

COMPUTATIONAL STUDIES OF ORPHAN G PROTEIN-COUPLED RECEPTORS

Thesis by

Jiyoung Heo

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy



California Institute of Technology

Pasadena, California

2007

(Defended October 30, 2006)

© 2007

Jiyoung Heo

All Rights Reserved

Acknowledgments

I am very fortunate to have enjoyed my graduate studies under the wonderful research environment at Caltech. Working with many intelligent and knowledgeable people here has been another great honor to me. First, I would like to thank my advisor Bill Goddard for his support during my PhD studies. He was always willing to discuss my research work with me and to guide my research direction. His long-time passion for science was inspiring. I also thank Mel Simon for suggesting this great project, for helpful discussion and for letting me use his facilities. I thank my committee members, Professor Jack Beauchamp, Doug Rees and Judy Campbell who came across campus on candidacy and proposition exams, discussing my research work and providing valuable feedback.

I am indebted to many people in the Goddard group. I especially thank Nagarajan Vaidehi for managing the project until she left the group. Bill gave me the big picture and insights, and she helped me a lot to start my research project and figure out practical problems. I thank Wely, Rene and Spencer for their past contribution to the methods that I used for my research. I thank John (Wendel), Victor and Pete for their effort in developing methods, Scott at UNC for collaboration on screening, Ravi for helping me to start NAMD runs, and other former and present Biogroup people for valuable discussions.

Three terrific Korean post-docs in the Simon group were essential in my experimental work. I especially thank Sang-kyou for training me to perform biology experiments. I learned most experimental techniques from him. I also thank Jong-Ik and Keum-Joo for additional help and comments in experiments.

I thank my 056C officemates Julius, Sam, Victor, Santiago, Candy and John (Keith) for support and friendship. I also thank WAG Koreans, Seung-Soon, Yunhee, Hyon-Jee and Eun-Jung (a pseudo-WAG Korean). I will not forget the time we chatted in Korean about science and

everything. Thanks to my friends that I met outside the lab, In-Jung and Junghi, I could feel the Korean culture (called “Jung”) that I have missed during my stay in the U.S.

I am so grateful to my parents, my sisters, and my brother who have been always supportive through my whole life. I would like to thank my parents-in-law for taking care of their grandson in Korea. Without their sacrifice, I could not be devoted in my research to finish my PhD studies.

Finally, my deepest gratitude goes to Nam-Joon, my husband, for his steadfast love and support.

Abstract

G protein-coupled receptors (GPCRs) play an essential role in cell communications and sensory functions. Consequently, they are involved in wide variety of diseases and are targets for many drug therapies. Particularly important is the large number of orphan GPCRs, which may play important, albeit unknown, functions in various cells. To understand their respective physiological roles, it is important to identify their endogenous ligands, and to find small molecule ligands that would serve as selective agonists or antagonists. The *mas-related gene G protein-coupled receptors* (Mrg receptors) belong to the orphan GPCR family, which is expressed in a specific subset of sensory neurons known to detect painful stimuli, suggesting that they could be involved in pain sensation or modulation.

The primary focus of this thesis is to predict the 3D structure and binding site of Mrg receptors and to identify novel ligands that would be potential agonists or antagonists. We predict the 3D structure for the mouse MrgC11 (mMrgC11) and the binding site for five chiral FMRF-NH₂ ligands. We correctly predict the relative binding observed for these five ligands. We find that Tyr110 (TM3), Asp161 (TM4), and Asp179 (TM5) are particularly important to binding the ligands. Subsequently, we carry out mutagenesis experiments followed by intracellular calcium release assays that demonstrate the dramatic decrease in activity for the Y110A, D161A, and D179A mutants predicted by our model.

The all-atom molecular dynamics simulation of the mMrgC11/F-(D)M-R-F-NH₂ complex structure in explicit water and infinite lipid membrane system shows that some conformational fluctuations are present, but no significant instability is detected, thus validating our structure prediction method.

The virtual screening with the combination of QSPR and docking methods is carried out for the predicted mMrgC11 receptor. The compounds showing the antagonistic effect are

identified by competitive functional assays. These hit compounds are certainly good starting points in designing better agonists or antagonists.

The binding site of rat MrgA receptor that shows differential binding between adenine and guanine is also predicted. The predicted binding affinity correlates with the availability of the hydrogen bonds to two Asn residues, which would be primary mutation candidates to validate the structure.

Table of Contents

Acknowledgments	iii
Abstract	v
Table of Contents	vii
Figures and Tables	xi
Chapter 1 Introduction	
1.1 G protein-coupled receptors	2
1.2 Orphan GPCRs and deorphanization	4
1.3 The 3D structure of GPCR and molecular modeling	7
1.3.1 Hydrophobicity scale: TM prediction from the primary sequence	8
1.3.2 Force field	9
1.3.3 Molecular mechanics	11
1.3.4 Molecular dynamics	14
1.3.5 Molecular docking	16
1.4 Outlines of Thesis	22
References	23
Chapter 2 Prediction of the 3D Structure for FMRF-amide Peptides Bound to Mouse MrgC11 Receptor with Subsequent Experimental Verification	
2.1 Introduction	26
2.2 Computational methods	27
2.2.1 Structure predictions of the Mrg receptor	27
2.2.2 Docking predictions with peptide ligands	31
2.3 Experimental procedures	37
2.3.1 <i>In vitro</i> mutagenesis	37
2.3.2 Cell culture and transfection	37

2.3.3 Biotinylation and immunoprecipitation	37
2.3.4 Intracellular calcium assay	38
2.4 Results and discussion	39
2.4.1 Characteristics of the predicted mMrgC11 receptor structure	39
2.4.2 Description of the peptide binding sites	43
2.4.3 Mutagenesis experimental results	49
2.4.4 Prediction of the structure of the mMrgA1 receptor and the binding site for ligands	54
2.4.5 Comparison of Mrg sequences	57
2.5 Summary and conclusions	58
References	60
Supporting figures and tables	63
Chapter 3 Molecular Dynamics Simulation of Mouse MrgC11 Receptor with Bound F-(D)M-R-F-NH₂ in Explicit Lipid/Water Environment	
3.1 Introduction	75
3.2 Simulation procedure	76
3.2.1 Setup of lipid and water environment	76
3.2.2 Molecular dynamics simulations	77
3.3 Results and discussion	79
3.3.1 Comparison between initial and final structures	79
3.3.2 Dynamic behavior in receptor conformation during MD simulation	81
3.3.3 Binding mode of F-(D)M-R-F-NH ₂ after equilibration	83
3.3.4 Time profile of receptor-ligand interactions	87
3.3.5 Time profile of interhelical interactions	95
3.4 Summary and conclusions	98
References	100

Chapter 4 Virtual Ligand Screening of Chemical Libraries for Mouse MrgC11 Receptor:**Combination of QSPR and Docking Methods**

4.1 Introduction	102
4.2 Materials and methods	104
4.2.1 Prescreening of compounds in chemical libraries	104
4.2.2 Chemical libraries	109
4.2.3 Molecular docking	110
4.2.4 Selection of final hits	111
4.2.5 Intracellular calcium release assay	113
4.2.6 Virtual screening of tetra-peptide binding site	113
4.3 Results and discussion	115
4.3.1 Hit compounds from virtual screening	116
4.3.2 Experimental activity test	116
4.3.3 Refined docking of MOL282 and design of its derivatives	118
4.3.4 Virtual screening for F-(D)M-R-F-NH ₂ bound site	122
4.4 Summary and conclusions	125
References	126
Supporting figures	129

Chapter 5 Prediction of the 3D Structure of Rat MrgA G Protein-Coupled Receptor and**Identification of its Binding Site**

5.1 Introduction	139
5.2 Materials and methods	140
5.2.1 Molecular modeling of receptor structure	140
5.2.2 QM calculation of ligand tautomers	141
5.2.3 Prediction of the adenine binding site	143
5.2.4 Refinement of the binding mode of adenine	145

5.2.5 Docking of other adenine derivatives	145
5.3 Results and discussion	146
5.3.1 Characteristics of receptor structure	146
5.3.2 QM results of ligand tautomers	149
5.3.3 Binding modes of adenine and other ligands	149
5.3.4 Comparison of the adenine binding site in rMrgA to the nucleotide binding sites in adenosine receptors and purinergic receptors	159
5.3.5 Comparison to other MrgA orthologs	161
5.4 Summary and conclusions	162
References	163
Supporting figures and tables	166

Appendix A Stability of Oxidized Base and its Mispair in DNA: Quantum Mechanics

Calculation and Molecular Dynamics Simulation

Figures and Tables

Figure 1.1 Various ways in which membrane proteins associate with the lipid bilayer	2
Figure 1.2 Schematic diagram of the general structure of G protein-coupled receptors	3
Figure 1.3 Classical examples of GPCR signaling	6
Table 1.1 Eisenberg hydrophobicity scale	9
Figure 1.4 Hydrophobicity profile for mouse MrgC11 sequence set	10
Figure 1.5 Schematic representations of the six key contributions of molecular mechanics force field	12
Figure 1.6 Construction of molecular surface in 2D	17
Figure 1.7 A binding site represented as a collection of overlapping spheres	18
Figure 1.8 Matching algorithm in DOCK	20
Figure 1.9 Atom pre-organization and anchor selection	21
Figure 2.1 Predicted transmembrane (TM) regions	28
Figure 2.2 Scanning regions used to determine the binding sites for the mMrgC11 receptor	32
Figure 2.3 Comparison of the predicted 3D structure for the RFa/mMrgC11 complex with the x-ray crystal structure of retinal/rhodopsin	40
Figure 2.4 Interhelical hydrogen bond networks in the mMrgC11 receptor	41
Figure 2.5 Aromatic interactions in TM regions of mMrgC11 receptor	43
Figure 2.6 Predicted 5 Å binding pocket of the RFa and RF dipeptide agonists	44
Figure 2.7 Predicted 3D structure for the FMRFa/mMrgC11 complex	45
Figure 2.8 Predicted 5 Å binding site to mMrgC11 of the agonist tetra-peptides	47
Figure 2.9 Predicted 5 Å binding pocket of the non-agonist tetra-peptides	48
Figure 2.10 Expression of mMrgC11 wild type and mutant receptors in the Flp-In293 cells	50
Table 2.1 The EC ₅₀ values of various peptide ligands	51

Table 2.2 Binding constants (EC50 values in nM) of mutant mMrgC11 receptors from intracellular calcium assays	53
Figure 2.11 Comparison between mMrgC11 and mMrgA binding sites	55
Figure S2.1 Multiple sequence alignment for mMrgC11 with 27 homologous sequences	63
Figure S2.2 Hydrophobicity profile for mMrgC11 sequence set	66
Figure S2.3 Multiple alignments of 39 verified Mrg sequences	67
Table S2.1 Hit sequences from independent BLAST search of each TM	71
Table S2.2 Calculated energies of configurations generated in combinatorial rotations of TM3, 5 and 6	73
Table S2.3 Calculated binding energy and its component contribution for ligands in mMrgC11	74
Figure 3.1 Fully solvated mMrgC11/F-(D)M-R-F-NH ₂ complex in the membrane	78
Figure 3.2 The mMrgC11/F-(D)M-R-F-NH ₂ complex structure after 7 ns run	80
Figure 3.3 The RMSD fluctuation of C α atoms with respect to the final 7 ns structure	82
Figure 3.4 The 5 Å binding site of F-(D)M-R-F-NH ₂ in mMrgC11 receptor	84
Figure 3.5 Water molecules in 5 Å binding pocket	85
Figure 3.6 The RMSD fluctuation for ligand heavy atoms	86
Figure 3.7 Time profile of intermolecular hydrogen bond distance	88
Figure 3.8 Time profile of centroid-to-centroid distance between two aromatic residues	93
Figure 3.9 Interhelical hydrogen bond networks in the mMrgC11 receptor	96
Figure 3.10 Time profile of the distance between residues residing in different helices	97
Table 4.1 Electron-density-derived TAE descriptors; $\rho(r)$ represents the electron density distribution	105
Figure 4.1 TAE local average ionization potential (PIP) surface property and its histogram distribution	106
Figure 4.2 Delaunay tessellation of a collection of random points in 2D	107

Figure 4.3 Geometric criteria for the hydrogen bonds	111
Figure 4.4 The 5 Å binding pocket of mMrgC11 receptor optimized with the di-peptide agonist, R-F-OH	115
Table 4.2 Inhibitory constant 50 % (IC50) of hit compounds	117
Figure 4.5 Compounds showing the inhibitory effect	117
Figure 4.6 Histograms of energy and RMSD distribution for 7,776 conformations of MOL282 in grid search	119
Figure 4.7 The 5 Å binding pocket of MOL282 in mMrgC11 receptor	120
Figure 4.8 Suggested better binders derived from MOL282	121
Figure 4.9 The 5 Å binding pocket of mMrgC11 receptor optimized with the tetra-peptide agonist, F-(D)M-R-F-NH ₂	122
Figure 4.10 The 5 Å binding sites of the best three hit compounds	123
Figure S4.1 Hit compounds from the first ligand set after docking	129
Figure S4.2 Hit compounds from the second set after docking	133
Figure S4.3 Hit compounds after virtual screening for the tetra-peptide binding site	135
Figure 5.1 Sequence alignment provided as an input for the homology modeling of rMrgA	140
Figure 5.2 Ligand compounds used in docking studies for the rMrgA receptor	142
Figure 5.3 Putative binding sites predicted from the HierDock scanning procedure	144
Figure 5.4 Predicted 3D structure of rMrgA receptor	147
Figure 5.5 Interhelical hydrogen bonds in rMrgA receptor	148
Figure 5.6 Predicted 5 Å binding pockets of adenine and guanine in the rMrgA receptor	151
Table 5.1 Decomposition of total intermolecular interaction between ligand and rMrgA receptor	153
Figure 5.7 The 5 Å binding pockets for various ligands in the rMrgA receptor	154
Figure 5.8 The 5 Å binding pockets of adenosine phosphates in the rMrgA receptor	155

Figure 5.9 Comparison of calculated binding energies with the experimental inhibition constants for rMrgA ligands	157
Table 5.2 Computational alanine-scanning results (SCAM) for adenine/rMrgA	158
Figure 5.10 Sequence alignment of rat MrgA receptor with other receptors known to bind adenine components of ligands	160
Figure S5.1 Multiple sequence alignment of rat MrgA with mouse MrgAs using Clustal-W	166
Table S5.1 The Gibbs free energies calculated from QM for various tautomeric forms of 1MA and 6BAP	167