

Rowan University

Rowan Digital Works

Theses and Dissertations

6-18-2019

An in silico study of G protein-coupled-receptor activation, specifically in the corticotropin releasing factor receptor and the glucagon-like peptide receptor

Nicolas Angelo Scorese
Rowan University

Follow this and additional works at: <https://rdw.rowan.edu/etd>



Part of the [Bioinformatics Commons](#), and the [Pharmacy and Pharmaceutical Sciences Commons](#)

Recommended Citation

Scorese, Nicolas Angelo, "An in silico study of G protein-coupled-receptor activation, specifically in the corticotropin releasing factor receptor and the glucagon-like peptide receptor" (2019). *Theses and Dissertations*. 2695.

<https://rdw.rowan.edu/etd/2695>

This Thesis is brought to you for free and open access by Rowan Digital Works. It has been accepted for inclusion in Theses and Dissertations by an authorized administrator of Rowan Digital Works. For more information, please contact graduateresearch@rowan.edu.

**AN *IN SILICO* STUDY OF G PROTIEN-COUPLED-RECEPTOR ACTIVATION,
SPECIFICLLY IN THE CORTICOTROPIN RELEASING FACOTR RECEPTOR
AND THE GLUCAGON-LIKE PEPTIDE RECEPTOR**

by

Nicolas A. Scorese

A Thesis

Submitted to the

Department of Molecular and Cellular Biosciences

College of Science and Mathematics

In partial fulfillment of the requirement

For the degree of

Master of Science in Bioinformatics

at

Rowan University

May 24, 2019

Thesis Chair: Chun Wu, Ph.D.

© 2019 Nicolas Angelo Scorese

Acknowledgments

I would like to thank my research advisor and mentor Dr. Chun Wu. He has helped and guided me not only in my research, but all throughout my academic career. His endless dedication in the lab and his numerous accomplishments has inspired me to pursue a career in the field. I would also like to thank Dr. Hickman for his guidance in my undergraduate career, and the members of my committee, Dr. Ben Carone and Dr. Thomas Keck, for their support. A final thank you goes to the members of Dr. Wu's research group.

This work was supported by Rowan Startup and SEED grant and the National Science Foundation under Grant NSF ACI-1429467 and XSEDE MCB177088.

Abstract

Nicolas Scorese

AN *IN SILICO* STUDY OF G PROTEIN-COUPLED-RECEPTOR ACTIVATION,
SPECIFICALLY IN THE CORTICOTROPIN RELEASING FACTOR RECEPTOR AND
THE GLUCAGON-LIKE PEPTIDE RECEPTOR

2018-2019

Chun Wu, Ph.D.

Master of Science in Bioinformatics

The drug discovery process is an extremely long and expensive process that modern computational methods help to alleviate. Through the use of computational methods, we provide information and insight into the activation methods of class B GPCRs so that future drugs can be developed to have less side effects. The first study focuses on the corticotropin releasing factor receptor, which is a good drug target for anxiety and depression. A mechanism of activation was theorized which focuses less on molecular switches (as has been the focus of several papers) and more on large scale conformation at the intracellular region of the receptor and the C-terminal helix. We also developed a homology model for the complete receptor, which previously did not exist. The second study focused on the glucagon-like-peptide receptor which is a good drug target for treating type 2 diabetes. Here we explored the difference between full agonist activation and biased agonist activation. A distinctive conformational change of the C-terminal helix in the biased system was linked to allowing G protein docking, while blocking arrestin proteins from docking. Our findings elucidate details on GPCR activation which can be used to develop more efficient drugs on these receptors and provides insight into developing more specific drugs on other class B GPCRs.

Table of Contents

Abstract.....	iv
List of Figures.....	ix
List of Tables.....	xi
Chapter 1: Introduction to Computational Drug Design.....	1
1.1 The Need for a Computational Approach.....	1
1.2 Methodologies.....	2
1.2.1 Bioinformatics.....	2
1.2.2 Homology Modeling.....	3
1.2.3 Molecular Docking.....	3
1.2.4 Molecular Dynamics Simulations.....	4
1.2.5 MMGBSA Binding Energy Calculations.....	5
1.3 Thesis Outline.....	5
Chapter 2: Binding of Peptide Agonist Urocortin and Small Molecule Antagonist CP 376395 to the CRF1 Receptor Probed by Molecular Dynamics Simulations.....	8
2.1 Introduction.....	8
2.1.1 G Protein-Coupled-Receptors.....	8
2.1.2 The Corticotropin-Releasing-Factor Receptor.....	9
2.1.3 Molecular Switches.....	11
2.1.4 Experimental Overview.....	11
2.2 Methods.....	12
2.2.1 Generating the Homology Model.....	12
2.2.2 Protein and Ligand Preparation.....	12
2.2.3 Docking.....	13

Table of Contents (Continued)

2.2.4 Molecular Dynamics Simulation.....	13
2.2.5 Trajectory clustering analysis.....	13
2.2.6 Simulation Interaction Diagram (SID) Analysis.....	14
2.2.7 Simulation Event Analysis.....	15
2.2.8 Comparison to Other GPCRs.....	15
2.2.9 Matlab Scripting.....	15
2.3 Results.....	16
2.3.1 I-TASSER Generated a Homology Model for CRF1R.....	16
2.3.2 The Antagonist and Agonist Model Produced Different Poses.....	18
2.3.3 Flexibility was Observed in the TMD, ECD, and Helices 8.....	19
2.3.4 Different Types of Protein-Ligand Interactions were Observed.....	22
2.3.5 Different Structural Elements and RMSF Values were Observed....	25
2.3.6 The Radius Gyration Data Showed Shift in the Agonist System....	29
2.3.7 Trajectory Clustering Analysis Revealed Three Molecular Switches.....	29
2.3.8 Comparison to Other GPCRs.....	35
2.4 Discussion.....	37
2.5 Conclusions.....	42
Chapter 3: GLP-1 Receptor in Complex with a Full Agonist and a Biased Agonist Probed by Molecular Dynamics Simulations for the Development of more Specific Drugs on Type 2 Diabetes.....	44
3.1 Introduction.....	44

Table of Contents (Continued)

3.1.1 G Protein-Coupled-Receptors.....	44
3.1.2 The Glucagon-Like-Peptide Receptor.....	45
3.1.3 Full vs. Biased Agonism.....	46
3.1.4 Experimental Overview.....	46
3.2 Methods.....	47
3.2.1 Protein and Ligand Preparation.....	47
3.2.2 Docking.....	47
3.2.3 Molecular Dynamics Simulations.....	49
3.2.4 Trajectory Clustering Analysis	49
3.2.5 Simulation Interaction Diagram (SID) Analysis.....	49
3.2.6 Comparison to Other GPCRs.....	50
3.2.7 MMGBSA Analysis.....	50
3.3 Results	51
3.3.1 Conformational Differences Between Full and Biased Agonist Systems.....	51
3.3.2 Different Types of Protein-Ligand Interactions	54
3.3.3 Secondary Structures Revealed Differences Among the Three Systems.....	56
3.3.4 The Protein RMSF Data Showed Different Fluctuation Levels.....	57
3.3.5 Comparison to Solved GPCRs Provided Insight Into Biased Activation.....	58
3.3.6 The Biased Agonist Bound More Stably Then the Full Agonist.....	62

Table of Contents (Continued)

3.4 Discussion.....	63
3.5 Conclusions.....	68
References.....	70
Appendix A: Binding of Peptide Agonist Urocortin and Small Molecule Antagonist CP 376395 to the CRF1 Receptor Probed by Molecular Dynamics Simulations.....	74
Appendix B: GLP-1 Receptor in Complex with a Full Agonist and a Biased Agonist Probed by Molecular Dynamics Simulations for the Development of more Specific Drugs on Type 2 Diabetes.....	139

List of Figures

Figure	Page
Figure 1. Full-length sequence of Human CRF1R.....	10
Figure 2. Snake diagram of human CRF1R.....	10
Figure 3. Comparison between our homology model and the crystal structures (PDB 3EHU and 4K5Y).....	17
Figure 4. Docked peptide agonist urocortin and small molecule antagonist CP 376395...18	
Figure 5. Root Mean Square Deviation (RMSD).....	20
Figure 6. Peptide agonist and small molecule antagonist most abundant cluster.....	21
Figure 7. Pairwise helical comparison of the transmembrane region.....	22
Figure 8. Ligand-protein interactions for peptide agonist system.....	24
Figure 9. Ligand-protein interaction for small molecule antagonist.....	24
Figure 10. Secondary structure elements (SSE) for agonist and antagonist.....	27
Figure 11. C α RMSF diagram of small molecule and peptide agonist	28
Figure 12. Ligand RMSF for small molecule antagonist.....	28
Figure 13. Radius gyration of peptide agonist and small molecule antagonist.....	29
Figure 14. Polar lock in small molecule and peptide agonist.....	31
Figure 15. Polar lock distances small molecule and peptide agonist for 3 trajectories....	32
Figure 16. Rotamer Toggle switch in peptide agonist and small molecule.....	33
Figure 17. Rotamer toggle switch torsion angles for each residue.....	34
Figure 18. Comparison between our peptide agonist complex and the G protein docked Glucagon-like-peptide 1 receptor.....	35

List of Figures (Continued)

Figure 19. Comparison between our small molecule antagonist complex and the G protein docked Glucagon-like peptide-1 receptor.....	36
Figure 20. Full length sequence of human GLP1R.....	45
Figure 21. Snake plot of full-length human GLP1R	46
Figure 22. Docked full agonist (red) and biased agonist (green).....	48
Figure 23. Root Mean Square Deviation (RMSD).....	52
Figure 24. Most abundant conformations.....	53
Figure 25. Protein-ligand interactions.....	55
Figure 26. Secondary structure elements (SSE).....	56
Figure 27. C α Root Mean Square Fluctuation (RMSF) of all three systems.....	58
Figure 28. Comparison between our full agonist system aligned with the G-protein and the solved GLP1 receptor bound to the G-protein.....	59
Figure 29. Comparison between our biased agonist system aligned with the G-protein and the solved GLP1 receptor bound to the G-protein.....	60
Figure 30. Comparison between our full unbiased agonist system aligned with beta arrestin and the rhodopsin receptor bound to beta arrestin.....	61
Figure 31. Comparison between our biased agonist system aligned with beta arrestin and the rhodopsin receptor bound to beta arrestin.....	62

List of Tables

Table	Page
Table 1. Protein-Ligand Contacts between peptide ligand and CRF1R	23
Table 2. Mean values (\AA) for the RMSF of each portion of the ligand-receptor complexes.....	26
Table 3. Mean values (\AA) for the RMSF of each portion of the ligand-receptor complexes.....	57
Table 4. MMGBSA binding values for full agonist and biased agonist systems.....	63

Chapter 1

Introduction to Computational Drug Design

1.1 The Need for a Computational Approach

Drug development from preclinical research to FDA approval is a long and expensive process, spanning an average of 12 years and over a billion dollars. A large portion of this time is spent in the lab where tens of thousands of drug compounds are created and screened. The California Biomedical Research Association discussed this timeline and claims only about 5 in 5,000 drugs will enter clinical trials (human testing). If you're lucky, then one of these compounds will be approved by the FDA¹.

Creating and screening compounds is the most time consuming portion of drug discovery¹ and can be very expensive. Since the drug discovery process is so labor intensive, computational tools are developed to accelerate the time frame and reduce cost significantly. These tools can quickly predict the binding behavior of a specific target with a library of compounds. Computer-aided drug design (CADD) can follow either structure-based drug design (SBDD) or ligand-based drug design (LBDD). In SBDD, the binding site is identified and used to evaluate ligands and predict their protein-ligand interactions. Ligands can be docked and scored in a virtual screen. In LBDD, the target protein is unavailable and information is taken from a number of ligands. A relevant receptor target can be used to evaluate properties associated with biological activation. Quantitative structure-activity relationships (QSAR) and pharmacophore modeling can be utilized in a virtual screen².

Tools like Maestro and Visual Molecular Dynamics (VMD) can be used to analyze and build high resolution structures to make visualization easier. These tools can be used to refine the crystal structure of proteins and ligands. Refinement consists of a various optimizations including the addition of hydrogen atoms, optimization of hydrogen bonds, assigning proper bond orders, and fixing atomic clashes³. High throughput screening (HTS) can then be used to quickly filter through ligands to determine their binding affinity. Binding energy calculations further narrow down the number of ligands that become lead compounds in the drug development process.

Overall, computational methods are greatly beneficial to the field in two main ways: to more quickly and cheaply determine suitable drug candidates, and to allow for greater visualization of how the ligands interact with the target receptor. This thesis utilizes computational methods to better understand the activation mechanisms of class B G-protein-coupled receptors (GPCRs) bound to peptide ligands. Two studies are performed: the first on the corticotropin-releasing-factor 1 receptor (CRF1R) bound to the peptide agonist urocortin, and the second on the glucagon-like-peptide 1 receptor (GLP1R) bound to both a full agonist and a biased agonist. The findings of these two studies will aid in the development of more specific drugs targeting these receptors.

1.2 Methodologies

1.2.1 Bioinformatics. Bioinformatics can be defined as the conceptualization of biological data using informatics techniques such as computer science, mathematics, and statistics. The predominant analyses that bioinformatics focuses on deals with large datasets associated with the structures of macromolecules, full genomic sequences, and

genomic experimental results. Bioinformatics can also be useful in the drug design process through sequence alignment techniques, homology modeling methods, and large-scale analyses⁴.

1.2.2 Homology modeling. High resolution structures are critical in rational drug design to provide more meaningful results and conclusions. However, there are not always high resolution structures available for a given receptor. When this is the case, homology modeling can be utilized to create a structure based on homologs of the receptor. I-TASSER is one such tool that is used for homology modeling and works by creating a structure from its FASTA file sequence based on templates from the Protein Data Bank⁵ (PDB). PDB is an online database composed of 3D structures of proteins which have been obtained from either NMR or X-ray crystallography experiments. I-TASSER works by comparing an inputted sequence with the sequences found in PDB to generate possible structures. Based on the sequence and the templates from PDB, theoretical models can be built with a higher degree of certainty due to structural similarity of these proteins from an evolutionary standpoint⁶. Maestro also has a useful tool for homology modeling called the Protein Preparation Wizard. This tool checks for inconsistencies in the structure and has the power to edit physiochemical properties such as charge and hydrogen bonding as well as optimize the overall geometry of the complex.

1.2.3 Molecular docking. Molecular docking software is used to determine if a compound can bind to a specific target. There are a number of areas where this technique can be utilized. Virtual screens can utilize molecular docking software to help identify lead compounds which is a much faster process than wet lab testing⁷. Molecular docking software can generate the molecular surface for the receptor based on high resolution

structures from the homology model. The binding site is then marked with potential binding locations represented by spheres. The best scored orientation is outputted as the binding pose⁹⁻¹⁰. There are several types of scoring functions that can be used. One example is a force-field based approach which derives physical-based functional forms from experimental data to estimate binding affinity. Another approach is empirical scoring which simplifies the parameters of force-field based approaches to determine approximately which interactions are more favorable. This technique is consequently less accurate, but substantially faster. A 3rd approach is knowledge-based scoring functions which is based on binding interactions that are known to be more frequent than expected by random distribution. This method has been shown to be faster than force-field based approaches and less prone to over-fitting complications that are associated with empirical approaches¹¹. Maestro's Glide docking tool was used to dock the ligands in this study. Extra precision (XP) was used to dock the small molecule ligands and standard precision-peptide (SP-peptide) was used to dock the peptide ligands.

1.2.4 Molecular dynamics simulations. Molecular dynamics simulations, first developed in the 1970s¹², mimic the way molecules move and behave in a three-dimensional model which can give a deeper understanding of the binding interactions between ligand and receptor. Some common molecular dynamics force-fields include AMBER¹³ and NAMD¹⁴ which only differ by the approach they take in setting up parameters. Properties considered during simulations include atomic diameters, bonds connecting atoms, bond angles, and electric charges. Environmental factors are also considered such as pH, temperature, and pressure. Force vectors are calculated to define the direction and distance of the movement of the molecules in the simulation¹⁵.

Molecular dynamics simulations are run until an equilibrium is reached. A series of snapshots can be produced to create a trajectory of the system as it moves throughout the simulation to make analysis easier. Once complete, this opens up the doors to further analysis on the complex to better understand the binding energies and interactions of the system.

1.2.5 MMGBSA binding energy calculations. Molecular mechanics generalized Born surface area (MMGBSA¹⁶) can be used to calculate binding energies for a given system. The binding free energy is calculated using the following equation: $\Delta G_{\text{bind}} = G_c - (G_p + G_L)$. G_c represents the protein-ligand complex's free energy, G_p is the free energy of the protein, and G_L is the free energy of the ligand. These binding energies can provide important details about how a given ligand interacts with its receptor which can be used to better understand its activation mechanism. MMGBSA analysis considers three main components: electrostatic, van der Waals, and surface area¹⁷. This analysis is performed on ligand only, receptor only, and receptor-ligand complexes to determine total binding energy. The results allow us to determine how stable a given ligand can bind to a given receptor. This technique has uses in the drug development process to help screen for compounds that have the highest binding stability.

1.3 Thesis Outline

In chapter 2, the binding of a peptide agonist and a small molecule antagonist to the corticotropin-releasing factor receptor type 1 (CRF1R) is analyzed. The CRF receptor is a Class B G protein-coupled receptor (GPCR) that is activated by a peptide hormone (CRF) for stress responses. Although CRF1R is a good drug target for treating depression, inflammation, and anxiety, there is no FDA approved drug. There is also no high

resolution structure of CRF1R in complex with the peptide hormone and its activation mechanism remains elusive. In this study, we explore conformational changes of CRF1R in complex with a peptide agonist (Urocortin) and small molecule antagonist (CP 376395). The structure model of CRF1R was constructed using I-TASSER which generated several models based on the top ten PDB hits. The model selected prioritized PDB 4K5Y because it represented inactivated CRF1R. Urocortin (PDB 2RMG) and CP 376395 were docked to the receptor. The docked systems were subjected to a total of 4 μ S ($2 \times 2 \mu$ S) molecular dynamics simulations, followed by trajectory clustering and simulation interaction diagram analysis. We examined three molecular switches (Ionic lock, Polar lock, and Rotamer Toggle switch) that were thought to play key roles in the activation of the receptor and show different conformations between peptide and small molecule systems. The extracellular loop 3 (EL3) and helix 8 also showed high flexibility between agonist and antagonist conformations. Finally, we compared our receptor conformations to existing GPCRs docked to the G-Protein to predict activation. Our findings point toward a new mechanism of activation favoring large scale conformational changes of the C-terminal helix and transmembrane helices as opposed to molecular switches which have been the focus of several previous studies. Our findings also provide insights for developing this class of drugs on CRF1R.

In chapter 3, the activation mechanism of a biased agonist is analyzed in the glucagon-like-peptide-1 receptor (GLP1R). The GLP1 receptor is a class B G protein-coupled-receptor (GPCR) which controls insulin secretion. For this reason, it is a good drug target for type 2 diabetes (the condition in which the body cannot produce enough insulin). Agonist drugs can be used to stimulate insulin secretion, but are typically not

specific to the G protein pathway which can lead to side effects. Biased agonists are specific to one particular pathway, but their activation mechanism is poorly understood. In this study, we aim to develop a better understanding of biased activation so that future drugs can be designed to be more specific and have less side effects. We propose that key conformational changes in the C-terminal helix of the biased system allow for binding of the G protein, but block binding of arrestin proteins. The structure of the GLP1 receptor was obtained from the Protein Data Bank (PDB 6b3j). The full agonist was obtained from PDB 5nx2 and the biased agonist was obtained from PDB 6b3j. A full agonist, biased agonist and APO form system was prepared and run under MD simulations for 2 μ s each. Trajectory clustering analysis, simulation interaction diagram analysis, and MMGBSA analysis was performed. Comparisons were made to G protein docked and arrestin docked structures obtained from PDB to determine activation. It was concluded that the C-terminal helical conformation of the biased system was responsible to the specificity toward the G protein pathway. Our results provide insight into developing a new, more efficient class of drugs on the GLP1 receptor.

Chapter 2

Binding of Peptide Agonist Urocortin and Small Molecule Antagonist CP 376395 to the CRF1 Receptor Probed by Molecular Dynamics Simulations

2.1 Introduction

2.1.1 G protein-coupled-receptors. G protein-coupled-receptors (GPCRs) are known for being the largest family of cell surface receptors. They share a structural similarity with each other in that they all have seven transmembrane helices connected by alternating intracellular and extracellular loops, an extracellular N-terminus, and an intracellular C-terminus. In order for a GPCR to enter a state of activation, it must undergo a molecular switch. A molecular switch can be defined as non-covalent intramolecular interactions that have to be disturbed in order for activation to occur¹⁸. GPCRs can be broken up into five distinct classes: Class A (rhodopsin), Class B (secretin), Class C (glutamate), Class D (adhesion), and Class E (frizzled)¹⁹. They are important drug targets due to the way they initiate signaling cascades which spread throughout the body. Class A GPCRs are the largest and most understood class of GPCR (containing over 700 receptors). Class B GPCRs are less understood, which is why they were chosen as the focus for this paper. Class B GPCRs are distinguished by their two domains: a large extracellular domain (ECD) which plays an important role in activation, and a helical seven transmembrane domain (TMD). Altogether, there are 15 known receptors in the Class B family that are important drug targets for diabetes, osteoporosis, hypercalcemia and more. Unlike Class A receptors, there is little known about the activation mechanism of Class B GPCRs beyond a general binding mechanism for peptide hormones²⁰. When binding, the C-terminal end of the peptide attaches to the

extracellular domain initiating a conformational change that allows the N-terminal end of the peptide to bind inside the seven transmembrane pocket. This conformational change allows for interaction with the G protein¹⁹.

2.1.2 The corticotropin-releasing-factor receptor. The corticotropin-releasing factor receptor type 1 (CRF1R) is a Class B GPCR that controls how the body responds to stress and is predominantly found in the central nervous system²¹. For this reason, CRF1R is a good drug target for things such as anxiety, depression, inflammation, and other stress related issues²². While the activation mechanism remains largely unknown, Seidal²⁰ proposes that agonists adopt different folds than antagonists to stabilize the transmembrane domain. The agonists produce a wider pose than antagonists, which affects the activation of the receptor. These differences occur at helices VI and VII and involve a bending around the glycine hinges. Seidal termed these two distinct poses as ‘wide’ (for agonists) and ‘compact’ (for antagonists). Seidal also confirmed that both agonist and antagonist models maintained integrity and flexibility during the simulation.

N-term					
MGGHPQLRLV	KALLLLGLNP	VSASLQDQHC	ESLSLASNIS	GLQCNASVDL	IGTCWPRSPA
10	20	30	40	50	60
				TM1	
GQLVVRPCPA	FFYGVRYNTT	NNGYRECLAN	GSWAARVNYS	ECQEILNEEK	KSKVHYHVAV
70	80	90	100	110	120
		ICL1	TM2		ECL1
IINYLGHGIS	LVALLVAFVL	FLRLRSIRCL	RNIHWNLIS	AFILRNATWF	VVQLTMSPEV
130	140	150	160	170	180
TM3				ICL2 TM4	
HQSNVGCWRL	VTAAYNYFHV	TNFFWMFGE	CYLHTAIVLT	YSTDRLRKWM	FICIGWGVPF
190	200	210	220	230	240
	ECL2	TM5			
PIIWAIGK	LYYDNEKWF	GKRPVYTDY	IYQGPMLVL	LINFILFNI	VRILMTKLRA
250	260	270	280	290	300
ICL3 TM6			ECL3	TM7	
STTSETIQYR	KAVKATLVLL	PLLGITYMLF	FVNPGEDEV	RVVFIYFNSF	LESFQGFVVS
310	320	330	340	350	360
H8			C-term		
VFYCFLNSEV	RSAIRKRWHR	WQDKHSIRAR	VARAMSIPTS	PTRVSFHSIK	Q S T A V
370	380	390	400	410	

Figure 1. Full-length sequence of Human CRF1R taken from GPCRdb²³⁻²⁴.

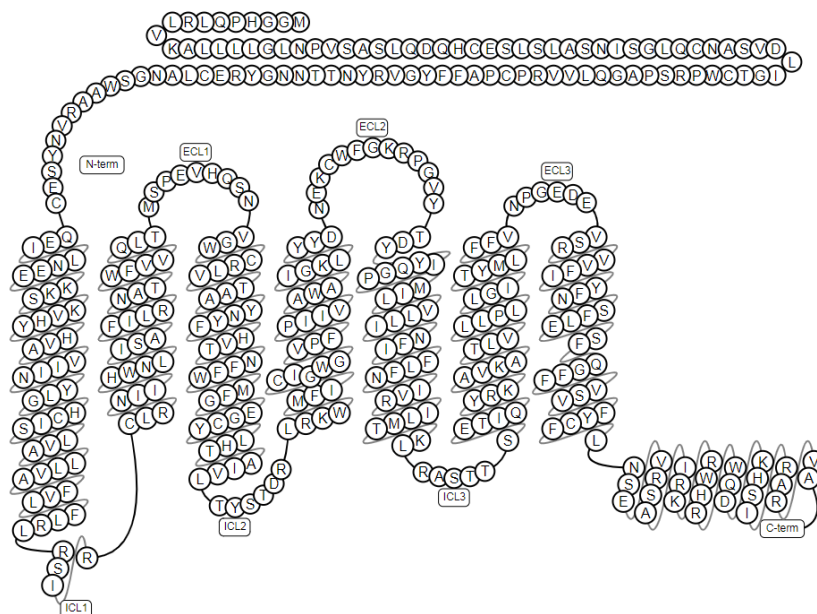


Figure 2. Snake diagram of human CRF1R taken from GPCRdb²³⁻²⁴. Generic numbering of this receptor can be found in Appendix A, table 5.

2.1.3 Molecular switches. We also examined three molecular switches and a conserved sequence motif found in CRF1R. First, the Ionic lock which is an interaction located between Arg151 on TM3 and Glu209 on TM3. The Polar lock is also located between TM2 and TM3, but is an interaction that occurs between His155 and Glu209 (see appendix A, table 5 for generic numbering). The Rotamer Toggle switch is located between TM5 and TM6 at residues Tyr327, Leu323, and Phe284. GWGxP is a conserved sequence motif found among all class B GPCRs that we examine for structure/rotational differences between agonist and antagonist bound systems. Xu²⁵ examined conformational differences at these locations between different small molecule antagonists and apo-form CRF1R. Xu concluded that the ionic lock is broken during antagonist binding, but remains formed in the apo form. Xu found that the Polar lock was unaffected by the binding of antagonists. Xu noted a distinct rotational difference of the Rotamer Toggle switch between apo-CRF1R and antagonist-bound CRF1R. Finally, Xu noted that the GWGxP motif plays no role in the inactivation of CRF1R. In this paper, we aim to examine these locations on a peptide agonist and small molecule antagonist system for conformational differences.

2.1.4 Experimental overview. In this study, we explore how these conformational changes affect the activation of CRF1R with a peptide agonist (Urocortin) and small molecule antagonist (CP 376395). We aim to expand the knowledge on class B GPCR activation mechanisms through examining the binding poses with the above mentioned ligands under molecular dynamics simulations. The PDB database does not have the complete structure of CRF1R, but does have its transmembrane domain and extracellular domain individually. The full structure

(obtained from I-TASSER) was compared to crystal structures from PDB IDs 4K5Y and 3EHU respectively so that the ligands could be docked. The peptide ligand was also obtained from PDB (2RMG). The small molecule antagonist CP 376395 was obtained from the 4K5Y system. Ligands were docked in Maestro and molecular dynamics simulation, Trajectory Clustering Analysis, and Simulation Interaction Diagram (SID) analyses were performed for both systems. Our systems were compared to other class B GPCRs bound to the G protein for comparison.

2.2 Methods

2.2.1 Generating the homology model: The sequence for CRF1R was downloaded and I-TASSER²⁶ was used to create several models based on the top ten PDB hits. The model selected prioritized PDB 4K5Y because it represented the transmembrane domain of inactivated CRF1R. This model was then opened in Maestro for further preparation.

2.2.2 Protein and ligand preparation: Maestro's Protein Preparation Wizard²⁷ was used to prepare the CRF1R model for docking and simulations. Preprocessing was performed on the protein which corrected the bond orders, added hydrogens and disulfide bonds where necessary, and removed water where appropriate. The model then had its charge state optimized followed by a restrained minimization. A 3D structure for the peptide agonist (Urocortin) was obtained from the protein data bank (PDB ID 2RMG). The crystal structure of the small molecule antagonist (CP 376395) was obtained from the transmembrane structure (PDB ID 4K5Y). The same preprocessing, optimization, and minimization as before was performed on these ligands.

2.2.3 Docking: The docking of the peptide ligands to the receptor was done in Maestro using Glide Docking. First the receptor grid was generated, then the ligands and receptors were specified. Default settings were used under the SP-Peptide mode to dock the prepared peptide ligand using an OPLS3 force field²⁸⁻³⁰. The docked complex was then loaded into VMD to optimize the binding pose. The C-terminal end of the peptide ligand was oriented to bind to the ECD in correspondence with Grace et al.³¹⁻³² who suggested that, due to the hydrophobic nature of the region, the C-terminal end expresses receptor specific binding to the ECD. The small molecule ligand was already in complex with the receptor in the 4K5Y structure.

2.2.4 Molecular dynamics simulation: The receptor-ligand complexes from the docking step were used to construct molecular dynamic simulation systems. The complexes were aligned in a membrane set to the helices of the transmembrane domain. The complexes were then solvated using a water box with a predefined SPC water model. To help neutralize the system, 0.15M NaCl was added to the system. An OPLS3 force field was used to build these systems using the Desmond System Builder in Maestro. The complexes ran for a total of 4 μ S under these conditions.

2.2.5 Trajectory clustering analysis: The Desmond³³ trajectory tool in maestro was used to group the structures that were produced during the final 100 ns of the simulation. Backbone RMSD was selected as a structural similarity metric and hierarchical clustering was selected with average linkage. The merging distance cutoff was set to 2.5 Å. Frequency was set to 2 with the number of frames set to 250 and number of threads set to 2. This calculation was performed for both agonist and antagonist structures. Structures of the most abundant cluster (those with a frequency of

2% or greater) were saved for further analysis. The most abundant structure for each complex was then aligned and superimposed to compare conformation differences. The three regions to compare were the extracellular domain, the transmembrane bundle, and helices 8.

2.2.6 Simulation interaction diagram (SID) analysis: The SID tool can be found in maestro under the tasks menu and was used to analyze the interactions between the protein and ligand. Calculations performed include: Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF), secondary structure changes, and protein-ligand contacts. The RMSD calculation measures the displacement change of atoms for the entire trajectory with respect to the reference frame. The RMSD equation used for this calculation is as follows:

$$RMSD_x = \sqrt{\frac{1}{N} \sum_{i=1}^N (r_i'(t_x) - r_i(t_{ref}))^2}$$

Where N is the number of atoms, t is time, and r is the position of the atom. RMSF analyzes changes along a protein chain or molecule. The RMSF equation used for this calculation is as follows:

$$RMSF_i = \sqrt{\frac{1}{T} \sum_{t=1}^T \langle (r_i'(t) - r_i(t_{ref}))^2 \rangle}$$

Where T is the trajectory, t is time, r is residue position and the < and > signs indicate that the average of the square distance is taken. The secondary structure changes are monitored for the entirety of the simulation. Structures recorded include alpha-helices and beta strands. The simulation also records how many protein-ligand contacts exist and

of what kind (such as hydrophobic, H-bond, etc.). For the peptide complex, a protein only analysis was performed since the program could not discern the difference between the protein receptor and the peptide ligand. Its protein contacts were determined using Protein Interaction Analysis under the biologics tasks menu. The resulting data was used to make a table of interactions which was used to make a contact map in VMD.

2.2.7 Simulation event analysis: The SID analysis calculated total RMSD for the entire complex. To calculate select regions (i.e. the extra cellular domain, transmembrane domain, and helices 8) the systems were aligned in maestro with their respective starting positions based solely on their transmembrane domain. Maestro's simulation event analysis tool was then used to calculate RMSD of each region separately for the entire trajectory. Radius of gyration was also used on select regions.

2.2.8 Comparison to other GPCRs: Five different PDB entries (3sn6³⁴, 5g53³⁵, 5vai³⁶, 5uz7³⁷, 6b3j³⁸) were aligned based on their transmembrane domain and then superimposed on our agonist and antagonist systems to look for key conformational differences associated with G protein docking. The key region associated with the ability for the G-protein to dock to the receptor is the C-terminal helix 8. Figures were generated comparing the C-terminal region of all five PDB entries with our agonist and antagonist systems.

2.2.9 Matlab scripting: Scripts were developed in Matlab for the purpose of generating graphs for RMSD data, Radius of Gyration, and molecular switches. The RMSD script took the RMSD of the ligand, transmembrane, N-terminal, and C-terminal domains obtained from the simulation event analysis data and plotted them all together to compare these regions between the agonist and antagonist systems. The radius of

gyration script generated two histograms comparing the N-terminal regions in one graph, and the C-terminal region in the other. The script for the molecular switches was developed to analyze all clusters in the simulation, not just the most abundant. The script worked by loading in the trajectories of both complexes, averaging the values to smooth out the data, and creating a graph of distance (Å) versus time (ns) in the case of figures 13 and 15 or torsion angle (degrees) versus time (ns) in the case of figures 17-19. These graphs aided in showing the conformation of key residues over the entire simulation.

2.3 Results

2.3.1 I-TASSER generated a homology model for CRF1R. Crystal structures exist for the transmembrane and extracellular region individually, but not together (PDB ID 3EHU and 4K5Y). **Figure 3** compares both crystal structures with the TASSER model. The small molecule ligand is already docked in the crystal structure and aided in justifying our docking position of the ligand into our model. Docking orientation of the peptide ligand was justified by comparing to GLP1 in **figure S3 and S4**.

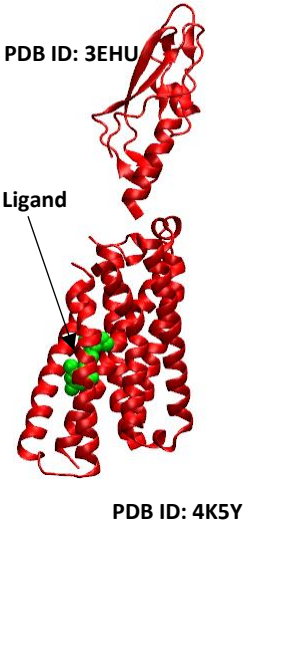
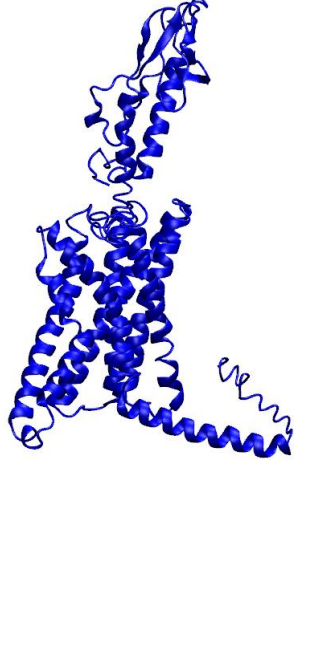
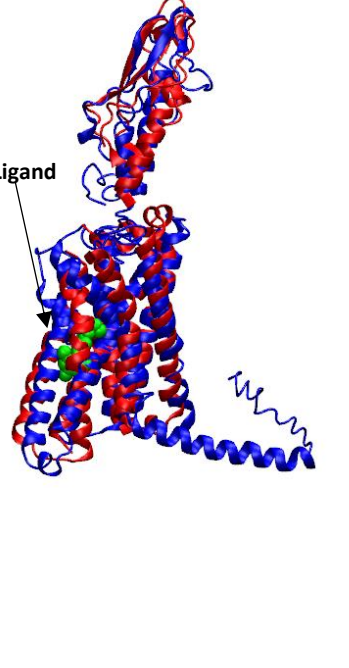
Crystal Structure	TASSER Model	Superimposition
 <p>PDB ID: 3EHU</p> <p>Ligand</p> <p>PDB ID: 4K5Y</p>		 <p>Ligand</p>

Figure 3. Comparison between our homology model and the crystal structures (PDB 3EHU and 4K5Y).

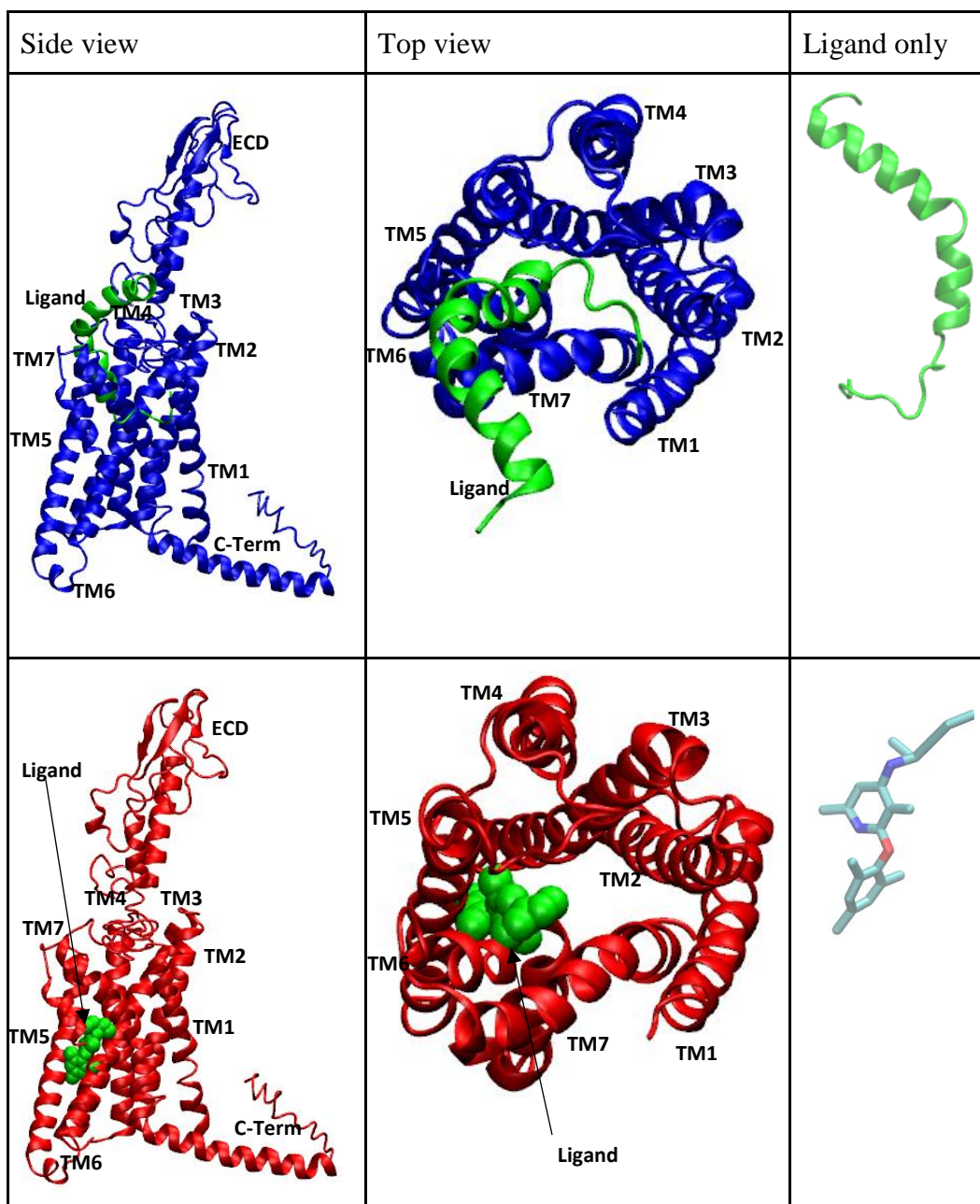


Figure 4. Docked peptide agonist urocortin and small molecule antagonist CP 376395 to the I-TASSER generated model.

2.3.2 The antagonist model and the agonist model produced different poses.

Seidel²⁰ proposed distinct conformational differences of helices 6 and 7 between agonist and antagonist systems. Agonist should produce a “wide” receptor pose and antagonists should produce a “compact” receptor pose. **Figure S6** highlights these two helices and

reveals that indeed the antagonist model produces the aforementioned compact pose and the agonist model produces the wide pose. We compared our docking position to Jazayeri et. al.³⁹ who docked a peptide ligand into another GPCR, GLP1 (Figure S3 and S4).

2.3.3 Flexibility was observed in the TMD, ECD, and helices 8. In the transmembrane domain, two distinct groups are seen. In one group, TMs 2, 3 and 4 retain their position in both systems and develop kinks in the agonist. The other group contain helices 1 and 5-7. In this group, the kinking is similar in both systems, but the helices are significantly shifted in the small molecule. Most helices showed a large conformational shift (**Figure 7**). The RMSD data from this region showed a great deal of difference between agonist and antagonist systems (**Figure 5**). The extracellular domain also showed a great deal of flexibility. The agonist conformation showed the extracellular domain adopt an open conformation by folding to the right, whereas the antagonist conformation adopted a more vertical conformation (**Figure 6**). The RMSD data shows more flexibility in the agonist system than the antagonist (**Figure 5**). Helices 8 showed a different conformation for each systems: the peptide agonist tilted upward, the small molecule antagonist tilted downward (**Figure 6**).

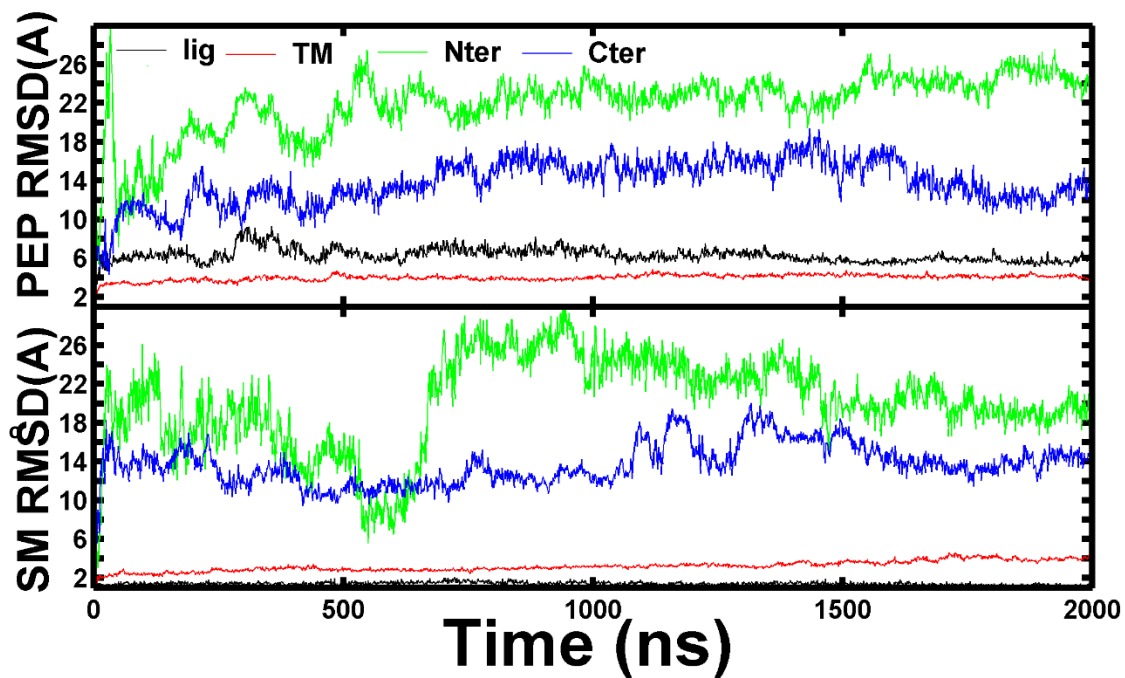


Figure 5. Root Mean Square Deviation (RMSD). Used to measure the average change in displacement of a selection of atoms for a particular frame with respect to a reference frame. It is calculated for all frames in the trajectory.

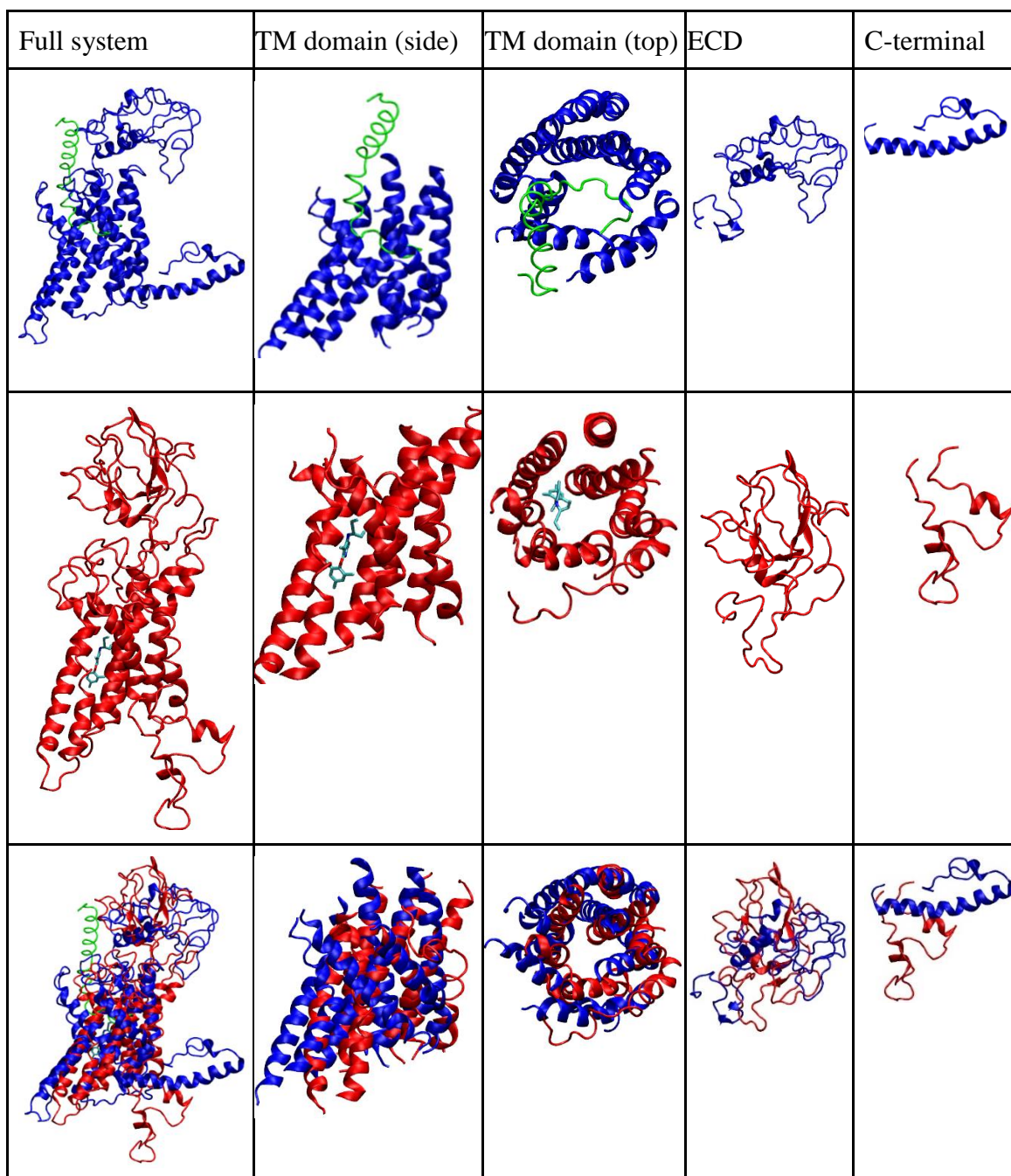


Figure 6. Peptide agonist and small molecule antagonist most abundant conformation for full system, 7TM domain, extracellular domain, and the C-terminal Helix 8 after **2000ns**. Blue = agonist protein conformation, Red = antagonist protein conformation, Green = peptide ligand.

All Helices		TM 2-4		TM 1, 5-7		
TM 1 and 7		TM 2 and 3		TM 4 and 5	TM 5 and 6	
TM1	TM2	TM3	TM4	TM5	TM6	TM7

Figure 7. Pairwise helical comparison of the transmembrane region. Images taken from the most abundant conformations between agonist and antagonist systems. Blue = peptide agonist, red = small molecule antagonist.

2.3.4 Different types of protein-ligand interactions were observed. The peptide agonist ligand experienced primarily hydrogen bonding and van der Waals interactions. Salt bridges were formed at residues 153D and 365E (**Figure 8**). The small molecule antagonist experienced a mixture of hydrophobic, hydrogen bonding, and polar interactions (**Figure 9**).

Table 1
Protein-Ligand Contacts between peptide ligand and CRF1R

Receptor Residue	Peptide Residue	Interaction
4 H	38 L	VDW
10 V	31 A	VDW
13 L	31 A	VDW
118 V	30 Q	HB
120 V	30 Q	HB, VDW
122 I	30 Q	HB
153 D	4 I	HB, Salt, VDW
156 D	4 I	HB
157 G	6 L	VDW
194 R	9 D	HB
224 Y	8 L	HB
231 N	9 D	VDW
365 E	21 Q	HB, Salt
366 D	25 R	HB
374 I	18 L	VDW

Residue number and 1-letter amino acid code for both receptor and ligand with their interaction type. VDW = Van Der Waals, HB = Hydrogen bonding, Salt = Salt Bridge.

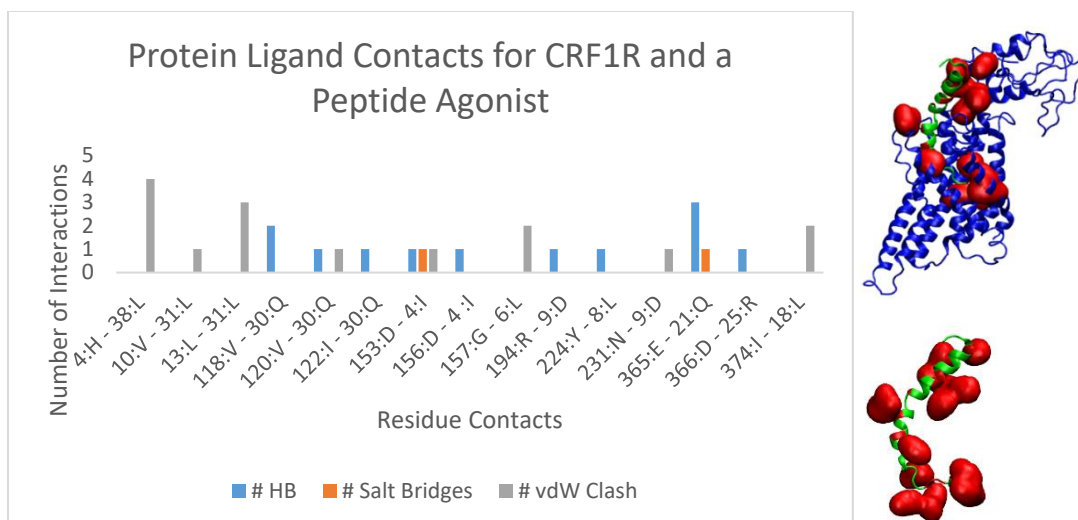


Figure 8. Ligand-protein interactions for peptide agonist system. For the residue contacts, bottom numbers represent residues on the receptor. Top numbers represent residues on the ligand. Residues that interact with the peptide are depicted by the red structures.

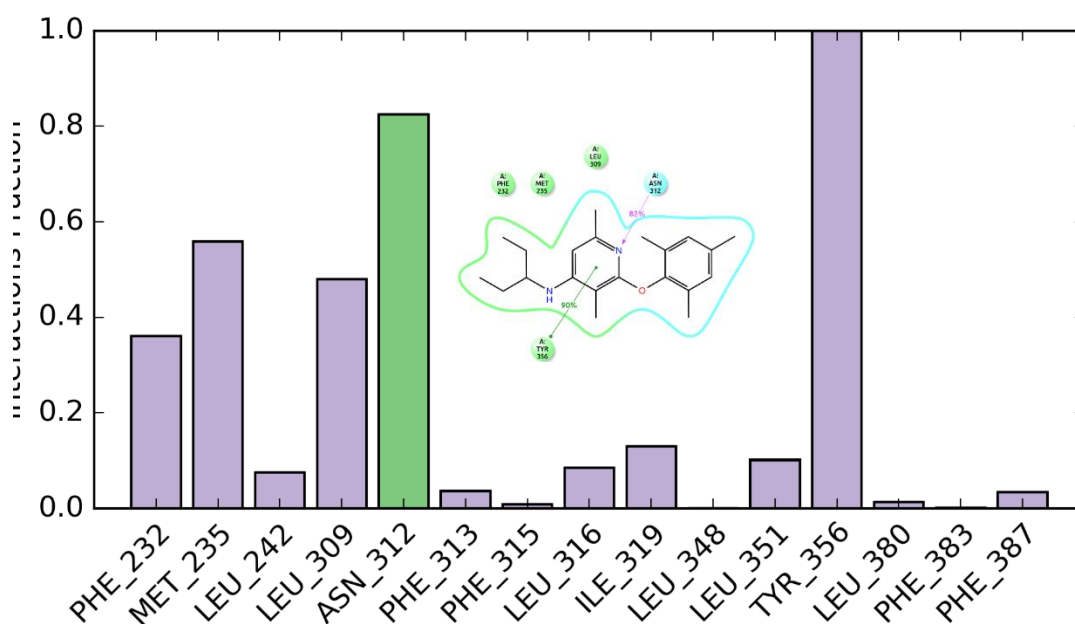


Figure 9. Ligand-protein interaction for small molecule antagonist. The blue molecules indicate polar interaction and the green molecules indicate hydrophobic interactions. On the histogram, the purple bars indicate hydrophobic interactions and the green bar indicates H-bonding interactions.

2.3.5 Different structural elements and RMSF values were observed. The small molecule system displays significantly more beta sheets than the peptide system. Differences also arise amongst the alpha helices between both systems (**Figure 10**). Notable differences include: significantly more beta sheets in the extracellular domain of the antagonist, different levels of helical structure and kinks in the transmembrane helices, and significantly less helical structure in the C-terminal region of the antagonist system. Our protein RMSF data shows that the receptor undergoes slightly different dynamic responses for both of the ligands tested. The protein C α Root Mean Square Fluctuation (RMSF) values for the peptide agonist and small molecule antagonist complexes are depicted in **Figure 11**. The mean values for each region can be found in **Table 1**. The C and N terminal regions proved, as expected, to be regions of high flexibility and as such exhibited high RMSF values. The TM region proved to be of similar flexibility to the loops, having RMSF values very close together. This is consistent with Seidel's³ findings in that the data supports both agonist and antagonist models maintain integrity and flexibility during simulation. Also consistent with Seidel's³ findings is Extracellular loop 3's (E3) RMSF value. E3 showed a much higher RMSF value than any of the other loops, indicating high levels of flexibility. Seidel proposed that the receptor takes on a wide pose (in agonists) and a compact pose (in antagonists) about TM 6 and 7. High flexibility on the loop that connects these two helices supports this claim.

Table 2

Mean values (\AA) for the RMSF of each portion of the ligand-receptor complexes.

Domain	2RMG (Ago)	Small Molecule
N-terminal	5.85	5.19
TM1	3.48	3.93
I1	3.17	3.27
TM2	2.99	2.39
E1	2.17	1.51
TM3	2.44	2.72
I2	1.96	1.74
TM4	1.91	1.80
E2	2.25	2.27
TM5	2.70	2.49
I3	1.99	2.36
TM6	2.51	2.31
E3	4.75	3.90
TM7	2.74	2.36
C-terminal	5.53	4.45
All I's	2.37	2.46
All E's	3.06	2.57
All Loops	2.72	2.51
All TM's	2.68	2.57
Overall	3.10	2.85

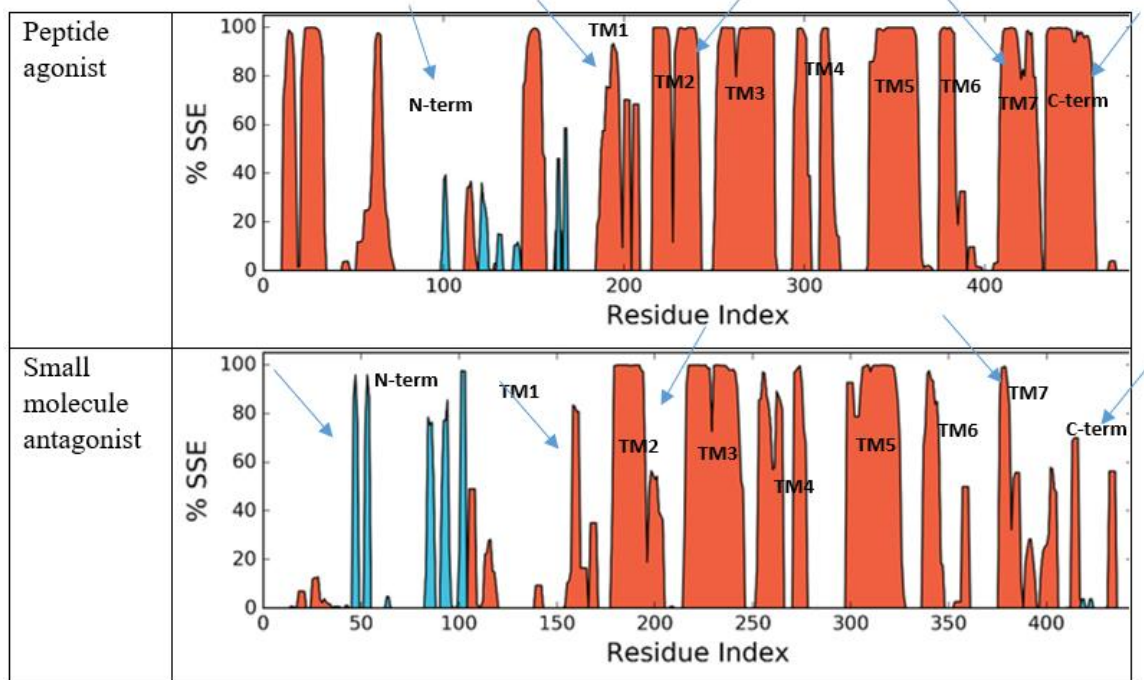


Figure 10. Secondary structure elements (SSE) for agonist and antagonist. Orange represents alpha helices. Blue represents beta sheets. N-terminal region: 0-145, TM1: 146-170, TM2: 178-205, TM3: 215-248, TM4: 255-281, TM5: 298-330, TM6: 339-362, TM7: 370-395, C-terminal region: 396-444. Key regions of difference are indicated by the arrows.

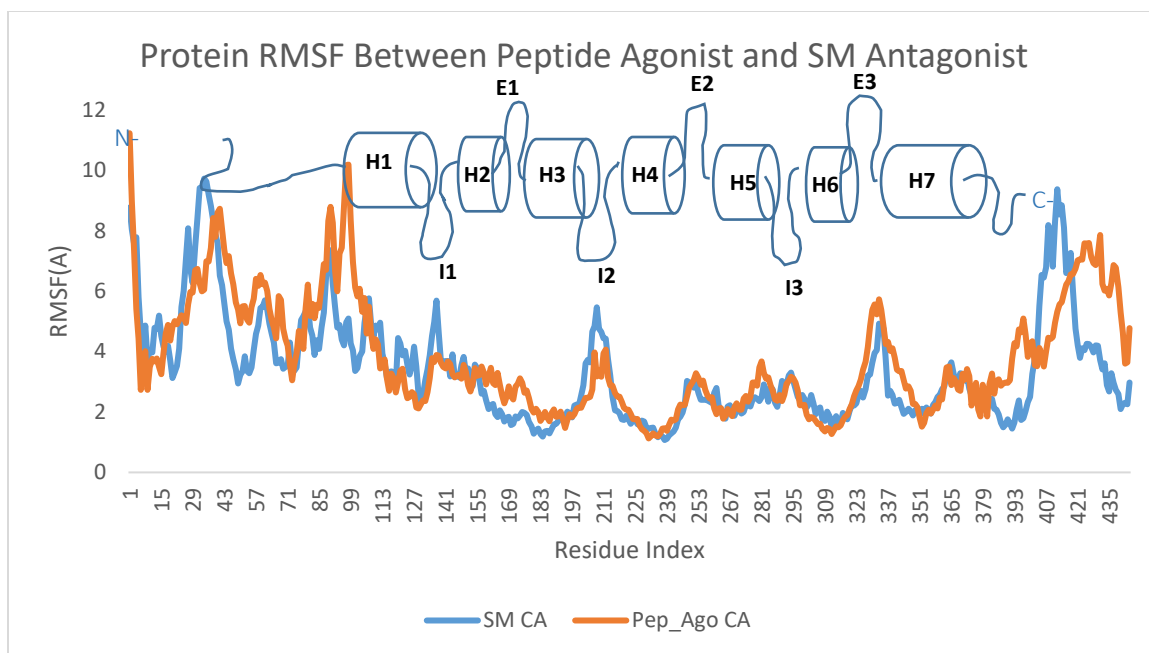


Figure 11. α RMSF diagram of small molecule and peptide agonist. Data taken after **2000ns** to measure flexibility at each region. N-terminal region: 1-102, C-terminal region: 393-444. **Helices:** H1: 103-144, H2: 149-175, H3: 185-219, H4: 226-254, H5: 268-298, H6: 304-332, H7: 339-392. **Intracellular loops:** I1: 145-148, I2: 220-225, I3: 299-303. **Extracellular loops:** E1: 176-184, E2: 255-267, E3: 333-338.

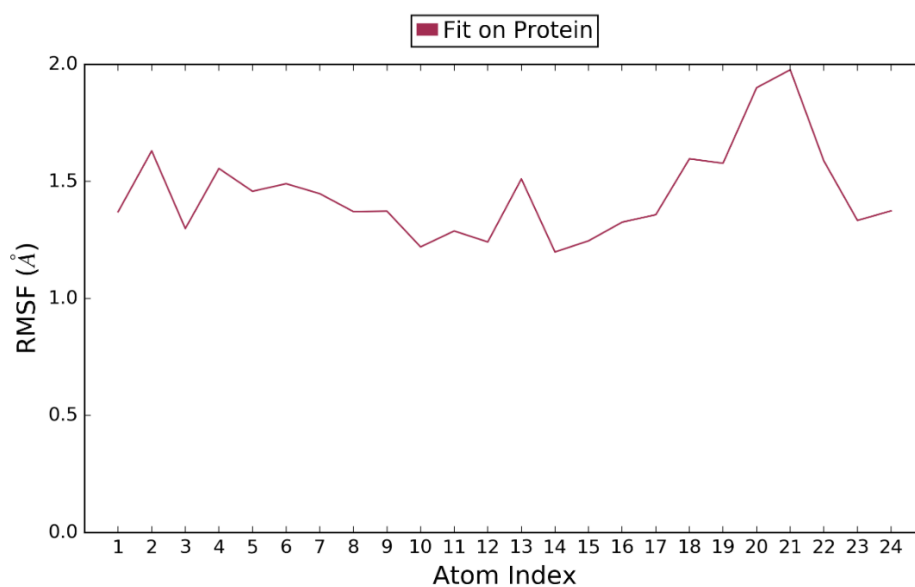


Figure 12. Ligand RMSF for small molecule antagonist to measure flexibility of each atom in the ligand throughout the entire simulation.

2.3.6 The radius gyration data showed shift in the agonist system. Both the N-terminal region (extracellular domain) and the C-terminal region (helix 8) were examined. In both regions, the peptide agonist system had overall higher radius gyration values producing a rightward shift (**Figure 13**).

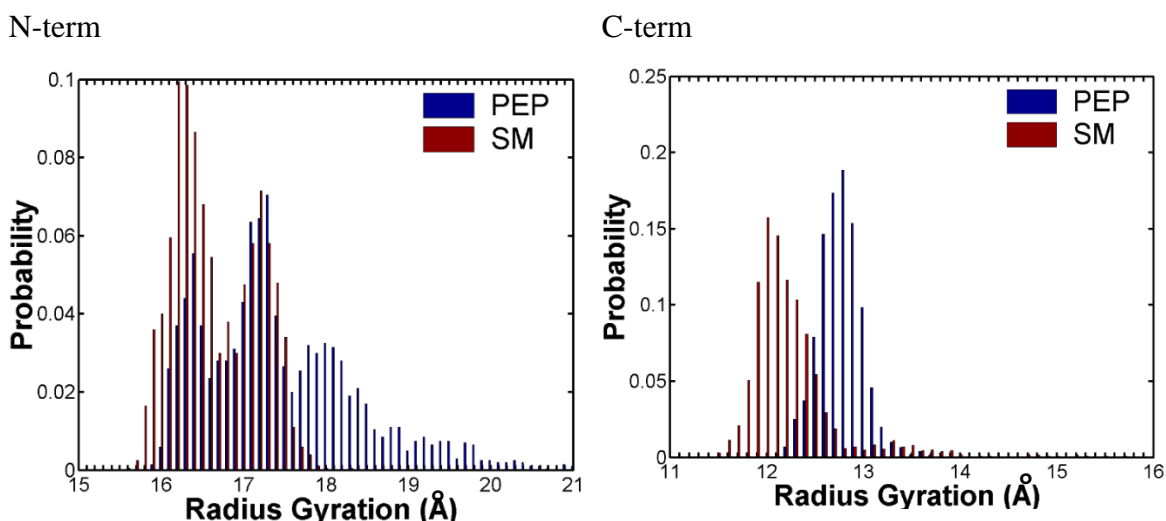


Figure 13. Radius gyration of peptide agonist and small molecule antagonist. Used to determine average size in angstroms of N-terminal region (left) and C-terminal region (right) of the receptor for the peptide agonist and small molecule antagonist systems. Blue = peptide agonist system, red = small molecule antagonist system.

2.3.7 Trajectory clustering analysis revealed three molecular switches.

Molecular switches are conformational changes induced by non-covalent interactions that cause the molecule to enter an activated state¹⁸. We were able to examine these molecular switches by superimposing the agonist and antagonist complexes of their most abundant cluster. In total, three molecular switches were studied: the Ionic lock, the Polar lock, and the Rotamer Toggle switch²⁵. The ionic lock, located between Arg180 and Glu238, was shown to be formed in both systems (**Figures S9 and S10**). To test whether the Ionic lock was formed for the entire simulation, a Matlab script was used to generate a graph of

distance (Å) versus time (ns) for the entire run. The graph clearly shows that the lock remains unbroken for the entire simulation. The polar lock, located between His180 and Glu238, appears to be broken in the agonist, but at least partially formed in the antagonist (**Figures 14 and 15**). Matlab was again used to test the entire simulation. The resulting graph showed that the small molecule antagonist consistently maintained the bond while the peptide agonist fluctuated, but was broken for a majority of the simulation. The two systems showed a distinct rotational difference in the Rotamer Toggle switch located at residues Tyr356, Phe313, and Leu352 (**Figure 16 and 17**). Residues Tyr356 and Phe313 showed the most rotational change. Leu352 showed little change between the two systems. We also examined a sequence motif GWGxP that is conserved among class B GPCRs to determine if it serves as a molecular switch in the activation of CRF1R. Little if any change was observed between agonist and antagonist systems (**Figure S11**).

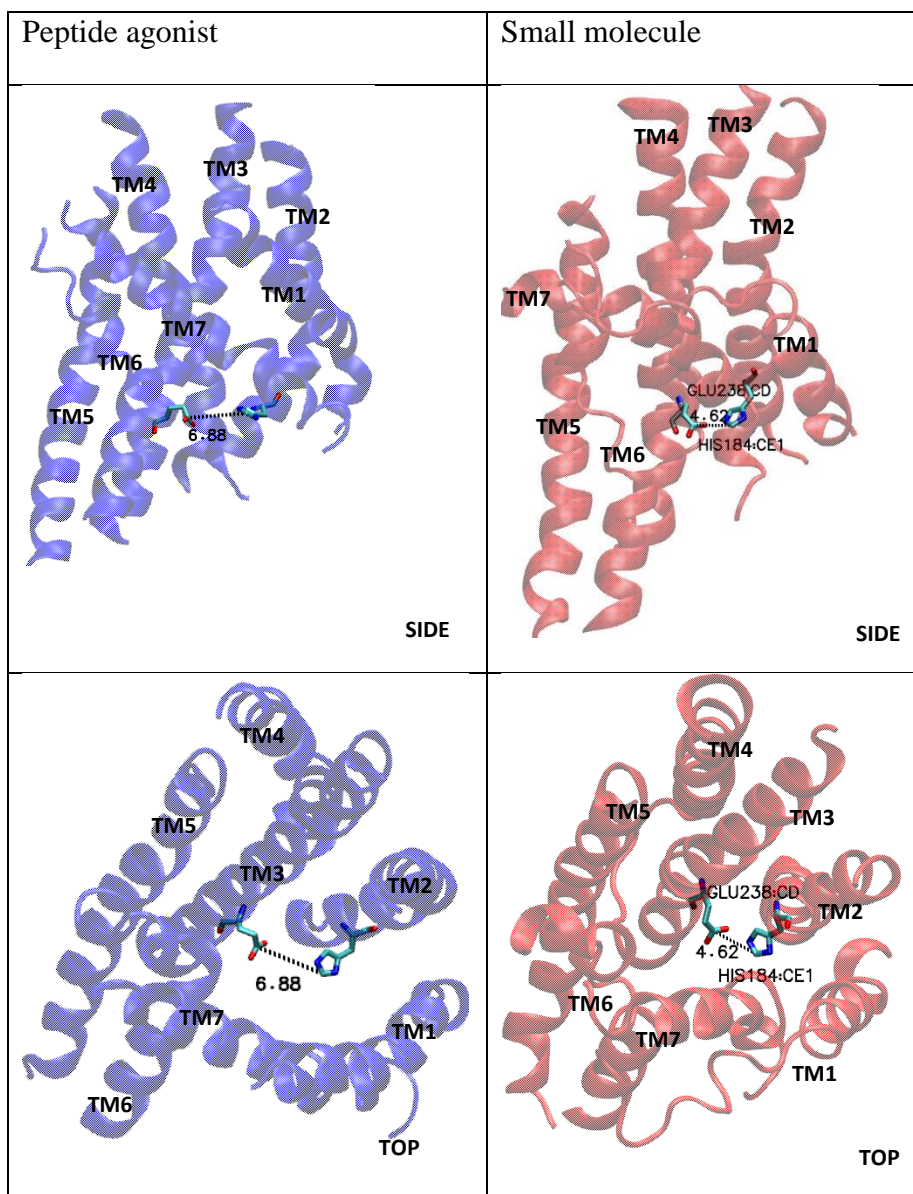


Figure 14. Polar lock in small molecule and peptide agonist. Displays location in reference to the transmembrane domain in both systems. Images taken from the first trajectory for the most abundant conformation obtained from trajectory clustering analysis. The Polar lock is located in the transmembrane domain on TM2 and TM3 at residues His155 and Glu209. See appendix A, Table 5 for generic numbering. Peptide agonist distance = 6.87. Small molecule distance = 6.31.

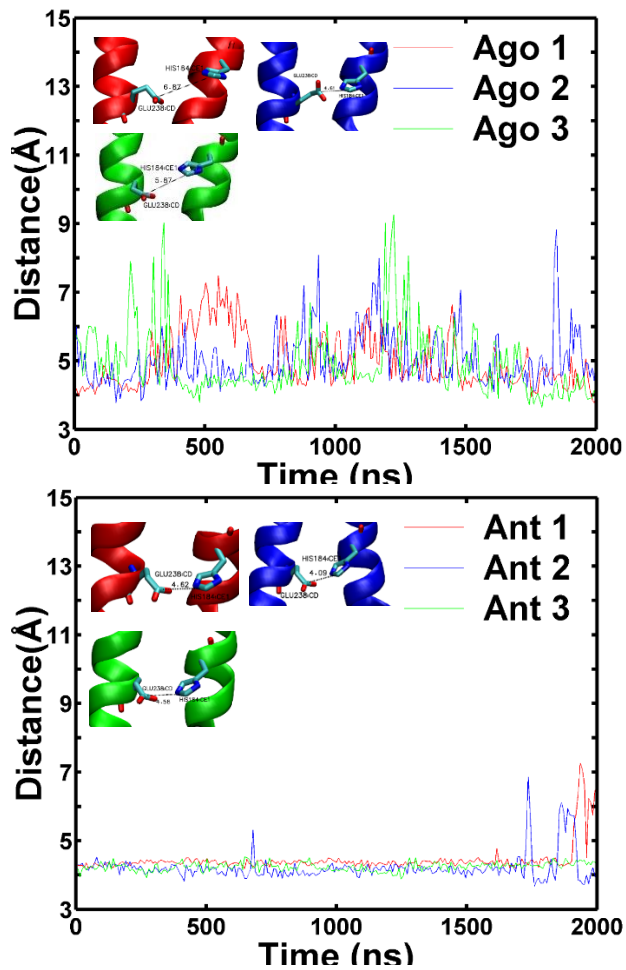


Figure 15. Polar lock distances small molecule and peptide agonist for 3 trajectories. Data taken after **2000ns** of simulation for all conformations for 3 trajectory simulations. Red = simulation trajectory 1, blue = simulation trajectory 2, green = simulation trajectory 3. Average distance for **Agonist** trajectory 1 = 6.87Å, trajectory 2 = 4.61 Å, trajectory 3 = 5.87 Å. Average distance for **Antagonist** trajectory 1 = 4.62 Å, trajectory 2 = 4.09 Å, trajectory 3 = 4.58 Å.

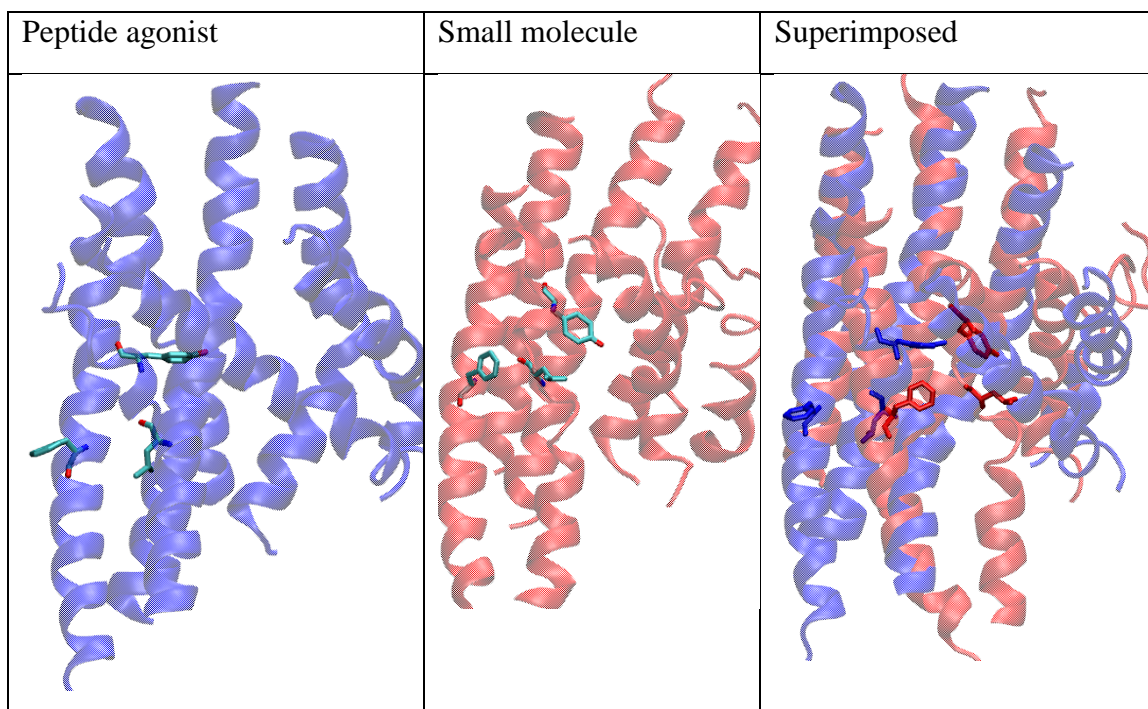


Figure 16. Rotamer Toggle switch in peptide agonist and small molecule. Displays location in reference to overall receptor after **2000ns** for the most abundant conformation obtained from trajectory clustering analysis. The Rotamer Toggle switch is located on TM5 and TM6 at residues Phe284, Leu323 and Tyr327. See appendix A, Table 5 for generic numbering.

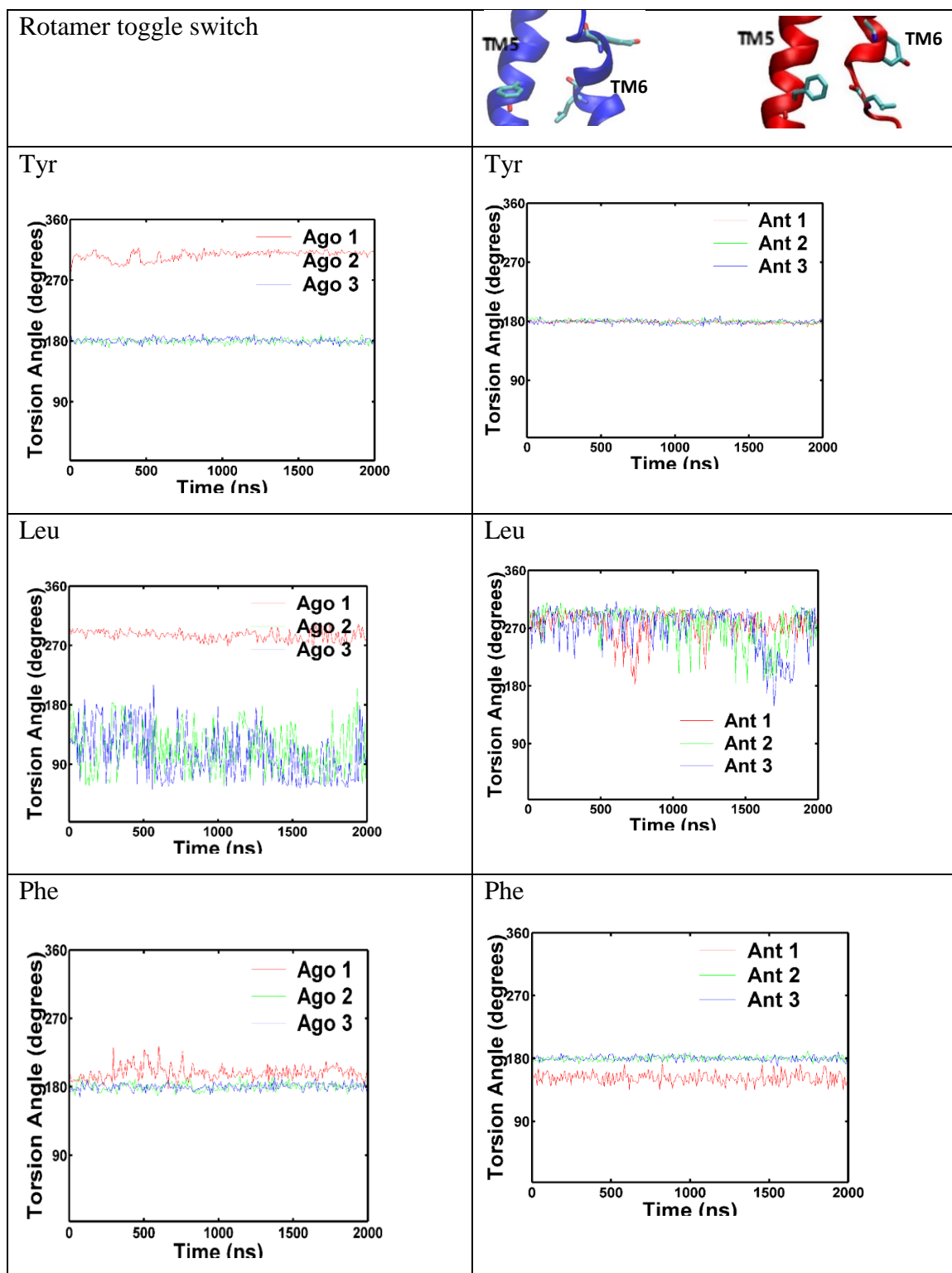


Figure 17. Rotamer toggle switch torsion angles for each residue. Depicts each residue Tyr356, Leu352, and Phe313 for both agonist and antagonist systems across 3 simulation trajectories. See appendix A, Table 5 for generic numbering. Red = simulation trajectory 1, green = simulation trajectory 2, blue = simulation trajectory 3.

2.3.8 Comparison to other GPCRs. Agonists appear to adopt a relatively conserved conformation to allow for G protein docking (**Figure S12**). After superimposing our agonist system on other G protein docked GPCRs, it was revealed that our agonist system adopts this same conformation. The c-terminal end opens up and the intracellular side of the transmembrane domain spreads to allow space for the G protein to dock. In our antagonist system, the c-terminal end uncoils and angles downward as well as the intracellular side of the transmembrane domain not spreading outward (**Figure 18, 19, S13-S25**). This would close off the site and prevent G protein docking as there would be no room for the G protein to fit.

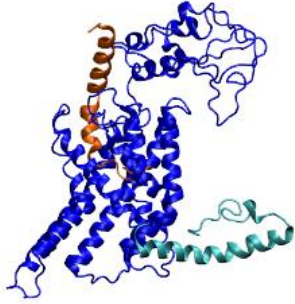
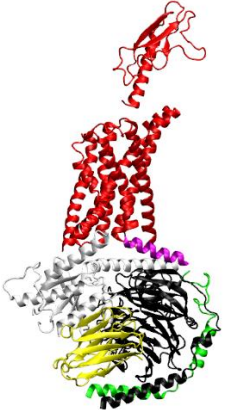
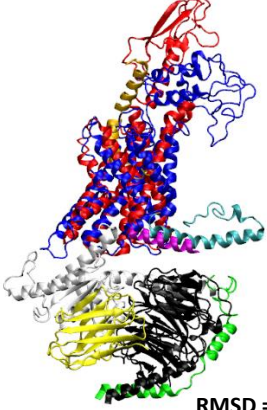
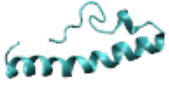

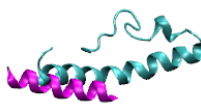
Our Peptide Agonist System	PDB ID: 6B3J	Superimposition
		
		

Figure 18. Comparison between our peptide agonist complex and the solved structure of the G protein docked Glucagon-like-peptide 1 receptor. Structure obtained from PDB 6B3J. Top panels compare the full receptor and bottom panels compare the C-terminal helices.

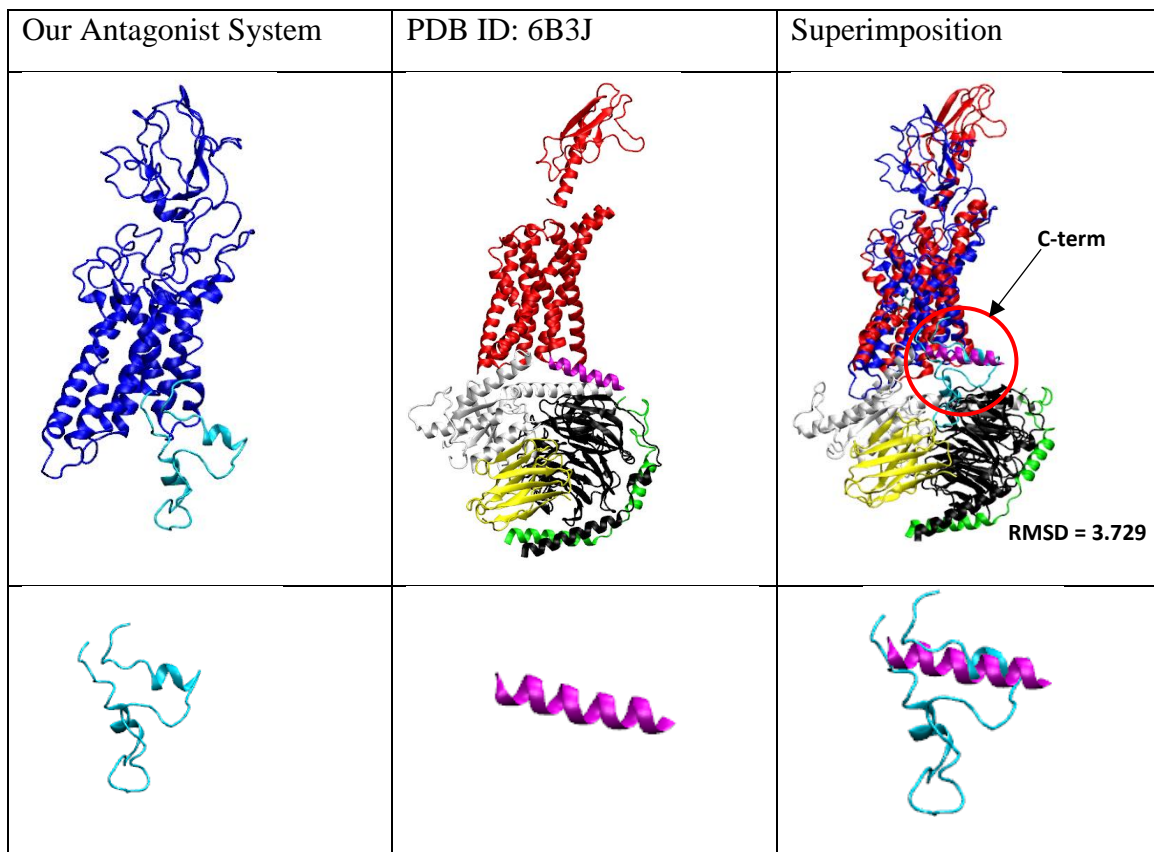


Figure 19. Comparison between our small molecule antagonist complex and the solved structure of the G protein docked Glucagon-like peptide-1 receptor. Structure obtained from PDB 6B3J. Top panels compare the full receptor and bottom panels compare the C-terminal helices.

2.4 Discussion

The CRF1 receptor is a good drug target for stress related issues such as depression, inflammation, and anxiety. A lack of FDA approved drugs targeting this receptor is due to a lack of knowledge on how class B GPCRs activate as well as a lack of a high resolution crystal structure for CRF1R bound to its peptide hormone. This makes it difficult to create ligands that are selective to only this receptor. Several previous studies have examined the CRF1 receptor binding to small molecule antagonist ligands,^{22, 25, 40} but there is a lack of studies on CRF1 bound to peptide ligands. Seidal²⁰ gained insight into the conformational differences between bound peptide agonist and peptide antagonist structures, but more studies need to be performed to gain a better understanding of the activation mechanisms. In our study, we examined how CRF1R interacts with a peptide agonist and a small molecule antagonist and propose key conformations and molecular switches that may play a role in the activation of the receptor. We develop a new possible mechanism of activation focusing less on molecular switches as previous studies have done.

Since there was no complete homology model for the CRF1R in the protein database, one had to be created. I-TASSER was used to create the homology model and did so using the top ten PDB templates for threading (Table S1). TASSER provided several possible models. The model we selected was the one that prioritized PDB 4K5Y because 4K5Y is the crystal structure of inactivated CRF1R, without the extracellular domain. This model would allow us to determine the effects that our ligands have on the activation of the receptor. We compared this model to the crystal structure for the transmembrane domain (PDB ID: 4K5Y) and extra cellular domain (PDB ID: 3EHU) in figure 2. This generic pose was used as the starting position, which quickly changed conformation upon docking of the ligands. Since 4K5Y contains the docked ligand CP

376395, this was used to justify our docking orientation. The justification of our docking orientation for the peptide ligand comes from PDB 5NX2 which docked a truncated peptide with the GLP-1 receptor (figure S4). The N-terminal region was placed in the binding pocket and the C-terminal region was oriented to interact with the extracellular domain. In terms of overall conformational change, distinct formations were associated with whether the protein-ligand complex was agonized or antagonized. We compare the conformational difference of each region of interest (extracellular domain, transmembrane domain, and helix 8) in figure 5. For the extracellular domain, the agonist conformation bends outward, opening up the transmembrane region for activation while the antagonist conformation maintained a vertical formation, closing off and preventing activation of the receptor. Our RMSD data in figure 4 further makes the distinction between agonist and antagonist conformations. The transmembrane region is more fully analyzed in figure 6. The conformational shift is very distinct in helices 1, and 5-7 between agonist and antagonist structures. Another important conformational difference was in helices 8. The agonist had a more rigid helix 8 angled upward aiming at the extracellular domain. The antagonist showed to be in a more relaxed, uncoiled state and angled downward. This indicates that the conformational change in the extracellular domain in the agonist complex induces the activation of helix 8.

We found that the peptide agonist and small molecule antagonist experience different types of protein-ligand interactions. Figure 7 depicts which residues are interacting with each other and what type of interaction there is for the agonist system. There are primarily hydrogen binding and van der Waals interactions, but notably there are two salt bridges that form in the agonist system which do not appear in the antagonist

system. Figure 8 depicts the protein-ligand interactions that exist between the small molecule and the receptor. A notable difference here is the existence of a polar interaction with the small molecule that does not exist in the agonist system. These differences could contribute to the activation/deactivation of the receptor.

Our secondary structure data in figure 9 summarizes structural differences between agonist bound and antagonist bound systems. Firstly, the extracellular domain of the agonist system display much more helical structure than in the antagonist system. The antagonist system also displays much more beta sheet structure in this region than the agonist system. This could indicate that helical structure in the extracellular domain is linked to activation and beta sheets are linked to inactivation. Also the secondary structure elements highlight the various kinks and shifts of the transmembrane region that we saw in figure 6. Finally, the C-terminal region was significantly less structured in the antagonist than in the agonist.

Our data is also consistent with Seidal's²⁰ findings. He proposed that agonists adopt different folds than antagonists in order to stabilize the transmembrane domain. This difference occurs at TM 6 and 7 to produce a wide pose for agonists and a compact pose for antagonists. Our RMSF data in Table 1 and figure 10 shows that the E3 loop (the loop that connects TM 6 and 7) has a higher flexibility than any other region in the transmembrane domain. This high flexibility, along with the images we generated of the transmembrane region in figure S6, reinforces his findings that antagonists develop a compact pose and agonists develop a wide pose. Furthermore, our findings also indicate that the extracellular domain also plays a role in the development of these poses. The extracellular domain opens, activating helix 8 which spreads the receptor into its wide

pose. This is important to note because it could affect interaction with the G-protein, which we examine later on.

Other important regions to look at when examining activation are molecular switches, conformational changes induced by non-covalent interactions. The three molecular switches we focused on were the ionic lock, polar lock, and rotamer toggle switch which Xu identified as being likely linked to activation²⁵. The ionic lock, located between TM2 and TM3 at Arg151 and Glu209, (see appendix A, Table 5 for generic numbering) was formed in the small molecule antagonist and peptide agonist complexes. This indicates that this lock likely does not play an important role in the activation of the receptor (figure S9-S10). The Polar lock is located between TM2 and TM3 at residues His155 and Glu209. In their most abundant conformations, this bond is formed in the antagonist complex, but broken in the agonist complex suggesting that the formation of this bond could play a role in inhibiting this receptor (figure 13-14), however it not believed to be a major contributor due to high variation in distance lengths. After running three separate trajectories, the polar lock ranges from 4 to 9Å in varying durations in the agonist system and remains fairly constant at 4Å in the antagonist (figure S29 and S54). The last molecular switch we looked at was the rotamer toggle switch located between TM5 and TM6 at Phe284, Leu323, and Tyr327 (figure 15-16). We examined the different rotational differences of these residues and initially determined that the agonist complexes showed different torsions to the antagonist. Residues Phe313 and Tyr356 show nearly a 180 degree rotational difference. However, inconsistencies were found after running three trajectories and it was determined that the rotamer toggle switch is relatively flexible in both agonist and antagonist systems (figures S30 and S56). The

GWGxP conserved sequence motif was also examined for differences between our two systems (figure S11). Xu²⁵ concluded that it does not play a role in antagonist binding as there were no noticeable χ_1 changes between the apo form and the antagonist bound form. Our results agreed with his conclusion. In our agonist bound model, there were also no noticeable changes, indicating that this conserved sequence also plays no role in agonist binding. Overall, it was determined that molecular switches, while likely contributing to activation, do not play as critical a role as originally thought.

Most important to activation of the receptor is its ability to interact with the G-protein. CRF1R interacts with a Gs protein (stimulating), so we found several receptors that have been docked to Gs protein in PDB for comparison. Upon superimposition of five different receptors, it was revealed that there is a high degree of conservation in the structure of the receptors to allow for docking of the G-protein. Our agonist and antagonist systems were then superimposed individually to each of the five receptors to gain insight into how our systems might interact with the G-protein. Our peptide agonist system adopted a similar conformation to all five receptors in comparison indicating that the peptide agonist signals for interaction with the G-protein. The key region to look at when making this claim is the C-terminal helix. The helix must undergo a conformational change to allow room for the G-protein to dock. We superimposed all C-terminal helices with our agonist system helix which showed that the agonist system forms a similar conformation. Furthermore, it was noted that the bottom region of the receptor also spreads itself, allowing for more room for the G protein. Our antagonist system does not match the others, indicating it does not allow for interaction with the G-protein. The C-terminal helix becomes uncoiled and angles downward, causing a clash with the

superimposed G-proteins. This clash indicates that the helix would prevent the G-protein from being able to dock to CRF1R, thus preventing activation. The bottom region of the receptor also does not spread like the others do, further justifying the importance of this conformational shift. The large scale conformational changes associated with the transmembrane helices and helix 8 are believed to be the most important changes in receptor activation.

2.5 Conclusions

The CRF1 receptor would be a good drug target for conditions such as depression, anxiety, and inflammation since the receptor is associated with stress response. Unfortunately, there are not currently any FDA approved drugs that target this receptor due to a lack of a high resolution structure and understanding of the activation method. Previous studies on the CRF1R primarily focused on small molecule antagonist complexes, or distinguishing conformational differences between peptide agonists and peptide antagonists. We expand upon Seidel's²⁰ study by exploring conformational changes associated with CRF1R in complex with a peptide agonist (Urocortin), and small molecule antagonist (CP 154526). Previous studies on activation have focused on molecular switches. While likely a contributing factor, we propose that these molecular switches are not as critical as previously thought. We propose that activation is linked to large scale conformational changes associated with the transmembrane helices and the intracellular C-terminal helix. We first generated a homology model for CRF1R using I-TASSER and compared to crystal structures from the Protein Data Bank (PDB 4K5Y and 3EHU). The ligands were then individually docked to the receptor and subjected to a total of 4 μ S (2 x 2 μ S) molecular dynamics simulations, followed by trajectory clustering and

simulation interaction diagram analysis. Our RMSF data was consistent with Seidel's²⁰ findings and we showed distinct conformational differences between agonist and antagonist complexes in the transmembrane domain, extracellular domain, and helices 8. Inconsistencies in our molecular switch data suggest that they play a far less significant role in activation than previously thought. Upon comparing our systems to other receptors, it was determined that they all adopt a conserved conformation to allow for G-protein docking. One key aspect of this conformation is the C-terminal helix, which must angle out to allow space for the G-protein. Additionally, the bottom of the receptor must spread to help provide room for the G protein. Only our agonist system adopts these conserved conformation. Our findings point toward a new mechanism of activation and provide insight into developing drugs to target this class of receptor.

Chapter 3

GLP-1 Receptor in Complex with a Full Agonist and a Biased Agonist Probed by Molecular Dynamics Simulations for the Development of more Specific Drugs on Type 2 Diabetes

3.1 Introduction

3.1.1 G protein-coupled-receptors. The largest family of cell surface receptors are G protein-coupled-receptors (GPCRs). The general structure of a GPCR consists of an extracellular region, a transmembrane region composed of seven helices, and an intracellular C-terminus. Another defining characteristic of GPCRs is their ability to interact with G proteins to stimulate signaling pathways. GPCRs can be sub-classified into five groups, the largest of which is class A rhodopsin composed of over 700 receptors. Other classes include class B secretin, class C glutamate, class D adhesion, and class E frizzled. Structural difference between these classes affect the ligand recognition mechanisms of the receptor and therefore affect their ability to transduce signals across the membrane¹⁹. Class B GPCRs are generally less understood when compared to class A, and were therefore chosen as the focus for this paper. A defining feature of class B GPCRs are their large extracellular domain which plays an integral role in ligands detection and binding. The extracellular domain specifically helps with peptide ligand binding. The C-terminal end of the peptide is first bound to the extracellular region which is used to guide the ligand into the binding pocket. This induces conformational changes to the receptor as a whole. Overall, there are 15 known class B GPCRs that control a number of functions in the body. They can be good drug targets to treat conditions ranging from diabetes to osteoporosis²⁰.

3.1.2 The glucagon-like-peptide receptor. The glucagon-like-peptide-1 receptor (GLP1R) is a class B GPCR that stimulates the secretion of insulin. It is primarily expressed in the pancreas, but can also be found in the heart and brain³⁹. For this reason, the GLP1R is a good drug target for type 2 diabetes. Individuals with diabetes have problems maintaining insulin levels. Insulin is important for transporting glucose into your cells so they can be converted into energy. Type 1 diabetes is a condition in which the individual cannot produce their own insulin. Type 2 diabetes, often called adult onset diabetes and associated with obesity, is where the body is capable of producing its own insulin, but is not producing it in high enough concentrations. Developing drugs on this class of receptor to stimulant insulin secretion can help improve the lives of the many inflicted with this condition.

N-term					
MAGAPGPLRL	ALLLLGMVGR	AGPRPQGATV	SLWETVQKWR	EYRRQCQRSL	TEDPPPATDL
10	20	30	40	50	60
FCNRTFDEYA	CWPDGEPGSF	VNVSCPWYLP	WASSVPQGHV	YRFCTAEGWL	LQKDNSSLPW
70	80	90	100	110	120
	TM1				ICL1 TM2
RDLSECEESK	RGERSSPEEQ	LLFLYIIYTV	GYALSFSALV	IASAILLGFR	HLHCTRNYIH
130	140	150	160	170	180
		ECL1		TM3	
LNLFASFILR	ALSVFIKDA	LKWMYSTAAQ	QHQWDGLLSY	QDSLSCRLVF	LLMQYCVAN
190	200	210	220	230	240
		ICL2 TM4			ECL2
YYWLLVEGVY	LYTLLAFSVL	SEQWIFRLYV	SIGWGVPLLF	VVPWGIVKYL	YEDEGCWTRN
250	260	270	280	290	300
TM5				ICL3 TM6	
SNMNYWLIIR	LPILFAIGVN	FLIFVRVICI	VVSKLKANLM	CKTDIKCRLA	KSTLTLIPLL
310	320	330	340	350	360
	ECL3 TM7			H8	
GTHEVIFAFV	MDEHARGTLR	FIKLFTELSF	TSFQGLMVAI	LYCFVNNEVQ	LEFRKSWERW
370	380	390	400	410	420
	C-term				
RLEHLHIQRD	SSMKPLKCPT	SSLSSGATAG	SSMYTATCQA	S C S	
430	440	450	460		

Figure 20. Full length sequence of human GLP1R. Data taken from GPCRdb²³⁻²⁴.

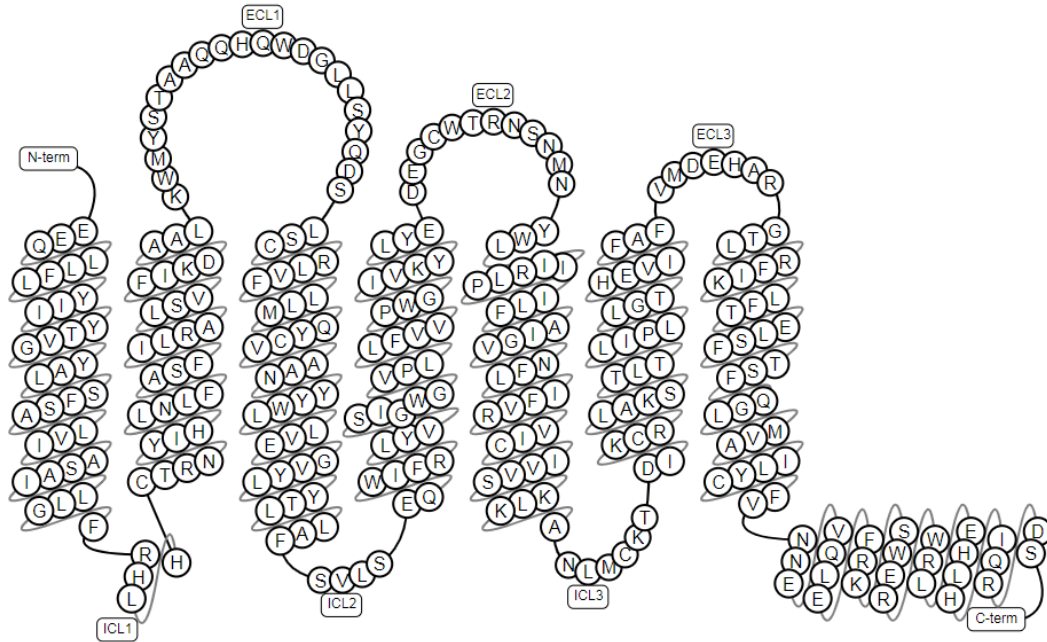


Figure 21. Snake plot of full-length human GLP1R. Data taken from GPCRdb²³⁻²⁴.

3.1.3 Full vs. biased agonism. GPCR signaling mechanisms are not limited to a single pathway. There are several pathways that can be activated by a ligand, most notably the G protein pathway and the β -arrestins pathway. Full agonist ligands activate the whole receptor and all signaling pathways associated with it. Drugs modeled after full agonists can have inadvertent side effects due to this unspecific signaling pathway. Biased agonism is where the ligand is specific to one particular signaling pathway. If the activation mechanism of biased agonism can be more fully understood, drugs can be developed to be specific to one pathway therefore reducing the amount of side effects the drug causes. In this study, we examine exendin-P5 (Exp5) which has been shown to be biased to the G protein pathway³⁸.

3.3.4 Experimental overview. Here we explore the conformational changes associated with the GLP1 receptor bound to a full agonist (truncated GLP1 peptide PDB ID: 5NX2) and a biased agonist (Exp5 PDB ID: 6B3J) to develop a more detailed

understanding of the activation mechanism of biased agonists. Molecular dynamics simulations were performed on these systems as well as the unbound form (APO) and analyses were performed. Analyses include simulation interaction diagram, trajectory cluster analysis, and MMGBSA analysis. Our findings help to elucidate the activation mechanism of biased agonists which can lead to the development of more specific drugs for the treatment of type 2 diabetes.

3.2 Methods

3.2.1 Protein and Ligand Preparation: The structure of the GLP1 receptor was obtained from the protein data bank (PDB) from PDB ID: 5NX2. The full agonist ligand was also obtained from this entry and the biased ligand was obtained from PDB ID: 6B3J. Maestro's Protein Preparation Wizard²⁷ was used to prepare the models for docking. Operations performed included a preprocessing step which added hydrogens where appropriate, corrected bond orders, and removed water where appropriate. This was followed by a charge state optimization and then a restrained minimization.

3.2.2 Docking: Maestro's Glide Docking feature was used to dock the ligands to the receptor. To do this, a grid is first generated and the ligands and receptors are specified. The biased agonist was docked using standard precision peptide mode (SP-peptide) in an OPLS3 field. The peptide was oriented such that the C-terminal end interacts with the extracellular domain (ECD) and the N-terminal end interacts with the binding pocket³¹⁻³². The truncated peptide full agonist was small enough the Maestro was

able to detect it as a small molecule ligand and was therefore docked using extra precision (XP) mode.

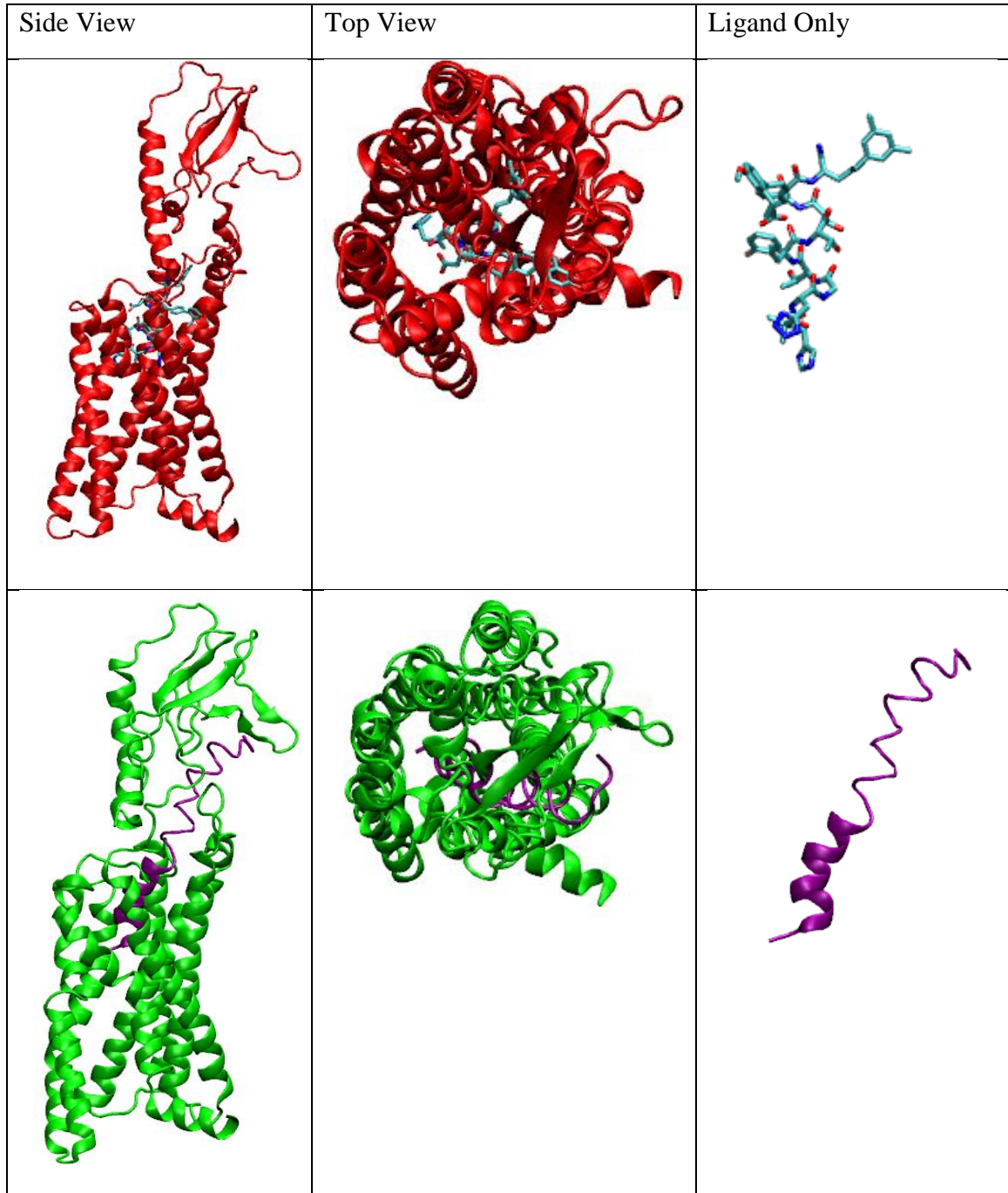


Figure 22. Docked full agonist (red) and biased agonist (green). Ligands docked into the GLP1 receptor obtained from PDB 6b3j.

3.2.3 Molecular Dynamics Simulations: Three systems, full agonist docked, biased agonist docked, and APO form were run under molecular dynamics (MD) simulations. Systems were aligned in a membrane which was set to the helices of the transmembrane region. They were then solvated in a water box using the SPC water model. 0.15M NaCl was added to help neutralize the system. The Desmond System Builder in Maestro was used to generate an OPLS3 force field to build these systems. MD simulations ran for a total of 6 μ S (2 μ S per system).

3.2.4 Trajectory Clustering Analysis: The structures of the final 100ns of each simulation was grouped using the Desmond trajectory tool. The structural similarity metric selected was backbone RMSD and average linkage was selected for hierarchical clustering. Merging distance cutoff was set to 2.5 Å. Frequency was set to 2. The number of frames was set to 250. Number of threads was set to 2. Structures who showed a frequency greater than 2% were saved for further analysis. The most abundant cluster for each system was aligned and superimposed to compare structural differences in the ECD, transmembrane domain (TMD), and C-terminal helix 8.

3.2.5 Simulation Interaction Diagram (SID) Analysis: Maestro's SID tool is used to analyze interactions between the protein and ligand. SID includes Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF), secondary structure analysis, and protein-ligand interactions. RMSD measures the displacement change of atoms for the entire trajectory with respect to the reference frame and is calculated using the below equation where N is the number of atoms, t is time, and r is the position of the atom.

$$RMSD_x = \sqrt{\frac{1}{N} \sum_{i=1}^N (r'_i(t_x) - r_i(t_{ref}))^2}$$

RMSF analyzes changes along a protein chain or molecule. The RMSF equation used for this calculation is shown below where T is the trajectory, t is time, r is residue position, and the < and > signs indicate that the average of the square distance is taken.

$$RMSF_i = \sqrt{\frac{1}{T} \sum_{t=1}^T \langle (r'_i(t)) - r_i(t_{ref}) \rangle^2}$$

Secondary structures are monitored for the entire trajectory. Alpha-helical structures and beta strands can be observed as well as kinks that develop over time. SID also analyzes the types of protein-ligand interactions that occur over time (i.e. H-bonds, van der Waals, etc.).

3.2.6 Comparison to other GPCRs: Five different complexes were taken from the Protein Data Bank (PDB IDs: 6b3j³⁸, 5uz7³⁷, 5vai³⁶, 5g53³⁵, 3sn6³⁴) that were docked to the G protein. Each complex was aligned with our systems individually and superimposed within Maestro to examine our systems ability to interact with the G protein. Comparisons were also made to a beta arrestin bound GPCR (PDB ID: 4zwj⁴¹). Structures were aligned in Maestro and optimized in VMD.

3.2.7 MMGBSA Analysis: MMGBSA was used to calculate binding energies of the last 100 ns of each system. The default procedures were used when performing this analysis. First the receptor was minimized, then the ligand was minimized, followed by a receptor-ligand complex minimization. The equation for total binding free energy is:

$$\Delta G (\text{bind}) = E_{\text{complex}} - (E_{\text{ligand (minimized)}} + E_{\text{receptor (minimized)}})$$

3.3 Results

3.3.1 Conformational differences between full and biased agonist systems.

There were three regions which showed conformational differences which include the ECD, TMD, and the C-terminal helix. In the most abundant conformation, obtained from the trajectory clustering analysis, the full agonist showed the ECD tilted toward the right (**Figure 24**). The TMD showed a shift in helices focusing on the intracellular region. The region appears to have spread and opened up, which is typical of agonists. The C-terminal helix angled out and also seemed to spread. For the APO form, the ECD maintained a more vertical conformation. The TMD and C-terminal helix did not spread and remained closed off. The biased system's ECD was more vertical than the APO form, but not as tilted as the full agonist complex. The TMD maintained an open, yet narrower, conformation and the C-terminal helix inverted and angled inward. Our RMSD data shows that both biased and APO systems stay within a certain range once they level off. The full agonist system levels off quickly, but jumps up toward the end of the simulation when the intracellular region widens (**Figure 23**).

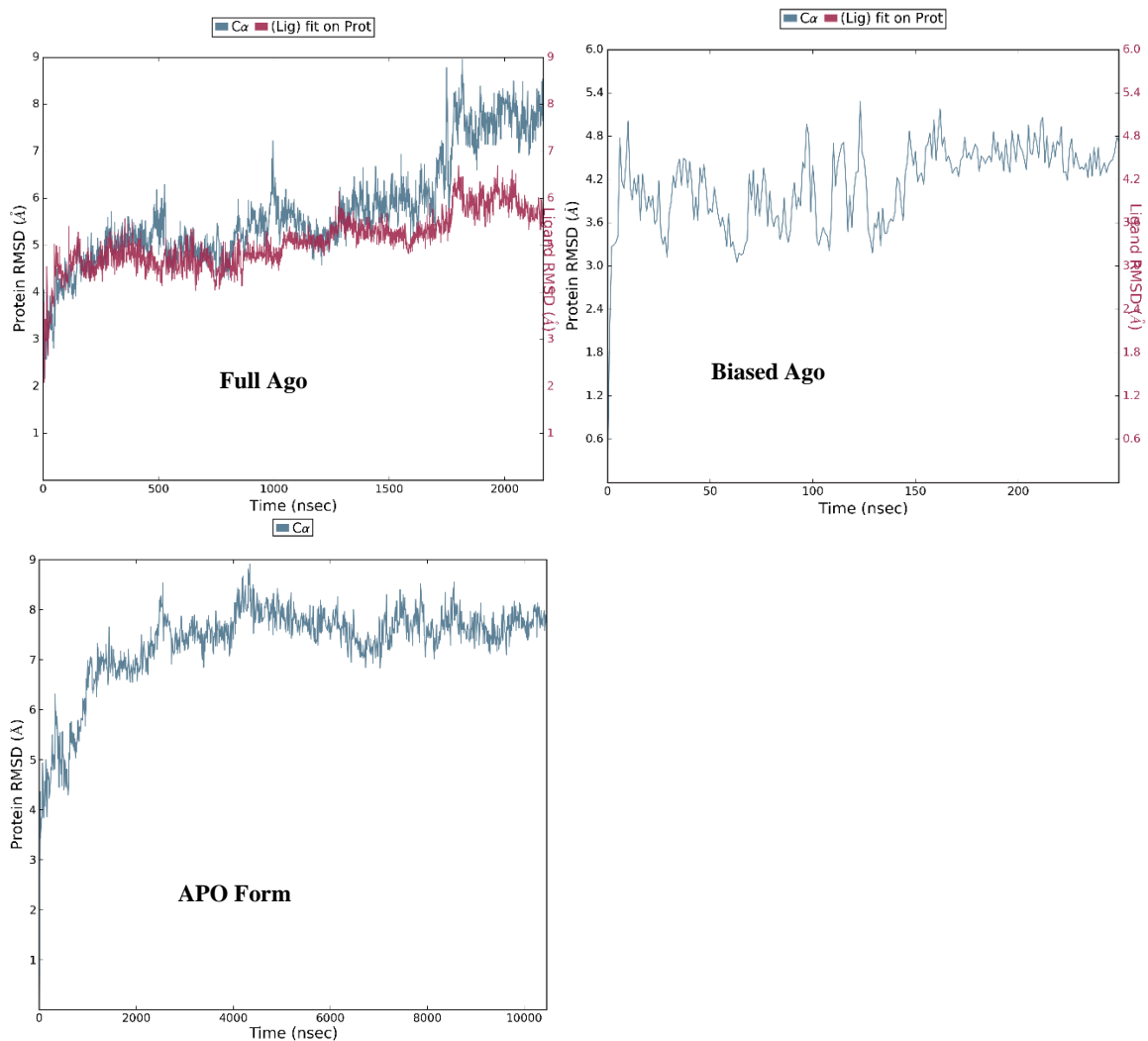


Figure 23. Root Mean Square Deviation (RMSD). Full unbiased agonist (left), biased agonist (right), APO form (bottom). Used to measure the average change in displacement of a selection of atoms for a particular frame with respect to a reference frame. It is calculated for all frames in the trajectory. Blue = protein receptor, red = ligand.

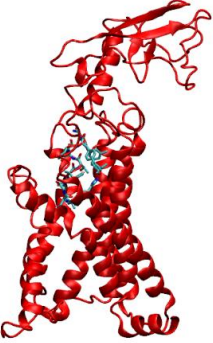


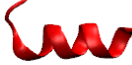
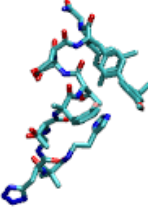
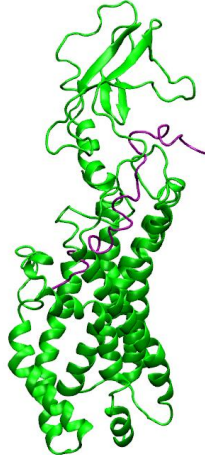
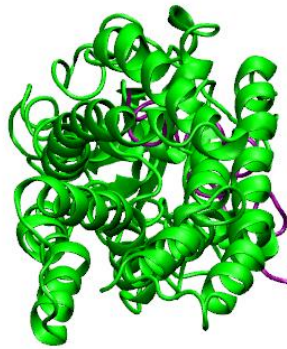



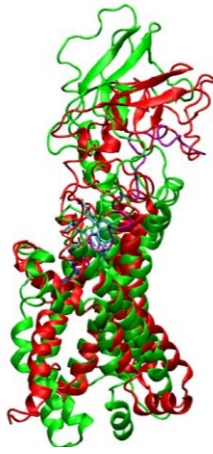
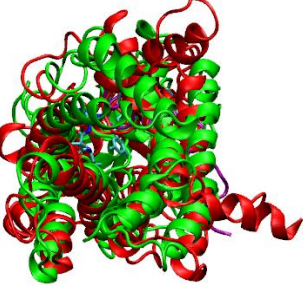
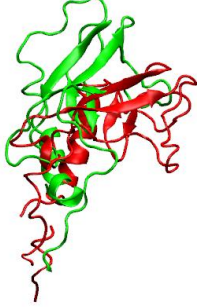

Full System	TM domain (bottom)	ECD	C-terminal	Ligand Only
				
				
				

Figure 24. Most abundant conformations. Pairwise comparison of the most abundant conformations for full agonist system (red) and biased agonist system (green) obtained from the trajectory clustering analysis. Comparisons broken down into key regions: TMD, ECD, C-terminal, and ligand.

3.3.2 Different types of protein-ligand interactions. Overall, the biased agonist had more interaction with the receptor since it was a much larger molecule. The full agonist ligand experienced predominantly hydrophobic, polar, and charged reactions, but also experienced salt bridges and pi-stacking. The biased ligand most predominantly experienced hydrophobic interactions, but also experienced polar, charged, salt bridges, and pi-stacking (**Figure 25**).

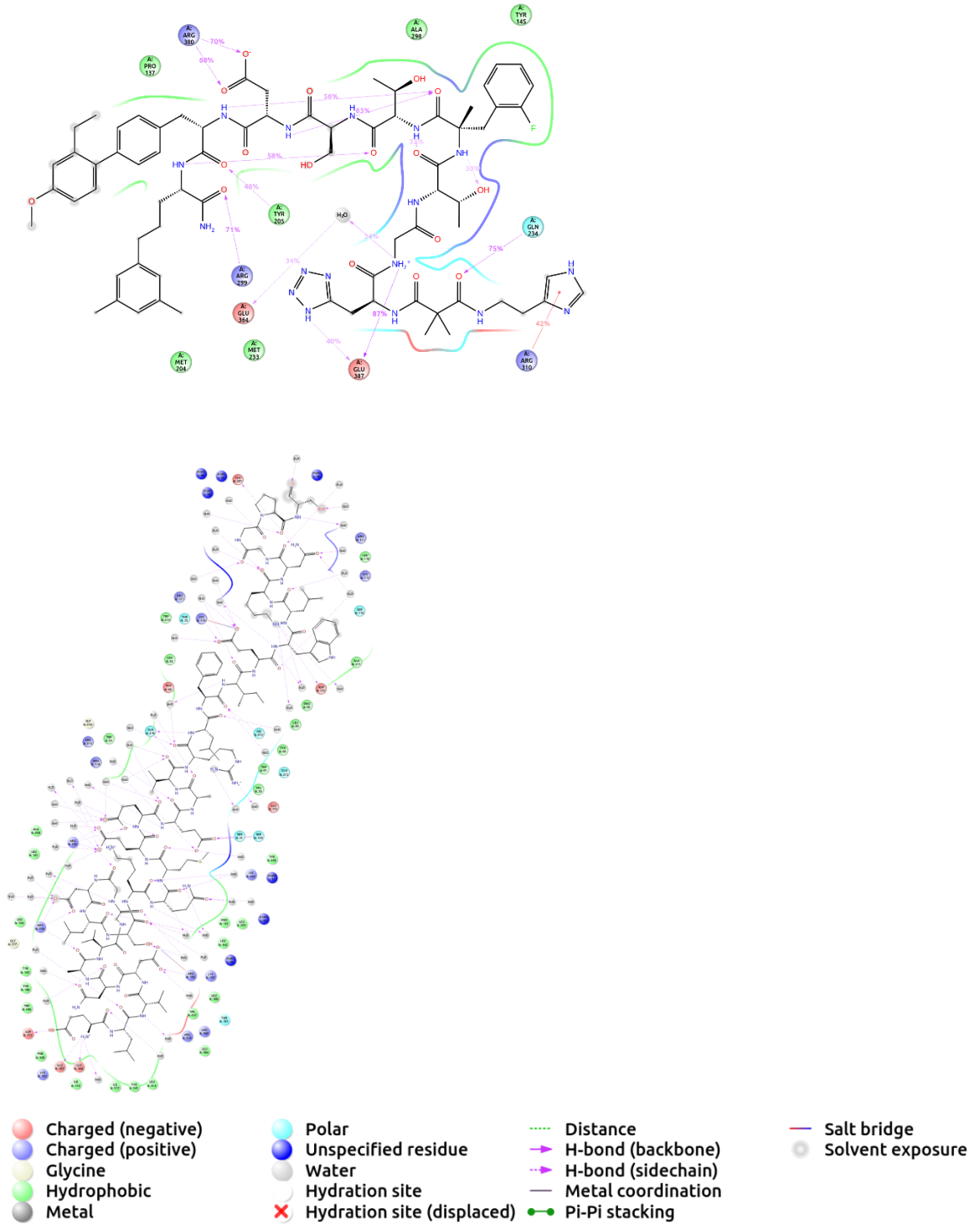


Figure 25. Protein-ligand interactions. Full agonist (left) and biased agonist (right).

3.3.3 Secondary structures revealed differences among the three systems. The

full agonist and biased agonist systems showed very similar secondary structures for the whole simulation. The major difference appears in the C-terminal helix 8. The biased system shows significantly less helical structure then the full agonist system. The APO form showed less helical structure across the TMD (Figure 26).

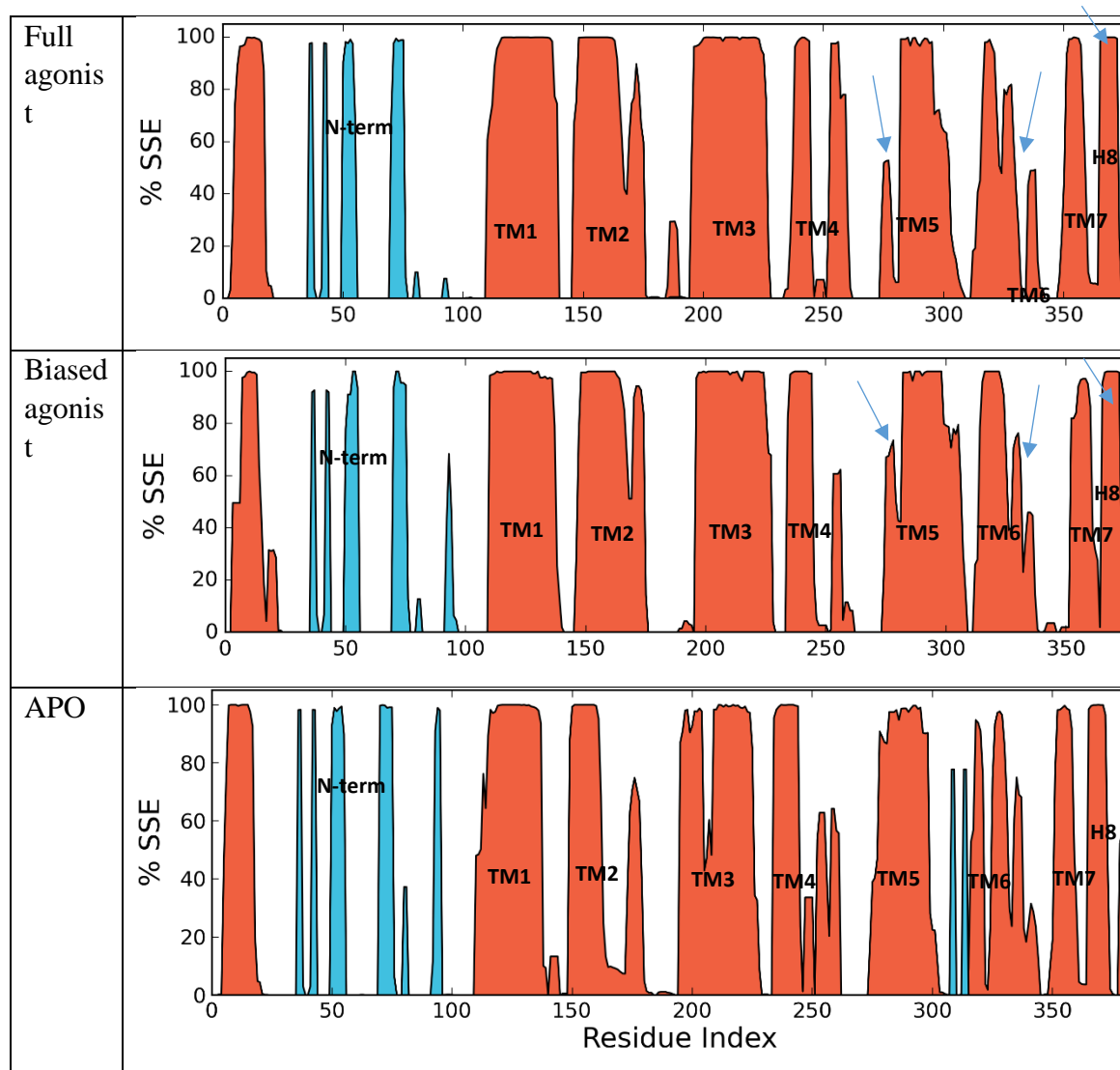


Figure 26. Secondary structure elements (SSE). Comparison for full agonist, biased agonist, and APO systems. Orange = alpha helices, blue = beta sheets. Key structural differences are indicated by arrows.

3.3.4 The protein RMSF data showed different fluctuation levels. The full agonist system overall had the highest RMSF values. The APO form overall had the lowest RMSF values, however certain regions of the biased system had lower RMSF than the APO. These regions include the extracellular loops and the transmembrane helices (Table 1). This comparison is visualized in figure 27.

Table 3

Mean values (Å) for the RMSF of each portion of the ligand-receptor complexes.

Domain	Full Ago	Apo	Biased Ago
N-terminal	3.62	2.75	2.67
TM1	2.40	1.75	1.14
I1	3.62	1.76	2.74
TM2	2.35	1.60	1.20
E1	3.02	2.19	1.83
TM3	1.57	1.46	1.33
I2	3.82	2.17	2.90
TM4	1.99	1.47	1.28
E2	2.57	1.97	1.63
TM5	2.10	1.93	1.72
I3	5.85	3.79	6.00
TM6	4.11	1.81	2.94
E3	2.40	2.56	1.98
TM7	1.88	1.67	1.09
C-terminal	3.95	2.87	3.33
All I's	4.43	2.58	3.88
All E's	2.66	2.24	1.81
All Loops	3.55	2.41	2.85
All TM's	2.34	1.67	1.53
Overall	3.02	2.12	2.26

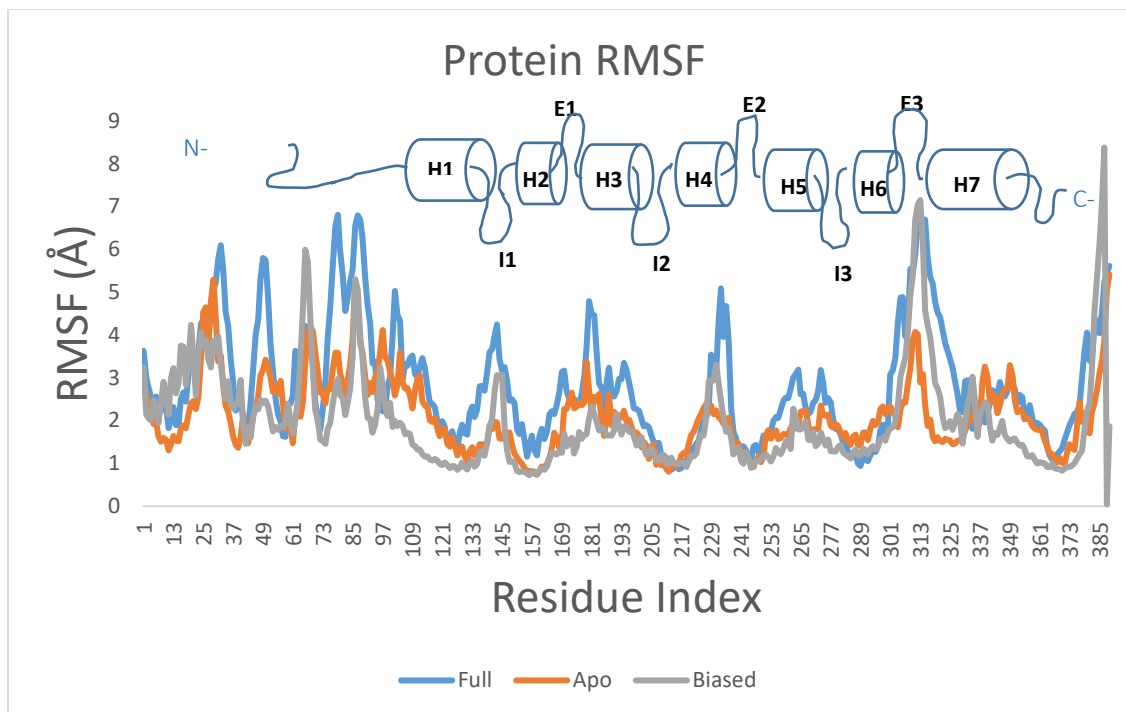


Figure 27. $C\alpha$ Root Mean Square Fluctuation (RMSF) of all three systems. Used to determine flexibility of each region throughout the simulation.

3.3.5 Comparison to solved GPCRs provided insight into biased activation.

Several solved G protein docked GPCRs (PDB IDs: 6b3j, 5uz7, 5vai, 5g53, 3sn6) were aligned and superimposed with our systems in maestro. The full agonist system adopted a very similar conformation to the solved structures. The ECD folded in the same direction, the TMD spread in a similar fashion, and the C-terminal helix moved to a similar conformation (**Figure 28**). The Biased agonist system adopted a different conformation than the solved structures. The ECD was more vertical, the TMD showed less spreading but angled out, and the C-terminal helix inverted in on itself to face the other way. With all these changes, the biased system still does not clash with the G protein (**Figure 29**).

When compared to the beta arrestin bound GPCR, the full agonist adopted similar

conformations and did not clash with the arrestin. The C-terminal helix of our biased system did clash with the beta arrestin (**Figure 30 and 31**).

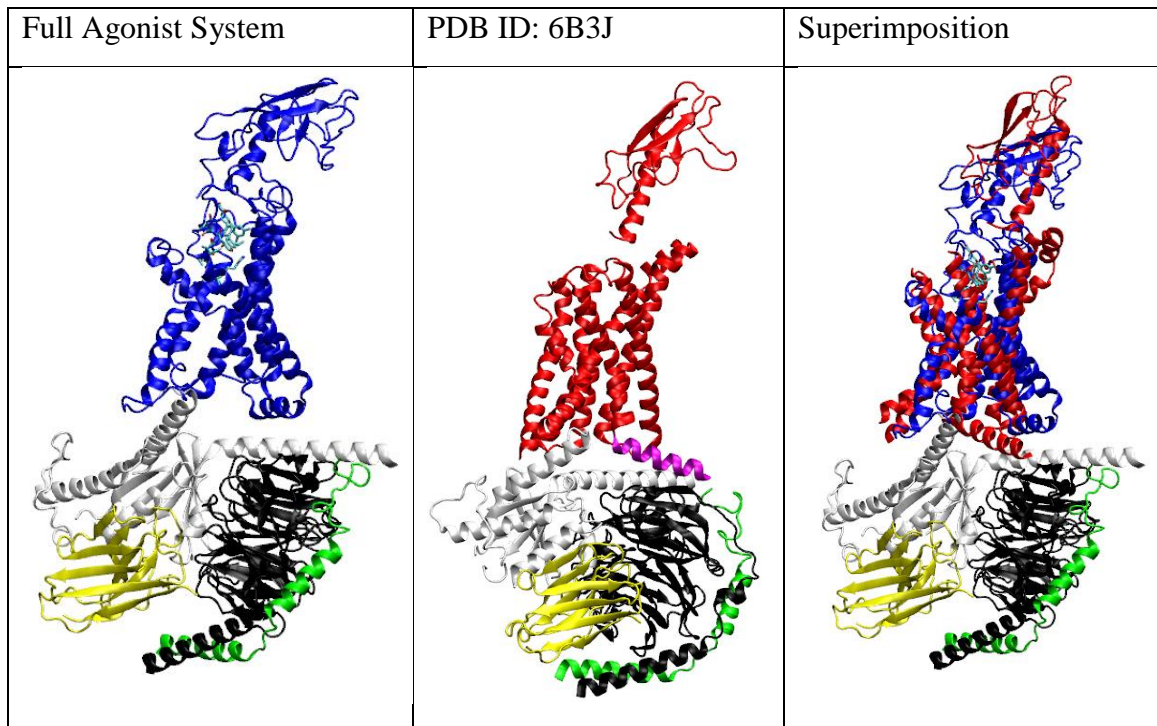


Figure 28. Comparison between our full agonist system aligned with the G-protein and the solved GLP1 receptor bound to the G-protein. Structure obtained from PDB 6B3J.

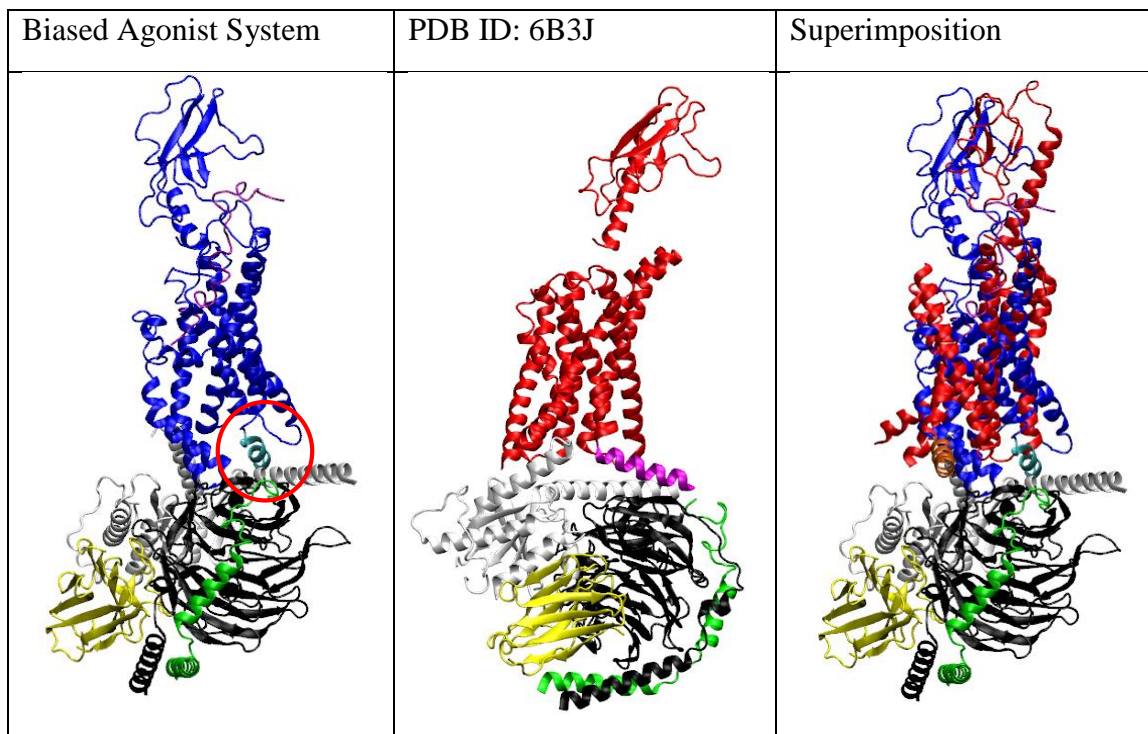


Figure 29. Comparison between our biased agonist system aligned with the G-protein and the solved GLP1 receptor bound to the G-protein. Structure obtained from PDB 6B3J. The C-terminal helix of the biased system is circled in red.

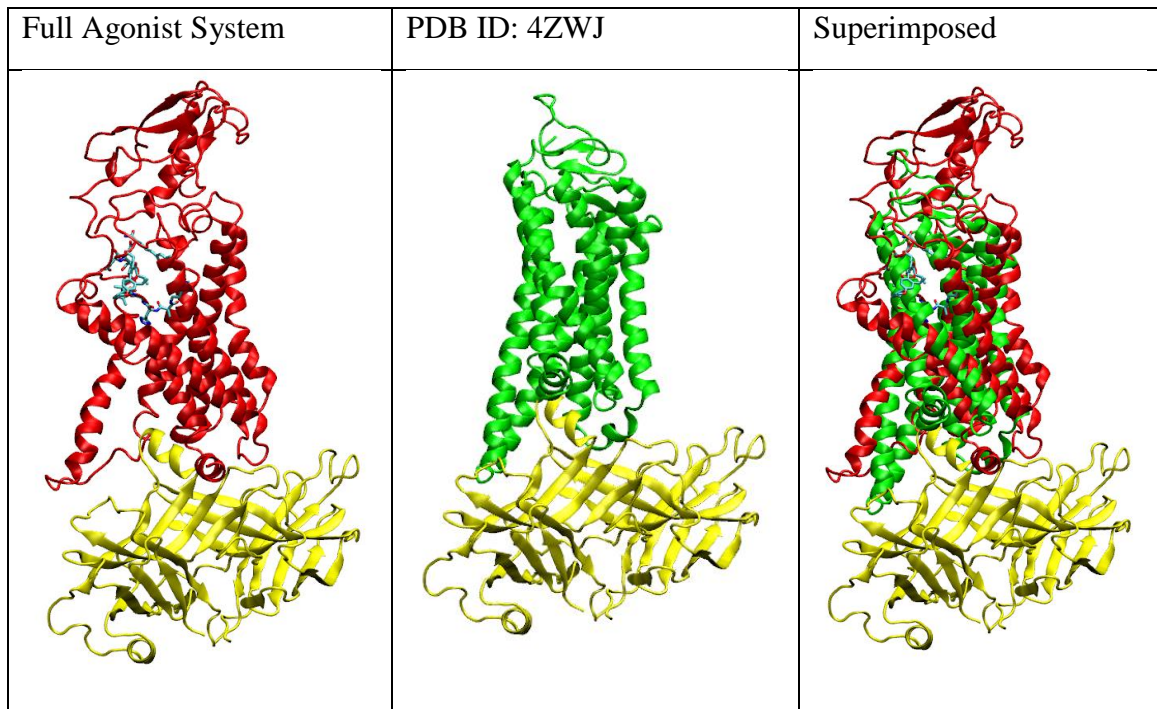


Figure 30. Comparison between our full unbiased agonist system aligned with beta arrestin and the rhodopsin receptor bound to beta arrestin. Solved structure of the rhodopsin receptor bound to beta arrestin (PDB 4zwj) obtained from PDB.

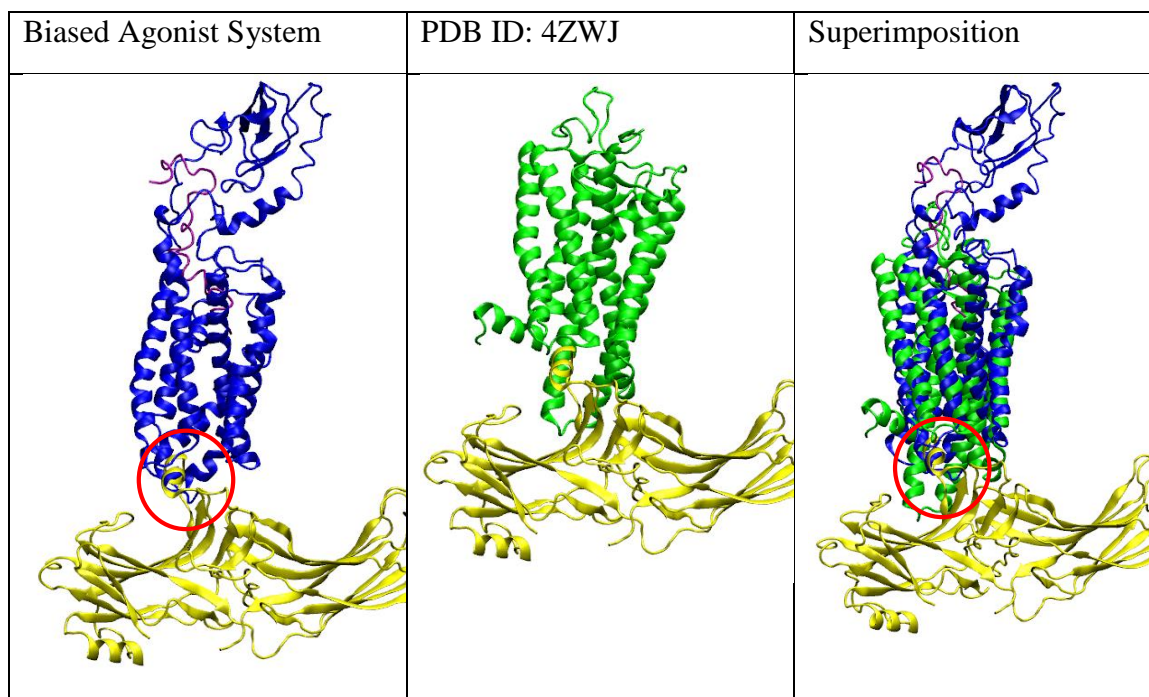


Figure 31. Comparison between our biased agonist system aligned with beta arrestin and the rhodopsin receptor bound to beta arrestin. Solved rhodopsin receptor bound to beta arrestin (4zwj) obtained from PDB. The C-terminal helix of the biased system is circled in red.

3.3.6 The biased agonist bound more stably than the full agonist. Table 2

summarizes the average MMGBSA values and standard deviations of the two systems.

Three categories were calculated from the outputted data from the MMGBSA analysis:

van der Waals (VDW), electrostatic (ELE), and hydrophobic. These values were totaled

up to give total binding energy. The biased system experienced more stable binding in

each of the categories.

Table 4

MMGBSA binding values for full agonist and biased agonist systems.

	VDW	ELE	Hydrophobic	Total
Full Ago	-121 ± 7.94	-80.1 ± 27.9	-99.6 ± 6.77	-300 ± 33.7
Biased Ago	-211 ± 15.7	-226.1 ± 51.4	-147.9 ± 9.13	-585 ± 51.6

Average MMGBSA values for both systems. Values are measured in kcal/mol. VDW = Van der Waals, ELE = electrostatic.

3.4 Discussion

The GLP1 receptor is a good drug target for type 2 diabetes because it controls insulin secretion. Individuals inflicted with type 2 diabetes are incapable of producing a sufficient amount of insulin required by their body. For this reason, agonist drugs can be used on the GLP1 receptor to raise insulin levels. However, current agonists used are full agonists, meaning they activate all signaling pathways of GLP1R. This could lead to the drug causing unwanted side effects due to not being specific for the desired pathway. A more suitable drug molecule would be a biased agonist, specifically one that is biased toward the G protein pathway. Unfortunately the mechanism of action of biased agonists are poorly understood. In this study, we aim to develop a more detailed understanding of this activation mechanism so that future drugs on this receptor will be more efficient and have less side effects. We propose that the key conformational change associated with biased agonism lies in the way the intracellular region opens to allow for G protein docking. Specifically, a conformational change in the C-terminal helix could play the most important role in biased activation.

The structure for GLP1R was obtained from the Protein Data Bank (PDB) from PDB ID 6b3j. The full agonist we used was a truncated GLP1 hormone which was obtained from PDB ID 5nx2. The biased agonist we used exendin-P5, which was known to be biased toward the G protein pathway, was obtained from PDB ID 6b3j. Ligands were docked to the receptor (Figure 22) and then run under MD simulations for 2 μ s each. An APO form (receptor only) was also examined under MD simulations for comparison purposes. Upon inspection of the most abundant conformation that resulted from these simulations, distinct conformational differences were observed (Figure 24). As expected, the full agonist system showed a wide spreading of the intracellular region to allow room for G protein docking. The APO form also behaved as expected and did not show this characteristic spreading and even appeared to close off slightly. This would indicate an inability to bind to the G protein. The biased system displayed a very interesting structure. The intracellular region formed a more compact pocket for the G protein. Most interesting is the conformation of the C-terminal helix. Instead of spreading outward like the full agonist system, it inverted and flipped directions. When we compared our structures to docked G protein complexes (Figures 28 and 29), it was revealed that these conformational differences do not clash with the G protein. It was therefore theorized that this compact opening of the intracellular region combined with the inverted C-terminal helix conformation are responsible for the specificity toward the G protein pathway. When compared to structures docked to beta arrestin proteins (Figures 30 and 31) this was confirmed showing a clash between the C-terminal helix in the biased system with the arrestin protein. The full agonist system did not show any clashes indicating it is full capable of binding.

Our RMSD data supports what we see from the MD simulations (Figure 23). For the full agonist, we see relatively constant values until the end of the simulation where we see a sharp spike. This spike is indicative of the spreading of the intracellular region. The RMSD data for the full agonist ligand shows that it remains fairly constant throughout the simulation. For the biased system, we were only able to do a protein only analysis due to limitations with Maestro's SID tool. The receptor complex fluctuates between 3 and 5 Å throughout the simulation, and becomes more rigid toward the end. The APO form as expected showed the least amount of deviation once it reached equilibrium. The tighter conformation of the biased system can be attributed to how tight the receptor binds to the ligand. There are significantly more protein-ligand interactions acting on the biased ligand than the full agonist ligand (Figure 25). This is due largely to the much greater size of the biased ligand. The MMGBSA results further confirm these conclusions. The total MMGBSA value for the biased system is almost twice as energetically favorable when compared to the full agonist which indicated greater stability (Table 4).

Our secondary structure elements graph (Figure 26) further clarifies conformational differences. The full agonist and biased agonist systems showed very similar levels of alpha helices and beta sheet formation. This was expected as both ligands are known agonists. Minor deviations can be seen in TM5 and TM6, but the most critical difference is in helix 8 (C-terminal helix). This change is attributed with the conformational flip in the biased system. Our RMSF data in figure 27 and table 3 elucidates further on the movement of the receptors. The full agonist system showed the highest overall RMSF values, which agrees with our MD simulation results since it observed the greatest conformational change. The APO form showed the overall lowest

RMSF values indicating the lowest amount of conformational change. This agrees with our MD results and was the expected outcome due to not being bound to a ligand. The biased system overall showed RMSF values in between these two systems, but important regions to note are the high RMSF values at the C-terminal region as well as intracellular loop 3. This coincides with our hypothesis that these conformational differences are critical in the activation mechanism of biased agonists.

The full agonist system was shown to adopt similar conformations to solved structures and did not clash with the G protein, indicating it is able to bind to the G protein. When compared to arrestin bound GPCRs, it is still found to not clash, indicating an ability to bind to arrestin protein as well. We can then link this general wide spreading of the receptor to correspond with full agonism. Our biased system was shown to not clash with the G protein, indicating an ability to bind to it. Its C-terminal helix however was shown to clash with the arrestin protein, indicating it does not bind to beta arrestin. The conformational change of the C-terminal helix can then be attributed to blocking arrestin binding. This agrees with information found in the literature⁴¹⁻⁴⁴. When a G protein docks to a receptor, it hangs off to one side. A docked arrestin is more centered underneath the receptor and has a helical structure pointing out the top end. Our results show that the wide spread that we observe in the full agonist system is general enough to accommodate both types of signaling molecules. Since the biased system remains more compact, this forces the signaling molecule to dock off-centered. This allows the G protein to still be able to dock, while at the same time prevents the arrestin from binding since it no longer has the room to dock directly underneath the receptor. The C-terminal helix plays the most important role in this mechanism since the inversion of the helix

allows for interaction with the G protein, but blocks the site for the arrestin protein by blocking that top helical structure of beta arrestin.

3.5 Conclusion

The GLP1 receptor is a class B GPCR whose main function is to stimulate insulin secretion. For this reason, GLP1 is a good drug target for treating type 2 diabetes, which is the condition in which the body cannot produce a sufficient amount of insulin to allow glucose to enter the cells. To treat this, agonist compounds can be used to stimulate the production and secretion of more insulin. However, most agonists are not specific to just one pathway and therefore can be associated with side effects. Special types of agonists called biased agonists are specific to one pathway, and would therefore lower the amount of side effects and improve the drug. Unfortunately, an understanding of the mechanism of activation of biased agonists remains elusive. The goal of this study was to develop a more detailed understanding of biased agonism so that improved, more specific drugs can be developed. Our studies have shown distinct conformations of the intracellular region of the receptor and the C-terminal helix that we propose to play a critical role in biased activation. We set up three systems for comparison: a full agonist, biased agonist, and APO form system. All complexes were run under MD simulations for 2 μ s and trajectory clustering analysis was performed to determine the most abundant conformation. It is here that we visually saw the conformational differences between the two systems. The full agonist spread wide enough to allow for G protein docking which also supported docking of beta arrestins indicating activation of both signaling pathways. The biased system formed a more compact conformation and observed a critical helical switch of the C-terminus. The C-terminal helix conformation allowed for G protein activation, but blocked the arrestin protein from binding indicating it only activates the G protein signaling pathway. Simulation Interaction Diagram analyses were also performed to

determine protein-ligand interactions, RMSD, RMSF, and secondary structure elements, all of which supported our hypothesis. Finally, MMGBSA analysis was performed to measure the binding affinity of the two ligands. It was determined that the biased ligand bound much more stability, further enticing the development of drugs of this class. Future studies can test other receptors and other ligands to confirm that these conformations are conserved. Our findings provide insight into developing a new, more efficient class of drugs on the GLP1 receptor.

References

1. Association, C. B. R., New Drug Development Process. **2019**.
2. Macalino, S. J. Y.; Gosu, V.; Hong, S.; Choi, S., Role of computer-aided drug design in modern drug discovery. *Archives of Pharmacal Research* **2015**, *38* (9), 1686-1701.
3. Madhavi Sastry, G.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W., Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *Journal of Computer - Aided Molecular Design* **2013**, *27* (3), 221-34.
4. Luscombe, N. M.; Greenbaum, D.; Gerstein, M., What is bioinformatics? A proposed definition and overview of the field. *Methods of information in medicine* **2001**, *40* (4), 346-58.
5. Westbrook, J.; Feng, Z.; Chen, L.; Yang, H.; Berman, H. M., The Protein Data Bank and structural genomics. *Nucleic acids research* **2003**, *31* (1), 489-491.
6. Withana-Gamage, T. S.; Hegedus, D. D.; Qiu, X.; Wanasundara, J. P. D., In Silico Homology Modeling To Predict Functional Properties of Cruciferin. *Journal of Agricultural and Food Chemistry* **2011**, *59* (24), 12925-12938.
7. Sergio, F. S.; Nuno, M. F. S. A. C.; Pedro, A. F.; Maria Joao, R., Virtual Screening in Drug Design and Development. *Combinatorial Chemistry & High Throughput Screening* **2010**, *13* (5), 442-453.
8. Sousa, S. F.; Ribeiro, A. J. M.; Coimbra, J. T. S.; Neves, R. P. P.; Martins, S. A.; Moorthy, N. S. H. N.; Fernandes, P. A.; Ramos, M. J., Protein-Ligand Docking in the New Millennium – A Retrospective of 10 Years in the Field. *Current medicinal chemistry* **2013**, *20* (18), 2296-2314.
9. Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrín, P. C.; Mainz, D. T., Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *Journal of Medicinal Chemistry* **2006**, *49* (21), 6177-6196.
10. Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S., Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of Medicinal Chemistry* **2004**, *47* (7), 1739-1749.
11. Du, X.; Li, Y.; Xia, Y.-L.; Ai, S.-M.; Liang, J.; Sang, P.; Ji, X.-L.; Liu, S.-Q., Insights into Protein-Ligand Interactions: Mechanisms, Models, and Methods. *Int J Mol Sci* **2016**, *17* (2), 144.
12. McCammon, J. A.; Gelin, B. R.; Karplus, M., Dynamics of folded proteins. *Nature* **1977**, *267* (5612), 585-590.
13. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A., A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *Journal of the American Chemical Society* **1995**, *117* (19), 5179-5197.
14. Kalé, L.; Skeel, R.; Bhandarkar, M.; Brunner, R.; Gursoy, A.; Krawetz, N.; Phillips, J.; Shinozaki, A.; Varadarajan, K.; Schulten, K., NAMD2: Greater Scalability for Parallel Molecular Dynamics. *Journal of Computational Physics* **1999**, *151* (1), 283-312.
15. Reza Kalani, M.; Tajkhorshid, E., Molecular Dynamics: The Computational Molecular Microscope. *Razavi Int J Med* **2014**, *2* (3), e20117.

16. Kollman, P., Free energy calculations: Applications to chemical and biochemical phenomena. *Chemical Reviews* **1993**, *93* (7), 2395-2417.
17. Hou, T.; Wang, J.; Li, Y.; Wang, W., Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *Journal of chemical information and modeling* **2011**, *51* (1), 69-82.
18. Kobilka, B. K.; Deupi, X., Conformational complexity of G-protein-coupled receptors. *Trends in Pharmacological Sciences* **2007**, *28* (8), 397-406.
19. Pal, K.; Melcher, K.; Xu, H. E., Structure and mechanism for recognition of peptide hormones by Class B G-protein-coupled receptors. *Acta Pharmacologica Sinica* **2012**, *33* (3), 300-311.
20. Seidal, L.; Zarzycka, B.; Zaidi, S. A.; Katritch, V.; Coin, I., Structural insight into the activation of a class B G-protein-coupled receptor by peptide hormones in live human cells. *Elife* **2017**, *6*, 25.
21. Kean, J.; Bortolato, A.; Hollenstein, K.; Marshall, F. H.; Jazayeri, A., Conformational thermostabilisation of corticotropin releasing factor receptor 1. *Scientific Reports* **2015**, *5*, 11.
22. Teleb, M.; Kuppast, B.; Spyridaki, K.; Liapakis, G.; Fahmy, H., Synthesis of 2-imino and 2-hydrazono thiazolo[4,5-d]pyrimidines as corticotropin releasing factor (CRF) antagonists. *European Journal of Medicinal Chemistry* **2017**, *138* (Supplement C), 900-908.
23. Pándy-Szekeres, G.; Munk, C.; Tsonkov, T. M.; Mordalski, S.; Harpsøe, K.; Hauser, A. S.; Bojarski, A. J.; Gloriam, D. E., GPCRdb in 2018: adding GPCR structure models and ligands. *Nucleic acids research* **2017**, *46* (D1), D440-D446.
24. Isberg, V.; de Graaf, C.; Bortolato, A.; Cherezov, V.; Katritch, V.; Marshall, F. H.; Mordalski, S.; Pin, J.-P.; Stevens, R. C.; Vriend, G.; Gloriam, D. E., Generic GPCR residue numbers – aligning topology maps while minding the gaps. *Trends in Pharmacological Sciences* **2015**, *36* (1), 22-31.
25. Xu, J. L.; Wang, Z. H.; Liu, P.; Li, D. M.; Lin, J. P., An insight into antagonist binding and induced conformational dynamics of class B GPCR corticotropin-releasing factor receptor 1. *Mol. Biosyst.* **2015**, *11* (7), 2042-2050.
26. Zhang, Y., I-TASSER server for protein 3D structure prediction. *Bmc Bioinformatics* **2008**, *9*.
27. Sastry, G. M.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W., Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J Comput Aided Mol Des* **2013**, *27* (3), 221-34.
28. Harder, E.; Damm, W.; Maple, J.; Wu, C.; Reboul, M.; Xiang, J. Y.; Wang, L.; Lupyan, D.; Dahlgren, M. K.; Knight, J. L.; Kaus, J. W.; Cerutti, D. S.; Krilov, G.; Jorgensen, W. L.; Abel, R.; Friesner, R. A., OPLS3: A Force Field Providing Broad Coverage of Drug-like Small Molecules and Proteins. *Journal of Chemical Theory and Computation* **2016**, *12* (1), 281-296.
29. Shivakumar, D.; Williams, J.; Wu, Y.; Damm, W.; Shelley, J.; Sherman, W., Prediction of Absolute Solvation Free Energies using Molecular Dynamics Free Energy Perturbation and the OPLS Force Field. *Journal of Chemical Theory and Computation* **2010**, *6* (5), 1509-1519.
30. Jorgensen, W. L.; Maxwell, D. S.; TiradoRives, J., Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society* **1996**, *118* (45), 11225-11236.

31. Grace, C. R.; Perrin, M. H.; Cattle, J. P.; Vale, W. W.; Rivier, J. E.; Riek, R., Common and divergent structural features of a series of corticotropin releasing factor-related peptides. *J Am Chem Soc* **2007**, *129* (51), 16102-14.
32. Grace, C. R.; Perrin, M. H.; Gulyas, J.; Digruccio, M. R.; Cattle, J. P.; Rivier, J. E.; Vale, W. W.; Riek, R., Structure of the N-terminal domain of a type B1 G protein-coupled receptor in complex with a peptide ligand. *Proceedings of the National Academy of Sciences of the United States of America* **2007**, *104* (12), 4858-63.
33. In *Proceedings of the 2006 ACM/IEEE conference on Supercomputing*, Tampa, Florida, ACM: Tampa, Florida, 2006; p 746.
34. Rasmussen, S. G.; DeVree, B. T.; Zou, Y.; Kruse, A. C.; Chung, K. Y.; Kobilka, T. S.; Thian, F. S.; Chae, P. S.; Pardon, E.; Calinski, D.; Mathiesen, J. M.; Shah, S. T.; Lyons, J. A.; Caffrey, M.; Gellman, S. H.; Steyaert, J.; Skiniotis, G.; Weis, W. I.; Sunahara, R. K.; Kobilka, B. K., Crystal structure of the beta2 adrenergic receptor-Gs protein complex. *Nature* **2011**, *477* (7366), 549-55.
35. Carpenter, B.; Nehme, R.; Warne, T.; Leslie, A. G.; Tate, C. G., Structure of the adenosine A(2A) receptor bound to an engineered G protein. *Nature* **2016**, *536* (7614), 104-7.
36. Zhang, Y.; Sun, B.; Feng, D.; Hu, H.; Chu, M.; Qu, Q.; Tarrasch, J. T.; Li, S.; Sun Kobilka, T.; Kobilka, B. K.; Skiniotis, G., Cryo-EM structure of the activated GLP-1 receptor in complex with a G protein. *Nature* **2017**, *546* (7657), 248-253.
37. Liang, Y. L.; Khoshouei, M.; Radjainia, M.; Zhang, Y.; Glukhova, A.; Tarrasch, J.; Thal, D. M.; Furness, S. G. B.; Christopoulos, G.; Coudrat, T.; Danev, R.; Baumeister, W.; Miller, L. J.; Christopoulos, A.; Kobilka, B. K.; Wootten, D.; Skiniotis, G.; Sexton, P. M., Phase-plate cryo-EM structure of a class B GPCR-G-protein complex. *Nature* **2017**, *546* (7656), 118-123.
38. Liang, Y. L.; Khoshouei, M.; Glukhova, A.; Furness, S. G. B.; Zhao, P.; Clydesdale, L.; Koole, C.; Truong, T. T.; Thal, D. M.; Lei, S.; Radjainia, M.; Danev, R.; Baumeister, W.; Wang, M. W.; Miller, L. J.; Christopoulos, A.; Sexton, P. M.; Wootten, D., Phase-plate cryo-EM structure of a biased agonist-bound human GLP-1 receptor-Gs complex. *Nature* **2018**, *555* (7694), 121-125.
39. Jazayeri, A.; Rappas, M.; Brown, A. J. H.; Kean, J.; Errey, J. C.; Robertson, N. J.; Fiez-Vandal, C.; Andrews, S. P.; Congreve, M.; Bortolato, A.; Mason, J. S.; Baig, A. H.; Teobald, I.; Dore, A. S.; Weir, M.; Cooke, R. M.; Marshall, F. H., Crystal structure of the GLP-1 receptor bound to a peptide agonist. *Nature* **2017**, *546* (7657), 254-258.
40. Dunlop, B. W.; Binder, E. B.; Iosifescu, D.; Mathew, S. J.; Neylan, T. C.; Pape, J. C.; Carrillo-Roa, T.; Green, C.; Kinkead, B.; Grigoriadis, D.; Rothbaum, B. O.; Nemeroff, C. B.; Mayberg, H. S., Corticotropin-Releasing Factor Receptor 1 Antagonism Is Ineffective for Women With Posttraumatic Stress Disorder. *Biological Psychiatry* **2017**, *82* (12), 866-874.
41. Kang, Y.; Zhou, X. E.; Gao, X.; He, Y.; Liu, W.; Ishchenko, A.; Barty, A.; White, T. A.; Yefanov, O.; Han, G. W.; Xu, Q.; de Waal, P. W.; Ke, J.; Tan, M. H.; Zhang, C.; Moeller, A.; West, G. M.; Pascal, B. D.; Van Eps, N.; Caro, L. N.; Vishnivetskiy, S. A.; Lee, R. J.; Suino-Powell, K. M.; Gu, X.; Pal, K.; Ma, J.; Zhi, X.; Boutet, S.; Williams, G. J.; Messerschmidt, M.; Gati, C.; Zatsepin, N. A.; Wang, D.; James, D.; Basu, S.; Roy-Chowdhury, S.; Conrad, C. E.; Coe, J.; Liu, H.; Lisova, S.; Kupitz, C.; Grotjohann, I.; Fromme, R.; Jiang, Y.; Tan, M.; Yang, H.; Li, J.; Wang, M.; Zheng, Z.; Li, D.; Howe, N.; Zhao, Y.; Standfuss, J.; Diederichs, K.; Dong, Y.; Potter, C. S.; Carragher, B.; Caffrey, M.; Jiang, H.; Chapman, H. N.; Spence, J. C.; Fromme, P.; Weierstall, U.; Ernst, O. P.; Katritch, V.; Gurevich, V. V.; Griffin, P. R.; Hubbell, W. L.; Stevens, R. C.; Cherezov, V.; Melcher, K.; Xu, H. E., Crystal structure of

- rhodopsin bound to arrestin by femtosecond X-ray laser. *Nature* **2015**, 523 (7562), 561-7.
42. Scheerer, P.; Sommer, M. E., Structural mechanism of arrestin activation. *Current Opinion in Structural Biology* **2017**, 45, 160-169.
 43. Gurevich, V. V.; Gurevich, E. V., Molecular Mechanisms of GPCR Signaling: A Structural Perspective. *International journal of molecular sciences* **2017**, 18 (12), 2519.
 44. Jean-Charles, P.-Y.; Kaur, S.; Shenoy, S. K., G Protein-Coupled Receptor Signaling Through β -Arrestin-Dependent Mechanisms. *Journal of cardiovascular pharmacology* **2017**, 70 (3), 142-158.

Appendix A

Binding of Peptide Agonist Urocortin and Small Molecule Antagonist CP 376395 to the CRF1 Receptor Probed by Molecular Dynamics Simulations

Table A1

Templates used by GPCR-I-TASSER

Rank	PDB Hit	Chain	GPCR	Iden 1	Iden 2	Cov.	Norm Z-score
1	5vai	R	Activated Glucagon-like peptide-1 receptor	0.28	0.29	0.84	1.95
2	5vai	A	Activated Glucagon-like peptide-1 receptor	0.28	0.29	0.83	3.91
3	5vai	R	Activated Glucagon-like peptide-1 receptor	0.27	0.29	0.83	2.86
4	4l6r	A	Class B human glucagon GPCR	0.28	0.27	0.82	3.02
5	4k5y	A	CRF1R in complex with antagonist	0.84	0.54	0.56	3.20
6	4l6r	A	Class B human glucagon GPCR	0.23	0.27	0.89	3.38
7	4l6r	A	Class B human glucagon GPCR	0.27	0.27	0.81	4.29
8	4l6r	A	Class B human glucagon GPCR	0.33	0.23	0.65	2.56
9	5vai	9	Activated Glucagon-like peptide-1 receptor	0.28	0.29	0.84	2.54
10	5nx2	A	GLP-1R with a truncated peptide agonist	0.27	0.26	0.84	1.79

(a) Rank of template represents top ten threading templates used by GPCR-I-TASSER

(b) Iden1 is the percentage sequence identity of the templates in the threading aligned region with the query sequence.

(c) Iden2 is the percentage sequence identity of the whole template chains with query sequence.

(d) GPCR is the type of G-protein-coupled-receptor correlating to the PDB ID.

(e) Cov. Represents the coverage of the threading alignment and is equal to the number of aligned residues divided by the length of query protein.

(f) Norm. Z-score is the normalized Z-score of the threading alignments. Alignment with a Normalized Z-score > 1 means a good alignment and vice versa.

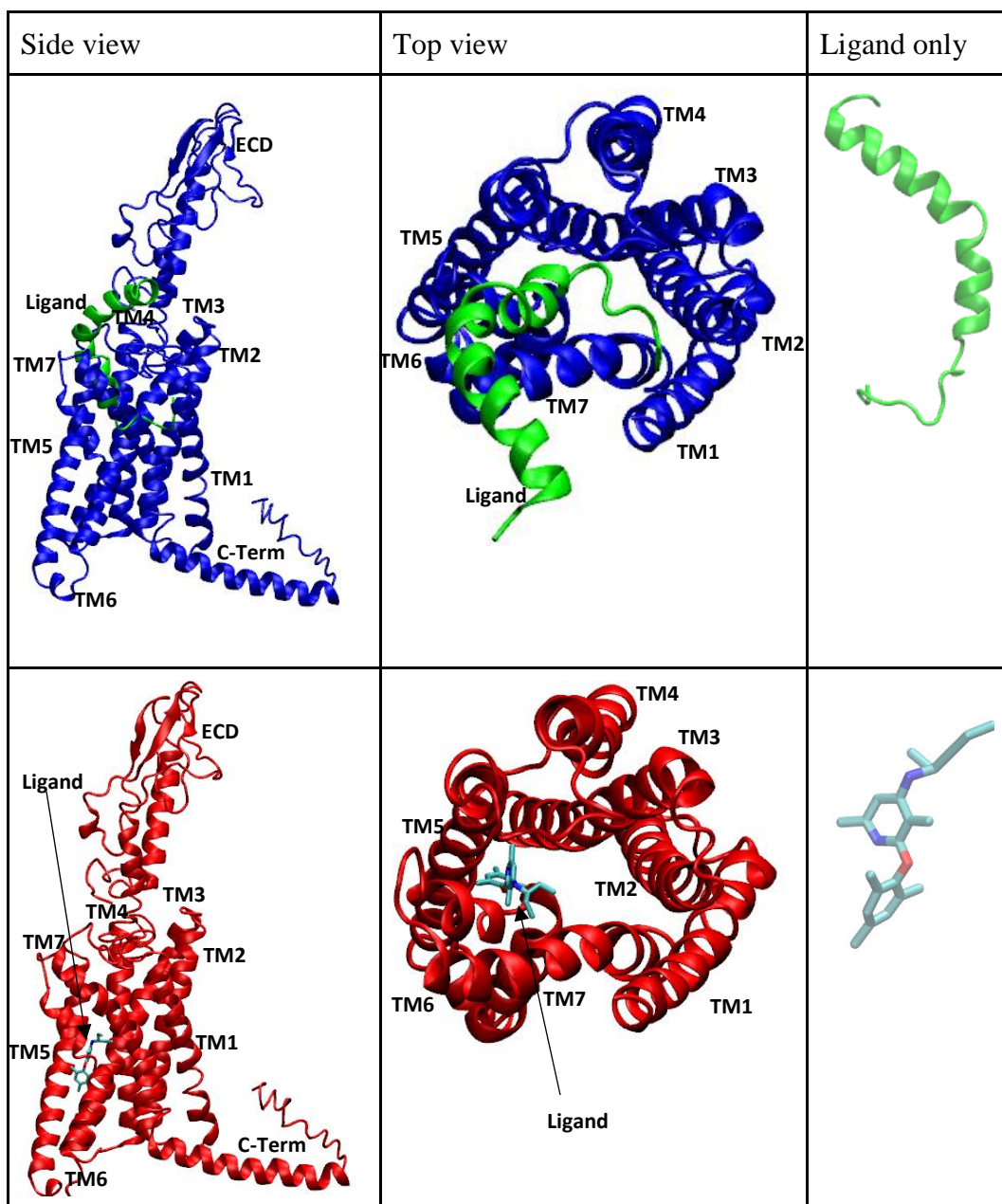


Figure A1. Docked peptide agonist urocortin and small molecule antagonist CP 376395 into the homology model generated by I-TASSER.


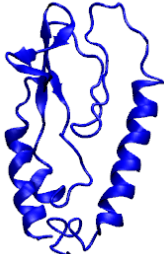
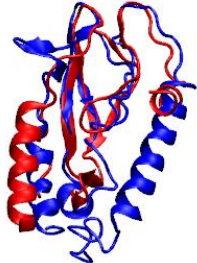

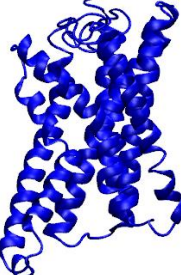
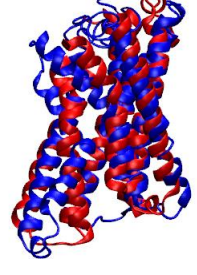

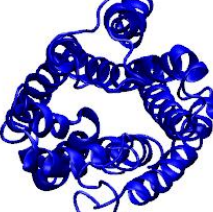
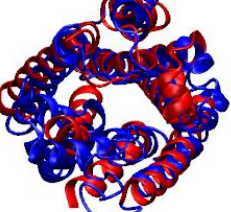
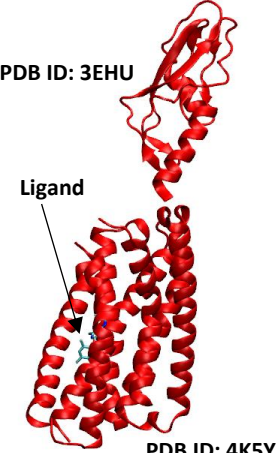
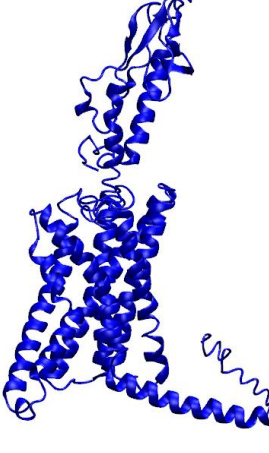
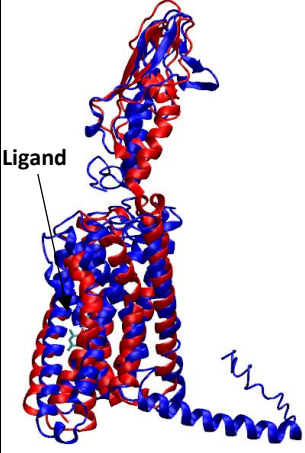
	Crystal Structure	TASSER Model	Superimposition
Extracellular Domain	 PDB ID: 3EHU		
Transmembrane Domain (side view)	 PDB ID: 4K5Y		
Transmembrane Domain (top view)	 PDB ID: 4K5Y		
Full System	 PDB ID: 3EHU Ligand PDB ID: 4K5Y		 Ligand

Figure A2. Pairwise comparison between crystal structures of CRF1R (ECD from PDB 3EHU and TMD from PDB 4K5Y) and our generated model for the CRF1 receptor from TASSER.

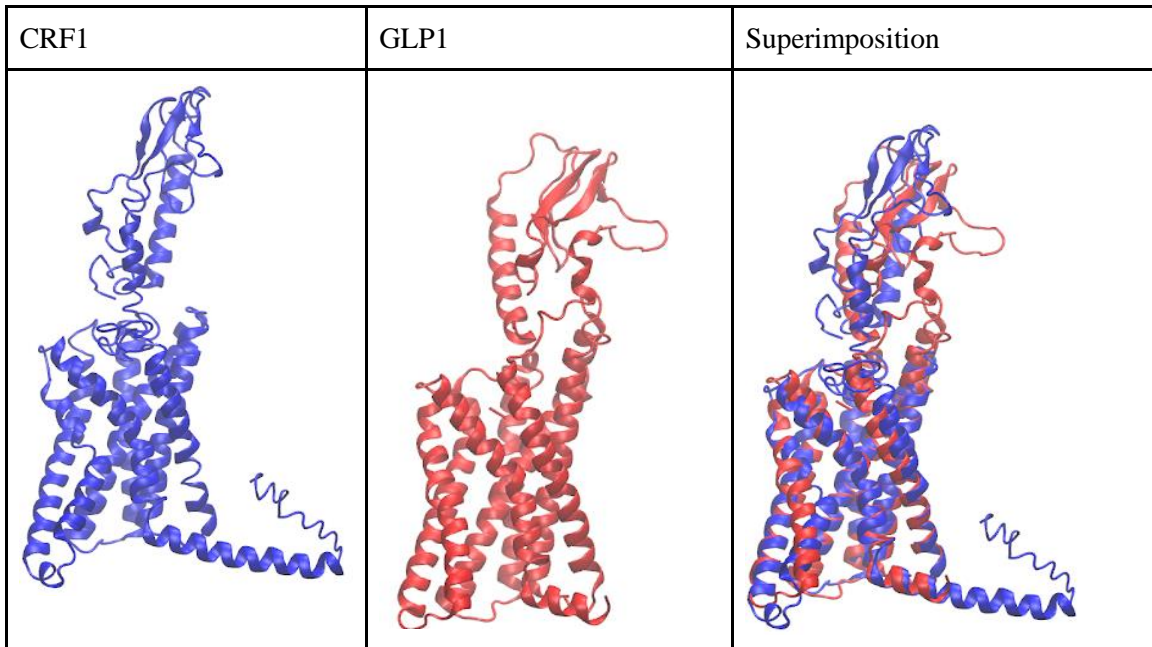


Figure A3. Structural similarity comparison between GLP1 structure (PDB 6B3J) obtained from the protein data bank and our generated model for the CRF1 receptor.

	Side view	Top view
CRF1R Peptide Agonist		
GLP1 with truncated peptide		

Figure A4. Comparison of CRF1R transmembrane domain bound to a peptide agonist and antagonist with GLP1 bound to a truncated peptide agonist. Peptides are denoted in yellow. The green tips on the CRF peptides denote the C-terminal end of the peptide. The blue tips denote the N-terminal end. The GLP1 truncated peptide also displays its side chains.

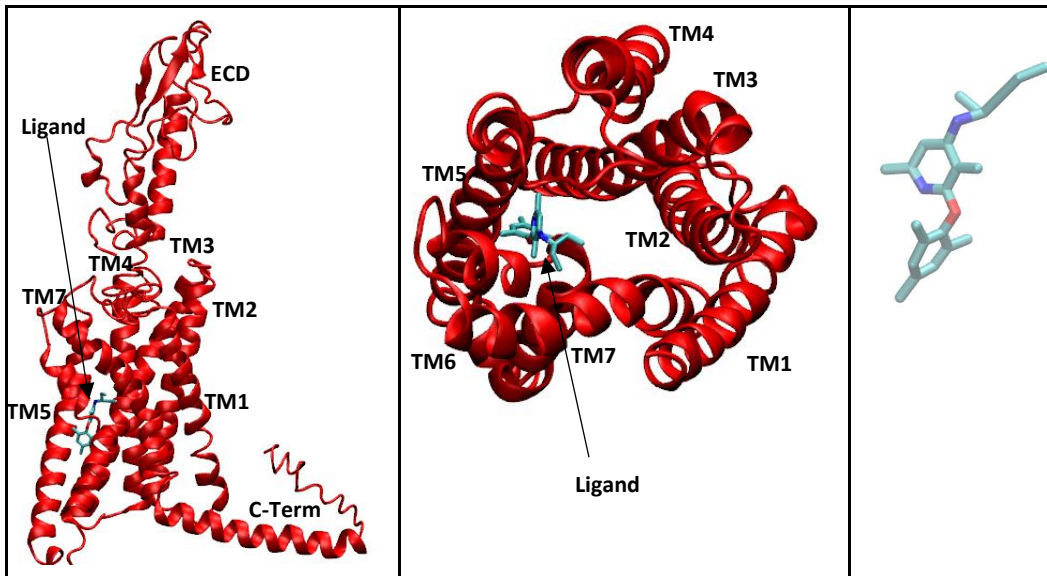


Figure A5. Docked small molecule CP 376395 antagonist with CRF1R obtained from PDB 4k5y.

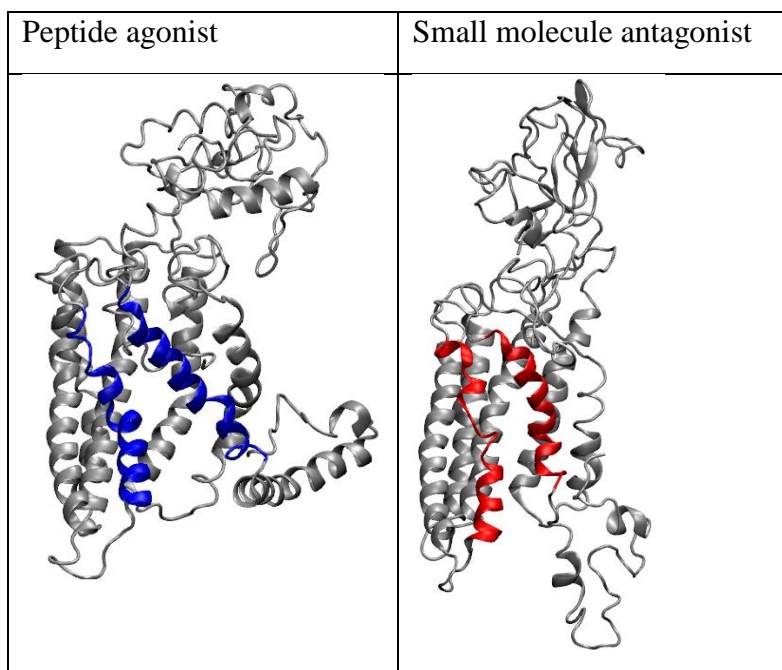


Figure A6. Comparison between wide pose observed in agonist systems and compact pose observed in antagonist systems. Highlighted are TM 6 and 7 to examine whether the receptor adopts a wide or compact pose.

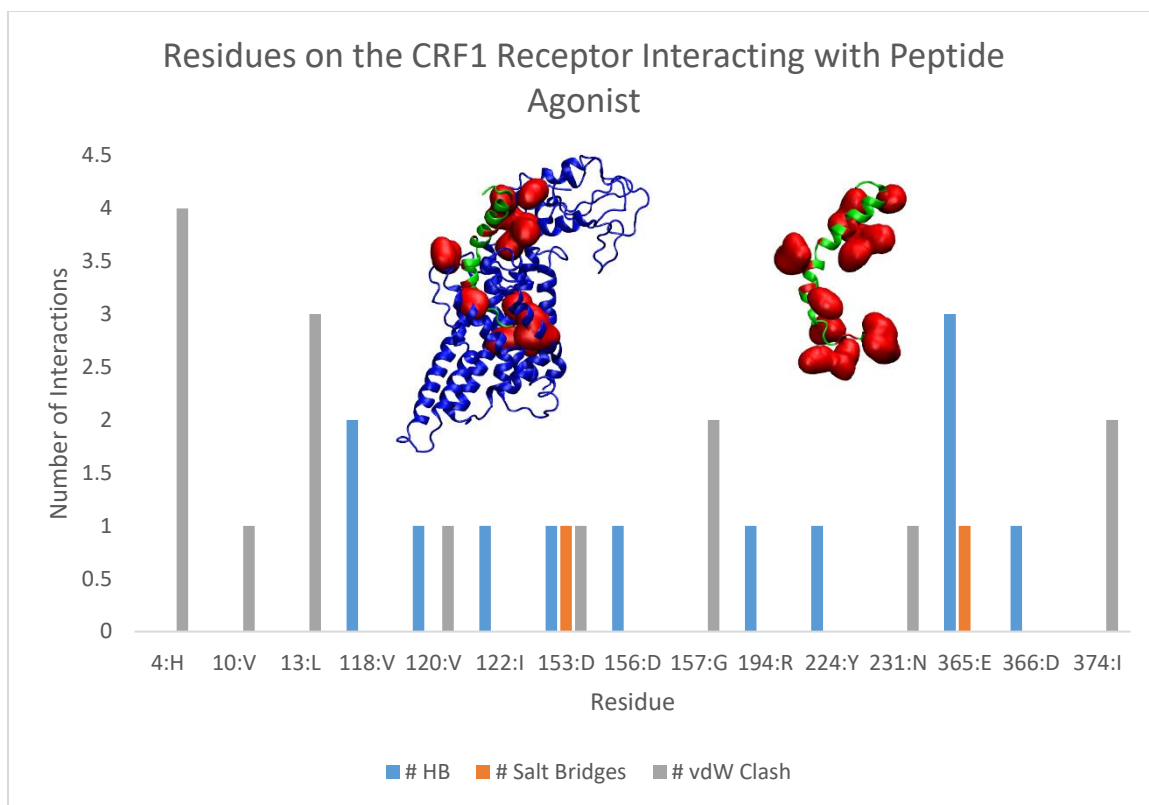


Figure A7. Ligand-protein interactions for peptide agonist system. Residues listed on the graph are the residues located on the receptor. Residues that interact with the peptide are depicted by the red structures.

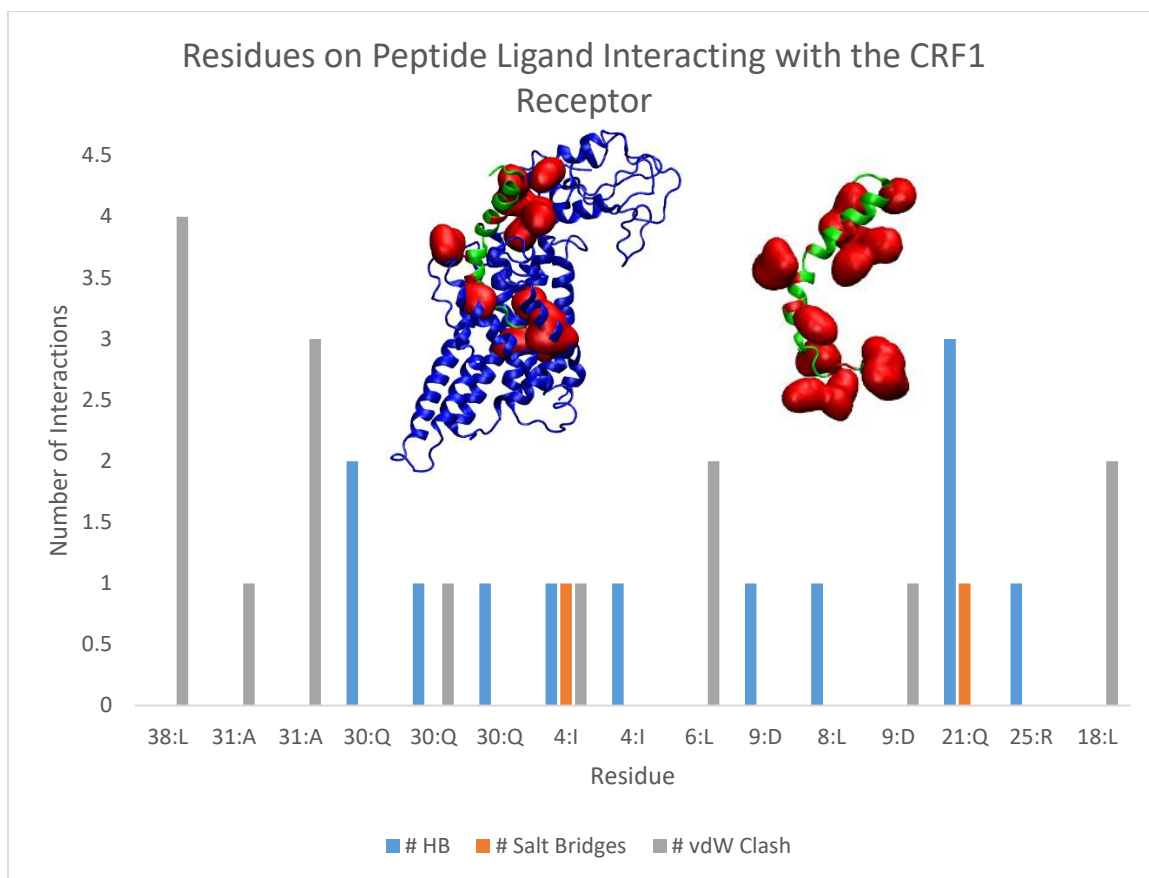


Figure A8. Ligand-protein interactions for peptide agonist system. Residues listed on the graph are the residues located on the ligand. Residues that interact with the peptide are depicted by the red structures.

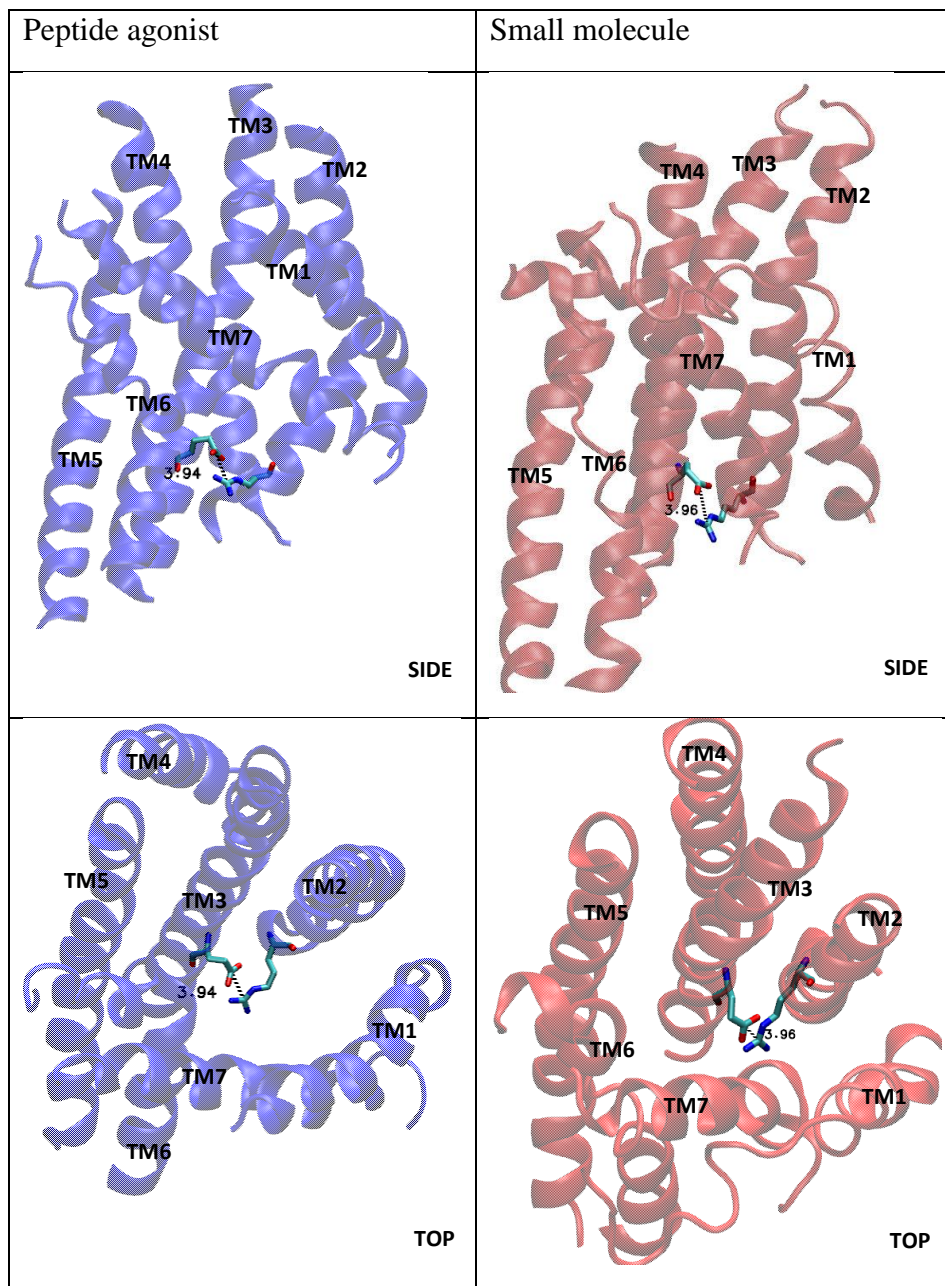


Figure A9. Ionic lock location in reference to the transmembrane domain in peptide agonist and small molecule antagonist systems after **2000ns** for the most abundant conformation. Images were taken from the first simulation trajectory. The ionic lock is located on TM2 and TM3 at residues Arg151 and Glu209. Peptide Agonist distance = 3.94. Small molecule distance = 3.96.

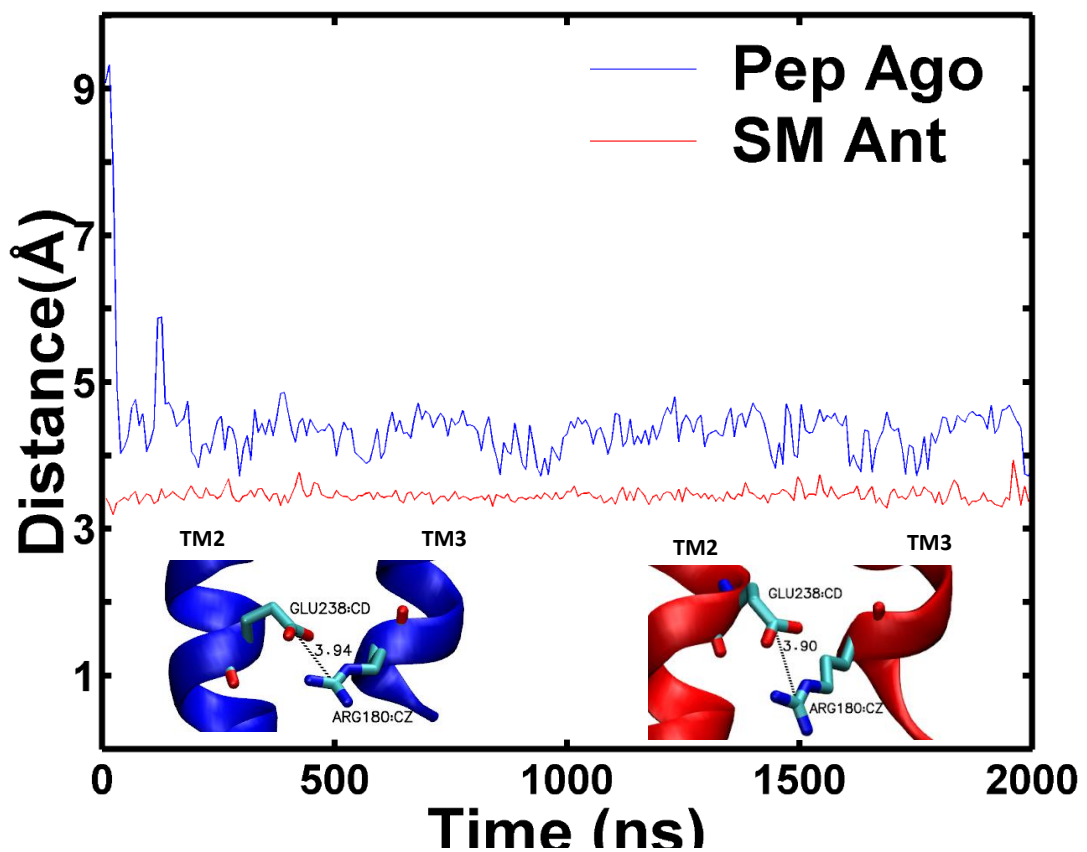


Figure A10. Ionic lock distances in small molecule and peptide agonist after 2000ns for the first simulation trajectory used for a preliminary analysis. Average distance of the ionic lock in the agonist was 3.94Å. Average distance for the antagonist was 3.90 Å.

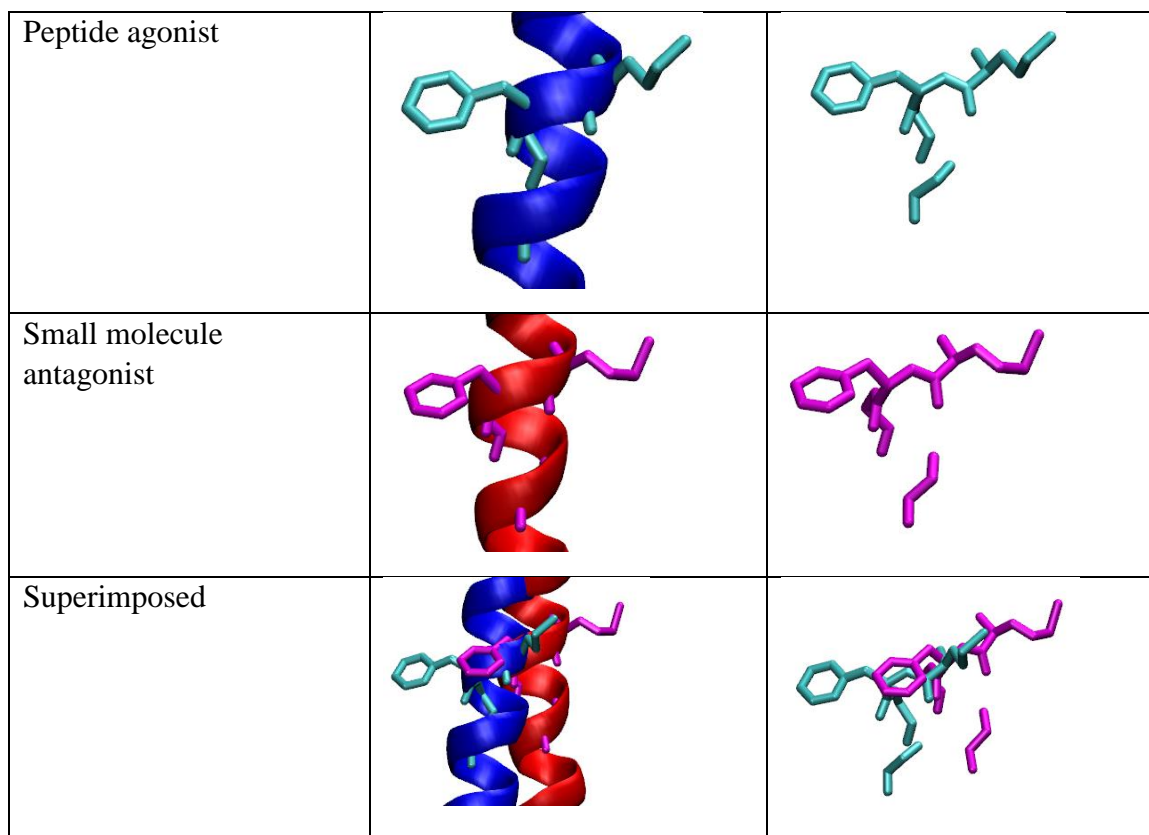


Figure A11. Comparing the GWGxP conserved motif in the agonist and antagonist systems for structural differences. It is located on TM4 between residues G235, W236, G237, and P239.

Table A2

GPCR PDB IDs

PDB ID	GPCR
3sn6	Adrenergic receptor
5g53	Adenosine A2A receptor
5uz7	Calcitonin receptor
5vai	Glucagon-like peptide-1 receptor
6b3j	Glucagon-like peptide-1 receptor

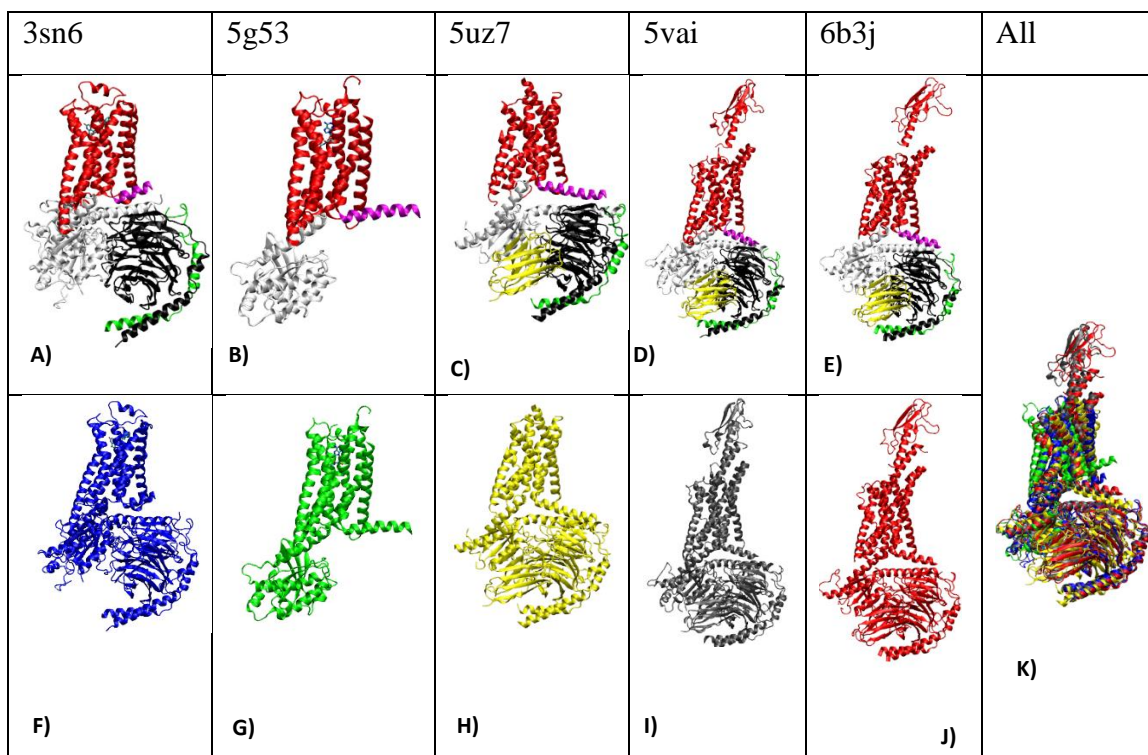


Figure A12. Comparison of different G-protein docked gpcrs. **A-E)** Red = transmembrane domain, magenta = C-terminal helix, white = alpha subunit of g protein, black = beta subunit of g protein, green = gamma subunit of g protein, yellow = nanobody. **F-K)** Red = 6b3j, blue = 3sn6, green = 5g53, yellow = 5uz7, gray = 5vai.

Table A3

Color code for figures A13 through A24

Color	Domain
Blue	Our agonist/antagonist transmembrane domain
Cyan	Our agonist/antagonist C-terminal helix
Orange	Peptide agonist ligand
Red	GPCR transmembrane domain
Magenta	GPCR C-terminal helix
White	Alpha subunit of G protein
Black	Beta subunit of G protein
Green	Gamma subunit of G protein
Yellow	Nanobody

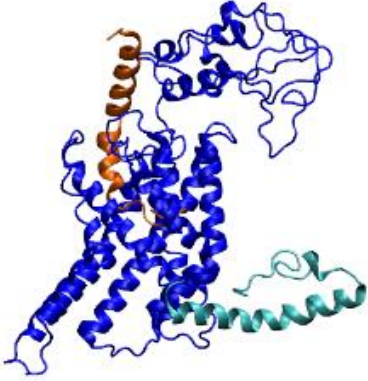
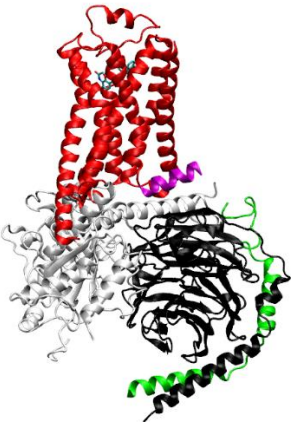
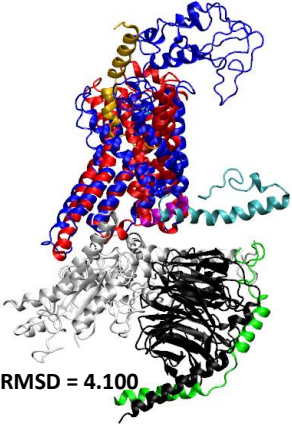


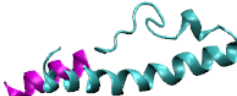
Our Peptide Agonist System	PDB ID: 3SN6	Superimposition
		
		

Figure A13. Comparison between our peptide agonist complex and the solved structure of a G protein docked Adrenergic receptor obtained from PDB. Top panels show the whole receptor complex, bottom panels show the C-terminal helix.

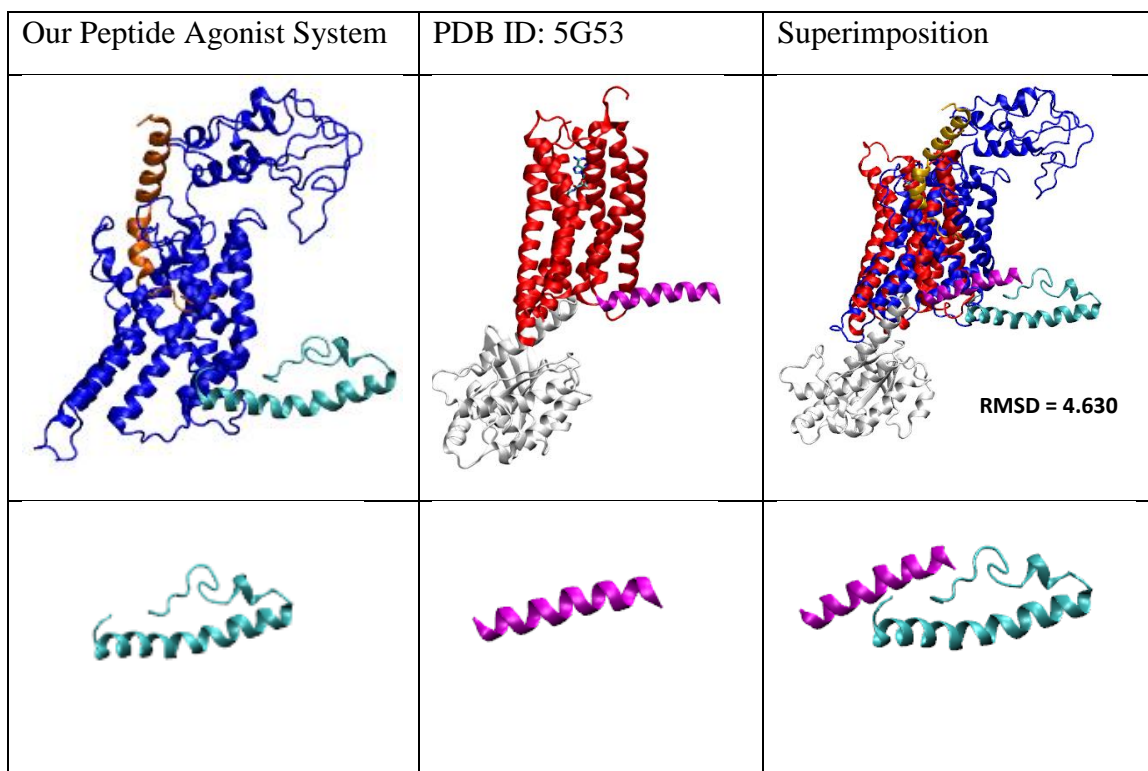


Figure A14. Comparison between our peptide agonist complex and the solved structure of a G protein docked Adenosine A2A receptor obtained from PDB. Top panels show the whole receptor complex, bottom panels show the C-terminal helix.

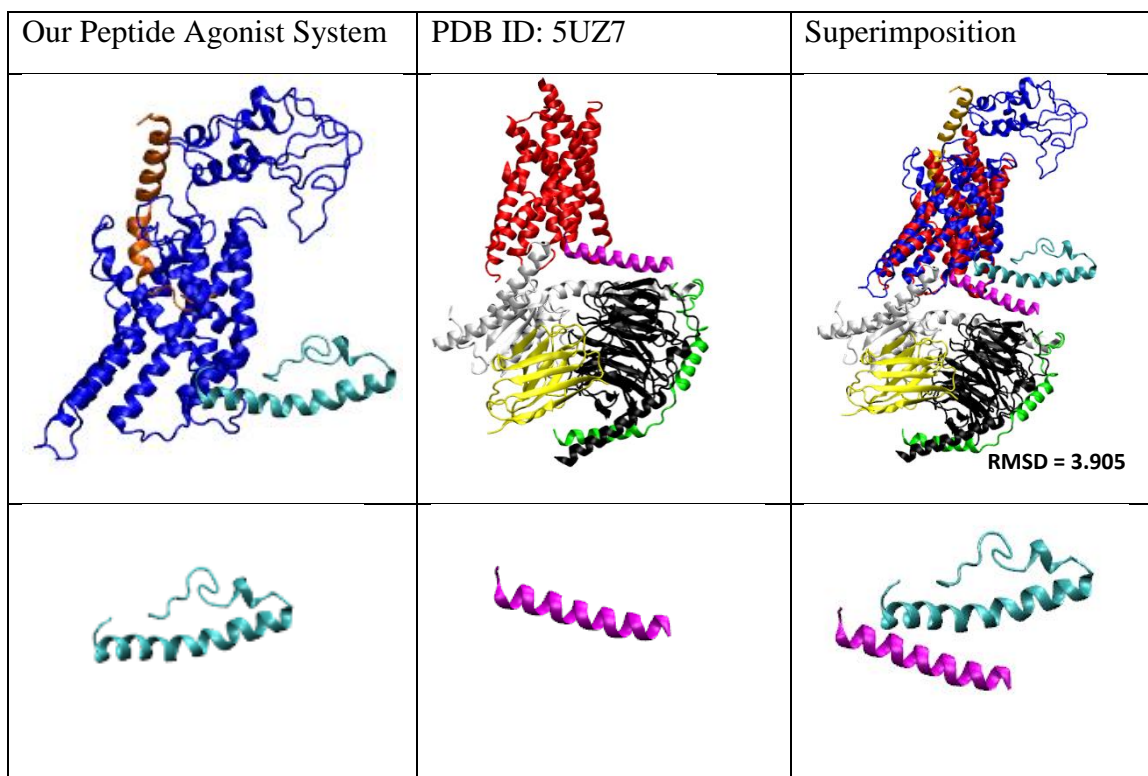


Figure A15. Comparison between our peptide agonist complex and the solved structure of G protein docked Calcitonin receptor obtained from PDB. Top panels show the whole receptor complex, bottom panels show the C-terminal helix.

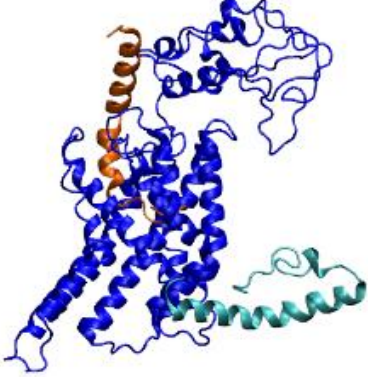
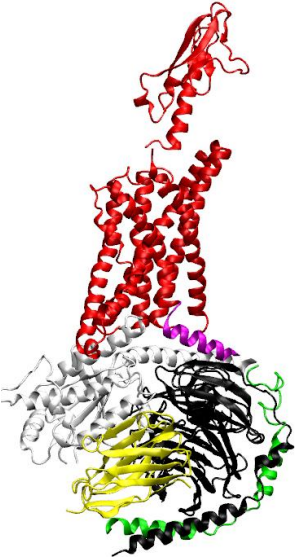
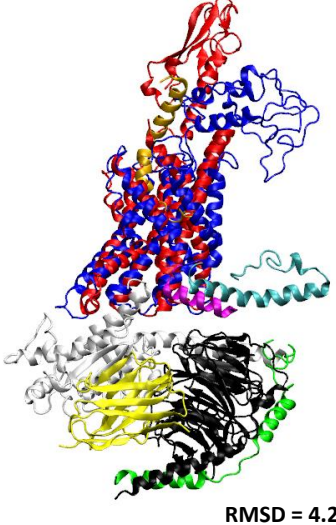
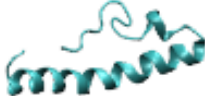

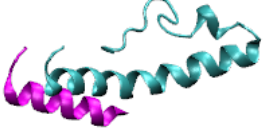
Our Peptide Agonist System	PDB ID: 5VAI	Superimposition
		
		

Figure A16. Comparison between our peptide agonist complex and the solved structure of G protein docked Glucagon-like peptide-1 receptor (5vai) obtained from PDB. Top panels show the whole receptor complex, bottom panels show the C-terminal helix.

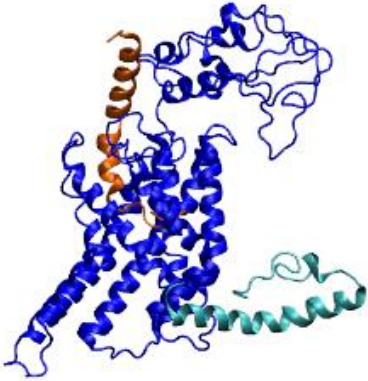
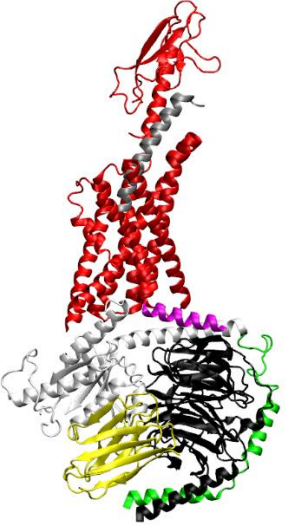
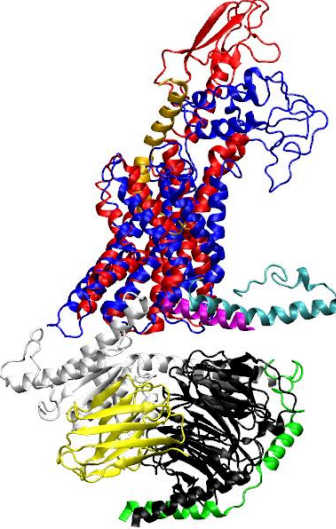
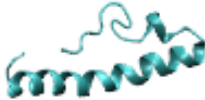

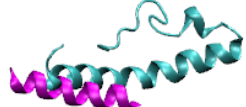
Our Peptide Agonist System	PDB ID: 6B3J	Superimposition
		 <p data-bbox="1203 863 1349 884">RMSD = 3.971</p>
		

Figure A17. Comparison between our peptide agonist complex and the solved structure of G protein docked Glucagon-like peptide-1 receptor (6b3j) obtained from PDB. Top panels show the whole receptor complex, bottom panels show the C-terminal helix.

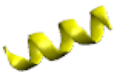
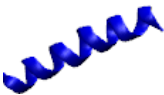

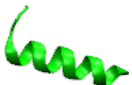

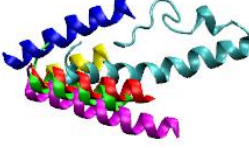
3sn6	5g53	5uz7	5vai	6b3j	All
					

Figure A18. Comparison of the C-terminal helices of the solved structures from figures S13-S17 with the C-terminal helix from our peptide agonist system (shown in cyan) after alignment.

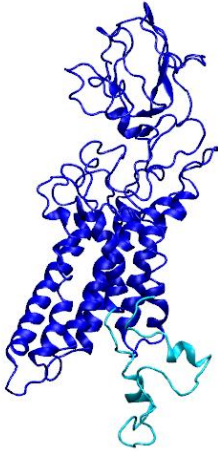
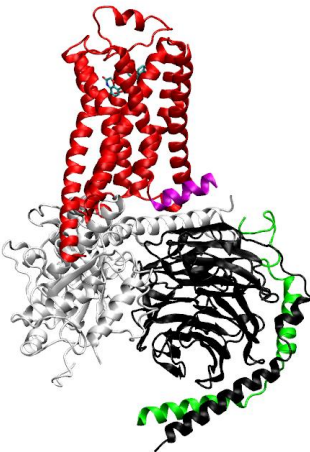
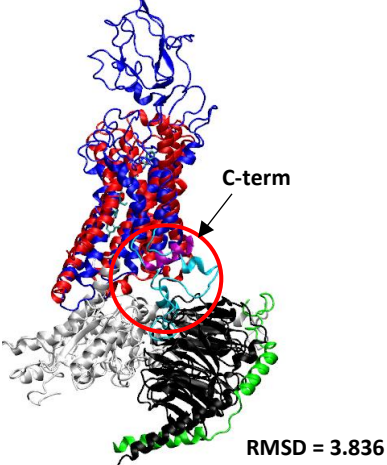



Our Antagonist System	PDB ID: 3SN6	Superimposition
		
		

Figure A19. Comparison between our small molecule antagonist complex and the solved structure of G protein docked Adrenergic receptor obtained from PDB. Top panels show the whole receptor complex, bottom panels show the C-terminal helix. The C-terminal helix is circled highlighting the clash with the G protein.

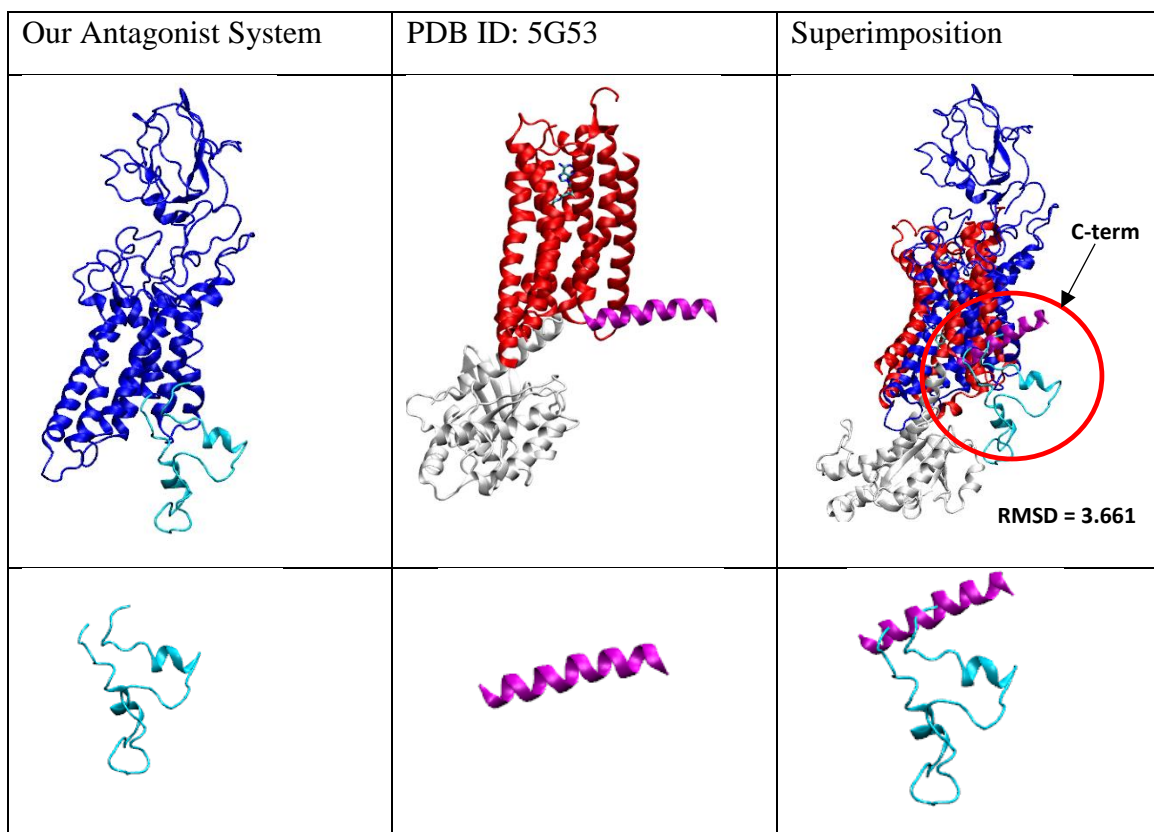


Figure A20. Comparison between our small molecule antagonist complex and the solved G protein docked Adenosine A2A receptor obtained from PDB. Top panels show the whole receptor complex, bottom panels show the C-terminal helix. The C-terminal helix is circled highlighting the clash with the G protein.

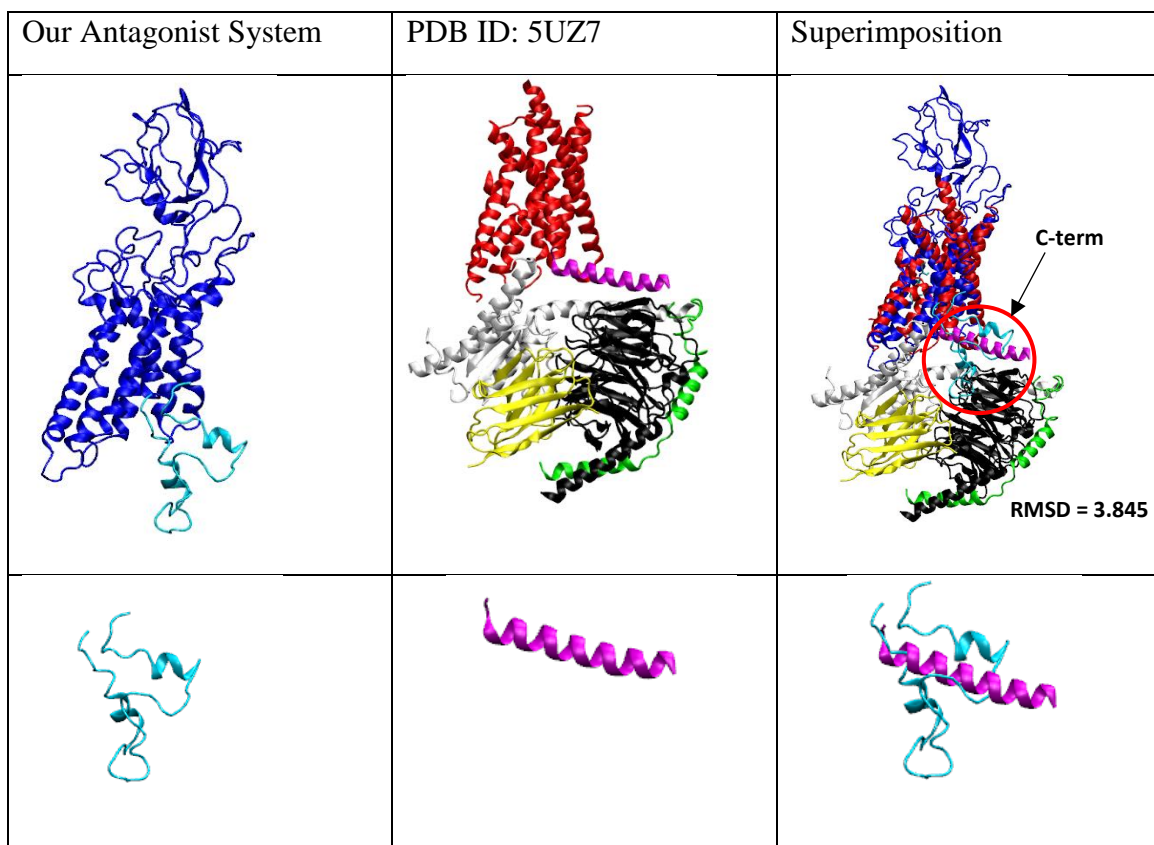


Figure A21. Comparison between our small molecule antagonist complex and the solved G protein docked Calcitonin receptor obtained from PDB. Top panels show the whole receptor complex, bottom panels show the C-terminal helix. The C-terminal helix is circled highlighting the clash with the G protein.

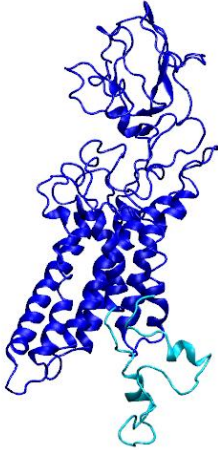
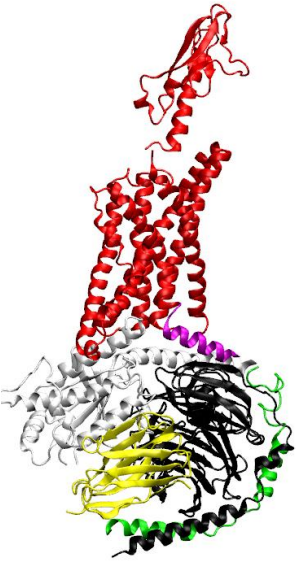
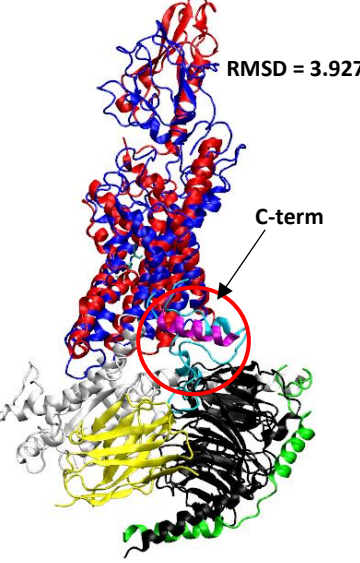


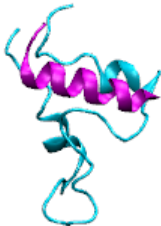
Our Antagonist System	PDB ID: 5VAI	Superimposition
		
		

Figure A22. Comparison between our small molecule antagonist complex and the solved structure of G protein docked Glucagon-like peptide-1 receptor (5vai) obtained from PDB. Top panels show the whole receptor complex, bottom panels show the C-terminal helix. The C-terminal helix is circled highlighting the clash with the G protein.

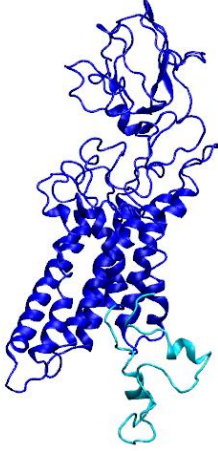
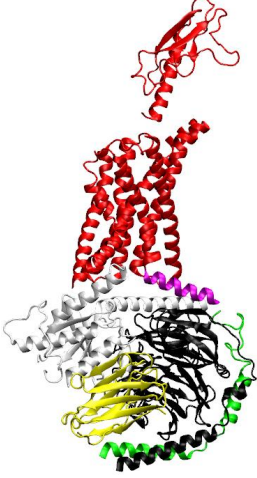
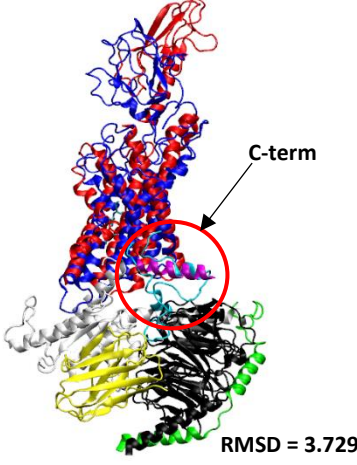



Our Antagonist System	PDB ID: 6B3J	Superimposition
		 <p data-bbox="1312 411 1382 432">C-term</p> <p data-bbox="1284 701 1419 722">RMSD = 3.729</p>
		

Figure A23. Comparison between our small molecule antagonist complex and the solved structure of G protein docked Glucagon-like peptide-1 receptor (6b3j) obtained from PDB. Top panels show the whole receptor complex, bottom panels show the C-terminal helix. The C-terminal helix is circled highlighting the clash with the G protein.

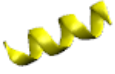
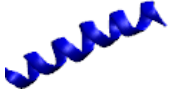
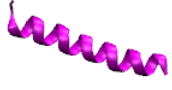
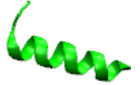

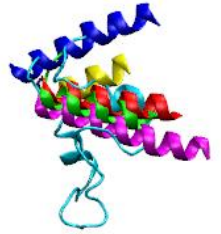
3sn6	5g53	5uz7	5vai	6b3j	All
					

Figure A24. Comparison of C-terminal helices from the solved structures from figures S19-S23 with the C-terminal helix from our small molecule antagonist system (shown in cyan) after alignment.

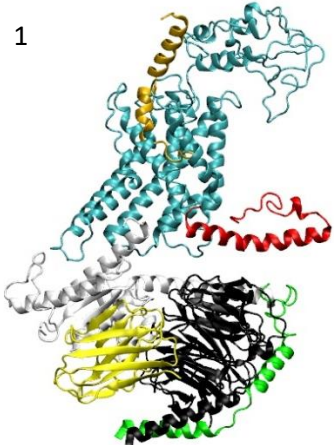
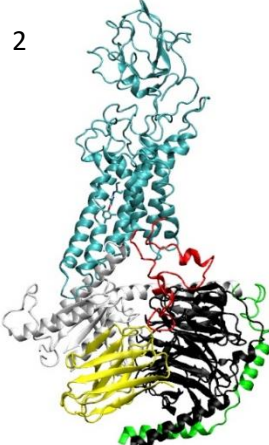
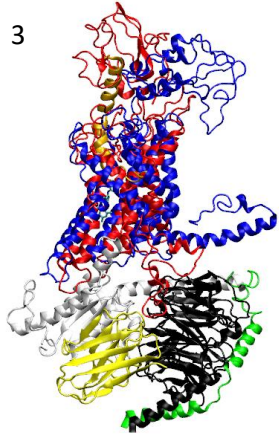


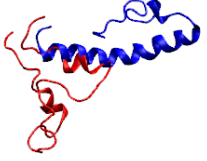
Peptide agonist	Small molecule antagonist	Superimposition
1 	2 	3 
4 	5 	6 

Figure A25. Comparison between our peptide agonist complex and small molecule antagonist complex bound to the G-protein, specifically looking at the C-terminal helix 8 show in red in panels 1 and 2. Panel 3 shows a superimposed comparison of the two complete systems. Panel 4 and 5 shows the agonist and antagonist C-terminal helices individually. Panel 6 compares the C-terminal helices superimposed.

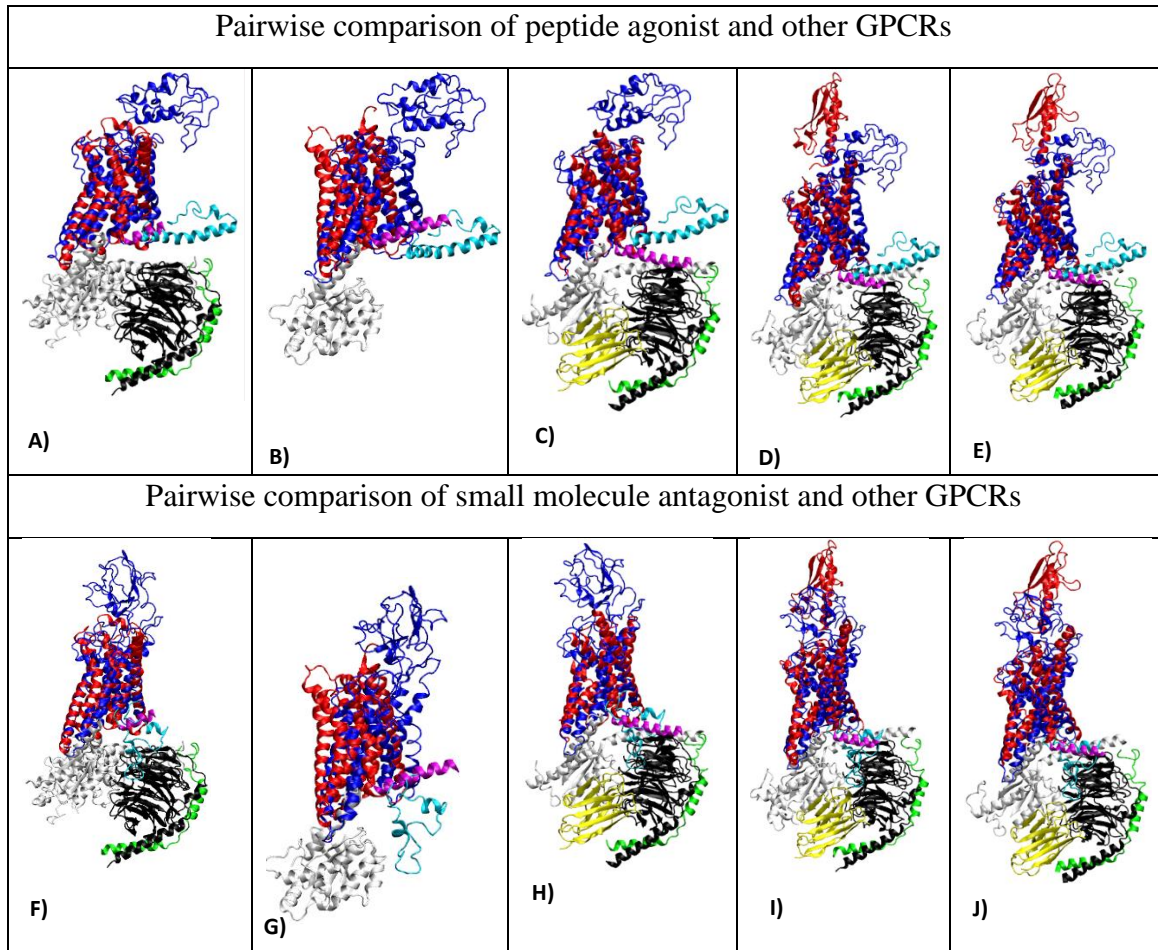


Figure A26. Pairwise comparison between our systems and the five systems from figure 18. Blue = transmembrane domain of our system, cyan = helix 8 from our system, red = other transmembrane domain, magenta = other helix 8, white = alpha component of G protein, black = beta component of G protein, green = gamma component of G protein, yellow = nanobody on G protein. **A)** Our peptide agonist system compared to PDB entry 3sn6. **B)** Peptide agonist compared to PDB entry 5g53. **C)** Peptide agonist compared to PDB entry 5uz7. **D)** Peptide agonist compared to PDB entry 5vai. **E)** Peptide agonist compared to PDB entry 6b3j. **F)** Our small molecule antagonist system compared to PDB entry 3sn6. **G)** Small molecule system compared to PDB entry 5g53. **H)** Small molecule compared to PDB entry 5uz7. **I)** Small molecule compared to PDB entry 5vai. **J)** Small molecule compared to PDB entry 6b3j.

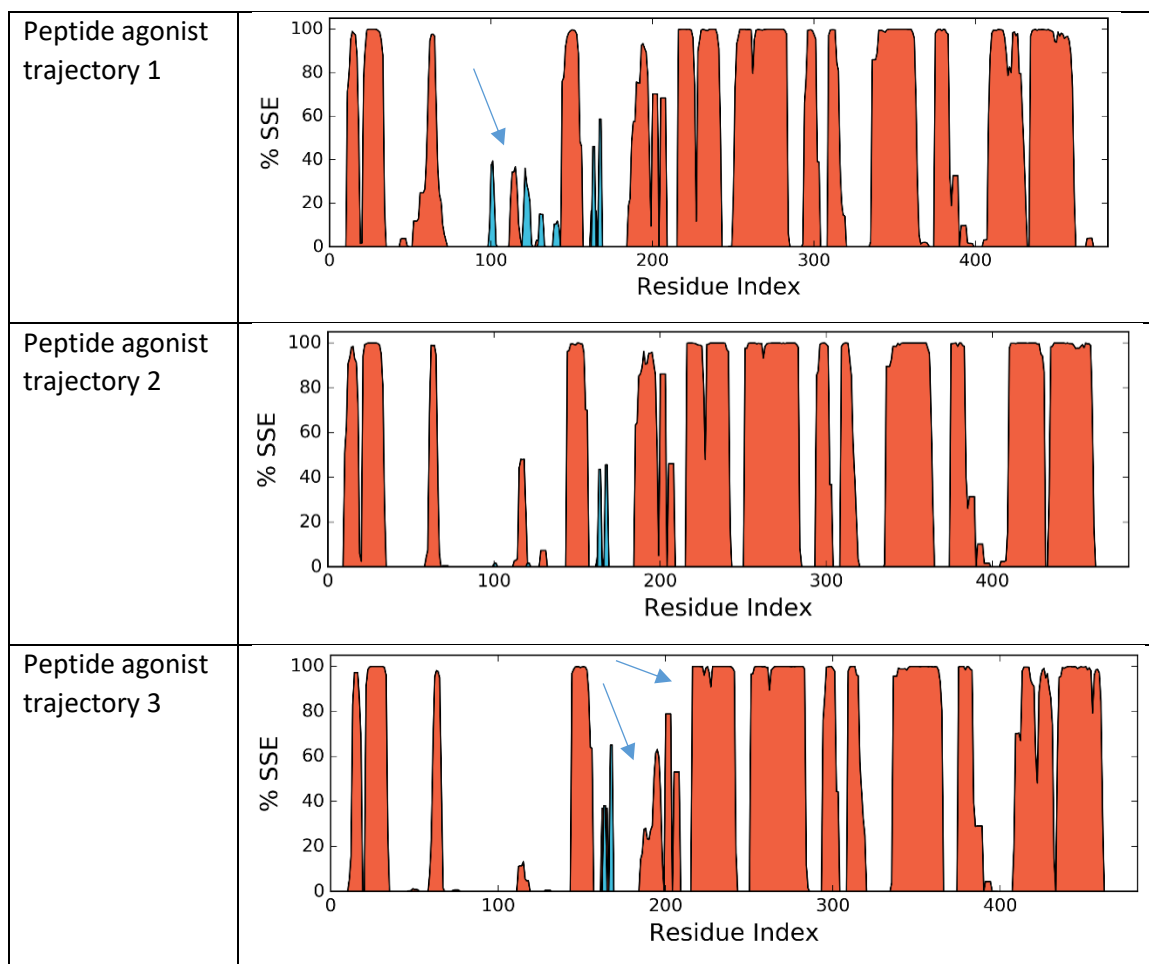


Figure A27. Comparison of secondary structural elements between the three simulation trajectories of the peptide agonist system. Orange = alpha helical structure, blue = beta sheets. N-terminal region: 0-145, TM1: 146-170, TM2: 178-205, TM3: 215-248, TM4: 255-281, TM5: 298-330, TM6: 339-362, TM7: 370-395, C-terminal region: 396-444. Arrows indicate regions of variance.

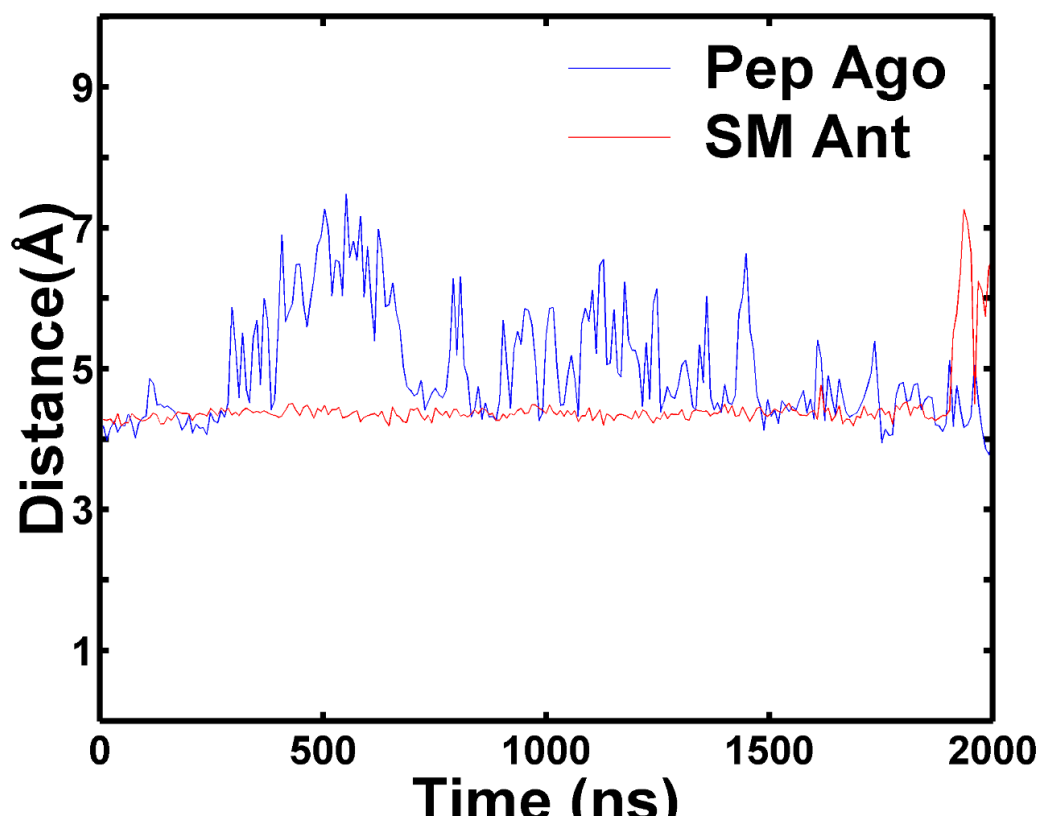


Figure A28. Comparison of the polar lock distances of the first trajectory for agonist and antagonist systems for a preliminary analysis.

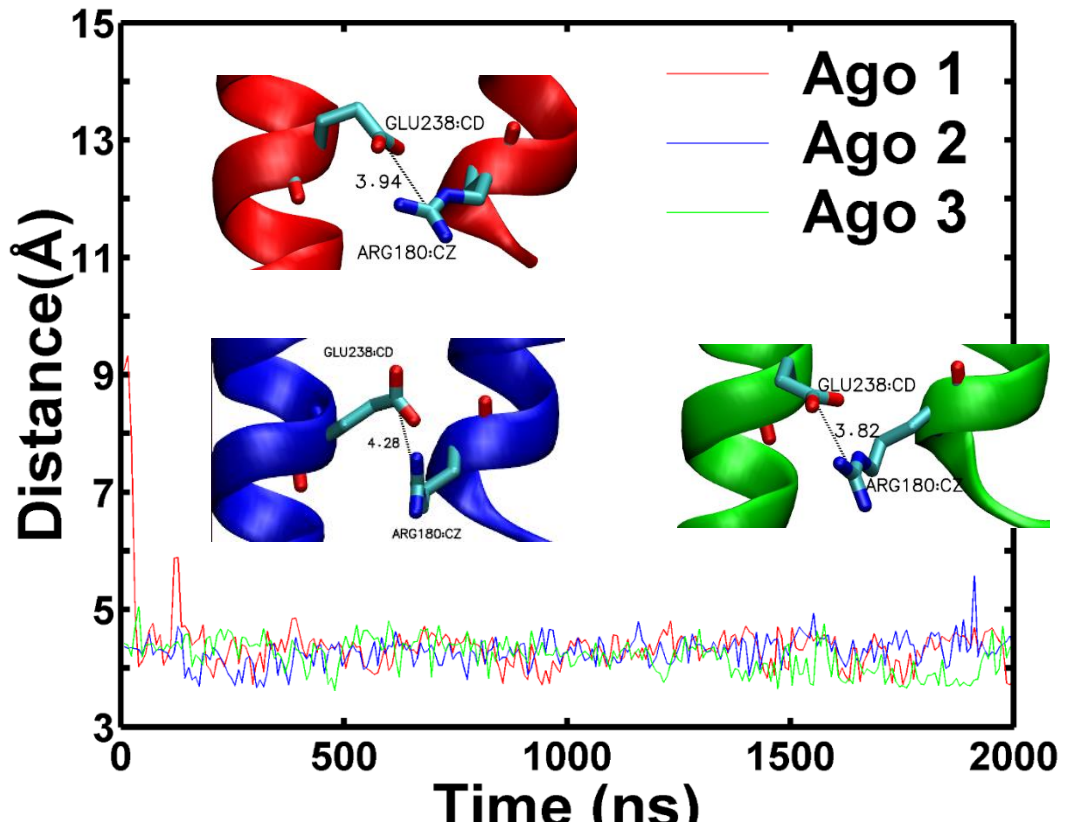


Figure A29. Comparison of the ionic lock distances in angstroms across all three simulation trajectories for 2000ns. Average distance for the 1st simulation = 3.94Å, 2nd simulation = 4.28Å, 3rd simulation = 3.84 Å.

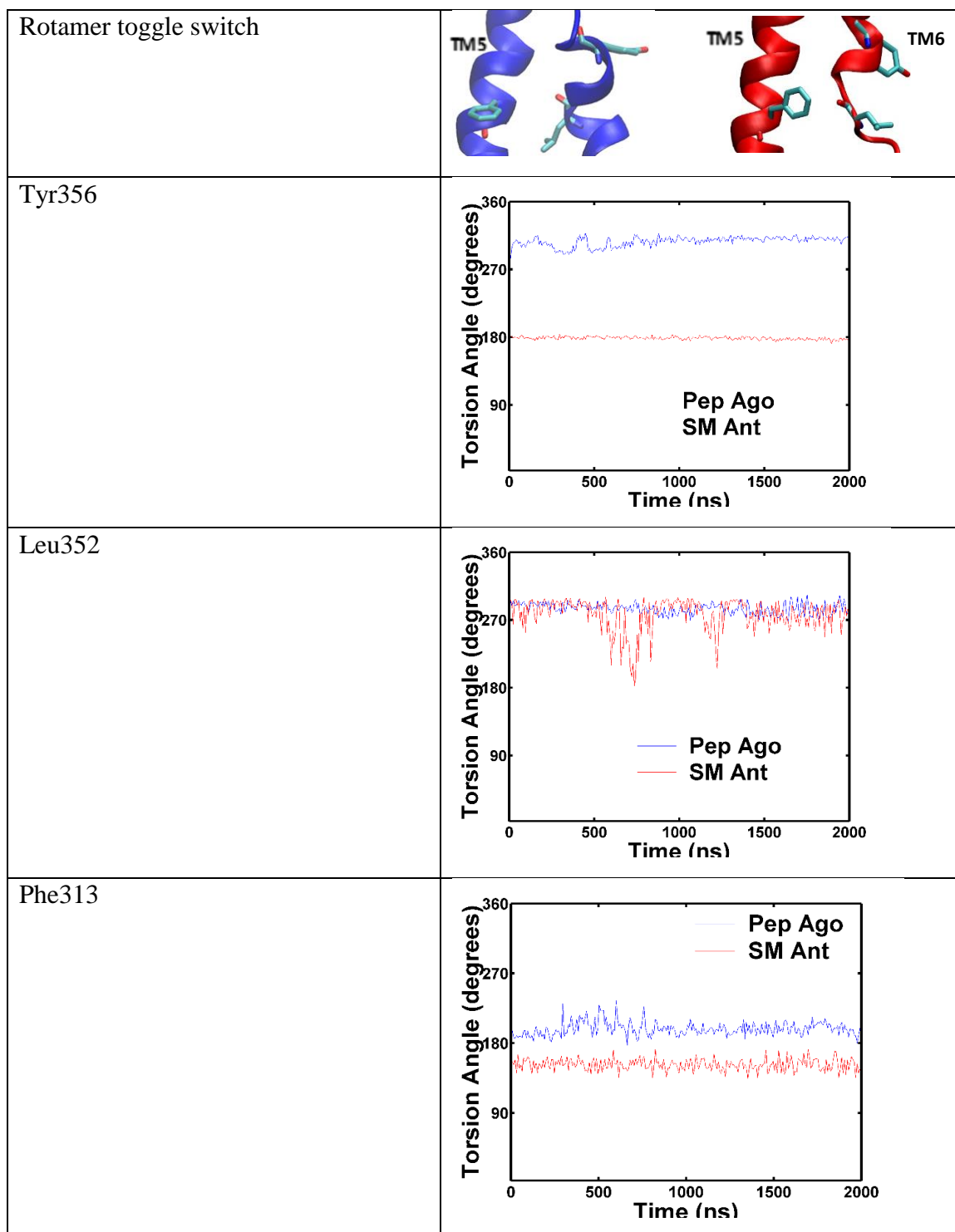


Figure A30. Comparison of the rotamer toggle switch rotational differences of each residue for the first trajectory for agonist and antagonist systems for a preliminary analysis.

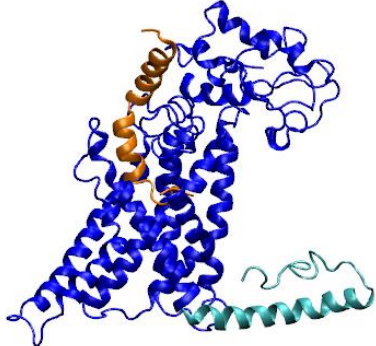
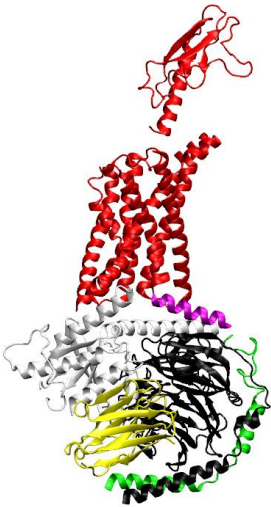
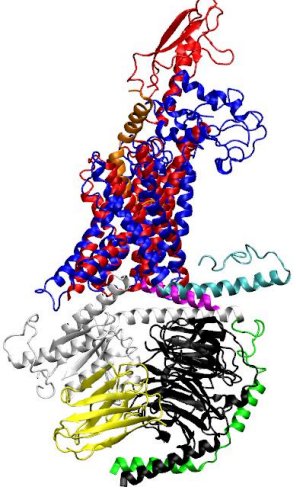
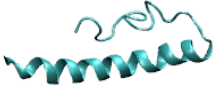

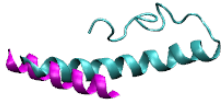
Our Peptide Agonist System	PDB ID: 6B3J	Superimposition
		
		

Figure A31. Comparison between our second trajectory peptide agonist complex and the solved structure of G protein docked Glucagon-like peptide-1 receptor (6b3j) obtained from PDB. Top panels show the full complex, bottom panels show the C-terminal helix.

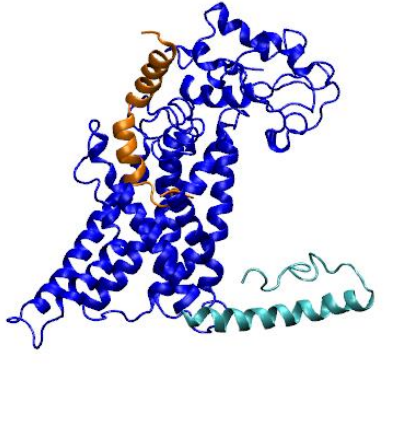
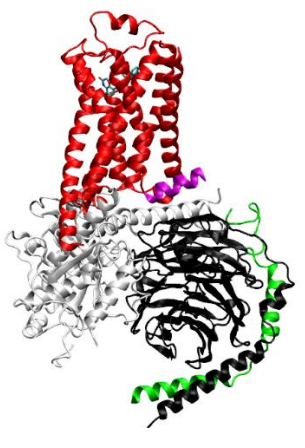
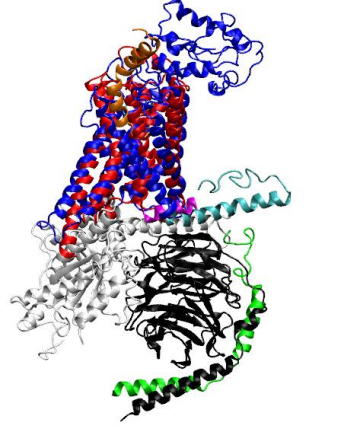
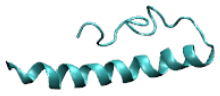

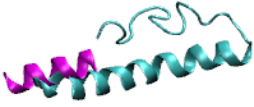
Our Peptide Agonist System	PDB ID: 3SN6	Superimposition
		
		

Figure A32. Comparison between our second trajectory peptide agonist complex and the solved structure for G protein docked Adrenergic receptor obtained from PDB. Top panels show the full complex, bottom panels show the C-terminal helix.

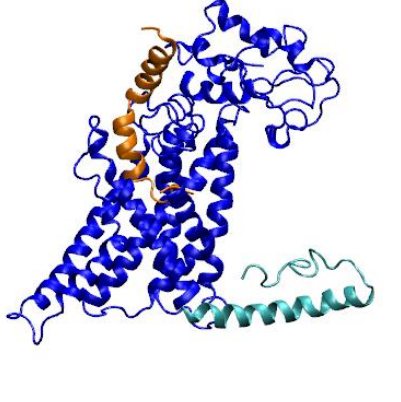
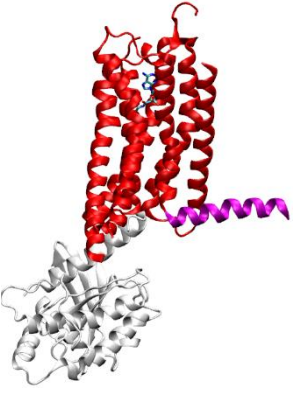
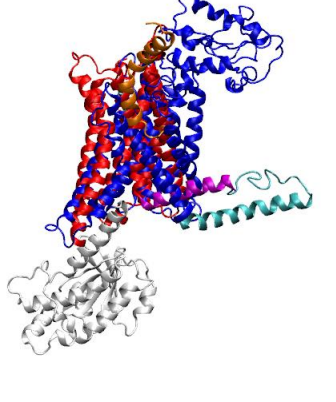
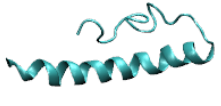

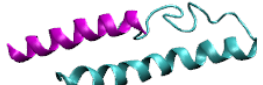
Our Peptide Agonist System	PDB ID: 5G53	Superimposition
		
		

Figure A33. Comparison between our second trajectory peptide agonist complex and the solved structure for G protein docked Adenosine A2A receptor obtained from PDB. Top panels show the full complex, bottom panels show the C-terminal helix.

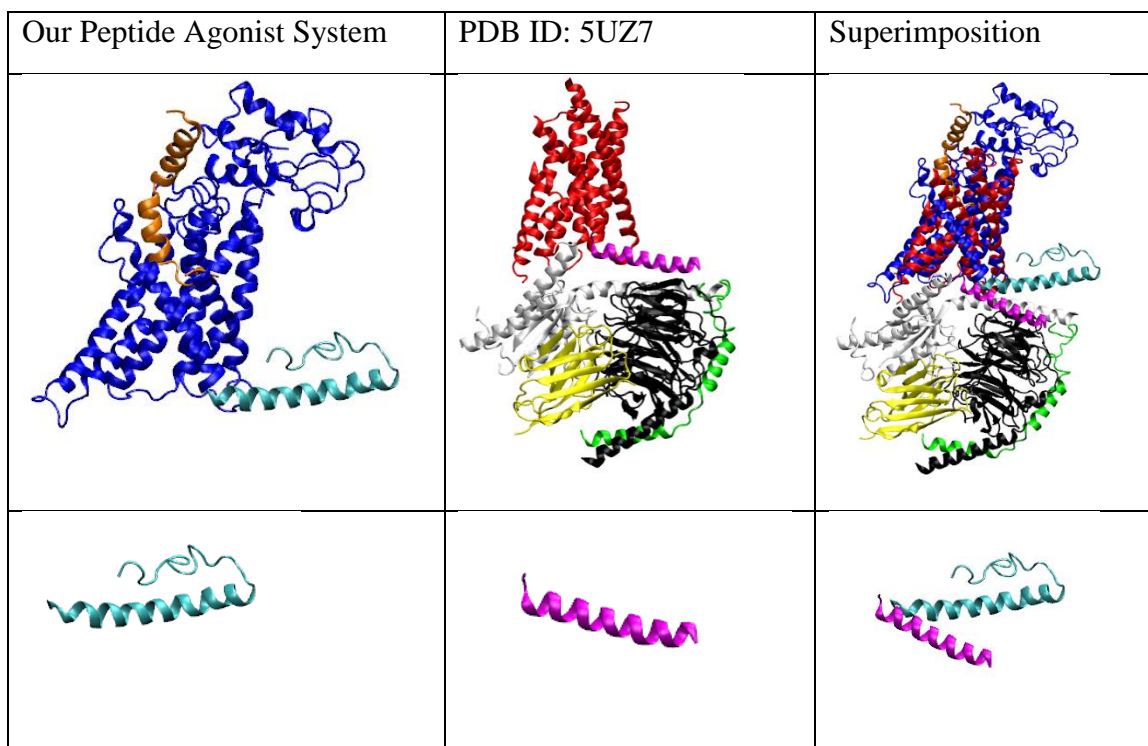


Figure A34. Comparison between our second trajectory peptide agonist complex and the solved structure for G protein docked Calcitonin receptor obtained from PDB. Top panels show the full complex, bottom panels show the C-terminal helix.

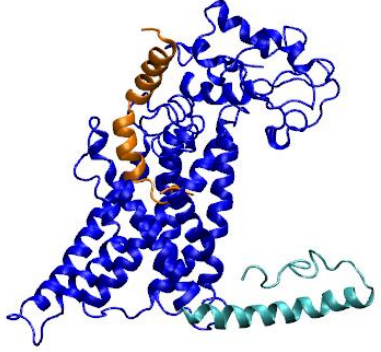
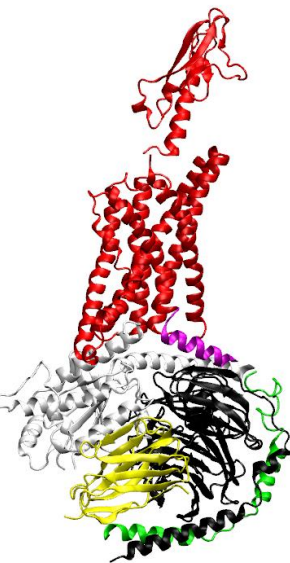
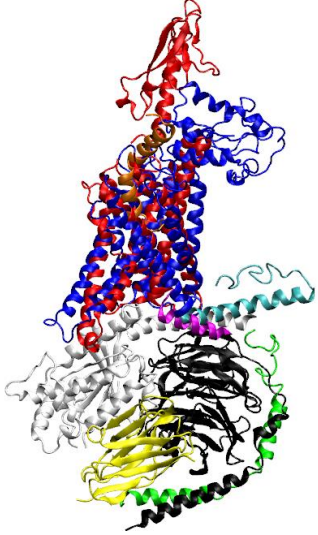
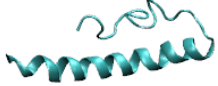

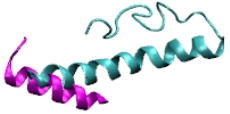
Our Peptide Agonist System	PDB ID: 5VAI	Superimposition
		
		

Figure A35. Comparison between our second trajectory peptide agonist complex and the solved structure for G protein docked Glucagon-like peptide-1 receptor (5vai) obtained from PDB. Top panels show the full complex, bottom panels show the C-terminal helix.

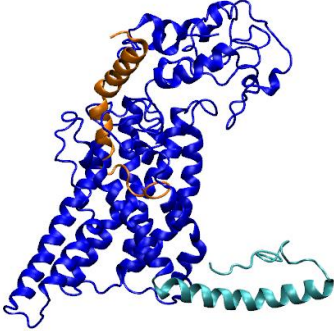
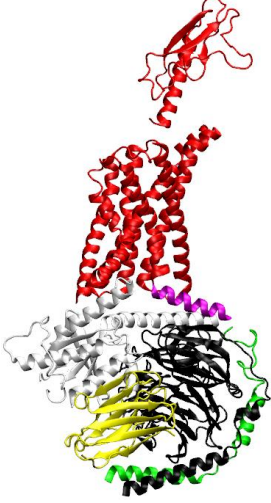
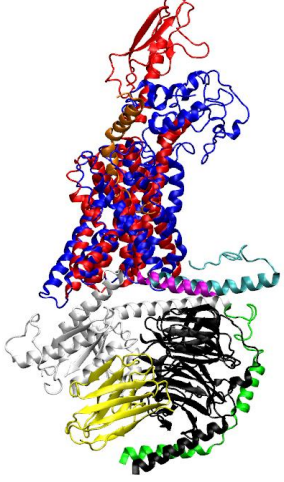


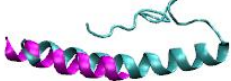
Our Peptide Agonist System	PDB ID: 6B3J	Superimposition
		
		

Figure A36. Comparison between our third trajectory peptide agonist complex and the solved structure for G protein docked Glucagon-like peptide-1 receptor (6b3j) obtained from PDB. Top panels show the full complex, bottom panels show the C-terminal helix.

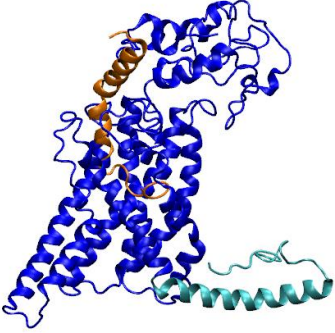
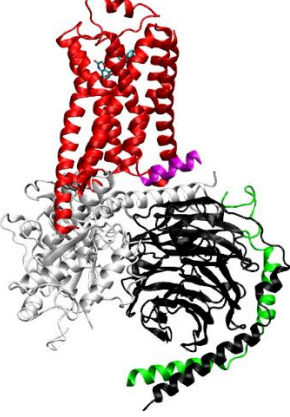
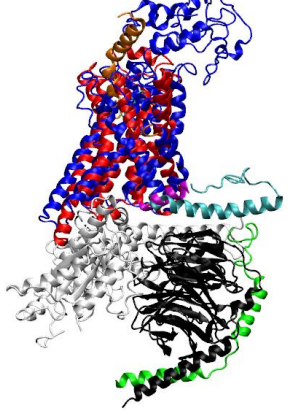



Our Peptide Agonist System	PDB ID: 3SN6	Superimposition
		
		

Figure A37. Comparison between our third trajectory peptide agonist complex and the solved structure for G protein docked Adrenergic receptor obtained from PDB. Top panels show the full complex, bottom panels show the C-terminal helix.

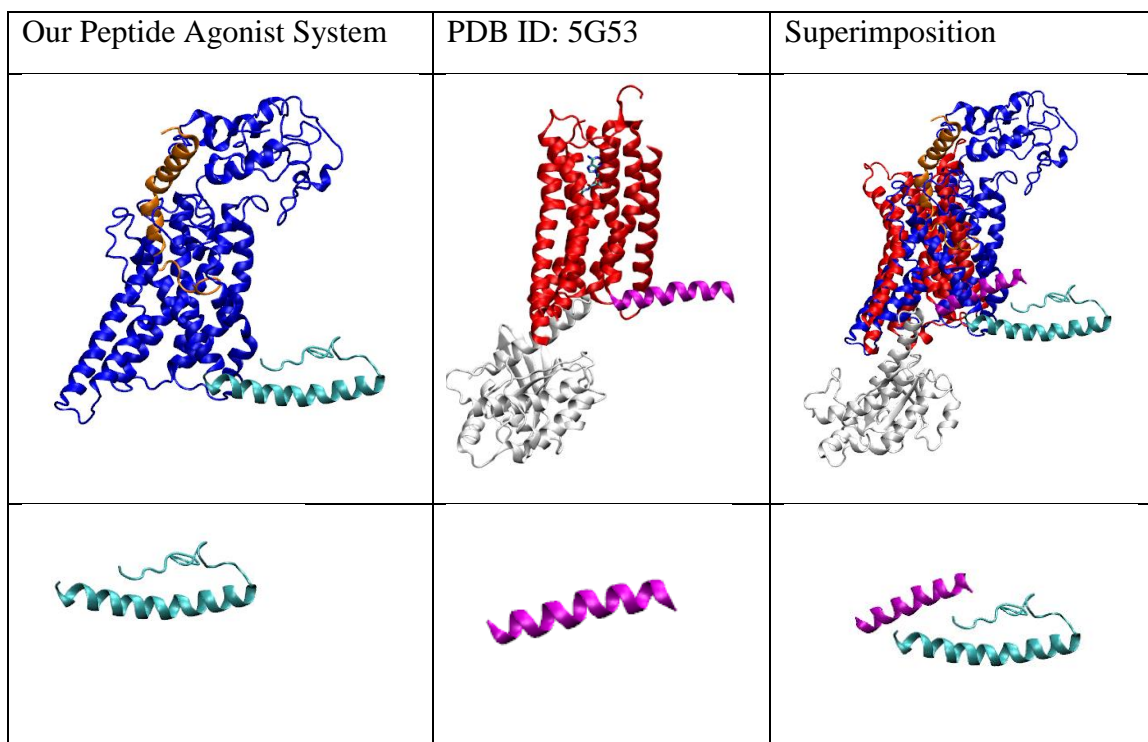


Figure A38. Comparison between our third trajectory peptide agonist complex and the solved structure for G protein docked Adenosine A2A receptor obtained from PDB. Top panels show the full complex, bottom panels show the C-terminal helix.

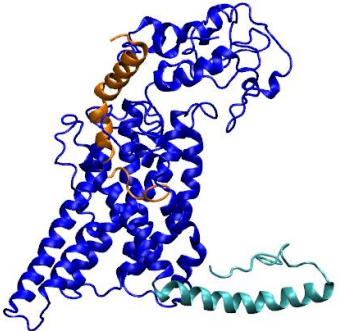
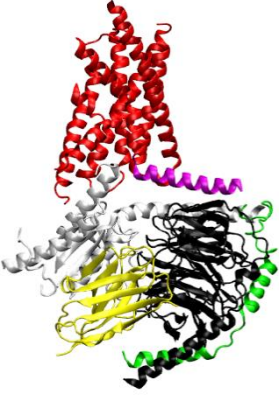
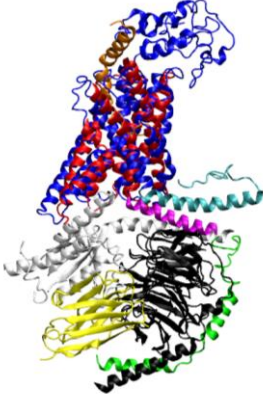

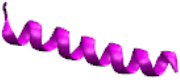

Our Peptide Agonist System	PDB ID: 5UZ7	Superimposition
		
		

Figure A39. Comparison between our third trajectory peptide agonist complex and the solved structure for G protein docked Calcitonin receptor obtained from PDB. Top panels show the full complex, bottom panels show the C-terminal helix.

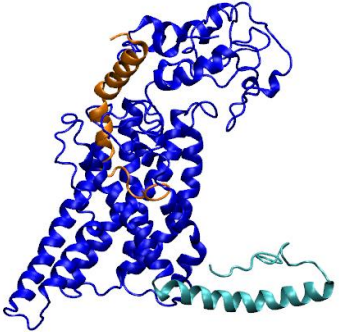
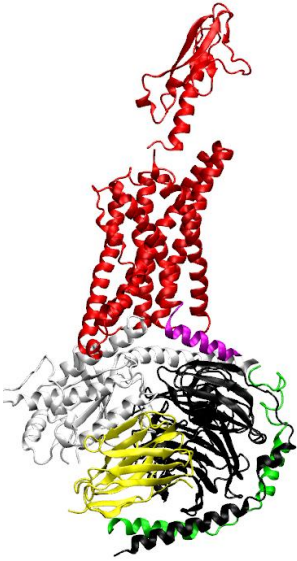
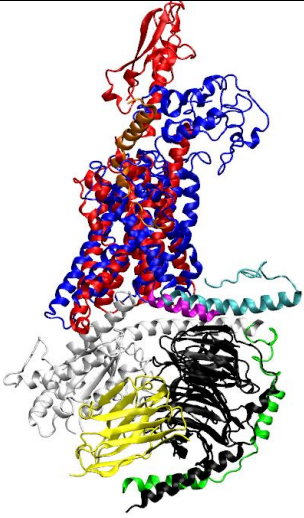
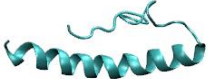

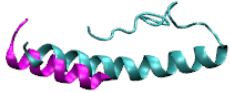
Our Peptide Agonist System	PDB ID: 5VAI	Superimposition
		
		

Figure A40. Comparison between our third trajectory peptide agonist complex and the solved structure for G protein docked Glucagon-like peptide-1 receptor (5vai) obtained from PDB. Top panels show the full complex, bottom panels show the C-terminal helix.


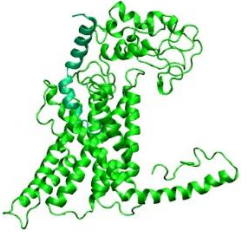
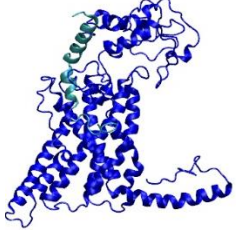
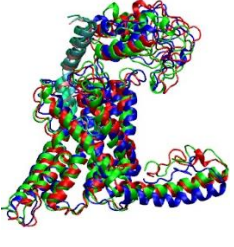
Agonist Trajectory 1	Agonist Trajectory 2	Agonist Trajectory 3	Superimposition
			

Figure A41. Comparison between all three trajectories of our peptide agonist systems after 2000ns of simulation. Red = trajectory 1, Green = Trajectory 2, Blue = Trajectory 3.

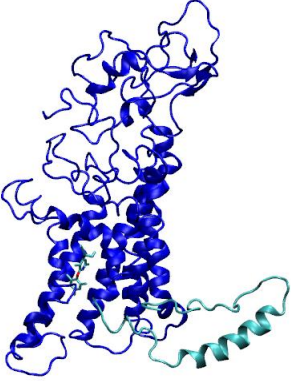
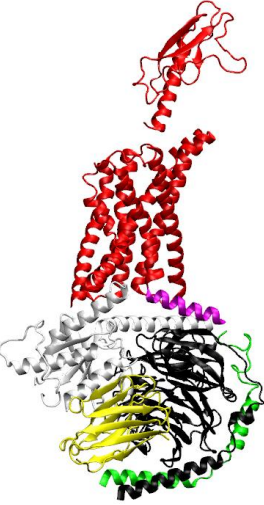
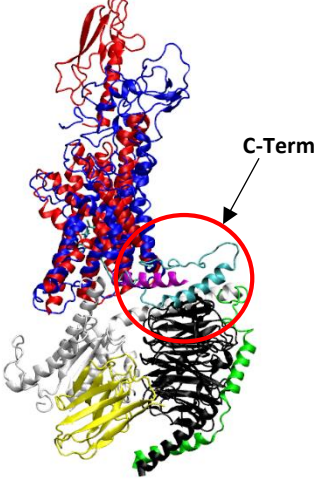
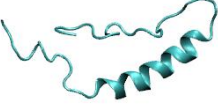

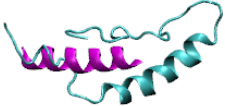
Our Antagonist System	PDB ID: 6B3J	Superimposition
		
		

Figure A42. Comparison between our second trajectory small molecule antagonist complex and the solved structure of G protein docked Glucagon-like peptide-1 receptor (6b3j) obtained from PDB. Top panels show the full complex, bottom panels show the C-terminal helix. The C-terminal helix is circled clashing with the G protein.

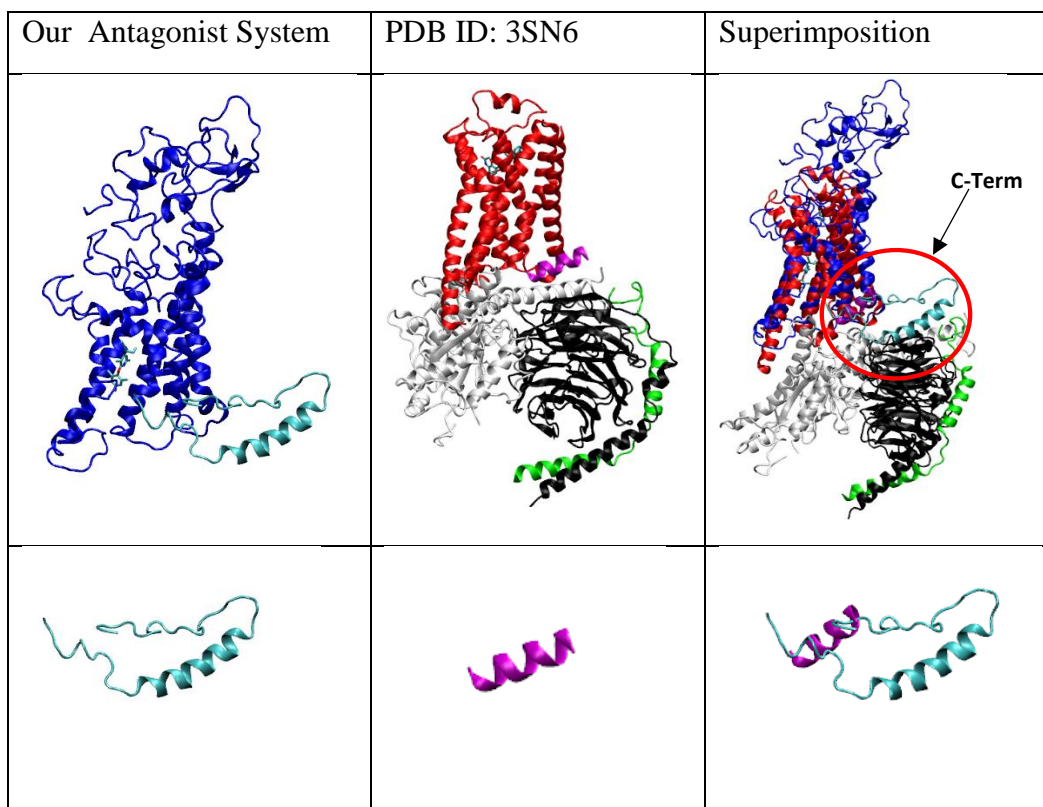


Figure A43. Comparison between our second trajectory small molecule antagonist complex and the solved structure of G protein docked Adrenergic receptor obtained from PDB. Top panels show the full complex, bottom panels show the C-terminal helix. The C-terminal helix is circled clashing with the G protein.

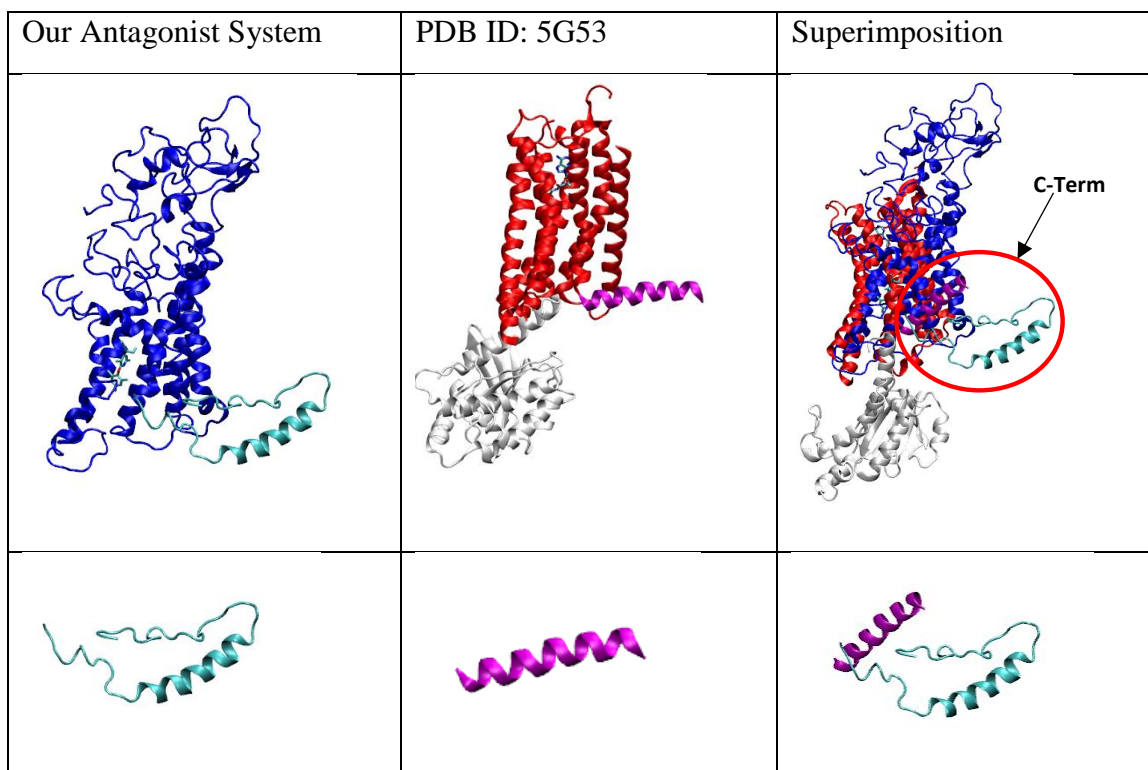


Figure A44. Comparison between our second trajectory small molecule antagonist complex and the solved structure for G protein docked Adenosine A2A receptor obtained from PDB. Top panels show the full complex, bottom panels show the C-terminal helix. The C-terminal helix is circled clashing with the G protein.

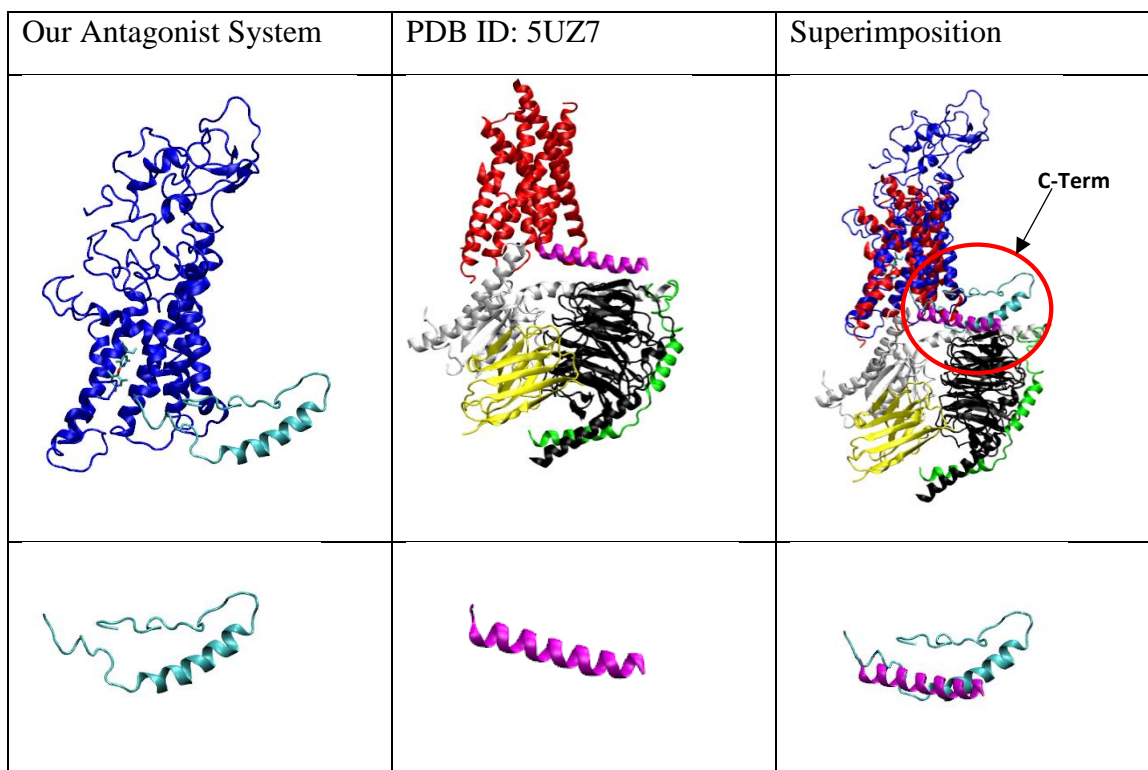


Figure A45. Comparison between our second trajectory small molecule antagonist complex and the solved structure for G protein docked Calcitonin receptor obtained from PDB. Top panels show the full complex, bottom panels show the C-terminal helix. The C-terminal helix is circled clashing with the G protein.

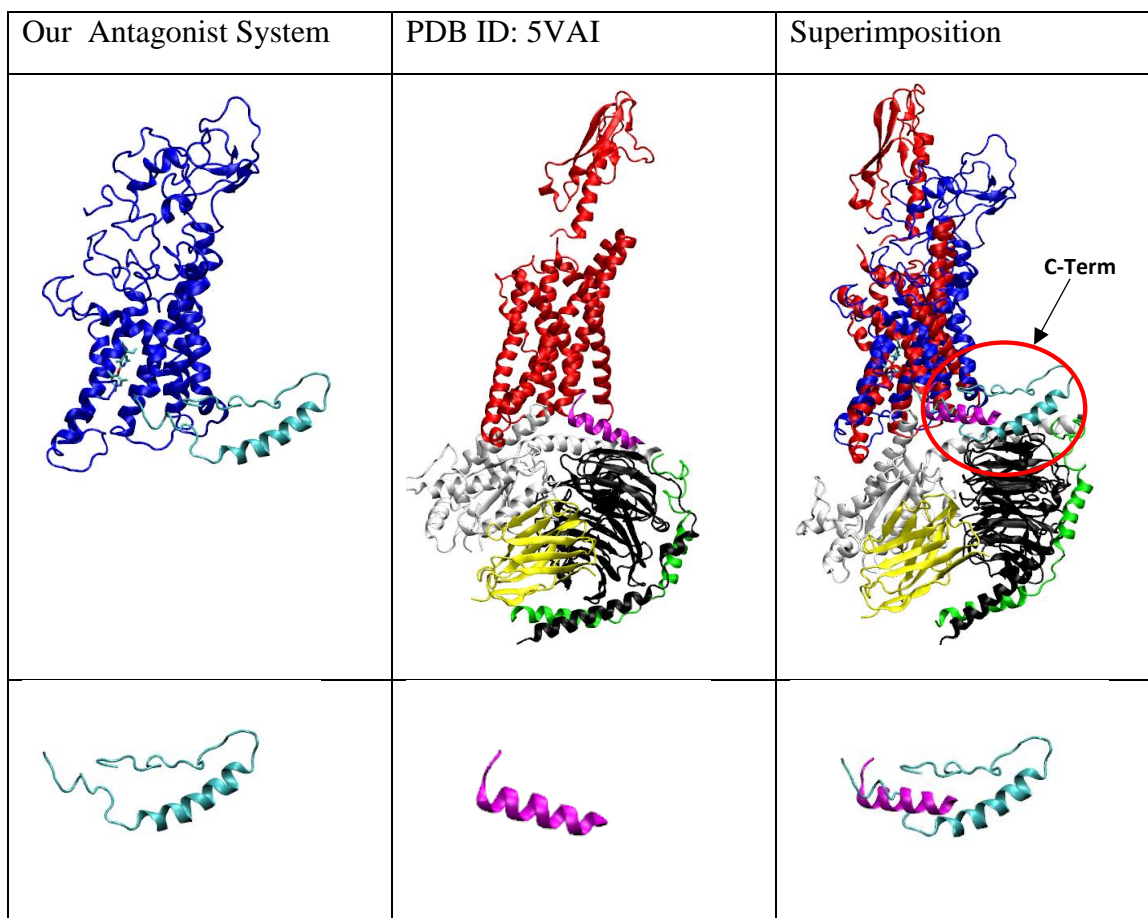


Figure A46. Comparison between our second trajectory small molecule antagonist complex and the solved structure for G protein docked Glucagon-like peptide-1 receptor (5vai) obtained from PDB. Top panels show the full complex, bottom panels show the C-terminal helix. The C-terminal helix is circled clashing with the G protein.

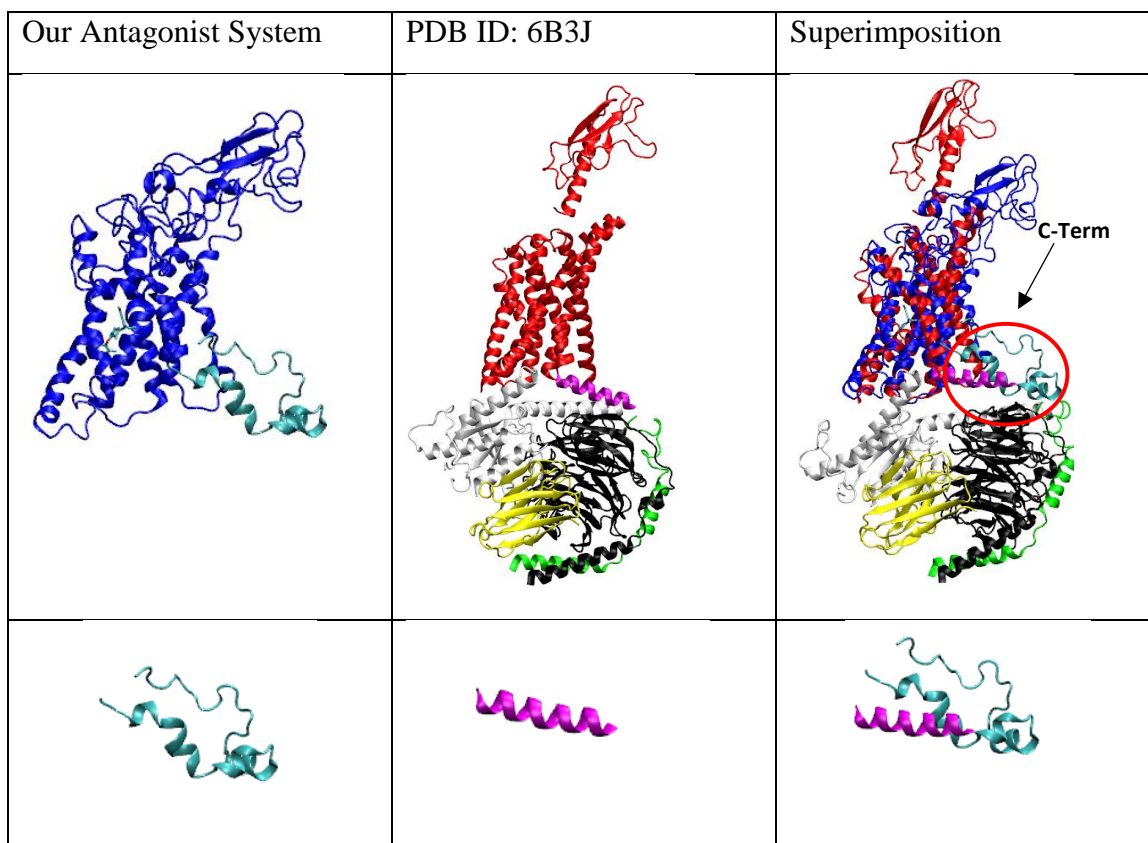


Figure A47. Comparison between our third trajectory small molecule antagonist complex and the solved structure for G protein docked Glucagon-like peptide-1 receptor (6b3j) obtained from PDB. Top panels show the full complex, bottom panels show the C-terminal helix. The C-terminal helix is circled clashing with the G protein.

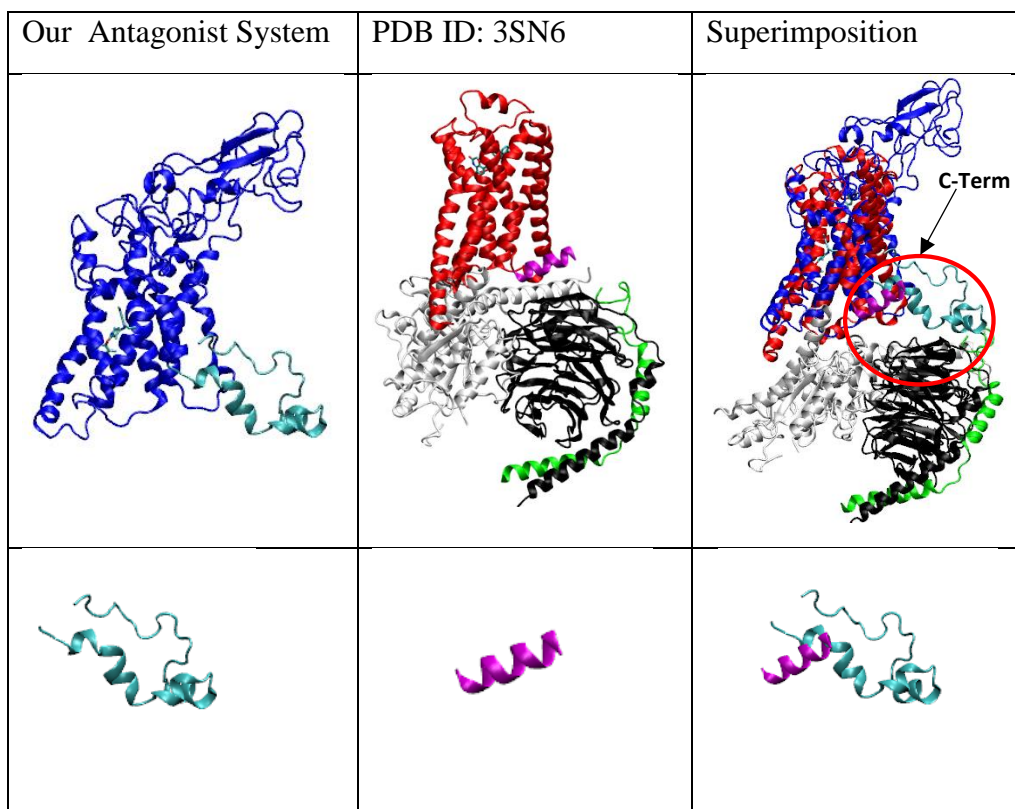


Figure A48. Comparison between our third trajectory small molecule antagonist complex and the solved structure for G protein docked Adrenergic receptor obtained from PDB. Top panels show the full complex, bottom panels show the C-terminal helix. The C-terminal helix is circled clashing with the G protein.

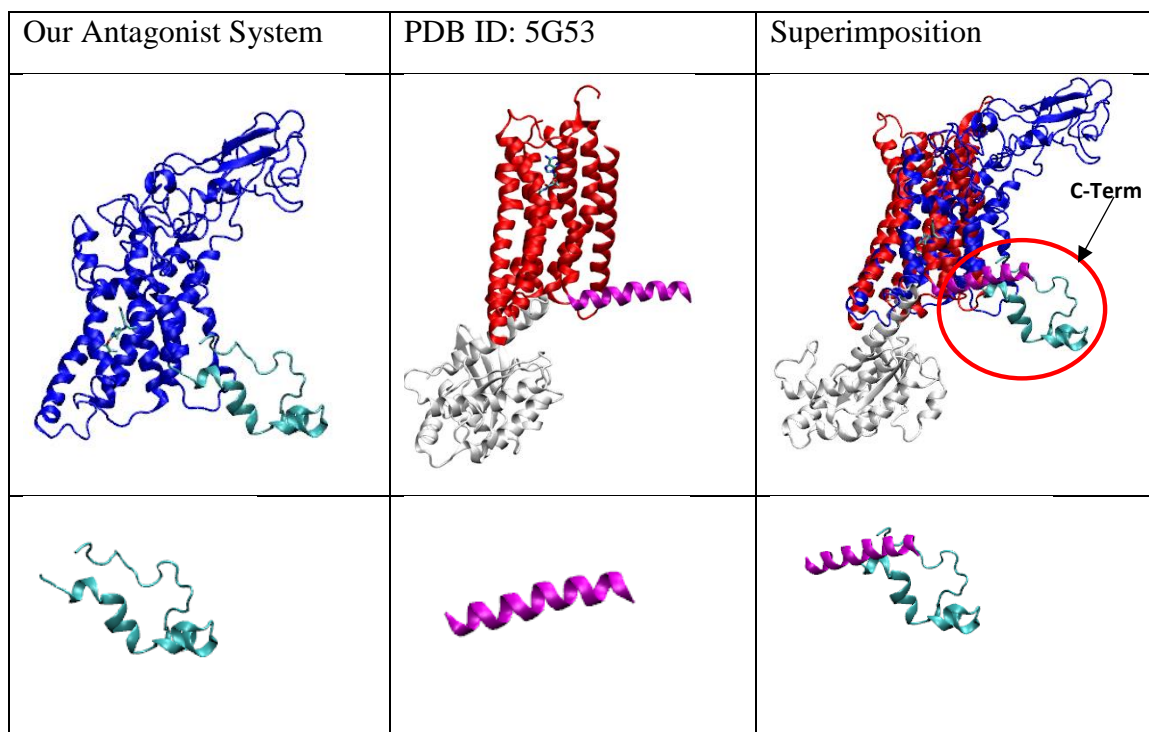


Figure A49. Comparison between our third trajectory small molecule antagonist complex and the solved structure for G protein docked Adenosine A2A receptor obtained from PDB. Top panels show the full complex, bottom panels show the C-terminal helix. The C-terminal helix is circled in red.

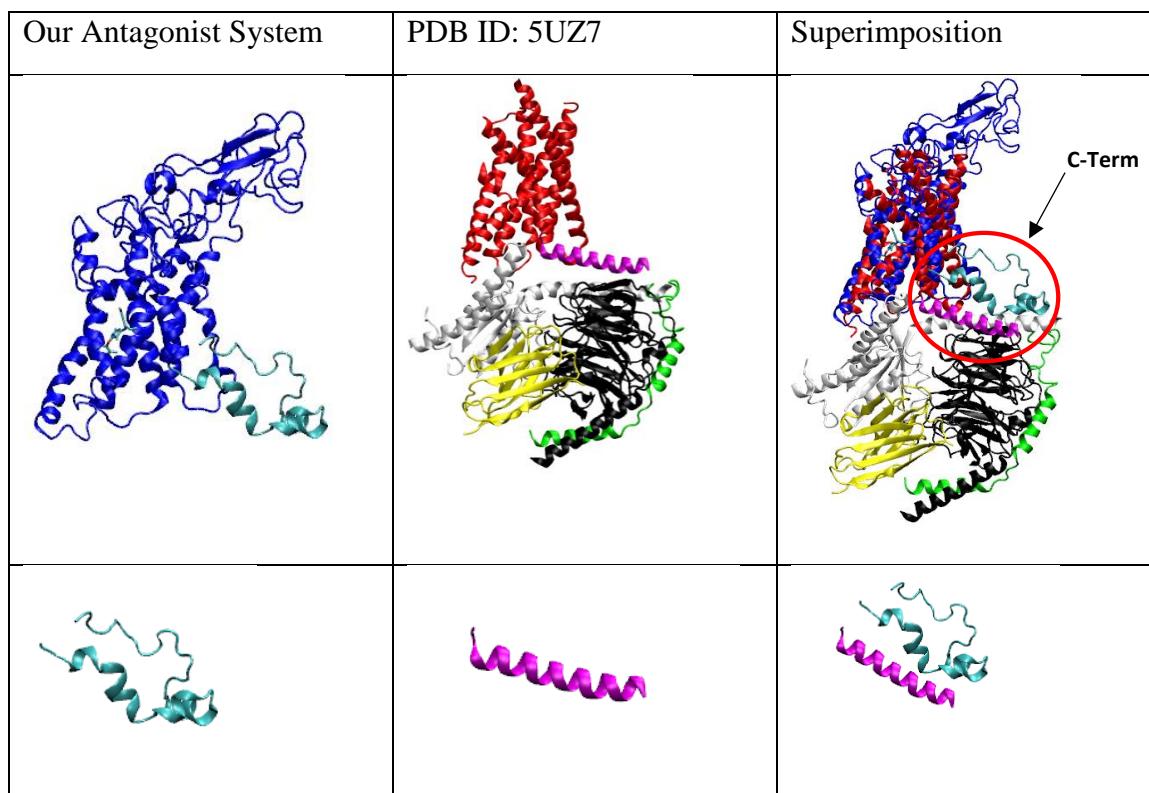


Figure A50. Comparison between our third trajectory small molecule antagonist complex and the solved structure for G protein docked Calcitonin receptor obtained from PDB. Top panels show the full complex, bottom panels show the C-terminal helix. The C-terminal helix is circled in red.

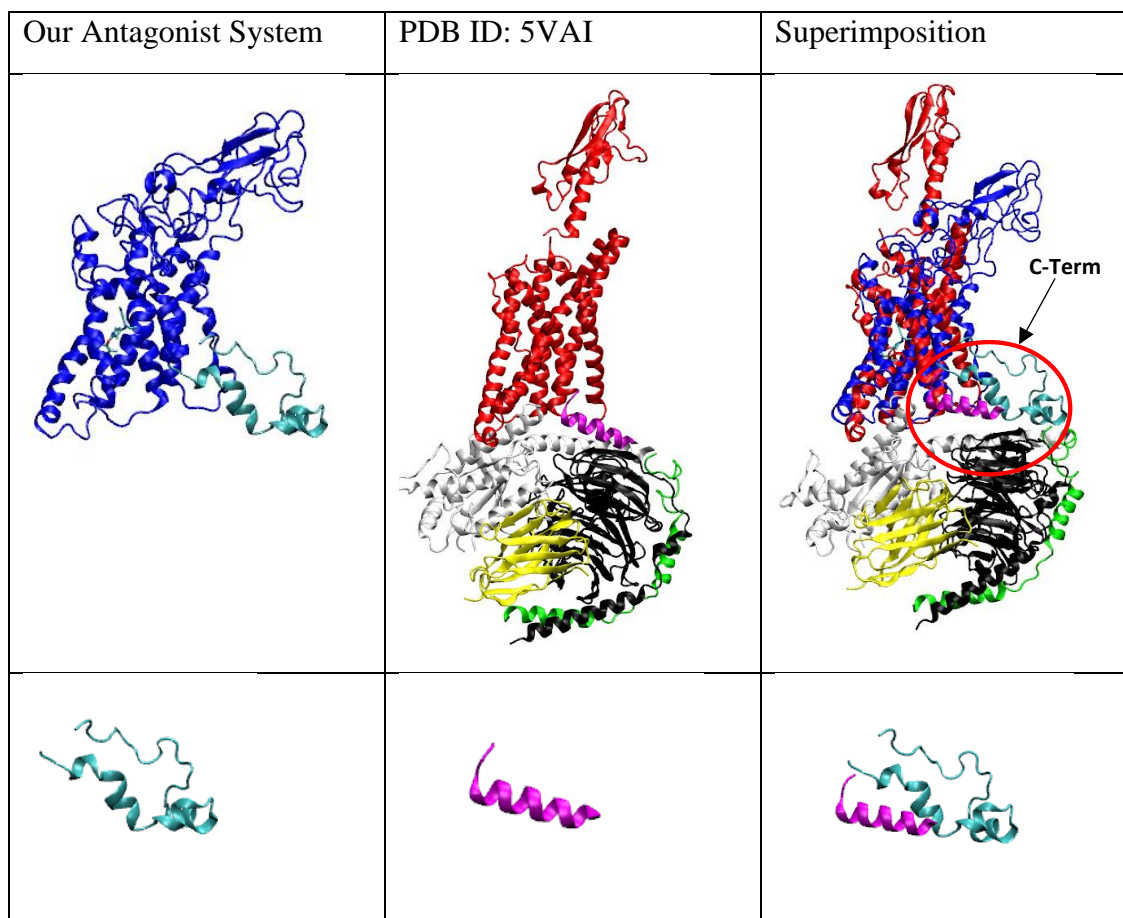


Figure A51. Comparison between our third trajectory small molecule antagonist complex and the solved structure for G protein docked Glucagon-like peptide-1 receptor (5vai) obtained from PDB. Top panels show the full complex, bottom panels show the C-terminal helix. The C-terminal helix is circled in red.

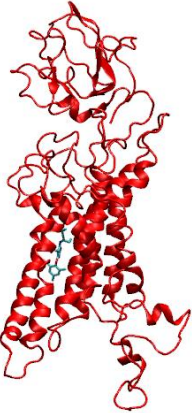
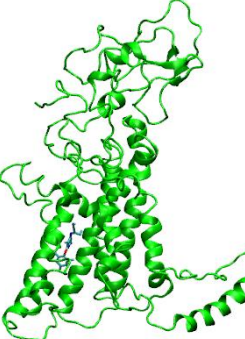
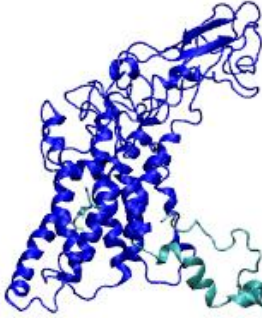
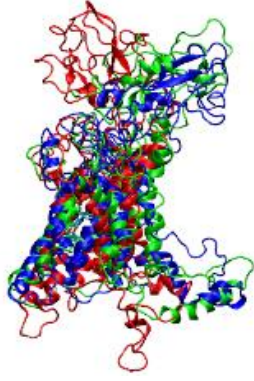
Antagonist Trajectory 1	Antagonist Trajectory 2	Antagonist Trajectory 3	Superimposition
			

Figure A52. Comparison between all three trajectories of our small molecule antagonist systems after 2000ns of simulation. Red = trajectory 1, Green = Trajectory 2, Blue = Trajectory 3.

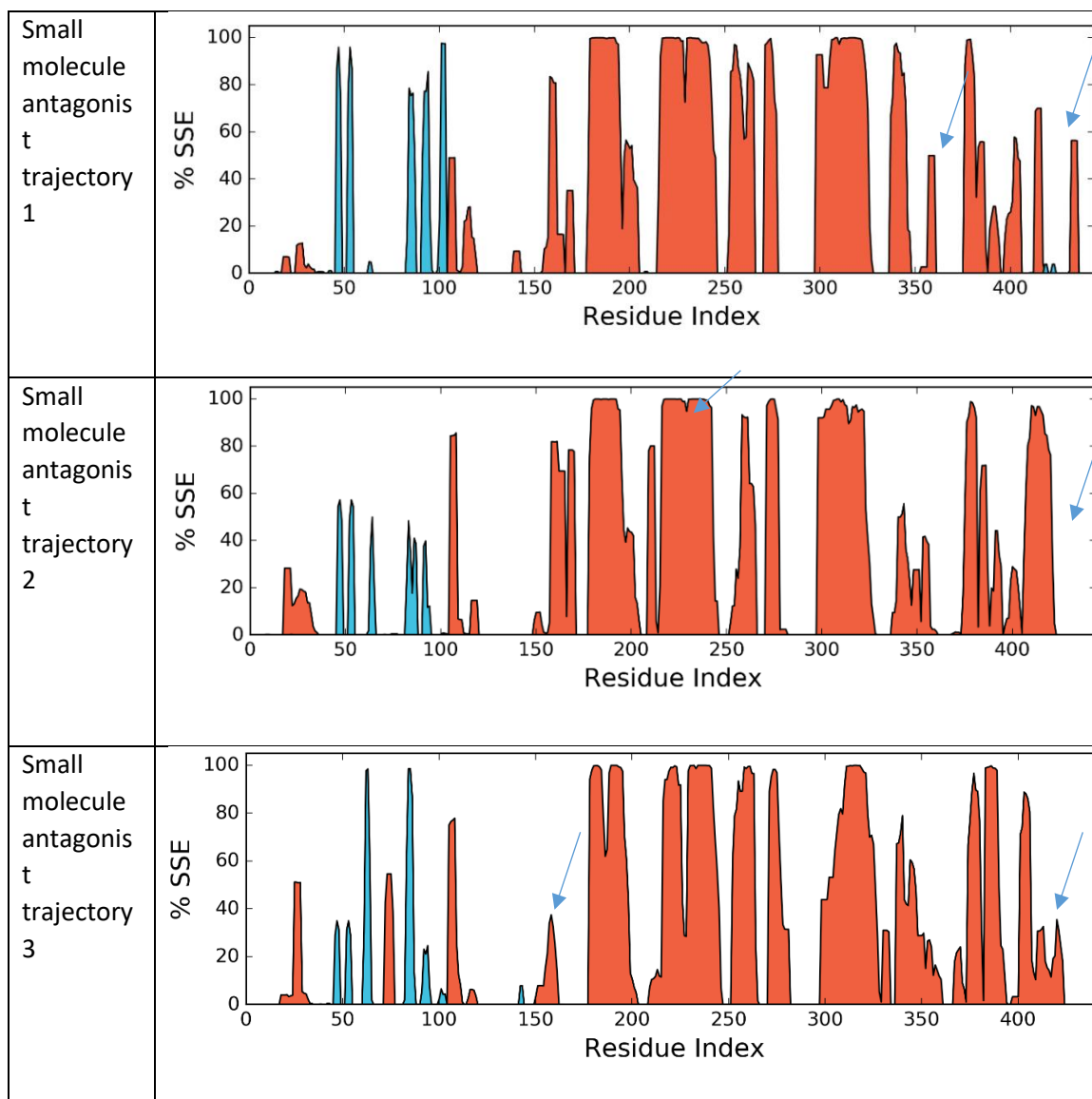


Figure A53. Comparison of the secondary structural element between the three trajectories of the small molecule antagonists. Orange = alpha helical structures, blue = beta sheet formation. N-terminal region: 0-145, TM1: 146-170, TM2: 178-205, TM3: 215-248, TM4: 255-281, TM5: 298-330, TM6: 339-362, TM7: 370-395, C-terminal region: 396-444. Arrows indicate regions of variance.

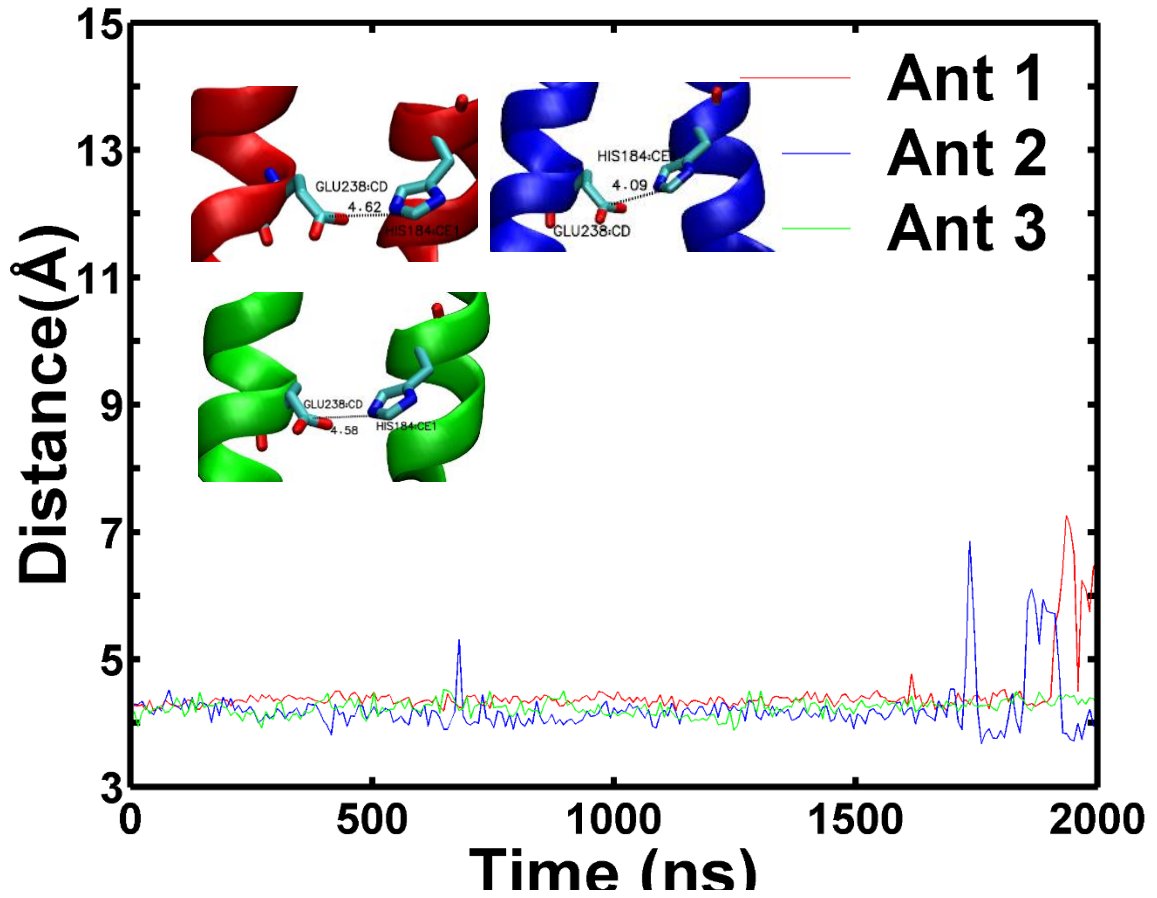


Figure A54. Comparison of the polar lock distances in angstroms across all three antagonist trajectory simulations. Average distance of trajectory 1 = 4.62 Å, trajectory 2 = 4.09 Å, trajectory 3 = 4.58 Å.

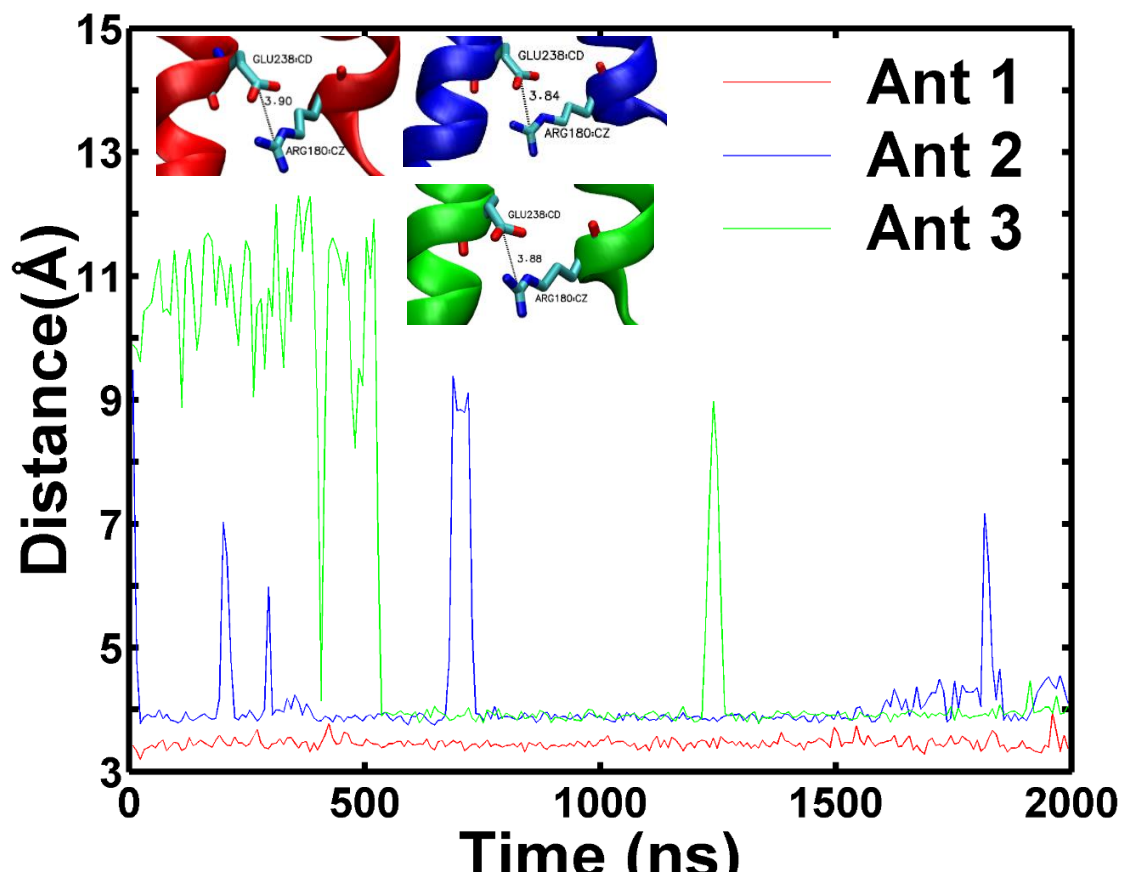


Figure A55. Comparison of the ionic lock distances across all three antagonist trajectories. Average distance for trajectory 1 = 3.90 Å, trajectory 2 = 3.84 Å, trajectory 3 = 3.88 Å.

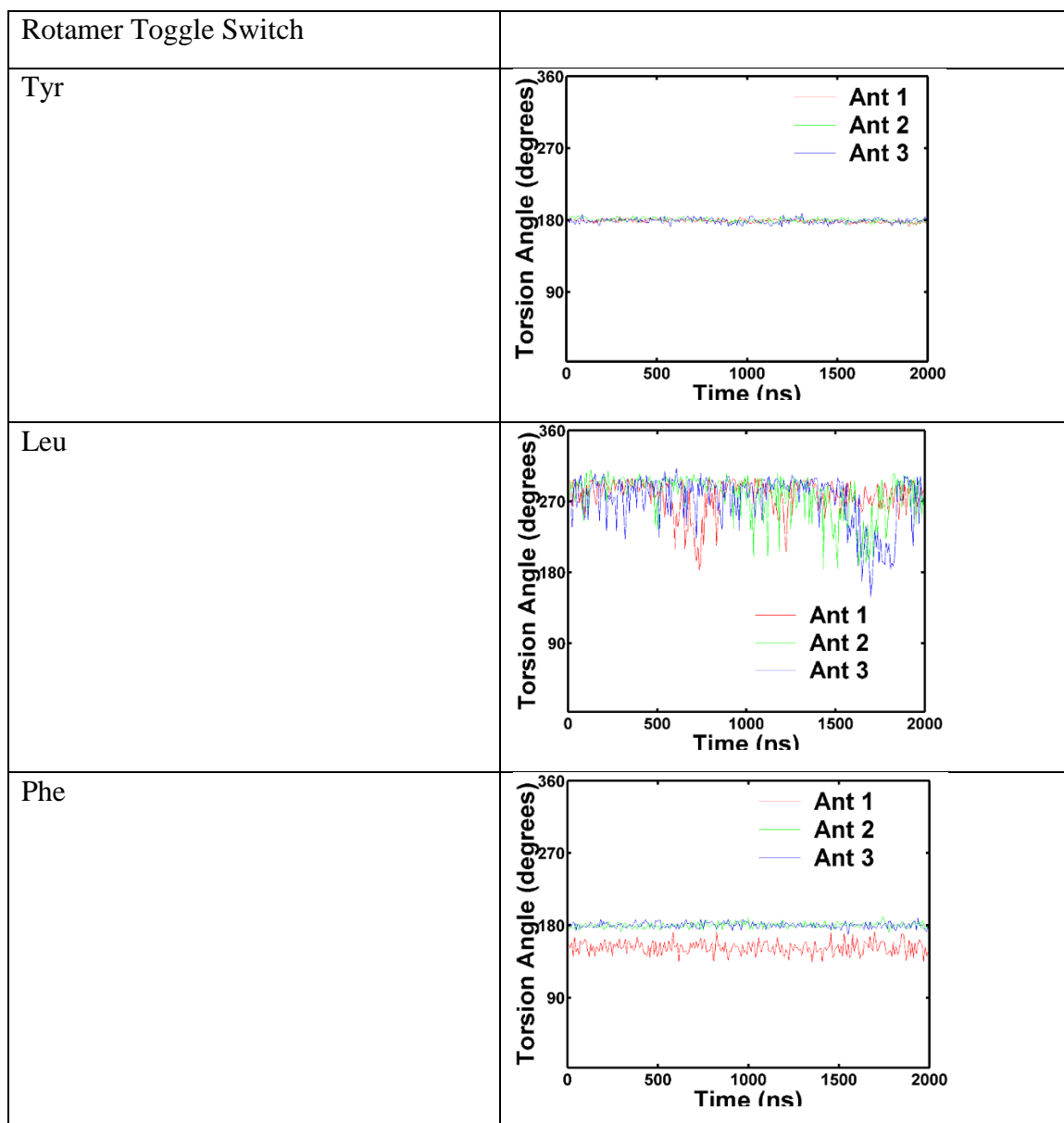


Figure A56. Comparison of the rotamer toggle switches torsion angles across all three antagonist simulation trajectories after 2000ns. Red = trajectory 1, green = trajectory 2, blue = trajectory 3.

Table A5

Generic numbering of residues for the CRF1R, Glucagon receptor, Oxytocin receptor, Adrenoceptor, Adenosine A2a receptor.

CRF1R TM1	Glucagon Receptor TM1	Oxytocin receptor TM1	Adrenoceptor TM1	Adenosine A2a TM1
1x19	1x25	1x32	1x25	1x27
1x20	1x26	1x33	1x26	1x28
1x21	1x27	1x34	1x27	1x29
1x22	1x28	1x35	1x28	1x30
1x23	1x29	1x36	1x29	1x31
1x24	1x30	1x37	1x30	1x32
1x25	1x31	1x38	1x31	1x33
1x26	1x32	1x39	1x32	1x34
1x27	1x33	1x40	1x33	1x35
1x28	1x34	1x41	1x34	1x36
1x29	1x35	1x42	1x35	1x37
1x30	1x36	1x43	1x36	1x38
1x31	1x37	1x44	1x37	1x39
1x32	1x38	1x45	1x38	1x40
1x33	1x39	1x46	1x39	1x41
1x34	1x40	1x47	1x40	1x42
1x35	1x41	1x48	1x41	1x43
1x36	1x42	1x49	1x42	1x44
1x37	1x43	1x50	1x43	1x45
1x38	1x44	1x51	1x44	1x46
1x39	1x45	1x52	1x45	1x47
1x40	1x46	1x53	1x46	1x48
1x41	1x47	1x54	1x47	1x49
1x42	1x48	1x55	1x48	1x50
1x43	1x49	1x56	1x49	1x51
1x44	1x50	1x57	1x50	1x52
1x45	1x51	1x58	1x51	1x53
1x46	1x52	1x59	1x52	1x54
1x47	1x53	1x60	1x53	1x55
1x48	1x54		1x54	1x56
1x49	1x55		1x55	1x57
1x50	1x56		1x56	1x58
1x51	1x57		1x57	1x59
1x52	1x58		1x58	1x60
1x53	1x59		1x59	TM2

Table A5 (Continued)

1x54	1x60		1x60	2x37
1x55	TM2	TM2	TM2	2x38
1x56	2x37	2x37	2x37	2x39
1x57	2x38	2x38	2x38	2x40
1x58	2x39	2x39	2x39	2x41
1x59	2x40	2x40	2x40	2x42
1x60	2x41	2x41	2x41	2x43
TM2	2x42	2x42	2x42	2x44
2x37	2x43	2x43	2x43	2x45
2x38	2x44	2x44	2x44	2x46
2x39	2x45	2x45	2x45	2x47
2x40	2x46	2x46	2x46	2x48
2x41	2x47	2x47	2x47	2x49
2x42	2x48	2x48	2x48	2x50
2x43	2x49	2x49	2x49	2x51
2x44	2x50	2x50	2x50	2x52
2x45	2x51	2x51	2x51	2x53
2x46	2x52	2x52	2x52	2x54
2x47	2x53	2x53	2x53	2x55
2x48	2x54	2x54	2x54	2x551
2x49	2x55	2x55	2x55	2x56
2x50	2x56	2x56	2x551	2x57
2x51	2x57	2x57	2x56	2x58
2x52	2x58	2x58	2x57	2x59
2x53	2x59	2x59	2x58	2x60
2x54	2x60	2x60	2x59	2x61
2x55	2x61	2x61	2x60	2x62
2x56	2x62	2x62	2x61	2x63
2x57	2x63	2x63	2x62	2x64
2x58	2x64	2x64	2x63	2x65
2x59	2x65	2x65	2x64	2x66
2x60	2x66	2x66	2x65	TM3
2x61	2x67	TM3	2x66	3x21
2x62	TM3	3x19	TM3	3x22
2x63	3x18	3x20	3x21	3x23
TM3	3x19	3x21	3x22	3x24
3x22	3x20	3x22	3x23	3x25
3x23	3x21	3x23	3x24	3x26
3x24	3x22	3x24	3x25	3x27
3x25	3x23	3x25	3x26	3x28
3x26	3x24	3x26	3x27	3x29

Table A5 (Continued)

3x27	3x25	3x27	3x28	3x30
3x28	3x26	3x28	3x29	3x31
3x29	3x27	3x29	3x30	3x32
3x30	3x28	3x30	3x31	3x33
3x31	3x29	3x31	3x32	3x34
3x32	3x30	3x32	3x33	3x35
3x33	3x31	3x33	3x34	3x36
3x34	3x32	3x34	3x35	3x37
3x35	3x33	3x35	3x36	3x38
3x36	3x34	3x36	3x37	3x39
3x37	3x35	3x37	3x38	3x40
3x38	3x36	3x38	3x39	3x41
3x39	3x37	3x39	3x40	3x42
3x40	3x38	3x40	3x41	3x43
3x41	3x39	3x41	3x42	3x44
3x42	3x40	3x42	3x43	3x45
3x43	3x41	3x43	3x44	3x46
3x44	3x42	3x44	3x45	3x47
3x45	3x43	3x45	3x46	3x48
3x46	3x44	3x46	3x47	3x49
3x47	3x45	3x47	3x48	3x50
3x48	3x46	3x48	3x49	3x51
3x49	3x47	3x49	3x50	3x52
3x50	3x48	3x50	3x51	3x53
3x51	3x49	3x51	3x52	3x54
3x52	3x50	3x52	3x53	3x55
3x53	3x51	3x53	3x54	3x56
3x54	3x52	3x54	3x55	TM4
3x55	3x53	3x55	3x56	4x38
3x56	3x54	3x56	TM4	4x39
TM4	3x55		4x38	4x40
4x41	3x56		4x39	4x41
4x42	TM4	TM4	4x40	4x42
4x43	4x41	4x38	4x41	4x43
4x44	4x42	4x39	4x42	4x44
4x45	4x43	4x40	4x43	4x45
4x46	4x44	4x41	4x44	4x46
4x47	4x45	4x42	4x45	4x47
4x48	4x46	4x43	4x46	4x48
4x49	4x47	4x44	4x47	4x49
4x49I	4x48	4x45	4x48	4x50

Table A5 (Continued)

4x50	4x49	4x46	4x49	4x51
4x51	4x491	4x47	4x50	4x52
4x52	4x50	4x48	4x51	4x53
4x53	4x51	4x49	4x52	4x54
4x54	4x52	4x50	4x53	4x55
4x55	4x53	4x51	4x54	4x56
4x56	4x54	4x52	4x55	4x57
4x57	4x55	4x53	4x56	4x58
4x58	4x56	4x54	4x57	4x59
4x59	4x57	4x55	4x58	4x60
4x60	4x58	4x56	4x59	4x61
4x61	4x59	4x57	4x60	4x62
4x62	4x60	4x59	4x61	4x63
4x63	4x61	4x60	4x62	TM5
4x64	4x62	4x61	4x63	5x36
4x65	4x63	4x62	4x64	5x37
4x66	4x64	4x63	TM5	5x38
4x67	4x65	4x64	5x36	5x39
4x68	4x66	4x65	5x37	5x40
TM5	4x67		5x38	5x41
5x40	4x68		5x39	5x411
5x41	TM5	TM5	5x40	5x42
5x42	5x37	5x33	5x41	5x43
5x421	5x38	5x34	5x42	5x44
5x43	5x39	5x35	5x43	5x45
5x44	5x40	5x36	5x44	5x46
5x45	5x41	5x37	5x45	5x461
5x46	5x42	5x38	5x46	5x47
5x47	5x421	5x39	5x461	5x48
5x48	5x43	5x40	5x47	5x49
5x49	5x44	5x41	5x48	5x50
5x50	5x45	5x42	5x49	5x51
5x51	5x46	5x43	5x50	5x52
5x52	5x47	5x44	5x51	5x53
5x53	5x48	5x45	5x52	5x54
5x54	5x49	5x46	5x53	5x55
5x55	5x50	5x461	5x54	5x56
5x56	5x51	5x47	5x55	5x57
5x57	5x52	5x48	5x56	5x58
5x58	5x53	5x49	5x57	5x59
5x59	5x54	5x50	5x58	5x60

Table A5 (Continued)

5x60	5x55	5x51	5x59	5x61
5x61	5x56	5x52	5x60	5x62
5x62	5x57	5x53	5x61	5x63
5x63	5x58	5x54	5x62	5x64
5x64	5x59	5x55	5x63	5x65
5x65	5x60	5x56	5x64	5x66
5x66	5x61	5x57	5x65	5x67
5x67	5x62	5x58	5x66	5x68
5x68	5x63	5x59	5x67	5x69
5x69	5x64	5x60	5x68	5x70
TM6	5x65	5x61	5x69	5x71
6x25	5x66	5x62	5x70	5x72
6x26	5x67	5x63	5x71	5x73
6x27	5x68	5x64	5x72	5x74
6x28	5x69	5x65	5x73	TM6
6x29	5x70	5x66	5x74	6x21
6x30	5x71	5x67	5x75	6x22
6x31	TM6	5x68	5x76	6x23
6x32	6x28	5x69	TM6	6x24
6x33	6x29	5x70	6x24	6x25
6x34	6x30	5x71	6x25	6x26
6x35	6x31	5x72	6x26	6x27
6x36	6x32	5x73	6x27	6x28
6x37	6x33	5x74	6x28	6x29
6x38	6x34	5x75	6x29	6x30
6x39	6x35	5x76	6x30	6x31
6x40	6x36	TM6	6x31	6x32
6x41	6x37	6x24	6x32	6x33
6x42	6x38	6x25	6x33	6x34
6x43	6x39	6x26	6x34	6x35
6x44	6x40	6x27	6x35	6x36
6x45	6x41	6x28	6x36	6x37
6x46	6x42	6x29	6x37	6x38
6x47	6x43	6x30	6x38	6x39
6x48	6x44	6x31	6x39	6x40
6x49	6x45	6x32	6x40	6x41
6x50	6x46	6x33	6x41	6x42
6x51	6x47	6x34	6x42	6x43
6x52	6x48	6x35	6x43	6x44
6x53	6x49	6x36	6x44	6x45
TM7	6x50	6x37	6x45	6x46

Table A5 (Continued)

7x28	6x51	6x38	6x46	6x47
7x29	6x52	6x39	6x47	6x48
7x30	6x53	6x40	6x48	6x49
7x31	TM7	6x41	6x49	6x50
7x32	7x26	6x42	6x50	6x51
7x33	7x27	6x43	6x51	6x52
7x34	7x28	6x44	6x52	6x53
7x35	7x29	6x45	6x53	6x54
7x36	7x30	6x46	6x54	6x55
7x37	7x31	6x47	6x55	6x56
7x38	7x32	6x48	6x56	6x57
7x39	7x33	6x49	6x57	6x58
7x40	7x34	6x50	6x58	6x59
7x41	7x35	6x51	6x59	6x60
7x42	7x36	6x52	6x60	6x61
7x43	7x37	6x53	6x61	TM7
7x45	7x38	6x54	TM7	7x30
7x46	7x39	6x55	7x30	7x31
7x47	7x40	6x56	7x31	7x32
7x48	7x41	6x57	7x32	7x33
7x49	7x42	6x58	7x33	7x34
7x50	7x43	6x59	7x34	7x35
7x51	7x45	6x60	7x35	7x36
7x52	7x46	6x61	7x36	7x37
7x53	7x47	TM7	7x37	7x38
7x54	7x48	7x30	7x38	7x39
7x55	7x49	7x31	7x39	7x40
7x56	7x50	7x32	7x40	7x41
H8	7x51	7x33	7x41	7x42
8x47	7x52	7x34	7x42	7x43
8x48	7x53	7x35	7x43	7x45
8x49	7x54	7x36	7x45	7x46
8x50	7x55	7x37	7x46	7x47
8x51	7x56	7x38	7x47	7x48
8x52	H8	7x39	7x48	7x49
8x53	8x47	7x40	7x49	7x50
8x54	8x48	7x41	7x50	7x51
8x55	8x49	7x42	7x51	7x52
8x56	8x50	7x43	7x52	7x53
8x57	8x51	7x45	7x53	7x54
8x58	8x52	7x46	7x54	7x55

Table A5 (Continued)

8x59	8x53	7x47	7x55	7x56
8x60	8x54	7x48	H8	H8
8x61	8x55	7x49	8x47	8x47
8x62	8x56	7x50	8x48	8x48
8x63	8x57	7x51	8x49	8x49
8x64	8x58	7x52	8x50	8x50
8x65	8x59	7x53	8x51	8x51
8x66	8x60	7x54	8x52	8x52
8x67	8x61	7x55	8x53	8x53
8x68	8x62	7x56	8x54	8x54
8x69	8x63	H8	8x55	8x55
8x70	8x64	8x47	8x56	8x56
8x71	8x65	8x48	8x57	8x57
8x72	8x66	8x49	8x58	8x58
	8x67	8x50	8x59	8x59
	8x68	8x51		8x60
	8x69	8x52		8x61
	8x70	8x53		8x62
	8x71	8x54		8x63
	8x72	8x55		8x64
		8x56		8x65
		8x57		8x66
		8x58		8x67
				8x68
				8x69

Appendix B

GLP-1 Receptor in Complex with a Full Agonist and a Biased Agonist Probed by Molecular Dynamics Simulations for the Development of more Specific Drugs on Type 2 Diabetes

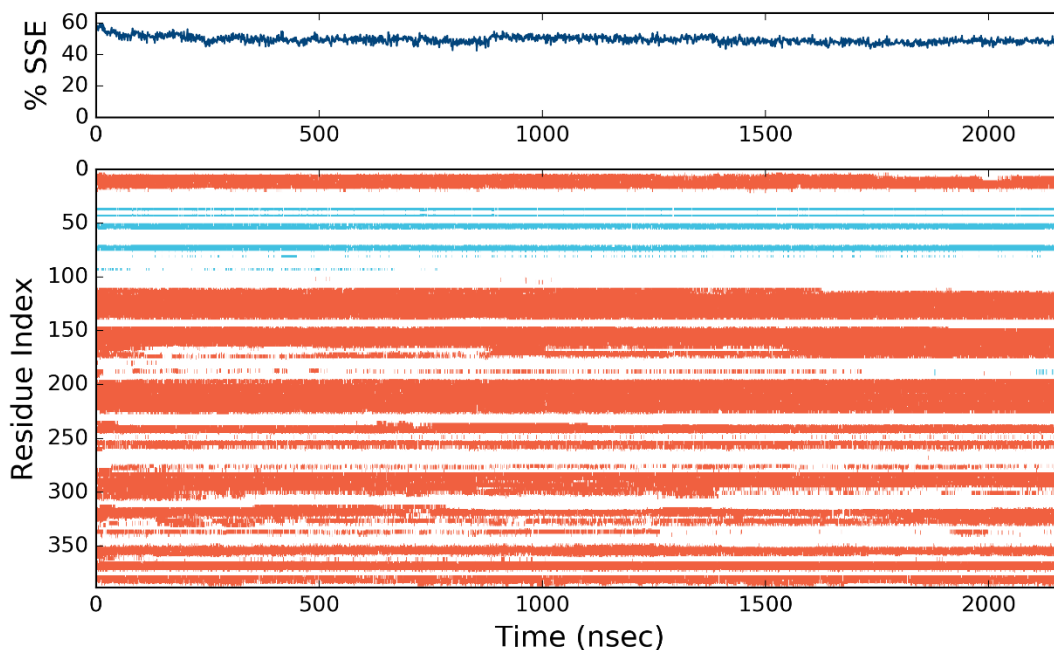


Figure B1. Protein secondary structures timeline for the full agonist system. The top plot summarizes the SSE composition for each trajectory frame over the course of the simulation. The bottom plot monitors each residue and its SSE over time.

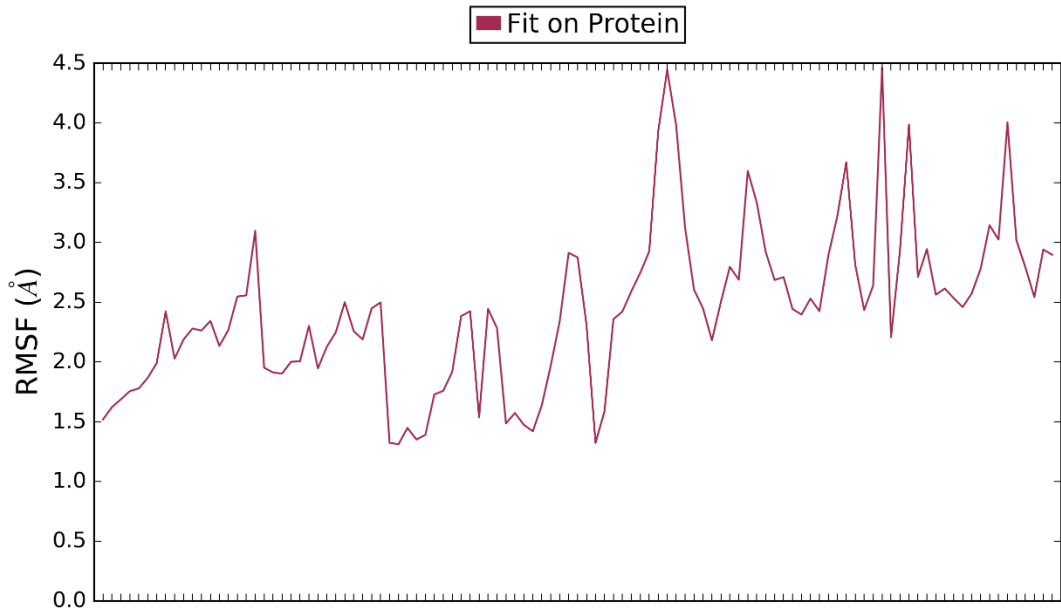


Figure B2. Ligand RMSF for the full unbiased agonist. Ligand fluctuations broken down by atom.

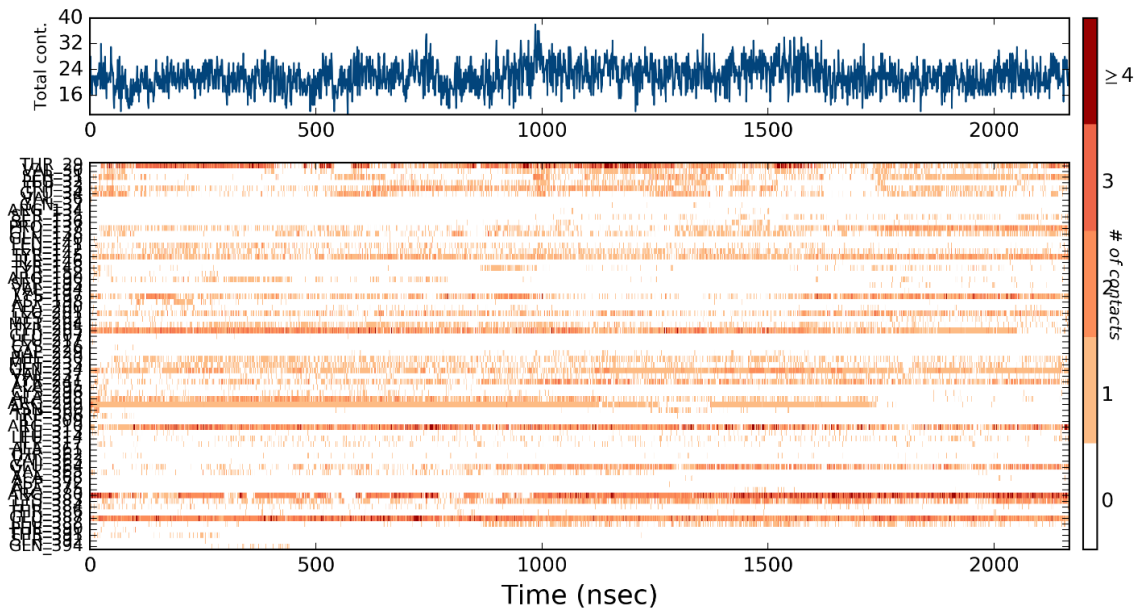


Figure B4. Protein-ligand contacts for the full agonist over time. The top panel shows the total number of specific contacts the protein makes with the ligand over time. The bottom panel shows which residues interact with the ligand in each trajectory frame. Darker orange indicates residues making more than one specific contact.

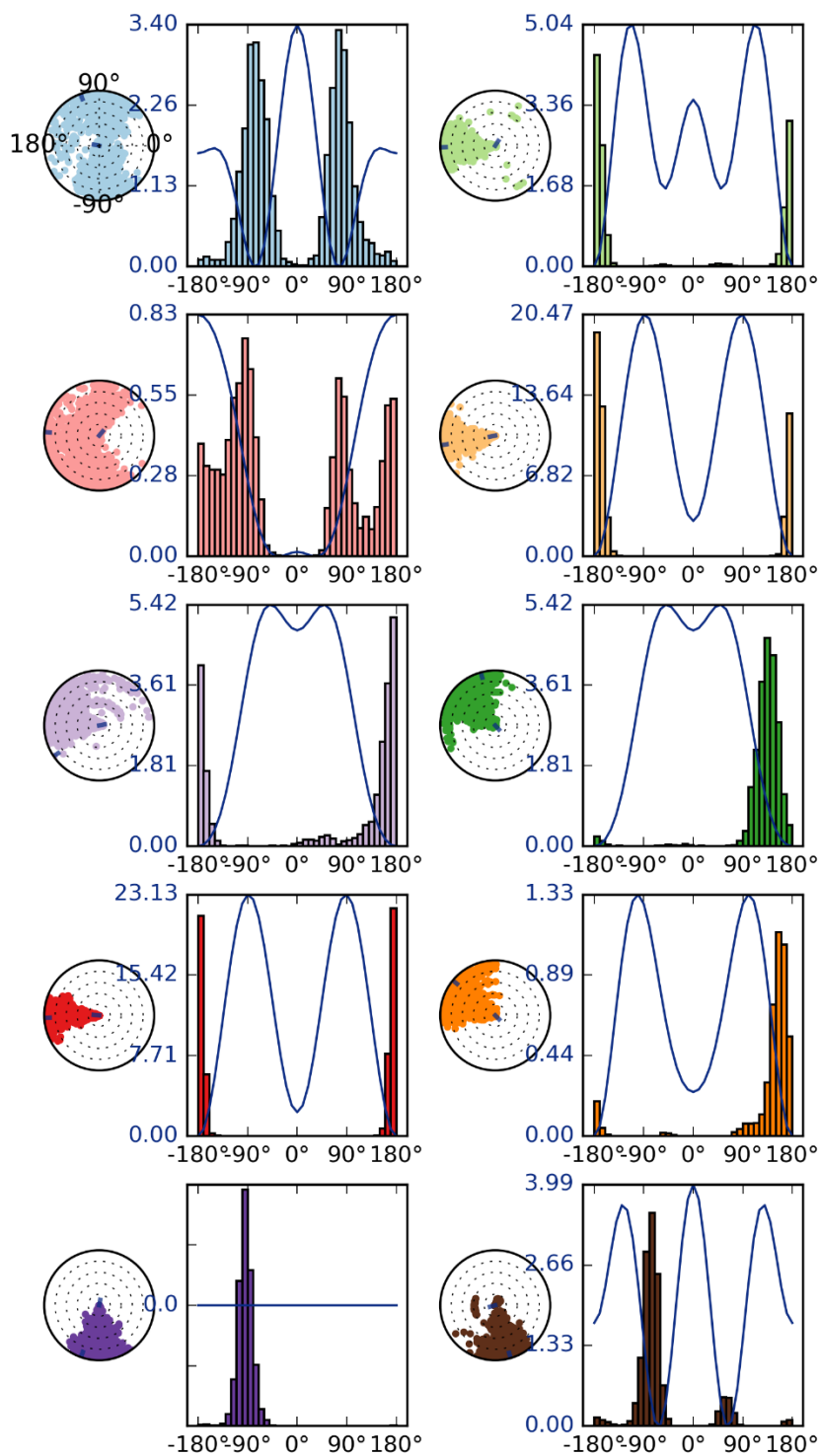


Figure B5. Ligand torsions profile for the full unbiased agonist. Summarizes the conformational evolution of every rotatable bond in the ligand throughout the trajectory.

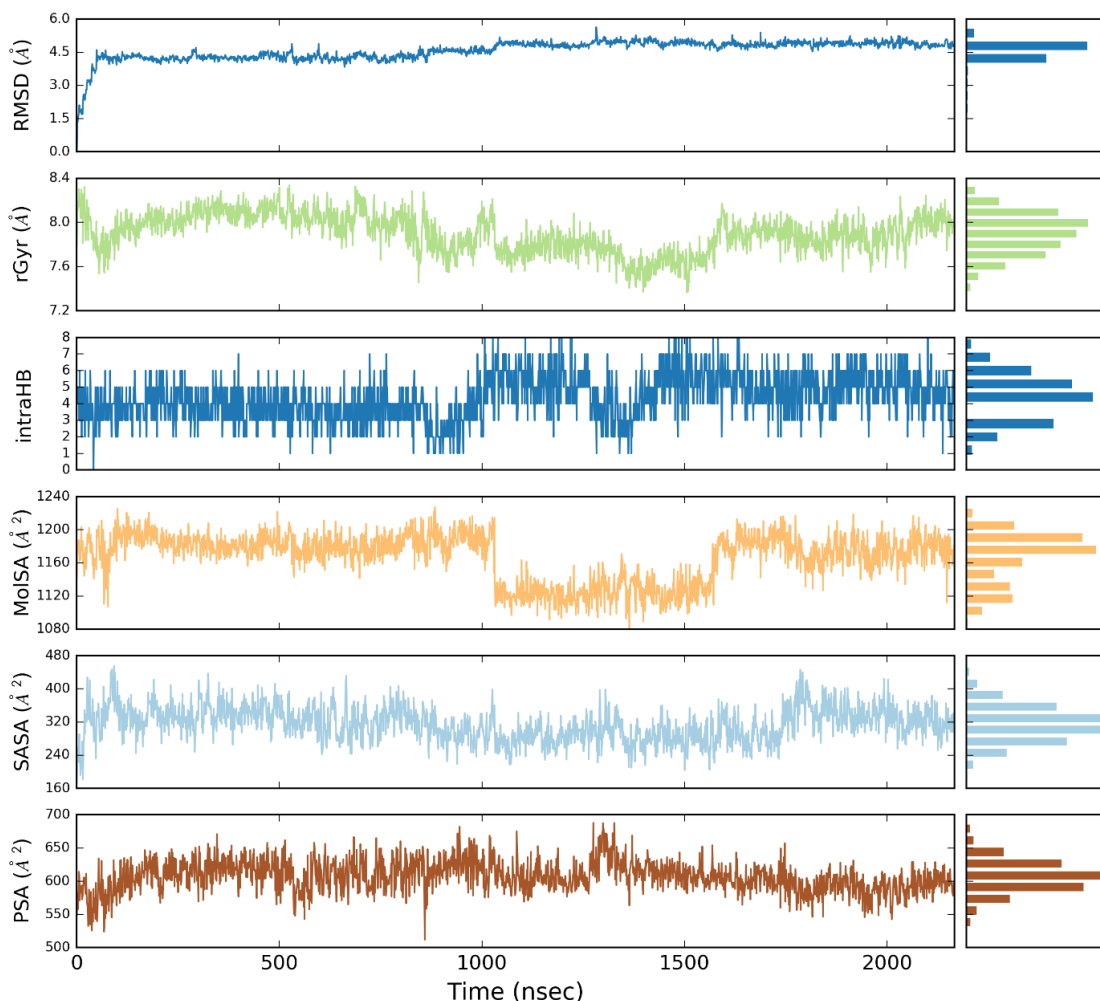


Figure B6. Ligand properties of the full agonist. **Ligand RMSD**: Root mean square deviation of a ligand with respect to the reference conformation (typically the first frame is used as the reference and it is regarded as time $t=0$). **Radius of Gyration (rGyr)**: Measures the 'extendedness' of a ligand, and is equivalent to its principal moment of inertia. **Intramolecular Hydrogen Bonds (intraHB)**: Number of internal hydrogen bonds (HB) within a ligand molecule. **Molecular Surface Area (MolSA)**: Molecular surface calculation with 1.4 Å probe radius. This value is equivalent to a van der Waals surface area. **Solvent Accessible Surface Area (SASA)**: Surface area of a molecule accessible by a water molecule. **Polar Surface Area (PSA)**: Solvent accessible surface area in a molecule contributed only by oxygen and nitrogen atoms.

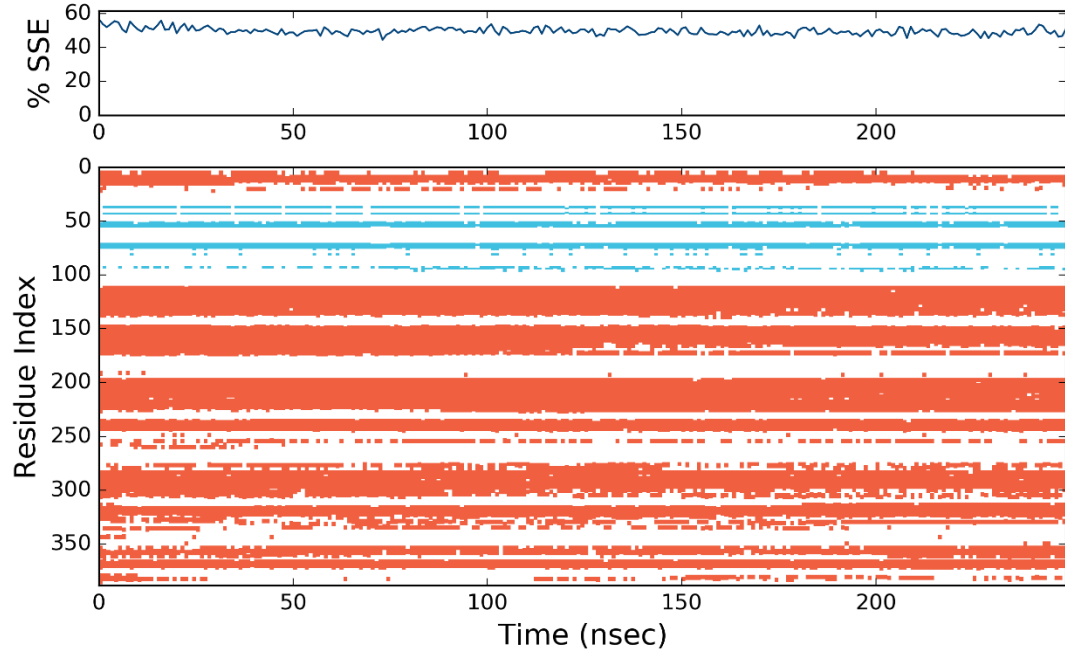


Figure B7. Protein secondary structures timeline for the biased agonists system. The top plot summarizes the SSE composition for each trajectory frame over the course of the simulation. The bottom plot monitors each residue and its SSE over time.

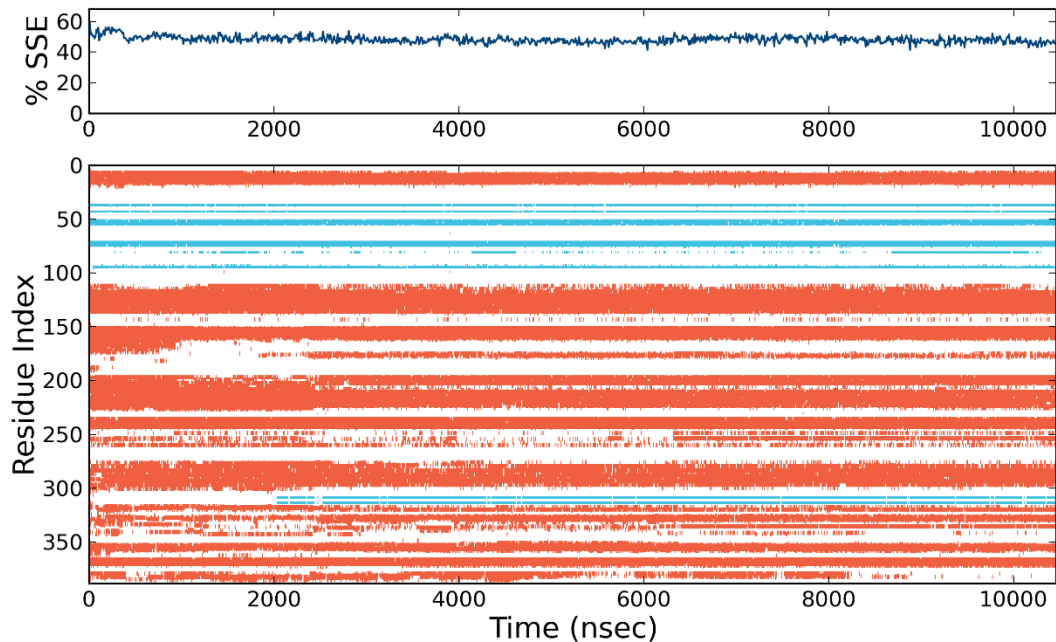


Figure B8. Protein secondary structures timeline for the APO form system. The top plot summarizes the SSE composition for each trajectory frame over the course of the simulation. The bottom plot monitors each residue and its SSE over time.

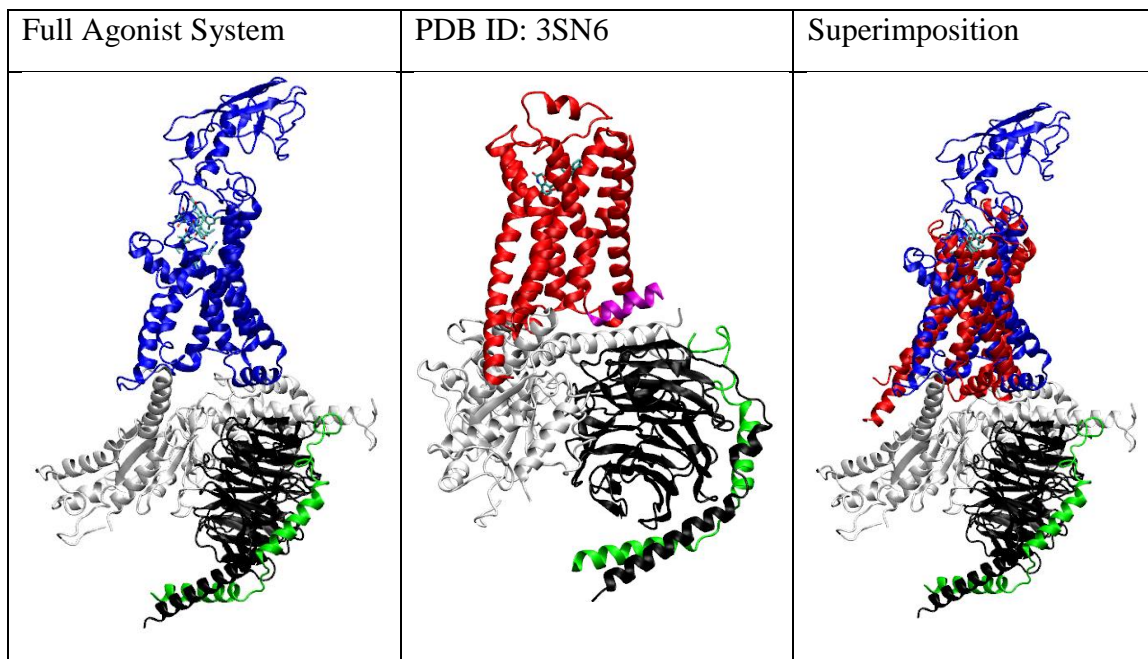


Figure B9. Comparison between our full unbiased agonist system and the solved structure of G protein docked Adrenergic receptor obtained from PDB.

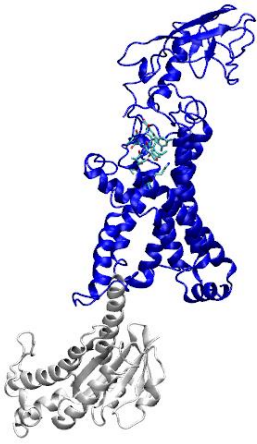
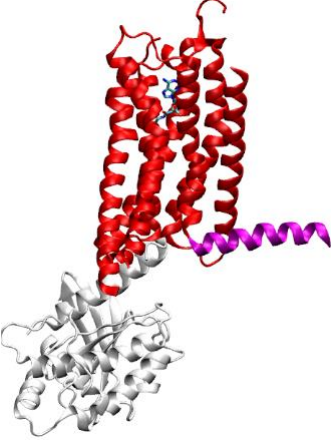
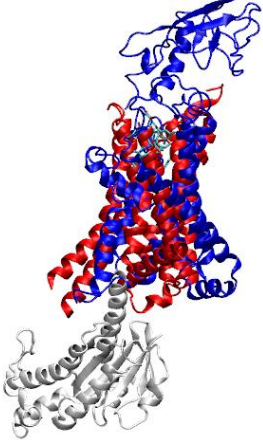
Full Agonist System	PDB ID: 5G53	Superimposition
		

Figure B10. Comparison between our full unbiased agonist complex and the solved structure for G protein docked Adenosine A2A receptor obtained from PDB.

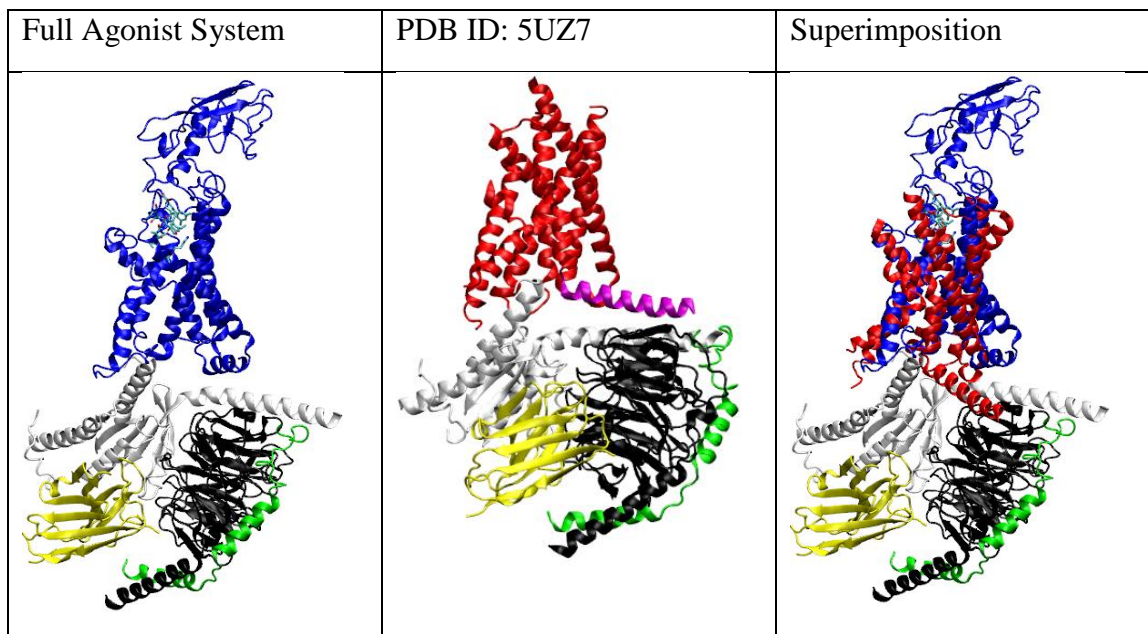


Figure B11. Comparison between our full unbiased agonist complex and the solved structure for G protein docked Calcitonin receptor obtained from PDB.

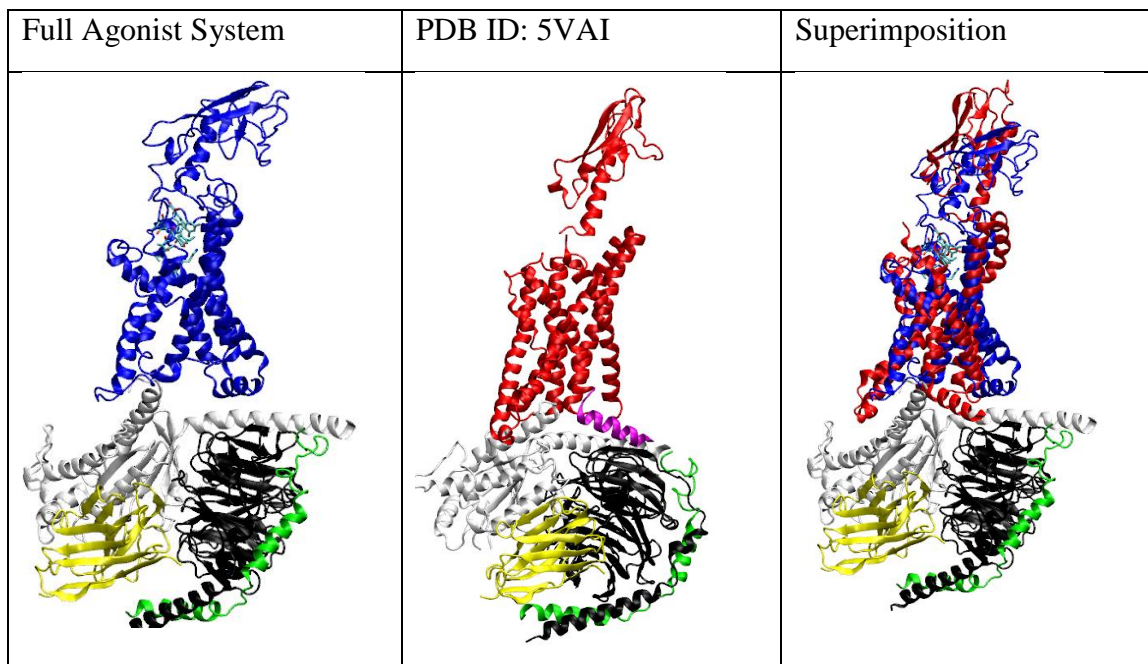


Figure B12. Comparison between our full unbiased agonist complex and the solved structure for G protein docked Glucagon-like peptide-1 receptor (5vai) obtained from PDB.

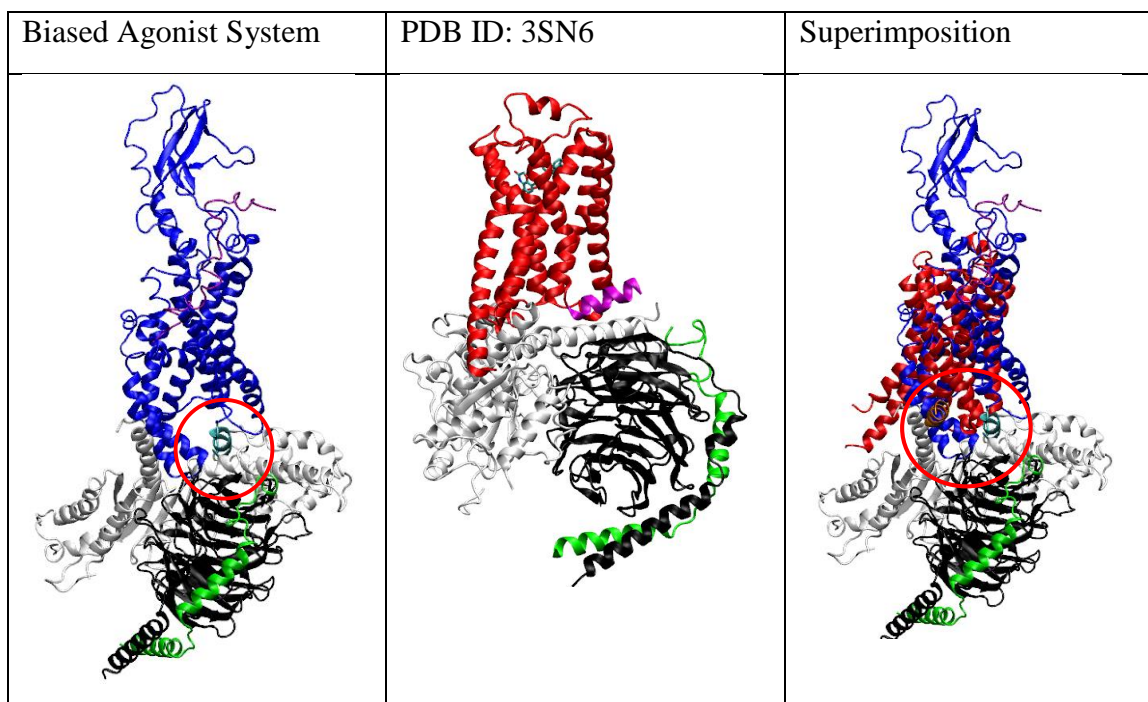


Figure B13. Comparison between our biased agonist complex and the solved structure for G protein docked Adrenergic receptor obtained from PDB. The C-terminal helix is circled in red and is shown to not clash with the G protein.

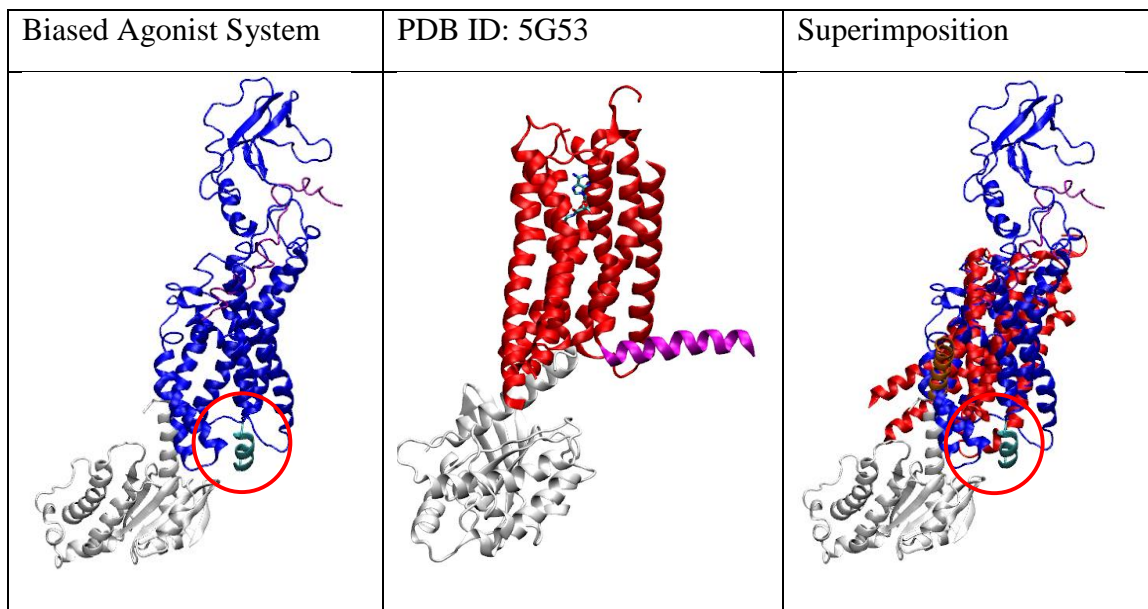


Figure B14. Comparison between our biased agonist complex and the solved structure for G protein docked Adenosine A2A receptor obtained from PDB. The C-terminal helix is circled in red and shown to not clash with the G protein.

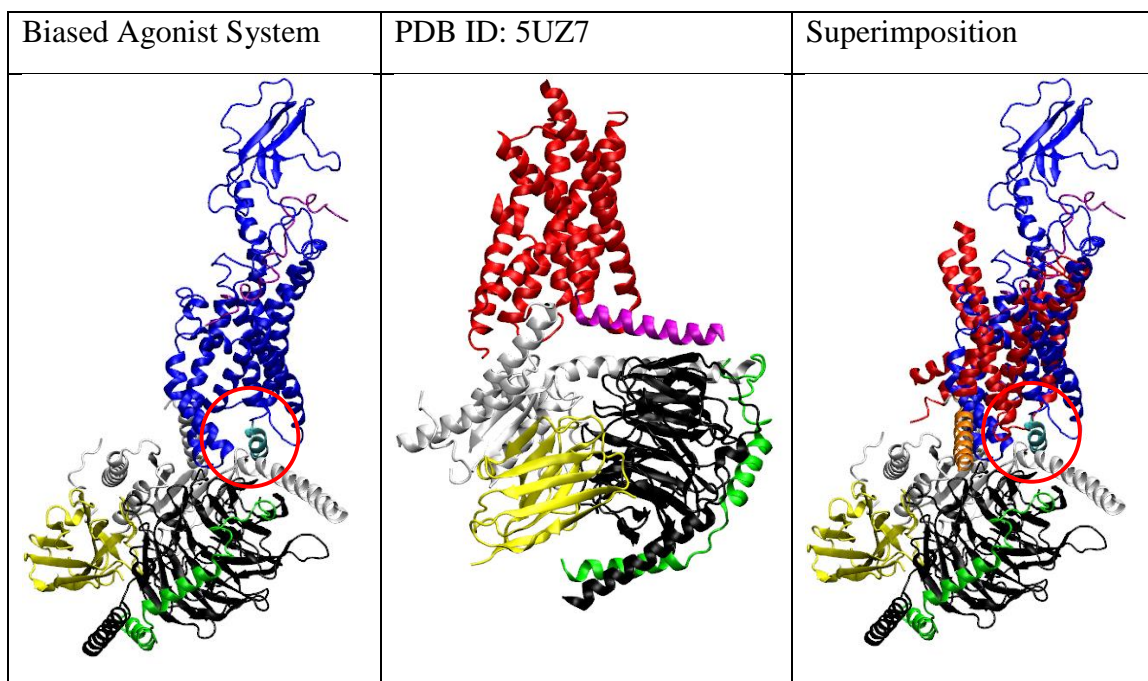


Figure B15. Comparison between our biased agonist complex and the solved structure for G protein docked Calcitonin receptor obtained from the PDB. The C-terminal helix is circled in red and shown to not clash with the G protein.

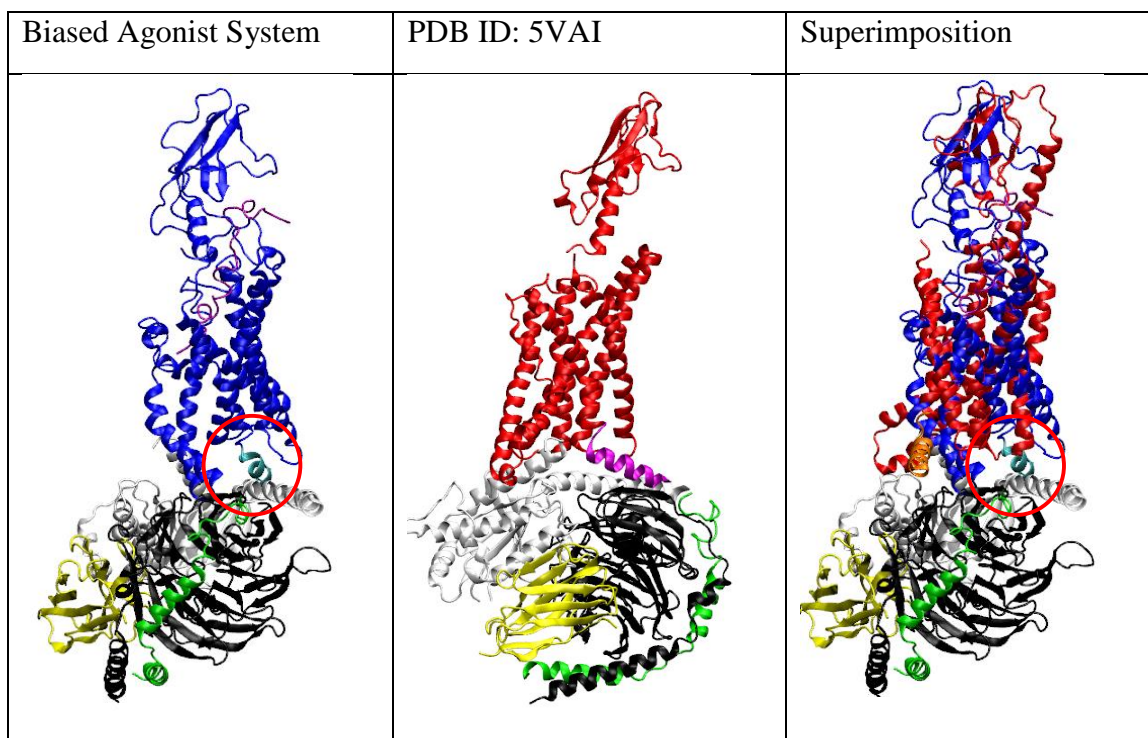


Figure B16. Comparison between our biased agonist complex and the solved structure of G protein docked Calcitonin receptor obtained from PDB. The C-terminal helix is circled in red and shown to not clash with the G protein.

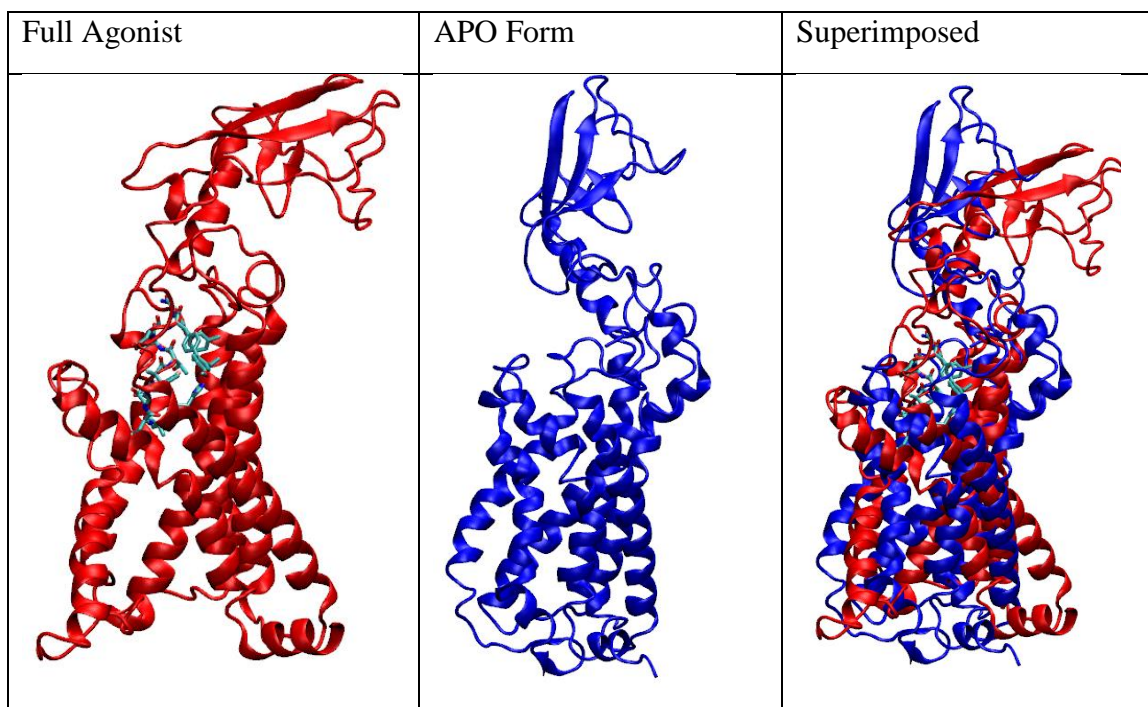


Figure S17. Comparison of conformational change between full unbiased agonist and APO systems after MD simulations. The most abundant conformations are shown here.

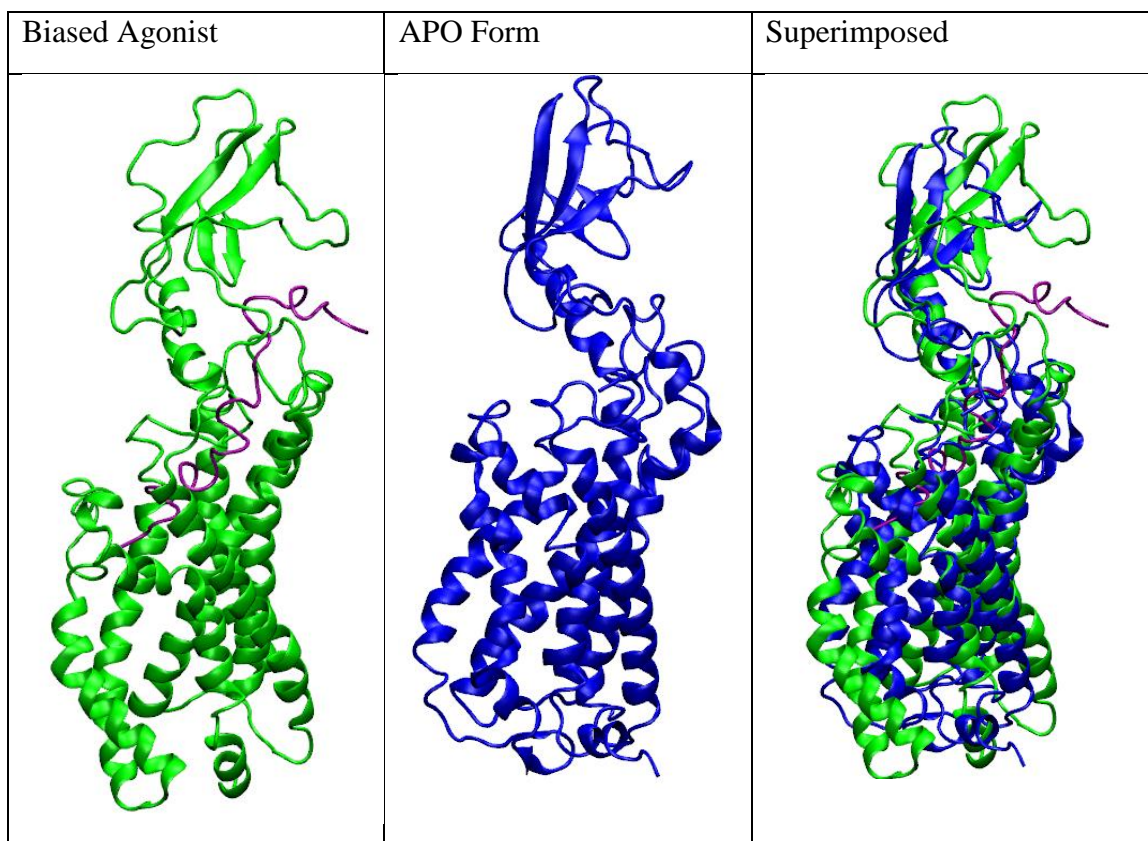


Figure B18. Comparison of conformational change between biased agonist and APO systems after MD simulations. The most abundant conformations are shown here.



Figure B19. Pairwise comparison of all ECDs after MD simulations to observe conformational differences.

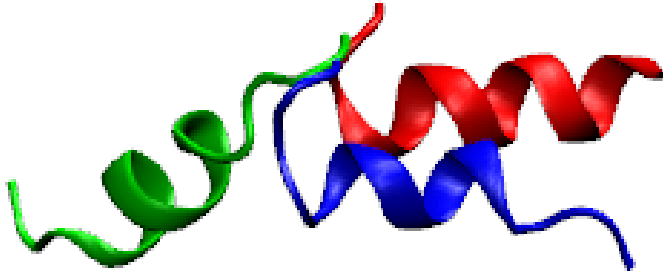


Figure B20. Comparison of all C-terminal helices to observe conformational differences. Red = full unbiased agonist, Blue = APO, Green = biased agonist.