2) hydrogen bonding with the most conserved TM4 residue, Trp158. In the hCCR1 structure we observed this same interaction, albeit at a larger distance of 5.56 Å. Comparison to the crystal structure of BR shows a hydrogen

bond between this Trp on TM4 and an Asn on TM2. In our work, mCCR1 is the only structure where we observe such an Asn-Trp interaction.

The rCCR1 has 18% overall sequence identity to BR, and 22% in the TM region. The main chain RMSD is 5.91 Å, and an overlay of the two structures is shown in Figure 5-11. This value is larger than the difference between hCCR1 and rCCR1 (4.63 Å from Figure 5-12), which has a sequence identity of 83% in the TM region, and 79% overall. Thus, the human and mouse receptor structures are structurally more similar (2.91 Å) than any other combination. This is a surprising result, as we would expect the rat and mouse structures (Figure 5-13) to be most alike, especially considering the high degree of sequence similarity (91% in the TM regions). These fundamental structural differences could not be elucidated as clearly using the typical homology modeling approach, even for highly homologous receptors such as mCCR1 and rCCR1. Yet, helical bending in rCCR1 occurs at the same (conserved) prolines in mCCR1. Sequence alignment to hCCR1 (Figure 5-14) shows that all such prolines are indeed conserved. Tables 4-1 shows the tilt and bend angles for each helix.

Docking of BX471 to rCCR1

HierDock[6-8] is used to predict the binding site of the antagonist BX471 in rCCR1. The procedure is identical to that used for mCCR1 (Chapter 4) binding site determination.

Binding site determination via coarse-grain docking

A total of 11625 spheres partitioned over 49 $(10Å)^3$ boxes were obtained for rCCR1 (Figure 5-5). After excluding binding regions in the loops and EC/IC boundaries, DOCK[9] was used to perform coarse-grain docking in each box. The best conformations in each box are hierarchically scored, yielding a total of 191 spheres merged from regions 27, 31, and 32 (Figure 5-6).

Predicted Binding Site of BX471

BX471 (Figure 5-15) was docked to rCCR1 as a starting point to gain insight into differential binding between human and rodent structures. Validation of this structure through function prediction has not been accomplished as of yet, due to the dearth of antagonist binding affinities published in the literature.

Predicted Binding Site for BX471 (K_i = 100 nM) and Comparison to hCCR1 and mCCR1

The binding domain of BX471 in rCCR1 involves contributions from TMs 2-7. Figure 5-16 shows the different interactions in the binding cavity that stabilize this docked conformation. Interestingly, the most dominant contributions in this cavity are attributable to pi-pi stacking interactions with Phe172 (TM4) and Phe260 (TM6). This mode of interaction was not observed in the hCCR1 binding site, because Phe172 is replaced by Ser172, and Phe260 substituted by Leu260. However, the calculated binding

energy of BX471 in rCCR1 is actually lower than BX471 in hCCR1, due to the rather large (5.30 kcal/mol) repulsion caused by Tyr113 (TM3) in rCCR1. This is in stark contrast to the hCCR1 binding site, where Tyr113 stacked directly below the plane of the piperidine ring and contributed most favorably to the binding energy. We attempted to reassign the side chain of Tyr113, but were unable to find a rotamer that was less repulsive. Furthermore, the orientation of this ligand in mCCR1 was such that this pi-pi stacking could not occur, despite the fact that both residues are conserved in the binding site, with similar side chain placements.

There are a number of other stabilizing interactions in the binding cavity. Namely, the chlorine substituent weakly interacts with the Leu203 (TM5) side chain, while the urea carbonyl hydrogen bonds to the Ala168 (TM3) main chain. It is not clear what the ester moiety is interacting with in the receptor, although the close proximity of the urea group can easily allow for intra-molecular hydrogen bonding. The fluorine substituent also loosely polarizes the Thr286 (TM7) side chain at a distance of 4.64 Å. Finally, 95% of the binding energy is recovered by summing the cavity and solvation energies.

Table 5-2 shows the correlation between our calculated binding energies and the experimental binding affinities across human, mouse, and rat CCR1. The differential between mCCR1 and rCCR1 is 6.94 kcal/mol, and of that Tyr113 (TM3) is 5.30 kcal/mol more favorable in rCCR1. The stabilizing stacking interactions of Phe172 (TM4) and Phe260 (TM6) contribute an additional 0.69 kcal/mol, which in total accounts for 86% of the calculated differential. Addition of the solvation energy (1.00 kcal/mol) returns 100% of the entire difference between mCCR1 and rCCR1.

The differential between hCCR1 and rCCR1 is 10.87 kcal/mol, of which the difference in Tyr113 (TM3) interaction contributes 4.74 kcal/mol. The interaction energy of BX471 with Gly168 (TM4) in hCCR1 is -5.66 kcal/mol, whereas the energy of BX471 with Ala168 in rCCR1 is -2.59 kcal/mol. This could possibly be due to the fact that the interaction is longer range in rCCR1 (3.16 vs. 4.88 Å). Addition of these major contributors and other weak interactions in the binding cavity yields an 8.63 kcal/mol, or 79% of the differential. Including the solvation term returns 91% of the differential between hCCR1 and rCCR1. Based on this information it is not entirely clear which residues in rCCR1 could be mutated to enhance the binding of BX471, as the major contributor (Tyr113) is a conserved residue. Approximately 28% of the differential comes from the Gly168Ala mutation in the rodent species, indicating that this residue could be a suitable candidate for point mutation (i.e., Ala168Gly). This mutant would be expected to increase binding affinity of BX471 to rCCR1.

Conclusions

MembStruk v4.0 was used to build a structural model of rCCR1. Characterization of the apo receptor yielded a fewer number of interhelical contacts when compared to both hCCR1 and mCCR1. HierDock was used to locate the antagonist binding site to which BX471 was docked. Analysis of the binding cavity illustrates the importance of two aromatic residues, Phe172 (TM4) and Phe260 (TM6) in ligand recognition for rodent receptors. Mapping these residues to the mCCR1 binding site shows that they are not in fact contributing comparably to the binding energy, due to the change in ligand orientation. The most important residue in the hCCR1 binding site, Tyr113 (TM3), has large repulsive character in rCCR1, and correspondingly lowers the binding energy. Our ranking of binding energies is entirely consistent with the experimental trend in binding affinities, reflecting additional confidence in our methodology. As a result, this work has provided new insight into the differential binding modes of CCR1 antagonists.

References

[1] Floriano, W. B., Vaidehi, N., Singer, M., Shepherd, G., and Goddard, W. A. III (2000) *Proc. Natl. Acad. Sci. USA* **97**, 10712-10716.

[2] Vaidehi, N., Floriano, W. B., Trabanino, R., Hall, S. E., Freddolino, P., Choi, E. J., Zamanakos, G., and Goddard, W. A. III (2002) *Proc. Natl. Acad. Sci. USA* **99**, 12622-12627.

[3] Trabanino, R. J., Hall, S. E., Vaidehi, N., Floriano, W. B., Kam, V. W. T., and Goddard, W. A. III (2004) *Biophys. J.* **86**, 1904-1921.

[4] Schertler, G. F. X. (1998) *Eye*. **12**, 504-510.

[5] Palczewski, K., Kumasaka, T., Hori, T., Behnke, C., Motoshima, H., Fox, B., Trong, I., Teller, D., Okada, T., Stenkamp, R., Yamamoto, M., and Miyano, M. (2000) *Science* **289**, 739-745.

[6] Vaidehi, N., Floriano, W. B., Trabanino, R., Hall, S. E., Freddolino, P., Choi, E. J., Zamanakos, G., and Goddard, W. A. III (2002) *Proc. Natl. Acad. Sci. USA* **99**, 12622-12627.

[7] Kekenes-Huskey, P. M., Vaidehi, N., Floriano, W. B., and Goddard, W. A. III (2003) *J. Chem. Phys. B* **107**, 11549-11557.

[8] Datta, D., Vaidehi, N., Zhang, D., and Goddard, W. A. III (2004) *Prot. Sci.* **13**, 2693-2705.

[9] Ewing, T. A., and Kuntz, I. D. (1997) *J. Comput. Chem.* **18**, 1175-1189.

Table 5-1: Bends and tilts for each helix in rCCR1.

Helix Number	Bend Angle (degrees)	Tilt Angle (degrees)
1	15.6	33.5
2	12.4	23.5
3	5.6	23.6
4	20.5	22.1
5	7.6	28.1
6	26.9	13.9
7	10.7	26.3

Table 5-2: Correlation between experimental BX471 binding affinities and calculated binding energies.

	Human	**Rat**	**Mouse**
Experimental K_i (nM)	1	100	380
Calculated BE (kcal/mol)	-39.25	-28.38	-21.44

>rCCR1 NP_020542 (LOC57301), mRNA, [Rattus norvegicus]
MEISNITETYPTTTEYDYGDSTPCQKTDVRAFGAGLLPPLYSFVFIIGVVGNILVILVLM
QHRRLQSMTSIYLFNLAVSDLVFLFTLPFWIDYKLKDNWVFGDAMCKLLSGFYYLGLYSE
IFFIILLTIDRYLAIVHAVFSLRARTVTFGIITSIIIWALAILASIPALCFFKAQWEFTH
HTCSPHFPDESLKTWKRFQALKLNLLGLILPLLVMIICYAGIIRILLRRPNEKKAKAVRL
IFAITLLFFLLWTPYNLTVFVSAFQDVLFTNQCEQSKQLDLAIQVTEVIAYTHCCVNPII
YVFVGERFRKYLRQLFQRHVAIPLAKWLPFFSVDQLERTSSLTPSTGEHELSGGF

Figure 5-1: rCCR1 primary sequence in FASTA format.

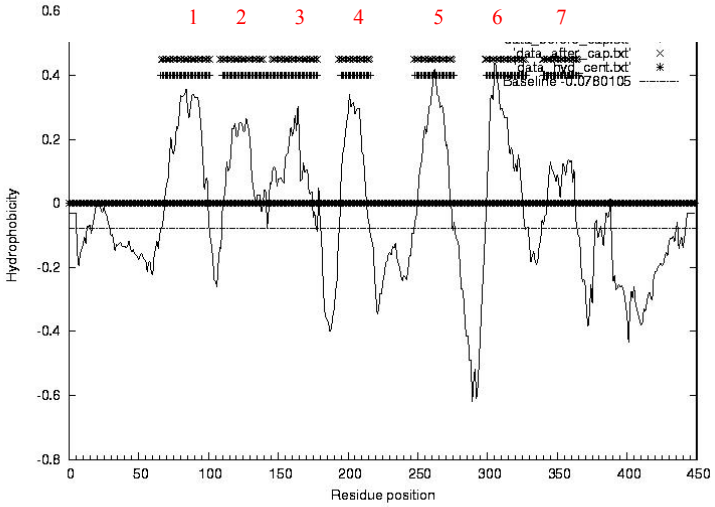


Figure 5-2: rCCR1 hydrophobicity profile.
Indicated in red are TM domains 1-7 at window size 12.

NT:	**1**	MEISNITETYPTTTEYDYGDSTPCQKTDVR	
TM1:	**31**	AFGAGLLPPLYSFVFIIGVVGNILVILVLMQH	**62**
IC1:		RRLQS	
TM2:	**68**	MTSIYLFNLAVSDLVFLFTLPFWIDYKLKDN	**98**
EC1:		WVFGD	
TM3:	**104**	AMCKLLSGFYYLGLYSEIFFIILLTIDRYLAIVH	**137**
IC2:		AVFSLRARTVTF	
TM4:	**150**	GIITSIIIWALAILASIPALCFFK	**173**
EC2:		AQWEFTHHTCSPHFPDESLKTW	
TM5:	**196**	KRFQALKLNLLGLILPLLVMIICYAGIIR	**224**
IC3:		ILLRRPNEKKAKAV	
TM6:	**239**	RLIFAITLLFFLLWTPYNLTVFVSVFQDVL	**268**
EC3:		FTNQCEQSKQ	
TM7:	**279**	LDLAIQVTEVIAYTHCCVNPIIYVFVG	**305**
CT:		ERFRKYLRQLFQRHVAIPLAKWLPFLSVDRLERTSSLTPSTGEHELSGGF	**355**

Figure 5-3: rCCR1 TM regions.
NT = N-terminus, CT = C-terminus, IC = intracellular, and EC = extracellular domains. Highlighted in red are the hydrophobic center predictions for each helix.

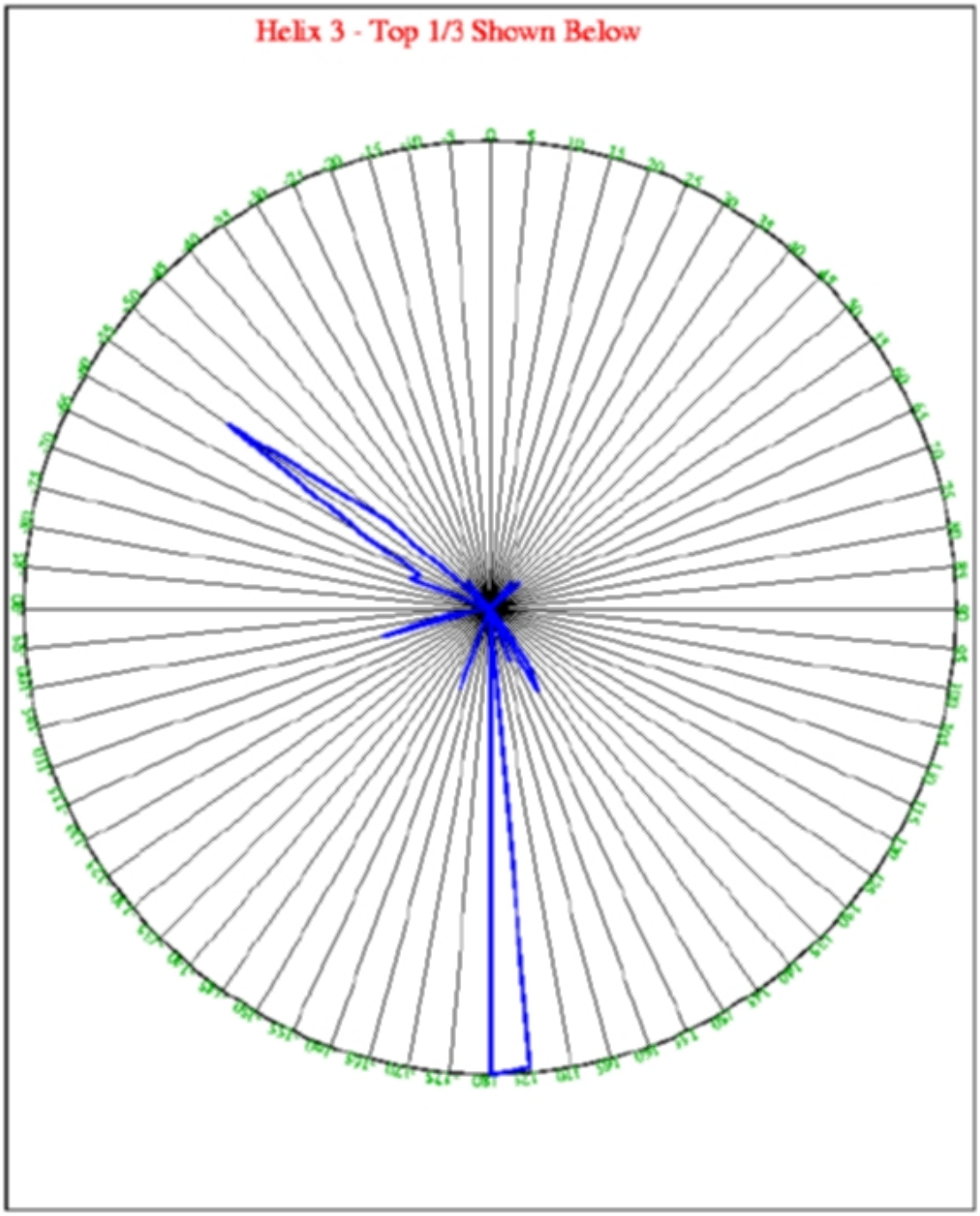


Figure 5-4: rCCR1 TM3 rotational optimization using RotMin.
After conducting a global energy scan for TM3, minima at -55° and 180° are observed. Analysis of each conformation shows that the 180° rotation leads to the formation of favorable polar interactions in the receptor.

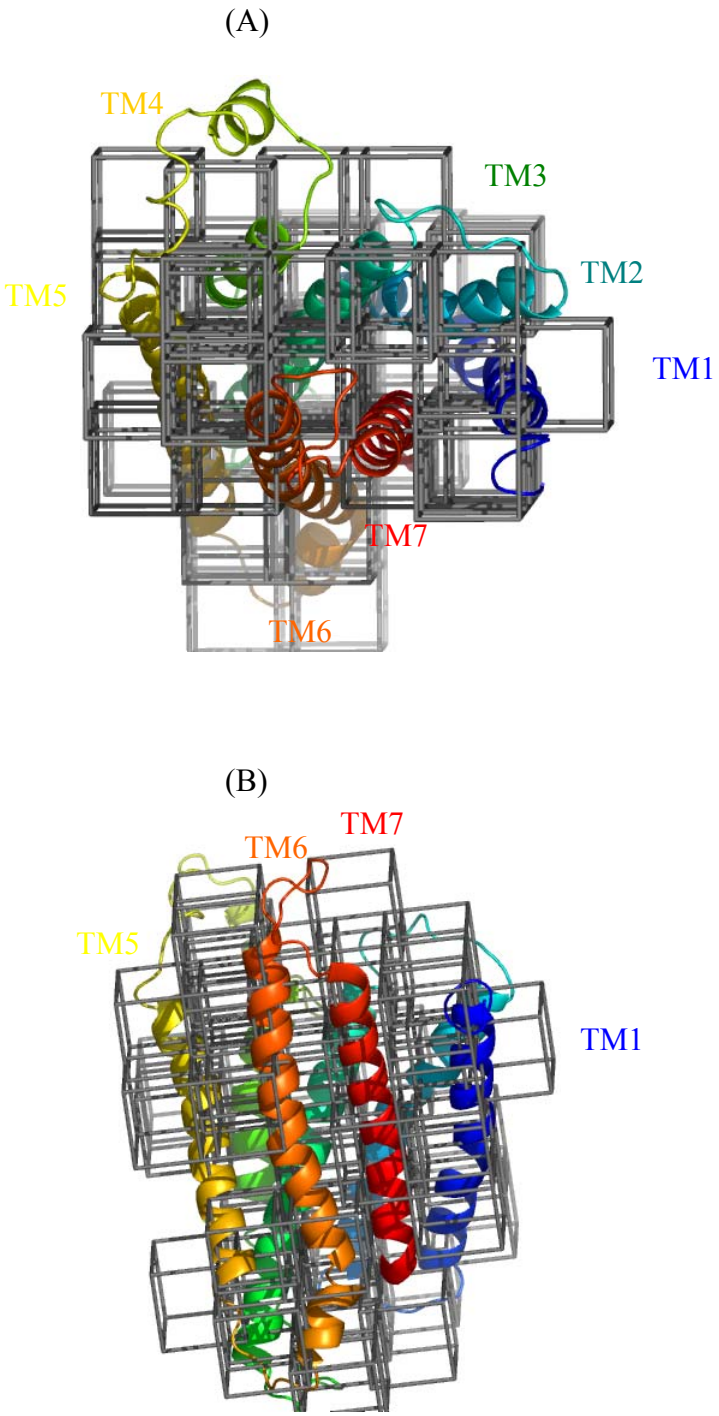
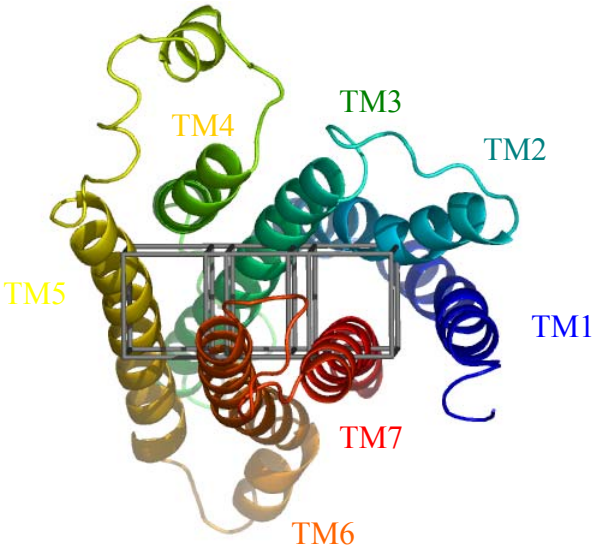


Figure 5-5: rCCR1 putative binding sites.
After scanning the molecular surface of the receptor, 49 boxes (each 1000 $Å^3$) containing a total of 11363 spheres are generated. Shown in (A) is the top view, and (B) the side view. Note the convention used here is the extracellular side on "top" and thus the intracellular interface is below. Also provided is the color code used for helices in the bundle.

(A)



(B)

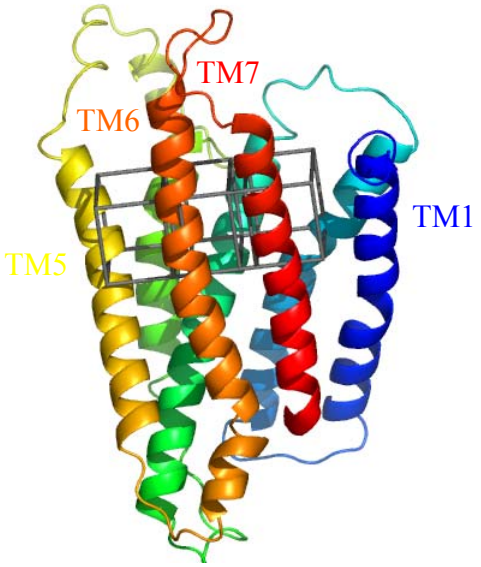


Figure 5-6: rCCR1 binding sites for docking.
From the 49 putative binding sites, we determine our refined binding site to be a horizontal mode across 3 contiguous regions (27, 31, and 32). Shown in (A) is the top view, and (B) the side view.

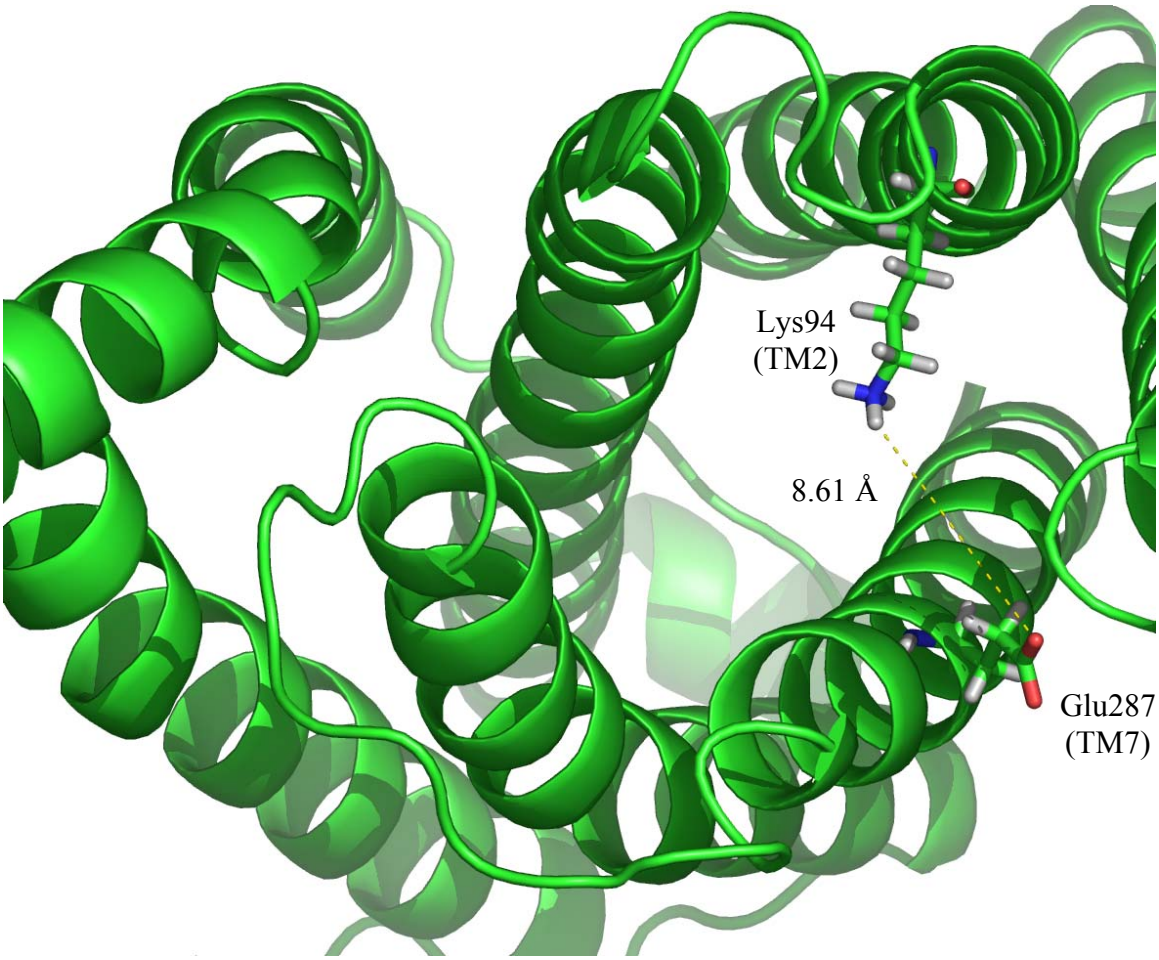


Figure 5-7: rCCR1 salt bridge (TM2-7).
Loose, water mediated salt bridge between Lys94 (TM2) and Glu287 (TM7).

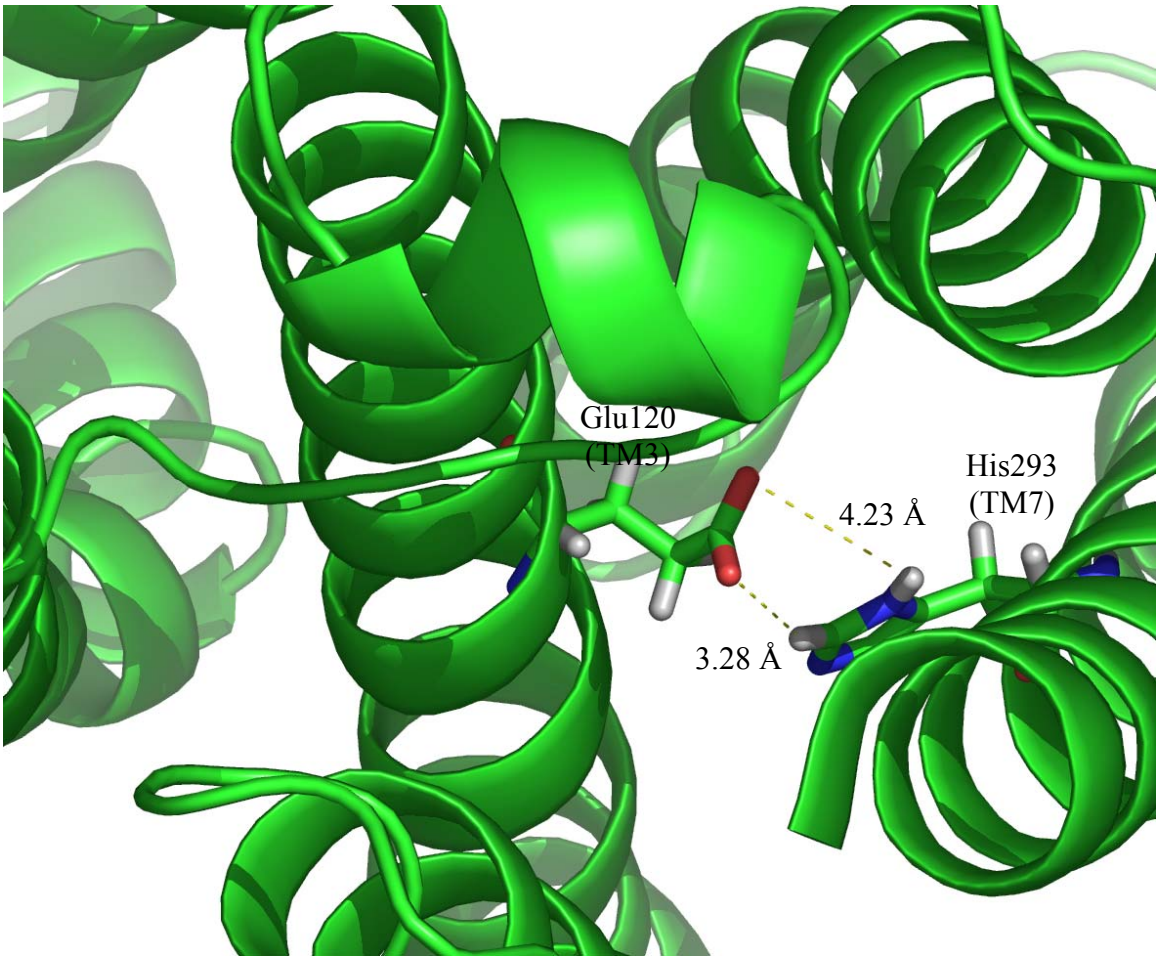


Figure 5-8: rCCR1 salt bridge (TM3-7).
Salt bridge between Glu120 (TM3) and His293 (TM7).

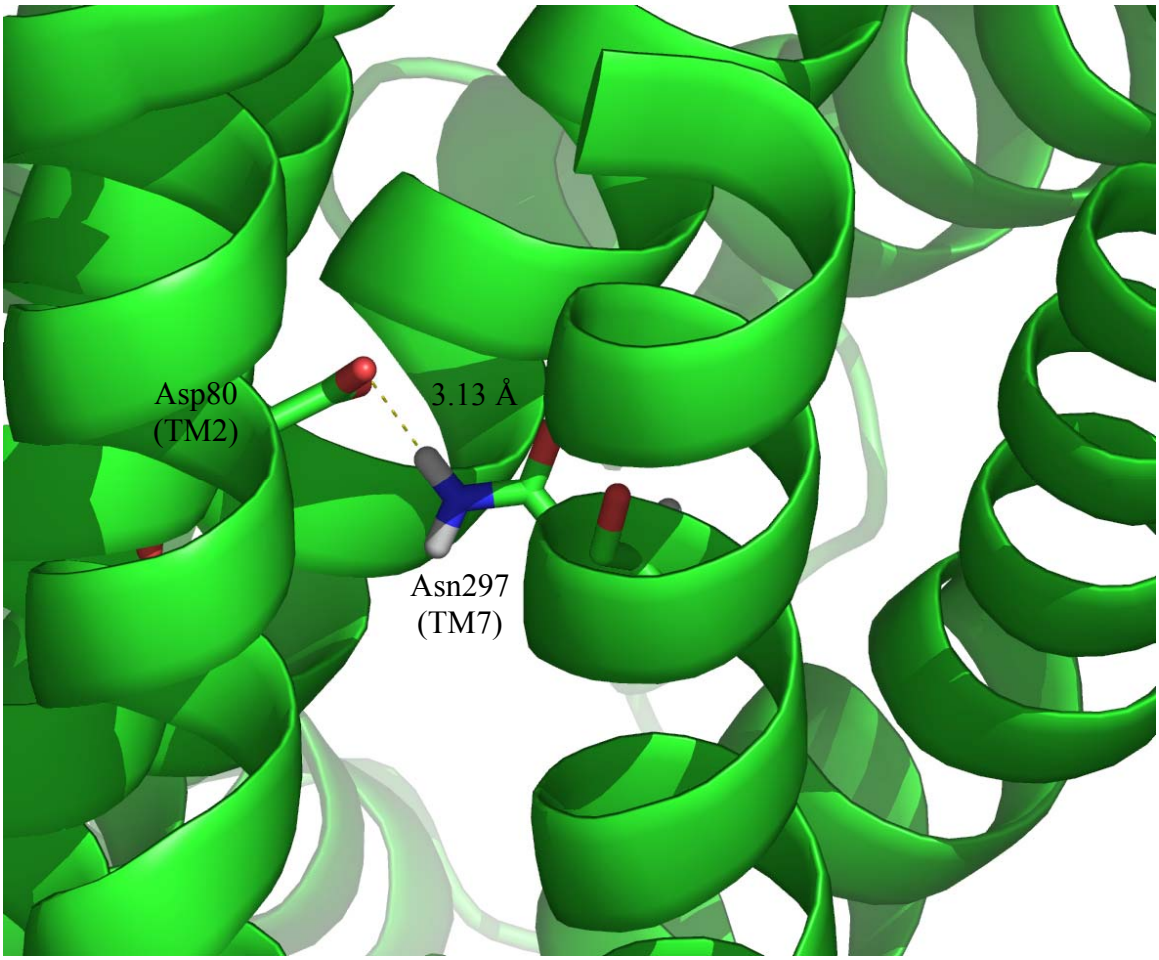


Figure 5-9: rCCR1 inter-helical hydrogen bond (TM2-7).
Hydrogen bond between Asp80 (TM2) and Asn297 (TM7). These are conserved residues across many GPCRs.

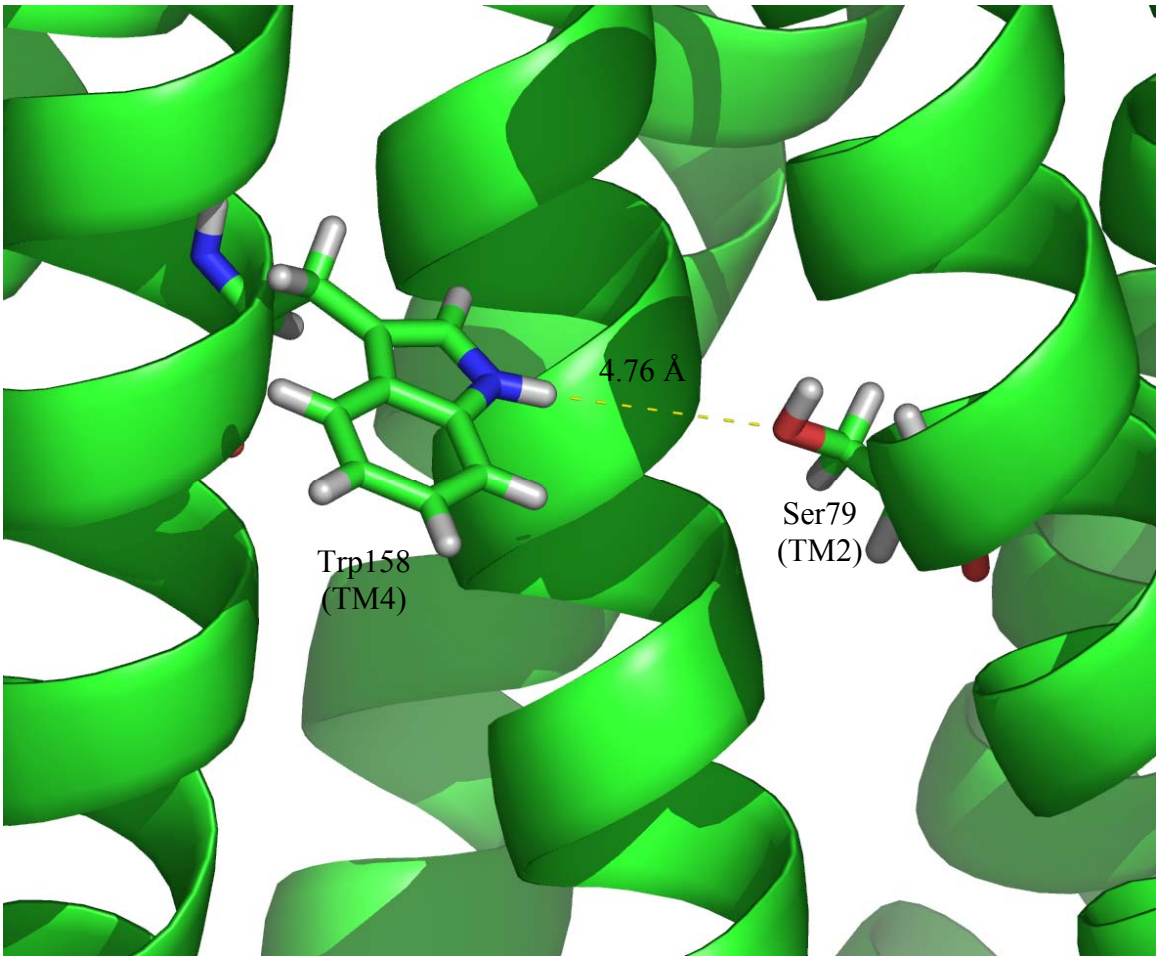
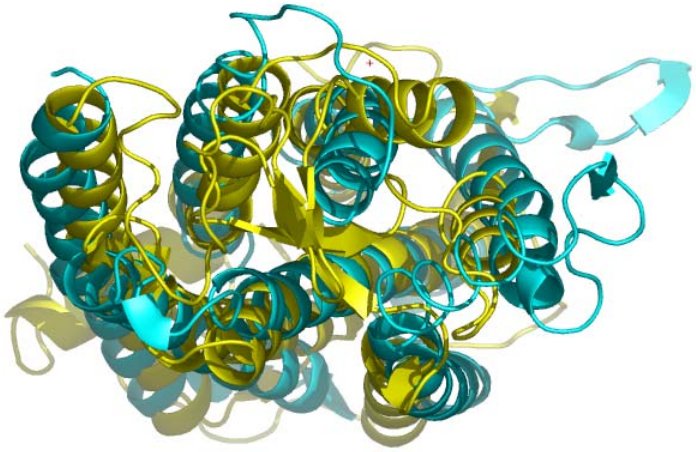
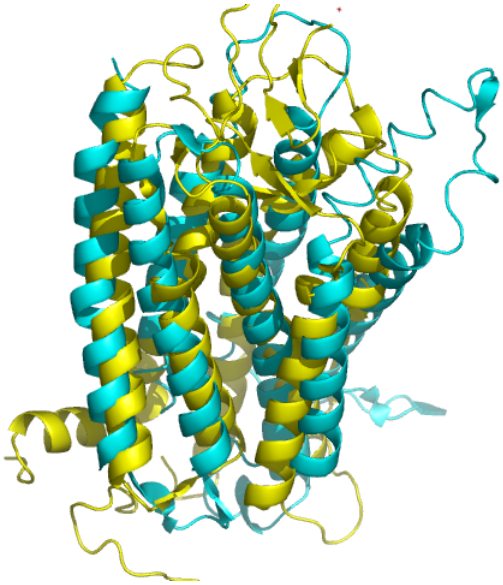


Figure 5-10: rCCR1 inter-helical hydrogen bond (TM2-4).
Hydrogen bond between Ser79 (TM2) and Trp158 (TM4). Trp158 is the most conserved TM4 residue across all GPCRs.

(A)



(B)



rCCR1

BR

Figure 5-11: Overlay of the rCCR1 and BR structures.
Predicted rCCR1 structure superimposed on experimental BR structure (PDB #1F88). The main chain RMSD is 5.91 Å.

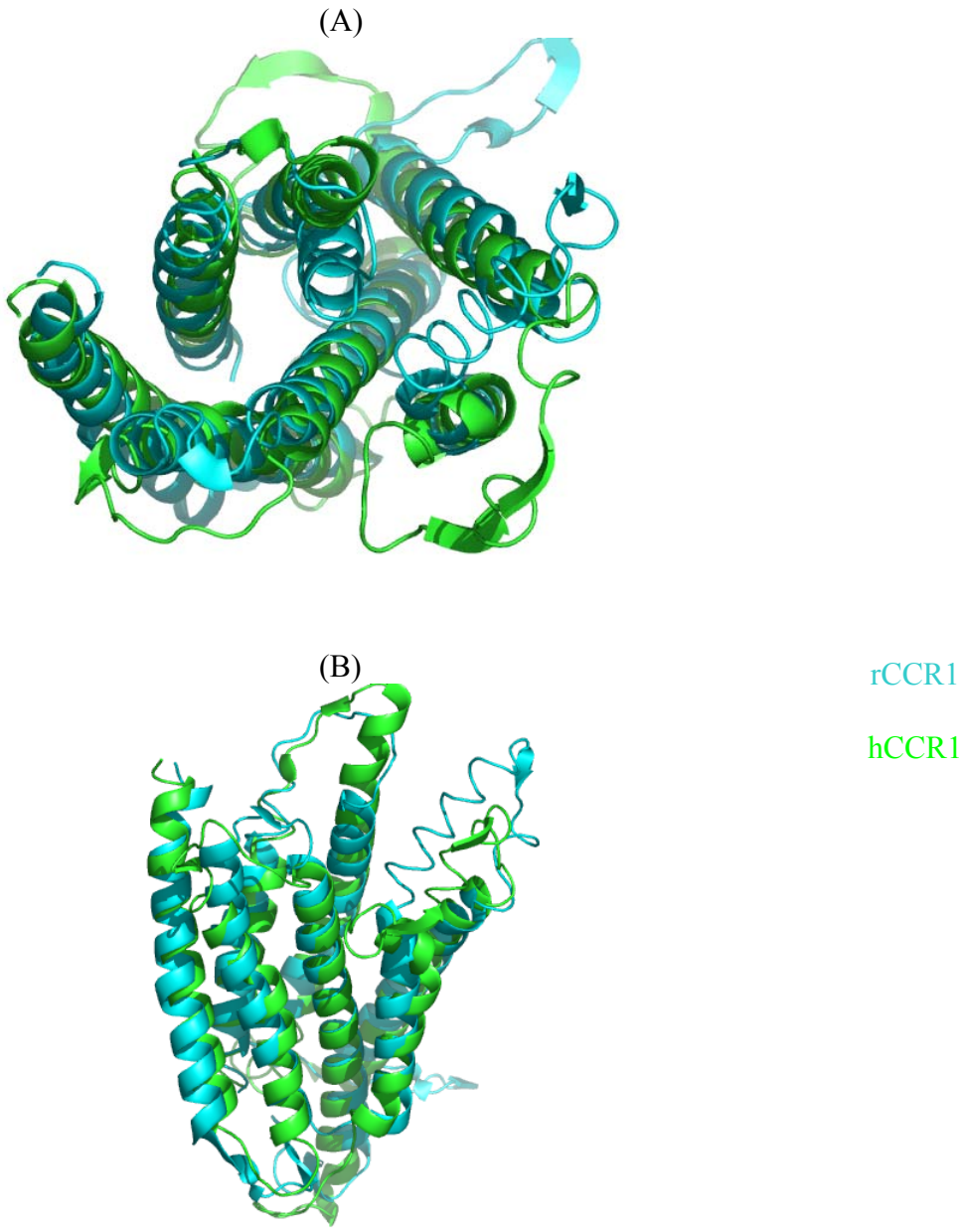


Figure 5-12: Overlay of the rCCR1 and hCCR1 structures.
Predicted rCCR1 structure superimposed on predicted hCCR1 structure. The main chain RMSD is 4.63 Å.

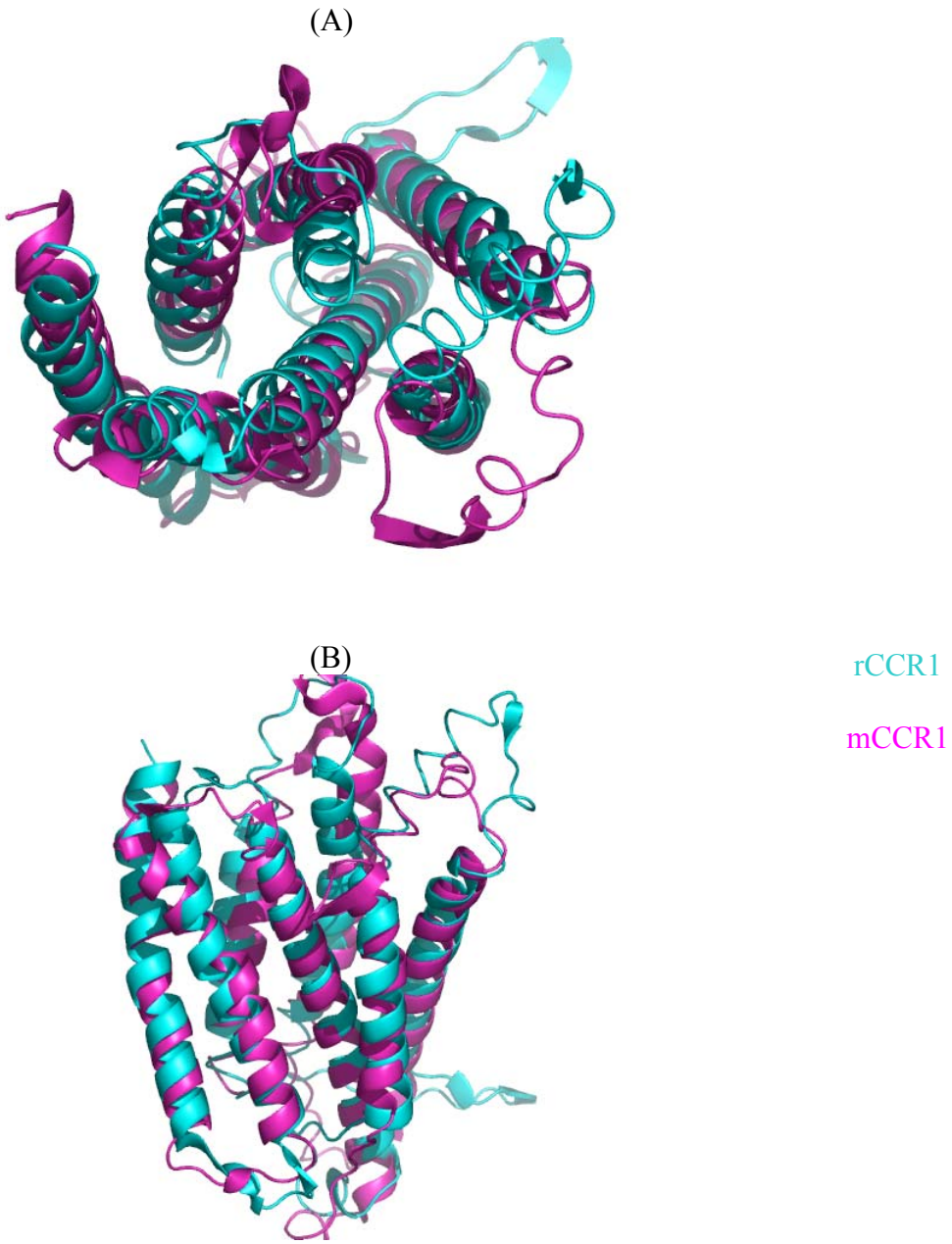


Figure 5-13: Overlay of the rCCR1 and mCCR1 structures.
Predicted rCCR1 structure superimposed on predicted mCCR1 structure. The main chain RMSD is 4.32 Å.

```
CLUSTAL W (1.82) multiple sequence alignment


hCCR1           AFGAQLLPPLYSLVFVIGLVGNILVVLVLVQYMTSIYLLNLAISDLLFLFTLPFWIDYKL 60
rCCR1           AFGAGLLPPLYSFVFIIGVVGNILVILVLMQHMTSIYLFNLAVSDLVFLFTLPFWIDYKL 60
                **** *******:**:**:******:***:*:******:***:***:**************

hCCR1           KDDAMCKILSGFYYTGLYSEIFFIILLTIDRYLAIVHGVITSIIIWALAILASMPGLYFS 120
rCCR1           KDNAMCKLLSGFYYLGLYSEIFFIILLTIDRYLAIVHGIITSIIIWALAILASIPALCFF 120
                **:****:****** ***********************:**************:*.* *

hCCR1           KKLFQALKLNLFGLVLPLLVMIICYTGIIKRLIFVIMIIFFLFWTPYNLTILISVFQD-- 178
rCCR1           KKRFQALKLNLLGLILPLLVMIICYAGIIRRLIFAITLLFFLLWTPYNLTVFVSVFQDVL 180
                ** ********:**:**********:***:****.* ::***:********:::*****

hCCR1           LDLAVQVTEVIAYTHCCVNPVIYAFVG 205
rCCR1           LDLAIQVTEVIAYTHCCVNPIIYVFVG 207
                ****:***************:**.***
```

Figure 5-14: TM sequence alignment of hCCR1 and rCCR1.
Sequence alignment of predicted TM regions of hCCR1 and rCCR1. The identity is 83%; 79% in overall sequence.

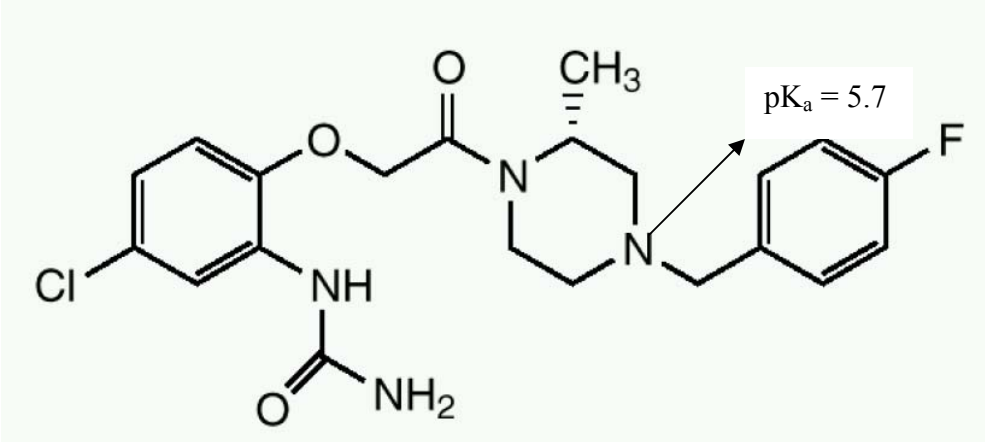


Figure 5-15: Structure of BX471 antagonist.

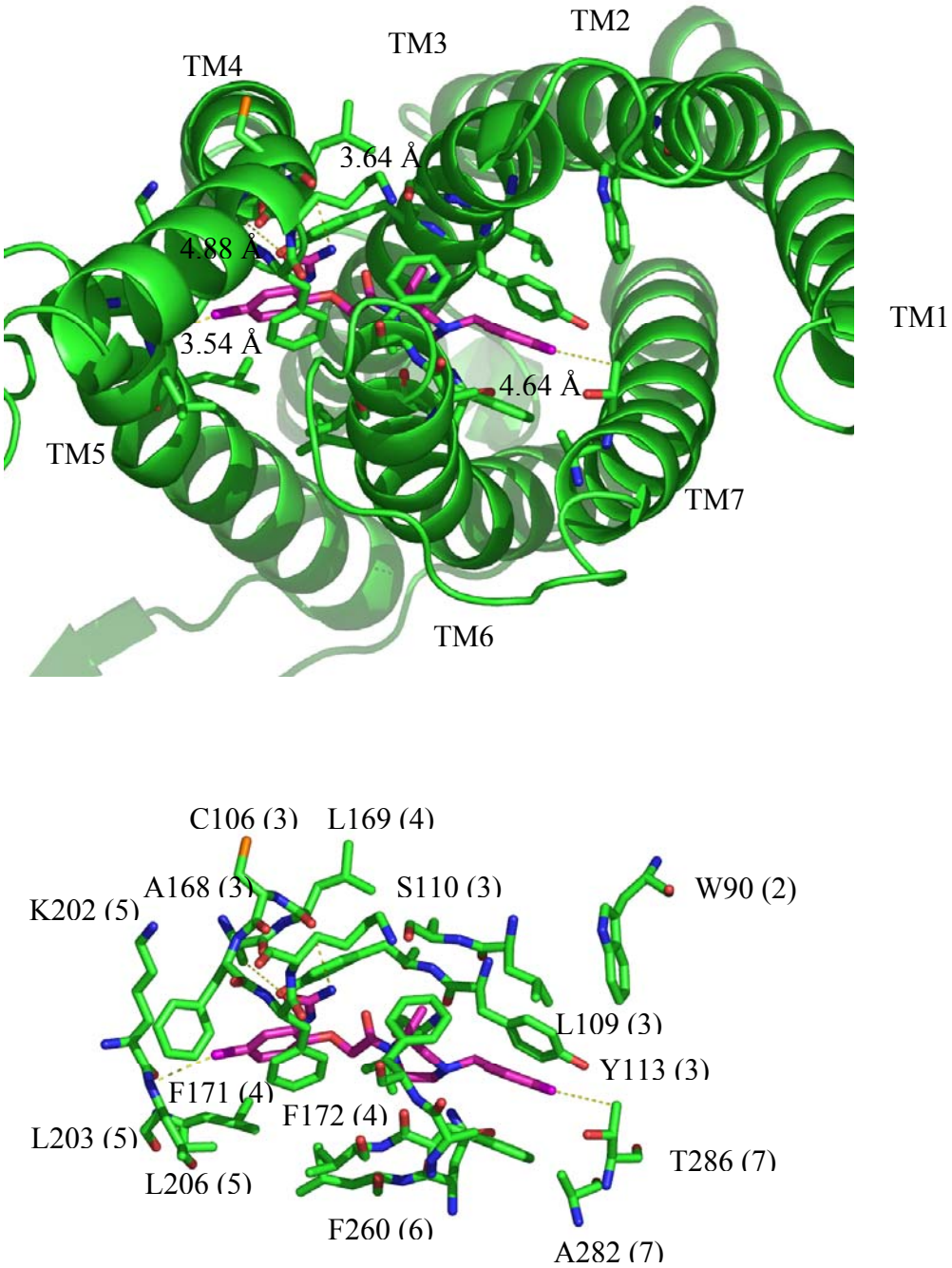


Figure 5-16: Interactions in the 5Å BX471 binding cavity (top).
The ligand is anchored by Ala168, Leu203, Leu169, and Thr286.