

**MOLECULAR MODELING OF MUSCARINIC ACETYLCHOLINE  
RECEPTORS: STRUCTURAL BASIS OF LIGAND-RECEPTOR  
INTERACTIONS**

**CHIN SEK PENG**

**THESIS SUBMITTED IN FULFILMENT OF THE  
REQUIREMENTS FOR THE DEGREE OF DOCTOR OF  
PHILOSOPHY**

**FACULTY OF MEDICINE  
UNIVERSITY OF MALAYA  
KUALA LUMPUR**

**2015**

**UNIVERSITI MALAYA**  
**ORIGINAL LITERARY WORK DECLARATION**

Name of Candidate: CHIN SEK PENG

Registration/Matric No: MHA 090025

Name of Degree: Ph. D.

Title of Project Paper/Research Report/Dissertation/Thesis (“this Work”):

MOLECULAR MODELING OF MUSCARINIC ACETYLCHOLINE RECEPTORS:  
STRUCTURAL BASIS OF LIGAND-RECEPTOR INTERACTIONS

Field of Study: Molecular modeling

I do solemnly and sincerely declare that:

- (1) I am the sole author/writer of this Work;
- (2) This Work is original;
- (3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
- (4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
- (5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya (“UM”), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
- (6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate’s Signature

Date

Subscribed and solemnly declared before,

Witness’s Signature

Date

Name:

Designation:

## ABSTRACT

Muscarinic acetylcholine receptors (mAChRs), specifically of the M<sub>1</sub> subtype, have been the focus of significant drug discovery and development due to their potential roles in the pathophysiology of several central nervous system disorders, such as, Alzheimer's disease. Because of the conserved orthosteric binding pocket of mAChRs, identification of the selective activators/modulators have not been realized, and this often lead to undesired side effects from off-target activation. In this thesis, structural and dynamics studies of the M<sub>1</sub> mAChR using computational approaches are presented. Homology models of the M<sub>1</sub> mAChR were constructed and virtual screening experiments showed that the models could efficiently differentiate agonists from decoys, with the TM5-modified models also giving good agonist/antagonist selectivity. Molecular dynamics simulations further allowed the characterization of the dynamics profiles of different mAChR subtypes, bound to an agonist or antagonist and in apo form, leading to the elucidation of ligand affinity, selectivity, and possible allosteric pocket formation. Lastly, the models together with the crystal structures of the M<sub>2</sub> and M<sub>3</sub> mAChRs were used in virtual screening to identify potential selective M<sub>1</sub> mAChR binders. Of the 19 hits identified, 11 ligands targeted the orthosteric cavity, 7 portray bitopic characteristic, and 1 was found to preferentially sit on top of the orthosteric site. Together, this study demonstrates that computational tools can be applied to provide insight in understanding the structural basis of ligand-receptor interactions and the dynamics patterns of different mAChR subtypes, and can aid the discovery of potential M<sub>1</sub> mAChR selective hits.

## ABSTRAK

Muskarinik asetilkolin reseptor (mAChRs), khususnya  $M_1$  subjenis telah menjadi fokus penting dalam penemuan ubat disebabkan oleh peranannya dalam patofisiologi berkaitan dengan gangguan sistem saraf pusat, seperti, penyakit Alzheimer. Oleh kerana kesemua mAChRs subjenis mempunyai poket orthosterik yang serupa, pengaktif yang khusus untuk subjenis tertentu tidak dapat dicapai, dan sering membawa kepada kesan sampingan yang tidak diingini akibat daripada pengaktifan luar sasaran. Dalam tesis ini, kajian struktur dan dinamik  $M_1$  mAChR menggunakan pendekatan komputasi dibentangkan. Model homologi  $M_1$  mAChR telah dibina dan eksperimen pemeriksaan maya menunjukkan bahawa model tersebut mampu membezakan agonis daripada umpan, dengan model dimana TM5 diubahsuai juga memberikan prestasi yang baik dalam membezakan agonis daripada antagonis. Dinamik simulasi molekul telah membolehkan pencirian profil dinamik subjenis mAChR yang terikat dengan agonis, antagonis dan dalam bentuk kosong, yang membawa kepada penerangan perhubungan antara ligan and reseptor dari segi penstrukturan, pemilihan antara subjenis, dan pembentukan poket allosterik. Akhir sekali, bersama-sama model dengan struktur kristal  $M_2$  dan  $M_3$  reseptor telah digunakan dalam eksperimen pemeriksaan maya untuk mengenal pasti ligan khusus untuk  $M_1$ . Daripada 19 ligan yang dikenal pasti, 11 ligan menyasar poket orthosterik, 7 menggambarkan ciri bitopik, dan 1 memilih rongga diatas poket orthosterik. Kajian ini menunjukkan bahawa cara komputasi boleh digunakan untuk memberi gambaran dalam memahami struktur asas interaksi ligan-reseptor dan corak dinamik subjenis mAChR yang berbeza, yang dapat membantu dalam penemuan ligan khusus untuk  $M_1$ .

## ACKNOWLEDGEMENT

I am deeply thankful to my supervisors, Dr. Michael Buckle and Prof. Stephen Doughty, for their patience, motivation, and efforts in critically correcting my writings. Thank you for granting me great freedom in my academic pursuits.

I am truly honored to have the opportunity to work in Dr. Elizabeth Yuriev and Dr. David Chalmers lab in Monash University, Melbourne. I appreciate their knowledgeability, enthusiasm, and hospitality throughout the short stay. Not to forget Dr. Fiona McRobb for her helpful comments and Samuel Robinson, my lunch-mate, for his friendship.

I have to thank my lab-mates (Choon Han, Yap, Asfa, Abbeer, Siew Liang) for their countless support throughout the difficult time. I am truly fortunate to become a part of this group and will never forget all the moments that we shared together. To Lei Cheng and Kah Peng, my friends of tears and laughter, no words can ever describe my appreciation towards your companion and trust. Thank you for always being there.

Thanks go also to my family (mama, gege, dajie, ling, and ting), for their understanding, sacrifices, and endless love. I love you all! (Papa, I wish you well and have a good life now. I miss you!) I owe a very special thanks to Shang Hong, who never giving up on me for all the troubles that I have created, never doubt the decision that I have made, and quietly walk with me through the ups and downs in life. Thank you!

Lastly, thanks to Ministry of Science, Technology and Innovation (MOSTI) for NSF scholarship, and University Malaya for postgraduate research grant (PS191-2009C). This is indeed one of the most exciting (and emotional) pages in this thesis. As for those whom my forgetful mind has absolutely forgot to name here, please accept my most humble apologies, and thank you!

## TABLE OF CONTENTS

	<b>PAGE</b>
<b>TITLE PAGE</b>	i
<b>ORIGINAL LITERACY WORK DECLARATION FORM</b>	ii
<b>ABSTRACT</b>	iii
<b>ABSTRAK</b>	iv
<b>ACKNOWLEDGEMENT</b>	v
<b>TABLE OF CONTENTS</b>	vi
<b>LIST OF FIGURES</b>	ix
<b>LIST OF TABLES</b>	xv
<b>LIST OF ABBREVIATIONS</b>	xvii
<b>LIST OF APPENDICES</b>	xix
<b>1. GENERAL INTRODUCTION</b>	
<b>1.1 Study scope and objectives</b>	2
<b>2. LITERATURE REVIEW &amp; BACKGROUND</b>	
<b>2.1 G protein-coupled receptors</b>	4
<b>2.2 Structural biology of GPCRs</b>	6
2.2.1 Architecture of GPCRs	7
2.2.2 Sequence conservation and structural microswitches	8
2.2.3 Activation of GPCRs	11
<b>2.3 Ligand classification</b>	13

<b>2.4</b>	<b>Muscarinic acetylcholine receptors</b>	15
2.4.1	Alzheimer's disease and M <sub>1</sub> mAChR	18
2.4.2	Allosteric ligands of M <sub>1</sub> mAChRs	20
<b>2.5.</b>	<b>Molecular modeling methods</b>	24
2.5.1	Computer-derived GPCR models	24
2.5.2	Modeling protein-ligand complexes	24
2.5.3	Molecular dynamics simulations	27
2.5.4	Molecular modeling of muscarinic receptors	30
<b>3.</b>	<b>TOWARDS AGONIST BOUND HOMOLOGY MODELS OF THE HUMAN M<sub>1</sub> mAChR</b>	
<b>3.1</b>	<b><math>\beta_2</math>-adrenergic based modeling</b>	34
<b>3.1.1</b>	<b>Introduction</b>	34
<b>3.1.2</b>	<b>Methods</b>	35
3.1.2.1	Model constructions	35
3.1.2.2	Model refinement and optimization	36
3.1.2.3	Model validation and evaluation	38
3.1.2.4	Induced-fit docking	39
3.1.2.5	Enrichment studies	40
<b>3.1.3</b>	<b>Results and discussion</b>	44
3.1.3.1	Sequence alignment and model construction	39
3.1.3.2	Molecular dynamics simulations	45
3.1.3.3	Stereochemical quality of the models	49
3.1.3.4	Model evaluation using docking	51
3.1.3.5	Enrichment studies	52
<b>3.1.4</b>	<b>Conclusion</b>	56
<b>3.2</b>	<b>M<sub>3</sub> mAChR based modeling</b>	
<b>3.2.1</b>	<b>Introduction</b>	57
<b>3.2.2</b>	<b>Methods</b>	58
3.2.2.1	Model construction	59
3.2.2.2	Binding site refinement	60
3.2.2.3	Generation of agonist-bound models	61
3.2.2.4	Enrichment and docking studies	62
<b>3.2.3</b>	<b>Results and discussion</b>	68
3.2.3.1	Model selection and validation	68
3.2.3.2	Enrichment and docking studies	71
3.2.3.3	Analysis of ligand binding modes	74
3.2.3.4	Challenges and limitations	83
<b>3.2.4</b>	<b>Conclusion</b>	87

<b>4.</b>	<b>EXPLORING THE DYNAMICS PROFILE OF mAChRs: A MEMBRANE BASED MOLECULAR DYNAMICS STUDY</b>	
<b>4.1</b>	<b>Introduction</b>	89
<b>4.2</b>	<b>Methods</b>	91
4.2.1	Receptor structures	91
4.2.2	Molecular docking	91
4.2.3	System setup	91
4.2.4	Molecular dynamics simulations	93
4.2.5	Analysis	94
<b>4.3</b>	<b>Results and discussion</b>	95
4.3.1	System stability and flexibility	95
4.3.2	Collective motions of the receptor structures	104
4.3.3	Hydrogen bond networks	118
4.3.4	Molecular Mechanics-Poisson Boltzmann Surface Area (MM-PBSA)	128
4.3.5	Pocket analysis	135
<b>4.4</b>	<b>Conclusion</b>	145
<b>5.</b>	<b>IDENTIFICATION OF POTENTIAL M<sub>1</sub> mAChR SELECTIVE LEADS: VIRTUAL SCREENING USING MULTIPLE RECEPTOR MODELS AND STRUCTURES.</b>	
<b>5.1</b>	<b>Introduction</b>	148
<b>5.2</b>	<b>Methods</b>	150
5.2.1	Estimation of binding free energy	153
<b>5.3</b>	<b>Results and discussion</b>	153
5.3.1	Model generation and selection	153
5.3.2	Binding modes of known M <sub>1</sub> selective ligands	156
5.3.3	Identification of virtual screening hits	160
5.3.4	Binding modes of selected virtual screening hits	164
5.3.5	Prime MM-GBSA	166
<b>5.4</b>	<b>Conclusion</b>	167
<b>6.</b>	<b>CONCLUSION &amp; FUTURE WORK</b>	169
<b>7.</b>	<b>REFERENCES</b>	172
<b>8.</b>	<b>APPENDICES</b>	197

## LIST OF FIGURES

	<b>PAGE</b>
1.1	Project workflow, current (grey box) and future (pink box) studies. 3
2.1	Classical signaling pathway of GPCRs. Ligand binding to a GPCR's extracellular region triggers changes to the receptor conformation, which causes the release of GDP and the uptake of GTP by the G protein, stimulating activation of associated signaling pathways. Termination of the signal is mediated via the hydrolysis of the GTP subunit, phosphorylation of the GRKs, and arrestin binding. 6
2.2	Conserved structural motifs in GPCRs. Motifs are colored in yellow and most conserved residue in each TM is colored in red. 11
2.3	Common activation mechanisms of GPCRs - slight rotation and upward movement of TM3, movement of TM5, rotation of TM6, and inward movement of TM7 and TM1. 12
2.4	Pharmacological effects of different ligands, adapted from Tate (2012). 13
2.5	Different classes of ligands with distinct mode of actions and key properties, adapted from Kruse et al. (2014). 14
2.6	Sequence alignment of five human muscarinic subtypes. Identical amino acids are captured in red boxes, while similar amino acids are in red alphabet letters. Regions corresponding to the seven transmembrane helices are shown with the symbol of a spring on top. GPCR structural fingerprints are highlighted with blue stars at the bottom of the sequence and green triangles correspond to the cysteine residues involved in the formation of disulfide bridges. 17
2.7	Structures of M <sub>1</sub> mAChR preferring allosteric ligands. 23
2.8	Schematic representation of force field interactions. Covalent bonds are indicated by solid lines and non-bonded interactions by a dashed line. 29
3.1	M <sub>1</sub> mAChR agonists used in the enrichment studies. 42
3.2	Sequence alignment between the $\beta_2$ AR receptor (top) and the M <sub>1</sub> mAChR (bottom). Identical amino acids are captured in red boxes, while similar amino acids are in red alphabet letters. Regions corresponding to the seven transmembrane helices are shown with the symbol of a spring on top. GPCR structural fingerprints are highlighted with blue stars at the bottom of the sequence. 44
3.3	The area per lipid of the <b>A.</b> Amber force field and <b>B.</b> CHARMM force field based simulations of the M <sub>1</sub> mAChR homology model as a function 46

	of time.	
3.4	RMSD of the backbone atoms of the M <sub>1</sub> mAChR homology model from the minimized starting structure as a function of time for <b>A.</b> Amber force field based simulations, <b>B.</b> CHARMM force field based simulations.	48
3.5	Secondary structure evolution of the M <sub>1</sub> mAChR homology model from the last 10 ns trajectory of <b>A.</b> Amber force field based simulations, <b>B.</b> CHARMM force field based simulations. Color code: purple, $\alpha$ -helix; blue, $3_{10}$ helix; red, $\pi$ -helix; cyan, turn; and white, coil. Full stretch of color bar, especially the purple indicated that the secondary structure of the helix was well maintained.	49
3.6	The Ramachandran plot of the <b>A.</b> lowest energy conformation of M <sub>1</sub> mAChR homology model from Amber force field based simulations. 99.7% of the residues are found within the favored and allowed regions with an outlier of V307, and <b>B.</b> lowest energy conformation of M <sub>1</sub> mAChR homology model from CHARMM force field based simulations. 100% of the residues are found within the favored and allowed regions. Disallowed regions are colored in light yellow.	50
3.7	Binding pose of ACh. <b>A.</b> A three-dimensional representation of a fully buried ACh molecule in the binding cavity of the lowest energy conformation of the M <sub>1</sub> mAChR homology model from the Amber force field based simulations. <b>B.</b> A two-dimensional interactions map of ACh and the binding site residues. Hydrogen bond: blue dashed line with arrow head directed towards the electron donor; Pi interactions: orange line with symbols indicating the interaction; ionic interaction: pink dashed line with arrow heads on both sides. Residues involved in hydrogen bond, polar/ionic interaction, are represented with magenta circles, while residues involved in van der Waals interactions are represented by green circles.	52
3.8	Receiver operating characteristic (ROC) enrichment plots for the different homology models	54
3.9	Orientations of the binding site residues for different models. <b>A.</b> Superposition of active site of the crude (green) and chamber (orange) models. <b>B.</b> Superposition of active site of the chamber (orange) and chamber+IFD (light orange) models.	56
3.10	Sequence alignment between the M <sub>3</sub> mAChR (top) and the M <sub>1</sub> mAChR (bottom). Identical amino acids are captured in red boxes, while similar amino acids are in red alphabet letters. Regions corresponding to the seven transmembrane helices are shown with the symbol of a spring on top. GPCR structural fingerprints are highlighted with blue stars at the bottom of the sequence and green triangles correspond to the cysteine residues involved in the formation of disulfide bridges.	60
3.11	Superposition of models 1 (yellow), 1A1 (blue) and 1A2 (purple) with the important interacting side chain orientations for each of the models shown in stick representation. Loops are not shown for the purpose of	62

clarity.

- 3.12 M<sub>1</sub> mAChR antagonists used in the docking studies. 66
- 3.13 Superposition of solved 3-D GPCR structures with model 1 and model 1A2. Color code: yellow, model 1; purple, model 1A2; pink, M<sub>2</sub> (PDB code: 3UON); orange, M<sub>3</sub> (PDB code: 4DAJ); light grey,  $\beta_2$ AR (PDB code: 3SN6); light green,  $\beta_2$ AR (PDB code: 2RH1); brown, Histamine H1 (PDB code: 3RZE); dark blue, A2A adenosine (PDB code: 3QAK). 70
- 3.14 Receiver operating characteristic (ROC) agonist enrichment plots using the property-matched decoy set (Set III) for the different homology models. 72
- 3.15 Receiver operating characteristic (ROC) plots showing the agonist/antagonist selectivity for the different homology models: yellow, model 1; blue, model 1A1; purple, model 1A2. 73
- 3.16 3-D representations of the superimposed antagonists docked at a probable allosteric site in models 1A1 and 1A2. **A:** side view of model 1A1, with TM1 on the right and the orthosteric site shown by atropine (as spheres). **B:** side view of model 1A2, with TM1 on the right and the orthosteric site shown by doxylamine (as spheres). **C, D:** top views from the extracellular surface of models 1A1 and 1A2, respectively, with the side chains of the non-conserved interacting residues Q177 (E175 in M<sub>2</sub>), L183 (F181 in M<sub>2</sub>, L225 in M<sub>3</sub>), E397 (N419 in M<sub>2</sub>, K522 in M<sub>3</sub>), W400 (W422 in M<sub>2</sub>) and E401 (T423 in M<sub>2</sub>) shown in stick representation (yellow). 75
- 3.17 The binding modes of representative agonists with model 1A2. **A.** ACh, **B.** carbachol, **C.** oxotremorine-M, and **D.** pilocarpine. The important interacting residues are shown in stick representation and labeled. For the purpose of clarity, ECLs, ICLs, TM1, TM6, and TM7 are not shown. Residues involved in hydrogen bonding, charged, or polar interactions are shaded in red. Residues involved in van der Waals interactions are shaded in green. Residues involved in pi interactions are shown with a blue ring. 77
- 3.18 The binding modes of representative antagonists with model 1. **A.** NMS, **B.** QNB, and **C.** propantheline. The important interacting residues are shown in stick representation and labeled. For the purpose of clarity, ECLs, ICLs, TM1, TM6, and TM7 are not shown. Residues involved in hydrogen bonding, charged, or polar interactions are shaded in red. Residues involved in van der Waals interactions are shaded in green. Residues involved in pi interactions are shown with a blue ring. 78
- 3.19 3-D representations of the superimposed ligands that successfully docked at the orthosteric site of models 1A2 and 1. **A:** side view of the agonists docked to model 1A2, with TM1 on the right. **B:** side view of the antagonists docked to model 1, with TM1 on the right. **C, D:** top views from the extracellular surface of models 1A2 and 1. The side chains of the important interacting residues are shown in stick representation and colored according to the type of interactions 81

(Hydrophobic, hydrogen bond, charged/polar, aromatic).

- 4.1 Overall view of the system that underwent MD simulations. **A.** Front cross section view. **B.** Top view, water molecules are not shown for the purpose of clarity. Receptor (white, cartoon representation, only apo form is shown as example) was inserted into the hydrated POPC lipid bilayer. Water and lipid molecules were shown in blue and yellow VDW representation, respectively. 93
- 4.2 Secondary structure evolution for systems that underwent MD simulations. **A.** M<sub>1</sub>-Apo; **B.** M<sub>1</sub>-ACh; **C.** M<sub>1</sub>-QNB; **D.** M<sub>1</sub>\*-Apo; and **E.** M<sub>1</sub>\*-ACh. Color code: purple,  $\alpha$ -helix; blue,  $3_{10}$  helix; red,  $\pi$ -helix; cyan, turn; and white, coil. Full stretch of color bar, especially the purple indicated that the secondary structure of the helix was well maintained. 98
- 4.3 Secondary structure evolution for systems that underwent MD simulations. **A.** M<sub>2</sub>-Apo; **B.** M<sub>2</sub>-ACh; **C.** M<sub>2</sub>-QNB; **D.** M<sub>3</sub>-Apo; **E.** M<sub>3</sub>-ACh; and **F.** M<sub>3</sub>-QNB. Color code: purple,  $\alpha$ -helix; blue,  $3_{10}$  helix; red,  $\pi$ -helix; cyan, turn; and white, coil. Full stretch of color bar, especially the purple indicated that the secondary structure of the helix was well maintained. 99
- 4.4 Radius of gyration calculated from MD trajectories. **A.** M<sub>1</sub>, **B.** M<sub>1</sub>\*, **C.** M<sub>2</sub>, and **D.** M<sub>3</sub>. The vertical lines separate between heating and equilibration phases with production phase. 100
- 4.5 Time series of RMSD from the minimized starting structure calculated using backbone atoms of TM domain, ECLs and ICLs. The vertical lines separate between heating and equilibration phases with production phase. **A.** M<sub>1</sub>-Apo; **B.** M<sub>1</sub>-ACh; **C.** M<sub>1</sub>-QNB; **D.** M<sub>1</sub>\*-Apo; and **E.** M<sub>1</sub>\*-ACh. 101
- 4.6 Time series of RMSD from the minimized starting structure calculated using backbone atoms of TM domain, ECLs and ICLs. The vertical lines separate between heating and equilibration phases with production phase. **A.** M<sub>2</sub>-Apo; **B.** M<sub>2</sub>-ACh; **C.** M<sub>2</sub>-QNB; **D.** M<sub>3</sub>-Apo; **E.** M<sub>3</sub>-ACh; and **F.** M<sub>3</sub>-QNB. 102
- 4.7 RMSF calculated for every residue fitted to the minimized starting structure in all the systems simulated (**A.** M<sub>1</sub>, **B.** M<sub>1</sub>\*, **C.** M<sub>2</sub>, and **D.** M<sub>3</sub>). ICLs and ECLs were indicated by shaded background. 103
- 4.8 Projection of trajectory snapshots onto the subspace spanned by principal modes PC1 and PC2 of backbone heavy atoms as determined from the MD simulations. **A.** M<sub>1</sub>, **B.** M<sub>1</sub>\*, **C.** M<sub>2</sub>, and **D.** M<sub>3</sub>. Red: receptor with ACh bound, green: QNB bound, and blue: apo. The arrow marked the starting point for the apo simulations. 106
- 4.9 Overlap between 8 PCA modes from different systems (Left: M<sub>1</sub>, right: M<sub>1</sub>\*). Similarity between the vector pair is sorted by color, where blue and red indicated least and highest similarity, respectively. RMSIP take values from 0 – 1, with value closer to 1 indicating greater similarity. 107

4.10	Overlap between 8 PCA modes from different systems (Left: M <sub>2</sub> , right: M <sub>3</sub> ). Similarity between the vector pair is sorted by color, where blue and red indicated least and highest similarity, respectively. RMSIP take values from 0 – 1, with value closer to 1 indicating greater similarity.	108
4.11	Mobility plots of first mode of principal component analysis. <b>A.</b> M <sub>1</sub> , <b>B.</b> M <sub>1</sub> <sup>*</sup> , <b>C.</b> M <sub>2</sub> , and <b>D.</b> M <sub>3</sub> . For the purpose of clarity, square fluctuations exceed 6 Å <sup>2</sup> (correspond to loop region where ICL3 was truncated) are not plotted.	110
4.12	Structural variations along the top two PCA modes. <b>A.</b> M <sub>1</sub> -Apo, <b>B.</b> M <sub>1</sub> -ACh, <b>C.</b> M <sub>1</sub> -QNB. The receptors are shown in tube, the arrows indicate directions and amplitudes. For the purpose of clarity, arrows are shown every 4 <sup>th</sup> residues and only if their lengths are more than 1 Å. Magenta: PC1, cyan: PC2.	114
4.13	Structural variations along the top two PCA modes. <b>A.</b> M <sub>1</sub> <sup>*</sup> -Apo, and <b>B.</b> M <sub>1</sub> <sup>*</sup> -ACh. The receptors are shown in tube, the arrows indicate directions and amplitudes. For the purpose of clarity, arrows are shown every 4 <sup>th</sup> residues and only if their lengths are more than 1 Å. Magenta: PC1, cyan: PC2.	115
4.14	Structural variations along the top two PCA modes. <b>A.</b> M <sub>2</sub> -Apo, <b>B.</b> M <sub>2</sub> -ACh, and <b>C.</b> M <sub>2</sub> -QNB. The receptors are shown in tube, the arrows indicate directions and amplitudes. For the purpose of clarity, arrows are shown every 4 <sup>th</sup> residues and only if their lengths are more than 1 Å. Magenta: PC1, cyan: PC2.	116
4.15	Structural variations along the top two PCA modes. <b>A.</b> M <sub>3</sub> -Apo, <b>B.</b> M <sub>3</sub> -ACh, and <b>C.</b> M <sub>3</sub> -QNB. The receptors are shown in tube, the arrows indicate directions and amplitudes. For the purpose of clarity, arrows are shown every 4 <sup>th</sup> residues and only if their lengths are more than 1 Å. Magenta: PC1, cyan: PC2.	117
4.16	M <sub>1</sub> hydrogen bonding networks. Solid line: inter-helical hydrogen bonds; dotted line: inter-helical hydrogen bonds mediated by water clusters. Internal water molecules were captured in 3-D diagrams for each of the systems, as red oxygen spheres.	124
4.17	M <sub>1</sub> <sup>*</sup> hydrogen bonding networks. Solid line: inter-helical hydrogen bonds; dotted line: inter-helical hydrogen bonds mediated by water clusters. Internal water molecules were captured in 3-D diagrams for each of the systems, as red oxygen spheres.	125
4.18	M <sub>2</sub> hydrogen bonding networks. Solid line: inter-helical hydrogen bonds; dotted line: inter-helical hydrogen bonds mediated by water clusters. Internal water molecules were captured in 3-D diagrams for each of the systems, as red oxygen spheres.	126
4.19	M <sub>3</sub> hydrogen bonding networks. Solid line: inter-helical hydrogen bonds; dotted line: inter-helical hydrogen bonds mediated by water clusters. Internal water molecules were captured in 3-D diagrams for each of the systems, as red oxygen spheres.	127

4.20	Decomposition of binding energy on a per residues basis into contributions from EEL+GB, NP, and vdW, upon the binding of <b>A.</b> ACh and <b>B.</b> QNB. First bar of each of the residues corresponding to M <sub>1</sub> , follow by M <sub>1</sub> *, M <sub>2</sub> , and M <sub>3</sub> of the subsequent bars in A and B, except that in B, which M <sub>1</sub> * is excluded.	132
4.21	ACh and QNB interactions with the receptors. <b>A.</b> M <sub>1</sub> *-ACh, <b>B.</b> M <sub>1</sub> -QNB, <b>C.</b> M <sub>2</sub> -ACh, <b>D.</b> M <sub>2</sub> -QNB, <b>E.</b> M <sub>3</sub> -ACh, and <b>F.</b> M <sub>3</sub> -QNB.	133
4.22	Transition of ACh from originally bound spot 1 (yellow) at the orthosteric site, to 2 (blue) and 3 (green) in M <sub>1</sub> -ACh simulations and establishes interactions with different set of residues.	134
4.23	Pockets found throughout the simulations for <b>A.</b> M <sub>1</sub> , <b>B.</b> M <sub>1</sub> *, <b>C.</b> M <sub>2</sub> , and <b>D.</b> M <sub>3</sub> . It is clear that there is a pocket extension from the primary orthosteric pocket in M <sub>1</sub> toward D2.50, where different patterns were found on other receptor subtypes (pointed by red arrow). Secondary pockets were found on top of the orthosteric site, involving ECLs.	137
4.24	Top view of the pockets found throughout the simulations for <b>A.</b> M <sub>1</sub> , <b>B.</b> M <sub>1</sub> *, <b>C.</b> M <sub>2</sub> , and <b>D.</b> M <sub>3</sub> . Pocket extension from the secondary pocket was shown (pointed red arrow).	138
4.25	<b>A.</b> Graphical and sequence representations of the non-conserved/identical residues at the extracellular vestibule of the receptors. Green: non-identical residues in the TM helices, red: non-identical residues in the loops, blue: highly conserved residues, and *: residues with side-chains facing the inner TM core. <b>B-D:</b> Electrostatic charge distribution calculated by APBS for M <sub>1</sub> , M <sub>2</sub> , and M <sub>3</sub> , respectively.	139
4.26	Dial plots of the average helix rotation angles in M <sub>1</sub> system during the MD simulations. Average helix Z-tilt angle (°) with respect to the apo receptors for each of the TM helices is shown underneath the dial plots.	141
4.27	Dial plots of the average helix rotation angles in M <sub>1</sub> * system during the MD simulations. Average helix Z-tilt angle (°) with respect to the apo receptors for each of the TM helices is shown underneath the dial plots.	142
4.28	Dial plots of the average helix rotation angles in M <sub>2</sub> system during the MD simulations. Average helix Z-tilt angle (°) with respect to the apo receptors for each of the TM helices is shown underneath the dial plots.	143
4.29	Dial plots of the average helix rotation angles in M <sub>3</sub> system during the MD simulations. Average helix Z-tilt angle (°) with respect to the apo receptors for each of the TM helices is shown underneath the dial plots.	144
5.1	Binding modes and 2-D interactions map of the known M <sub>1</sub> selective ligands. Red circle: -ve charged, light blue circle: polar, green circle: hydrophobic, green line: pi-pi stacking, red line: pi-cation, purple arrow: hydrogen bond (backbone), purple dash arrow: hydrogen bond (side-chain), red-blue line: salt bridge.	158

## LIST OF TABLES

	<b>PAGE</b>	
2.1	Modeling studies of muscarinic receptors.	31
3.1	Distribution of molecular properties and Tanimoto similarity scores for agonists and decoys.	41
3.2	Enrichment factors for the different homology models at x% of screened library.	54
3.3	Distribution of molecular properties and Tanimoto similarity scores for agonists and decoys.	65
3.4	RMSD values between the selected models and solved crystal structures.	68
3.5	PROCHECK and Ramachandran plot summary results for the template structure (4DAJ) and the selected models.	70
3.6	WHATCHECK Z-scores for quality assessment and statistical analysis for the template structure (4DAJ) and the selected models.	71
3.7	Analysis of data from enrichment studies for the crude and refined models.	72
3.8	Structural interaction fingerprints bit-string of bound agonists in all models.	79
3.9	Structural interaction fingerprints bit-string of bound agonists in comparison with antagonists.	79
3.10	An averaged structural interactions fingerprint calculated over all successfully docked poses of agonists and antagonists with the interacting residues in the TM domain.	82
4.1	Description of the systems that underwent MD studies.	92
4.2	Average area per lipid, radius of gyration and RMSD calculated over the trajectories for the different systems studied.	96
4.3	PCA fraction of variance for the different systems studied.	98
4.4	Hydrogen bond fractions of M <sub>1</sub> and M <sub>1</sub> * simulations.	119
4.5	Hydrogen bond fractions of M <sub>2</sub> and M <sub>3</sub> simulations.	120
4.6	Residues involved in inter-helical water-mediated hydrogen bonds.	122

4.7	Residues involved in inter-helical water-mediated hydrogen bonds.	123
4.8	Relative binding free energies of complexes estimated using MMPBSA for the different complexes studied.	129
5.1	Distribution of molecular properties for ligand library used in the virtual screening.	152
5.2	Enrichment results for M <sub>1</sub> homology models.	154
5.3	Enrichment results for M <sub>2</sub> crude and IFD structures.	154
5.4	Enrichment results for M <sub>3</sub> models and IFD structures.	155
5.5	GScore values of the known selective M <sub>1</sub> ligands showing their preferential docking to the M <sub>1</sub> model compared to the M <sub>2</sub> and M <sub>3</sub> models and Prime MM-GBSA estimations of the free energy changes for binding to the preferred M <sub>1</sub> model.	157
5.6	GScore values of the virtual screening hits showing their preferential docking to the M <sub>1</sub> models compared to the M <sub>2</sub> and M <sub>3</sub> structures and Prime MM-GBSA estimations of the free energy changes for binding to the preferred M <sub>1</sub> model.	162
5.7	Molecular properties of known M <sub>1</sub> selective ligands and virtual screening hits.	163

## LIST OF ABBREVIATIONS

		<b>PAGE</b>
mAChR	Muscarinic acetylcholine receptor	1
GPCRs	G-protein coupled receptors	1
ACh	Acetylcholine	1
TM	Transmembrane	4
GRKs	GPCR-regulating kinases	4
$\beta_2$ AR	$\beta_2$ -adrenergic	6
ECLs	Extracellular loops	7
ICLs	Intracellular loops	7
PAMs	Positive allosteric modulators	13
NAMs	Negative allosteric modulators	14
CNS	Central nervous system	18
AD	Alzheimer's disease	18
APP	Amyloid precursor protein	19
Glide	Grid-based ligand docking with energetics	25
IFD	Induced-fit docking	26
RMSD	Root mean square deviation	27
MD	Molecular dynamics	27
3-D	Three-dimensional	33
PDB	Protein Data Bank	34
POPC	palmitoyl-oleoyl-phosphatidyl-choline	36
EF	Enrichment factors	40
SDM	Site directed mutagenesis	51
ROC	Receiver Operating Characteristic	53

BEDROC	Boltzmann-enhanced discrimination of receiver operating characteristic	58
AUC	Area under the curve	58
NMS	<i>N</i> -methylscopolamine	76
QNB	(-)-3-quinuclidinyl benzilate	76
OPM	Orientations of Proteins in Membranes	92
GB	Generalized Born	94
PCA	Principal component analysis	94
RMSF	Root mean square fluctuation	96
RMSIP	Root mean square inner product	104
MM-PBSA	Molecular Mechanics-Poisson Boltzmann Surface Area	121
GScore	Glide score	152

## LIST OF APPENDICES

	<b>PAGE</b>
A. Enrichment results using decoy set I for the M <sub>1</sub> models developed based on the inverse agonist bound M <sub>3</sub> crystal (PDB code: 4DAJ).	197
B. Enrichment results for M <sub>1</sub> models based on other class A GPCR, developed following the methods used for muscarinic based modeling.	198
C. Enrichment results for MD snapshots extracted from the apo and ACh bound simulations of M <sub>1</sub> models.	199
E. PCA fraction of variance for the different systems studied.	200
D. Publications/conference papers	200

# CHAPTER 1

## GENERAL INTRODUCTION

Muscarinic acetylcholine receptors (mAChRs) are G protein-coupled receptors (GPCRs), which transduce signals across the cell membrane into the intracellular side upon binding of the endogenous neurotransmitter, acetylcholine (ACh) (Taylor & Brown, 2006). Ligand binding to GPCRs stabilizes different conformational states of the receptors. The recent determination of the crystal structures of some GPCRs in both active and inactive states has opened up a new phase of investigations aimed at comprehending the structural basis of ligand receptor interactions and the activation mechanism (Cherezov et al., 2007; Rasmussen et al., 2011; Haga et al., 2012; Kruse et al., 2012; Kruse et al., 2013). This breakthrough has also benefited the drug discovery program as the structures can now be used for *in silico* screening and lead optimization.

To date, five different subtypes of mAChRs ( $M_1$  to  $M_5$ ) have been identified and are widely expressed in various regions within the central and peripheral nervous systems. The  $M_1$  mAChR is predominantly found in the hippocampus, cerebral cortex, corpus striatum, and thalamus, and is physiologically linked to multiple functions such as synaptic plasticity, neuronal excitability, learning, and memory. Since the  $M_1$  mAChR plays an important role in learning and memory, it has long been a therapeutic target for neurodegenerative disorders such as Alzheimer's disease (Broadley & Kelly, 2001). However, the development of effective  $M_1$  mAChR agonists has not yet been realized,

due to poor subtype selectivity. Most of the muscarinic agonists identified to date suffer from adverse central and peripheral side effects.

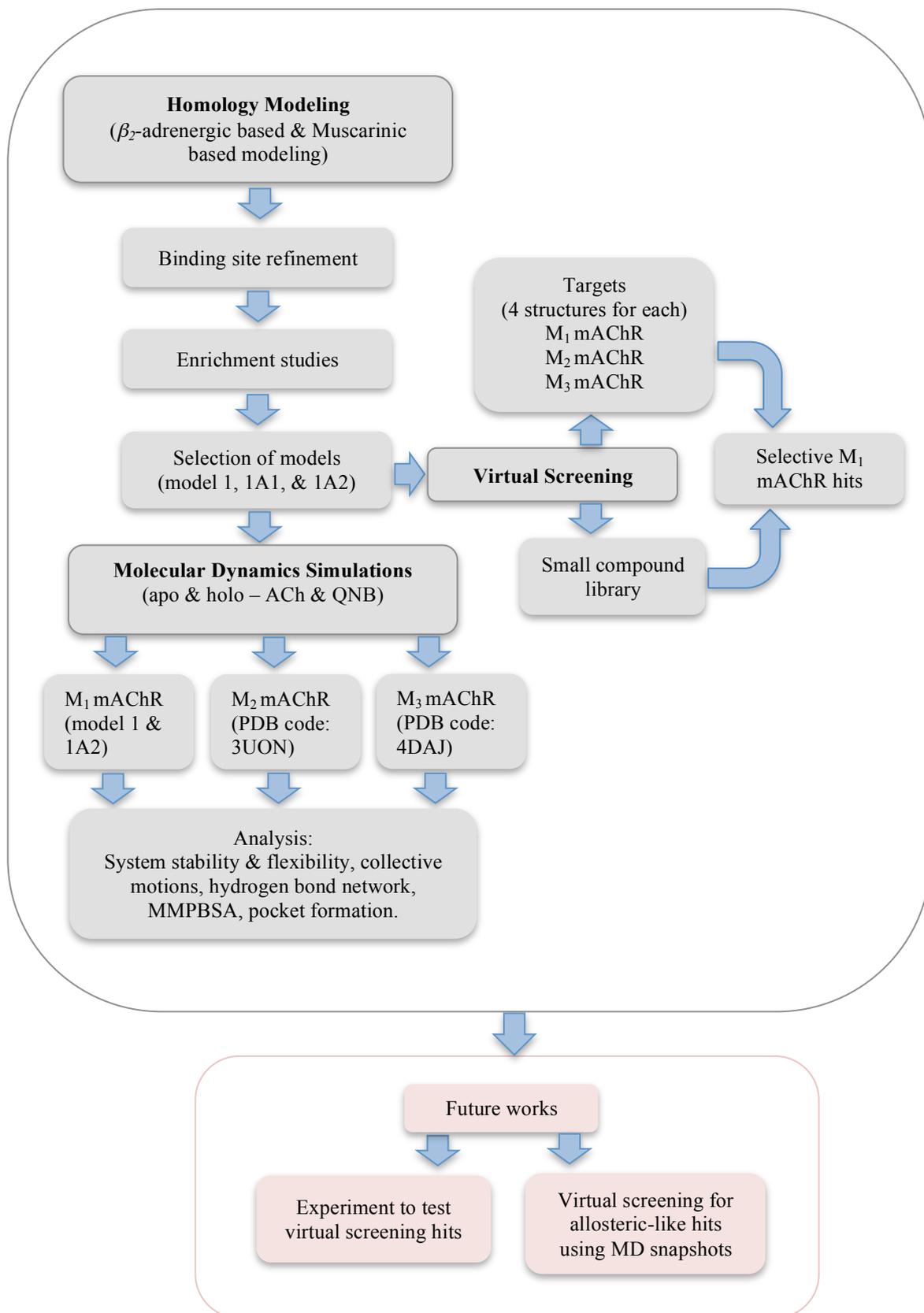
In the search for new lead compounds, structure-based computational approaches such as virtual screening have advantages over conventional drug discovery that depends on the synthesis and screening of vast numbers of compounds to optimize their activity profile, which is very time-consuming and expensive. In order to carry out virtual screening, a 3-D structure is needed. Since an M<sub>1</sub> mAChR crystal structure is not available yet, a homology model based on a closely related protein family has the potential to accelerate the drug discovery process for the M<sub>1</sub> mAChR and solve the subtype selectivity problem. A static model however gives little information about the dynamics of the structure, which is known to be important for its function, so structure and dynamics have been brought together in this work in order to give a better understanding of the receptor system.

## **1.1 Study scope and objectives**

The aim of this study is to understand and characterize the structural basis of ligand receptor interactions for the M<sub>1</sub> mAChR. To pursue the goal, three main objectives were established:

1. To construct M<sub>1</sub> mAChR models that are capable in agonist recognition.
2. To explore the dynamics profiles of different receptor subtypes.
3. To identify potential M<sub>1</sub> selective lead compounds through virtual screening.

To achieve these objectives, studies were planned and carried out according to the workflow shown in Figure 1.1 and detailed methodologies and findings for objective 1, 2, and 3 are described in chapters 3, 4, and 5, respectively.



**Figure 1.1.** Project workflow, current (grey box) and future (pink box) studies.

## CHAPTER 2

### LITERATURE REVIEW & BACKGROUND

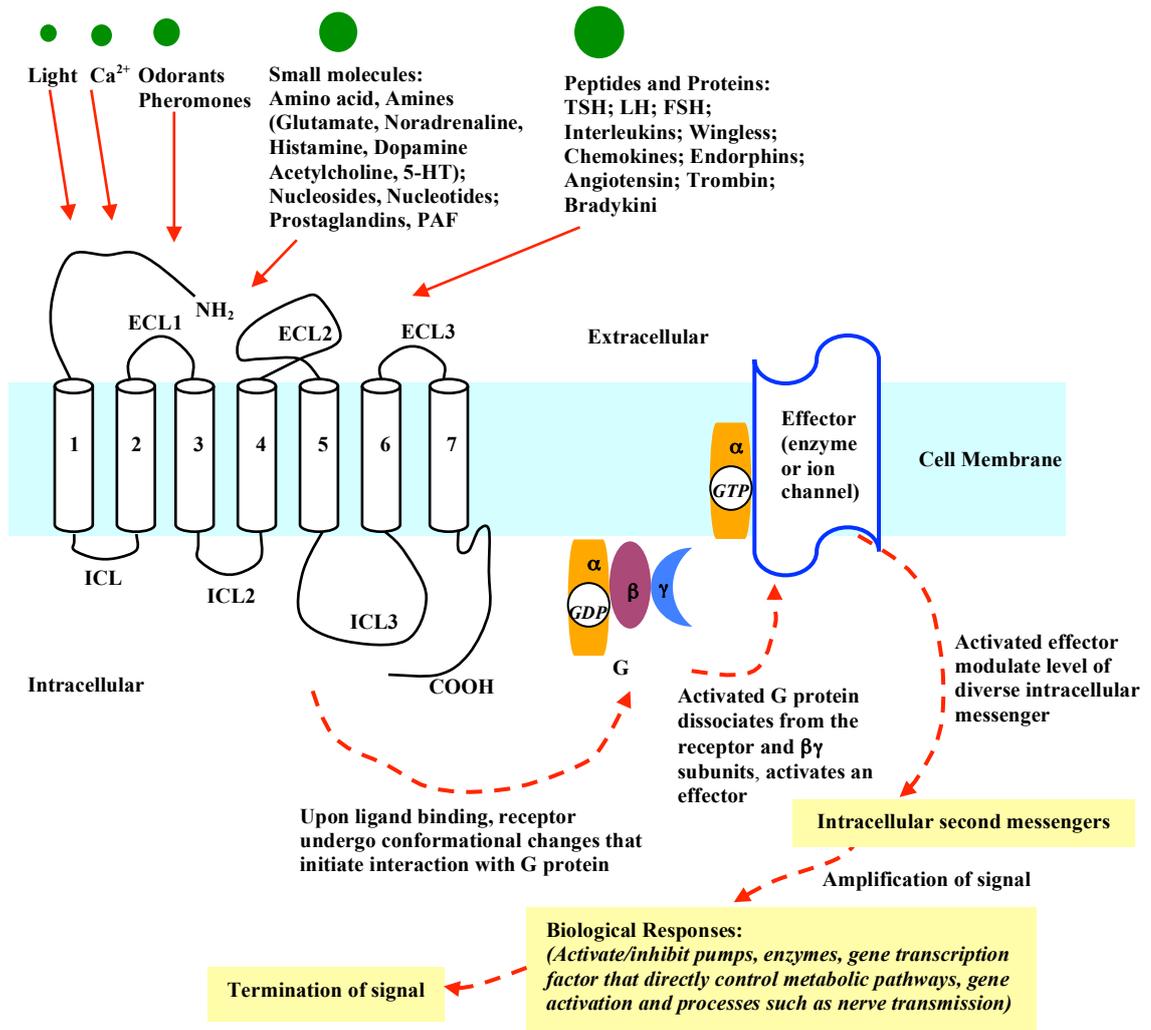
#### 2.1 G protein-coupled receptors

G protein-coupled receptors (GPCRs) represent the largest family of cell surface receptors. In light of their characteristic topology, which features one polypeptide chain that spans seven times through the membrane, they are also known as seven transmembrane (TM) receptors (Salon, Lodowski, & Palczewski, 2011). These receptors transduce extracellular signals into intracellular signaling cascades for vast array of natural ligands/transmitters, including ions, photon, biogenic amines, lipids, nucleotides, odorants, and peptides.

Ligand binding induces changes in the receptor conformation that are transmitted down to the cytoplasmic face of the protein, facilitating a coupling of the cytoplasmic face with an intracellular heterotrimeric G protein (GTP binding protein). The G protein, in turn, acts as an intracellular signal by activating or inhibiting intracellular effectors, which regulate the intracellular concentrations of secondary messengers, leading to various physiological responses (Strader, Fong, Tota, Underwood, & Dixon, 1994). The hydrolysis of the subunit-bound GTP to GDP, resulting in the re-association of the  $\alpha$  and  $\beta\gamma$  subunits, terminates the signal and completes the cycle. Alternatively, desensitization (functional uncoupling) is regulated by GPCR-regulating kinases (GRKs) and arrestins (Gurevich & Gurevich, 2006) (Figure 2.1).

Traditionally, most GPCR signaling pathways/activities were thought to have been mediated by the activation of G proteins. However, several studies pioneered by Robert Lefkowitz showed that the GPCR signaling pathway is also modulated by proteins known as arrestins (Lefkowitz & Shenoy, 2005). Arrestins form a small family of GPCR-binding proteins originally discovered for their role in receptor desensitization but now recognized as true adaptor proteins that transduce signals to multiple effector pathways.  $\beta$ -arrestin-dependent signal transduction has distinct biochemical and functional consequences from those mediated by G proteins and biased ligands that preferentially signal through either G protein- or  $\beta$ -arrestin-mediated pathways are of therapeutic importance. The concept of a biased ligand describes the ability of a ligand to selectively stabilize receptor conformations that stimulate or inhibit subsets of receptor activities/selectively activates only one signaling response (biased signaling) (Violin & Lefkowitz, 2007).

Phylogenetic classification of GPCRs categorizes them into five distinct subfamilies: Rhodopsin, Glutamate, Adhesion, Frizzled/Taste2 and Secretin (GRAFS nomenclature) (Schiöth & Fredriksson, 2005). Due to their roles in modulating tissue/cell physiology and homeostasis, signaling pathways associated with GPCRs are closely linked to diverse pathological processes in neural, cardiovascular, endocrine, immune systems, and cancer (Heng, Aibel, & Fussenegger, 2013), and make them attractive druggable targets for intense drug discovery efforts.



**Figure 2.1.** Classical signaling pathway of GPCRs. Ligand binding to a GPCR's extracellular region triggers changes to the receptor conformation, which causes the release of GDP and the uptake of GTP by the G protein, stimulating activation of associated signaling pathways. Termination of the signal is mediated via the hydrolysis of the GTP subunit, phosphorylation of the GRKs, and arrestin binding.

## 2.2 Structural biology of GPCRs

In the year 2000, the first high resolution X-ray structure of GPCRs - bovine rhodopsin, was made available (Palczewski et al., 2000). Rhodopsin remained as the only available X-ray crystal structure for a GPCR, until a structure of the human  $\beta_2$ -adrenergic ( $\beta_2$ AR) receptor was resolved in 2007 (Cherezov et al., 2007; Rasmussen et al., 2007). The

biggest achievement to date in this area came in 2011, when the first ternary complex of an agonist bound  $\beta_2$ AR with the G protein, capturing the active state of the receptor was released (Rasmussen et al., 2011). Throughout the last few years, many GPCR structures, especially class A (rhodopsin-like) GPCRs, have begun to emerge, including adenosine  $A_{2A}$ , chemokine CXCR4, dopamine D3, histamine H1, S1P<sub>1</sub>, muscarinic (M<sub>2</sub> and M<sub>3</sub>), opioid, neurotensin, and serotonin 5-HT<sub>1B</sub>, 5-HT<sub>2B</sub> structures. The discovery journey and timeline of the GPCR structures, and the limitations and challenges to obtain such structures have been reviewed and described in detail by Topiol and Sabio (2009); Flight (2013); Katritch, Cherezov, and Stevens (2013); Kobilka (2013). These GPCR crystal structures have provided a wealth of information to redefine our knowledge on how the receptors recognize a diverse range of ligands and transmit a signal across the cell membrane, and have made a very significant contribution towards understanding the structural basis of ligand-receptor interactions.

### 2.2.1 Architecture of GPCRs

The structure of GPCRs consists of three parts: the extracellular region – the N terminus and extracellular loops (ECL1 to ECL3); the TM core - seven alpha helices (TM1 to TM7); and the intracellular region - intracellular loops (ICL1 to ICL3), amphipathic helix (H8), and the C terminus (Figure 2.1). Generally, the extracellular region controls ligand access, the TM core binds ligands and transduces the signal to the intracellular region via conformational changes, and the intracellular region interacts with cytosolic signaling proteins.

The variations of the extracellular regions revealed by the GPCR X-ray structures imply a distinctive role of the ECLs in shaping the binding pocket and route for ligand entry. ECL2, which most often interacts with bound ligands, could possibly be involved in

early ligand recognition and subtype selectivity (Dror et al., 2011; Kruse et al., 2012). In contrast to ECL2, ECL1 and ECL3 are relatively shorter, and lack defined secondary structure elements. This region contributes to the receptor stability through disulfide bridges, especially the disulfide bridge between ECL2 and TM3 (extracellular side of the helix adjacent to the binding pocket), which limits the extent of the conformational changes in this portion during activation. The disulfide bridge between the ECLs also possibly dictates receptor function by restricting the conformational freedom of the loops (Venkatakrisnan et al., 2013).

Ligands of diverse shapes, sizes and chemical properties bind to the different members of the GPCR family at the same binding pocket within the TM core, with similar contacts, despite the diversity in the ligands. In particular, residues at positions 3.32, 3.33, 3.36, 6.48, 6.51 and 7.39 at TM3, 6 and 7 (see section 2.2.2 for the explanation of numbering convention) make consensus contacts with diverse ligands across class A GPCRs and the variation of the residues at these positions is responsible for the ligand specificity of the different receptors. While residues at other TM helices make contacts with specific ligands to different extents, water molecules are also observed to mediate indirect contacts between the ligands and receptors (Congreve, Langmead, Mason, & Marshall, 2011).

### 2.2.2 Sequence and structural conservation of GPCRs

GPCRs with diverse sequences share a similar overall architecture with the seven TM helices being held together by tertiary contacts. Several well conserved patterns have been observed in GPCR sequences (Mirzadegan, Benko, Filipek, & Palczewski, 2003), and are exemplified by the Ballesteros and Weinstein residue numbering scheme (Ballesteros & Weinstein, 1995). In this numbering method, the term X.YY is used,

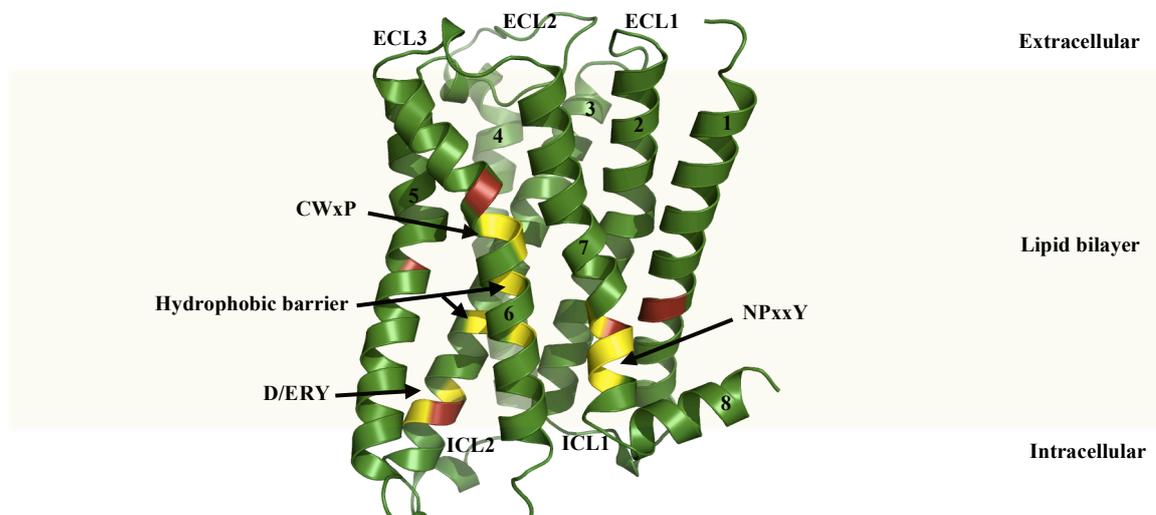
where X represents the number of the transmembrane domain and YY is the position of the residue within the transmembrane domain. GPCR amino acid sequences are aligned and the most conserved residue in each TM helix is assigned with the number of 50. Hence, each residue within the same helix is given a number relative to the most conserved position of 50. According to the Ballesteros and Weinstein nomenclature, N in TM 1, D in TM2, R in TM3, W in TM4, and P in TM5, TM6, and TM7 are numbered 50, as they are the most conserved residues in each of the TM helices. For ease of comparison, this scheme is also adopted in this thesis. The conserved sequence patterns in GPCRs are functionally important and are characterized as key structural motifs in GPCR structures or functional microdomains. The most important structural motifs (see Figure 2.2) are: the ionic lock (D/ERY motif), the transmission switch/rotamer toggle switch (CWxP motif), the hydrophobic cage (the conserved hydrophobic residues at position 3.43 and 6.44, with bulky hydrophobic residue at position 6.40 and 6.41), and the tyrosine toggle switch (NPxxY motif).

The presence of an ionic lock was first shown in the X-ray structure of bovine rhodopsin (Palczewski et al., 2000). The inactive state of bovine rhodopsin shows a strong intermolecular interaction between E<sup>3.49</sup>/R<sup>3.50</sup> of the conserved D/ERY motif in TM3 and E<sup>6.30</sup>/T<sup>6.34</sup> in TM6. It is proposed that the ionic lock is one of the critical constraints keeping the receptor in its inactive conformation and the disruption of the ionic lock forms part of the activation cascade to allow the interactions of the receptor with its cognate G protein upon activation. However, an ionic lock is present only in some of the crystal structures, such as those for dopamine D<sub>3</sub> and adenosine A<sub>2A</sub> (Chien et al., 2010; Dore et al., 2011).

TM5 and TM6 undergo large conformational changes upon agonist binding, facilitated

by relocation of the conserved W<sup>6.48</sup> and F<sup>6.44</sup> towards L<sup>5.51</sup> and P<sup>5.55</sup>, and of I/L<sup>3.40</sup> away from P<sup>5.55</sup>. Such movements are termed transmission switches and play a key role in linking the agonist binding site with TM5 and TM6 movement, through the rearrangement of the TM3-5-6 interface (Deupi & Standfuss, 2011). The conserved P<sup>6.50</sup> (part of the CWxP motif) contributes to the helix kink observed in TM6 and W<sup>6.48</sup> (floor of the orthosteric binding site), which is also part of the CWxP motif, plays a central role in this transmission switch through the side chain rotamer transition (hence the previous name of rotamer toggle switch). However, no major conformational change to W<sup>6.48</sup> has been seen in any of the active receptor structures, as proposed by biochemical experiments and computer simulations (Taddese, Simpson, Wall, Blaney, & Reynolds, 2013).

The hydrophobic barrier formed by L<sup>2.43</sup>, L<sup>2.46</sup>, L<sup>3.43</sup>, L<sup>3.46</sup>, M<sup>6.36</sup>, and M<sup>6.40</sup> hold TM3 and TM6 in place in the inactive state and are conserved throughout the class A GPCRs (Tehan, Bortolato, Blaney, Weir, & Mason, 2014). During activation, the rotation of TM6 disrupts the water-mediated hydrogen bond between W<sup>6.48</sup> and S<sup>7.45</sup> and reorganizes the ground state hydrogen bond network, which leads to the opening of the hydrophobic barrier. This further allows the rearrangement of Y<sup>7.53</sup> from the NPxxY motif and Y<sup>5.58</sup> to fill in the hydrophobic gap and extend the hydrogen bond network towards the D/ERY motif. It has been suggested that ligand binding induces local conformational changes that are transmitted to the cytoplasmic end of the receptor and changes in the water mediated hydrogen bond network around the NPxxY motif in TM7. In the inactive receptor structure, the side chain of Y<sup>7.53</sup> points toward TM1, TM2 or helix 8, in contrast to the active state receptor structure, where Y<sup>7.53</sup> side chain changes its rotamer conformation and points toward the middle core of the TM bundle, establishing stabilizing interactions with TM3 and TM6 (Trzaskowski et al., 2012).



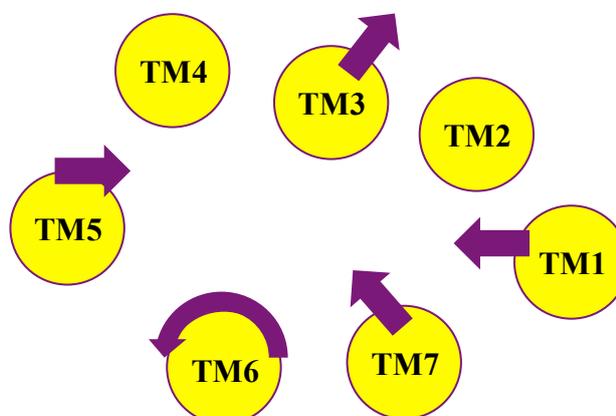
**Figure 2.2.** Conserved structural motifs in GPCRs. Motifs are colored in yellow and the most conserved residue in each TM is colored in red.

### 2.2.3 Activation of GPCRs

The currently available crystal structures of GPCRs are all solved in complex with agonist/antagonist/inverse agonist, and fall into three distinct conformations: an inactive state in complex with an antagonist or inverse agonist, an agonist-bound state without G protein or G protein surrogate, and a fully activated state – the ternary complex consisting of receptor, agonist and G protein/G protein surrogate. These conformations are thought to be linked by intermediate conformations. With the multiple active and inactive pairs of GPCR crystal structures, a common activation mechanism can be identified.

TM3 and TM6 are the functional hub of the common activation pathway as they not only form the ligand-binding pocket but also have direct contact with each of the TM

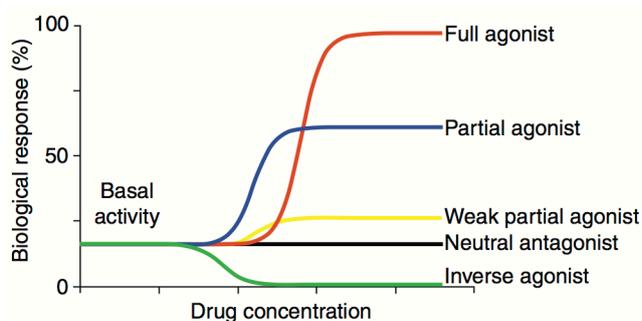
helices, except TM1. The common movements involved in the activation mechanisms are: slight rotation and upward movement of TM3, movement of TM5, rotation of TM6, and inward movement of TM7 and TM1 (Figure 2.3). The hydrophobic residues between TM3 and TM6, right in the TM core, facilitate the movement of TM3 and TM6, where they form a hydrophobic barrier in the inactive state structure to prevent the formation of the water channel seen in the active state structure. The breakage of the hydrophobic cage through the upward movement of TM3 and rotation of TM6 allows N<sup>7.49</sup> of the NPxxY motif to move up and in, making contacts with D<sup>2.50</sup> and the water channel within the TM core. Following the action of N<sup>7.49</sup>, Y<sup>7.53</sup> moves into the cytoplasmic cleft between TM3 and TM6, opened up by the movement of the residue at position 6.40 and makes water mediated hydrogen bonds with Y<sup>5.58</sup>. The breaking of the ionic lock further allows the movement of TM which opens up to a water channel that stabilizes the active conformation through extensive hydrogen bond networks from the extracellular side of the receptor all the way through to the intracellular side. All this movement ends up with the opening of the cytoplasmic side of the receptor, to allow the interaction of the G protein and transmit the signal cascade further. The mechanisms have been described in detail elsewhere (Tehan et al., 2014).



**Figure 2.3.** Common activation mechanism of GPCRs - slight rotation and upward movement of TM3, movement of TM5, rotation of TM6, and inward movement of TM7 and TM1.

### 2.3 Ligand classification

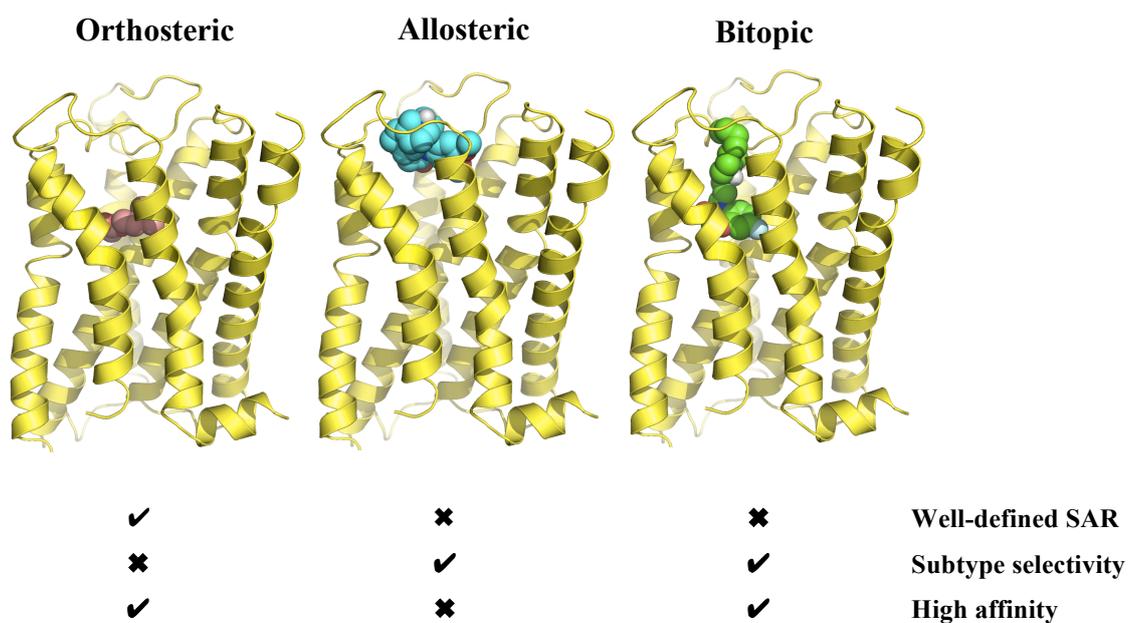
Ligand free GPCRs exhibit basal activity even in the absence of an activating ligand, due to the thermal excitation from the environment that provides enough energy to drive the sub-population towards the active conformation. The basal activity is enhanced when agonists bind to the receptor, reduced by inverse agonists, and unaffected by neutral antagonists. Hence ligands that interact with GPCRs are classified according to their activity exerted on the GPCR (Figure 2.4).



**Figure 2.4.** Pharmacological effects of different ligands, adapted from Tate (2012).

GPCR ligands can also be categorized according to the location of their binding site, that is, orthosteric, allosteric, or bitopic. Endogenous ligands and conventional agonists, antagonists, and inverse agonists typically occupy the orthosteric binding site. Because of the highly conserved orthosteric binding site across receptor subtypes, orthosteric ligands may lack subtype selectivity properties. Allosteric ligands are compounds that bind to a receptor, on a site other than the orthosteric site and alter the properties of a bound orthosteric ligand or mediate agonistic activity in the absence of the endogenous or orthosteric ligands (allosteric agonists). Positive allosteric modulators (PAMs) enhance the orthosteric activity of the orthosteric ligands, negative allosteric modulators

(NAMs) inhibit the activity and neutral allosteric ligands have no effect on the activity. Bitopic ligands are molecules consisting of two parts that are bridged by a linker and bind to the orthosteric and allosteric sites simultaneously. One part of the molecules is responsible for the activation and the other part is responsible for subtype selectivity (Mohr, Schmitz, Schrage, Trankle, & Holzgrabe, 2013). Each of these ligands has a different mode of action for targeting the receptors and has its advantages and disadvantages (Figure 2.5).



**Figure 2.5.** Different classes of ligands with distinct mode of actions and key properties, adapted from Kruse et al. (2014).

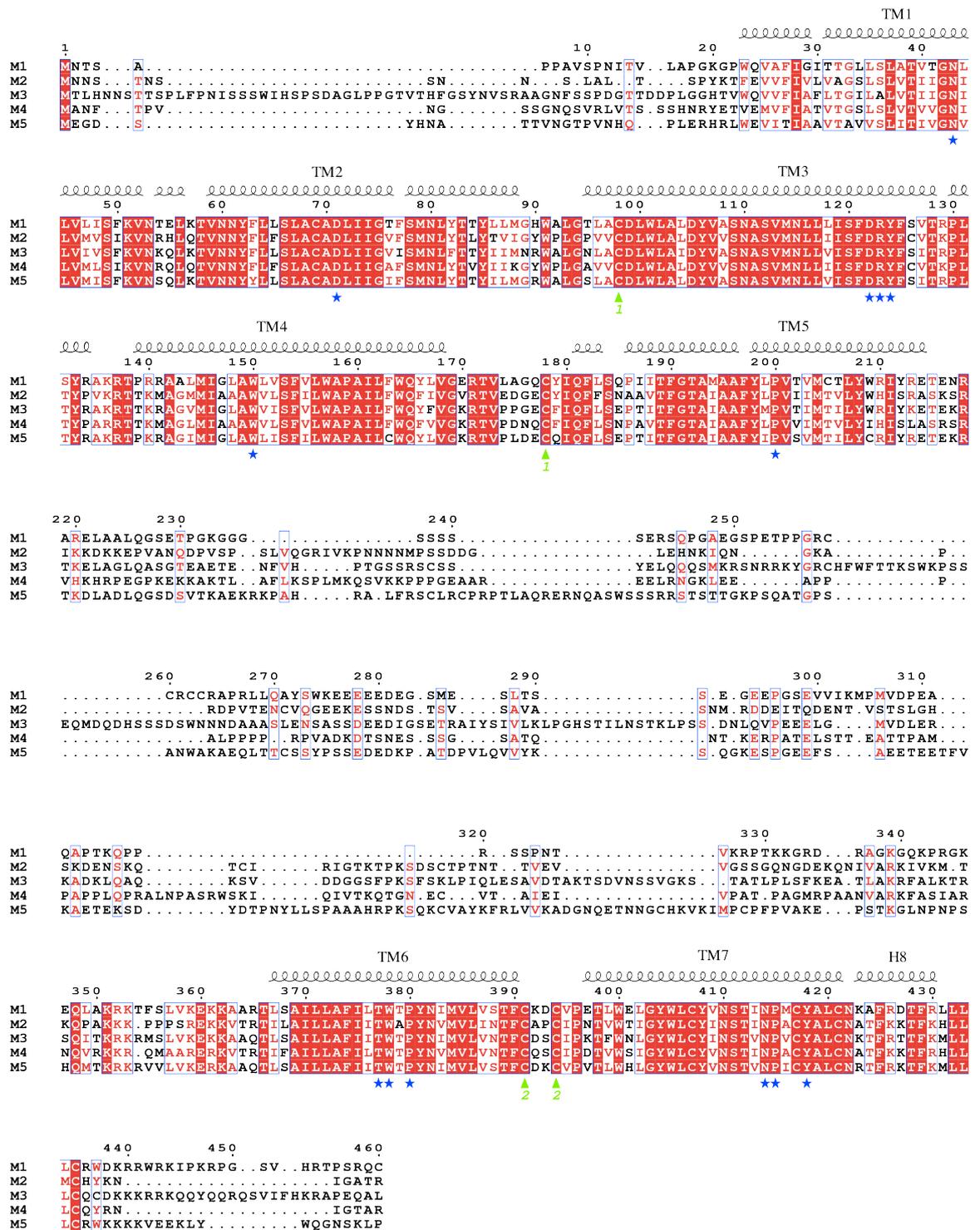
## 2.4 Muscarinic acetylcholine receptors

The mAChR family belongs to Family A/rhodopsin-like G protein-coupled receptors. The mAChRs mediate the effects of the neurotransmitter acetylcholine (ACh) in both the peripheral and the central nervous system and consist of five different subtypes, M<sub>1</sub> to M<sub>5</sub>. Upon activation, M<sub>1</sub>, M<sub>3</sub>, and M<sub>5</sub> receptor subtypes coupled with the G<sub>q/11</sub> family of G proteins, while M<sub>2</sub> and M<sub>4</sub> couple with the G<sub>i/o</sub> family of G proteins, and give two distinct outcomes, mobilization of intracellular calcium and blockage of voltage-gated calcium channels, respectively (Taylor & Brown, 2006). The sequences of the five mAChRs are highly conserved, with an average of more than 50% of sequence identity (Figure 2.6) and these mAChR subtypes share a highly conserved orthosteric binding pocket. Phylogenetic analysis indicates that the M<sub>2</sub> and M<sub>4</sub> subtypes are closer than the M<sub>1</sub>, M<sub>3</sub> and M<sub>5</sub> subtypes (Jiang et al., 2014).

Recently, X-ray crystal structures of mAChRs have provided novel insights about their structural features in their inactive (Haga et al., 2012; Kruse et al., 2012) and active (Kruse et al., 2013) conformations. Importantly, the M<sub>2</sub> mAChR structure in complex with an allosteric modulator has also been reported, providing detailed structural information about the interaction between an allosteric agent and the receptor for the first time (Kruse et al., 2013). The structures of the inactive M<sub>2</sub> and M<sub>3</sub> mAChRs resemble each other and with other biogenic amine GPCRs crystalized so far. Superposition of the M<sub>2</sub> structure in complex with antagonist QNB and the M<sub>3</sub> structure in complex with inverse agonist tiotropium indicates that the ligand configuration within the binding pockets is identical. Both binding pockets are buried within the TM core with residues forming the binding pockets residing on TM3, 4, 5, 6 and 7. These residues are fully conserved among all subtypes, which again explain the relatively lack of success to date in developing selective orthosteric ligands. These structures also

highlight several features, which are unique to the mAChR family, including a relatively large solvent accessible extracellular vestibule, which is occluded from the orthosteric site by a tyrosine lid. This outer cavity with less conserved residues represents a potential targeting surface for the development of subtype selective agents.

Similar to other active state GPCR structures, the iperoxo bound M<sub>2</sub> activated structure showed significant outward movement of the cytoplasmic end of TM6, and smaller outward movement of the C-terminal portion of TM5. Rearrangement of the highly conserved motifs – NPxxY (TM7) and D/ERY (TM3), was also observed (Kruse et al., 2013). Comparing with the structures of activated rhodopsin and  $\beta_2$ AR structures, the binding of iperoxo to the M<sub>2</sub> mAChR induced more profound structural changes to the binding pocket, leading to a significant contraction of the orthosteric binding site, which completely occluded the agonist ligand from the solvent. Despite the structural changes, iperoxo bound to the receptor with a set of polar contacts that resembled those found when QNB bound to the M<sub>2</sub> mAChR. A crystal structure has also been obtained for the iperoxo bound M<sub>2</sub> mAChR in complex with a positive allosteric modulator - LY2119620, which selectively enhances the affinity of iperoxo to the M<sub>2</sub> mAChR. This allosteric modulator was found to sit directly above the orthosteric agonist and established extensive interactions with the extracellular vestibule. Since there is no significant difference observed in the binding of iperoxo to the M<sub>2</sub> mAChR compared with the allosteric bound structure, the allosteric binding site is probably pre-formed after the binding of the orthosteric agonist. The structural changes and contractions related to iperoxo binding allow the allosteric modulator to interact with the receptor more extensively, and in return, enhance the orthosteric agonist affinity by stabilizing the active conformation and slowing down the agonist dissociation from the binding pocket (Kruse et al., 2013).



**Figure 2.6.** Sequence alignment of five human muscarinic subtypes. Identical amino acids are captured in red boxes, while similar amino acids are in red alphabet letters. Regions corresponding to the seven transmembrane helices are shown with the symbol of a spring on top. GPCR structural fingerprints are highlighted with blue stars at the bottom of the sequence and green triangles correspond to the cysteine residues involved in the formation of disulfide bridges.

#### 2.4.1 Alzheimer's disease and the M<sub>1</sub> mAChR

The M<sub>1</sub> subtype is mainly expressed postsynaptically in the central nervous system (CNS), specifically in the hippocampus, prefrontal cortex (areas critical for memory and learning) and striatum. The M<sub>2</sub> and M<sub>3</sub> subtypes are expressed both pre- and postsynaptically in the basal forebrain, thalamus, hippocampus, cardiac and smooth muscle tissues. The M<sub>4</sub> and M<sub>5</sub> subtypes are expressed presynaptically in the striatum, hippocampus, and substantia nigra, respectively. Due to their potential roles in the pathophysiology of CNS disorders such as depression, schizophrenia, Parkinson's disease and Alzheimer's disease (AD), modulation of the muscarinic system has been the focus for drug discovery efforts, particularly, the M<sub>1</sub> mAChR in targeting the treatment of AD (Broadley & Kelly, 2001).

AD is most commonly concomitant with four key hallmarks: cognitive impairment (deficits in learning and memory), cholinergic neuron dysfunction and death,  $\beta$ -amyloid plaque accumulation, and neurofibrillary tangle formation (due to tau-protein hyperphosphorylation) (Selkoe, 1991; Woolf, 1996; Fisher, 2008). The current primary treatment for AD symptoms based on the cholinergic hypothesis is the use of acetylcholinesterase inhibitors (inhibiting the enzyme responsible for ACh degradation), donepezil, galantamine, and rivastigmine, and the non competitive NMDA glutamate receptor antagonist, memantine (Zemek et al., 2014). Although the treatment is able to lessen the AD symptoms, none of the drugs are able to modify disease progression, a fact that has been the driving force for the ongoing research searching for new and potent anti-Alzheimer compounds. These treatments also often lead to undesirable cardiovascular (bradycardia) and gastrointestinal (hypersalivation, diarrhea) side effects because of their general activation. Alternatively, selective mAChR activation could possibly enhance and potentiate cholinergic signaling, by imitating ACh actions.

Selective activation of M<sub>1</sub> mAChRs is a prospective dual action treatment as they are not only capable of restoring cognitive functions but can also act as disease modifiers through promotion of the non-amyloidogenic processing of amyloid precursor protein (APP) and inhibition of tau hyperphosphorylation (Fisher et al., 2000; Hock et al., 2003; Caccamo et al., 2006), possibly reducing the amyloidogenic burden in the AD brains. M<sub>1</sub> mAChR-knockout mice show a series of cognitive deficits and impairments in long-term potentiation, showing that the M<sub>1</sub> subtype is physiologically linked to multiple functions such as synaptic plasticity, neuronal excitability, neuronal differentiation during early development, and learning and memory (Miyakawa, Yamada, Duttaroy, & Wess, 2001; Anagnostaras et al., 2003; Wess, 2004).

Stimulation of M<sub>1</sub> mAChR by agonists has been found to enhance sAPP $\alpha$  generation and reduce A $\beta$  production (Haring et al., 1995; Pittel, Heldman, Barg, Haring, & Fisher, 1996; Beach, Walker, Potter, Sue, & Fisher, 2001). The activation of protein kinase C upon M<sub>1</sub> mAChR stimulation may also promote the activity of  $\alpha$ -secretase (Cisse et al., 2011) and modulate the trafficking of APP from the Golgi/trans-Golgi network to the cell surface (Xu, Greengard, & Gandy, 1995). Some studies suggest that M<sub>1</sub> mAChR stimulation leads to the activation of ERK1/2, which can modulate  $\alpha$ -secretase activity and APP processing (Bigl & Rossner, 2003). Importantly, loss of M<sub>1</sub> mAChR has been shown to increase amyloidogenic APP processing in neurons and promotes brain A $\beta$  plaque pathology in a mouse model of AD (Davis, Fritz, Wess, Lah, & Levey, 2010).

The association of M<sub>1</sub> mAChR in AD is also manifested by its potential role in improving tau pathology. Stimulation of M<sub>1</sub> mAChR by two agonists, carbachol and AF102B, has been shown to decrease tau phosphorylation in PC12 cells (Sadot et al., 1996). Chronic treatment with AF267B also alleviates tau pathology in 3 $\times$ Tg AD mice,

through the activation of protein kinase C and inhibition of GSK-3 $\beta$  (Forlenza et al., 2000; Farias et al., 2004). A study has also observed that M<sub>1</sub> mAChR expression increased in the AD cortical tissue as a compensatory up-regulation due to the decreased of ACh levels and receptor functional activity was correlated negatively with the neuropathology severity (Overk et al., 2010). Hence, selective targeting of M<sub>1</sub> mAChRs could be beneficial in the treatment of AD.

Several mAChR agonists have been developed and clinically tested (Greenlee et al., 2001). Of these, xanomeline, a M<sub>1</sub>/M<sub>4</sub>-preferring agonist, has shown improved cognitive function in the patients but failed to achieve significant effects statistically and produced adverse cholinergic events including gastrointestinal distress, salivation, sweating and emesis (Mirza, Peters, & Sparks, 2003). Despite the unsuccessful phase III clinical trial, xanomeline has proved the therapeutic utility of the mAChR as a target for the treatment of cognitive problems. The functional involvement of M<sub>1</sub> mAChRs in the AD pathology and the potential of M<sub>1</sub> mAChR-targeted drugs for its treatment have been discussed and reviewed in detail elsewhere (Jiang et al., 2014).

#### 2.4.2 Allosteric ligands of M<sub>1</sub> mAChRs

The currently available therapies fail to relieve the symptoms and are often obstructed by dose-limiting side effects, emphasizing the need to develop novel efficacious cholinergic therapeutics without side effects mediated by peripherally located M<sub>2</sub> and M<sub>3</sub> receptors (Bymaster et al., 2003). In principle, selective M<sub>1</sub> agonists are more specific and less likely to induce side effects than anticholinesterase. However, no selective M<sub>1</sub> agonists have been developed successfully to date, as the targeting agonist binding site (primary/orthosteric) is highly conserved across all the mAChR subtypes.

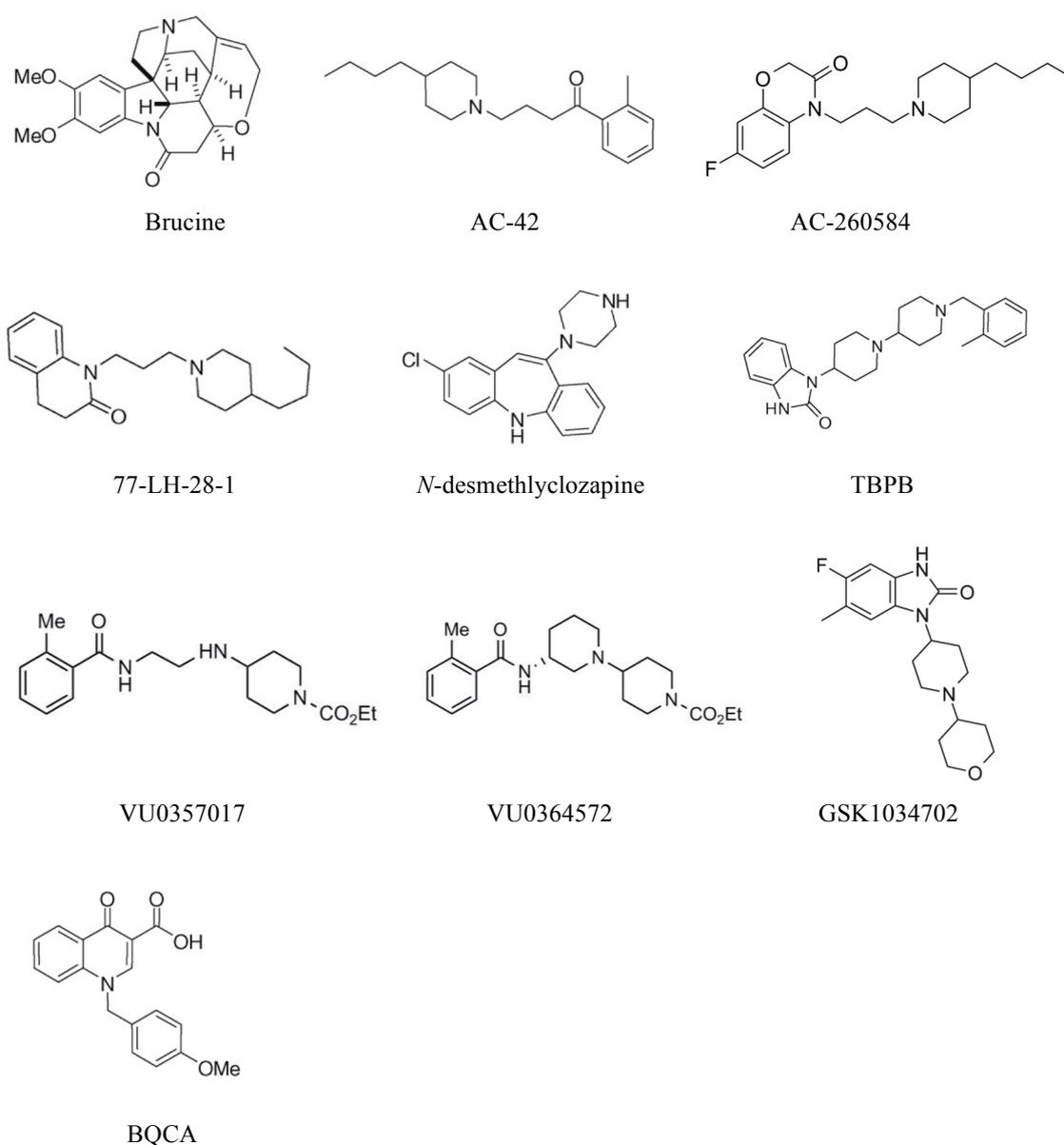
To overcome this selectivity obstacle, targeting on other non/less conserved (allosteric/secondary) sites, topologically distant from the orthosteric site, has emerged as a promising strategy. Allosteric ligands exert their effects through several mechanisms (Langmead & Christopoulos, 2006; Kenakin, 2012): direct activation (allosteric agonist), or modulation of the binding affinity/efficacy of the orthosterically bound ligand (positive allosteric modulator-PAM).

A number of novel and selective M<sub>1</sub> ligands (Figure 2.7) have been discovered and reviewed in detail by Conn, Christopoulos, and Lindsley (2009); Conn, Jones, and Lindsley (2009); Valant, Sexton, and Christopoulos (2009); Kuduk and Beshore (2012); Davie, Christopoulos, and Scammells (2013); Foster, Choi, Conn, and Rook (2014); Nickols and Conn (2014). The first reported M<sub>1</sub> mAChR PAM was brucine (Lazareno, Gharagozloo, Kuonen, Popham, & Birdsall, 1998). Even though it is a weak M<sub>1</sub> PAM, which induces a relatively small increase in the affinity of ACh for the M<sub>1</sub> mAChR and required high concentrations for its activity, it provided proof-of-concept for the development of M<sub>1</sub> subtype specific ligands (Lazareno et al., 1998). The discovery of AC-42 marked another breakthrough (Spalding et al., 2002; Langmead et al., 2006), as its mode of binding is rather distinct, occupying both the orthosteric and extracellular vestibule – the bitopic mode (Avlani et al., 2010). More potent analogs AC-260584 and 77-LH-28-1, from subsequent optimization, later superseded AC-42. While retaining their high M<sub>1</sub> selective profile, these ligands also showed possible bitopic binding modes. *N*-desmethylozapine is another M<sub>1</sub> agonist that demonstrates a non-classical orthosteric ligand binding mode and might be mechanistically similar to AC-42 (Sur et al., 2003; Spalding et al., 2006).

TBPB represents a structurally unique M<sub>1</sub> selective ligand compared to previously discovered compounds. It possesses agonistic ability and was shown to promote the non-amyloidogenic pathway of APP cleavage and decrease A $\beta$  production, demonstrating the prospective dual-action treatment by modulating the M<sub>1</sub> mAChR (Jones et al., 2008). Following functional high-throughput screening and subsequent diversity-oriented synthesis approach optimization, several compounds were identified as potent and highly selective to M<sub>1</sub> ligands (Lebois et al., 2010; Lebois et al., 2011). Of these, VU0357017 and VU0364572 behave as agonists and were classified as bitopic ligands. GSK has discovered and developed a number of noteworthy subtype specific agonists for the M<sub>1</sub> mAChR (Budzik, Garzya, Shi, Foley, et al., 2010; Budzik, Garzya, Shi, Walker, Lauchart, et al., 2010; Budzik, Garzya, Shi, Walker, Woolley-Roberts, et al., 2010; Johnson et al., 2010). A recent clinical study using GSK1034702 has successfully shown pro-cognitive efficacy improvement on episodic memory in human, in a nicotine abstinence model (Nathan et al., 2013).

Merck's BQCA is the first reported orally bioavailable M<sub>1</sub> allosteric ligand with absolute subtype selectivity, with no activity recorded for the rest of subtypes (Ma et al., 2009). It behaves both as a PAM and as an allosteric agonist (Canals et al., 2012) and promotes the non-amyloidogenic pathway of APP cleavage (Shirey et al., 2009). Subsequent optimization and extensive BQCA analog studies were carried out for further improvements (Kuduk, Chang, et al., 2010; Kuduk, Di Marco, Chang, et al., 2010; Kuduk, Di Marco, Cofre, Pitts, Ray, Ma, Wittmann, Seager, et al., 2010; Kuduk, Di Marco, Cofre, Pitts, Ray, Ma, Wittmann, Veng, et al., 2010; Kuduk, DiPardo, et al., 2010; Kuduk, Chang, et al., 2011; Kuduk, Di Marco, et al., 2011; Kuduk & Beshore, 2012; Kuduk et al., 2012; Uslaner et al., 2013). The structural basis of BQCA function and selectivity has been probed using modeling approaches recently. BQCA binding

site was found to partially overlap with the common allosteric site at the extracellular vestibule described for the M<sub>1</sub> mAChR. Y85<sup>2,64</sup>, Y179 and F182 in the ECL2, and E397<sup>7,32</sup> and W400<sup>7,35</sup> were identified as important residues contributed to the binding of BQCA. The high subtype selectivity of BQCA may be derived from either its additional contacts with the allosteric site or through a subtype-specific cooperativity mechanism (Abdul-Ridha et al., 2014).



**Figure 2.7.** Structures of M<sub>1</sub> mAChR preferring allosteric ligands.

## 2.5 Molecular modeling methods

### 2.5.1 Computer-derived protein models

Homology/comparative modeling is one of the most widely used and accurate approach to obtain protein models (Michino et al., 2009; Kufareva et al., 2011), apart from *ab-initio* modeling/topology based techniques (Vaidehi et al., 2002), and threading/fold recognition techniques (Zhang, Devries, & Skolnick, 2006). Homology modeling is based on the principal that evolutionary related proteins tend to fold into similar 3-D structures. Furthermore, protein structural signatures are preserved long after sequence modification through mutations, insertions, and deletions, and are therefore more conserved than the sequence. In homology modeling, the 3-D model of a protein (target) is built on the basis of amino acid sequence alignment with a related known structure (template). This approach follows a typical flow of 4 main steps: 1. Identification of a related protein with a solved experimental 3-D structure that can be used as template, 2. Mapping residues between the template and the target accordingly through sequence-structure alignment, 3. Building a 3-D model of the target based on the alignment, 4. Assessment of the model quality and refinement (Venclovas, 2012). In GPCR homology modeling, several practices are commonly applied and emphasized, such as attention being given to the conserved structural motif in the TM core during the sequence structure alignment, loop modeling being considered as a separate issue, refinement of the models using ligand guided optimization, induced-fit docking, or molecular dynamics simulations, and the use of multiple templates in modeling (Costanzi, 2012).

### 2.5.2 Modeling protein-ligand complexes

Once the model of the protein is built, ligand-docking algorithms predict the ligand-protein interactions through searching for the best steric and energetically favorable fit.

This includes electrostatics interactions, hydrogen bonding, van der Waals, hydrophobic interactions and the loss of entropy of the ligand upon binding. The basis of the docking methodology includes a search algorithm and an energy scoring function. Successful docking methods rely on both the ability to correctly predict the binding pose and the associated physical-chemical molecular interactions, and reliably distinguish between binders and non-binder and estimate their affinity (Guedes, de Magalhães, & Dardenne, 2014).

The protein is kept rigid while the ligand is granted full flexibility in most of the docking methods. Search algorithms can be categorized into three main groups: systematic, stochastic and deterministic. Glide (Grid-based ligand docking with energetics) (Friesner et al., 2004; Halgren et al., 2004) employs an exhaustive systematic search algorithm that uses a series of filters to explore for the best positions, orientations and conformations of a ligands, which are subsequently clustered and minimized within receptor energy grids. AutoDock uses a genetic algorithm (one of the methods of stochastic search) that randomly changes a ligand's degree of freedom at each step. The concept of a genetic algorithm is derived from the biological evolution theory, where different conformations of the ligands in a population are defined by a set of stated variables or genes, that describe the conformation of the ligands and its translation and orientation relative to the receptor (Morris et al., 1998).

During the docking procedure, many ligand poses are assessed and those that clash with the receptor are filtered out. For well-fitting ligands, a scoring function is used to discriminate between binders and non-binders. There are three types of scoring functions: force field based, empirical, and knowledge based. Empirical scoring functions are derived by fitting scoring function to experimental binding constants

obtained from a training set of protein-ligand complexes. AutoDock and Glide both implement empirical scoring functions. AutoDock makes use of the AMBER force field with energy terms empirically determined by linear regression analysis from a set of protein-ligands complexes of known binding constants (Morris et al., 1998). In Glide, the selected poses are refined using a Monte Carlo procedure and rescored using the GlideScore function, with force field based components and additional terms accounting for solvation and repulsive interactions (Friesner et al., 2004).

Refinement after the initial docking is usually considered to optimize the docking predictions, to address issues such as receptor flexibility and to correct the side chain conformation of the residues in the binding pocket. Several methods were developed for this purposes including a soft docking approach (Jiang & Kim, 1991), side chain refinement (Leach, 1994), the use of a hybrid map/explicit atom grid (Orry & Abagyan, 2012), induced-fit docking (Sherman, Day, Jacobson, Friesner, & Farid, 2005), and ensemble docking (Huang & Zou, 2007; Korb et al., 2012). Induced-fit docking (IFD) by Schrödinger considers both ligand and receptor flexibility using rigid receptor docking (Glide) and protein structure prediction (Prime) combinations. The degree of freedom of the receptor side chains is sampled and minor backbone movement is allowed. The ligand is first docked into the rigid receptor using a softened energy function followed by receptor side chain sampling and minimization of the complex for many different ligand poses to search for low free energy conformation of the receptor-ligand complex. A second round of docking is carried out using a hard potential function to continue sampling the ligand conformational space within the refined receptor environment. The complexes are then ranked using a composite scoring function. This induced-fit methodology that iteratively couples Glide and Prime has been shown to improve enrichment factor in virtual screening and is robust enough to

apply to a wide range of pharmaceutical relevant examples. The average ligand root mean square deviation (RMSD) for the traditional rigid receptor docking for 21 cases was 5.5 Å, while using IFD, the RMSD was 1.4 Å (Sherman, Beard, & Farid, 2006; Sherman, Day, Jacobson, Friesner, & Farid, 2006).

### 2.5.3 Molecular Dynamics Simulations

The molecular dynamics (MD) method is the most common method used for *in silico* studies of molecular motion and flexibility at the atomic level. It integrates Newton's second equation of motion numerically to simulate how biological systems evolve as a function of time. In MD, integrating Newton's law of motion generates successive atomic positions of the system. This will result in giving a trajectory that specifies how the positions and velocities of the atoms in the system vary as a function of time. The trajectory is obtained by solving the differential equations embodied in Newton's second law:

$$a_i(t) = \frac{d^2x_i}{dt^2} = \frac{F(x_i)}{m_i}$$

where  $a_i$  is the acceleration of particle  $i$  at time  $t$  determined by the force  $F(x_i)$  acting on particle  $i$  of mass  $m_i$  at position  $x_i$ . The force  $F(x_i)$  computed using molecular mechanics/classical approach is calculated from the derivative of the expression for potential energy as a function of position  $V(x_i)$ , which is described by a molecular mechanics force field. The force field is a compilation of atom types, parameters and equations. It is used to calculate the energy and geometry of a molecule. In classical force fields, a molecule is symbolized as a group of balls representing atoms with a fixed electronic distribution, joined together by springs representing bonds (Figure 2.8). The functional form of a force field that can be used to model single molecule or assemblies of atoms/molecules is:

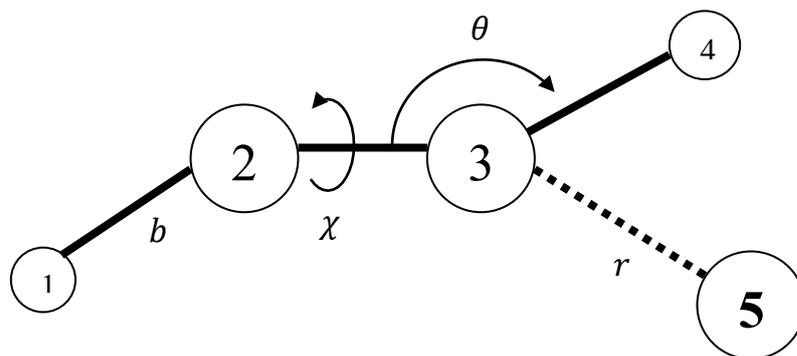
$$V = \sum_{i=1}^{N_{atom}} V_{bond}(i) + V_{angle}(i) + V_{dihedral}(i) + V_{non-bonded}(i)$$

This equation can be expressed as:

$$V(r^n) = \sum_{bonds} K_b (b - b_{eq})^2 + \sum_{angles} K_\theta (\theta - \theta_{eq})^2$$

$$+ \sum_{dihedrals} K_\chi [1 + \cos(n\chi - \delta)] + \sum_{nonbond} \epsilon \left[ \left( \frac{R_{minij}}{r_{ij}} \right)^{12} - \left( \frac{R_{minij}}{r_{ij}} \right)^6 \right] + \frac{q_i q_j}{\epsilon_1 r_{ij}}$$

where  $K_b$ ,  $K_\theta$ , and  $K_\chi$  are the bond, angle, and dihedral angle force constant, respectively;  $b$ ,  $\theta$ , and  $\chi$  are the bond length, bond angle, and dihedral angle, respectively. Coulomb and Lennard-Jones 6-12 terms contribute to the non-bonded interactions;  $\epsilon$  is the Lennard-Jones well depth and  $R_{min}$  is the distance at the Lennard-Jones minimum;  $q_i$  is the partial atomic charge,  $\epsilon_1$  is the effective dielectric constant, and  $r_{ij}$  is the distance between atoms  $i$  and  $j$ . In the CHARMM force field (force field employed in this work), a limited number of Urey-Bradley terms and improper dihedral angles are included in the function (MacKerell et al., 1998). The various contributions are schematically represented in Figure 2.8.



**Figure 2.8.** Schematic representation of force field interactions. Covalent bonds are indicated by solid lines and non-bonded interactions by a dashed line.

The velocity of individual atoms in a molecule at time  $t$  can be calculated by integrating the classical equations of motion for every atom of the system at every time step  $\partial t$ . With the use of integrators, the position of every atom in the system can be evaluated and updated as a function of time.  $\partial t$  is typically in the order of 1 or 2 femtosecond, for the integration of the equation of motion. The iteratively repeated calculations for a given number of steps eventually lead to a trajectory in time, which contain forces, positions and velocities of every atoms in the system, describing the evolution of a biomolecular system in atomistic detail (Nurisso, Daina, & Walker, 2012).

Using periodic boundary conditions is one of the practical considerations in MD, to simulate macroscopic properties with small numbers of particles and to avoid the artificial boundaries effects. In this method, the system will be surrounded with replicas of itself in all directions to give a periodic array of identical cells, representing a continuous and infinite space. During the simulation, a particle that leaves the box is replaced by an image particle, which enters from the opposite site. Thus, the number of particles in the system remains constant (Leach, 2001). The periodic boundary conditions method is associated with other approximations to treat non-bonded interactions. The Particle Mesh Ewald method (Darden, York, & Pedersen, 1993) is

commonly used to compute long-range electrostatic interactions, whilst van der Waals interactions can be truncated at a specific cut-off distance (typically between 8 and 10 Å), as they fall off quickly with distance. Another method used to reduce computing time is to constrain the degrees of freedom of covalent bond involving hydrogen bonds, as it has the highest frequency. The SHAKE (Ryckaert, Ciccotti, & Berendsen, 1977) algorithm is the most commonly used method, in which the equations of motion are constrained. There are three types of ensemble commonly employed in MD simulation, NVT, NVE and NPT. Parameters such as temperature (T), pressure (P), total energy (E), volume (V), and number of particle (N) are kept constant in different ensemble during the simulation. NPT is used widely as it can reproduce experimental conditions and is most directly relevant to experimental data as in this ensemble, the number of particles, temperature and pressure are constant. Practically, there are four phases of MD simulations, namely, system preparation, heating, equilibration, and production.

#### 2.5.4 Molecular modeling of muscarinic receptors

The highly conserved TM core and structural motifs of GPCRs have granted the use of computational approaches to predict the structures of other GPCR family members. Several muscarinic receptor models ( $M_1$  to  $M_5$ ) have been constructed over the past few years (Table 2.1). The majority of the models were used in conjunction with the pharmacology studies to elucidate receptor activation, selectivity, allosterism, bitopic binding, and dimerization, and others were used to understand the ligand binding (mainly antagonist), and to test the model reliability and ability in virtual screening for the antagonist like small molecules, by modeling approaches. A range of templates was used in modeling the structure for muscarinic receptors, including rhodopsin,  $\beta$ AR, Dopamine  $D_3$ , muscarinic  $M_2$ , and  $M_3$ , using a number of modeling software/programs such as MODELLER, Prime, MOE, ICM, and etc. The solved crystal structures of  $M_2$

and M<sub>3</sub> receptors have inspired further MD studies aiming to investigate the receptor activation and to shed light on the mechanisms of ligand (orthosteric and allosteric) binding and modulations, important for the design of mAChR subtype-selective ligands (Table 2.1).

**Table 2.1.** Modeling studies of muscarinic receptors.

Receptor	Template	Remark	Reference
M <sub>1</sub>	Rhodopsin	Modeling study to investigate whether rhodopsin-based GPCR homology models are reliable enough to be used for virtual screening of chemical databases – antagonist binding.	Bissantz, Bernard, Hibert, and Rognan (2003)
		MD study to gain insight into the dynamics and stability of the receptor.	Espinoza-Fonseca, Pedretti, and Vistoli (2008)
		Modeling study to compare protein- and ligand-based virtual screening techniques for identifying the ligands of GPCRs - antagonist binding.	Evers, Hessler, Matter, and Klabunde (2005)
		Modeling study to provide structural basis of ligands (agonists and antagonists) binding.	Peng, Vaidehi, Hall, and Goddard (2006)
		Pharmacology study of allosteric potentiation.	Ma et al. (2009)
		Modeling study to interpret the results of scanning and point mutagenesis study on the transmembrane domain.	Hulme, Lu, Saldanha, and Bee (2003)
		Modeling study to interpret the results of scanning and point mutagenesis study on the receptor.	Goodwin, Hulme, Langmead, and Tehan (2007)
$\beta_2$ AR	$\beta_2$ AR	$\beta_2$ AR - based GPCR homology modeling - antagonist binding.	McRobb, Capuano, Crosby, Chalmers, and Yuriev (2010)
		$\beta_2$ AR - based muscarinic receptors (M <sub>1</sub> -M <sub>5</sub> ) homology modeling – antagonist binding, effect of template choice.	Thomas et al. (2014)
		Pharmacology study of receptor activation and selectivity	Lebon, Langmead, Tehan, and Hulme (2009)
		Pharmacology study of allosterism and bitopic binding	Avlani et al. (2010)
		Pharmacology study of receptor activation	Kaye, Saldanha, Lu, and Hulme (2011)

		Pharmacology study of allosteric binding	Abdul-Ridha, Lane, et al. (2014); Abdul-Ridha, Lopez, et al. (2014)
	D <sub>3</sub>	Pharmacology study of allosterism and bitopic binding	Daval et al. (2012)
	M <sub>2</sub>	Homology modeling – antagonist binding	Jójárt, Balint, Balint, and Viskolcz (2012)
	M <sub>3</sub>	Pharmacology study of allosterism and bitopic binding	Daval et al. (2013)
M <sub>2</sub>	Rhodopsin	MD study – influence of different environments on M <sub>2</sub> receptor structure	Jöhren and Höltje (2005)
		MD study- conformational space of acetylcholine in M <sub>1</sub> , M <sub>2</sub> , and M <sub>5</sub> receptor models	Vistoli, Pedretti, Testa, and Matucci (2007)
	β <sub>2</sub> AR	Pharmacology study of allosterism and bitopic binding	Valant et al. (2008); Gregory, Hall, Tobin, Sexton, and Christopoulos (2010)
	M <sub>3</sub>	Modeling study to investigate the effect of template choice	Jakubik, Randakova, and Dolezal (2013)
M <sub>2</sub> crystal structure	-	MD study –to simulate receptor activation from inactive crystal structure via accelerated MD simulation, in contrast to microsecond timescale conventional MD simulations, which the receptor remained inactive.	Miao, Nichols, Gasper, Metzger, and McCammon (2013)
M <sub>2</sub> crystal structure	-	MD study – allosteric modulation, mechanisms that contribute to positive and negative allosteric modulation of ligand binding	Dror et al. (2013)
M <sub>3</sub>	Rhodopsin	MD study - to provide structural basis of ligands (antagonists) binding	Martinez-Archundia, Cordomi, Garriga, and Perez (2012)
	β <sub>1</sub> AR	Pharmacology study of dimerization	McMillin, Heusel, Liu, Costanzi, and Wess (2011)
M <sub>3</sub> crystal structure	-	MD study – ligand entry and dissociation pathway. Simulations revealed that tiotropium dissociates from M <sub>3</sub> receptors slower than from M <sub>2</sub> receptors – implication of the ‘kinetic selectivity’ of this drug for M <sub>3</sub> receptors despite similar equilibrium binding affinities for both subtypes.	Kruse et al. (2012)

## CHAPTER 3

### TOWARDS AGONIST BOUND HOMOLOGY MODELS OF THE HUMAN M<sub>1</sub> mAChR

The use of selective M<sub>1</sub> mAChR agonists has promising potential for the treatment of Alzheimer's disease since they are not only capable of restoring cognitive functions but can also act as disease modifiers through promotion of the non-amyloidogenic processing of amyloid precursor protein and inhibition of tau hyperphosphorylation (Fisher, 2008; Langmead, Watson, & Reavill, 2008; Davis, Fritz, Wess, Lah, & Levey, 2010). However, to date, such agents have not yet been used clinically due to a lack of subtype selectivity and/or intrinsic activity. To assist the design of new M<sub>1</sub> mAChR agonists, a reliable three-dimensional (3-D) structure of the receptor is required. In the absence of an experimental high-resolution structure, 3-D models can be obtained by homology modeling. A model that is able to identify agonists reliably can then be employed to retrieve allosteric modulator/understand allosteric binding, as the affinity of the allosteric modulator was found to be dependent on the occupancy of the orthosteric site (Dror et al., 2013). This chapter describes the modeling process towards agonist recognizable M<sub>1</sub> mAChR models using  $\beta_2$ -adrenergic ( $\beta_2$ AR) and M<sub>3</sub> mAChR crystal structures as templates and the limitations/challenges of GPCR modeling.

### 3.1 $\beta_2$ -adrenergic based modeling

#### 3.1.1 Introduction

The first crystal structure of a true GPCR was that of bovine rhodopsin, deposited into the Protein Data Bank (PDB) in the year 2000 (Palczewski et al., 2000), which has then been used extensively as a template to model other homologous GPCRs (Xhaard et al., 2005; Krystek, Kimura, & Tebben, 2006; Heo, Vaidehi, Wendel, & Goddard, 2007; Farce et al., 2008). Rhodopsin is formed by an extracellular N-terminal of five distorted strands, 7 transmembrane helices (TM1 to TM7) connected by alternating intracellular (ICL1 to ICL3) and extracellular (ECL1 to ECL3) loops, a disulfide bridge between ECL2 and TM3, and a cytoplasmic C-terminal containing an alpha helix parallel to the cell membrane. In 2007, a breakthrough was achieved when the crystal structure of the human  $\beta_2$ AR GPCR was solved at a resolution of 2.4 Å (Cherezov et al., 2007). Despite having similar topology and position of the ligand-binding site, several structural differences were spotted when comparing rhodopsin and adrenergic structures, notably in the N-terminus, several kinked TM helices, and in the ECL2, where an unusual pair of disulfide bridge and an extra helix was observed in  $\beta_2$ AR receptor (Rosenbaum et al., 2007). These apparent divergences have raised concerns that the rhodopsin-based homology models may suffer from shortcomings and incorrectness for ligand-docking studies (Cherezov et al., 2007). Functionally, rhodopsin is a light receptor. Unlike other GPCRs, rhodopsin binds its intrinsic ligand, 11-*cis*-retinal, covalently. It exhibits distinct functional and biochemical uniqueness compared to other GPCRs that bind to diffusible hormones and neurotransmitters, such as dopamine, serotonin, epinephrine, norepinephrine, histamine and acetylcholine. Thus, the  $\beta_2$ AR structure might become a better template than rhodopsin for homology modeling, especially for closely related members of the monoamine subfamily of GPCRs.  $M_1$  mAChR homology models have previously been constructed based on crystal structures of either bovine rhodopsin

(Bissantz, Bernard, Hibert, & Rognan, 2003; Hulme, Lu, Saldanha, & Bee, 2003; Evers, Hessler, Matter, & Klabunde, 2005; Peng, Vaidehi, Hall, & Goddard, 2006; Goodwin, Hulme, Langmead, & Tehan, 2007; Espinoza-Fonseca, Pedretti, & Vistoli, 2008) or the  $\beta_2$ AR (McRobb, Capuano, Crosby, Chalmers, & Yuriev, 2010; Thomas et al., 2014), with the latter emphasizing antagonist recognition, due to the fact that the  $\beta_2$ AR was in an inactive state with an inverse agonist (carazolol) bound to it when the crystal structure was determined. As the  $\beta_2$ AR crystal structure was the only available class A GPCR crystal structure when the project started, it was used as a template in the efforts to generate a model of the M<sub>1</sub> mAChR that is able to recognize agonists. The model was subsequently optimized and refined using molecular dynamics simulations. Model validation and prediction power were tested in stages using docking/enrichment studies.

### **3.1.2 Methods**

#### 3.1.2.1 Model construction

The construction of the M<sub>1</sub> mAChR homology model based on  $\beta_2$ AR receptor (PDB code: 2RH1) (Cherezov et al., 2007) followed well-established processes, involving: 1. identification of a suitable template; 2. retrieval of the target protein sequences and template crystal structure coordinates; 3. sequence alignment between the target and the template; 4. model building; 5. model refinement and optimization; 6. model validation and evaluation. The sequence of the M<sub>1</sub> mAChR (accession number: P11229) was retrieved from the Swiss-Prot (Boeckmann et al., 2003) sequence database. Sequence alignment was carried out using the Expresso structural sequence alignment program (Armougom et al., 2006), which takes structural aspects into account during the alignment process. The alignment was based on the classic GPCR structural fingerprints (asparagine in TM1, aspartic acid in TM2, the D/ERY motif in TM3, tryptophan in TM4 and the conserved proline residues in TM5, TM6, and TM7). The alignment was

inspected manually to avoid gaps in the TM domains and to enforce alignment of the highly conserved structural fingerprints of the class A GPCRs, including the conserved disulfide bridge. ESPript (Gouet, Courcelle, Stuart, & Métoz, 1999) was used to display the final sequence alignment of the  $\beta_2$ AR receptor, PDB code: 2RH1 (template) and the M<sub>1</sub> mAChR (target) with identical, similar, and structural fingerprints highlighted. The homology model was built using the SYBYL Biopolymer module based on the sequence alignment. The aligned main-chain coordinates of the template were copied over to the target, followed by building the structurally divergent regions (which were revealed as gaps in the alignment) using loop modeling. The side-chains were assigned using SCWRL4 (Krivov, Shapovalov, & Dunbrack, 2009).

#### 3.1.2.2 Model refinement and optimization

The homology model was refined and optimized by means of molecular dynamics (MD) simulations. Two all atom MD simulations of the M<sub>1</sub> mAChR model embedded in a fully hydrated lipid bilayer were carried out using Amber 10 (Case et al., 2010), based on two different sets of force field, Amber and CHARMM force field. For the system employing the Amber force field, the starting structure was generated using the following steps: 1. orient the model along the Z-axis using VMD (Humphrey, Dalke, & Schulten, 1996); 2. assign internal water molecules to the model using DOWSER (Zhang & Hermans, 1996); 3. create a water shell according to the protein shape using SOLVATE (Grubmueller, 1996) and remove the water shell in the TM domain; 4. insert the model into a lipid bilayer consisting of 256 palmitoyl-oleoyl-phosphatidyl-choline (POPC) molecules (Jo, Kim, & Im, 2007); 5. remove overlapping lipid and water molecules; 6. solvate the whole system according to the dimensions of the lipid bilayer using VMD. The model was positioned across the lipid bilayer, with the hydrophobic portion of the TM domain matched with the layer formed by the hydrophobic lipid

hydrocarbon tails and the intracellular loops (ICLs), extracellular loops (ECLs), and the N- and C- terminals sticking out from the lipid bilayer toward the water layers. After removing all the overlapping lipid and water molecules, the final system consisted of 25,578 water molecules, 209 POPC molecules, and 20 counter-ions, making a total of 112,054 atoms.

For the system based on the CHARMM force field, the starting structure was generated by the CHARMM-GUI (Jo, Kim, Iyer, & Im, 2008) and the constructed system had a total of 89,491 atoms, which contained 19,545 water molecules, 174 POPC molecules, and 117 counter ions. The Chamber program (Crowley, Williamson, & Walker, 2009) from AmberTools was used to convert the CHARMM psf, associated coordinated file, parameter and topology to a CHARMM force field enabled version of AMBER's prmtop and inpcrd.

A total of 27.65 ns (Amber force field based) and 63.65 ns (CHARMM force field based) MD simulations were performed using AMBER 10 (Case et al., 2008). For the Amber force field based system, the Amber ff99SB force field (Hornak et al., 2006) was used to describe the model, water molecules and the counter-ions. Since the force field for lipids is not available in Amber (at time when the simulations were carried out), force field parameters and RESP charges for lipids have been developed and calculated following Amber standard protocols. Antechamber was used to determine GAFF atom types for POPC, and 70 diverse conformations of POPC were selected randomly for charge calculations. After performing an ab-initio HF-SCF at 6-31g\* calculation using Gaussian 03 (Frisch et al., 2004), the partial charges were then extracted with Antechamber following the RESP protocol (Case et al., 2010). For the CHARMM force field based system, CHARMM36 (Kluda et al., 2010) was used to describe the system.

MD simulations of both systems were performed using constant pressure and temperature, NPT ensemble, maintaining the pressure and temperature at 1.0 atm and 310 K, respectively, by means of anisotropic pressure scaling and Langevin dynamics. The periodic boundary conditions based on the particle mesh Ewald method with a non-bonded cutoff of 8 Å were used. The integration time step was set at 2 fs and the SHAKE algorithm was used to constrain bond lengths involving hydrogen atoms. The protocol used for both systems was the same. The system underwent three stages of minimization using steepest descent and conjugate gradient with different parts of the system gradually released in stages. Then, it was slowly heated from 0 to 310 K within 150 ps with restraints on the model and the lipid molecules. After the heating stage, the system was equilibrated for 3.5 ns, with weak restraints on the model followed by another 24 ns and 60 ns of a relaxed MD run for Amber force field based and CHARMM force field based system, respectively. The complete system trajectory was collected every 2 ps for analysis. Trajectories analyses were performed using the PTRAJ modules of AMBER 11. Energy minimum conformations were extracted from the trajectories by computing the lowest potential energy of the system. Cluster analysis was carried using average linkage algorithm to produce four clusters for each selection by comparing RMS metric of the selected atoms. The clustering was based on different selection/region of the receptor, such as, whole receptor, ECLs, ECL2 and ECL3, orthosteric binding site, orthosteric binding site and ECLs, and TM only.

### 3.1.2.3 Model validation and evaluation

The stereochemical and residue contacts of the model were evaluated using PROCHECK (Laskowski, Macarthur, Moss, & Thornton, 1993) and ERRAT (Colovos & Yeates, 1993). Flexible ligand docking of a set of agonists and antagonist was performed using AutoDock4 (Morris et al., 2009) to check the ability of the model to

display correct binding modes. The coordinates of ACh were retrieved from PubChem (Bolton, Wang, Thiessen, & Bryant, 2008) and AutoDockTools was used to prepare the input files and docking grids. The grid box was set to cover the TM domain and AutoGrid was used to pre-calculate interaction energies of various ligand atom types with the receptor. AutoDock then used the grid maps to determine the total interaction energy of a ligand with the receptor. The Lamarckian genetic algorithm search method was employed with each population consisting of 300 individuals. The calculation was set to a maximum number of 2500000 energy evaluations with a maximum number of 2500000 generations (Morris et al., 1998). The ligand-receptor interactions were analyzed using DS Visualizer v3 (Accelrys Inc., San Diego, CA, USA).

#### 3.1.2.4 Induced-fit docking

Induced-fit docking (IFD) from Schrodinger suite 2011 (Schrödinger LLC, New York, USA) was used to refine the binding site residues side-chain by docking muscarine into the M<sub>1</sub> mAChR models generated before and after MD simulations. A box of 28 Å x 28 Å x 28 Å with the centroid set to the binding site (D105<sup>3.32</sup>, Y106<sup>3.33</sup>, W157<sup>4.57</sup>, T189<sup>5.39</sup>, T192<sup>5.42</sup>, W378<sup>6.48</sup>, Y381<sup>6.51</sup>, Y404<sup>7.39</sup> and Y408<sup>7.43</sup>) of the M<sub>1</sub> mAChR was used. The IFD protocol began with a constrained minimization of the receptor, followed by initial Glide docking of the ligand using softened potential, to allow more poses of ligand to be generated and collected. One round of Prime side-chain prediction were carried out for the residues found within 5 Å of the ligand followed by minimization of each receptor-ligand complex. Finally, the ligand was rigorously re-docked into the induced-fit receptor structure, without softened potential and binding energy or IFD score estimation was obtained (Sherman, Day, Jacobson, Friesner, & Farid, 2005). The final model was chosen after multiple iterations of docking and refinement of the receptor binding site.

### 3.1.2.5 Enrichment studies

A set of 51 reported agonists (Figure 3.1) for the M<sub>1</sub> mAChR was retrieved from GLIDA, the GPCR-ligand database (Okuno et al., 2008) and IUPHAR-DB (Sharman et al., 2011). The 3-D coordinates of the agonists were obtained from PubChem (Bolton et al., 2008). A set of 1000 drug-like decoy compounds (set I) was chosen randomly and downloaded from ZINC (Irwin & Shoichet, 2005) database, using the Lipinski rule of five as a selection criteria and with molecular weight ranging from 140-600 g/mol. The molecular properties and Tanimoto similarity scores of the agonists and decoys were calculated using Discovery Studio v3.1 (Accelrys Inc.) with predefined parameters. The properties of the decoys were found to be similar to those of the agonists (Table 3.1). All compounds were prepared using LigPrep to assign appropriate protonation states, generate tautomers and optimize geometry prior to the docking calculation. The decoy set enriched with the 51 agonists was docked into the M<sub>1</sub> models using Glide and ranked by GlideScore. The docking site was defined using a box of 28 Å x 28 Å x 28 Å covering the active site of M<sub>1</sub> as described earlier. One pose per ligand was collected for analysis. The enrichment factors (EF) were calculated at 2, 5 and 10% of the total database screened, using the following equation:

$$EF = (\text{Hits}_{\text{sampled}} / N_{\text{sampled}}) \div (\text{Hits}_{\text{total}} / N_{\text{total}})$$

**Table 3.1.** Distribution of molecular properties and Tanimoto similarity scores for agonists and decoys.

Library / Property	ALogP			Molecular Weight			Fractional polar surface area		
	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD
Agonists	1.3	1.0	2.0	241.2	206.3	97.5	0.2	0.2	0.1
Set I	1.2	1.3	1.7	238.2	204.2	92.6	0.3	0.2	0.2

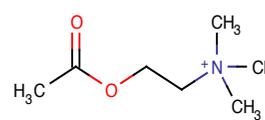
  

Library / Property	Hydrogen bond acceptor			Hydrogen bond donor			Rotatable bonds		
	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD
Agonists	3.2	3.0	1.5	0.4	0.0	0.6	3.5	3.0	2.3
Set I	2.9	3	1.7	1.0	1.0	0.9	3.3	3.0	2.0

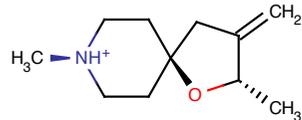
  

Library / Property	Number of aromatic rings			Number of rings			Tanimoto similarity score*		
	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD
Agonists	0.8	1.0	0.8	2.1	2.0	1.4	0.14	0.10	0.16
Set I	1.1	1.0	1.0	1.8	2.0	1.3	0.08	0.07	0.04

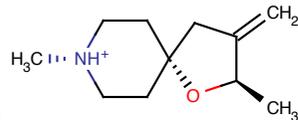
\*Tanimoto similarity scores take values from 0 – 1, with scores closer to 1 indicating greater similarity.



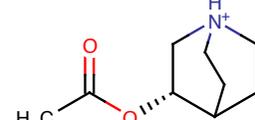
Acetylcholine



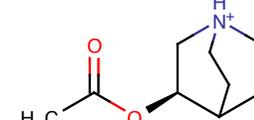
(-)-YM796



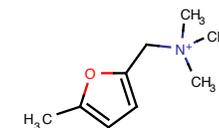
(±)-YM796



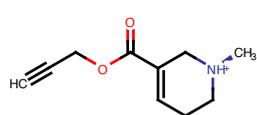
(+)-Aceclidine



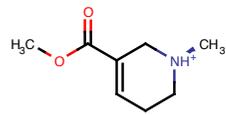
(-)-Aceclidine



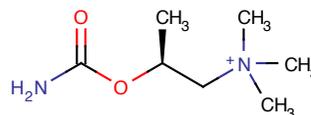
5-Methylfurfurethonium



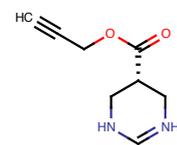
Arecaidine-propargyl-ester



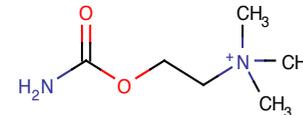
Arecholine



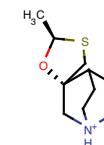
Bethanechol



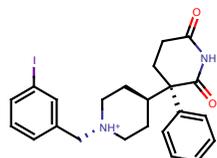
CDD0097



Carbachol



Cevimeline



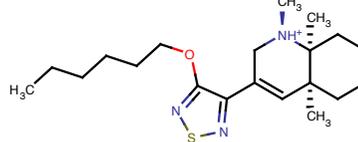
CHEMBL10278



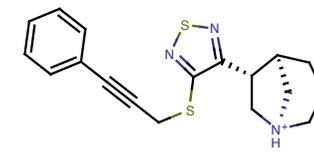
CHEMBL21322



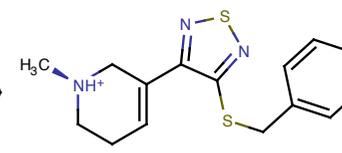
CHEMBL49915



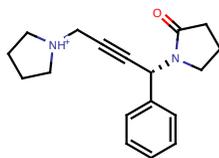
CHEMBL71380



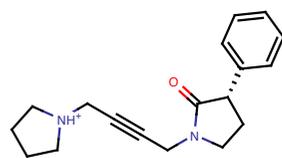
CHEMBL100147



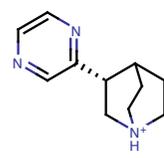
CHEMBL100253



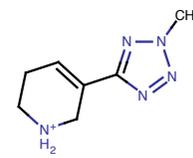
CHEMBL101645



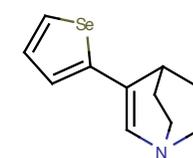
CHEMBL101646



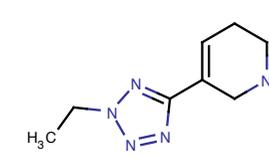
CHEMBL102560



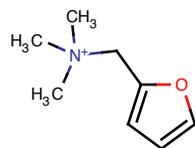
CHEMBL130625



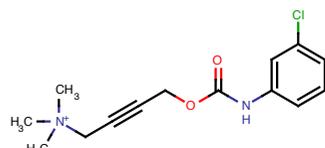
CHEMBL144123



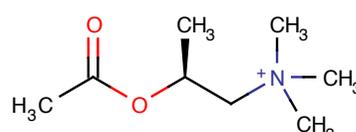
CHEMBL337676



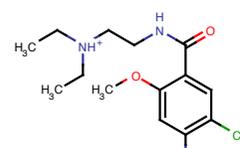
Furfurethonium



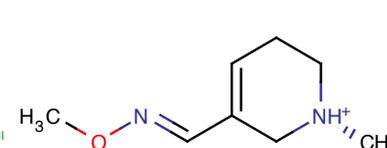
McN-A-343



Methacholine



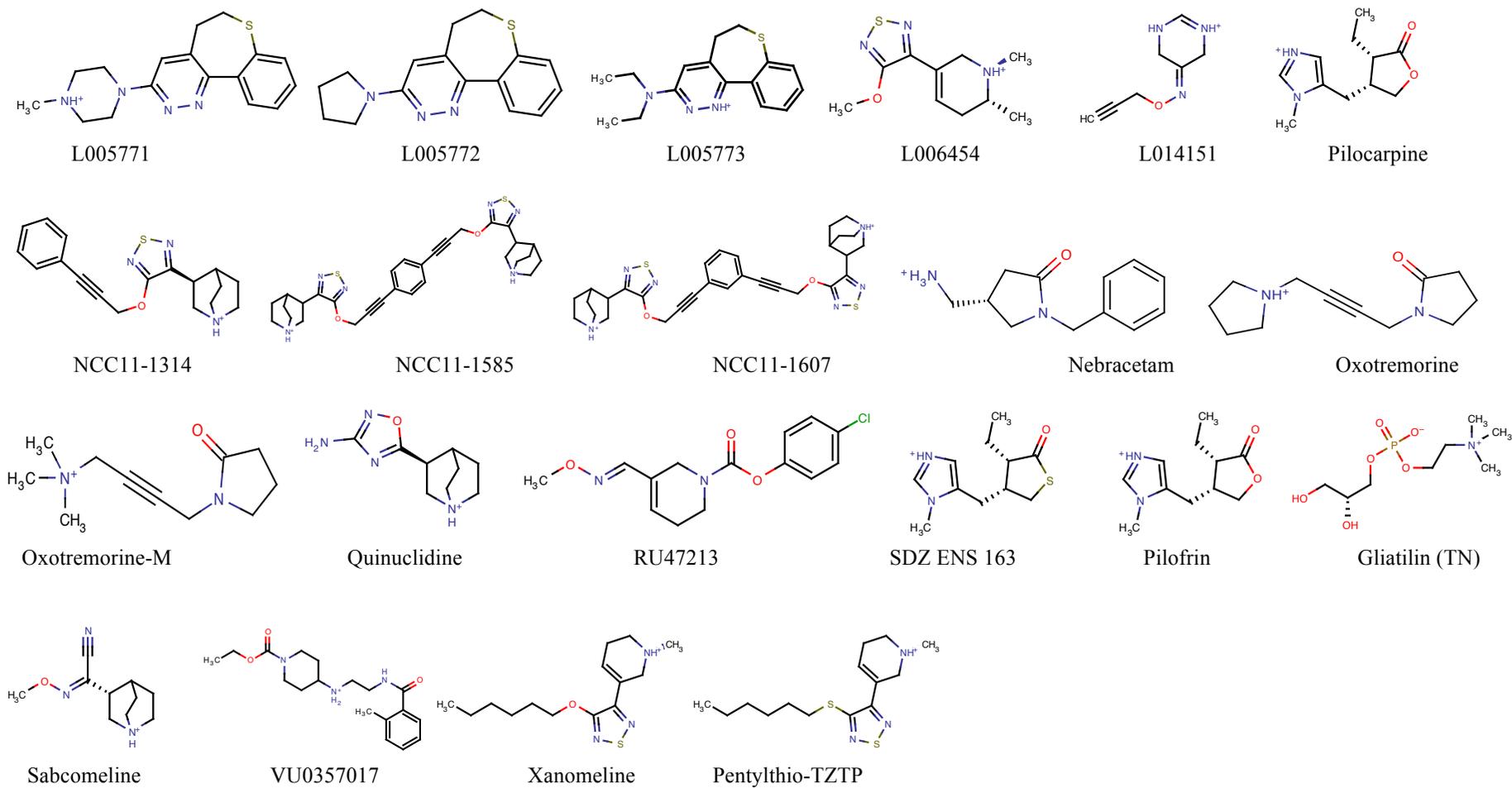
Metoclopramide



Milameline



Muscarine

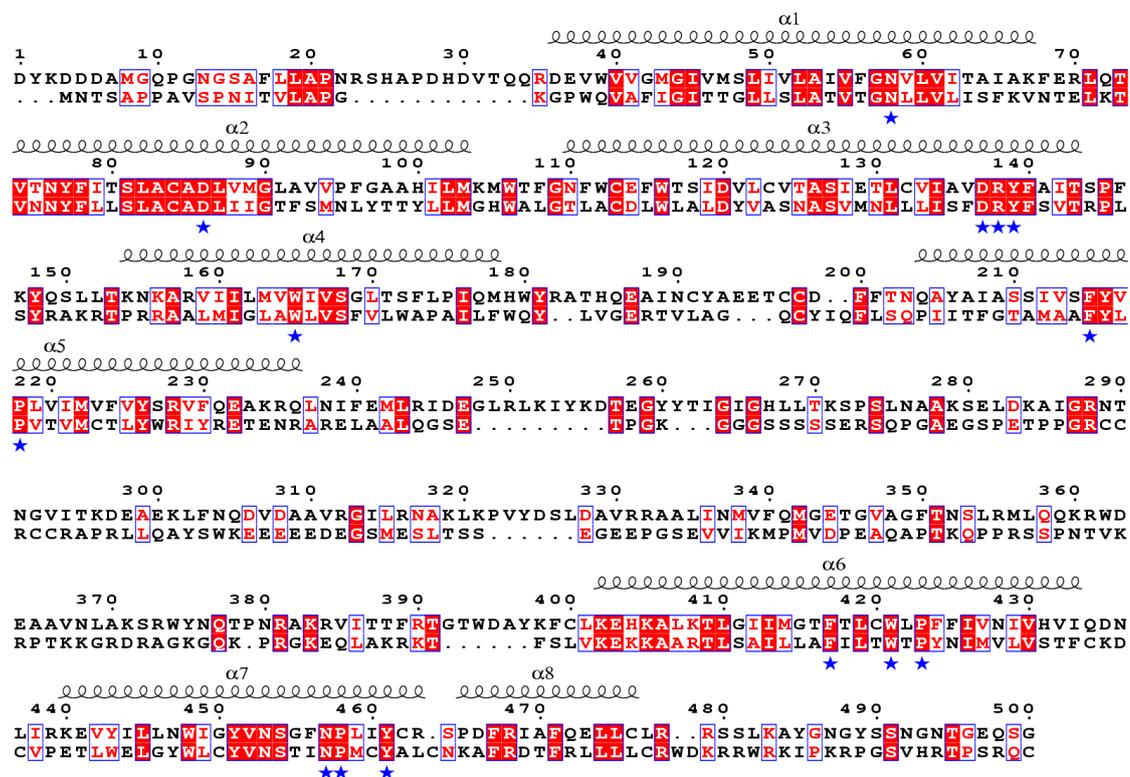


**Figure 3.1.** M<sub>1</sub> mAChR agonists used in the enrichment studies.

### 3.1.3 Results and Discussion

#### 3.1.3.1 Sequence alignment and model construction

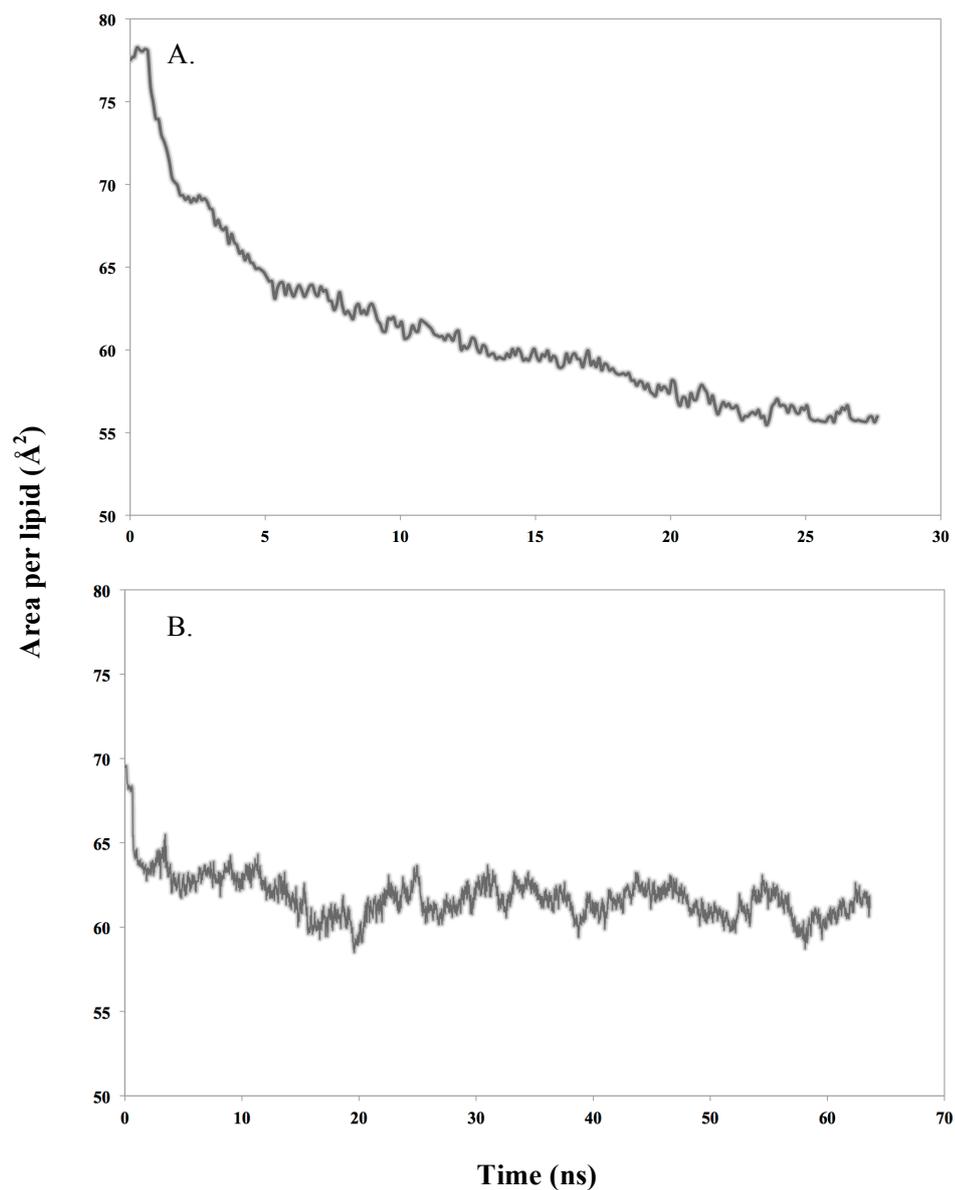
The sequence alignment guided with GPCR structural fingerprints showed that the template and target shared 35% of sequence identity in the TM domain, and 44% of sequence similarity for the whole receptor (Figure. 3.2). The Expresso score of the alignment was 77, suggesting that a good alignment was achieved (score > 50). The E-value from the Blast search gave a value of 5e-34, indicating that the  $\beta_2$ AR receptor in complex with an inverse agonist carazolol (PDB code: 2RH1) and the M<sub>1</sub> mAChR are closely related and have sufficient homology for modeling (Altschul et al., 1997). Based on this sequence alignment, an M<sub>1</sub> mAChR homology model was constructed and was further refined using MD simulations.



**Figure 3.2.** Sequence alignment between the  $\beta_2$ AR receptor (top) and the M<sub>1</sub> mAChR (bottom). Identical amino acids are captured in red boxes, while similar amino acids are in red alphabet letters. Regions corresponding to the seven transmembrane helices are shown with the symbol of a spring on top. GPCR structural fingerprints are highlighted with blue stars at the bottom of the sequence.

### 3.1.3.2 Molecular dynamics simulations

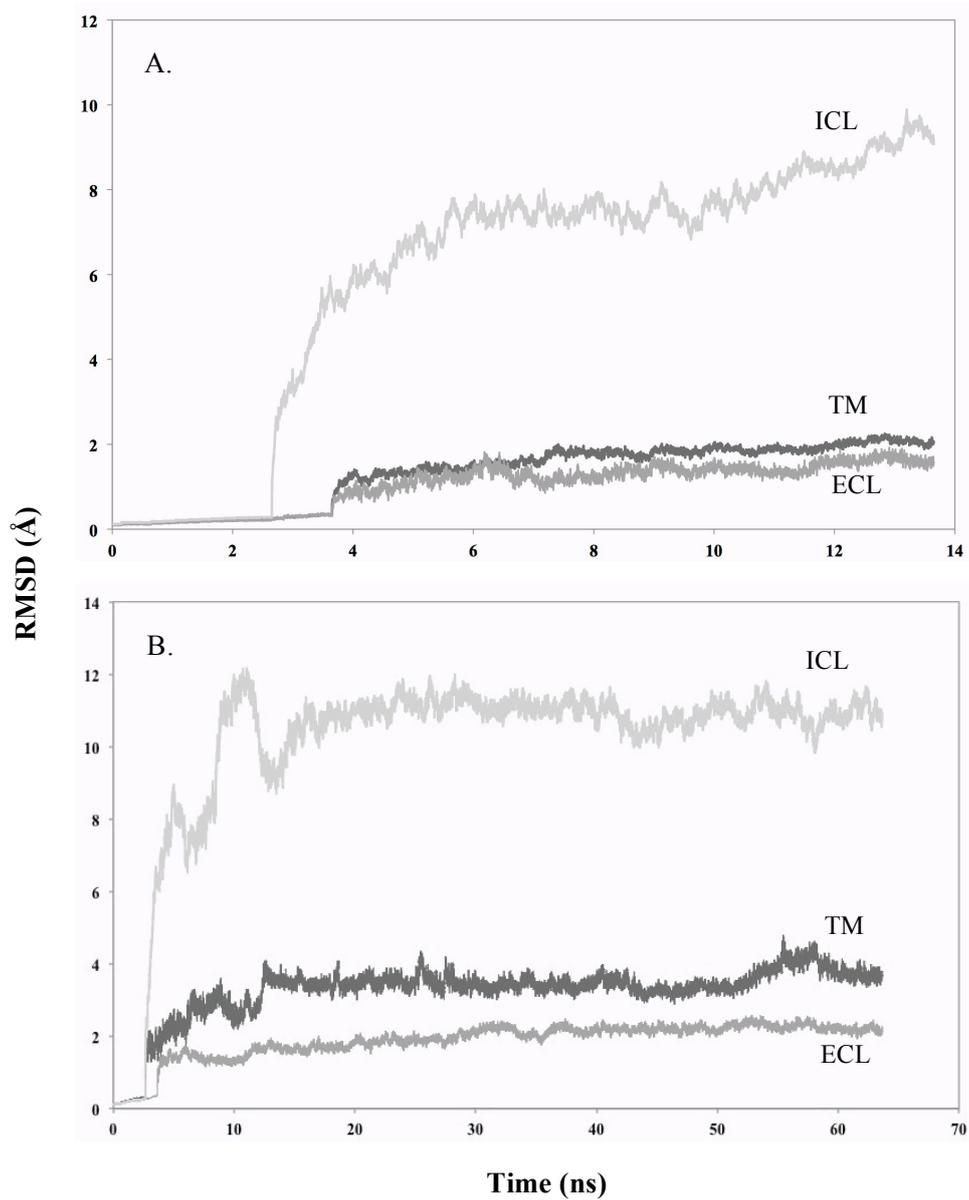
The behavior of the lipid bilayer in explicit membrane simulations plays a crucial role in defining the stability of the simulations (Filizola, Wang, & Weinstein, 2006). The area per lipid for the simulations employing the Amber and CHARMM force field, respectively, demonstrated two different outcomes, as shown in Figure 3.3. The area per lipid obtained from the Amber force field based simulation decreased over time, indicating shrinkage of the lipid bilayer. This artifact may affect the overall receptor structure as shrinking of the lipid bilayer compresses the receptor structure located in the middle of it and might further distort the structure, and hence the reliability of the observations from this set of simulations cannot be assured. The shrinkage of the lipid bilayer may be due to the force field parameterization or the fact that the simulations were run without using surface tension. However, the use of surface tension to restrain the lipid bilayer/artificially kept its area constant is still debatable, and its impacts on the resulting observations are unclear (Marrink & Mark, 2001; Benz, Castro-Roman, Tobias, & White, 2005; Jojart & Martinek, 2007). In contrast, the CHARMM force field based simulation, the area per lipid fluctuated around a mean value of about  $62 \text{ \AA}^2$ , close to the experimentally observed value ( $63 \text{ \AA}^2$ ). Due to this reason, the Amber force field based simulations were only continued until 24 ns of production run and the CHARMM force field based simulations were allowed to run until 60 ns.



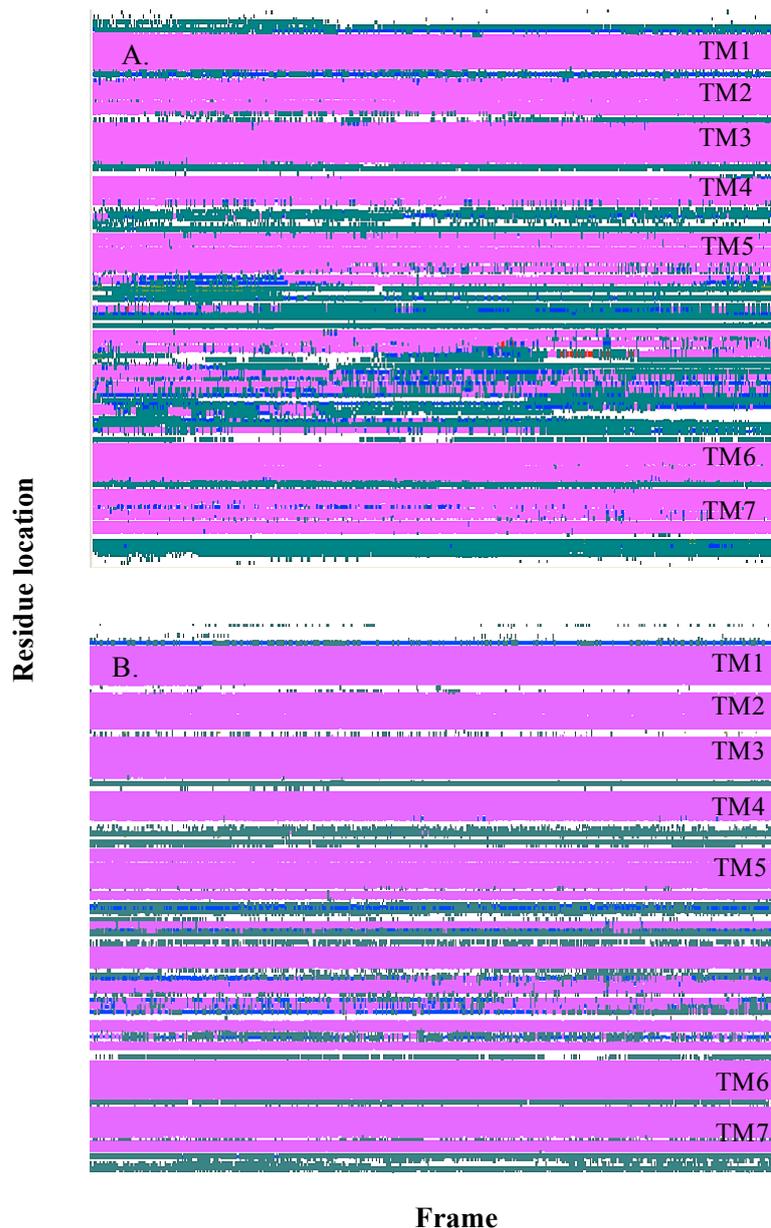
**Figure 3.3.** The area per lipid of the **A.** Amber force field and **B.** CHARMM force field based simulations of the M<sub>1</sub> mAChR homology model as a function of time.

The root mean square deviation (RMSD) for the M<sub>1</sub> mAChR homology model computed using the backbone atoms of the TM helices, ECLs and ICLs, with respect to the minimized starting structure at 310 K as a function of time is shown in Figure 3.4. For the RMSD analysis of the Amber force field based simulations, only the first 10 ns of the production run is shown, before the lipid bilayer started to shrink. The

simulations were well equilibrated after 4 ns and the RMSD of the backbone atoms of the TM helices remained close to 2 Å, achieving a plateau, over the last 6 ns and 50 ns, for the Amber force field based and CHARMM force field based simulations, respectively. By comparison, the RMSD of the ICLs showed noticeably greater disorder in both systems. The extracellular loops had a stability that was comparable to the TM helices. The reduced flexibility of ECL2 is due to the existence of a conserved disulfide bond between C98<sup>3.25</sup> in TM3 and C178 in ECL2 and the fact that they are relatively shorter in length, compared with the ICLs. Despite the structural fluctuations in the ICLs, there was no evidence of significant structural modifications in the TM domains. ECLs and ICLs seem to be important for holding the structure together and maintaining the TM domain, throughout the MD simulations. This is useful in avoiding severe distortions in the receptor structure. Considering that the ECLs and ICLs are non-regular secondary structures and were constructed using loop modeling based on lower sequence identity with the template in these regions, significant rearrangement was to be expected. Moreover, the flexibility of the ICLs, especially ICL3, is expected to be reduced in the presence of the G-protein. A closer inspection of the MD trajectories revealed that the secondary structures of the M<sub>1</sub> mAChR in the TM domain are well conserved throughout the simulations (Figures 3.5).



**Figure 3.4.** RMSD of the backbone atoms of the M<sub>1</sub> mAChR homology model from the minimized starting structure as a function of time for **A.** Amber force field based simulations, **B.** CHARMM force field based simulations. RMSD was computed for the TM core, ICLs and ECLs separately.

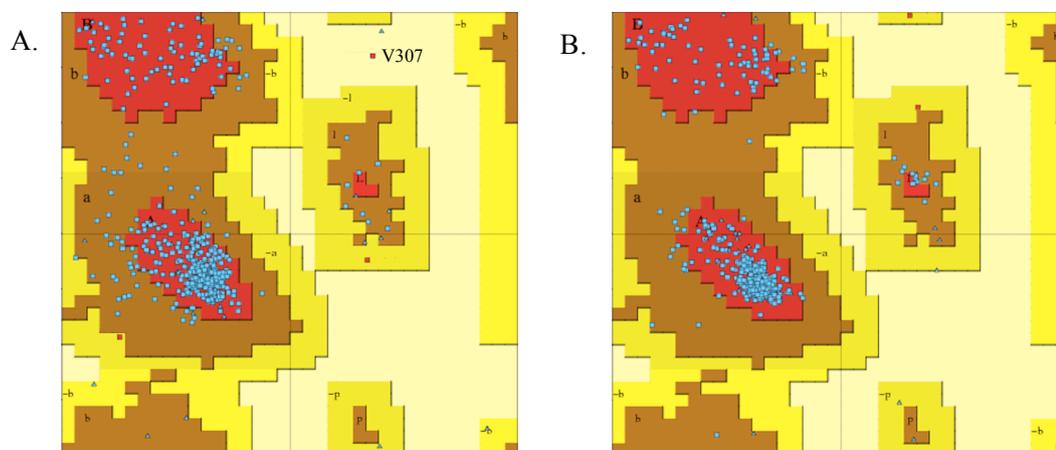


**Figure 3.5.** Secondary structure evolution of the M<sub>1</sub> mAChR homology model from the last 10 ns trajectory of **A.** Amber force field based simulations, **B.** CHARMM force field based simulations. Color code: purple,  $\alpha$ -helix; blue,  $3_{10}$  helix; red,  $\pi$ -helix; cyan, turn; and white, coil. Full stretch of color bar, especially the purple indicated that the secondary structure of the helix was well maintained.

### 3.1.3.3 Stereochemical quality of the models

The lowest energy conformation of the M<sub>1</sub> mAChR homology model extracted from the Amber force field based and CHARMM force field based MD simulations of the

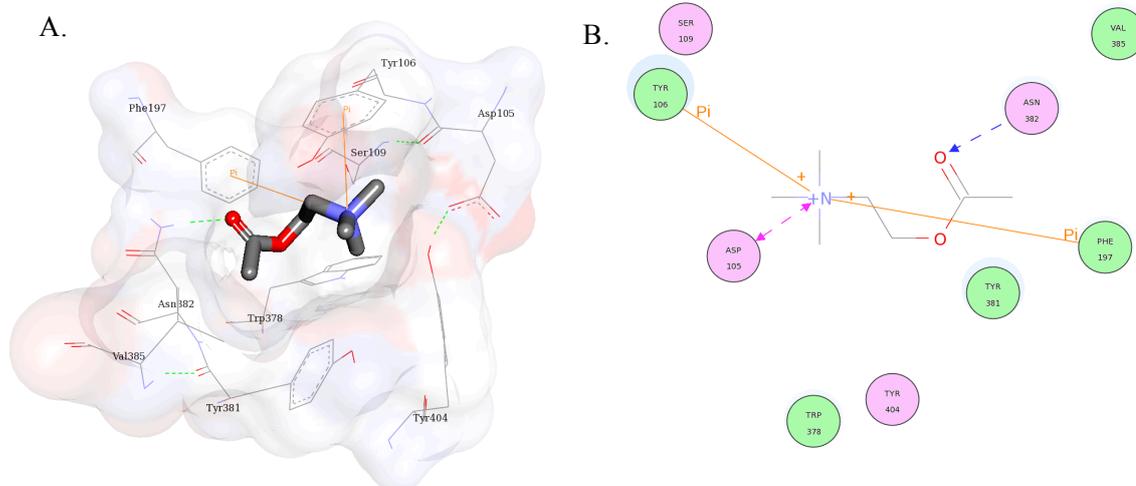
complete system achieved an overall quality score of 90 upon evaluation with ERRAT, which can be considered as acceptable since good high-resolution crystal structures usually give quality scores of 95 or higher (Colovos & Yeates, 1993). Evaluation of the stereochemical properties using PROCHECK showed that 99.7% of the residues in the M<sub>1</sub> mAChR model from the Amber force field based system fell in the favored and allowed regions and the only residue found in the disallowed region was V307, which is located in ICL3 (Figure 3.6). As for the CHARMM force field based system, 100% of the residues in the M<sub>1</sub> mAChR model were found within the favored and allowed regions (Figure 3.6). These analyses suggested that the M<sub>1</sub> mAChR models from both systems were of reliable and acceptable quality in terms of backbone conformation and overall atomic non-bonded interactions.



**Figure 3.6.** The Ramachandran plot of the **A.** lowest energy conformation of M<sub>1</sub> mAChR homology model from Amber force field based simulations. 99.7% of the residues are found within the favored and allowed regions with an outlier of V307, and **B.** lowest energy conformation of M<sub>1</sub> mAChR homology model from CHARMM force field based simulations. 100% of the residues are found within the favored and allowed regions. Disallowed regions are colored in light yellow.

#### 3.1.3.4 Model evaluation using docking

Overall, the docking results correlate well with the experimental site-directed mutagenesis (SDM) data as exemplified by the lowest energy conformation of the M<sub>1</sub> model from the Amber force field based simulations. ACh is known to adopt a conformation directing the protonated nitrogen of the choline moiety towards D105<sup>3.32</sup> and Y106<sup>3.33</sup>, and the ester moiety positioned towards T192<sup>5.42</sup> and N382<sup>6.52</sup>. These common patterns involved in ACh binding to the M<sub>1</sub> mAChR were observed in the docking studies (Figure 3.7). D105<sup>3.32</sup>, Y106<sup>3.33</sup>, S109<sup>3.36</sup>, F197<sup>5.47</sup>, Y381<sup>6.51</sup>, N382<sup>6.52</sup>, V385<sup>6.55</sup>, W378<sup>6.48</sup>, and Y404<sup>7.39</sup> were found in the close vicinity of ACh, forming the binding pocket. A key ionic interaction was formed between the protonated nitrogen in ACh and the highly conserved D105<sup>3.32</sup>. The protonated nitrogen also formed cation-pi interactions with the aromatic rings in the side chains of Y106<sup>3.33</sup> and F197<sup>5.47</sup>. In addition, a hydrogen bond was found between N382<sup>6.52</sup> and the carbonyl oxygen of ACh ester moiety. A network of hydrophobic residues that include Y106<sup>3.33</sup>, F197<sup>5.47</sup>, Y381<sup>6.51</sup>, V385<sup>6.55</sup>, and W378<sup>6.48</sup> appeared to participate in favorable van der Waals contacts with ACh (Figure 3.7). Based on SDM studies, D105<sup>3.32</sup>, S153<sup>4.53</sup>, W157<sup>4.57</sup>, P159<sup>4.59</sup>, I188<sup>5.38</sup>, T189<sup>5.39</sup>, T192<sup>5.42</sup>, A193<sup>5.43</sup>, F197<sup>5.47</sup>, Y381<sup>6.51</sup>, N382<sup>6.52</sup>, Y404<sup>7.39</sup> and Y408<sup>7.43</sup> have been identified as the key contributors for ligand binding. Alanine-substitution mutagenesis experiments on these residues have also resulted in the reduction of ACh binding affinity by at least 5-fold, together with S109<sup>3.36</sup>, N110<sup>3.37</sup>, L386<sup>6.56</sup> and C407<sup>7.42</sup> (Ward, Curtis, & Hulme, 1999; Allman, Page, Curtis, & Hulme, 2000; Lu, Saldanha, & Hulme, 2001; Bee & Hulme, 2007).



**Figure 3.7.** Binding pose of ACh. **A.** A three-dimensional representation of a fully buried ACh molecule in the binding cavity of the lowest energy conformation of the M<sub>1</sub> mAChR homology model from the Amber force field based simulations. **B.** A two-dimensional interactions map of ACh and the binding site residues. Hydrogen bond: blue dashed line with arrow head directed towards the electron donor; Pi interactions: orange line with symbols indicating the interaction; ionic interaction: pink dashed line with arrow heads on both sides. Residues involved in hydrogen bond, polar/ionic interaction, are represented with magenta circles, while residues involved in van der Waals interactions are represented by green circles.

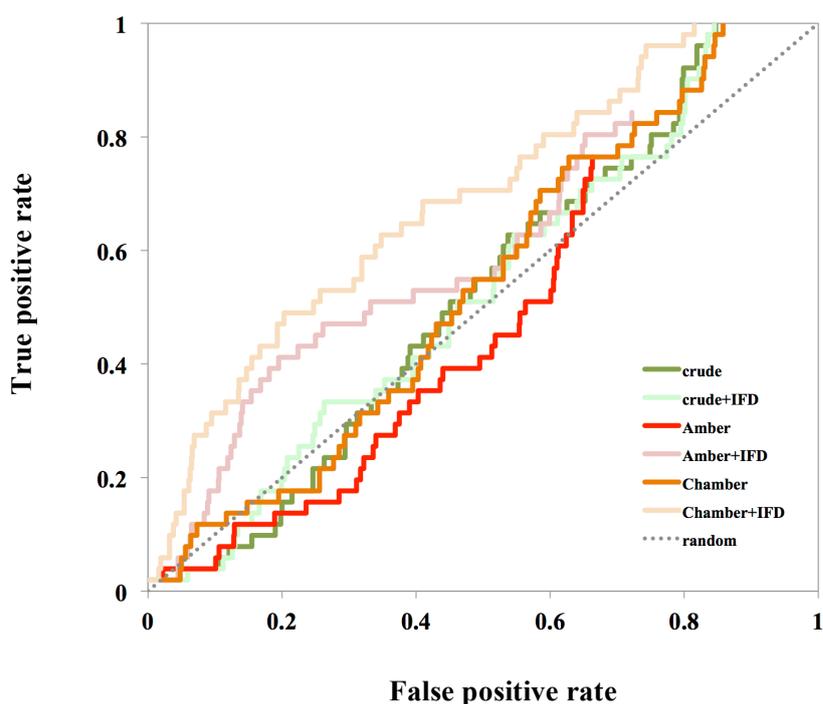
### 3.1.3.5 Enrichment studies

For the Amber force field based simulations, 7 minimum energy structures were extracted from the first 10 ns of the production run trajectory, while 6 minimum energy structures were extracted from the last 60 ns of the trajectory of the CHARMM force field based simulations. Cluster analysis then produced a further 80 conformations, 40 for each system. At first, ACh was docked to these 93 receptor structures using AutoDock and those structures that do not model the binding of ACh correctly were filtered out, leaving a total of 11 structures from the Amber force field based system and 8 structures from the CHARMM force field based system. Subsequently, 11 agonists were docked to these models, with the grid box covering only the TM domain. Two out

of 19 structures, where more than half of the ligands were bound to the orthosteric site, were selected for further enrichment studies. These 2 models (Amber and Chamber) are the lowest energy structure from the Amber force field based and CHARMM force field based systems, extracted at 11.560 ns and 12.324 ns, respectively.

Before evaluating the selected models with enrichment studies, IFD was employed to refine the binding site again. The crude model before MD simulations was also included in this refinement. For each of the IFD generated complexes, the position of the ligand and its interactions with the receptor residues are visually inspected. Figure 3.8 shows the Receiver Operating Characteristic (ROC) curve of 6 structures using virtual screening of a set of 1000 decoy ligands enriched with 51 known  $M_1$  agonists (Figure 3.1). The Chamber model after IFD outperformed the other models in the test. From the curve and the calculated EF (Table 3.2), it is obvious that a crude model is not suitable to be employed in virtual screening studies, even after IFD. The Amber model also does not show good enrichment, even though it passed the validation of docking with 11 agonists. This indicated that the binding site of Amber model is not suitable for more diverse structures of agonists. In fact, through pocket analysis using fpocket (Le Guilloux, Schmidtke, & Tuffery, 2009), it was shown that the volume of the binding pocket of the Amber model is relatively smaller than the Chamber model. This might be due to the shrinkage of the lipid bilayer that further compresses the structure. However, after IFD, the Amber model dramatically improved its prediction power and started to pick up active compounds at a better rate. Overall, IFD improved the enrichment for all the three selected models, but it is not encouraging. Through the enrichment studies, it highlights the importance of refinement steps and suggests that a crude model with only energy minimization is not enough to give reasonable structure for binding studies. It also

revealed the ability of IFD to quickly refine the binding pocket and produce better enrichment. Even though structures after MD simulations did not show a good enrichment curve, the side chain of W157<sup>4.57</sup> that was originally facing outward was flipped in facing the binding cavity, causing the bulky Y106<sup>3.33</sup> to be re-oriented to accommodate W157<sup>4.57</sup>. Energy minimization is not able to produce the same outcome (Figure 3.9).



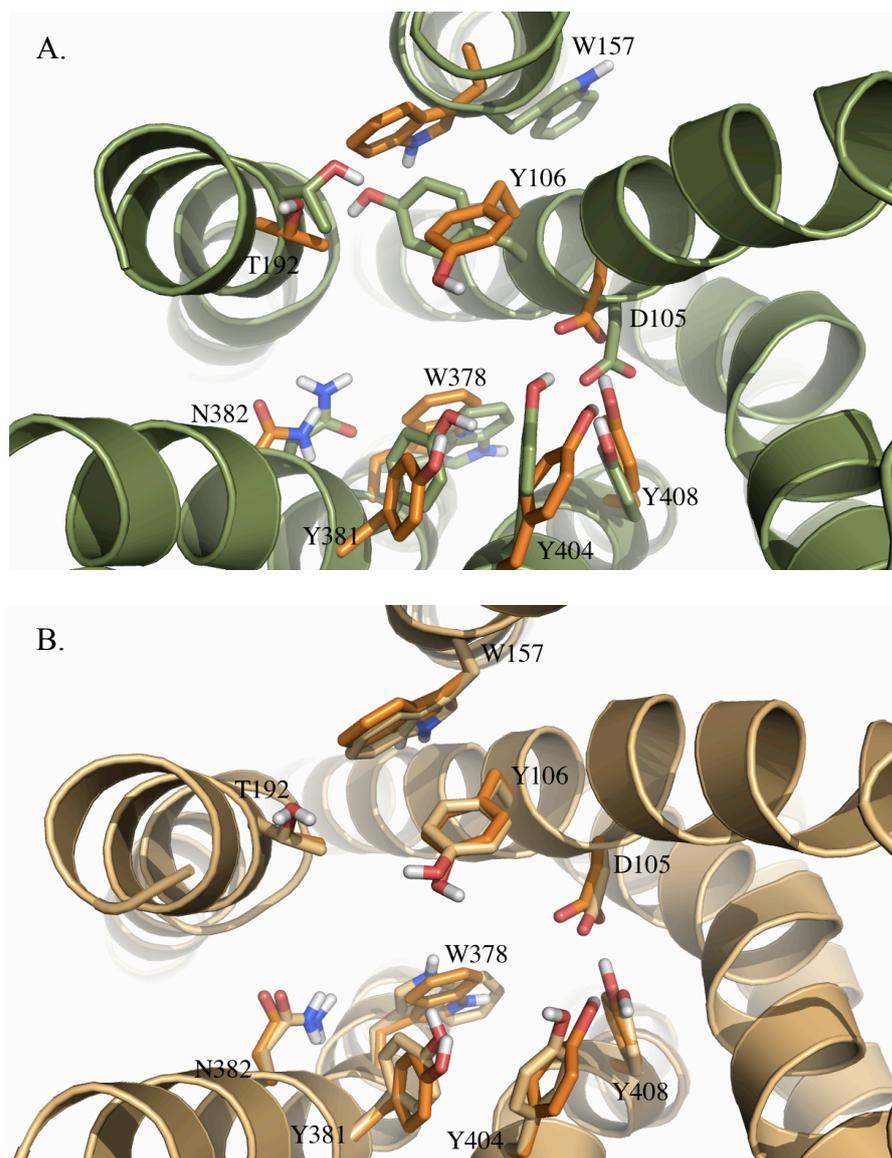
**Figure 3.8.** Receiver operating characteristic (ROC) agonist enrichment plots using decoy set I for the different homology models.

**Table 3.2.** Enrichment factors for the different homology models at  $x\%$  of screened library using decoy set I.

Model	Enrichment factor		
	2%	5%	10%
Crude	0.98	0.78	0.39
Crude+IFD	0.0	0.0	0.39
Amber	0.0	0.78	0.39
Amber+IFD	0.0	1.2	1.8
Chamber	0.0	0.78	1.2
Chamber+IFD	2.9	2.7	3.1

The relatively poor EF is probably due to the difficulties in generating the correct conformation for ECL2, which was found to play a role in binding. Besides, the differences of the residues forming the binding site between the template (2RH1) and the M<sub>1</sub> mAChR resulted in difficulties in optimizing the orientation of the side chain of binding site residues. This is proven when a bulky Y at position 404 in TM7 replaces a small residue such as N in 2RH1. IFD also generated more Y404<sup>7.39</sup> side chain orientations indicating that this particular residue required more refinement as compared to other binding site residues. Superposition of the binding site before and after IFD showed that the side chain orientation of Y404<sup>7.39</sup> and W378<sup>6.48</sup> differ greatly (Figure 3.9). W378<sup>6.48</sup> is known to be a toggle switch that triggers the conformational changes from inactive to active via changes in the side chain orientation. It was also observed that the changes of W378<sup>6.48</sup> side chain as a result of IFD refinement did produce different enrichment curves. Furthermore, to model the binding of agonist is challenging due the fact that the template (2RH1) used is an inactive structure and that the relatively smaller agonist to bind and to fill up the bigger pocket pre-form for antagonist to establish optimum interactions with the binding site residues.

Enrichment studies are important as an indicator of whether the generated model is suitable for virtual screening to find new leads. In an ideal case, a receptor structure or target should find about 80% of the actives at about 20% of the screened library. However, this ideal condition is normally achievable only by the experimentally solved structure, such as an X-ray crystallography 3-D structure with bound ligand. Expectedly, homology models usually display lower enrichment curves.



**Figure 3.9.** Orientations of the binding site residues for different models. **A.** Superposition of active site of the crude (green) and chamber (orange) models. **B.** Superposition of active site of the chamber (orange) and chamber+IFD (light orange) models.

### 3.1.4 Conclusion

Homology models of the  $M_1$  mAChR based on a crystal structure of the  $\beta_2$ AR receptor have been developed. The models have been optimized and refined with MD simulations, involving a complete solvated system of the  $M_1$  mAChR embedded in a lipid bilayer, and the trajectory has been analyzed. The lowest energy conformations from the MD simulations have been validated and evaluated using molecular docking.

Intermolecular ligand-receptor interactions at the binding site agreed well with the available site-directed mutagenesis data. However, the enrichment studies revealed the limitations of the models to separate the actives and decoys adequately, even after MD simulations. Models after IFD did show improvements but more related templates or template with higher sequence identity especially in the binding site are probably required.

## **3.2 M<sub>3</sub> mAChR based modeling**

### **3.2.1 Introduction**

Recently, crystal structures of the M<sub>2</sub> and M<sub>3</sub> mAChRs have been solved (Haga et al., 2012; Kruse et al., 2012). Since these receptors are expected to bear closer structural resemblance to the M<sub>1</sub> mAChR than bovine rhodopsin or the  $\beta_2$ AR receptor due to their high degrees of sequence similarity, they provide an opportunity to develop improved 3-D models of the M<sub>1</sub> mAChR. Although an M<sub>1</sub> mAChR model has already been generated based on the crystal structure of the human M<sub>2</sub> mAChR complexed with an antagonist (PDB code: 3UON), similar to previous M<sub>1</sub> mAChR models (Bissantz et al., 2003; Evers et al., 2005; McRobb et al., 2010; Thomas et al., 2014), it was only reported to be able to identify antagonists in virtual screening experiments (Jójárt, Balint, Balint, & Viskolcz, 2012), probably due to the fact that these models were all based on inactive structures.

Due to the unsatisfying results from the previous  $\beta_2$ AR-based M<sub>1</sub> models, efforts were continued to construct models of the human M<sub>1</sub> mAChR based on the crystal structure of the rat M<sub>3</sub> mAChR (PDB code: 4DAJ), modified using the agonist-bound crystal structure of a  $\beta_2$ AR receptor (PDB code: 3SN6) (Rasmussen, DeVree, et al., 2011) as a

guide, and refined by induced-fit docking (IFD) with acetylcholine (ACh). In contrast with previous  $\beta_2$ AR based modeling, the models were generated and refined using Schrödinger suite 2011 (instead of SYBYL, AutoDock4 and MD), and enrichment studies were carried out using two additional decoy sets (set II and III), for better and more accurate performance measurement. The abilities of these models to differentiate agonists from not only decoy molecules but also antagonists were investigated using additional parameters such as Boltzmann-enhanced discrimination of receiver operating characteristic (BEDROC) and ROC area under the curve (AUC). The binding poses of docked agonists and antagonists were also examined and expanded to the whole collection of agonists and antagonists to map the structural interaction fingerprint. To the best of my knowledge, these are the first reported  $M_1$  receptor models, which are able to identify agonists effectively, and thus can be used as targets for structure-based discovery of novel compounds for the treatment of Alzheimer's disease.

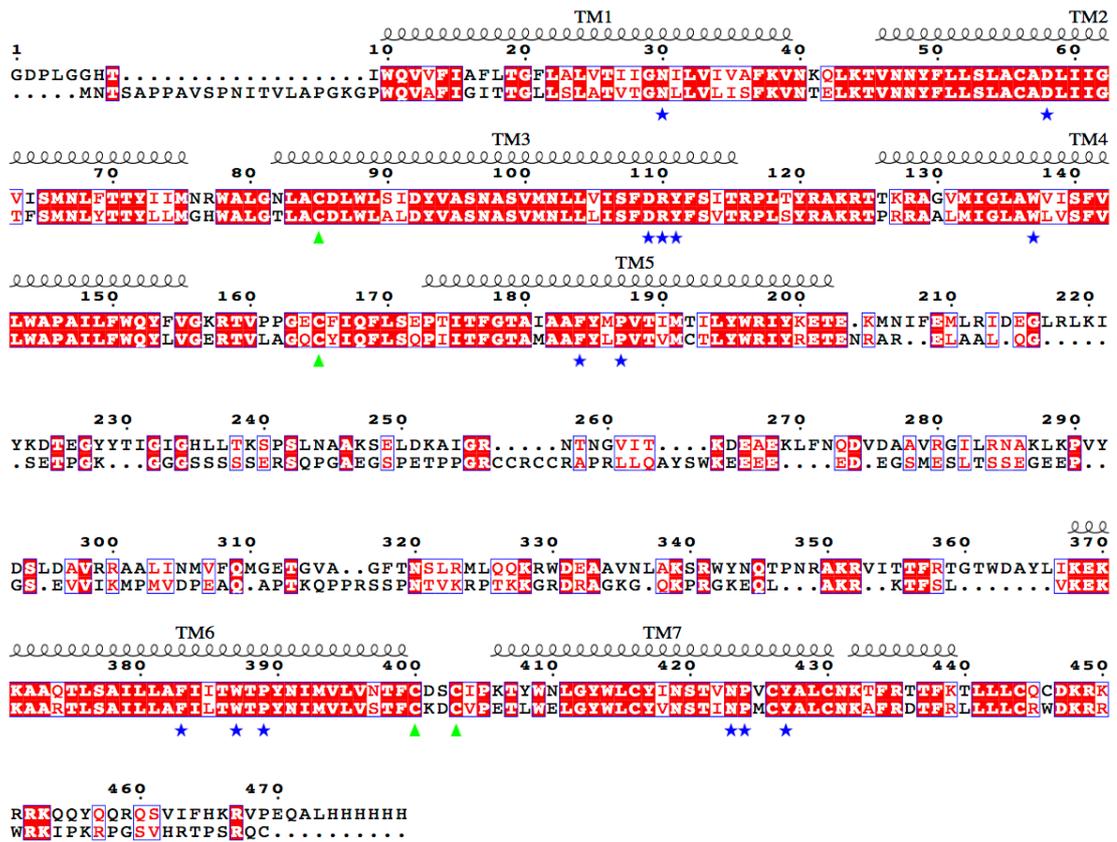
### **3.2.2 Methods**

All-atom molecular models were generated and IFD, docking and enrichment studies were performed using Schrödinger suite 2011 (Schrödinger LLC, New York, USA) with default settings and parameters, unless stated otherwise (Friesner et al., 2004; Halgren et al., 2004; Sherman et al., 2005) Visual inspections were carried out with the aid of Maestro v9.2 (Schrödinger LLC), and Discovery Studio Visualizer v3.1 (Accelrys Inc., San Diego, CA, USA). Binding mode analyses were performed using the scripts within Maestro and Discovery Studio Visualizer 2-D interaction diagrams. VMD (Humphrey et al., 1996) and PyMOL v1.3 (Schrödinger LLC) were used to produce 3-D figures. Structural validation of models was carried out using PROCHECK (Laskowski et al., 1993) and WHATCHECK (Hooft, Vriend, Sander, & Abola, 1996) For ease of comparison and standardization, residues are labeled using both their amino

acid sequence number and Ballesteros-Weinstein nomenclature, as a superscript extension (Ballesteros & Weinstein, 1995).

### 3.2.2.1 Model construction

The M<sub>3</sub> mAChR based homology modeling following the methods previously described in section 3.1.2.1. The sequence of the human M<sub>1</sub> mAChR (accession number: P11229) was retrieved from the Swiss-Prot sequence database. Guided alignment of the human M<sub>1</sub> receptor sequence and that obtained from the structure of the rat M<sub>3</sub> mAChR complexed with an inverse agonist (PDB code: 4DAJ), was carried out using the Expresso structural sequence alignment program. The Expresso score of the alignment was 99, suggesting a very good and reliable alignment (score > 50). ESPript was used to display the final sequence alignment (Figure 3.10). The M<sub>1</sub> and M<sub>3</sub> receptor sequences were found to share 53% sequence identity overall, with the seven TM domains sharing 79% sequence identity. In contrast, following a similar alignment of the human M<sub>1</sub> mAChR receptor sequence and that obtained from the crystal structure of the human M<sub>2</sub> mAChR complexed with an antagonist (PDB code: 3UON), this pair of sequences was found to share a lower sequence identity (45% overall, 69% for the seven TM domains). For this reason and also due to the fact that the M<sub>3</sub> receptor, like the M<sub>1</sub> receptor, is an “odd-numbered” stimulatory muscarinic receptor and so has greater functional similarity to the M<sub>1</sub> receptor, the M<sub>3</sub> receptor structure was used as the template for homology modeling. Following the sequence alignment, a crude homology model of the M<sub>1</sub> receptor was constructed and optimized using Prime v3.0 (Schrödinger LLC). The orientations of the amino acids in the putative binding pocket, which are completely conserved across mAChRs, were frozen and retained in the crude model. ECLs and ICLs connecting the TM domains were modeled according to the template structure except ICL3, which was excluded from the modeling.



**Figure 3.10.** Sequence alignment between the  $M_3$  mAChR (top) and the  $M_1$  mAChR (bottom). Identical amino acids are captured in red boxes, while similar amino acids are in red alphabet letters. Regions corresponding to the seven transmembrane helices are shown with the symbol of a spring on top. GPCR structural fingerprints are highlighted with blue stars at the bottom of the sequence and green triangles correspond to the cysteine residues involved in the formation of disulfide bridges.

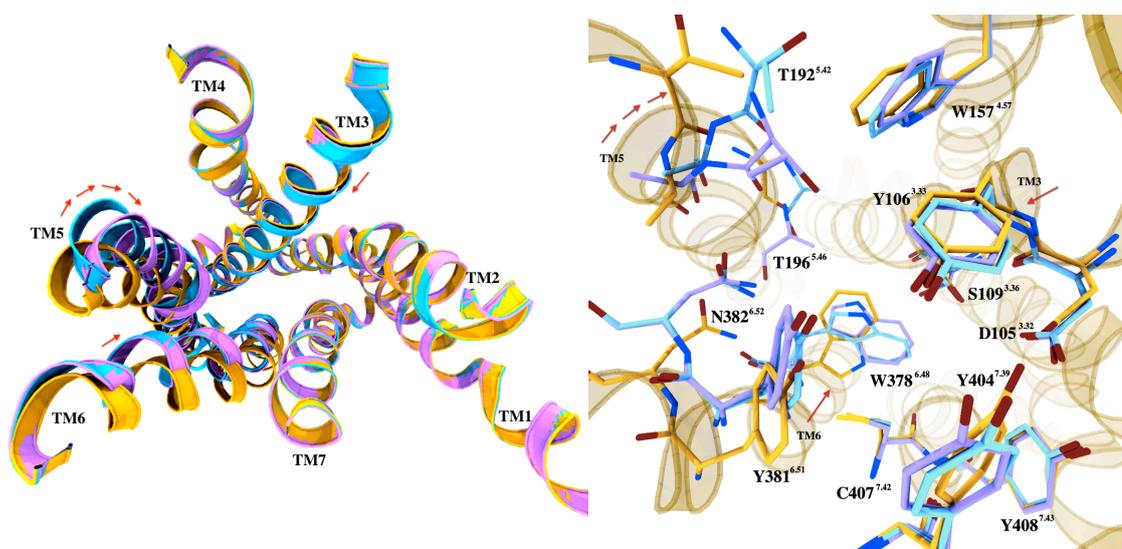
### 3.2.2.2 Binding site refinement

Since the shape of the binding pocket of the template is likely to have been influenced by the bound inverse agonist (tiotropium), binding site refinement was carried out by constructing a cubic box of 28 Å x 28 Å x 28 Å, centered on the centroid of selected known orthosteric site residues (D105<sup>3.32</sup>, Y106<sup>3.33</sup>, W157<sup>4.57</sup>, T189<sup>5.39</sup>, T192<sup>5.42</sup>, W378<sup>6.48</sup>, Y381<sup>6.51</sup>, Y404<sup>7.39</sup> and Y408<sup>7.43</sup>) of the  $M_1$  mAChR suggested by SDM studies (Lu & Hulme, 1999; Ward et al., 1999; Allman et al., 2000; Lu et al., 2001;

Spalding et al., 2006; Goodwin et al., 2007), followed by IFD of the endogenous muscarinic receptor agonist (ACh) using Glide v5.7 (Schrödinger LLC) and side chain refinement using Prime v3.0. The IFD protocol used was as described in section 3.1.2.4. After multiple iterations of docking and binding site refinement, an initial set of models was chosen on the basis of showing the expected ligand-receptor interactions. The resulting ligand-receptor conformations were inspected visually to ensure that the quaternary amine group of the docked ACh was directed towards the conserved D105<sup>3.32</sup> and the side chains of the orthosteric site residues were facing inwards towards the inner channel of the TM region. The best model (henceforth referred to as model 1) was then selected on the basis of enrichment study results.

#### 3.2.2.3 Generation of agonist-bound models

To generate an activated M<sub>1</sub> mAChR model, the best IFD-refined model (model 1), was superimposed on the activated  $\beta_2$ AR receptor structure (PDB code: 3SN6) and, as depicted in Figure 3.11, TM3, TM5 and TM6 were shifted laterally with reference to the  $\beta_2$ AR receptor structure using the GPCR Helix Manipulator module in Maestro v9.2, following the precedent set for the activation of a  $\beta_2$ AR receptor (Vilar et al., 2011). The side chain rotamer of T192<sup>5.42</sup> was then adjusted and the resulting structure subjected to IFD with ACh to produce model 1A1. Further modifications on model 1A1 were carried out whereby TM5 was rotated clockwise and tilted towards the inner channel of the TM region and the resulting structure was again subjected to IFD to give model 1A2. The models reported were selected on the basis of enrichment study results.



**Figure 3.11.** Superposition of models 1 (yellow), 1A1 (blue) and 1A2 (purple) with the important interacting side chain orientations for each of the models shown in stick representation. Loops are not shown for the purpose of clarity.

#### 3.2.2.4 Enrichment and docking studies

Enrichment studies were carried out to test the ability of the refined models to prioritize agonists over decoys. The agonists were retrieved and prepared as described in section 3.1.2.5. In addition to decoy set I, two extra sets of decoys were used in these studies. The Schrödinger set (Set II), containing 1000 drug-like ligands, was selected randomly from a library of one million compounds from [www.schrodinger.com](http://www.schrodinger.com) (Friesner et al., 2004; Halgren et al., 2004), whereas the property-matched set (Set III, provided by Dr. David Chalmers from Monash University, Melbourne), containing 1499 decoy molecules, was derived from the ZINC database (Irwin & Shoichet, 2005), based on having properties which were matched to those of each of the agonists, i.e. in numbers of heavy atoms (same as parent  $\pm 3$ ), H-bond donors (same as parent  $\pm 1$ ), H-bond acceptors (same as parent  $\pm 2$ ), rotatable bonds (same as parent  $\pm 2$ ), number of rings (same as parent  $\pm 1$ ), and with a Tanimoto similarity score  $< 0.5$  with respect to each other and the parents (to ensure topological diversity). The molecular properties and

Tanimoto scores of the agonists and decoys were calculated using Discovery Studio v3.1 (Accelrys Inc.) with predefined parameters. Although neither decoy set was substantially biased in terms of polarity (log P, numbers of donors and acceptors) or aromaticity (numbers of aromatic rings), compared with the agonists, Set III was more closely matched (Table 3.3). In contrast to Set II, which was not specifically chosen to match the agonists, Set III can be considered as efficient challengers for the agonists (having similar molecular properties without being chemically similar). Ligands were prepared using LigPrep v2.5 (Schrödinger LLC) to assign appropriate protonation states, generate tautomers and optimise geometry prior to docking calculations. The decoy sets enriched with the 51 agonists was flexibly docked into the M<sub>1</sub> receptor models using Glide v5.7 (Friesner et al., 2004; Halgren et al., 2004) and the bound poses were ranked by GlideScore. The grid box settings were the same as those used for IFD, accommodating ligands with a length of 20 Å or less and with a default inner box of 10 Å on each side. Glide Standard Precision scoring functions were used and one pose per ligand was collected for analysis. Visual inspection was carried out to ensure that the ligands were bound within the defined binding pocket and the structural interaction fingerprints were mapped to examine the expected important interactions reveal by SDM. Enrichment factors (EF) were calculated at 2%, 5% and 10% of the total library screened, using the following equation:

$$EF = (\text{Hits}_{\text{sampled}} / N_{\text{sampled}}) \div (\text{Hits}_{\text{total}} / N_{\text{total}})$$

where Hits is the number of actives and N is the number of compounds (actives and decoys).

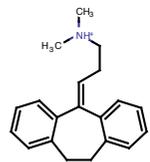
ROC curves were plotted as the true positive rate against the false positive rate (Triballeau, Acher, Brabet, Pin, & Bertrand, 2005). The Boltzmann-enhanced discrimination of receiver operating characteristic (BEDROC) (Truchon & Bayly, 2007) and ROC area under the curve (AUC) values were calculated to measure the early recognition performance and the overall predictive performance, respectively, of each model.

To assess the abilities of the models to differentiate agonists from antagonists, additional docking studies were carried out on a set of 50 reported antagonists (Figure 3.12), which were retrieved, downloaded and prepared as described earlier for the agonist set. The Mann-Whitney  $U$  test was used to compare the Gscore values obtained from docking studies on agonists and antagonists, with  $p < 0.01$  being considered to be significant.

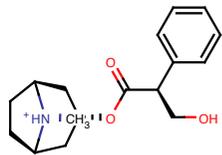
**Table 3.3.** Distribution of molecular properties and Tanimoto similarity scores for agonists and decoys.

Library / Property	ALogP			Molecular Weight			Fractional polar surface area		
	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD
Agonists	1.3	1.0	2.0	241.2	206.3	97.5	0.2	0.2	0.1
Set II (Schrödinger)	2.5	2.6	1.7	359.6	358.4	87.9	0.3	0.3	0.1
Set III (property-matched)	1.9	1.9	1.6	241.4	216.3	88.4	0.3	0.2	0.1
Library / Property	Hydrogen bond acceptor			Hydrogen bond donor			Rotatable bonds		
	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD
Agonists	3.2	3.0	1.5	0.4	0.0	0.6	3.5	3.0	2.3
Set II (Schrödinger)	4.7	4.0	2.1	1.7	2.0	1.1	5.2	5.0	2.9
Set III (property-matched)	3.2	3.0	1.3	0.6	0.0	0.8	3.4	3.0	2.1
Library / Property	Number of aromatic rings			Number of rings			Tanimoto similarity score*		
	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD
Agonists	0.8	1.0	0.8	2.1	2.0	1.4	0.14	0.10	0.16
Set II (Schrödinger)	2.5	3.0	1.1	3.2	3.0	1.2	0.13	0.11	0.06
Set III (property-matched)	1.4	1.0	1.1	2.1	2.0	1.4	0.13	0.11	0.06

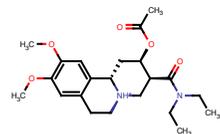
\*Tanimoto similarity scores take values from 0 – 1, with scores closer to 1 indicating greater similarity.



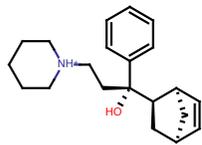
Amitriptyline



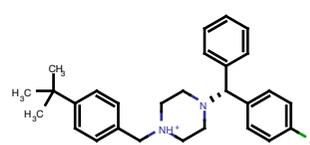
Atropine



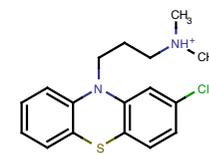
Benzquinamide



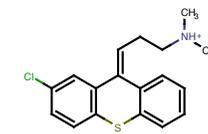
Biperidine



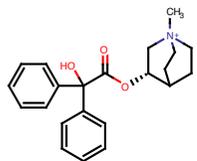
Buclizine



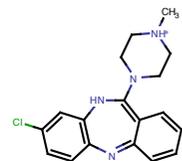
Chlorpromazine phenothiazine



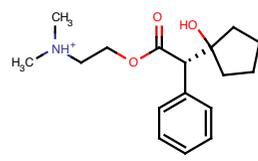
Chlorprothixene



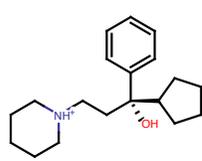
Clidinium



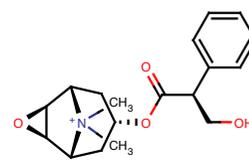
Clozapine



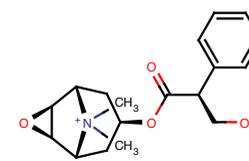
Cyclopentolate



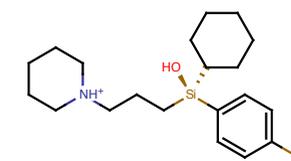
Cycrimine



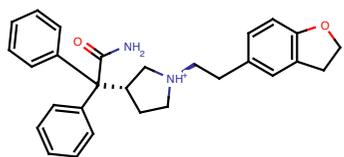
(+)-N-Methylscopolamine



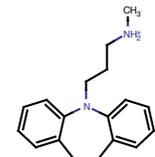
(-)-N-Methylscopolamine



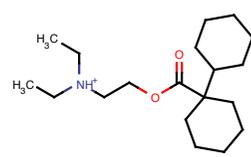
4-fluorohexahydrosiladifenidol



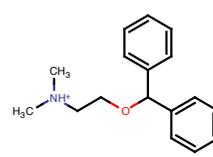
Darifenacin



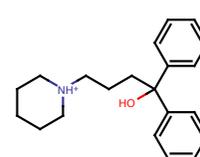
Desipramine



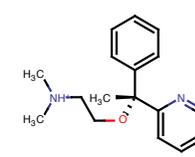
Dicyclomine



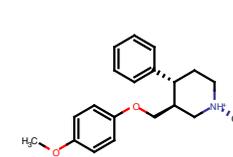
Diphenhydramine



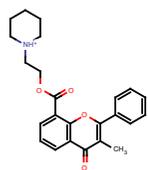
Diphenidol



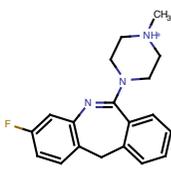
Doxylamine



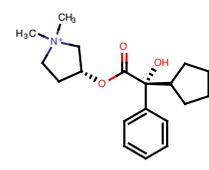
Femoxetine



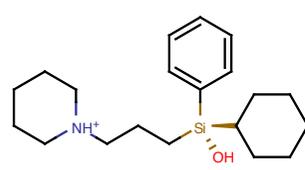
Flavoxate



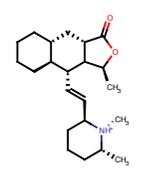
Fluperlapine



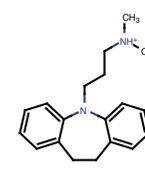
Glycopyrrolate



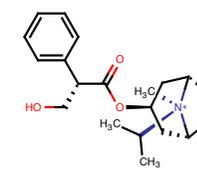
Hexahydrosiladifenidol



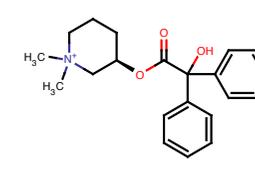
Himbacine



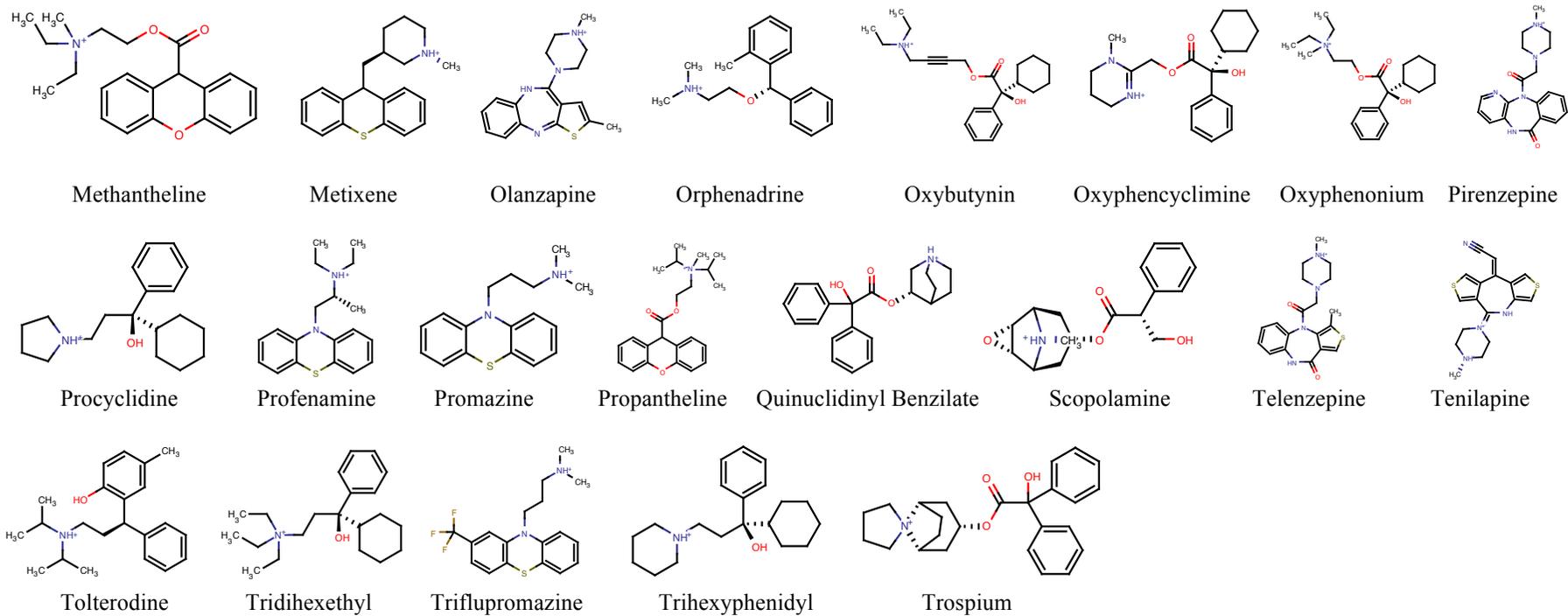
Imipramine



Ipratropium



Mepenzolic



**Figure 3.12.** M<sub>1</sub> mAChR antagonists used in the docking studies.

### 3.2.3 Results and discussion

#### 3.2.3.1 Model selection and validation

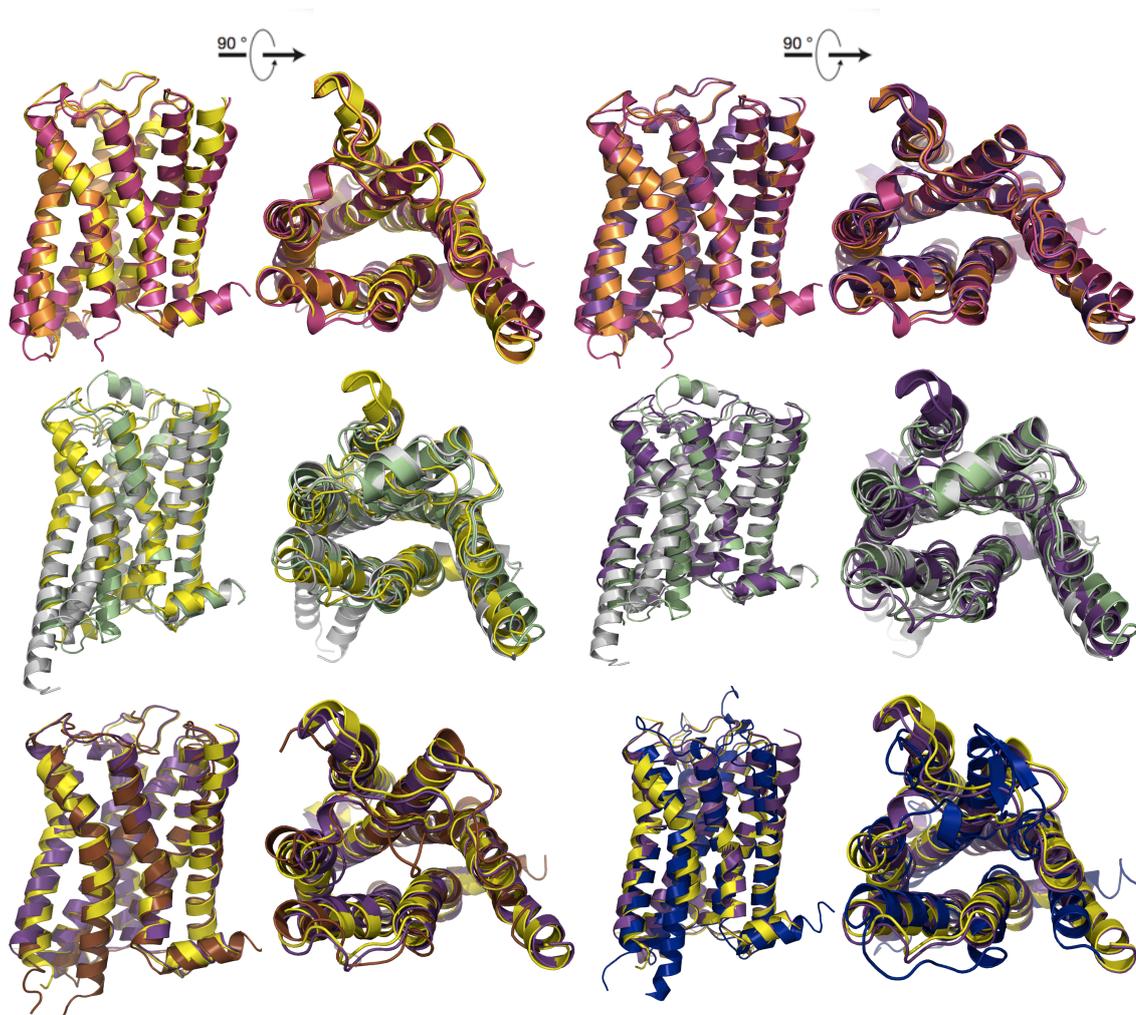
Models 1, 1A1 and 1A2 were selected from among 32 M<sub>1</sub> mAChR models (see appendix, Table A), based on their performance in the enrichment studies using decoy set I. For better and more accurate performance measurement, the selected models were further tested using decoy set II and III in the following enrichment studies. The models preserved the specific signature features of GPCRs and showed similar structural features to the M<sub>3</sub> mAChR template (PDB code: 4DAJ). Thus superposition of the generated models on the M<sub>3</sub> mAChR template, the agonist-bound  $\beta_2$ AR structure and various other solved GPCR crystal structures showed good overlap, with RMSD values in the range 0.1-2.5 Å (Table 3.4 and Figure 3.13). Models 1A1 and 1A2, which were obtained by modification using the agonist-bound crystal structure of a  $\beta_2$ AR receptor (PDB code: 3SN6) as a guide, followed by binding site refinement using IFD, possessed smaller binding pockets than model 1 (the best model obtained from IFD refinement alone) (Figure 3.11). In addition to the contraction of the binding cavity, the other prominent feature observed in agonist-bound structures is the outward movement of TM6 (Lebon, Warne, & Tate, 2012). However, models 1A1 and 1A2 did not reproduce this feature due to the nature of the induced-fit methodology used.

**Table 3.4.** RMSD values between the selected models and solved crystal structures.

Receptor (PDB code)	M <sub>2</sub> (3OUN)	M <sub>3</sub> (4DAJ)	$\beta_2$ AR (2RH1)	$\beta_2$ AR (3SN6)	H1 (3RZE)	A2A (3QAK)
Model 1	0.620	0.098	1.179	1.864	1.004	2.301
Model 1A1	0.892	0.204	1.064	1.458	1.173	2.307
Model 1A2	0.876	0.231	1.033	1.788	1.157	2.483

RMSD is given in Å.

The chosen M<sub>1</sub> receptor homology models were subjected to extensive validation analysis using PROCHECK and WHATCHECK. The results indicated that the models are reasonably good structurally with all the residues contributing substantially to ligand binding being found in the allowed regions of the Ramachandran plot (Table 3.5). Model 1 had no residues in the disallowed region and had the largest number of residues in the most favored regions. Model 1A1 had two residues (T215<sup>5.65</sup> and T366<sup>6.36</sup>) in the disallowed region, which are located near the end of TM5 and the beginning of TM6, where ICL3 was truncated. Model 1A2 had one residue (T366<sup>6.36</sup>) in the disallowed region. Neither of these residues is in the vicinity of the binding cavity of the receptor. All the models demonstrated WHATCHECK scores better or comparable with the template with no severe errors and all Z-scores being within the normal range, except for backbone conformation and inside/outside distribution (Table 3.6). The reason for the latter observation may simply be that TM helices are not very often seen in the database of solved protein structures and the database is not optimized for membrane proteins. Overall, the scores indicated that all the models have no discernable shortcomings structurally.



**Figure 3.13.** Superposition of solved 3-D GPCR structures with model 1 and model 1A2. Color code: yellow, model 1; purple, model 1A2; pink, M<sub>2</sub> (PDB code: 3UON); orange, M<sub>3</sub> (PDB code: 4DAJ); light grey,  $\beta_2$ AR (PDB code: 3SN6); light green,  $\beta_2$ AR (PDB code: 2RH1); brown, Histamine H1 (PDB code: 3RZE); dark blue, A2A adenosine (PDB code: 3QAK).

**Table 3.5.** PROCHECK and Ramachandran plot summary results for the template structure (4DAJ) and the selected models.

Model/ Structure	Ramachandran Plot (%)			
	Core	Allowed	Generous	Disallowed
4DAJ	94.6	4.8	0.5	0.0
Model 1	97.5	2.1	0.4	0.0
Model 1A1	95.9	2.9	0.4	0.8
Model 1A2	95.8	3.8	0.0	0.4

**Table 3.6.** WHATCHECK Z-scores for quality assessment and statistical analysis for the template structure (4DAJ) and the selected models.

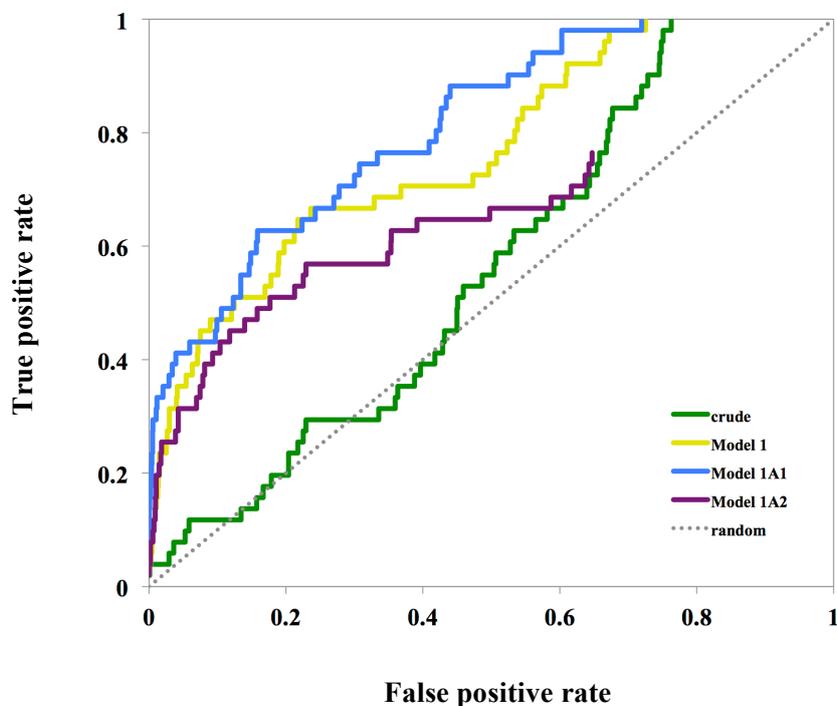
Model/ Structure	Structure Z-score						RMS Z-score			
	PQ	RPA	RN	BC	BL	BA	OAR	SCP	IDD	IOD
4DAJ	-0.5	-4.3	-2.6	-6.0	0.2	0.4	0.7	0.2	0.4	1.2
Model 1	-0.1	-2.3	-2.8	-5.8	0.7	1.1	1.1	1.1	0.9	1.2
Model 1A1	-0.5	-3.2	-2.7	-8.1	0.9	1.3	1.3	1.4	1.1	1.2
Model 1A2	-0.6	-3.5	-2.8	-9.8	0.8	1.3	1.4	1.6	1.0	1.3

PQ, second-generation packing quality; RPA, Ramachandran plot appearance; RN, chi1/chi2 rotamer normality; BC, backbone conformation; BL, bond lengths; BA, bond angles; OAR, omega angle restraints; SCP, side-chain planarity; IDD, improper dihedral distribution; and IOD, inside/outside distribution. RMS Z-value is expected to be 1.0 or close to 1.0, and Z-value above 4.0 and below -4.0 is very uncommon.

### 3.2.3.2 Enrichment and docking studies

The three refined models were found to give considerably better enrichment for 51 reported agonists than the crude model (Figure 3.14 and Table 3.7), yielding AUC values in excess of 0.70, except for model 1 in the case the Schrödinger decoy set (Set II), suggesting that they are all able to differentiate between agonists and decoys moderately accurately (Swets, 1988). Although both sets of decoys produced quite similar enrichment data, the results obtained from using Set III, were chosen for analyses, as the better match of the molecular properties between the decoys and the agonists, compared to Set II, is expected to give unbiased results and a good reflection of the actual performance of the models. The fact that the models produced similar enrichment values to that of the nonproperty-matched decoys imply that the models are good and preferring agonists even among the more challenging property-matched decoy molecules. Model 1A1 recorded the highest EF 2% and BEDROC value, which is considered to be the best indicator of the ability of a model to give early recognition of actives (Truchon & Bayly, 2007). Model 1A2 failed to bind 24% (12/51) of the larger agonists (gap at the end of curves, purple), including NCC 11-1585, NCC 11-1607 and pentylthio-TZTP (Figure 3.1), which are described as full agonists in the IUPHAR

database. This suggests that the smaller and tighter binding cavity of model 1A2 is unable to accommodate bulkier molecules.

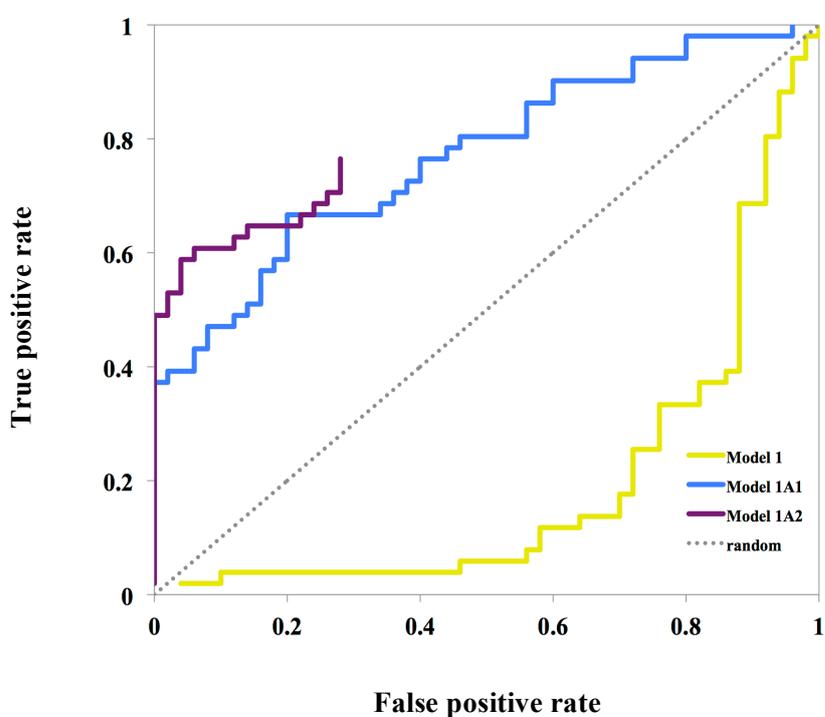


**Figure 3.14.** Receiver operating characteristic (ROC) agonist enrichment plots using the property-matched decoy set (Set III) for the different homology models.

**Table 3.7** Analysis of data from enrichment studies for the crude and refined models.

Decoy/model		Crude	1	1A1	1A2	
Set II (Schrödinger)	BEDROC	0.102	0.244	0.480	0.526	
	AUC	0.34	0.67	0.75	0.91	
	EF	2%	2.0	5.9	17	19
	EF	5%	1.6	3.5	7.8	9.4
	EF	10%	0.98	2.9	4.1	5.1
Set III (property-matched)	BEDROC	0.097	0.355	0.440	0.324	
	AUC	0.56	0.76	0.80	0.86	
	EF	2%	2.0	12	17	13
	EF	5%	1.6	7.1	8.2	6.3
	EF	10%	1.2	4.7	4.7	4.1

Docking of 50 reported antagonists resulted in 100% and 96% (48) of them being able to bind to models 1 and 1A1, respectively, whereas model 1A2 was only able to accommodate 28% (14) of the antagonists. However, it was found that the Gscore values for those antagonists that bound to model 1A1 tended to be significantly lower ( $p < 0.01$ ) than for the 51 agonists. A similar trend was observed for antagonist and agonist docking to model 1A2. Thus, on the basis of Gscore values, both models 1A1 and 1A2 exhibited a marked selectivity for agonists over antagonists, with excellent EF 2% (19 and 25, respectively) and BEDROC values (0.987 and 0.997, respectively). Conversely, in the case of model 1, the Gscore values for bound antagonists tended to be significantly higher ( $p < 0.01$ ) than for bound agonists, resulting in a reversal of the agonist/antagonist selectivity. These observations are clearly illustrated by the agonist/antagonist selectivity ROC plots shown in Figure 3.15. Taken together, these results suggest that models 1A1 and 1A2 possess near activated state character, whereas model 1 has inactive state character.



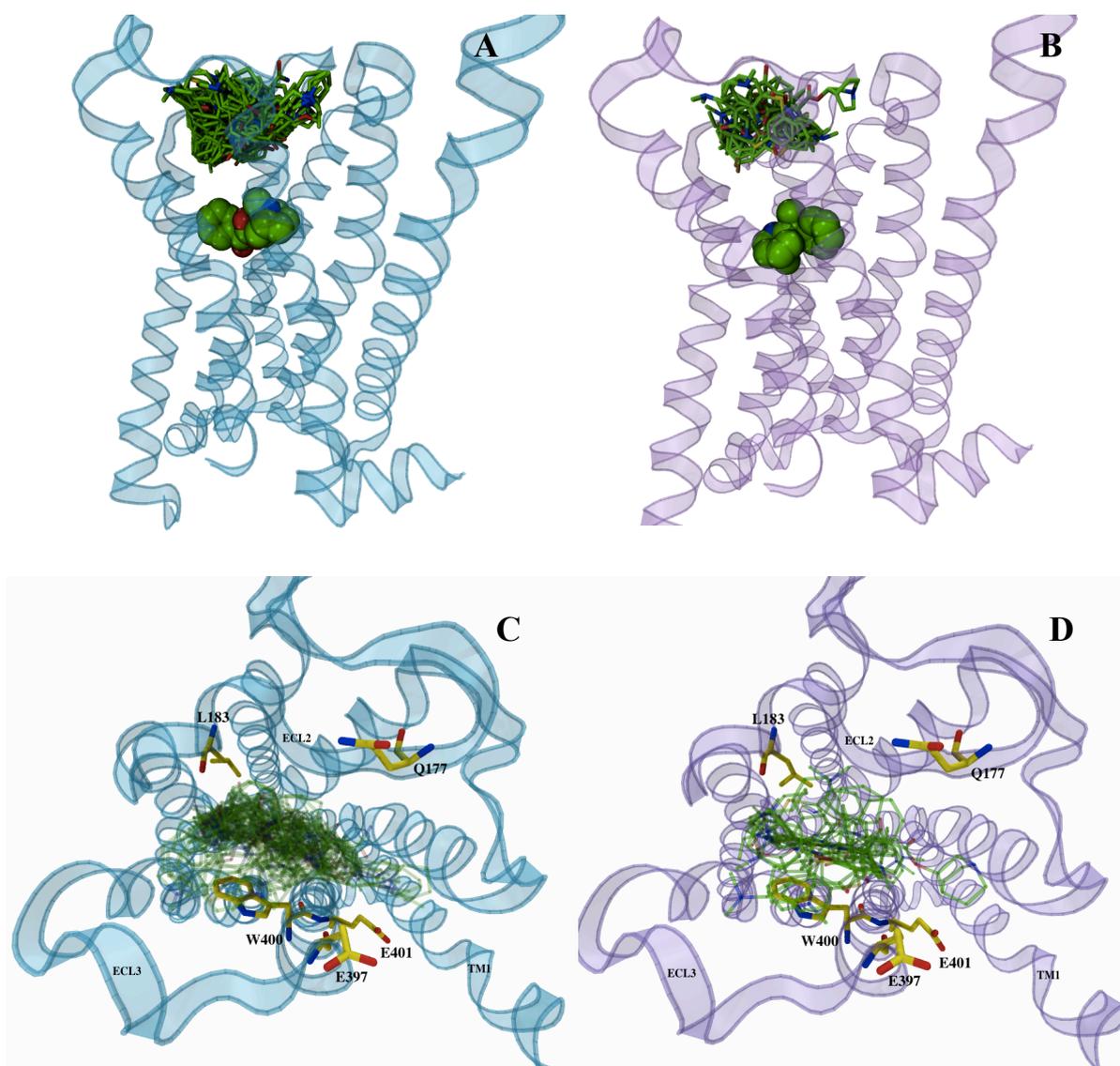
**Figure 3.15.** Receiver operating characteristic (ROC) plots showing the agonist/antagonist selectivity for the different homology models: yellow, model 1; blue, model 1A1; purple, model 1A2.

### 3.2.3.3 Analysis of ligand binding modes

Analysis of the antagonist binding locations for each of the models revealed that, whereas 88% (44/50) of the antagonists docked to the orthosteric site of model 1, in the activated models 1A1 and 1A2, nearly all of the antagonists (46/48 and 13/14, respectively) did not dock within the orthosteric site but instead were found at a site adjacent to the orthosteric site, towards the extracellular surface of the receptor and some distance away from the conserved D105<sup>3.32</sup> residue (Figure 3.16). This observation is consistent with the proposal that antagonists have low affinity for activated receptor conformations (Kenakin, 2004) and in fact may not fit within the smaller, activated orthosteric site. Furthermore, suboptimal binding at this external site may account for the lower Gscore values obtained for antagonists bound to the activated models. Interestingly, the location of this site seems to closely correspond to that of the allosteric site that has been identified in MD simulation studies of the binding of the inverse agonist, tiotropium, to M<sub>2</sub> and M<sub>3</sub> receptor models (Kruse et al., 2012), derived from their respective crystal structures. The fact that, in general, both the agonists and the antagonists bind within the orthosteric site of model 1, whereas only the agonists bind to this site in models 1A1 and 1A2, supports the proposition that model 1 resembles an inactive state, while models 1A1 and 1A2 more closely represent an activated state.

Of the two activated models, model 1A2 more strongly discriminates between agonists and antagonists. The examination of the binding orientation of agonists within this model revealed that 80% of them conformed to the expected binding modes, making interactions with the key residues identified by SDM experiments (Lu & Hulme, 1999; Ward et al., 1999; Allman et al., 2000; Lu et al., 2001; Spalding et al., 2006; Goodwin et al., 2007). This observation demonstrates that the models are not strongly biased

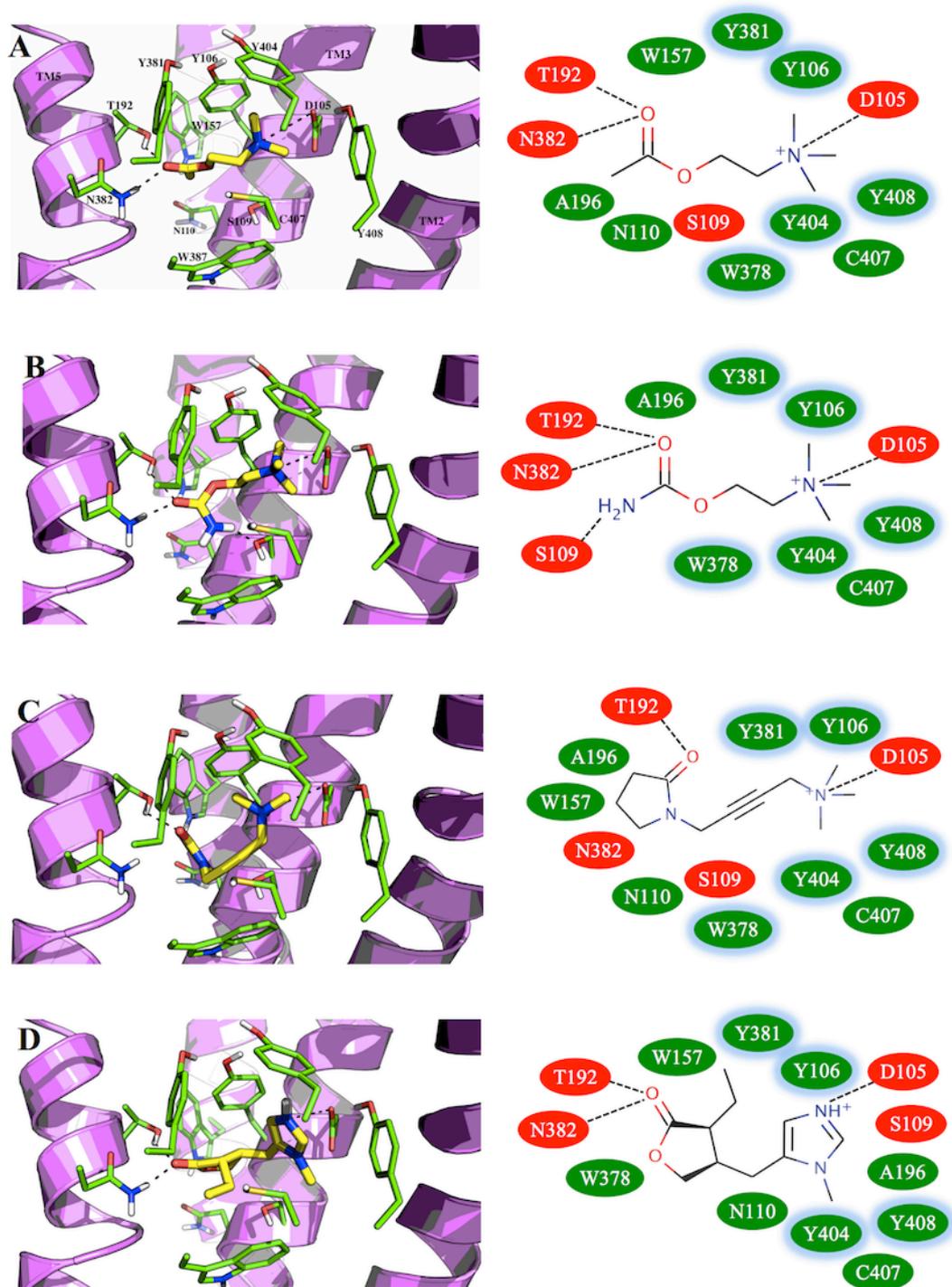
towards the structure of the agonist (ACh) that was used in the binding site refinement and are capable of correctly predicting the binding interactions of structurally diverse ligands. Similar results were also obtained for refined models 1 and 1A1.



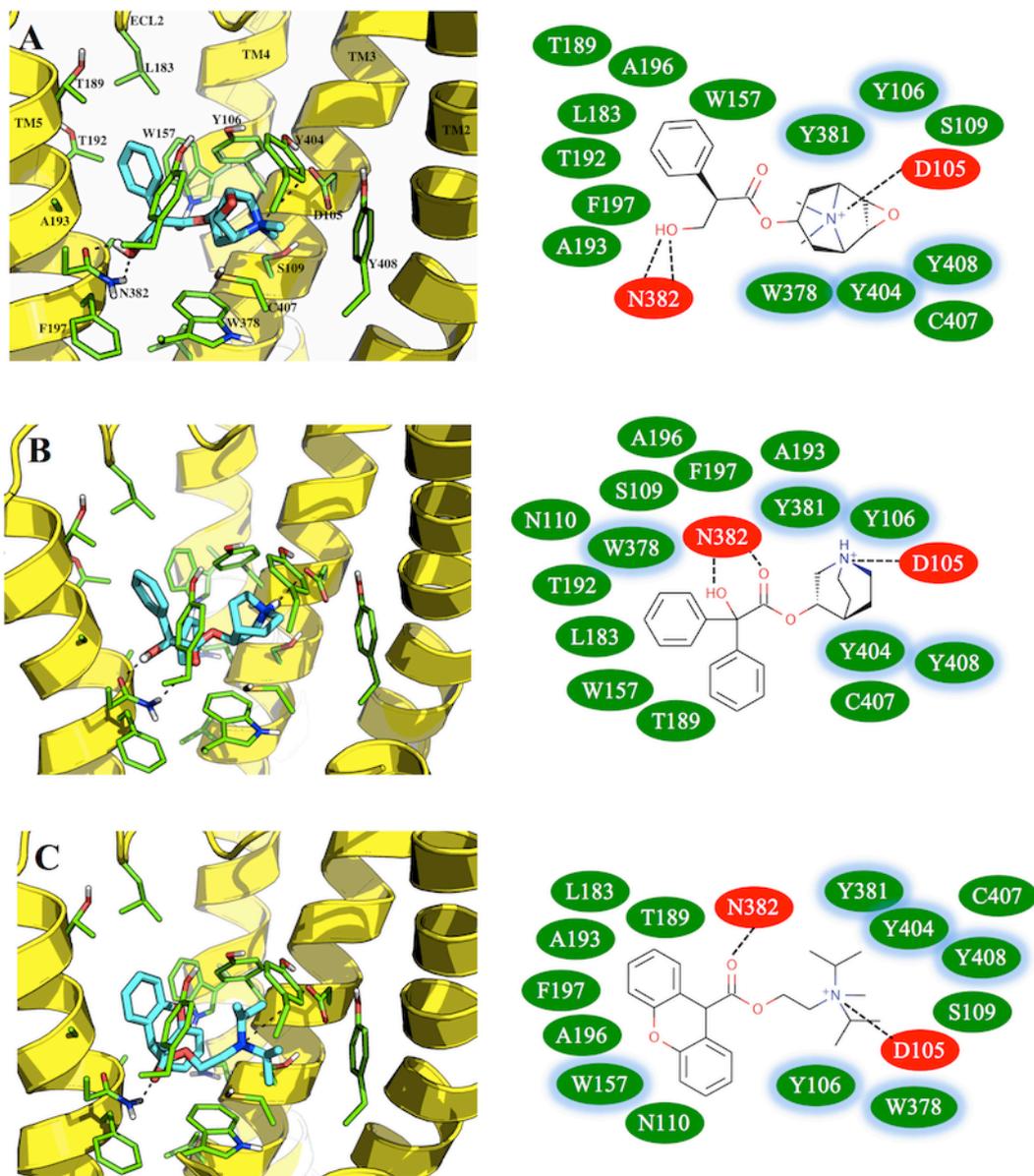
**Figure 3.16.** 3-D representations of the superimposed antagonists docked at a probable allosteric site in models 1A1 and 1A2. **A:** side view of model 1A1, with TM1 on the right and the orthosteric site shown by atropine (as spheres). **B:** side view of model 1A2, with TM1 on the right and the orthosteric site shown by doxylamine (as spheres). **C, D:** top views from the extracellular surface of models 1A1 and 1A2, respectively, with the side chains of the non-conserved interacting residues Q177 (E175 in M<sub>2</sub>), L183 (F181 in M<sub>2</sub>, L225 in M<sub>3</sub>), E397 (N419 in M<sub>2</sub>, K522 in M<sub>3</sub>), W400 (W422 in M<sub>2</sub>) and E401 (T423 in M<sub>2</sub>) shown in stick representation (yellow).

Figure 3.17 shows the interaction patterns for the representative agonists, ACh, carbachol, oxotremorine-M and pilocarpine docked with model 1A2. All the ligands form a polar charged interaction with the conserved D105<sup>3.32</sup> and cation-pi interactions with aromatic residues Y106<sup>3.33</sup>, W378<sup>6.48</sup>, Y381<sup>6.51</sup>, Y404<sup>7.39</sup> and Y408<sup>7.43</sup> through their positively-charged head groups, and hydrophobic interactions with the latter set of residues and others, including A196<sup>5.46</sup> and C407<sup>7.42</sup> (Table 3.8). The interactions within model 1A2 are in particularly close agreement with the SDM data, with hydrogen bonding between both T192<sup>5.42</sup> and N382<sup>6.52</sup> and the carbonyl oxygen atom in the agonist tail groups (Huang, Nagy, Williams, Peseckis, & Messer, 1999). Furthermore, the fact that, in all three models, Y381<sup>6.51</sup> is involved in pi-cation interactions with ACh suggests that the models contain important features of the activated state (Allman et al., 2000).

The binding modes of the representative antagonists, *N*-methylscopolamine (NMS), (-)-3-quinuclidinyl benzilate (QNB) and propantheline are shown in Figure 3.18. These compounds were found to bind to the orthosteric pocket in model 1 with the positively charged head groups interacting in a similar way to the agonists discussed above. The main differences were that the antagonists exhibited additional hydrophobic interactions with A193<sup>5.43</sup>, while hydrogen bonding interactions were only observed with N382<sup>6.52</sup> and not with T192<sup>5.42</sup> (Table 3.9). These differences are also in good agreement with SDM experiments, in which mutations to these residues were observed to have greater effects on the binding of antagonists than that of agonists (Huang et al., 1999; Allman et al., 2000). In the case of QNB, the docked binding mode reproduced that found in the crystal structure of its complex with the human M<sub>2</sub> receptor (PDB code: 3UON), in which both the carbonyl and hydroxyl groups were observed to hydrogen bond to the equivalent asparagine residue.



**Figure 3.17.** The binding modes of representative agonists with model 1A2. **A.** ACh, **B.** carbachol, **C.** oxotremorine-M, and **D.** pilocarpine. The important interacting residues are shown in stick representation and labeled. For the purpose of clarity, ECLs, ICLs, TM1, TM6, and TM7 are not shown. Residues involved in hydrogen bonding, charged, or polar interactions are shaded in red. Residues involved in van der Waals interactions are shaded in green. Residues involved in pi interactions are shown with a blue ring.



**Figure 3.18.** The binding modes of representative antagonists with model 1. **A.** NMS, **B.** QNB, and **C.** propantheline. The important interacting residues are shown in stick representation and labeled. For the purpose of clarity, ECLs, ICLs, TM1, TM6, and TM7 are not shown. Residues involved in hydrogen bonding, charged, or polar interactions are shaded in red. Residues involved in van der Waals interactions are shaded in green. Residues involved in pi interactions are shown with a blue ring.

**Table 3.8.** Structural interaction fingerprints bit-string of bound agonists in all models.

Ligands	Interacting residues												
	D105 <sup>3.32</sup>	Y106 <sup>3.33</sup>	S109 <sup>3.36</sup>	W157 <sup>4.57</sup>	T192 <sup>5.42</sup>	A193 <sup>5.43</sup>	A196 <sup>5.46</sup>	W378 <sup>6.48</sup>	Y381 <sup>6.51</sup>	N382 <sup>6.52</sup>	Y404 <sup>7.39</sup>	C407 <sup>7.42</sup>	Y408 <sup>7.43</sup>
ACh	001100	100110	000100	000100	000100	000100	000100	100110	100110	010100	100110	100100	000110
	001100	100110	000100	000100	000100	000100	100100	110100	100110	000100	100110	100100	000110
	001100	100110	000100	000100	010100	000000	100100	100110	100110	010100	100110	100100	000110
Carbachol	000100	100110	010100	000000	000000	000000	000000	100110	100110	000100	100110	100100	000100
	001100	100110	000100	000100	000000	000000	100100	100100	100110	000100	100110	100100	000110
	001100	100110	010100	000000	010100	000000	100100	100110	100110	010100	100110	100100	000110
Oxo-M	001100	100110	000100	100100	100100	100100	100100	100110	100110	000100	100110	100100	000110
	001100	100110	000100	100100	000100	000100	100100	100100	100110	000100	100110	100100	000110
	001100	100110	000100	100100	010100	000000	100100	100110	100110	000100	100110	100100	000110
Pilocarpine	011100	100110	000100	100100	000100	000100	100100	100110	100100	000100	100100	100100	000100
	011100	100110	000100	100100	100100	000100	100100	110100	100100	000100	100110	100100	000100
	001100	100110	000100	000100	010100	000000	100100	100100	100110	010100	100110	100100	000110

Hydrophobic, hydrogen bond, charge/polar, good contact, aromatic (pi-cation), aromatic (pi-pi). For each of the ligand, the first, second, and third lines of bit-string correspond to model 1, 1A1 and 1A2, respectively.

**Table 3.9.** Structural interaction fingerprints bit-string of bound agonists in comparison with antagonists.

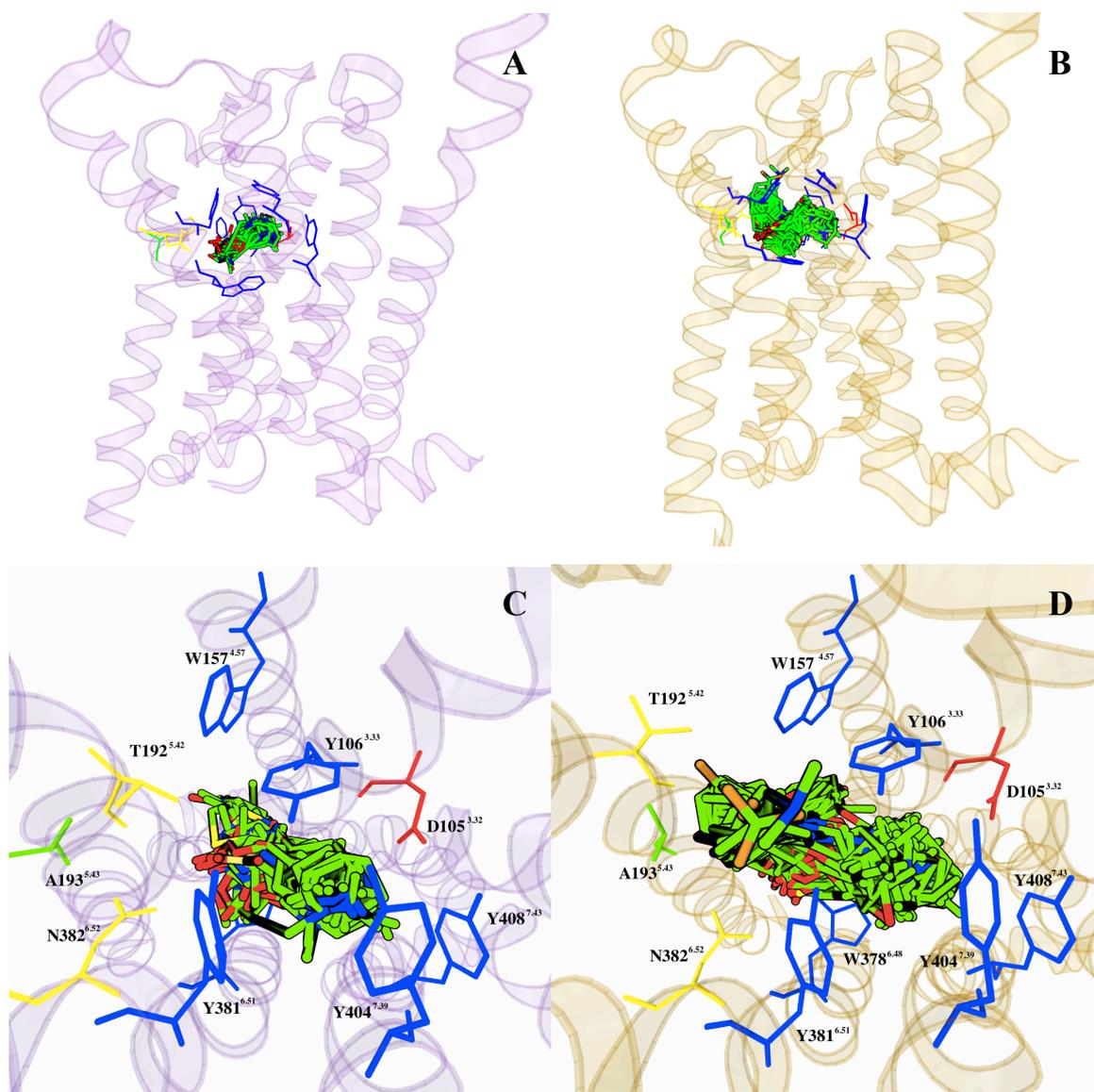
Ligands	Interacting Residues												
	D105 <sup>3.32</sup>	Y106 <sup>3.33</sup>	S109 <sup>3.36</sup>	W157 <sup>4.57</sup>	T192 <sup>5.42</sup>	A193 <sup>5.43</sup>	A196 <sup>5.46</sup>	W378 <sup>6.48</sup>	Y381 <sup>6.51</sup>	N382 <sup>6.52</sup>	Y404 <sup>7.39</sup>	C407 <sup>7.42</sup>	Y408 <sup>7.43</sup>
ACh	001100	100110	000100	000100	010100	000000	100100	100110	100110	010100	100110	100100	000110
Carbachol	001100	100110	010100	000000	010100	000000	100100	100110	100110	010100	100110	100100	000110
Oxo-M	001100	100110	000100	100100	010100	000000	100100	100110	100110	000100	100110	100100	000110
Pilocarpine	001100	100110	000100	000100	010100	000000	100100	100100	100110	010100	100110	100100	000110
NMS	001100	100110	000100	100100	100100	100100	100100	100110	100110	010100	100110	100100	000110
QNB	001100	100110	000100	100100	100100	100100	100100	100100	100110	010100	100110	100100	000110
Proprantheline	001100	100110	000100	100100	000000	100100	100100	100110	100110	010100	100110	100100	000110

Hydrophobic, hydrogen bond, charged/polar, good contact, aromatic (pi-cation), aromatic (pi-pi). The entries for the agonists are from model 1A2, whereas the entries for the antagonist are from model 1.

Superposition of all the ligands that successfully docked to the refined models, showed that the binding cavities for the agonists and antagonists overlap, but with the antagonist binding cavity extending towards the extracellular vestibule and ECL2 (Figure 3.19). Nevertheless, while the head groups for both agonists and antagonists were directed towards the conserved D105<sup>3.32</sup>, the tail groups positioned themselves differently, confirming observations made in a previous study (Goodwin et al., 2007).

Structural interaction fingerprinting (Mordalski, Kosciolk, Kristiansen, Sylte, & Bojarski, 2011) was used to further investigate the binding of agonists and antagonists to models 1 and 1A2 (Table 3.10). The interactions map was able to distinguish between the agonist and antagonist binding patterns and confirmed the importance of T192<sup>5.42</sup> and N382<sup>6.52</sup> in agonist and antagonist binding, respectively. Even though T192<sup>5.42</sup> was found within 4 Å of 76% of the docked antagonists, none of them established a hydrogen bond with this residue, in contrast with 18% of the agonists. On the other hand, almost half of the antagonists were found to form hydrogen bonds with N382<sup>6.52</sup>, whereas this interaction does not seem to be mandatory for the agonists. Furthermore, A193<sup>5.43</sup> was completely absent from agonist binding in model 1A2 and relatively unimportant in model 1, but contributed hydrophobic interactions with almost all the antagonists. L183, which extends downward from ECL2 towards the binding cavity, was observed as part of the binding pocket for 84% of the antagonists but only 18% for the agonists, while another ECL2 residue, I180, was found to be involved in the binding of antagonists but was absent from agonist binding. These residues emphasize the role of ECL2 in ligand binding, especially for the relatively larger antagonists. Overall, the results show that the binding modes of agonists are different to those of antagonists. This is in agreement with observations made on agonist-bound crystal structures of  $\beta_1$ AR and  $\beta_2$ AR receptors (Rasmussen, Choi, et al., 2011; Warne et

al., 2011).



**Figure 3.19.** 3-D representations of the superimposed ligands that successfully docked at the orthosteric site of models 1A2 and 1. **A:** side view of the agonists docked to model 1A2, with TM1 on the right. **B:** side view of the antagonists docked to model 1, with TM1 on the right. **C,** **D:** top views from the extracellular surface of models 1A2 and 1. The side chains of the important interacting residues are shown in stick representation and colored according to the type of interactions (Hydrophobic, hydrogen bond, charged/polar, aromatic).

**Table 3.10.** An averaged structural interactions fingerprint calculated over all successfully docked poses of agonists and antagonists with the interacting residues in the TM domain.

Residues	Interactions					
	Good Contact	Hydrophobic	H-bond acceptor	H-bond donor	Aromatic	Charged
D105 <sup>3.32</sup>	<b>0.83</b>	<b>0.0</b>	<b>0.13</b>	<b>0.0</b>	<b>0.0</b>	<b>0.80</b>
	<i>0.88</i>	<i>0.0</i>	<i>0.0</i>	<i>0.0</i>	<i>0.0</i>	<i>0.82</i>
	0.92	0.0	0.08	0.0	0.0	0.84
Y106 <sup>3.33</sup>	<b>0.83</b>	<b>0.63</b>	<b>0.03</b>	<b>0.0</b>	<b>0.68</b>	<b>0.0</b>
	<i>1.0</i>	<i>0.92</i>	<i>0.02</i>	<i>0.0</i>	<i>0.80</i>	<i>0.0</i>
	1.0	0.71	0.02	0.0	0.76	0.0
S109 <sup>3.36</sup>	<b>0.83</b>	<b>0.0</b>	<b>0.08</b>	<b>0.0</b>	<b>0.0</b>	<b>0.0</b>
	<i>0.88</i>	<i>0.0</i>	<i>0.12</i>	<i>0.0</i>	<i>0.0</i>	<i>0.0</i>
	0.96	0.0	0.22	0.0	0.0	0.0
W157 <sup>4.57</sup>	<b>0.75</b>	<b>0.33</b>	<b>0.0</b>	<b>0.0</b>	<b>0.0</b>	<b>0.0</b>
	<i>0.96</i>	<i>0.86</i>	<i>0.0</i>	<i>0.0</i>	<i>0.06</i>	<i>0.0</i>
	0.84	0.49	0.0	0.0	0.20	0.0
<b>T192<sup>5.42</sup></b>	<b>0.83</b>	<b>0.0</b>	<b>0.0</b>	<b>0.18</b>	<b>0.0</b>	<b>0.0</b>
	<i>0.76</i>	<i>0.44</i>	<i>0.0</i>	<i>0.0</i>	<i>0.0</i>	<i>0.0</i>
	0.25	0.18	0.0	0.0	0.0	0.0
<b>A193<sup>5.43</sup></b>	<b>0.0</b>	<b>0.0</b>	<b>0.0</b>	<b>0.0</b>	<b>0.0</b>	<b>0.0</b>
	<i>0.94</i>	<i>0.72</i>	<i>0.0</i>	<i>0.0</i>	<i>0.0</i>	<i>0.0</i>
	0.49	0.22	0.0	0.0	0.0	0.0
A196 <sup>5.46</sup>	<b>0.83</b>	<b>0.53</b>	<b>0.0</b>	<b>0.0</b>	<b>0.0</b>	<b>0.0</b>
	<i>0.90</i>	<i>0.80</i>	<i>0.0</i>	<i>0.0</i>	<i>0.0</i>	<i>0.0</i>
	0.69	0.31	0.0	0.0	0.0	0.0
W378 <sup>6.48</sup>	<b>0.68</b>	<b>0.38</b>	<b>0.0</b>	<b>0.0</b>	<b>0.48</b>	<b>0.0</b>
	<i>0.90</i>	<i>0.90</i>	<i>0.0</i>	<i>0.0</i>	<i>0.68</i>	<i>0.0</i>
	0.96	0.73	0.0	0.0	0.67	0.0
Y381 <sup>6.51</sup>	<b>0.95</b>	<b>0.45</b>	<b>0.0</b>	<b>0.0</b>	<b>0.58</b>	<b>0.0</b>
	<i>0.98</i>	<i>0.92</i>	<i>0.04</i>	<i>0.0</i>	<i>0.70</i>	<i>0.0</i>
	0.92	0.71	0.02	0.0	0.33	0.0
<b>N382<sup>6.52</sup></b>	<b>0.78</b>	<b>0.0</b>	<b>0.0</b>	<b>0.28</b>	<b>0.0</b>	<b>0.0</b>
	<i>0.78</i>	<i>0.0</i>	<i>0.0</i>	<i>0.34</i>	<i>0.0</i>	<i>0.0</i>
	0.53	0.0	0.0	0.06	0.0	0.0
Y404 <sup>7.39</sup>	<b>0.93</b>	<b>0.33</b>	<b>0.0</b>	<b>0.0</b>	<b>0.8</b>	<b>0.0</b>
	<i>1.0</i>	<i>0.54</i>	<i>0.0</i>	<i>0.0</i>	<i>0.42</i>	<i>0.0</i>
	1.0	0.41	0.0	0.0	0.27	0.0
C407 <sup>7.42</sup>	<b>0.83</b>	<b>0.35</b>	<b>0.0</b>	<b>0.0</b>	<b>0.0</b>	<b>0.0</b>
	<i>0.88</i>	<i>0.50</i>	<i>0.0</i>	<i>0.0</i>	<i>0.0</i>	<i>0.0</i>
	0.93	0.41	0.0	0.0	0.0	0.0
Y408 <sup>7.43</sup>	<b>0.73</b>	<b>0.08</b>	<b>0.0</b>	<b>0.0</b>	<b>0.73</b>	<b>0.0</b>
	<i>0.72</i>	<i>0.30</i>	<i>0.0</i>	<i>0.0</i>	<i>0.72</i>	<i>0.0</i>
	0.55	0.10	0.0	0.0	0.55	0.0

Upper row, red font: agonist binding in model 1A2; middle row, italics font: antagonist binding in model 1; bottom row, blue font: agonist binding in model 1. Good contact defined the residues found within 4 Å from the bound ligands. The averaged structural interactions fingerprint is calculated by dividing the total number of ligand in contact with a particular residue with the total number of successfully docked ligands.

#### 3.2.3.4 Challenges and limitations

Although the success of homology modeling highly depend on the sequence similarity, where only the closest phylogenetic relative is chosen, GPCR homology modeling often excluded from such rules, as they share a signature structural similarity among the family members – the 7 TM helices. The other reason for such exception arises because of the relatively lesser choice of templates in modeling GPCR. GPCR Dock 2010 Assessment states that only closest related templates produce satisfying outcome, and it is true in our attempt to model M<sub>1</sub> receptor using M<sub>3</sub> receptor crystal structure as template, compared to  $\beta_2$ AR based modeling. For better comparison, same methods were also employed to generate M<sub>1</sub> models based on other available class A GPCR crystal structures (including  $\beta_2$ AR) and tested with the enrichment studies. However, none of the models outperform model 1, 1A1 and 1A2, including the models based on 3SN6, which was used as a guide to generated the model 1A1 and 1A2 (see appendix, Table B).

The sequence identity between M<sub>1</sub> and M<sub>3</sub> receptors are 53% overall and 79% within the TM domain, vs. ~30% overall and 35% within the TM domain, between M<sub>1</sub> and  $\beta_2$ AR. While  $\beta_2$ AR and M<sub>1</sub> share overall structural architecture (just as other GPCR family members) and conserved motif within the TM domain, some of the residues such as W157<sup>4.57</sup>, Y404<sup>7.39</sup> and N382<sup>6.52</sup>, which is conserved among the muscarinic subtypes, are not retained in  $\beta_2$ AR. Replacing a polar aliphatic serine in  $\beta_2$ AR with a bulkier hydrophobic tryptophan is challenging and could ended with the indole ring of W157<sup>4.57</sup> facing outward, away from the inner TM channel. This is also observed in Lu et al., 2001, which was later remodeled to project the indole ring of W157<sup>4.57</sup> inward, toward the binding cavity in Goodwin et al., 2007. Similar situation occurred when hydrophobic Y404<sup>7.39</sup> with aromatic side chain in M<sub>1</sub> substitutes polar asparagine in

$\beta_2$ AR. The conserved N382<sup>6.52</sup> also replaced by phenylalanine in all catecholamine receptors. These three conserved residues contributed significantly to the ligands binding and hence their side chains orientations have definitive control on ligand binding modes.

The  $M_3$  receptor features an obvious outward bend at the extracellular end of TM4, which has not been detected in any other GPCR family member crystallized so far, suggesting the importance of this unusual feature of mAChRs in general. The model built upon  $\beta_2$ AR/rhodopsin has no clue to capture such distinct characteristic of the TM4, which might influence the binding, and the TM helices packing and conformation as a whole. Furthermore, highly conserved proline residues in GPCRs probably induce notable kinks in the TM helices and may serve important functional roles in signal transduction (Sansom & Weinstein, 2000). However, the position of the P159<sup>4.59</sup> in TM4, which is conserved among the muscarinic subtypes, is not conserved in  $\beta_2$ AR. Replacing leucine with proline resulted only a slight kink in the model.

GPCRs have diverse loops structures that connect the TM helices and differ greatly among all the GPCRs in the same class. The difficulties to model the loops including low sequence identity, inconsistent length, restricted position of ECL2 (due to the conserved disulfide bond), and substitution of smaller residues with bulkier one in the loop which resulted in the protrusion of the loop partly into the binding cavity. Apart from that, ECLs play a role in ligand binding and docking. The ECL2 of rhodopsin forms a short  $\beta$ -sheet that covers the 11-*cis* retinal, shielding the chromophore, and precludes the solvent expose. This resulted in the trouble to accurately dock the ligand into the binding site for the rhodopsin-based models (Bissantz et al., 2003). The conformation of  $M_3$  ECL2 in the crystal structure is significantly different from the

ECL2 in  $\beta_2$ AR, which  $M_3$  ECL2 has a shorter helical segment at the second half of the loop while  $\beta_2$ AR has a longer helical segment on the first half of the loop. The comparison between the crystal structures of  $M_2$  and  $M_3$  revealed that the ECLs conformations are unexpectedly preserved. The fact that the ECL2 might play a crucial part in subtype and agonist/antagonist selectivity, the loop was modeled directly using the template and confidently utilized to comprehend the receptor binding interactions.

In term of ligand binding, the quality of modeled structures based on rhodopsin is limited by the nature of rhodopsin and its covalently bound ligand, where other members of GPCR family bind to a diffusible ligand. In rhodopsin, the ligand retinal is entirely enclosed within the binding site cavity, formed between TM5 and TM6, while in the structures of the  $\beta_2$  (Cherezov et al., 2007; Rasmussen, DeVree, et al., 2011) and  $\beta_1$ AR (Warne et al., 2011) the binding sites are more open and the ligands are posed in a position forming interactions with TM3, TM5, and TM7. It has been challenging to correctly predict the binding mode of ligands, as revealed by the recently solved crystal structures of GPCRs. Although the orthosteric ligand binding site is well conserved across the family A ligands, the ligands differ considerably in their ways of interacting with the binding site (Congreve, Langmead, Mason, & Marshall, 2011). A closely related template within subtypes is definitely able to improve the quality and accuracy of the model, in the aspect of ligand binding.

The other failure of  $\beta_2$ AR-based models to successfully recognize true agonists vs. decoys also relied on the different of pocket size/volume in the agonist/antagonist bound receptor structure. The agonist bound  $\beta_2$ AR revealed a contracted binding pocket compared to more open binding pocket in antagonist bound  $\beta_2$ AR. However, the

interactions observed between agonist/antagonist with the receptor are of similar manner with only additional hydrogen bond formed between agonist and S203<sup>5.42</sup> and S207<sup>5.46</sup>. The strategy to further modify the models based on 3SN6 has proof to improve the enrichment for agonist retrieval. In early 2014, the agonist bound M<sub>2</sub> structure was made available (PDB code: 4MQS). Motivated to see the quality of M<sub>1</sub> model based on this newly release structure, M<sub>1</sub> models were generate using the same described methods in section 3.2.2 (except 3.2.2.3), and it is found that the models showed comparable performance with model 1, 1A1, and 1A2 (see section 5.3.1). The agonist bound M<sub>2</sub> portray contracted binding pocket as previously shown by the agonist bound  $\beta_2$ AR, and unexpectedly, iperoxo interactions patterns with the receptor resembled those in QNB bound in inactive M<sub>2</sub>.

After the MD simulations were carried out (see chapter 4), snapshots were extracted from the simulations trajectories to test the performance of the models using enrichment studies. Although some of the snapshots (especially those extracted from the MD simulations of M<sub>1</sub> models in complex with ACh) obtained better BEDROC and EF compared to the crude model (model without IFD), they do not outperform the models after IFD (models 1, 1A1, and 1A2) (see appendix, Table C). Despite MD simulations have been used to improve performance of model/crystal structure in the virtual screening against GPCRs (Tarcsey et al., 2013), in the case of muscarinic receptors modeling, this results emphasize that the used of IFD to refine the binding site is of fundamental importance in improving model quality and effectiveness in the virtual screening, as shown by many others modeling works involving GPCRs (McRobb et al., 2010; Vilar et al., 2011; Kołaczkowski, Bucki, Feder, & Pawłowski, 2013; Pala et al., 2013; Thomas et al., 2014). The resulting models from the binding site refinement seems to be more robust and do not favor only the ligand used during the IFD,

compared to the MD snapshots. The binding pocket of the receptor in complex with ACh was specific and mould according to the ACh throughout the simulations, reduced their capability to efficiently recognized other agonists used in the enrichment studies and to rank them better than the decoys. In contrast, expansion of the binding pocket of the snapshots from the apo simulations was observed in the MD simulations. The initial volumes of model 1 and 1A2 in the apo simulations are 655.92 Å<sup>3</sup> and 300.61 Å<sup>3</sup>, respectively. Throughout the simulations, the values increased and fluctuate around 815.34 Å<sup>3</sup> and 747.40 Å<sup>3</sup> for model 1 and 1A2, respectively.

### 3.2.4 Conclusion

GPCR models that can correctly identify active compounds and furthermore discriminate agonists from antagonists are important tools for rational drug design. Two main complementary approaches currently used for structure-based drug design against GPCRs are virtual screening and MD (Yuriev, Agostino, & Ramsland, 2011; Yuriev & Ramsland, 2013). In the context of muscarinic acetylcholine receptors, recent MD simulations of M<sub>2</sub> and M<sub>3</sub> receptors have provided enlightening insights into the binding pathways of agonist and antagonist ligands and have demonstrated the influence of binding site volume on the selectivity profiles of ligands of varying size (Kruse et al., 2012; Miao, Nichols, Gasper, Metzger, & McCammon, 2013).

In this study, refined human M<sub>1</sub> mAChR homology models were developed based on a crystal structure of an M<sub>3</sub> mAChR, which efficiently differentiate agonists from decoy molecules and exhibit high agonist/antagonist selectivity in docking studies. Model 1A2, in which the orientation of the T192<sup>5.42</sup> side chain was adjusted and the whole TM5 was rotated and tilted towards the inner channel of the TM region, gives the best

prediction of the binding poses and interactions of small to medium-sized agonists, although agonists that are larger or bulkier are unable to dock into its relatively tight binding pocket. On the other hand, models 1 and 1A1 are both able to bind the whole range of agonists, with model 1A1 giving good selectivity for agonists over antagonists and model 1 showing a high preference for antagonists over agonists. Since a single GPCR model is unlikely to be sufficient to completely represent the range of 'ligandable' conformations, the combined use of models of 1, 1A1 and 1A2 for structure-based virtual screening to identify potential M<sub>1</sub> mAChR agonists and differentiate them from antagonists is proposed. MD simulations are also required to investigate further the conformational changes that occur upon ligand binding and the factors that influence subtype selectivity.

## CHAPTER 4

### **EXPLORING THE DYNAMICS PROFILE OF mAChRs: A MEMBRANE BASED MOLECULAR DYNAMICS STUDY**

#### **4.1 Introduction**

Recently, crystal structures of M<sub>2</sub> and M<sub>3</sub> mAChRs were reported (Haga et al., 2012; Kruse et al., 2012). These crystal structures maintain high structural similarity between the subtypes, including the non-conserved intracellular and extracellular loops. The crystal structures also possess unique structural topology including a relatively large extracellular vestibule and a profound outward bend of the TM4 extracellular end, as compared to other GPCR crystal structures. Despite the structural conservation, the crystal structures are able to reveal important structural divergences between the M<sub>2</sub> and M<sub>3</sub> subtypes (Kruse et al., 2012). Experimental X-ray crystal structures are very helpful in understanding the fundamentals of structure function relationships, but such static structures have no dynamic information. Molecular dynamics (MD) simulations are able to track the time-dependent positions of all atoms in the system and describe the dynamics characteristics of a protein structure.

Various research groups have carried out MD simulations of mAChR (both homology models and experimental structures) in order to gain different insights such as influence of different environments on M<sub>2</sub> receptor model (Jöhren & Höltje, 2005), the conformational space of acetylcholine in M<sub>1</sub>, M<sub>2</sub>, and M<sub>5</sub> receptor models (Vistoli, Pedretti, Testa, & Matucci, 2007), ligand-receptor interactions – M<sub>3</sub> receptor model

with antagonist *N*-methylscopolamine (Martinez-Archundia, Cordomi, Garriga, & Perez, 2012), M<sub>5</sub> receptor model with antagonists SVT-40776 and solifenacin (Huang, Zheng, & Zhan, 2012), M<sub>1</sub> receptor model with acetylcholine (Espinoza-Fonseca, Pedretti, & Vistoli, 2008), ligand entry and dissociation pathways – tiotropium on M<sub>2</sub> and M<sub>3</sub> receptors (Kruse et al., 2012), M<sub>2</sub> receptor activation (Miao, Nichols, Gasper, Metzger, & McCammon, 2013), and allosteric modulation of M<sub>2</sub> receptor (Dror et al., 2013).

In the present work, MD simulations on M<sub>1</sub>, M<sub>2</sub> and M<sub>3</sub> receptor structures in apo and ligand-bound forms were performed in order to gain understanding of the dynamics and stability of these different systems and to provide structural information on ligand binding and selectivity. The ligands chosen were the native agonist, acetylcholine (ACh), and the classic antagonist, 3-quinuclidinyl benzilate (QNB), whose activities have been extensively investigated experimentally, particularly in site-directed mutagenesis (SDM) studies. Collective motions, hydrogen bond networks, free energy of binding and per residue decomposition, pocket formation were analysed and compared between receptor subtypes and forms (apo and holo). Several effects associated with the ligand binding were noticed, including suppression of overall receptor dynamics, changes in the binding pocket size, reorganization of the hydrogen bonding networks, and functionally relevant mobility of a number of the TM helices. Pockets formed at the extracellular vestibule in the simulations of receptors in complex with ligands emphasized the possibility of allosteric modulations in muscarinic receptors with subtle differences in the residues involved in forming these pockets for the different receptor subtypes suggesting potential for subtype selectivity.

## 4.2 Methods

### 4.2.1 Receptor structures

Homology models of the M<sub>1</sub> mAChR were generated from the recently-solved crystal structure of an M<sub>3</sub> muscarinic acetylcholine receptor (PDB code: 4DAJ) as the template, by carrying out induced-fit docking of acetylcholine, followed by further modifications using the active crystal structure of a  $\beta_2$ AR receptor (PDB code: 3SN6) as a guide, as previously described in Chapter 3 (section 3.2.2). Two homology models (chosen from those described in Chapter 3), M<sub>1</sub>, representing an inactive structure (model 1), and M<sub>1</sub>\*, representing an agonist bound structure (model 1A2), were included in this study. M<sub>2</sub> and M<sub>3</sub> mAChR structures were obtained from their respective crystal structures (PDB codes: 3UON and 4DAJ).

### 4.2.2 Molecular docking

Prior to MD simulations, receptor-ligand complexes were obtained by docking ACh and QNB into the receptor models using the Schrödinger suite 2011 (Schrödinger LLC, New York, USA), with the exception of the M<sub>2</sub>-QNB structure which was obtained directly from the crystal structure. Since the contracted binding pocket of model 1A2 was unable to accommodate QNB, this complex could not be modeled. The Ligprep and protein preparation wizard modules were used to prepare the ligands and the receptor models, respectively. Glide v5.7 (Schrödinger LLC) was then used to dock the ligand into the active site enclosed by a grid box, accommodating ligands with a length of 20 Å or less and with a default inner box of 10 Å on each side.

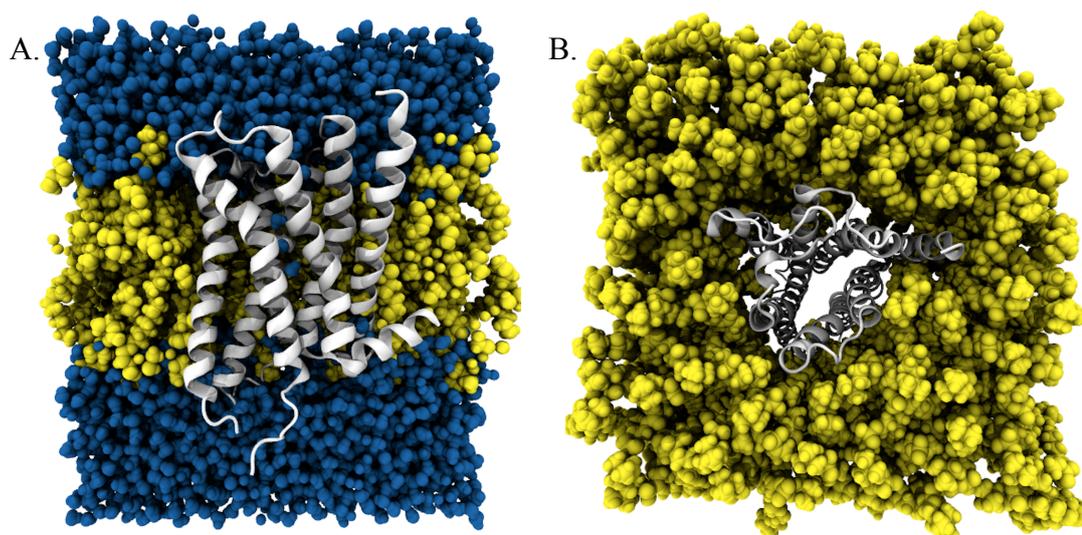
### 4.2.3 System setup

The apo and holo forms of the M<sub>2</sub> and M<sub>3</sub> receptor structures (see Table 1 for complete system description) were oriented according to their respective receptor's position in a

lipid bilayer from the Orientations of Proteins in Membranes (OPM) database (<http://opm.phar.umich.edu>) by superposition. M<sub>1</sub> receptor models built based on M<sub>3</sub> receptor were oriented according to M<sub>3</sub> receptor's position from OPM. Internal water molecules were assigned and added using Dowser (Zhang & Hermans, 1996). The complexes were then loaded into CHARMM-GUI server (Jo, Kim, & Im, 2007; Jo, Kim, Iyer, & Im, 2008) and the membrane builder tool was used to generate pre-equilibrated palmitoyl oleoyl phosphatidylcholine (POPC) lipid bilayer around the complexes. A rectangular box consisting of two layers of POPC lipids, along with a 15 Å thick layer of water molecules above and below the lipid layer, was built. An ion concentration of 0.15 M was included. All individual building parts were assembled together to produce complete systems (Figure 2) with varied total number of atoms (52000-61000), POPC molecules (138-168) and water molecules (9868-11592).

**Table 4.1.** Description of the MD systems.

Code	Simulated receptor
M <sub>1</sub> -Apo	Model 1 in empty form
M <sub>1</sub> -ACh	Model 1 in complex with ACh
M <sub>1</sub> -QNB	Model 1 in complex with QNB
M <sub>1</sub> <sup>*</sup> -Apo	Model 1A2 in empty form
M <sub>1</sub> <sup>*</sup> -ACh	Model 1A2 in complex with ACh
M <sub>2</sub> -Apo	M <sub>2</sub> crystal structure (PDB code: 3UON) in empty form
M <sub>2</sub> -ACh	M <sub>2</sub> crystal structure in complex with ACh
M <sub>2</sub> -QNB	M <sub>2</sub> crystal structure in complex with QNB
M <sub>3</sub> -Apo	M <sub>3</sub> crystal structure (PDB code: 4DAJ) in empty form
M <sub>3</sub> -ACh	M <sub>3</sub> crystal structure in complex with ACh
M <sub>3</sub> -QNB	M <sub>3</sub> crystal structure in complex with QNB



**Figure 4.1.** Overall view of the system that underwent MD simulations. **A.** Front cross section view. **B.** Top view, water molecules are not shown for the purpose of clarity. Receptor (white, cartoon representation, only apo form is shown as example) was inserted into the hydrated POPC lipid bilayer. Water and lipid molecules were shown in blue and yellow VDW representation, respectively.

#### 4.2.4 Molecular dynamics simulations

The Chamber program (Crowley, Williamson, & Walker, 2009) from AmberTools was used to convert the CHARMM psf, associated coordinated file, parameter and topology to a CHARMM force field enabled version of AMBER's prmtop and inpcrd. MD simulations were performed using PMEMD of Amber 10 (with CHARMM force field) (Case et al., 2008), with constant pressure and temperature, NPT ensemble, maintaining the pressure and temperature at 1.0 atm and 310 K, respectively, by means of anisotropic pressure scaling and Langevin dynamics. The periodic boundary conditions based on the particle mesh Ewald method with a non-bonded cutoff of 8 Å were used. The integration time step was set at 2 fs and the SHAKE algorithm was used to constrain bond lengths involving hydrogen atoms. The system underwent three stages of minimization using steepest descent and conjugate gradient with different parts of the system gradually released in stages. Then, it was slowly heated from 0 to 310 K within 150 ps with restraints on the receptor and the lipid molecules. After the heating stage,

the system was equilibrated for 3.5 ns, with weak restraints on the receptors followed by another 20 ns of a relaxed MD run. The complete system trajectory was collected every 2 ps for analysis.

#### 4.2.5 Analysis

Trajectory analyses were performed using the PTRAJ and cpptraj modules (Roe & Cheatham, 2013) of AmberTools 12. The relative binding free energy for each of the systems was evaluated using MM-PBSA approach as implemented in Amber. Per residue MM-GBSA energy decomposition was performed to calculate the contribution of each residue to the total binding energy of the complex. In Amber, only MM-GBSA is available for energy decomposition, where the electrostatic contribution to the solvation energy was calculated using the Generalized Born (GB) method instead of by solving the Poisson Boltzmann (PB) equation. Both analyses were carried out using MMPBSA.py module in Amber 12 (Miller et al., 2012). A total of 120 snapshots from the MD trajectories were included in the calculations. Area per lipid was determined using GridMAT-MD (Grid-based Membrane Analysis Tool) (Allen, Lemkul, & Bevan, 2009). Principal component analysis (PCA) was performed using Prody (Bakan, Meireles, & Bahar, 2011) and NMWiz (part of Prody) was used to generate the PCA diagrams in VMD. Pocket analysis was performed using MDpocket (Schmidtke, Bidon-Chanal, Luque, & Barril, 2011). Electrostatic charge distribution was calculated using the APBS (Baker, Sept, Joseph, Holst, & McCammon, 2001) plugin in PyMOL v1.3 (Schrödinger, LLC). Average helical rotation and tilt angle were calculated using TRAJELIX (Mezei & Filizola, 2006). Visualization and graphic images were produced using VMD (Humphrey, Dalke, & Schulten, 1996) and PyMOL v1.3. Throughout the chapter, residues are labelled using both their amino acid sequence number and

Ballesteros-Weinstein nomenclature, as a superscript extension (Ballesteros & Weinstein, 1995), for the ease of comparison and standardization.

### 4.3 Results and Discussion

#### 4.3.1 System stability and flexibility

A hydrated pre-equilibrated patch of POPC lipid bilayers was used to simulate the membrane environment of the receptors. The behaviour of the lipid bilayer is an adequate indicator to define the stability of a simulation in explicit membrane conditions (Filizola, Wang, & Weinstein, 2006). For all the systems simulated, the area of the xy plane per lipid fluctuated around  $63 \text{ \AA}^2$  (Table 4.2), which is close to the experimental value (Kucerka, Tristram-Nagle, & Nagle, 2005) and in agreement with recent MD studies (Plesnar, Subczynski, & Pasenkiewicz-Gierula, 2012; Larsson & Kasson, 2013). Analysis of secondary structure evolution throughout the MD simulation trajectories showed that  $\alpha$ -helices in the TM domains remained intact for all the systems.  $3_{10}$  helices were observed in ICL1, ICL2 and ECL2, in addition of  $\pi$ -helices in ICL2, in  $M_1$ ,  $M_1^*$  and  $M_3$  systems, interplaying with random coil and turn, while  $M_2$  in contrast showed no  $\pi$ -helices in ICL2 and  $3_{10}$  helices were also detected in ECL3 (Figures 4.2 and 4.3). The radius of gyration, which is an indicator for the compactness of receptor structures, showed that the binding of ACh and QNB to the receptor resulted in slightly more compact structures than for the apo forms, except for  $M_1$ -QNB, which was comparable with  $M_1$ -apo and  $M_2$  systems (Table 4.2 and Figure 4.4). The root mean square deviation (RMSD) values, which were computed using the backbone atoms of the TM helices, ECLs and ICLs with respect to the minimized starting structure were found to be less than  $2.5 \text{ \AA}$  for all the systems, except for ECL in  $M_1^*$ -apo (Table 4.2,

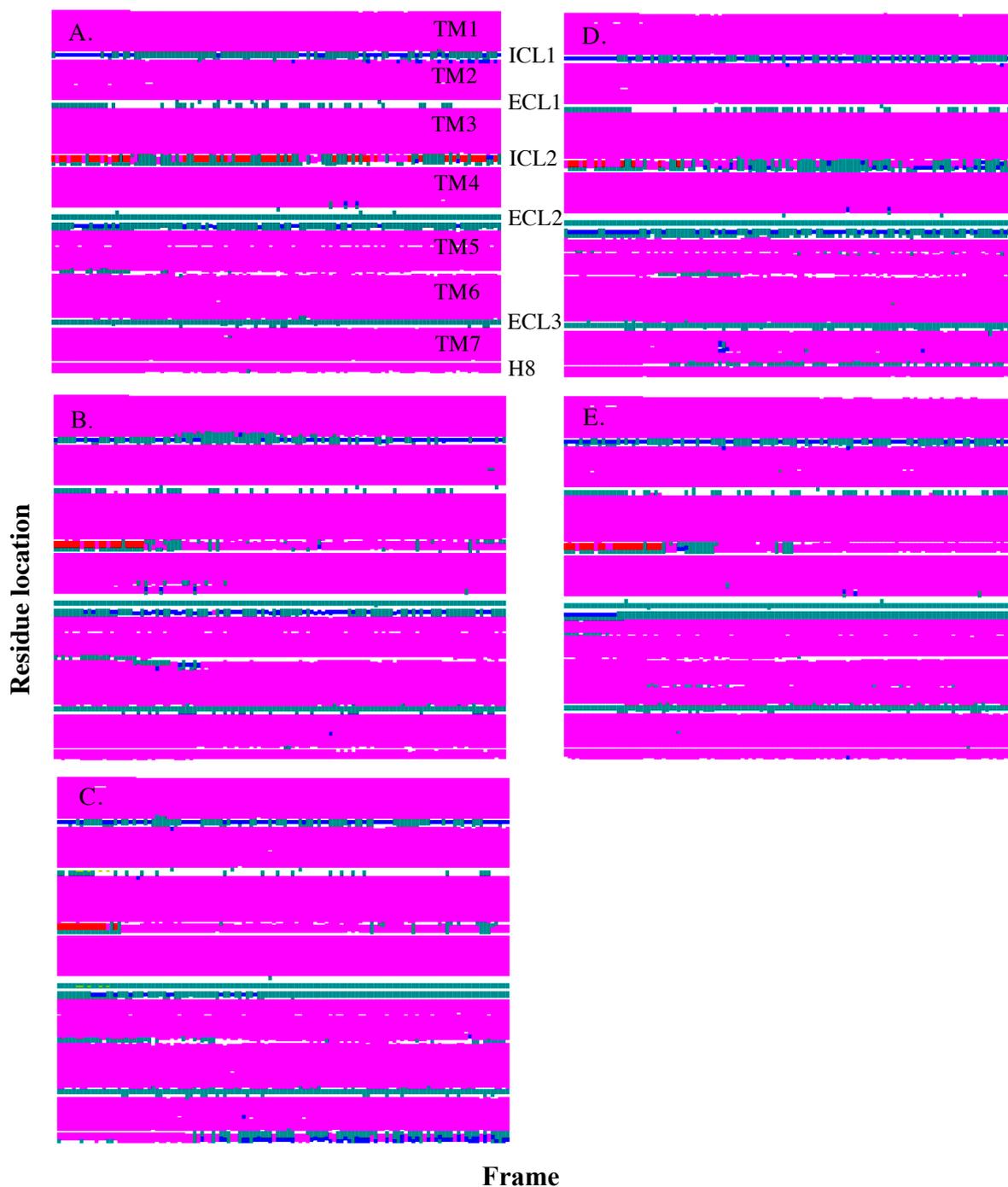
Figures 4.5 and 4.6). These findings showed that the systems were well equilibrated and do not deviate greatly from the initial starting structure.

**Table 4.2.** Average area per lipid, radius of gyration and RMSD calculated over the trajectories for the different systems studied.

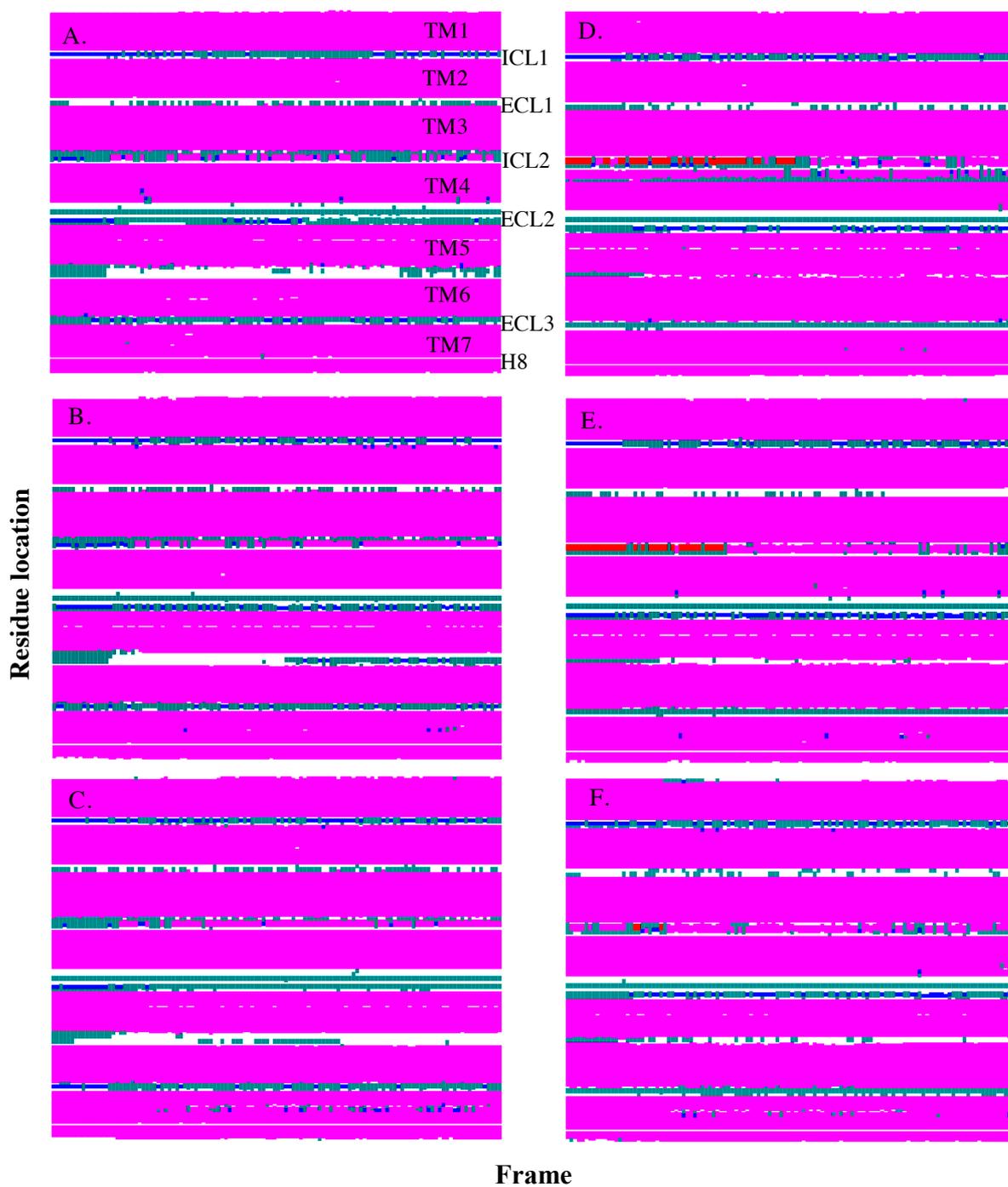
System	Area per lipid ( $\text{\AA}^2$ )	Radius of gyration ( $\text{\AA}$ )	RMSD ( $\text{\AA}$ )		
			TM	ECL	ICL
M <sub>1</sub> -Apo	62.9	19.0	1.6	1.7	1.4
M <sub>1</sub> -ACh	62.5	18.9	1.9	1.5	1.5
M <sub>1</sub> -QNB	62.7	19.0	2.0	1.2	1.8
M <sub>1</sub> <sup>*</sup> -Apo	63.7	18.9	2.3	2.6	1.6
M <sub>1</sub> <sup>*</sup> -ACh	63.9	18.6	1.9	2.2	1.6
M <sub>2</sub> -Apo	63.7	18.7	1.4	1.5	1.2
M <sub>2</sub> -ACh	63.7	18.7	1.5	1.9	1.2
M <sub>2</sub> -QNB	63.6	18.7	1.3	1.3	1.0
M <sub>3</sub> -Apo	63.5	19.1	1.9	1.1	1.9
M <sub>3</sub> -ACh	63.1	18.9	1.3	1.2	1.4
M <sub>3</sub> -QNB	62.6	18.9	1.4	1.4	1.1

The relative fluctuations of individual residue were quantified through calculation of the root mean square fluctuation (RMSF) for each residue. Expectedly, the RMSF of the backbone atoms of each of the receptors simulated after fitting to their individual starting structures showed high mobility in the loop regions (Figure 4.7). However, as a result of two stabilizing disulfide bonds, which limit the movement of the loops, ECL2 were found to be less mobile. The flexibility of the loops regions also influences the mobility of the extracellular and intracellular ends of the TM helices, which connected via alternating ICLs and ECLs. This is particularly apparent in the case of TM5 and TM6, for which the connecting loop region (ICL3) was not modeled. The flexibility plots were in agreement with the RMSD values, where ECLs and ICLs showed bigger fluctuations than TM domains, even though the deviations were small (Figures 4.5 and 4.6). For all the systems, the presence of ACh and QNB reduced the flexibility of the

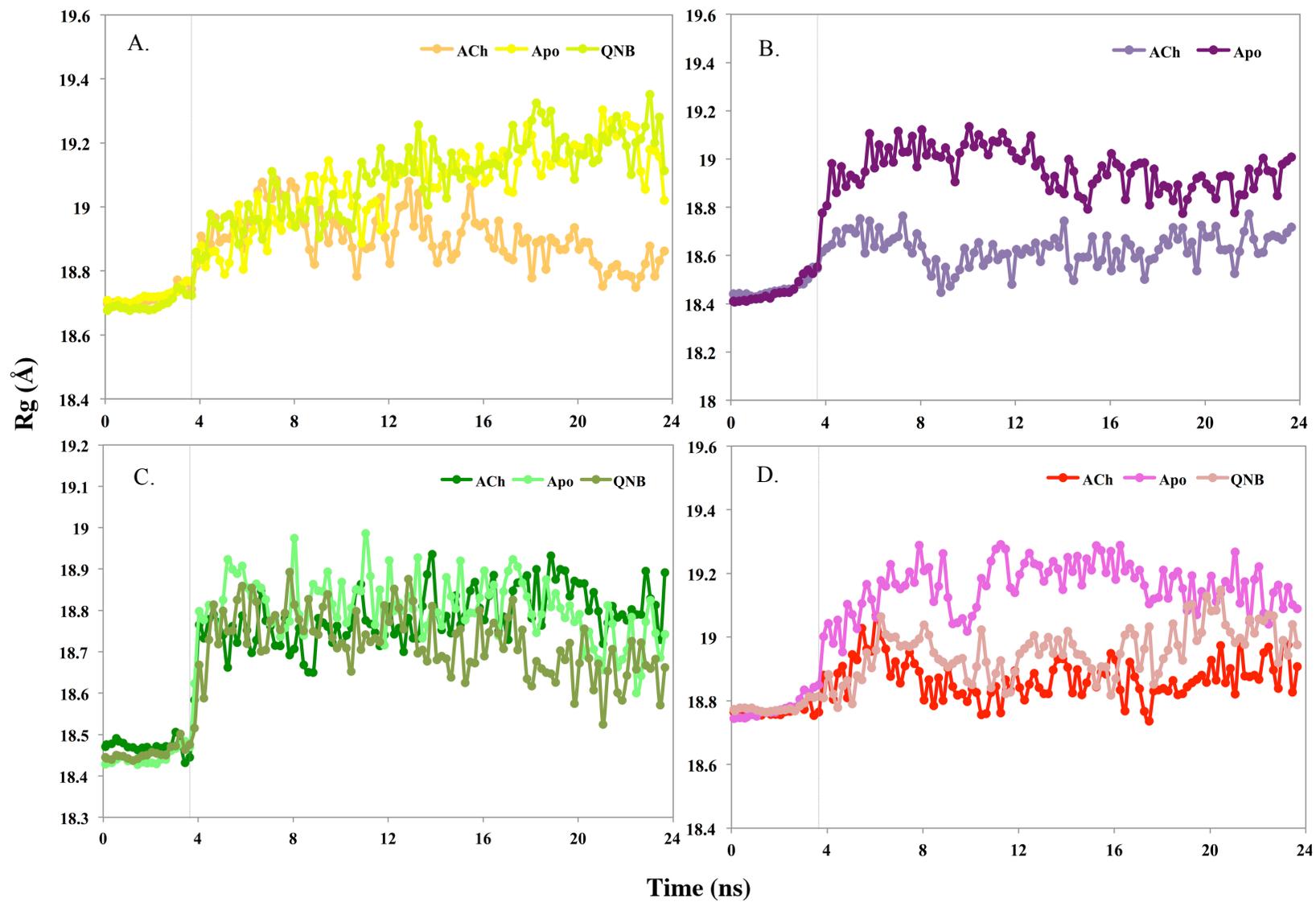
receptors. However, TM6 of M<sub>1</sub>-ACh, and ICL2 of M<sub>2</sub>-ACh and M<sub>2</sub>-QNB, showed greater mobility than the apo form. The RMSD plots also depicted smaller deviations for receptors in complex with ACh and QNB except for M<sub>1</sub>-ACh and M<sub>1</sub>-QNB, where the TM region possessed larger deviations compared to the apo form, suggesting local rearrangement of this region might have been taken place. Overall, all the receptors maintained their global folding patterns, in parallel with their largely preserved secondary structures and compactness of their TM domains.



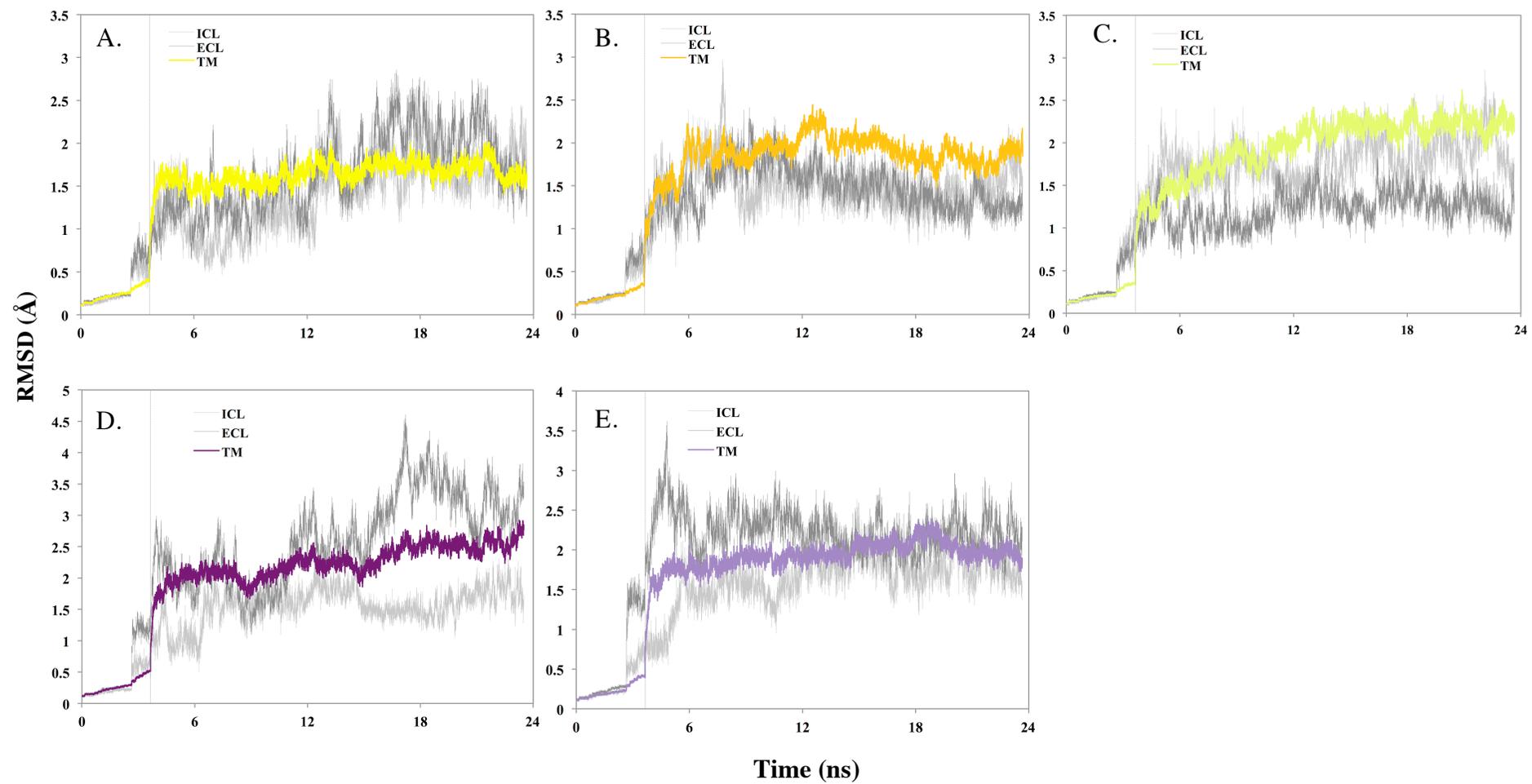
**Figure 4.2.** Secondary structure evolution for systems that underwent MD simulations. **A.**  $M_1$ -Apo; **B.**  $M_1$ -ACh; **C.**  $M_1$ -QNB; **D.**  $M_1^*$ -Apo; and **E.**  $M_1^*$ -ACh. Color code: purple,  $\alpha$ -helix; blue,  $3_{10}$  helix; red,  $\pi$ -helix; cyan, turn; and white, coil. Full stretch of color bar, especially the purple indicated that the secondary structure of the helix was well maintained.



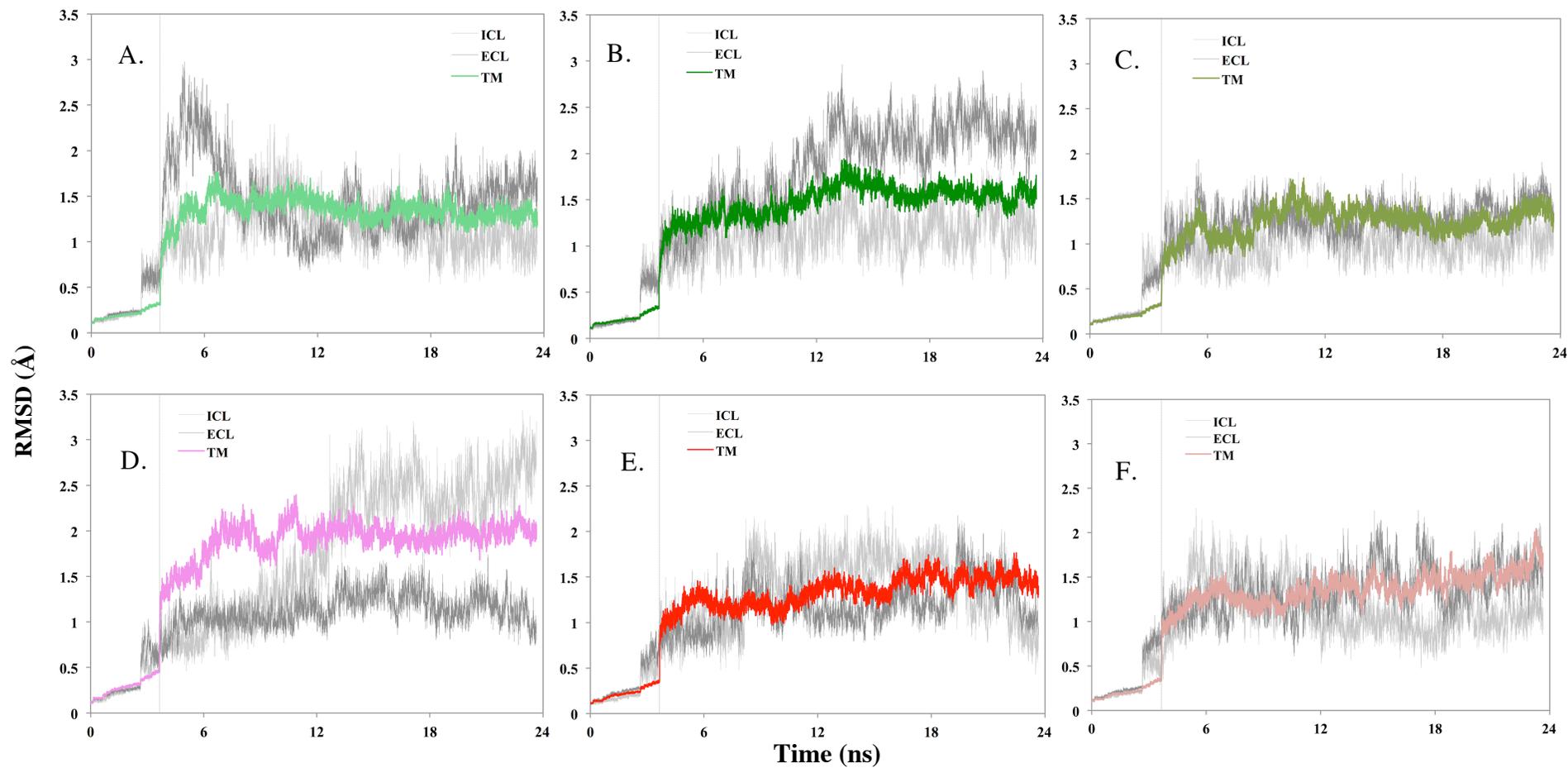
**Figure 4.3.** Secondary structure evolution for systems that underwent MD simulations. **A.** M<sub>2</sub>-Apo; **B.** M<sub>2</sub>-ACh; **C.** M<sub>2</sub>-QNB; **D.** M<sub>3</sub>-Apo; **E.** M<sub>3</sub>-ACh; and **F.** M<sub>3</sub>-QNB. Color code: purple,  $\alpha$ -helix; blue,  $3_{10}$  helix; red,  $\pi$ -helix; cyan, turn; and white, coil. Full stretch of color bar, especially the purple indicated that the secondary structure of the helix was well maintained.



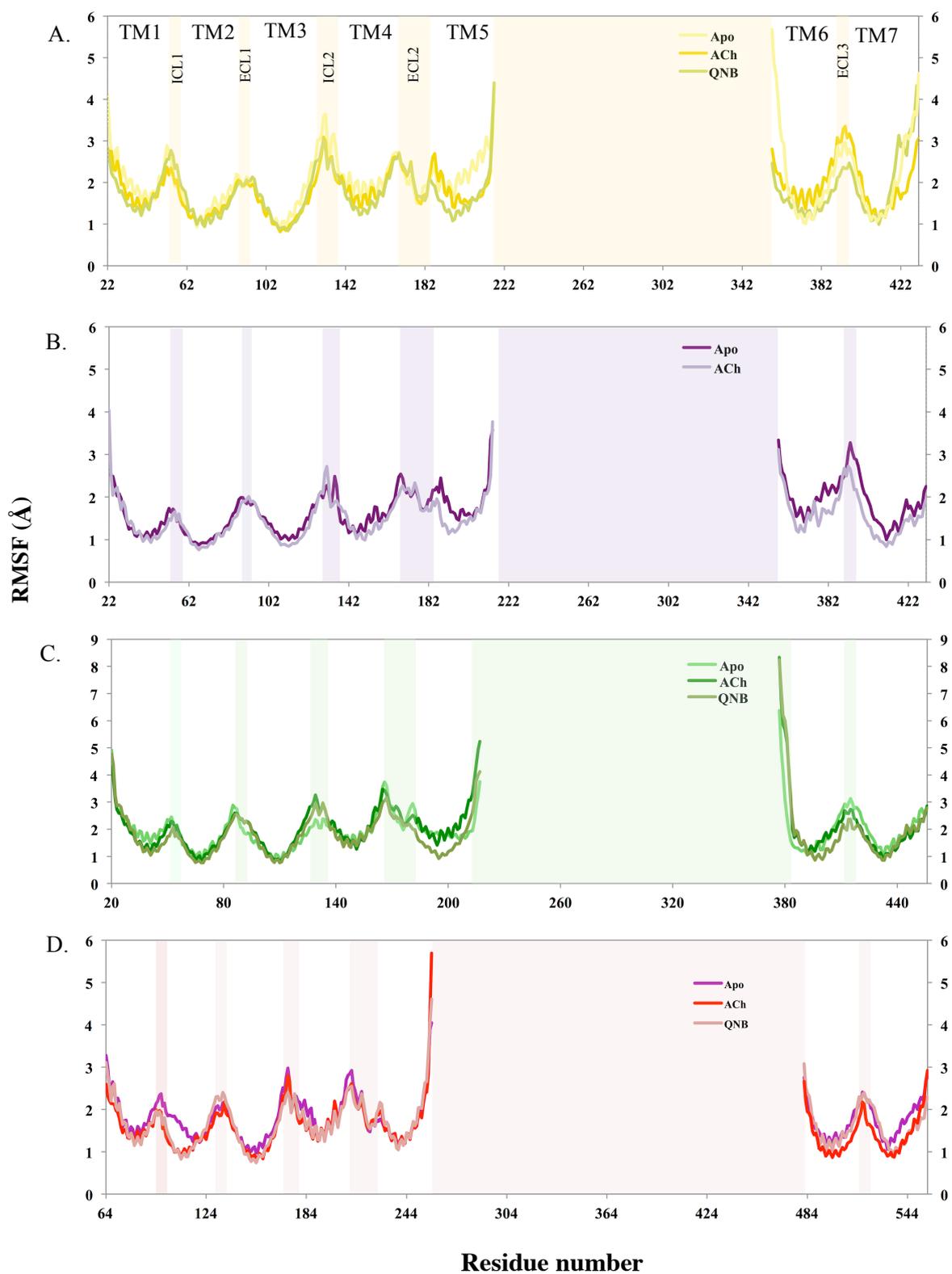
**Figure 4.4.** Radius of gyration calculated from MD trajectories. **A.**  $M_1$ , **B.**  $M_1^*$ , **C.**  $M_2$ , and **D.**  $M_3$ . The vertical lines separate between heating and equilibration phases with production phase.



**Figure 4.5.** Time series of RMSD from the minimized starting structure calculated using backbone atoms of TM domain, ECLs and ICLs. The vertical lines separate between heating and equilibration phases with production phase. **A.** M<sub>1</sub>-Apo; **B.** M<sub>1</sub>-ACh; **C.** M<sub>1</sub>-QNB; **D.** M<sub>1</sub>\*-Apo; and **E.** M<sub>1</sub>\*-ACh.



**Figure 4.6.** Time series of RMSD from the minimized starting structure calculated using backbone atoms of TM domain, ECLs and ICLs. The vertical lines separate between heating and equilibration phases with production phase. **A.** M<sub>2</sub>-Apo; **B.** M<sub>2</sub>-ACh; **C.** M<sub>2</sub>-QNB; **D.** M<sub>3</sub>-Apo; **E.** M<sub>3</sub>-ACh; and **F.** M<sub>3</sub>-QNB.



**Figure 4.7.** RMSF calculated for every residue fitted to the minimized starting structure in all the systems simulated (A. M<sub>1</sub>, B. M<sub>1</sub><sup>\*</sup>, C. M<sub>2</sub>, and D. M<sub>3</sub>). ICLs and ECLs were indicated by shaded background.

### 4.3.2 Collective motions of the receptor structures

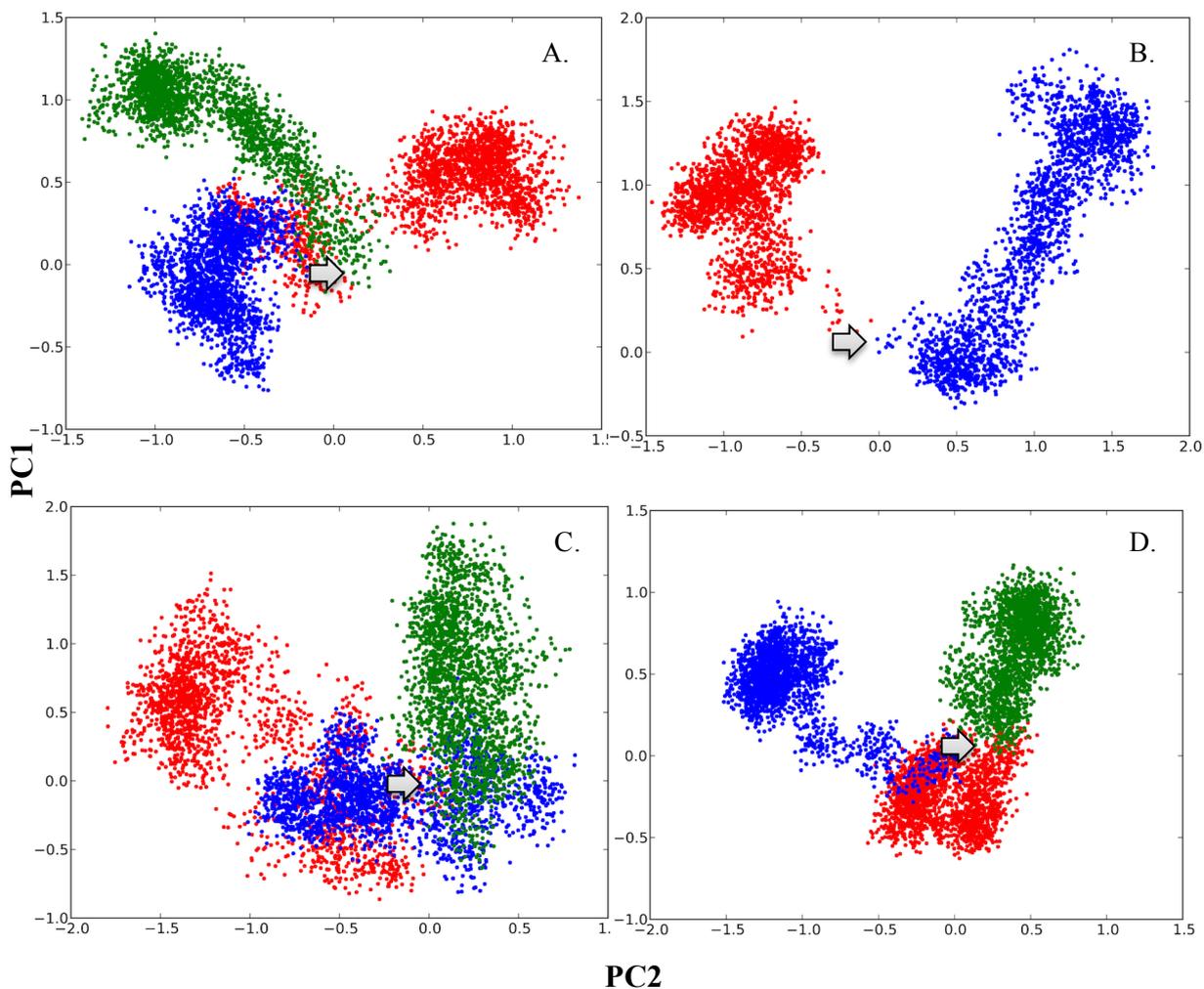
To better understand the global motion of the receptors, principal component analyses (PCA) at 10 ps intervals along the simulation trajectories were used to examine the correlated backbone motions of the receptor structures. For all the systems simulated, the first PCA mode (PC1) refers to the direction of maximal variance, followed by PCA mode 2 (PC2), and etc. The discussion was focused on the PC1 with largest fraction contribution (see appendix, Table D). Two-dimensional projections along the major PCA modes were plotted to represent and compare sample distribution in the configuration space (Figure 4.8). The matrix of inner product in a graphical form is illustrated in Figure 4.9 and 4.10, to show the overlap of the top eight modes. Square displacement plots illustrating structural mobility per residue are shown in Figure 4.11 and the eigenvectors of PC1 and PC2 depicting the correlated backbone motion of the structures are shown in Figures 4.12 to 4.16.

Projections of each simulation onto the vectors produced from PCA resulted in a map of conformational space that illustrated the differences between the systems (Figure 4.8). The plots clearly reveal three separate regions of conformational space occupied by the different forms of each receptor model, suggesting that subtle structural changes occur upon the binding of ACh and QNB. All simulations drifted away from the starting point along distinct paths with M<sub>2</sub> showing greater overlap between the apo, ACh- and QNB-bound forms. The simulations of the M<sub>1</sub>\*- and M<sub>3</sub>-apo systems were more divergent than those of their respective ligand-bound systems, in contrast to what was observed for the simulations involving the M<sub>1</sub> and M<sub>2</sub> structures (Figure 4.8).

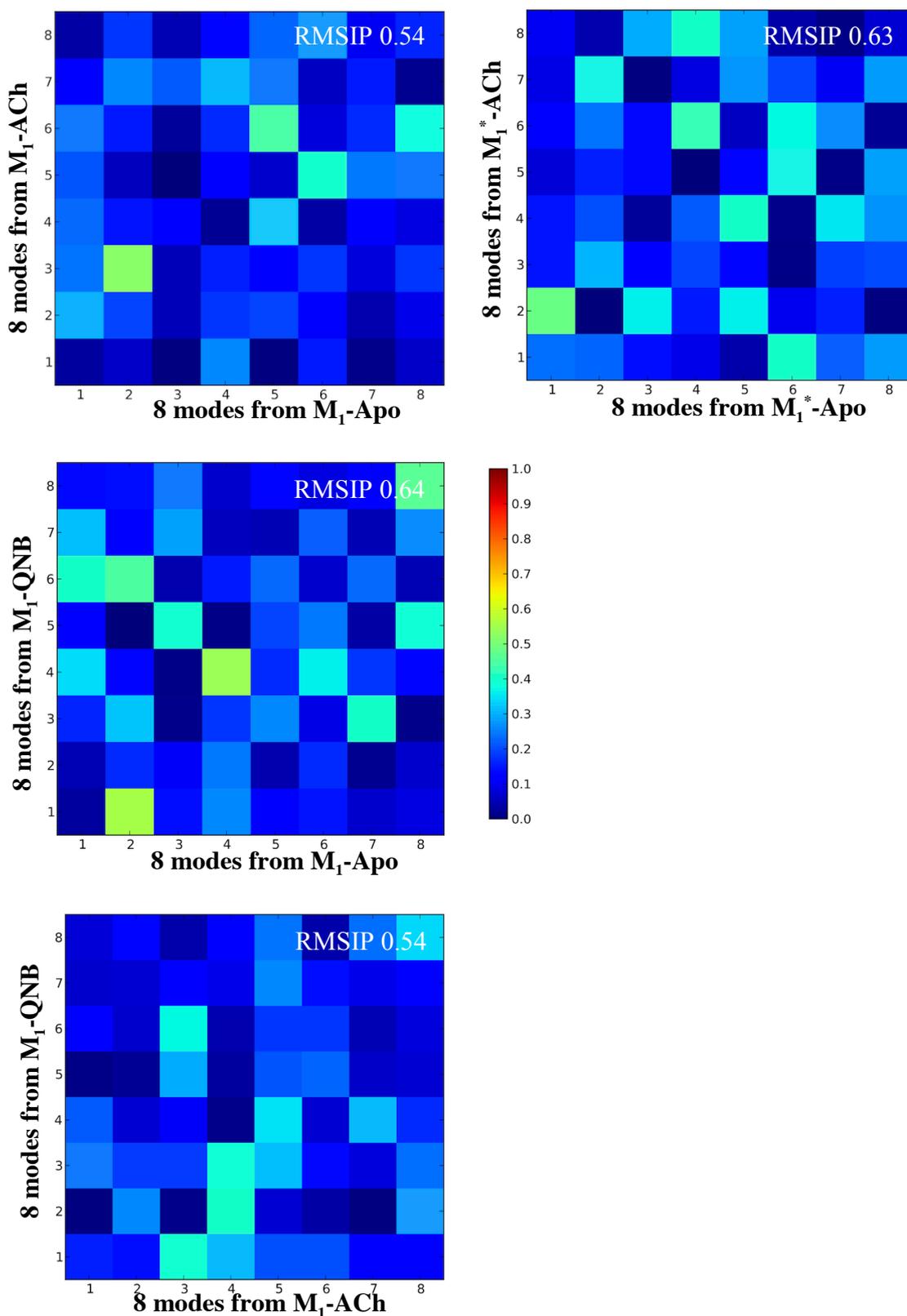
The inner product matrices and root mean square inner product (RMSIP) for the top eight modes was plotted and calculated to assess the similarity between the sets of eigenvectors generated from simulations of the different receptor systems. The RMSIP

for all the simulations were almost the same, in the range from 0.54 to 0.68 (Figure 4.9 and 4.10). In all the systems simulated, the dynamic behaviors of the apo forms were closer to those of the QNB-bound forms. This might be due to the apo forms resembling the inactive forms and QNB not only stabilizing these forms but also steering their dynamic behaviours towards that of their respective apo forms.

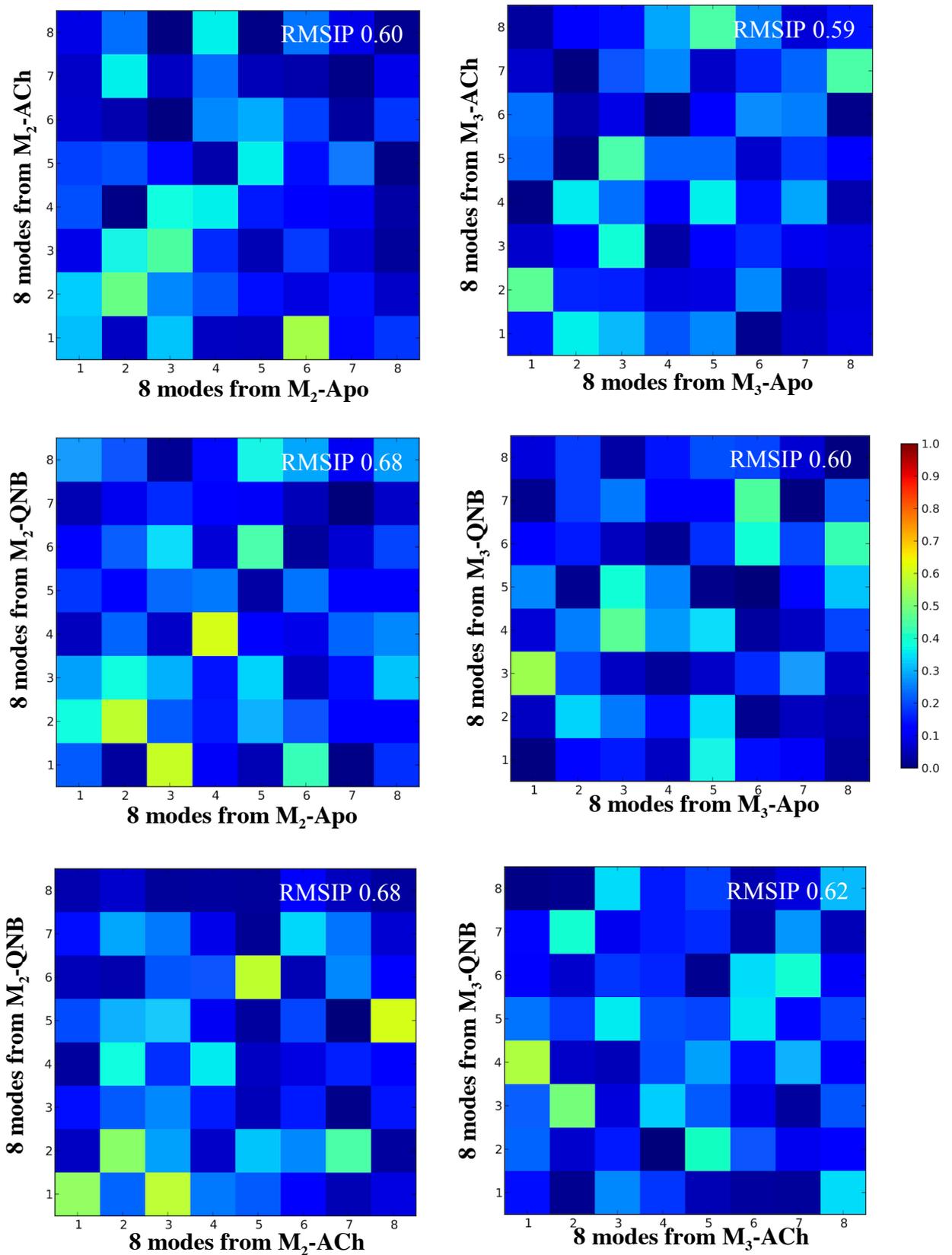
Generally, PCA revealed rather complex collective motions in all the systems simulated. A range of varied helical motions was observed, including bending, twisting, and rocking (animation data not shown). However, it is obvious that the highly dynamic regions were dominated by ICLs and ECLs, sections on TM helices, which connected to them, and the N and C termini of the helices (Figure 4.11). The core sections of the TM helices were more stable than the extracellular and cytoplasmic end of the helices. This is also in accord with the RMSF analyses (Figure 4.7). In most cases, larger and more varied movements were observed in the apo form, especially at the extracellular site and overall dynamics were suppressed upon ACh or QNB binding. Motions in the apo systems led to more open extracellular vestibules compared to the holo systems. The changes of the binding pocket size were particularly apparent from the top view of the receptor structures.



**Figure 4.8.** Projection of trajectory snapshots onto the subspace spanned by principal modes PC1 and PC2 of backbone heavy atoms as determined from the MD simulations. **A.**  $M_1$ , **B.**  $M_1^*$ , **C.**  $M_2$ , and **D.**  $M_3$ . Red: receptor with ACh bound, green: QNB bound, and blue: apo. The arrow marked the starting point for the apo simulations.



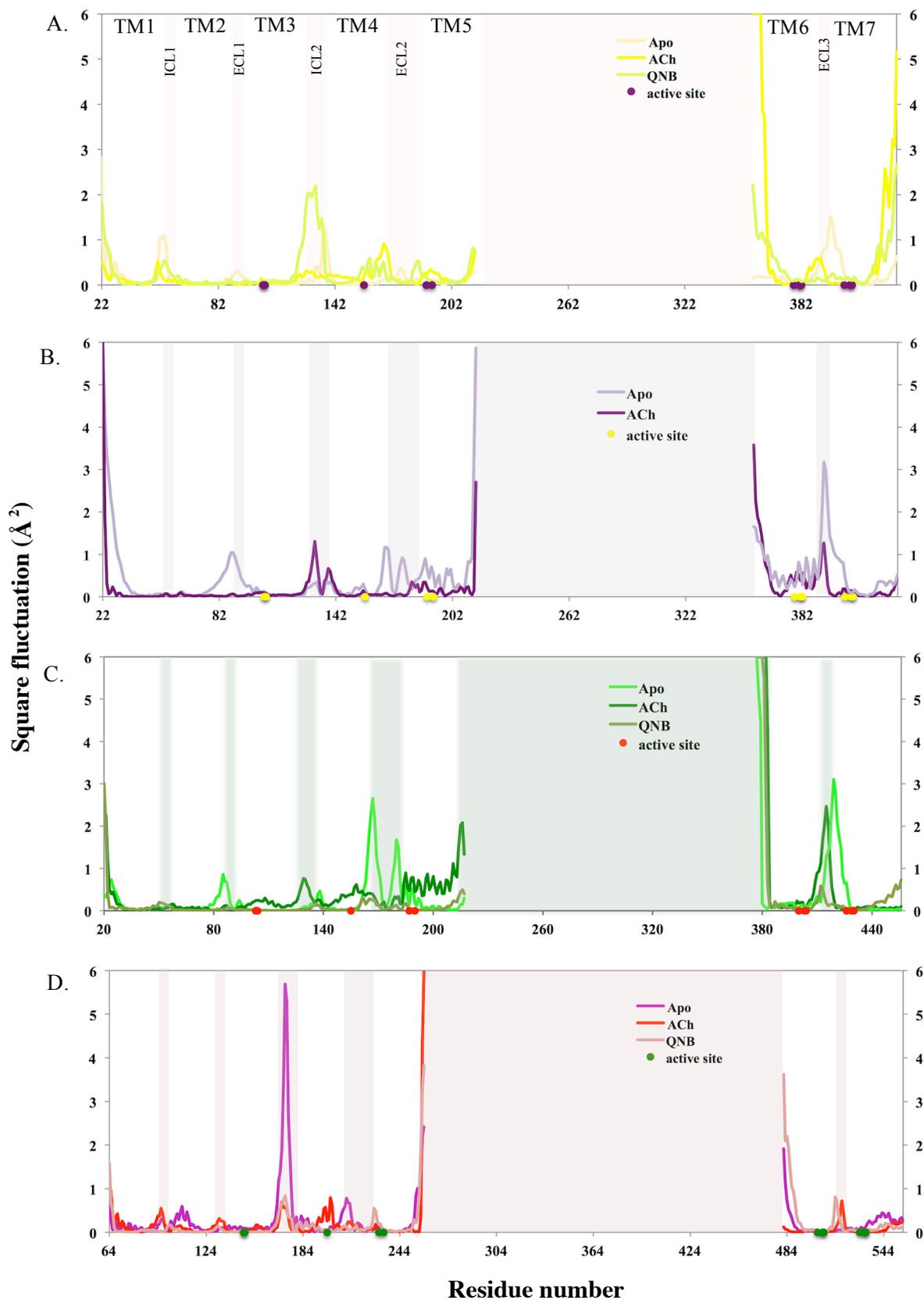
**Figure 4.9.** Overlap between 8 PCA modes from different systems (Left:  $M_1$ , right:  $M_1^*$ ). Similarity between the vector pair is sorted by color, where blue and red indicated least and highest similarity, respectively. RMSIP take values from 0 – 1, with value closer to 1 indicating greater similarity.



**Figure 4.10.** Overlap between 8 PCA modes from different systems (Left:  $M_2$ , right:  $M_3$ ). Similarity between the vector pair is sorted by color, where blue and red indicated least and highest similarity, respectively. RMSIP take values from 0 – 1, with value closer to 1 indicating greater similarity.

PC1 of the  $M_1$ -apo system captured high mobility in the region corresponding to loops, TM1, TM5, TM6, the extracellular end of TM7 and H8. The binding of ACh stabilized most of these regions and mobilized TM4, whereas QNB binding restored the dynamics to being similar to those found for the apo system, except in the extracellular loops and TM4 (extracellular end). The dynamics of the  $M_1^*$ -apo system reflected by PC1 were found to be similar to those of the  $M_1$ -apo system, except for the ECLs which were moving in opposite directions. Generally, the movements coupled along PC1 for the  $M_1^*$ -ACh system were smaller in amplitude as compared to the apo system, but TM6 mobility was relatively higher than for other systems (Figure 4.11).

The  $M_2$ -QNB system, which was obtained directly from the crystal structure (PDB code: 3UON), demonstrated lower mobility in the TM domain compared to the apo and ACh-bound systems. The binding of ACh to  $M_2$  was found to induce higher mobility in TM3, TM4 and TM5, where TM3 was one of the least mobile TM helices, besides TM2, for all the systems simulated. With the removal of QNB, the stability of the TM domain remained, but the loop regions became more flexible. It is therefore clear that ACh and QNB favour different receptor conformations as the binding of ACh increases flexibility of the TM segments, which were found stable in the QNB-bound system. This increased flexibility of TM3, TM4 and TM5 in the presence of ACh may allow this region to sample active state conformations more effectively. The most distinctive feature coupled along PC1 in the  $M_2$ -apo system is the large movement of TM1 (N-terminus), ECL2, ECL3 and the parts of TM5 and TM6 connected to it, the loop part where the ICL3 is truncated at TM6 and the beginning of ECL1. These movements were also observed in PC1 for the  $M_2$ -ACh system, except for the movement of ECL1 and ECL2. For the  $M_2$ -QNB system, large movements were also observed in TM1 (N-terminus part), H8 (C-terminus part) and the tails of TM5 and TM6.



**Figure 4.11.** Mobility plots of first mode of principal component analysis. **A.** M<sub>1</sub>, **B.** M<sub>1</sub><sup>\*</sup>, **C.** M<sub>2</sub>, and **D.** M<sub>3</sub>. For the purpose of clarity, square fluctuations exceed 6  $\text{\AA}^2$  (correspond to loop region where ICL3 was truncated) are not plotted.

While the M<sub>2</sub> system exhibited extensive overlap in the conformational space and dynamics of the different simulated forms (Figure 4.8) and similar fluctuation patterns in the TM domain between the M<sub>2</sub>-apo and M<sub>2</sub>-QNB systems (Figure 4.11), this was not the case for the M<sub>3</sub> system, despite the fact that they were both obtained from crystal structures of receptors in complex with a blocker. The absence of QNB resulted in greater mobility in TM1, TM2, TM3, TM7 and H8, which were stabilized by the binding of QNB or ACh. PC1 of M<sub>3</sub>-ACh was somewhat similar to that of M<sub>3</sub>-QNB, except that it possessed a less mobile TM6 tail and a more dynamic TM1 (N-terminus) and TM4. The projection plot showed that both QNB and ACh indeed stabilized conformations that were not sampled by the apo receptor, in contrast to what was observed for the M<sub>2</sub> system (Figure 4.8).

The presence of ACh or QNB in the binding cavity notably reduced the mobility of the helices, resulting in changes to the binding site volume. TM4 was found to be sensitive to the presence of ACh as larger fluctuations were observed in the M<sub>1</sub>-, M<sub>2</sub>- and M<sub>3</sub>-ACh systems (Figure 4.11). The extracellular site sequence of TM4 is critical for agonist and antagonist binding. Among the residues residing in TM4, W<sup>4.57</sup> and P<sup>4.59</sup>, whose mutation not only reduced ACh and QNB affinity, but also decrease signalling efficacy, play an important role in the receptor activation (Lu, Saldanha, & Hulme, 2001). It may therefore be surmised that these ACh-induced fluctuations in TM4, reflecting different receptor conformations than those favoured by QNB, are needed to accommodate ACh. However, this phenomenon was not observed in the M<sub>1</sub>\*-ACh system, which was modeled to represent the activated form of M<sub>1</sub>. The extracellular site of the highly mobile TM6 was moving outward in the apo form of M<sub>1</sub>\*, and the binding of ACh reversed this action. As a consequence, TM5, which was moving inwards in the apo form, gave way to TM6 by moving sideways. Together with the inward movement

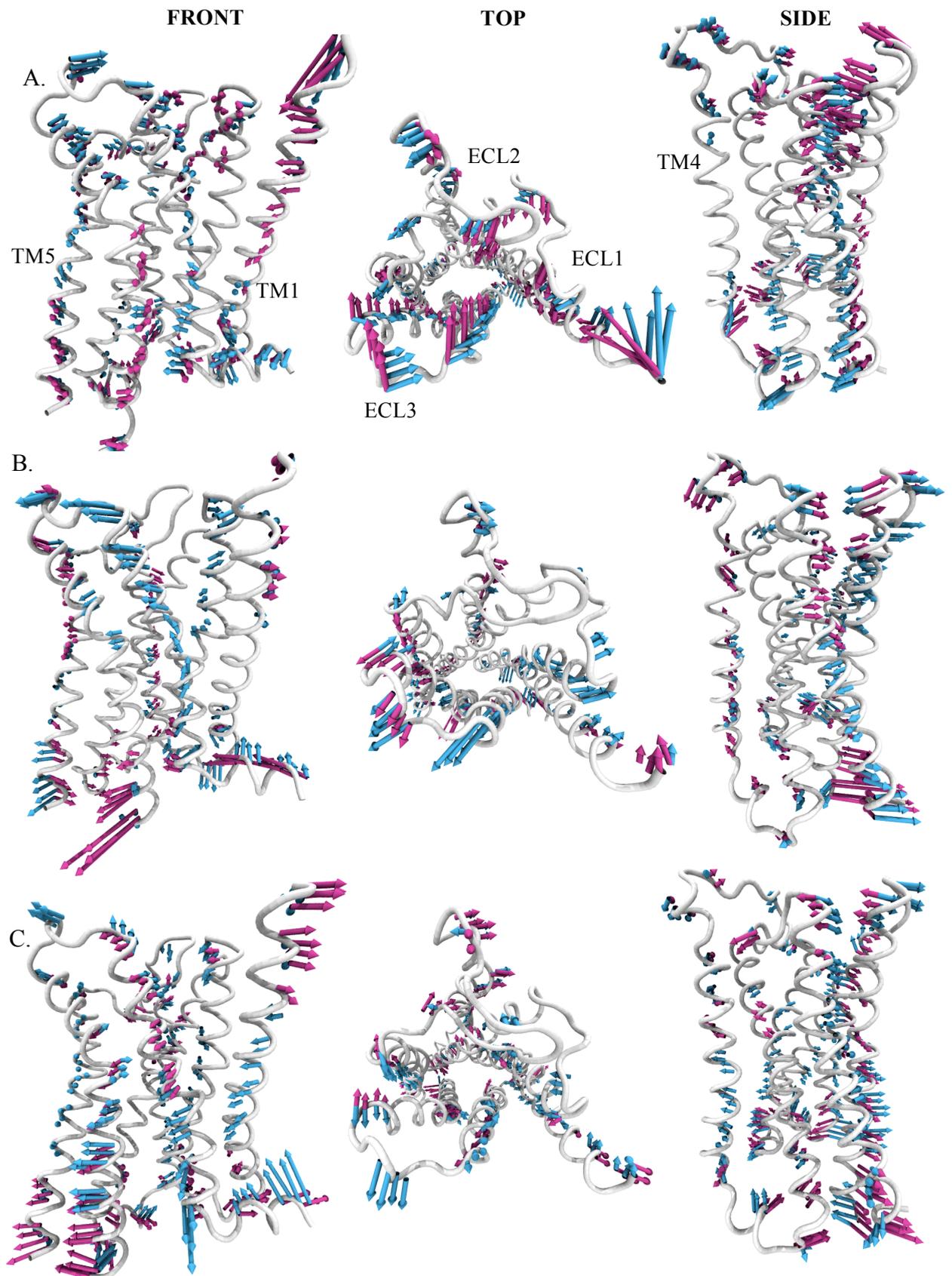
of TM7, this resulted in a smaller, compressed binding cavity in the  $M_1^*$ -ACh system (Figure 4.13 and animation data not shown). The binding of ACh to the  $M_1^*$  structure mobilized TM5, TM6 and TM7 but TM1 (except the N-terminus portion), TM2, TM3 and TM4 were relatively less mobile (Figure 4.11 and 4.13).

TM5 was found to be relatively more flexible, in the  $M_1$ -ACh,  $M_1^*$ -ACh and  $M_2$ -ACh systems, compared to the corresponding apo- and QNB-bound systems. This is in accordance with SDM studies, which suggest that TM5 rotation optimized the anchoring of ACh and this action is part of the activation process (Allman, Page, Curtis, & Hulme, 2000). Even though TM5 in the  $M_3$ -ACh systems was relatively stable compared to the other ACh-bound systems, a slight rotation in the extracellular segment of TM 5 was observed. Comparing all the simulated systems, TM1 (except the N-terminus) and TM2 were found to be the most rigid helices, in all forms of the receptors, followed by TM3. TM4 appeared to be rigid in apo forms and was sensitive to the binding of ligands. While TM1 and TM2 did not take part in the binding of ACh and QNB, the important residues, D<sup>3.32</sup>, W<sup>4.57</sup>, and P<sup>4.59</sup>, reside on TM3 and TM4, respectively. Slight fluctuations in TM3 and TM4 reflect their functional relevance in the binding of ligands. The rigidity of TM1 and TM2 might be due to inter-helical interactions which help to maintain the global folding pattern of the receptor.

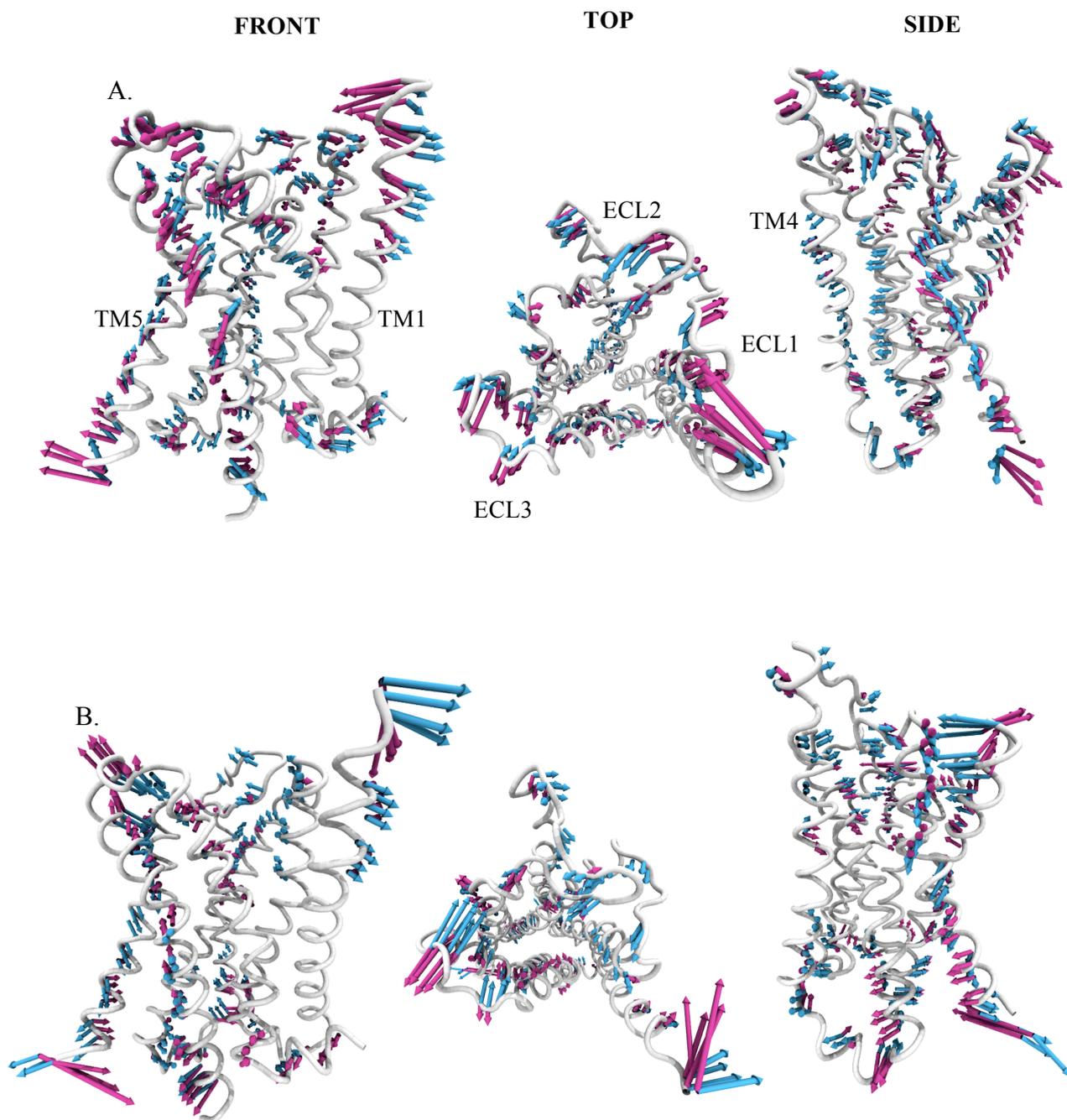
It has been proposed that the mobilization of TM6 and TM7 with respect to TM3 are essential in stabilizing the activated state of mAChRs, through rearrangement of the hydrogen bonding networks (Hulme, Lu, Saldanha, & Bee, 2003). It is known that TM6 and TM7 contain 2 important switches, the transmission switch (CWxP) and the tyrosine toggle switch (NPxxY), respectively. TM6 also forms an ionic lock switch with TM3 at the cytoplasmic end of the helices. Additionally, Y<sup>6.51</sup> in TM6 plays a key role in

receptor function and activation, while rotation of TM6 has also been implicated in receptor activation (Venkatakrishnan et al., 2013). Hence the higher mobility observed in TM5, TM6 and TM7 in the  $M_1^*$ -ACh system, captured through PC1, suggest that the model closely resembles an activated state.

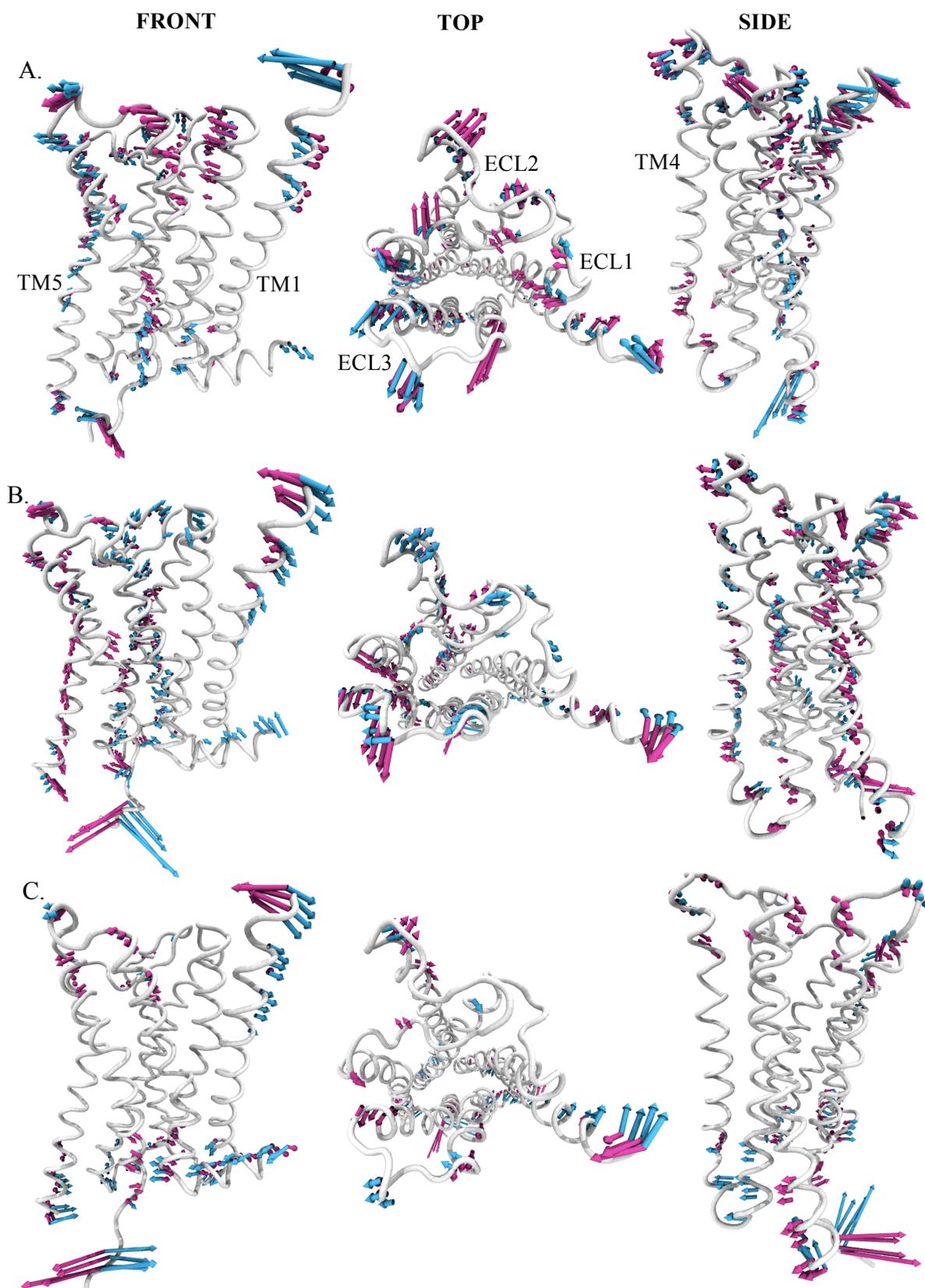
The dynamics profiles of the cytoplasmic ends of TM5 and TM6 upon the binding of ACh or QNB may be functional relevant as they are connected by ICL3 which is involved in coupling to the G protein. Conformational plasticity of the receptors triggered by ligand binding has been shown experimentally and theoretically (Wess, Han, Kim, Jacobson, & Li, 2008; Rasmussen, Choi, et al., 2011; Novikov, Sivozhelezov, & Shaitan, 2013; Xu, Li, Sun, Li, & Hou, 2013). In this study, the cytoplasmic end of TM6 was found to be highly mobile, especially upon the binding of ACh to the  $M_1$  structure, which also mobilized H8 to give space to accommodate the flexible TM6 tail (animation data not shown). Unexpectedly, the complete reverse was found for the  $M_3$  structure as the binding of ACh did not induce high flexibility in the TM6 tail, but this motion was observed in the presence of QNB instead. Large movements of ICL2 were also observed in the ACh-bound forms of all the models except for the  $M_3$ -ACh system. The ECLs were also observed to be highly mobile throughout the simulations and the binding of ligands suppressed the mobility of ECL2 (Figure 4.12-4.16). Experimental studies have shown that the ECLs are clearly functionally relevant, especially ECL2, which has been proved to be the gatekeeper for ligand binding in the  $M_2$  receptor (Avlani et al., 2007). Besides their role in ligand recognition and allosteric modulations, they might also influence receptor activation and signaling (Peeters, van Westen, Li, & AP, 2011; Gil-Mast, Kortagere, Kota, & Kuzhikandathil, 2013; Seibt et al., 2013).



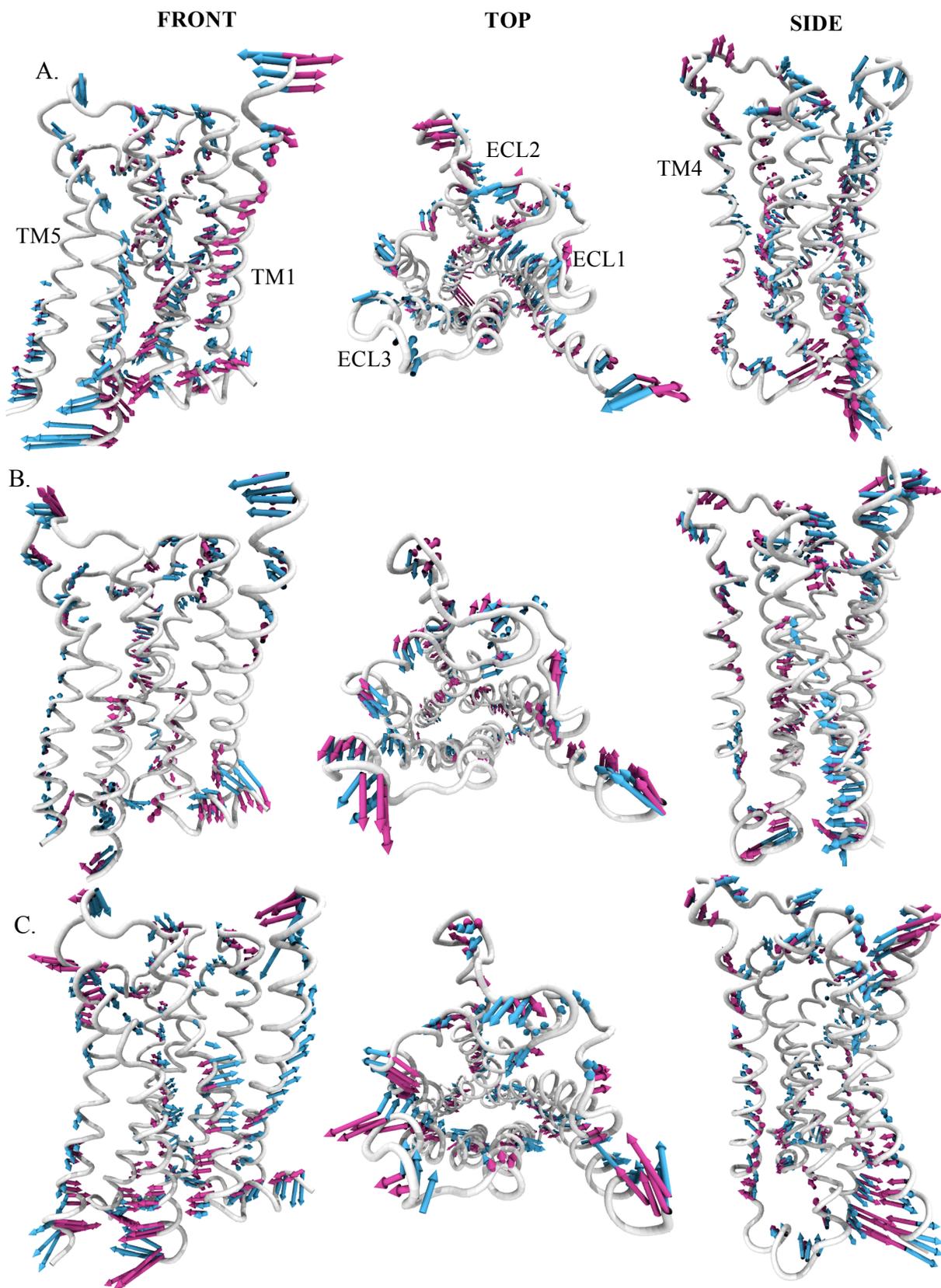
**Figure 4.12.** Structural variations along the top two PCA modes. **A.**  $M_1$ -Apo, **B.**  $M_1$ -ACh, **C.**  $M_1$ -QNB. The receptors are shown in tube, the arrows indicate directions and amplitudes. For the purpose of clarity, arrows are shown every 4<sup>th</sup> residues and only if their lengths are more than 1 Å. Magenta: PC1, cyan: PC2.



**Figure 4.13.** Structural variations along the top two PCA modes. **A.**  $M_1^*$ -Apo, and **B.**  $M_1^*$ -ACh. The receptors are shown in tube, the arrows indicate directions and amplitudes. For the purpose of clarity, arrows are shown every 4<sup>th</sup> residues and only if their lengths are more than 1 Å. Magenta: PC1, cyan: PC2.



**Figure 4.14.** Structural variations along the top two PCA modes. **A.**  $M_2$ -Apo, **B.**  $M_2$ -ACh, and **C.**  $M_2$ -QNB. The receptors are shown in tube, the arrows indicate directions and amplitudes. For the purpose of clarity, arrows are shown every 4<sup>th</sup> residues and only if their lengths are more than 1 Å. Magenta: PC1, cyan: PC2.



**Figure 4.15.** Structural variations along the top two PCA modes. **A.**  $M_3$ -Apo, **B.**  $M_3$ -ACh, and **C.**  $M_3$ -QNB. The receptors are shown in tube, the arrows indicate directions and amplitudes. For the purpose of clarity, arrows are shown every 4<sup>th</sup> residues and only if their lengths are more than 1 Å. Magenta: PC1, cyan: PC2.

### 4.3.3 Hydrogen bond networks

Hydrogen bond networks are essential molecular interactions maintaining the individual helical structure and the overall architecture of the TM bundle in the ground and activated states. Upon ligand binding, hydrogen bond networks may be either strengthened, resulting in stabilization of the receptor conformation, or disrupted, facilitating receptor activation. Hydrogen bond analyses on the MD trajectories were carried out and their fractions, which reflect the percentage of conservation, are reported in Tables 4.4 and 4.5. The simulations and PCA show that TM1 to TM4, especially TM1 and TM2, undergo much less movement than the other TMs. These regions are likely to be stabilized by a network of hydrogen bonding interactions involving residues S<sup>1.43</sup>, T<sup>1.46</sup>, D<sup>2.50</sup>, N<sup>2.59</sup>, S<sup>7.46</sup>, T<sup>7.47</sup>, N<sup>7.49</sup>, and Y<sup>7.53</sup>. The first two of these residues located near to the conserved N<sup>1.50</sup> and the last two residues are part of the NPxxY motif. Analysis of the hydrogen bond networks also highlighted the important role of the conserved residue D<sup>2.50</sup> in establishing extensive hydrogen bond networks in all the systems simulated, mutation of which leads to a complete abolishment of signaling (Bee & Hulme, 2007). However, the role of N<sup>1.50</sup> was surprisingly found to be less important in this. In the M<sub>1</sub>\*-ACh system, direct interactions between D<sup>2.50</sup> and N<sup>7.49</sup> existed for about 60% of the simulation time but was almost absent in the apo simulation, as opposed to the interactions between residues N<sup>1.50</sup> and N<sup>7.49</sup>. Similarly, direct interactions between D<sup>2.50</sup> and N<sup>7.49</sup> were found in the M<sub>3</sub> holo system. Several experimental studies have suggested the additional role of N<sup>7.49</sup> in stabilizing an active conformation of the M<sub>1</sub> receptor (Lu et al., 2001) and have observed direct interactions between D<sup>2.50</sup> and N<sup>7.49</sup> in the active state of GPCRs, through side chain conformation change in N<sup>7.49</sup> (Govaerts et al., 2001; Urizar et al., 2005).

**Table 4.4.** Hydrogen bond fractions of M<sub>1</sub> and M<sub>1</sub>\* simulations.

	Hydrogen bonding pair		Location	ACh	Apo	QNB
M <sub>1</sub>	D71 <sup>2.50</sup> _OD1/OD2	S411_OG	TM2-TM7	0.9580	0.9635	0.9440
	N80_OD1	S36_OG	TM2-TM1	0.9110	0.7905	0.9090
	T39_OG1	T412_OG1	TM1-TM7	0.4355	0.8430	0.7965
	D71 <sup>2.50</sup> _OD1/OD2	Y418_OH	TM2-TM7	0.2505	0.0010	0.7605
	D71 <sup>2.50</sup> _OD1/OD2	N414 <sup>7.49</sup> _ND2	TM2-TM7	0.7385	0.6360	0.6570
	S411_O	N43 <sup>1.50</sup> _ND2	TM7-TM1	0.0950	0.2135	0.0430
	ACh_OD2	S109 <sup>3.36</sup> _OG	TM3	0.0165	-	-
	ACh_OD2	N382 <sup>6.52</sup> _ND2	TM6	0.0005	-	-
	ACh_O2	WAT_O	-	1.0000	-	-
	ACh_O1	WAT_O	-	0.0040	-	-
	N382 <sup>6.52</sup> _OD1	QNB_O1	TM6	-	-	0.7845
	QNB_O1/O2	N382 <sup>6.52</sup> _ND2	TM6	-	-	0.6560
	D105 <sup>3.32</sup> _OD1/OD2	QNB_N	TM3	-	-	0.2615
	Y404 <sup>7.39</sup> _OH	QNB_N	TM7	-	-	0.0005
	WAT_O	QNB_N	-	-	-	0.5210
	M <sub>1</sub> *	N80_OD1	S36_OG	TM2-TM1	0.8790	0.9043
T39_OG1		T412_OG1	TM1-TM7	0.8340	0.4650	
N115_OD1		S66_OG	TM3-TM2	0.8325	0.8181	
D71 <sup>2.50</sup> _OD1/OD2		S411_OG	TM2-TM7	0.9530	0.9783	
N410_OD1		T377_OG1	TM7-TM6	0.7830	0.1285	
S411_O		N43 <sup>1.50</sup> _ND2	TM7-TM1	0.2105	0.3345	
N43 <sup>1.50</sup> _OD1		N414 <sup>7.49</sup> _ND2	TM1-TM7	0.0215	0.5854	
N43 <sup>1.50</sup> _OD1		Y418_OH	TM1-TM7	0.0115	0.0015	
D71 <sup>2.50</sup> _OD1/OD2		N414 <sup>7.49</sup> _ND2	TM2-TM7	0.6190	0.1552	N/A
D71 <sup>2.50</sup> _OD1/OD2		S112_OG	TM2-TM3	0.4945	0.0015	
ACh_O2		T192 <sup>5.42</sup>	TM5	0.1185	-	
ACh_O2		Y381 <sup>6.51</sup> _OH	TM6	0.0990	-	
ACh_O2		N382 <sup>6.52</sup> _ND2	TM6	0.0895	-	
ACh_O2		Y106 <sup>3.33</sup> _OH	TM3	0.0030	-	
ACh_OD1/O2		S109 <sup>3.36</sup> _OG	TM4	0.0010	-	
ACh_O2		Y404 <sup>7.39</sup> _OH	TM7	0.0005	-	
ACh_O2	WAT_O	-	0.3135	-		

Only conserved and orthosteric site residues are labelled with Ballesteros-Weinstein scheme, as a superscript extension.

**Table 4.5.** Hydrogen bond fractions of M<sub>2</sub> and M<sub>3</sub> simulations.

	Hydrogen bonding pair		Location	ACh	Apo	QNB
M <sub>2</sub>	D69 <sup>2.50</sup> _OD1	S433_OG	TM2-TM7	0.9715	0.9320	0.9630
	T37_OG1	T434_OG1	TM1-TM7	0.8140	0.8225	0.7885
	S433_O	N42 <sup>1.50</sup> _ND2	TM7-TM1	0.1550	0.0735	0.2290
	D69 <sup>2.50</sup> _OD2	N436 <sup>7.49</sup> _ND2	TM2-TM7	0.6625	0.6595	0.7250
	ACh_O2	W400 <sup>6.48</sup> _NE1	TM6	0.1690	-	-
	ACh_O2	N404 <sup>6.52</sup>	TM6	0.1330	-	-
	ACh_O1	S107 <sup>3.36</sup> _OG1	TM3	0.0005	-	-
	ACh_O2	WAT_O	-	0.6620	-	-
	ACh_O2	WAT_O	-	0.0065	-	-
	N404 <sup>6.52</sup> _OD1	QNB_O1	TM6	-	-	0.8360
	QNB_O1/O2	N404 <sup>6.52</sup> _ND2	TM6	-	-	0.6920
	S107 <sup>3.36</sup> _OG	QNB_N	TM3	-	-	0.2730
	D103 <sup>3.32</sup> _OD1/OD2	QNB_N	TM3	-	-	0.3670
	WAT_O	QNB_N	-	-	-	0.0785
	M <sub>3</sub>	T81_OG1	T537_OG1	TM1-TM7	0.8470	0.8965
D113 <sup>2.50</sup> _OD1/OD2		S536_OG	TM2-TM7	0.9795	0.8230	0.9555
D113 <sup>2.50</sup> _OD1/OD2		N539 <sup>7.49</sup> _ND2	TM2-TM7	<b>0.6410</b>	0.1400	<b>0.8365</b>
D113 <sup>2.50</sup> _OD2		Y543_OH	TM2-TM7	0.0005	0.1215	0.0150
S536_O		N85 <sup>1.50</sup> _ND2	TM7-TM1	0.0875	0.0870	0.2780
ACh_O2		N507 <sup>6.52</sup> _ND2	TM6	0.0575	-	-
ACh_O2		Y148 <sup>3.33</sup> _OH	TM3	0.0055	-	-
ACh_O2		WAT_O	-	0.8805	-	-
ACh_O1		WAT_O	-	0.0130	-	-
N507 <sup>6.52</sup> _OD1		QNB_O1	TM6	-	-	0.8675
QNB_O2		N507 <sup>6.52</sup> _ND2	TM6	-	-	0.6740
D147 <sup>3.32</sup> _OD1/OD2		QNB_N	TM3	-	-	0.8615
S151 <sup>3.36</sup> _OG	QNB_N	TM3	-	-	0.0225	

Only conserved and orthosteric site residues are labelled with Ballesteros-Weinstein scheme, as a superscript extension.

The water-mediated inter-helical hydrogen bonds with greater than 50% of occupancy during the simulations were also analysed. It was found that all the simulated systems had dissimilar patterns of hydrogen bond networks (Tables 4.6 and 4.7, Figures 4.16 to 4.19). In particular, the hydrogen bond network patterns for the holo forms of the receptors were different from their respective apo forms, supporting the notion that ligand binding destabilizes contacts between the inner parts of the TM core and thus stabilization occurs through different hydrogen bonding pair residues. The M<sub>1</sub>-, M<sup>\*</sup><sub>1</sub>- and M<sub>2</sub>-ACh systems showed less extensive networks compared to the corresponding apo and QNB-bound forms, while the holo forms of the M<sub>3</sub> structure had hydrogen bond networks which were highly similar. The binding of ACh to the receptors tends to break hydrogen bond networks, including those between TM 1-2-7, TM 2-3-6-7 (Tables 4.6 and 4.7), while QNB mostly enhances these networks, holding the receptors in stable conformations. The water-mediated hydrogen bonding network involving D<sup>2.50</sup>-W<sup>6.48</sup>-N<sup>7.45</sup>, was broken in the holo form of all the structures, indicating that this network of interactions probably holds the receptors in the apo form (Tables 4.6 and 4.7). The water clusters and networks found were also reported from analyses of 11 GPCR crystal structures, suggesting some of the water molecules are conserved, including those present in the environment of D<sup>2.50</sup> of the N/SLxxxD motif in TM2, R<sup>3.50</sup> of the D/ERY motif in TM3, W<sup>6.48</sup> of the CWxPF/Y motif in TM6 and N<sup>7.49</sup>, Y<sup>7.53</sup> of the NPxxY motif in TM7 (Angel, Chance, & Palczewski, 2009). Water-mediated hydrogen bonding networks are shown to be functionally important, as they involve many highly conserved residues (Standfuss et al., 2011; Deupi et al., 2012). This is also evident in this work where most of the water molecules are located in close proximity to the conserved residues (or adjacent residues) and motifs such as D/ERY, CWxPF/Y and NPxxY to mediate hydrogen bonds (Tables 4.6 and 4.7). Site-directed mutagenesis studies have revealed that the residues, D<sup>2.50</sup>, D<sup>3.32</sup>, Y<sup>3.33</sup>, S<sup>3.36</sup>, S<sup>3.39</sup>, D<sup>3.49</sup>, R<sup>3.50</sup>, W<sup>6.48</sup>,

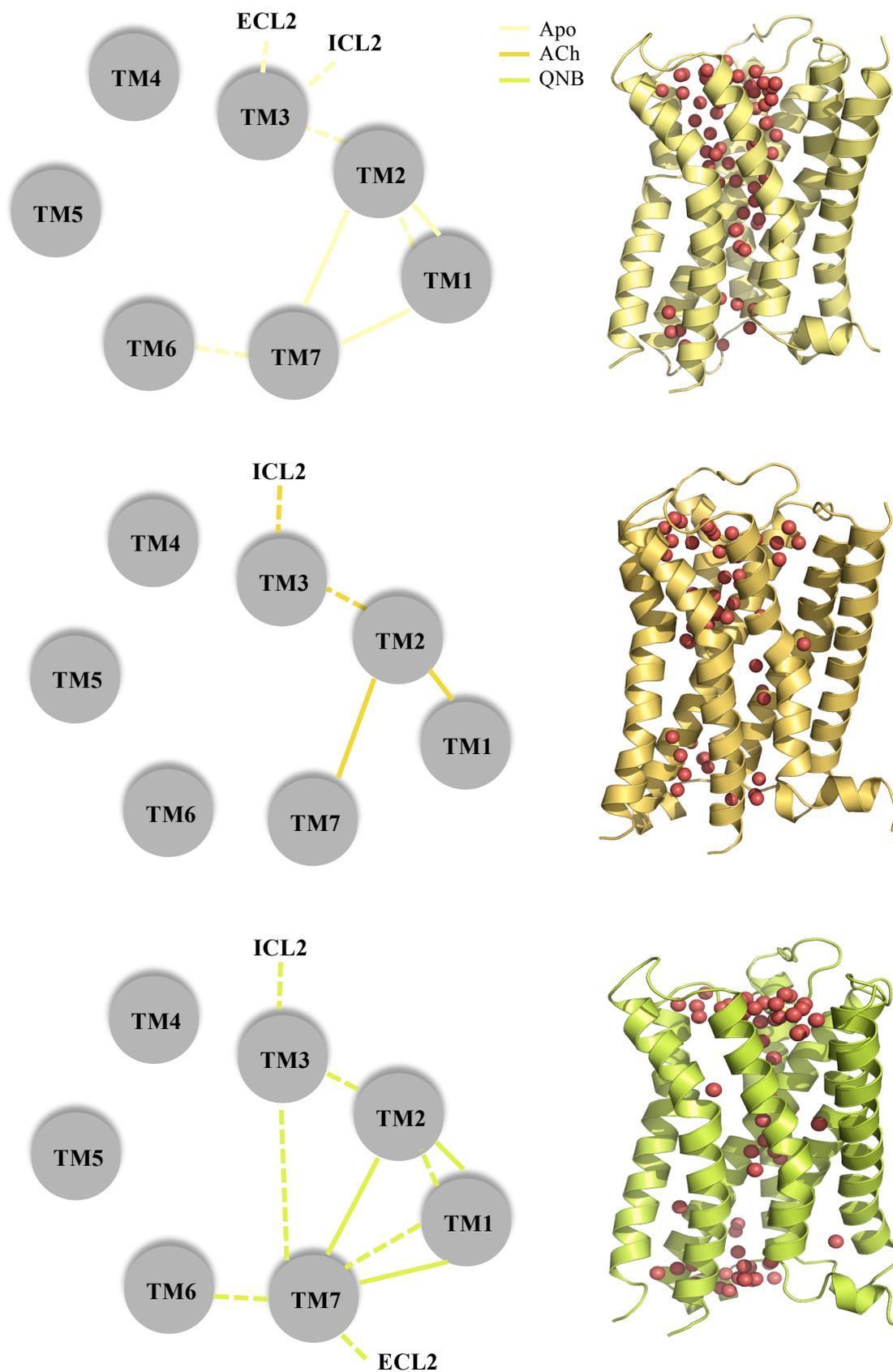
Y<sup>6.51</sup>, N<sup>6.52</sup>, N<sup>7.49</sup> and Y<sup>7.53</sup>, are involved in the binding, signalling and activation of the muscarinic receptors (Blüml, Mutschler, & Wess, 1994; Spalding, Burstein, Henderson, Ducote, & Brann, 1998; Huang, Nagy, Williams, Peseckis, & Messer, 1999; Lu & Hulme, 1999; Ward, Curtis, & Hulme, 1999; Allman et al., 2000; Lu et al., 2001; Spalding et al., 2006; Bee & Hulme, 2007; Goodwin, Hulme, Langmead, & Tehan, 2007). The involvement of the highly conserved residues reflects their importance in stabilizing the inactive state conformations through water-mediated contacts between highly conserved motifs. Reorganization in these contacts is expected upon ligand binding, either to further stabilize the receptor conformation or to facilitate receptor activation (Pardo, Deupi, Dolker, Lopez-Rodriguez, & Campillo, 2007).

**Table 4.6.** Residues involved in inter-helical water-mediated hydrogen bonds.

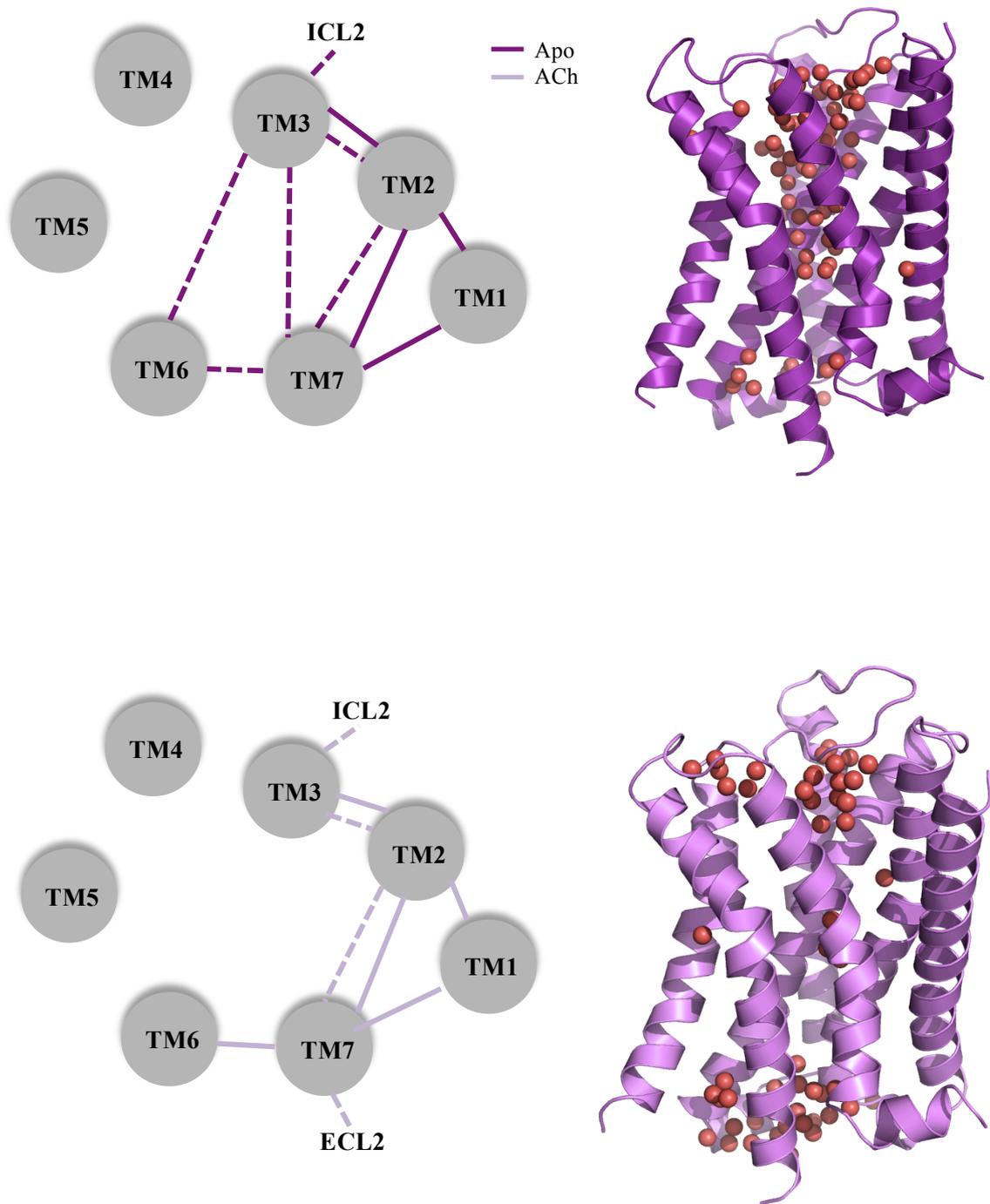
Systems	Residues
M <sub>1</sub> -Apo	N43 <sup>1.50</sup> D71 <sup>2.50</sup> D105 <sup>3.32</sup> S112 <sup>3.39</sup> Y381 <sup>6.51</sup> N382 <sup>6.52</sup> N410 <sup>7.45</sup> Y418 <sup>7.53</sup> D105 <sup>3.32</sup> I180 <sup>ECL2</sup> S184 <sup>ECL2</sup> N60 <sup>2.39</sup> D122 <sup>3.49</sup> R123 <sup>3.50</sup> R134 <sup>ICL2</sup>
M <sub>1</sub> -ACh	D71 <sup>2.50</sup> S112 <sup>3.39</sup> N60 <sup>2.39</sup> D122 <sup>3.49</sup> R137 <sup>ICL2</sup> T138 <sup>ICL2</sup>
M <sub>1</sub> -QNB	Y82 <sup>2.61</sup> Q24 <sup>1.31</sup> E401 <sup>7.36</sup> I180 <sup>ECL2</sup> D71 <sup>2.50</sup> S112 <sup>3.39</sup> N410 <sup>7.45</sup> D105 <sup>3.32</sup> Y106 <sup>3.33</sup> Y408 <sup>7.43</sup> Y404 <sup>7.39</sup> T377 <sup>6.47</sup> Y381 <sup>6.51</sup> N60 <sup>2.39</sup> D122 <sup>3.49</sup> R123 <sup>3.50</sup> R134 <sup>ICL2</sup> R135 <sup>ICL2</sup>
M <sub>1</sub> <sup>*</sup> -Apo	E401 <sup>7.36</sup> Y381 <sup>6.51</sup> D71 <sup>2.50</sup> L67 <sup>2.46</sup> N414 <sup>7.49</sup> N410 <sup>7.45</sup> W378 <sup>6.48</sup> S112 <sup>3.39</sup> S109 <sup>3.36</sup> D105 <sup>3.32</sup> Y404 <sup>7.39</sup> N60 <sup>2.39</sup> D122 <sup>3.49</sup> R123 <sup>3.50</sup> R137 <sup>ICL2</sup>
M <sub>1</sub> <sup>*</sup> -ACh	Y82 <sup>2.61</sup> D105 <sup>3.32</sup> D71 <sup>2.50</sup> N414 <sup>7.49</sup> N60 <sup>2.39</sup> D122 <sup>3.49</sup> R123 <sup>3.50</sup> R137 <sup>ICL2</sup> E401 <sup>7.36</sup> Y179 <sup>ECL2</sup> Y404 <sup>7.39</sup> I180 <sup>ECL2</sup>

**Table 4.7.** Residues involved in inter-helical water-mediated hydrogen bonds.

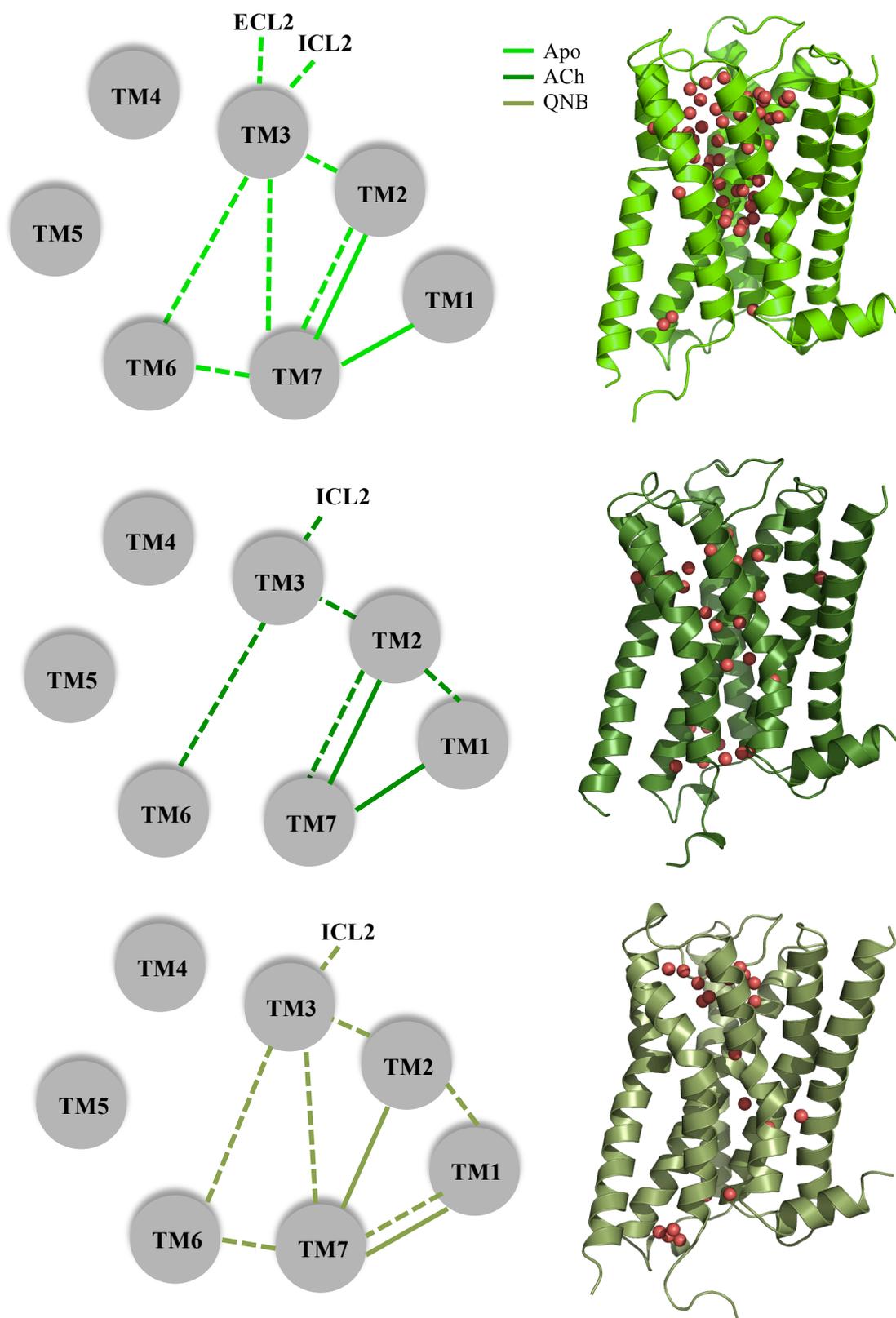
Systems	Residues
M <sub>2</sub> -Apo	D69 <sup>2.50</sup> Y440 <sup>7.53</sup> S110 <sup>3.39</sup> W400 <sup>6.48</sup> Y430 <sup>7.43</sup> N432 <sup>7.45</sup> I392 <sup>6.40</sup>
	N58 <sup>2.39</sup> D120 <sup>3.49</sup> R121 <sup>3.50</sup> R135 <sup>ICL2</sup>
	D103 <sup>3.32</sup> Y430 <sup>7.43</sup> Y403 <sup>6.51</sup>
	Y104 <sup>3.33</sup> N404 <sup>6.52</sup>
	D97 <sup>3.26</sup> R169 <sup>ECL2</sup>
M <sub>2</sub> -ACh	A30 <sup>1.39</sup> N78 <sup>2.59</sup>
	N41 <sup>1.50</sup> D69 <sup>2.50</sup> Y440 <sup>7.53</sup>
	N58 <sup>2.39</sup> D120 <sup>3.49</sup> R121 <sup>3.50</sup> R135 <sup>ICL2</sup>
	Y104 <sup>3.33</sup> N404 <sup>6.52</sup>
M <sub>2</sub> -QNB	N41 <sup>1.50</sup> Y440 <sup>7.53</sup> D69 <sup>2.50</sup> S110 <sup>3.39</sup>
	N58 <sup>2.39</sup> D120 <sup>3.49</sup> R121 <sup>3.50</sup> R135 <sup>ICL2</sup>
	D103 <sup>3.32</sup> Y430 <sup>7.43</sup>
	N410 <sup>6.58</sup> Y104 <sup>3.34</sup> Y426 <sup>7.39</sup>
M <sub>3</sub> -Apo	D113 <sup>2.50</sup> N535 <sup>7.45</sup> N539 <sup>7.49</sup>
	Y529 <sup>7.39</sup> I222 <sup>ECL2</sup> Y506 <sup>6.51</sup> T502 <sup>6.47</sup> G528 <sup>7.38</sup>
	N102 <sup>2.39</sup> R165 <sup>3.50</sup> R179 <sup>ICL2</sup>
	Y148 <sup>3.33</sup> N513 <sup>6.58</sup>
M <sub>3</sub> -ACh	S108 <sup>2.45</sup> S154 <sup>3.39</sup>
	Y148 <sup>3.33</sup> Y529 <sup>7.38</sup> I222 <sup>ECL2</sup> Y506 <sup>6.51</sup>
	N102 <sup>2.39</sup> D164 <sup>3.49</sup> R165 <sup>3.50</sup> R179 <sup>ICL2</sup>
	T502 <sup>6.47</sup> G528 <sup>7.38</sup>
M <sub>3</sub> -QNB	N85 <sup>1.50</sup> D113 <sup>2.50</sup> Y543 <sup>7.53</sup>
	Y148 <sup>3.33</sup> Y529 <sup>7.38</sup> I222 <sup>ECL2</sup> W525 <sup>7.35</sup> Y506 <sup>6.51</sup>
	T502 <sup>6.47</sup> G528 <sup>7.38</sup> N513 <sup>6.58</sup>
	N102 <sup>2.39</sup> D164 <sup>3.49</sup> R165 <sup>3.50</sup>



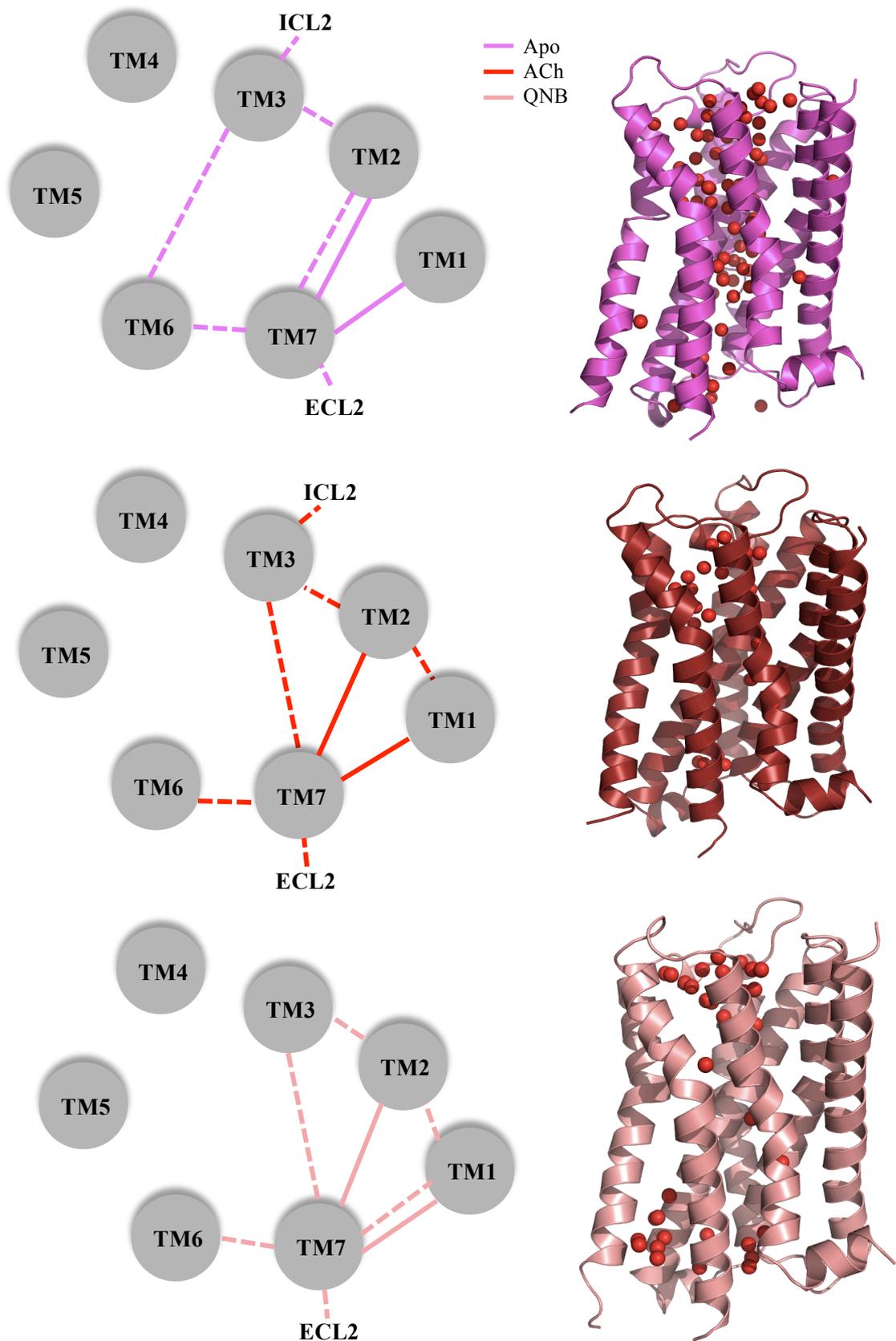
**Figure 4.16.**  $M_1$  hydrogen bonding networks. Solid line: inter-helical hydrogen bonds; dotted line: inter-helical hydrogen bonds mediated by water clusters. Internal water molecules were captured in 3-D diagrams for each of the systems, as red oxygen spheres.



**Figure 4.17.**  $M_1^*$  hydrogen bonding networks. Solid line: inter-helical hydrogen bonds; dotted line: inter-helical hydrogen bonds mediated by water clusters. Internal water molecules were captured in 3-D diagrams for each of the systems, as red oxygen spheres.



**Figure 4.18.**  $M_2$  hydrogen bonding networks. Solid line: inter-helical hydrogen bonds; dotted line: inter-helical hydrogen bonds mediated by water clusters. Internal water molecules were captured in 3-D diagrams for each of the systems, as red oxygen spheres.



**Figure 4.19.** M<sub>3</sub> hydrogen bonding networks. Solid line: inter-helical hydrogen bonds; dotted line: inter-helical hydrogen bonds mediated by water clusters. Internal water molecules were captured in 3-D diagrams for each of the systems, as red oxygen spheres.

#### 4.3.4 Molecular Mechanics-Poisson Boltzmann Surface Area (MM-PBSA)

To explore the differences in the binding of ACh and QNB to the receptors, the relative free energy of binding was computed from enthalpy contributions, using snapshots extracted from the trajectories, following the MM-PBSA approach (Srinivasan, Cheatham, Cieplak, Kollman, & Case, 1998; Kollman et al., 2000). The entropy contribution was not included in the calculations due to the high computational costs involved and difficulties in its prediction (Gohlke & Case, 2004; Hou, Wang, Li, & Wang, 2011; Homeyer & Gohlke, 2012; Yuriev & Ramsland, 2013). Limiting our considerations to enthalpic contributions was sufficient given the aim was to reveal interaction features (contribution of different energy terms and key contributing residues) rather than to obtain the absolute Gibbs energy, following the precedent of previous MM-PBSA studies (Aruksakunwong et al., 2007; Cang, Sponer, & Cheatham, 2011; Sanders, Wampole, Thakur, & Wickstrom, 2013; Shen et al., 2013; Xu et al., 2013; Zhu, Yu, Hao, Yang, & Yang, 2013).

Table 4.8 lists the contributions to the binding free energy for the different systems studied. In all cases, favourable contributions to the binding arose from VDW interactions and the non-polar part of the solvation free energy, as opposed to unfavourable total electrostatic contributions (EEL+EPB). According to the MM-PBSA calculations, ACh and QNB formed more energetically favourable complexes to the M<sub>3</sub> structure, compared to the others. Since the VDW and ENPOLAR components have a similar contribution to the binding, the total electrostatic contributions became the main cause of differences in the binding free energy.

**Table 4.8.** Relative binding free energies of complexes estimated using MM-PBSA for the different complexes studied.

Complex	EEL	vdW	EPB	ENPOLAR	$\Delta E_{\text{binding}}$
M <sub>1</sub> -ACh	-40.76	-21.29	50.39	-3.37	-15.03 (0.46)
M <sub>1</sub> *-ACh	-41.47	-22.74	51.74	-3.37	-15.84 (0.77)
M <sub>2</sub> -ACh	-22.40	-19.97	32.87	-3.53	-13.02 (0.48)
M <sub>3</sub> -ACh	26.71	-21.51	-19.80	-3.40	-18.01 (0.49)
M <sub>1</sub> -QNB	-49.22	-44.23	72.50	-4.94	-25.89 (0.68)
M <sub>2</sub> -QNB	-52.71	-43.83	71.28	-4.93	-30.19 (0.46)
M <sub>3</sub> -QNB	1.28	-43.72	14.94	-4.91	-32.41 (0.43)

The EEL and vdW represent the electrostatic and van der Waals contributions from MM, respectively. EPB stands for PB electrostatic contribution to the solvation free energy, and ENPOLAR is the nonpolar contribution to the solvation free energy.  $\Delta E_{\text{binding}}$  (in kcal/mol, binding energy neglecting the contribution of entropy) is the final estimated binding free energy calculated from the terms above. The values in brackets are the standard error of the calculation.

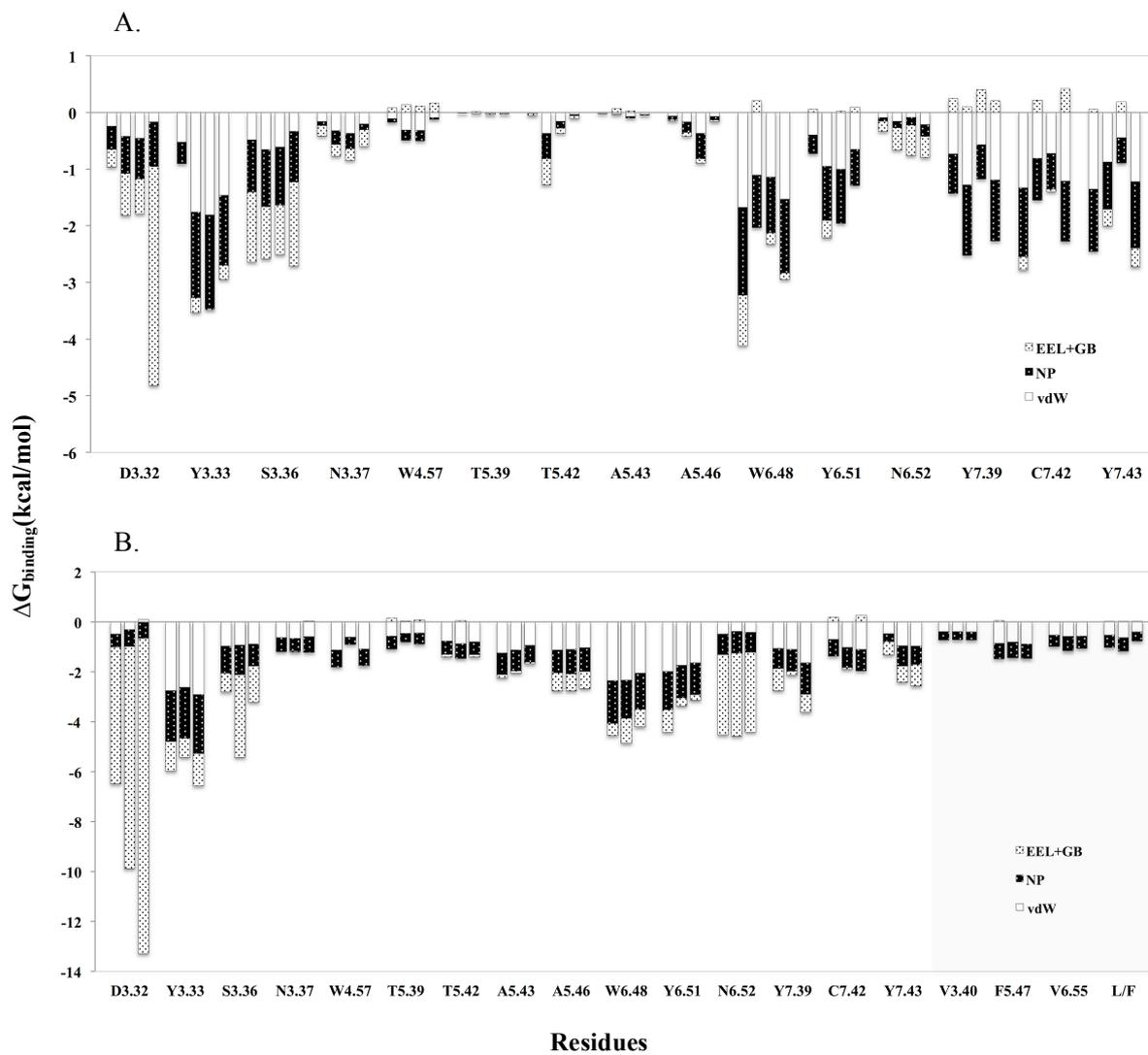
To gain further insight into the contribution of individual residues to binding, free energy decomposition was performed. Direct comparison of the contribution of residues towards the binding with the SDM/experimental data was not possible since the binding energy calculation does not include the entropy contribution, but it was qualitatively useful in a comparative way. Figure 4.20 depicts the decomposition of the binding free energy value on a per residue basis into vdW, nonpolar contributions to the solvation free energy (NP) and the sum of electrostatic interactions (EEL) and electrostatic contribution to the solvation free energy components (GB). The major favourable energy contributions in ACh binding originate predominantly from residues D<sup>3.32</sup>, Y<sup>3.33</sup>, S<sup>3.36</sup>, W<sup>6.48</sup>, Y<sup>6.51</sup>, Y<sup>7.39</sup>, C<sup>7.42</sup> and Y<sup>7.43</sup>. NP and vdW contributes mainly to the binding free energy, except for D<sup>3.32</sup>, where EEL+GB dominated. Even though ACh bound to the same binding site in all the muscarinic subtypes, it seems to be moving flexibly in the binding site (animations data not shown) and making contacts with the key residues with varying strengths. The choline head group was observed to be directed towards the conserved residue D<sup>3.32</sup>, while the tail groups hydrogen bonded with a different set of

residues and water molecules in a swapping pattern (Figure 4.21). The main binding free energy contributors ( $\geq 2.5$  kcal/mol) for the  $M_1$  receptor were  $S^{3.36}$ ,  $W^{6.48}$ ,  $C^{7.42}$  and  $Y^{7.43}$ , while for the  $M_3$  structure, there were additional contributions from  $D^{3.32}$  and  $Y^{3.33}$ . The main binding free energy contributors for the  $M_1^*$  and  $M_2$  structures were  $Y^{3.33}$  and  $S^{3.36}$ . However,  $T^{5.42}$  was found to give a relatively larger contribution to the free energy binding in the  $M_1^*$ -ACh system than in any other systems simulated and  $Y^{6.51}$  in the  $M_1^*$ -ACh system gave favourable EEL+GB components but not in the others (Figure 4.20). SDM studies have shown that  $T^{5.42}$  affects binding of most agonists but not of antagonists (Huang et al., 1999). As the agonist ACh is much smaller than the bulky antagonist QNB, interactions between  $T^{5.42}$  and ACh can only be seen when contraction of the ligand binding pocket has occurred as a result of an inward shift and rotation of TM5 (Allman et al., 2000), as exemplified in the agonist-bound  $\beta_2$ AR receptor structure (Rasmussen, DeVree, et al., 2011). Furthermore, mutation of  $Y^{6.51}$  has shown that it plays an important role in both the inactive and active state of muscarinic receptors, where the benzene ring of  $Y^{6.51}$  may form a cation-pi interaction with the positively charged head group of ACh that contributes to the activated state of the receptor but not the ground state (Ward et al., 1999).

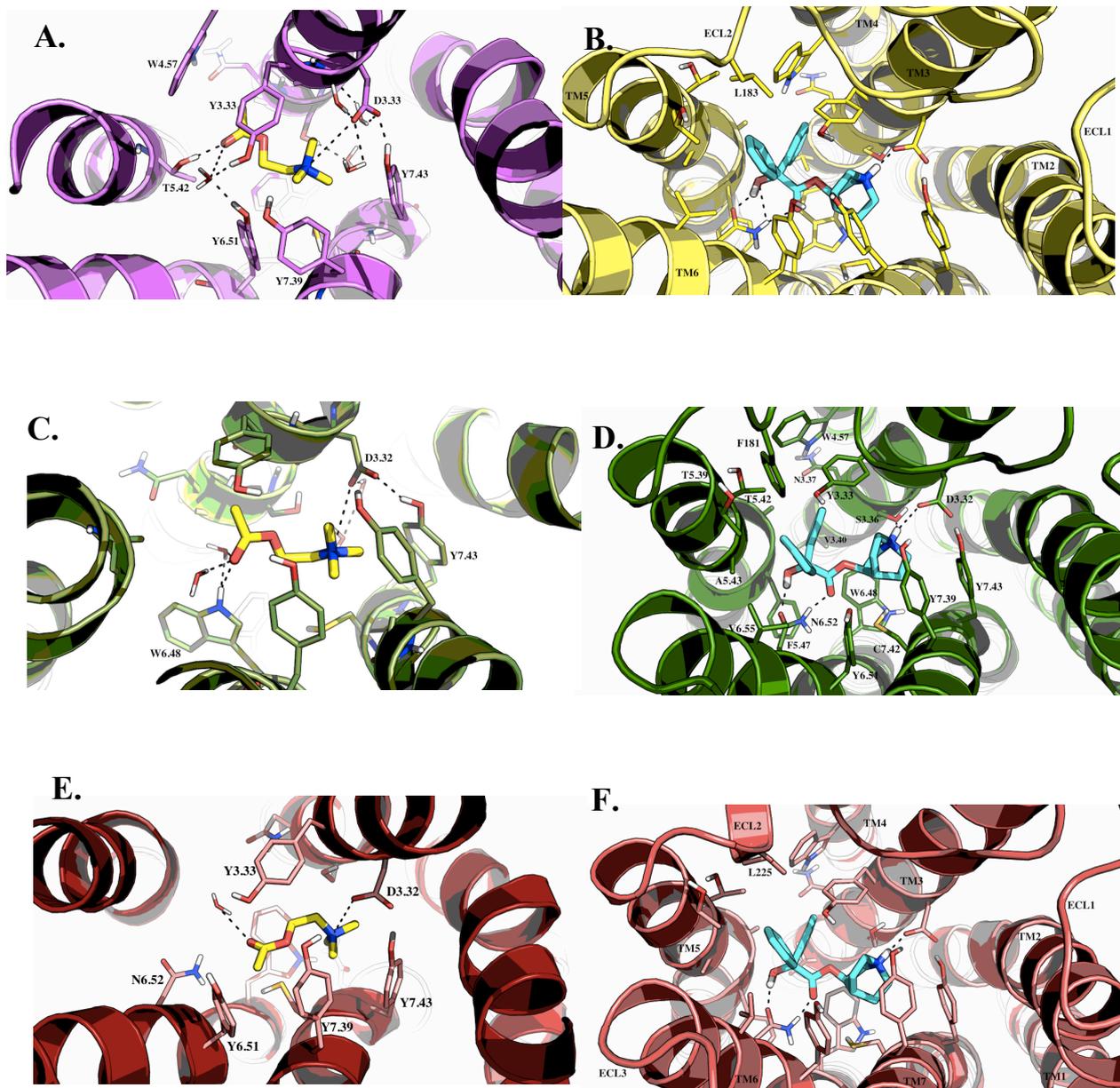
QNB bound to the same site as ACh and interacted with the same range of residues but more strongly (Figure 4.21). However, there were also residues, which only appeared to make contacts with QNB and not ACh (Figure 4.20). Among these residues, L/F refers to residues L183, F181 and L225 in the  $M_1$ ,  $M_2$  and  $M_3$  receptors, respectively, located in ECL2, and was found to give favourable vdW and NP components.  $D^{3.32}$  interacted more efficiently with QNB and made the largest contribution to the binding free energy in all the systems simulated. As seen with ACh binding, NP and vdW contributed mainly to the binding free energy, except for  $D^{3.32}$ , and  $N^{6.52}$ . The contributions of each

residue among the receptor subtypes simulated were quite even as QNB is bulkier than ACh and thus the range of orientations within the binding pocket is much more limited. N<sup>6.52</sup> made a larger contribution for QNB than for ACh, with EEL+GB being the dominant term, as it is constantly making hydrogen bonds with O1/O2 tail group of QNB (Tables 4.4 and 4.5). This is in agreement with SDM data where mutation of N<sup>6.52</sup> to alanine was shown to significantly alter the binding of QNB but have little effect on the binding of ACh (Blüml et al., 1994; Huang et al., 1999). It is therefore possible that N<sup>6.52</sup> is strongly hydrogen bonded with the ester group of QNB but not of ACh (or to a much lesser extent). Comparing the total binding free energy of QNB and ACh suggested that the latter binds rather weakly. Hence, it seems that binding of the bulky QNB efficiently locks the receptor in an inactive conformation and blocks activation-related contraction of the pocket.

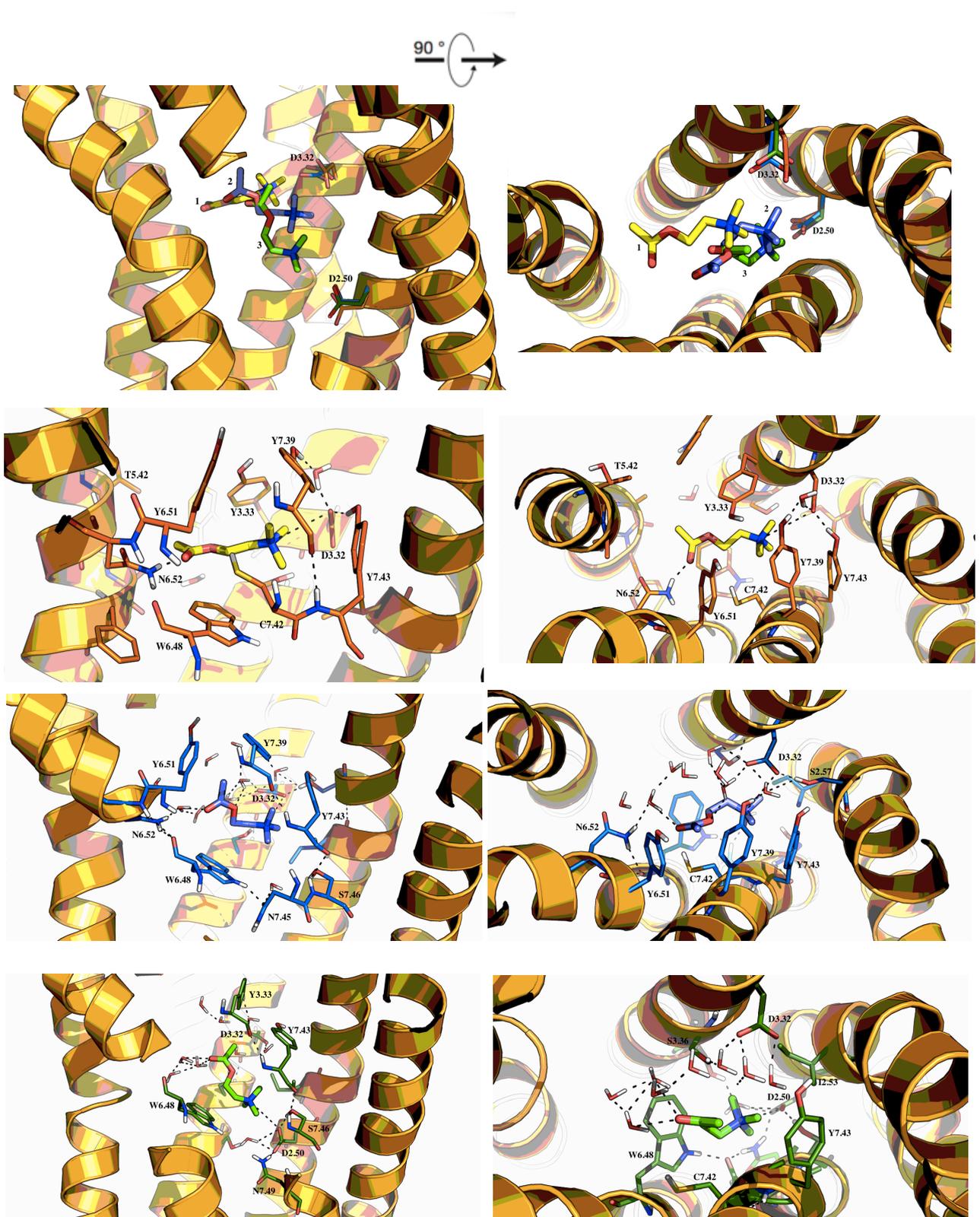
The contributions of D<sup>3.32</sup> to the binding of ACh and QNB in all the systems simulated were quite different, especially in the EEL+GB term. The distance between the OD1/OD2 of the D<sup>3.32</sup> side chain to the protonated nitrogen atom of ACh was in the range 4.1 – 4.7 Å for the M<sub>1</sub>\*, M<sub>2</sub> and M<sub>3</sub> structures, whereas it increased to ~6.3 Å for the M<sub>1</sub> structure, explaining the reason for the lower energy contribution from D<sup>3.32</sup>. Unexpectedly, the ACh bound to M<sub>1</sub> gradually fell into a sub-pocket near to D<sup>2.50</sup> and some distance away from D<sup>3.32</sup> (Figure 4.22). This is also confirmed in the pocket analyses (see section 4.3.5). Similarly, the increasing strength of the contributions of D<sup>3.32</sup>, as seen when comparing the M<sub>1</sub>-, M<sub>2</sub>- and M<sub>3</sub>-QNB systems, correlated with the distance between OD1/OD2 of the D<sup>3.32</sup> side chain to the protonated nitrogen atom of QNB decreasing successively from ~4.5 to ~3.6 Å.



**Figure 4.20.** Decomposition of binding energy on a per-residue basis into contributions from EEL+GB, NP, and vdW, upon the binding of **A.** ACh and **B.** QNB. First bar of each of the residues corresponding to  $M_1$ , follow by  $M_1^*$ ,  $M_2$ , and  $M_3$  of the subsequent bars in A and B, except that in B, which  $M_1^*$  is excluded.



**Figure 4.21.** ACh and QNB interactions with the receptors. **A.**  $M_1^*$ -ACh, **B.**  $M_1$ -QNB, **C.**  $M_2$ -ACh, **D.**  $M_2$ -QNB, **E.**  $M_3$ -ACh, and **F.**  $M_3$ -QNB.



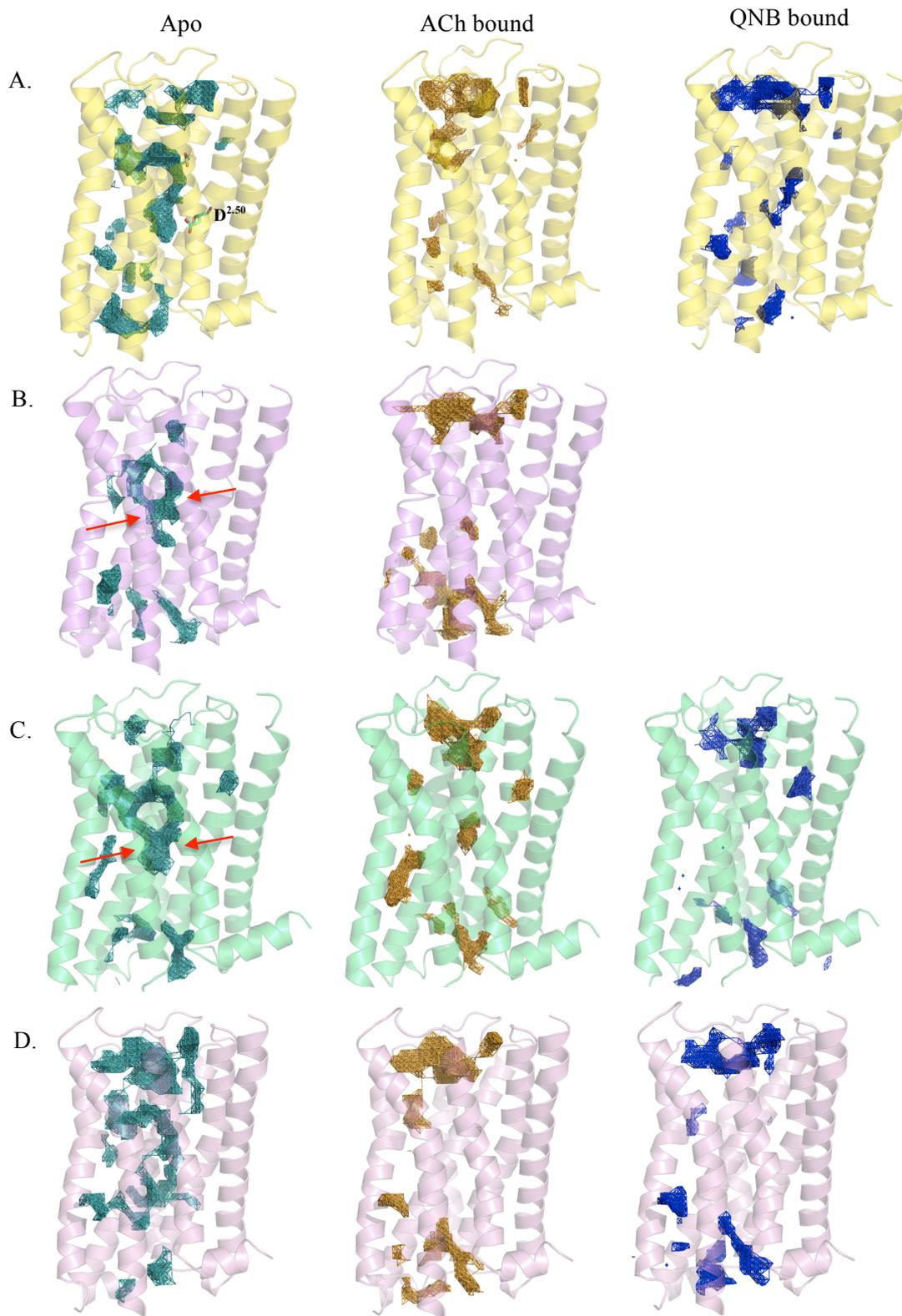
**Figure 4.22.** Transition of ACh from originally bound spot 1 (yellow) at the orthosteric site, to 2 (blue) and 3 (green) in M<sub>1</sub>-ACh simulations and establishes interactions with different set of residues.

#### 4.3.5 Pocket analysis

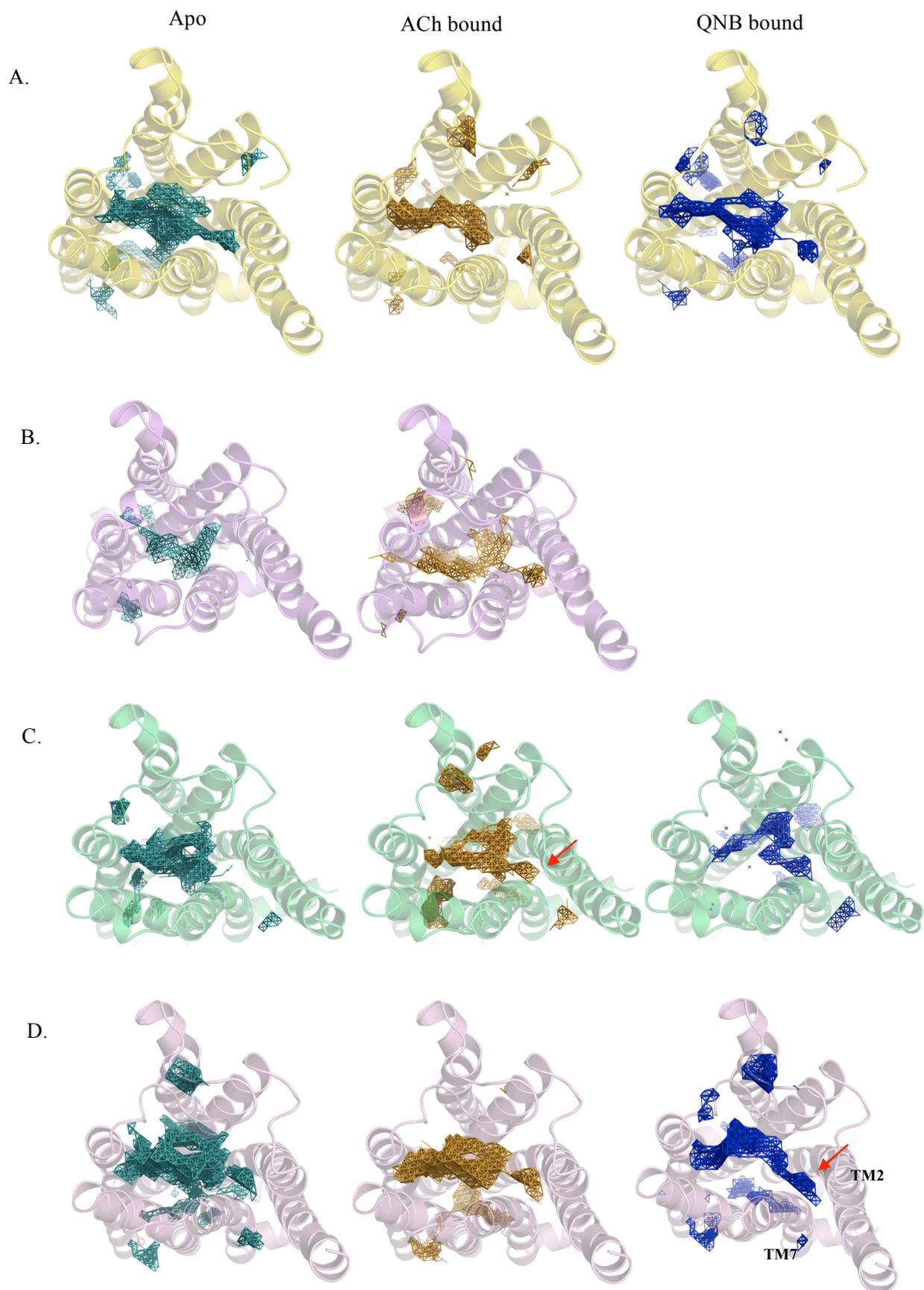
MDpocket was used to detect cavities or pockets formed in the receptors throughout the MD simulations. From the analysis, pockets formed at 50% frequency from all the apo systems were very similar, as shown in Figures 4.23 and 4.24. The pockets formed on the extracellular side and internal channel where the orthosteric binding pockets reside. The orthosteric pocket extended downwards and towards TM6 and TM2, TM6, and TM4, for the M<sub>1</sub>, M<sub>2</sub> and M<sub>3</sub> structures, respectively. It is also obvious that the M<sub>1</sub> structure has a pocket extending towards residue D<sup>2.50</sup>, into which ACh was found to be moving and making interactions with D<sup>2.50</sup>, I<sup>2.53</sup>, S<sup>3.39</sup>, L<sup>6.46</sup>, N<sup>7.45</sup> and S<sup>7.46</sup> (Figures 4.22 and 4.23). The pocket extensions for the M<sub>1</sub>\* and M<sub>2</sub> structures in the apo form were somewhat different in that the extension was from both sides (Figure 4.23). The pockets formed after the binding of ACh and QNB were also compared in order to search for possible allosteric/secondary binding sites. The holo forms of the receptors showed more pockets being formed on top of the orthosteric site compared to the apo forms, whereas almost all the pockets in the internal channel disappeared. For M<sub>1</sub>- and M<sub>2</sub>-ACh a pocket was formed between TM3 and TM4, underneath the first half of ECL2, which was in the same location as a pocket that was found for M<sub>3</sub>-apo (Figure 4.23). However, these pockets were rather small for binding purposes compared to the pockets found on top and to the left of the orthosteric sites. A pocket extension from the top toward the cavity between TM2 and TM7 was also found for the M<sub>2</sub> and M<sub>3</sub> structures and was more noticeable in the holo forms of receptors. The binding of QNB formed pockets on top of the orthosteric site as well, but to a lesser extent compared to apo or ACh-bound structures. It is known that many GPCRs enhance agonist binding affinity through an allosteric mechanism (Shoichet & Kobilka, 2012). Hence this explained the observations above where the pockets formed on top of the orthosteric site upon the

ACh binding are probably required for allosteric action or act as intermediate/secondary binding sites.

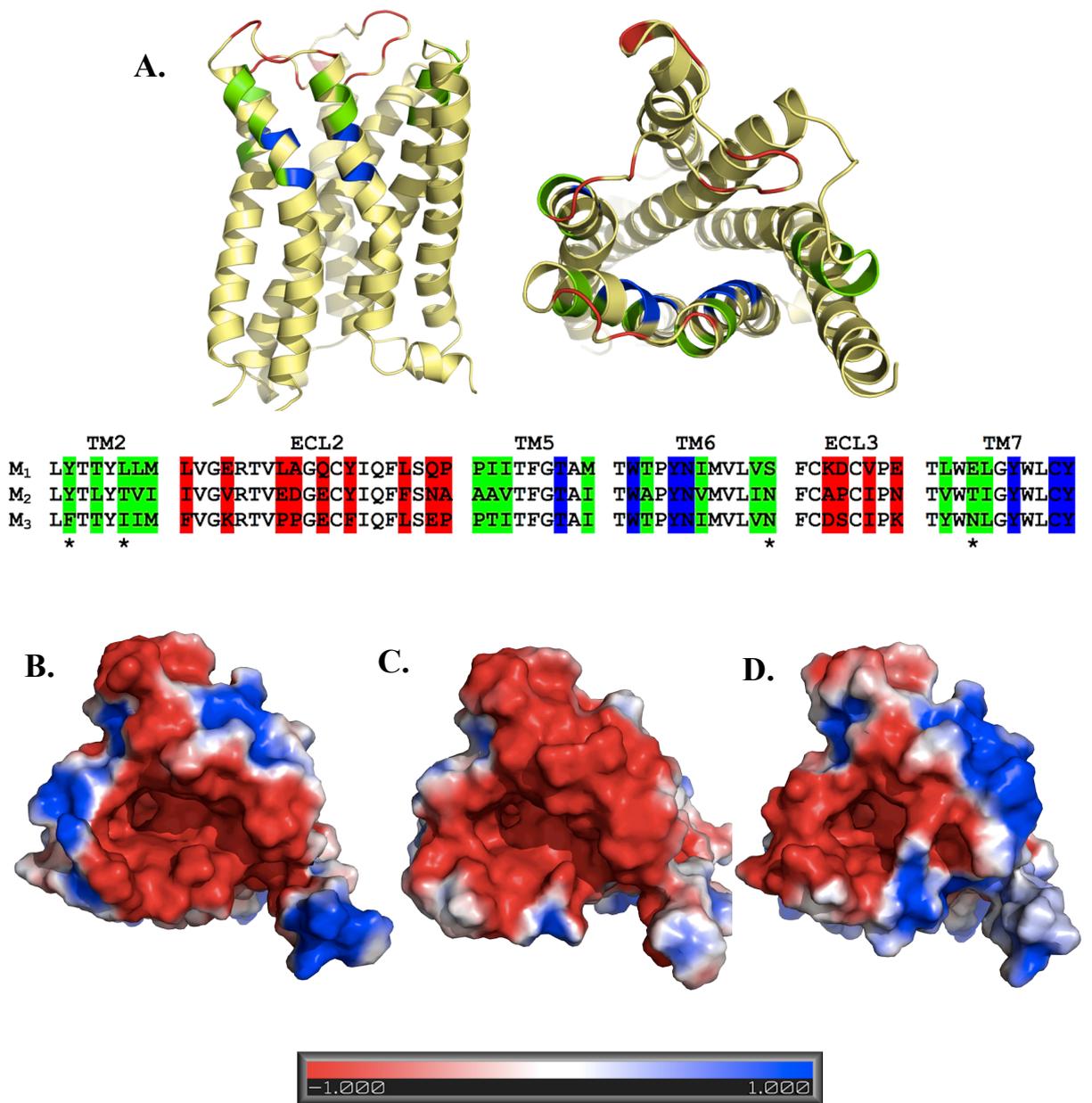
Inspecting the receptor sequences and structures adjacent to the secondary pockets also showed that, apart from those of the non-conserved ECL2 and ECL3, there were 4 residues (Y/F<sup>2.61</sup>, L/T/I<sup>2.65</sup>, S/N<sup>6.58</sup>, E/T/N<sup>7.36</sup>) whose side chains were positioned facing towards the internal channel (Figure 4.25). These residues with different physicochemical side chain properties, which reside on the extracellular portion of TM2, TM5, TM6 and TM7, may influence the overall packing of the receptors and initial recognition. Computation of the electrostatic potential distribution of the receptors using APBS revealed that, they are not identical, especially at the extracellular surface, despite the high sequence similarities within the subtypes. The differences in the surface charges may arise from the non-conserved extracellular loops and non-identical residues. These secondary pockets with subtle differences of residues may account for subtype selectivity and are worth further exploration. It has been shown for the D3 receptor structure that an extracellular extension of the core binding pocket forms a second binding site which account for D2 and D3 subtype selectivity (Chien et al., 2010). The second binding pocket in D3 comprising the junction of ECL1 and ECL2 and the interface of TM2, TM7 and TM1, is similar to that identified in this study for the M<sub>2</sub> and M<sub>3</sub> structures. The presence of a secondary binding site located between the ECL2, ECL3 and TM7 has also been observed for  $\beta$ -adrenergic receptors, where ligands are transiently retained in the early steps of the binding process and which is involved in the ligand binding/unbinding pathway (González, Perez-Acle, Pardo, & Deupi, 2011). It has been also demonstrated for M<sub>2</sub> and M<sub>3</sub> receptors that ECLs played an important role in determining the kinetic selectivity of a drug through secondary site binding (Kruse et al., 2012; Dror et al., 2013).



**Figure 4.23:** Pockets found throughout the simulations for **A.** M<sub>1</sub>, **B.** M<sub>1</sub><sup>\*</sup>, **C.** M<sub>2</sub>, and **D.** M<sub>3</sub>. It is clear that there is a pocket extension from the primary orthosteric pocket in M<sub>1</sub> toward D2.50, where different patterns were found on other receptor subtypes (pointed by red arrow). Secondary pockets were found on top of the orthosteric site, involving ECLs.

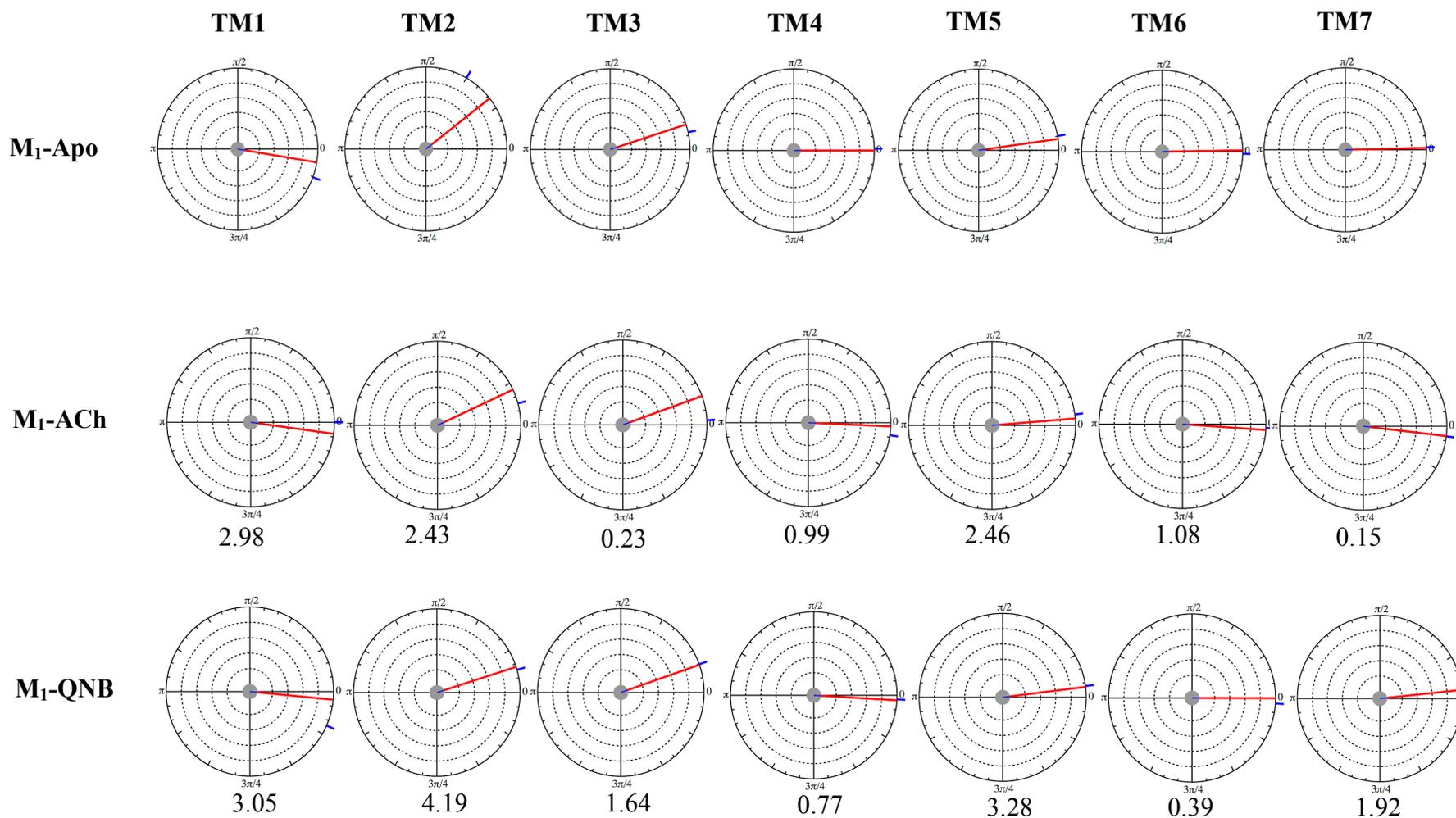


**Figure 4.24:** Top view of the pockets found throughout the simulations for **A.** M<sub>1</sub>, **B.** M1\*, **C.** M<sub>2</sub>, and **D.** M<sub>3</sub>. Pocket extension from the secondary pocket was shown (pointed red arrow).

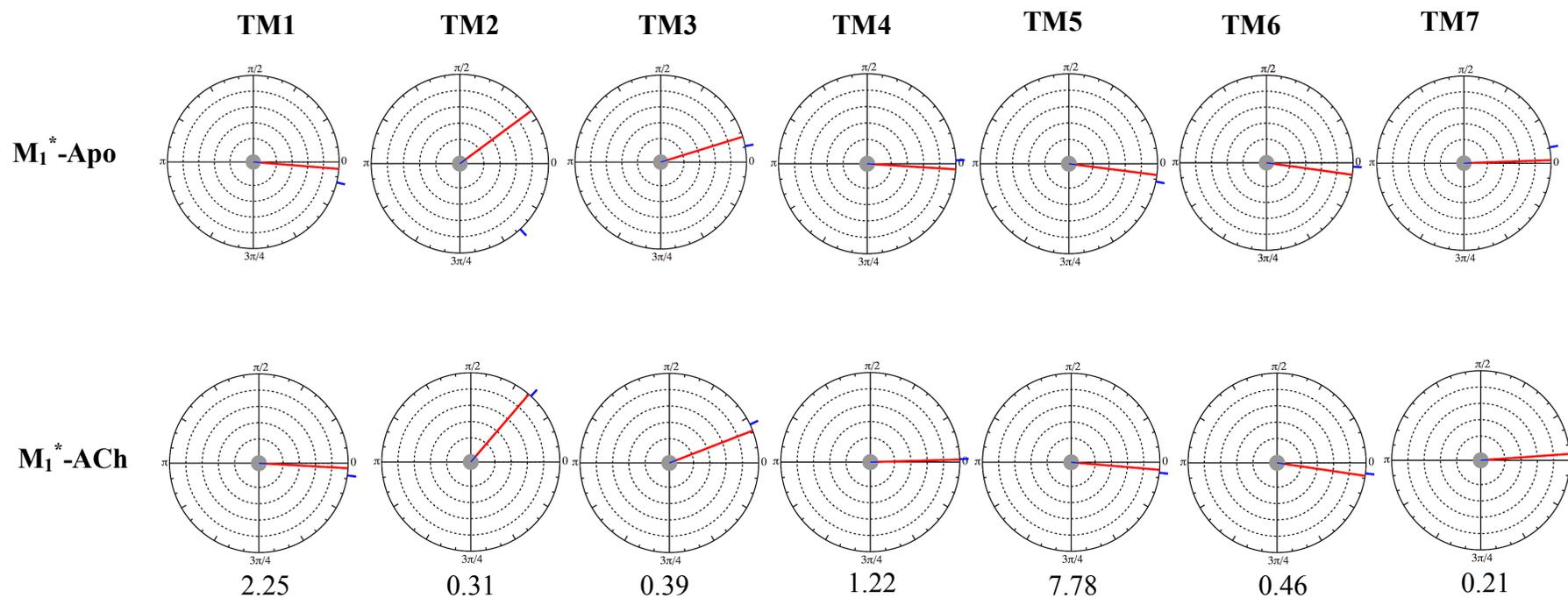


**Figure 4.25:** **A.** Graphical and sequence representations of the non-conserved/identical residues at the extracellular vestibule of the receptors. Green: non-identical residues in the TM helices, red: non-identical residues in the loops, blue: highly conserved residues, and \*: residues with side-chains facing the inner TM core. **B-D:** Electrostatic charge distribution calculated by APBS for M<sub>1</sub>, M<sub>2</sub>, and M<sub>3</sub>, respectively.

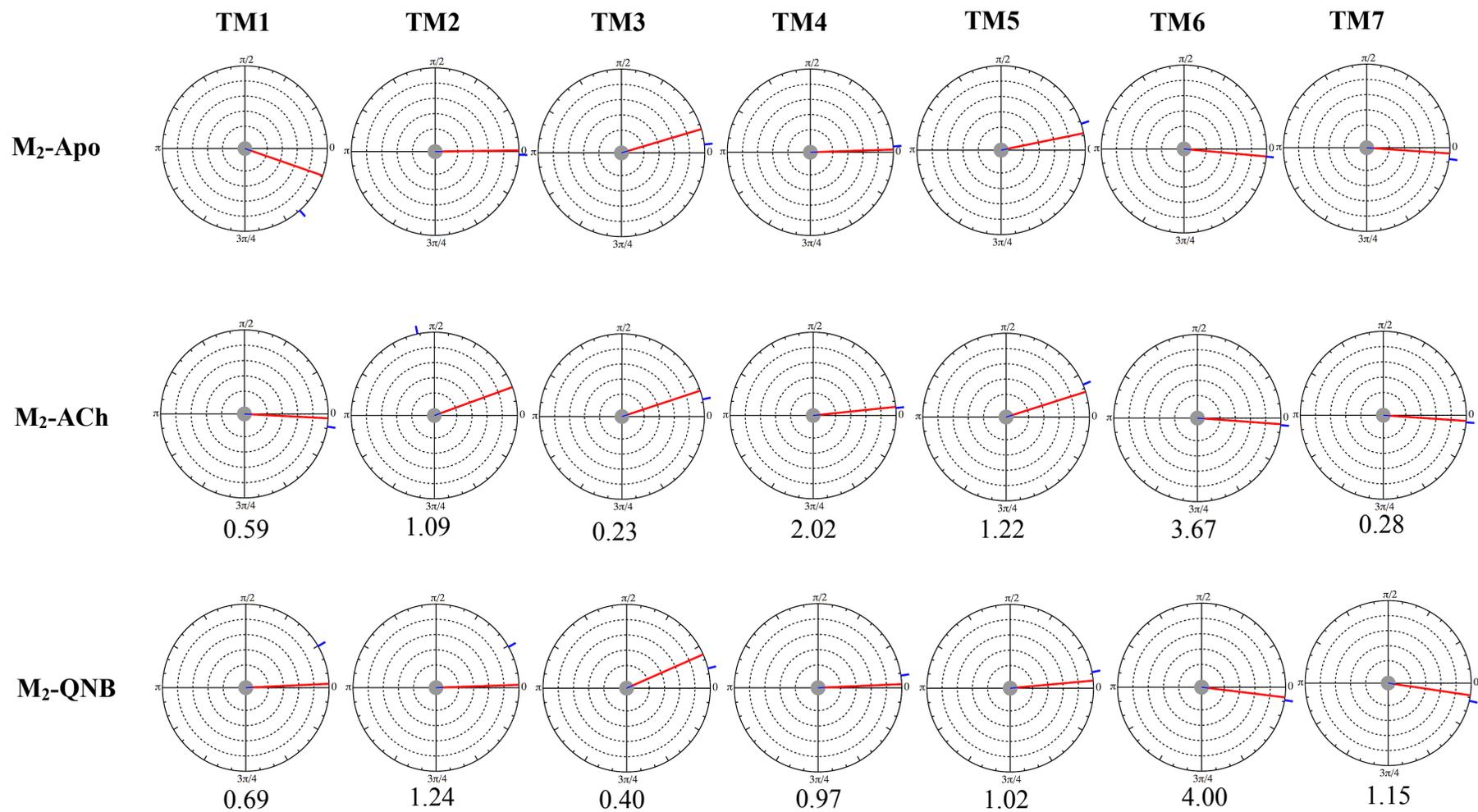
Apart from assessing the pockets formed throughout the simulations, rotations around the helical axes and helix tilt angles with respect to bilayer normal were also investigated for each of the systems to elucidate the structural and dynamic effects induced by ligands on the receptor features. For each of these, the average values from the trajectories of the apo and holo forms of the receptors were calculated using TRAJELIX (Figures 4.26 to 4.29). Larger helical rotations ( $>10^\circ$ ) for the ACh- and QNB-bound structures, with respect to their apo counterparts, were observed in TM1 (M<sub>2</sub>- and M<sub>3</sub>-QNB) and TM2 (M<sub>1</sub>-, M<sub>1</sub>\*-, and M<sub>2</sub>-ACh), helices that are not involved in the binding of either ACh or QNB. Regarding the tilt angles, changes induced by ACh and QNB were smaller ( $<5^\circ$ ), except for TM5 in the case of the M<sub>1</sub>\* structure. Binding of ACh induced larger changes in these parameters compared to QNB, although ACh is structurally smaller than QNB. This may be explained by the greater ability of ACh, observed in this study, to move flexibly within the binding pocket and interact with surrounding water molecules, compared to the less flexible and bulkier QNB. Since no significant changes in the secondary structure of the helices were observed, the effect of the presence of ACh and QNB on the receptors seems to be more on the dynamic aspect than the structural.



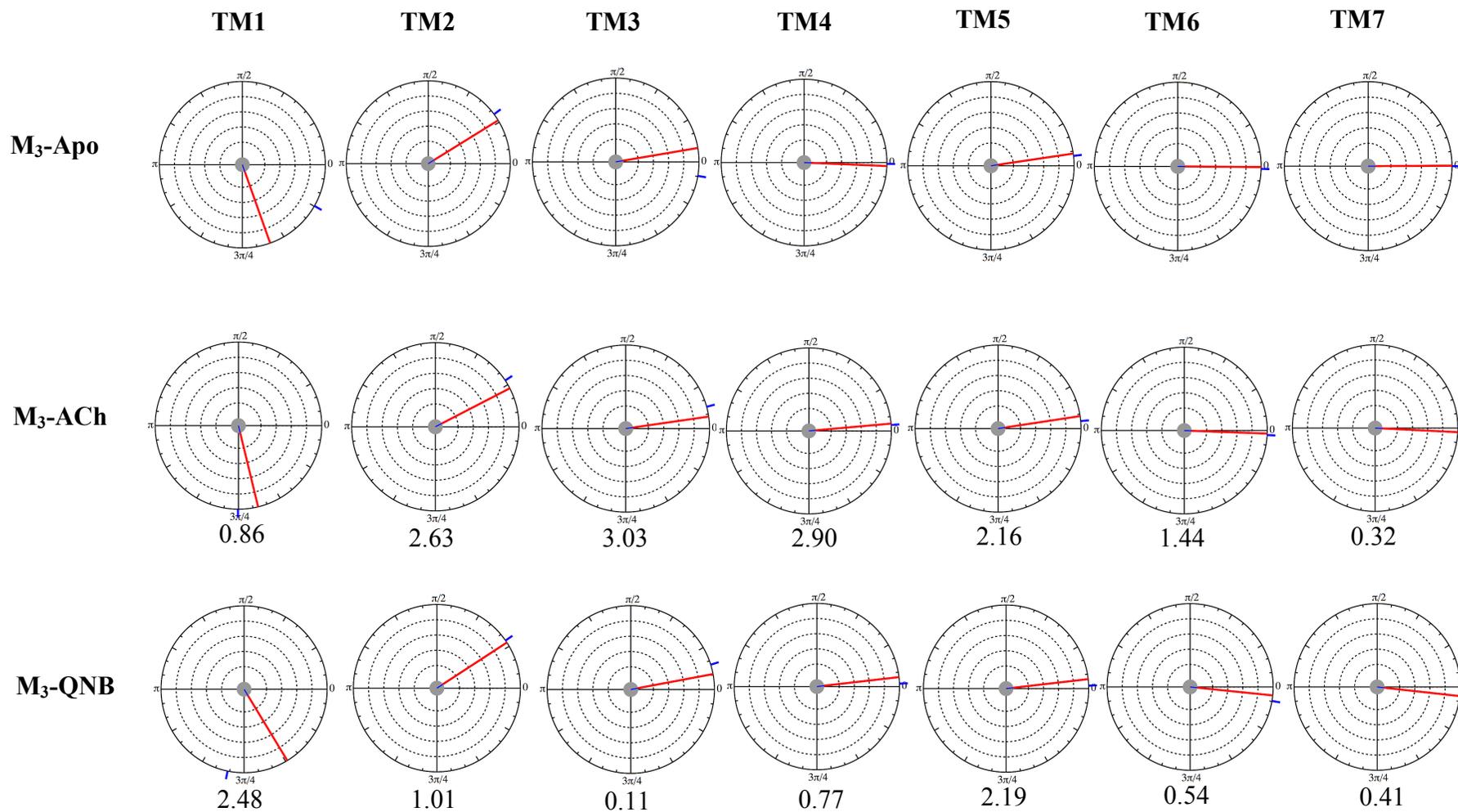
**Figure 4.26.** Dial plots of the average helix rotation angles in M<sub>1</sub> system during the MD simulations. Average helix Z-tilt angle (°) with respect to the apo receptors for each of the TM helices is shown underneath the dial plots.



**Figure 4.27.** Dial plots of the average helix rotation angles in  $M_1^*$  system during the MD simulations. Average helix Z-tilt angle ( $^\circ$ ) with respect to the apo receptors for each of the TM helices is shown underneath the dial plots.



**Figure 4.28.** Dial plots of the average helix rotation angles in M<sub>2</sub> system during the MD simulations. Average helix Z-tilt angle (°) with respect to the apo receptors for each of the TM helices is shown underneath the dial plots.



**Figure 4.29.** Dial plots of the average helix rotation angles in M<sub>3</sub> system during the MD simulations. Average helix Z-tilt angle (°) with respect to the apo receptors for each of the TM helices is shown underneath the dial plots.

#### 4.4 Conclusion

Using conventional MD simulations, this study provides detailed information about the structural and dynamic features of M<sub>1</sub>, M<sub>2</sub> and M<sub>3</sub> receptors in the absence and the presence of an agonist (ACh) or an antagonist (QNB), to help comprehend the relationship between receptor dynamics and function. Despite the fact that the different receptor subtypes are highly homologous, the simulations highlight important differences in their dynamic behaviour.

Larger and varied movements were observed in the apo form, especially at the extracellular site, and overall dynamics were suppressed upon ACh or QNB binding. Several structural rearrangements were observed upon the binding of ligands, including stabilization of the highly mobile ECLs, functionally relevant mobility of a number of TM helices such as TM3, TM4, and TM6 and the reorganization of the hydrogen bonding networks upon the binding of ACh and QNB. Although ACh and QNB bound to the same binding pockets, D3.32 was found to make stronger interactions with the QNB head group and N6.52 made a larger contribution for QNB than for ACh. Formation of secondary pockets above the orthosteric binding pocket upon the binding of ACh were observed, suggesting possible allosteric action for the receptor. Subtle differences in the residues involved in forming these pockets for the different receptor subtypes point to possibilities for subtype selectivity.

During the course of the simulations, no ionic lock formation was observed in the apo or antagonist bound receptor and the structures in complex with ACh after the simulations did not reveal significant outward kink of the cytoplasmic termini of TM6 (although TM6 was observed to be highly dynamic in PCA), as observed in the few crystal structures of activated GPCRs that have so far been made available.

Nevertheless, it is interesting to note that ACh induced more pronounced structural perturbations of the structures than QNB, despite being a smaller molecule. Although the simulation times were relatively short to explore the conformational changes involved in the receptor activation or to simulate receptor activation from an inactive receptor structure, they are sufficient given that the aim of the study was simply to explore the dynamics and stability of different receptor subtypes in apo and ligand bound forms, following the precedent of previous MD works (Chen, Cao, Chen, & Chen, 2013; Balupuri & Sobhia, 2014; Helal, Darwish, & Hammad, 2014; Patra, Maharana, Dehury, & De, 2014).

In agreement with the previously reported simulations of GPCRs, a full agonist stabilizes the receptor and results in an near agonist-bound conformation but the binding of the ligand is not sufficient for initialization of the major structural changes observed upon comparison of the crystal structures of  $\beta_2$ AR receptors in their active and inactive sub-states (Martínez-Archundia & Correa-Basurto 2013, Miao et al. 2013, Novikov et al. 2014, Vaidehi et al. 2014). Miao et al. (2013) have showed that the M<sub>2</sub> receptor remained inactive even after microsecond timescale conventional MD. The binding of agonists therefore seems to open a new path in the conformational space of the receptor, through which the receptor acquires a higher probability compared to the apo form to sample active sub-states.

During the preparation of this thesis, the first structure of an activated mAChR was made available (Kruse et al. 2013). The structure shows activation-related structural changes and highlights the important role of TM6, which, due to its mobility, structurally links up three regions of the receptor: the extracellular vestibule, the orthosteric binding pocket, and the intracellular surface. Unexpectedly, the bound ligand

(iperoxo) established contacts, which resembled those observed with QNB. Despite the fact that GPCRs probably have common activation mechanisms, the accompanying structural changes are still dependent on the exact nature of the ligand involved, and so other agonists, including ACh, might show differences compared to those observed for the iperoxo-bound structure (Kruse et al. 2013).

## CHAPTER 5

### IDENTIFICATION OF POTENTIAL M<sub>1</sub> mAChR SELECTIVE LEADS: VIRTUAL SCREENING USING MULTIPLE RECEPTOR MODELS AND STRUCTURES

#### 5.1 Introduction

Structure-based drug discovery has made promising developments in the past 30 years, benefiting from the advancement in high performance computing. Identification of ligands through docking/virtual screening has been widely used and relatively successful (Lounnas et al., 2013), where each molecule from a library is docked into a binding site then scored and ranked accordingly to its complementarity to the receptor binding site. Selection of molecules is carried out based on ranking and is subsequently tested in experiments. Despite its effectiveness and widespread practice, the problem of accurately and efficiently modeling receptor flexibility and conformational heterogeneity is difficult to overcome, as virtual screening involves docking of multiple diverse ligands, each of them potentially having a different preference to a distinct receptor conformation. Because receptors are known to be dynamic and flexible, functionally relevant receptor conformations influence ligand binding and determine the success of structure based virtual screening. The paradigm shift from key and lock to induced fit concept (Koshland, 1995) demands new methods that can incorporate receptor plasticity. Ensemble dockings (combined used of multiple experimental structures or homology models) have shown to improve enrichment factors and ability to retrieve diverse set of ligands (Fan et al., 2009; Novoa, Pouplana, Barril, & Orozco, 2010; Vinh, Simpson, Scammells, & Chalmers, 2012; Xu & Lill, 2013). Induce-fit

docking (IFD), is another novel method that takes into account both the ligand and receptor flexibility and can be used to generate receptor ensembles. The use of IFD ensembles in virtual screening has proved to significantly increase enrichment factors (Sherman, Beard, & Farid, 2006) and has gained in popularity in modeling GPCRs (McRobb, Capuano, Crosby, Chalmers, & Yuriev, 2010; Vilar et al., 2011; Kołaczkowski, Bucki, Feder, & Pawłowski, 2013; Pala et al., 2013; Thomas et al., 2014).

The challenge to solve the 3-D structure of receptors, especially GPCRs, has not limited the use of virtual screening in drug discovery, as homology models have been used in the absence of crystal structures. In fact, there is an increasing number of studies using GPCR homology models for structure based virtual screening (Katritch, Rueda, Lam, Yeager, & Abagyan, 2010; Dong et al., 2013; Heifetz et al., 2013; Kiss, Jójárt, Schmidt, Kiss, & Keserű, 2014; Vass, Schmidt, Horti, & Keseru, 2014). Recently GPCR-based drug discovery has gained intense interest due to advancements in high-resolution structure determinations that have successfully produced 3-D structures of the receptors. Not only do these structures provide information directly applicable to drug discovery, making identification of novel ligands through virtual screening effectively possible (de Graaf & Rognan, 2008; Kolb et al., 2009; Katritch, Jaakola, et al., 2010; Mysinger et al., 2012; Kruse, Weiss, et al., 2013; Weiss et al., 2013), they also serve as a platform to rectify and improve theoretical structures derived from homology modeling (Michino et al., 2009; Kufareva et al., 2011; Kufareva, Katritch, Stevens, & Abagyan, 2014). The existence of highly homologous muscarinic acetylcholine receptor (mAChR) subtypes have hindered the drug discovery and development progress such that none of the reported muscarinic agonists and antagonists are particularly selective for a certain subtype, hence causing undesirable side effects. However, in the past few years, driven

by continuing efforts, several selective muscarinic ligands have been developed to preferential recognize distinct mAChR subtypes. These ligands exert their effects as allosteric agonists or positive/negative allosteric modulators (PAMs/NAMs), and some have bitopic characteristics (Kuduk & Beshore, 2012; Davie, Christopoulos, & Scammells, 2013; Foster, Choi, Conn, & Rook, 2014; Nickols & Conn, 2014). The recently solved crystal structures of M<sub>2</sub> and M<sub>3</sub> receptors (Haga et al., 2012; Kruse et al., 2012; Kruse, Ring, et al., 2013) have made virtual screening and subtype selectivity assessment possible.

In the absence of an M<sub>1</sub> crystal structure, several M<sub>1</sub> models based on different templates and modifications were developed. Together with the M<sub>2</sub> and M<sub>3</sub> crystal structures, a total of 12 targets were used in these virtual screening studies to identify potential M<sub>1</sub> selective ligands. A small subset of a compound library consisting of 5915 small molecules was docked to each of the targets. By comparing the docking results from the 12 targets, 19 compounds that showed the largest differences in terms of GScore and ranking were selected. Among these 19 compounds, 7 were bound to both the orthosteric/primary site and secondary site; 11 docked to the orthosteric site and 1 was found to preferably sit on top of the orthosteric site, between the extracellular loop (ECL) 2 and 3. The molecular properties of the compounds were accessed and 8 of the compounds showed characteristic that do not conform to the central nervous system (CNS) drug properties. The final selection of the ligands was subsequently subjected to Prime MM-GBSA calculations.

## **5.2 Methods**

Following the methods described in Chapter 3 (section 3.2.2, except 3.2.2.3), further M<sub>1</sub> and M<sub>3</sub> homology models based on the newly published agonist bound M<sub>2</sub> crystal

structure (PDB code: 4MQS) were developed. The crystal structures of antagonist bound M<sub>2</sub> receptor (PDB code: 3UON), agonist bound M<sub>2</sub> receptor (4MQS), and inverse agonist bound M<sub>3</sub> receptor (PDB code: 4DAJ) together with the models developed, underwent IFD using ACh to sample different binding site residue configurations to those found in the originally crystal structures. After multiple iterations of docking and binding site refinement from the IFD, complexes were chosen by visual inspection on the basis of showing the expected ligand-receptor interactions, where the amine of the docked ACh was directed towards the conserved D<sup>3.32</sup> and the side chains of the orthosteric site residues were all facing inwards towards the inner channel of the transmembrane (TM) region. The chosen IFD structures were then subjected to enrichment dockings as described in Chapter 3 (section 3.2.2.4) and 12 structures (4 for each subtype) that showed good enrichment factor and BEDROC values were selected for the subsequent virtual screening and selectivity assessment.

Prior to the virtual screening, six known M<sub>1</sub> selective ligands were chosen and were docked to the receptors as a benchmark and starting point to access the binding of these ligands to the M<sub>1</sub>, M<sub>2</sub>, and M<sub>3</sub> receptors. The small molecule library was comprised of a small subset from the drug-like ZINC compounds and Schrodinger drug-like ligand decoys set, which were previously used in the enrichment studies (decoy set I, II, and III). An additional set of ligands contain an amino group (provided by Dr. David Chalmers from Monash University, Melbourne) was also retrieved from ZINC database, using the same methods as in the retrieval of decoy set III (see section 3.2.2.4). The properties of the ligand library were calculated using Discovery Studio v3.1 (Accelrys Inc.) with predefined parameters and are shown in Table 5.1. Glide was used to perform virtual screening on 5915 small molecules against 12 mAChR targets, where the ligands were flexibly docked into each of the receptor structure with identical Glide docking

settings and parameters. Glide score (GScore), a modified and expanded version of the ChemScore scoring function, was used to score and rank the ligands. For each of the ligands, the pose that gave the maximum score when bound to a particular receptor subtype was selected, allowing each of the docked ligand to choose its preferred receptor conformations (among 4 structures for each subtype). The molecules with the largest GScore difference in favour of the M<sub>1</sub> receptor were identified and inspected.

All-atom molecular models were generated and protein and ligand preparations, IFD, docking, enrichment studies, and ligand physicochemical properties inspection were performed using Schrödinger suite 2011 (Schrödinger LLC, New York, USA) with default settings and parameters, unless stated otherwise (Friesner et al., 2004; Halgren et al., 2004; Sherman, Day, Jacobson, Friesner, & Farid, 2006). The detailed methodology is previously described in Chapter 3 (section 3.2.2). Visual inspections and binding modes analyses were carried out with the aid of Maestro v9.2 (Schrödinger LLC) and PyMOL v1.6.9.0 (Schrödinger LLC) was used to produce 3-D figures. For ease of comparison and standardization, residues are labeled using Ballesteros-Weinstein nomenclature, as a superscript extension (Ballesteros & Weinstein, 1995).

**Table 5.1.** Distribution of molecular properties for ligand library used in the virtual screening.

Property	Ligand library		
	Mean	Median	SD
Molecular weight	264.2	249.3	99.6
ALogP	1.6	1.7	1.8
Fractional polar surface area	0.3	0.2	0.1
Hydrogen bond acceptor	3.2	3.0	1.8
Hydrogen bond donor	1.1	1.0	0.9
Rotatable bonds	3.7	3.0	2.4
Number of aromatic rings	1.6	1.0	1.1
Number of rings	2.3	2.0	1.4

### 5.2.1 Estimation of binding free energy

The Prime MM-GBSA module of the Schrödinger suite (Schrödinger LLC, New York, USA) was used to predict the binding free energy ( $\Delta G_{\text{Bind}}$ ) for a chosen set of ligand-receptor complexes using the following equation:

$$\Delta G_{\text{Bind}} = G_{\text{complex}} - (G_{\text{protein}} + G_{\text{ligand}})$$

where  $G_{\text{complex}}$  is the optimized free energy for the complex,  $G_{\text{protein}}$  and  $G_{\text{ligand}}$  are the optimized free energy for the free protein and free ligand, respectively. Each energy term was calculated by a combination of molecular mechanics energy, implicit solvation energy and surface area energy. Residues in the binding pocket of the receptor were treated as flexible.

## 5.3 Results and discussions

### 5.3.1 Model generation and selection

A total of 12 targets consisting of different mAChR subtypes selected based on their enrichment docking performance were used in the virtual screening studies (Tables 5.2 to 5.4). The enrichment factors and BEDROC values for the  $M_1$  models are comparable or better than the  $M_2$  and  $M_3$  crystal structures. The performance of the  $M_1$  models has been discussed previously in Chapter 3 (section 3.2.3.2 and 3.2.3.3). The ability of the new  $M_1$  model (model 2) generated based on the agonist bound  $M_2$  crystal structure (4MQS) to separate actives from decoys is comparable with the models 1, 1A1, and 1A2 (Table 5.2). Similar with model 1A2, it fails to accommodate all the agonists, due to its contracted binding pocket. The superposition of models 1A1 and 1A2 to the agonist bound  $M_2$  crystal structure showed good overlap, with RMSD less than 1 Å (0.945 and 0.994 Å, respectively). IFD improved the dockability of the structure significantly, where all the crude models/structures without IFD optimization performed badly in the enrichment study. It is also apparent that the agonist bound

models/structures were not able to dock all the agonists due to their contracted binding pockets (Tables 5.2 to 5.4).

**Table 5.2.** Enrichment results for M<sub>1</sub> homology models.

M <sub>1</sub> models based on inverse agonist bound M <sub>3</sub> crystal structure (4DAJ) and modified using agonist bound $\beta_2$ AR crystal structure (3SN6)						
Model	EF2%	EF5%	EF10%	BEDROC	AUC	# of bound active
Crude	2.0	1.6	1.2	0.097	0.56	51
<b>1</b>	<b>12</b>	<b>7.1</b>	<b>4.7</b>	<b>0.355</b>	<b>0.76</b>	<b>51</b>
<b>1A1</b>	<b>17</b>	<b>8.2</b>	<b>4.7</b>	<b>0.440</b>	<b>0.80</b>	<b>51</b>
<b>1A2</b>	<b>13</b>	<b>6.3</b>	<b>4.1</b>	<b>0.324</b>	<b>0.86</b>	<b>39</b>
M <sub>1</sub> models based on agonist bound M <sub>2</sub> crystal structure (4MQS)						
Model	EF2%	EF5%	EF10%	BEDROC	AUC	# of bound active
Crude	2.9	2	2	0.166	0.94	25
IF1	9.9	5.5	3.7	0.303	0.93	32
IF2	13	5.9	4.3	0.347	0.95	31
IF3	14	7.1	4.5	0.368	0.96	31
IF4	8.9	5.5	3.5	0.285	0.93	32
IF5	11	5.1	3.5	0.303	0.94	31
IF6	9.9	6.7	3.7	0.316	0.96	28
<b>IF7<sup>^</sup></b>	<b>16</b>	<b>6.7</b>	<b>4.1</b>	<b>0.385</b>	<b>0.94</b>	<b>32</b>

<sup>^</sup>Referred as model 2 in text.

Models chosen and used in the virtual screening are shown in bold.

**Table 5.3.** Enrichment results for M<sub>2</sub> crude and IFD structures.

Agonist bound M <sub>2</sub> crystal structure (4MQS)						
Structure	EF2%	EF5%	EF10%	BEDROC	AUC	# of bound active
Crude	4.9	3.5	2.4	0.215	0.80	39
<b>IF18</b>	<b>15</b>	<b>6.7</b>	<b>3.7</b>	<b>0.377</b>	<b>0.91</b>	<b>36</b>
IF19	9.9	6.3	4.5	0.344	0.91	36
<b>IF21</b>	<b>13</b>	<b>7.1</b>	<b>4.7</b>	<b>0.371</b>	<b>0.85</b>	<b>44</b>
IF24	8.9	7.4	4.1	0.337	0.82	45
IF25	9.9	5.9	3.5	0.297	0.89	35
IF28	3.0	3.9	3.3	0.211	0.76	44
Antagonist bound M <sub>2</sub> crystal structure (3UON)						
Structure	EF2%	EF5%	EF10%	BEDROC	AUC	# of bound active
Crude	5.9	3.1	1.8	0.172	0.65	51
<b>IF2</b>	<b>6.9</b>	<b>6.3</b>	<b>3.7</b>	<b>0.301</b>	<b>0.74</b>	<b>51</b>
<b>IF4</b>	<b>5.0</b>	<b>3.1</b>	<b>2.7</b>	<b>0.196</b>	<b>0.66</b>	<b>51</b>
IF12	5.9	4.0	2.4	0.138	0.63	51
IF13	2	2.7	1.6	0.127	0.63	51

IFD Structures chosen and used in the virtual screening are shown in bold.

**Table 5.4.** Enrichment results for M<sub>3</sub> models and IFD structures.

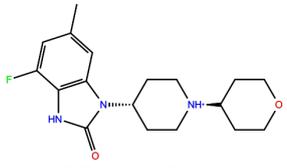
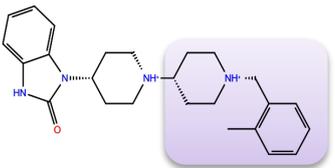
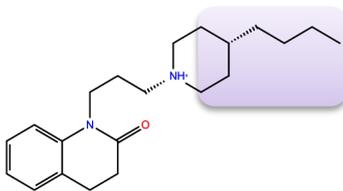
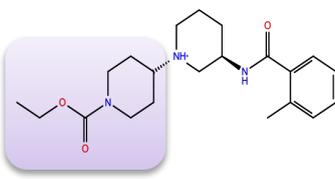
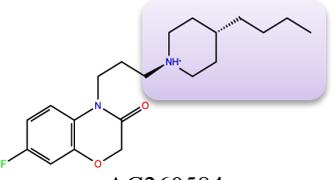
M <sub>3</sub> models based on agonist bound M <sub>2</sub> crystal structure (4MQS)						
Structure	EF2%	EF5%	EF10%	BEDROC	AUC	# of bound active
Crude	5.9	4.3	2.9	0.220	0.85	38
<b>IF1</b>	<b>13</b>	<b>6.7</b>	<b>4.7</b>	<b>0.380</b>	<b>0.92</b>	<b>35</b>
IF2	11	6.3	4.3	0.331	0.92	34
IF3	7.9	6.3	3.5	0.284	0.86	38
IF4	12	6.7	4.1	0.328	0.88	37
IF6	7.9	4.3	3.5	0.266	0.84	40
IF7	8.9	5.9	3.5	0.299	0.82	42
<b>IF9</b>	<b>13</b>	<b>7.1</b>	<b>4.3</b>	<b>0.357</b>	<b>0.83</b>	<b>43</b>
IF10	11	5.1	2.9	0.291	0.89	36
IF14	7.9	4.3	2.7	0.221	0.83	38
IF15	7.9	4.7	2.7	0.239	0.79	44
IF16	8.9	4.7	2.9	0.238	0.78	44
IF17	4	3.5	2.4	0.187	0.76	44
Inverse agonist bound M <sub>3</sub> crystal structure (4DAJ)						
Structure	EF2%	EF5%	EF10%	BEDROC	AUC	# of bound active
Crude	4	2.4	1.2	0.129	0.56	51
<b>IF1</b>	<b>14</b>	<b>6.3</b>	<b>4.3</b>	<b>0.368</b>	<b>0.79</b>	<b>50</b>
<b>IF2</b>	<b>11</b>	<b>7.1</b>	<b>4.9</b>	<b>0.363</b>	<b>0.78</b>	<b>51</b>
IF3	11	6.3	5.1	0.363	0.77	51
IF5	6.9	5.5	4.5	0.321	0.76	51
IF6	5.9	5.9	4.7	0.272	0.76	51
IF7	5	3.1	3.5	0.228	0.72	51
IF8	6.9	3.9	4.3	0.273	0.72	51
IF9	7.9	5.1	4.5	0.297	0.74	51
IF10	5.9	3.9	4.1	0.250	0.73	51
IF11	11	5.5	3.9	0.316	0.71	51

IFD Structures/models chosen and used in the virtual screening are shown in bold.

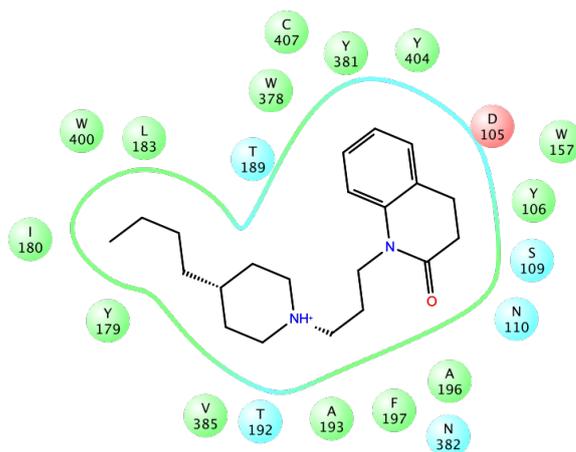
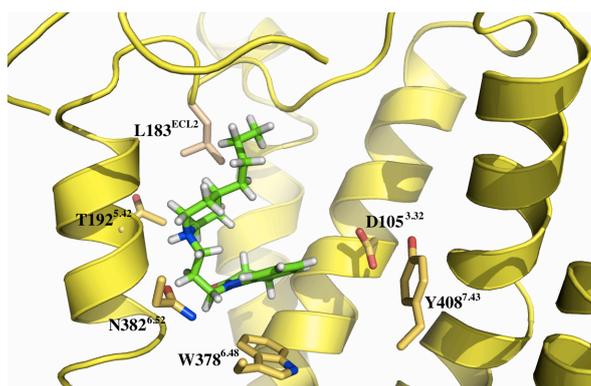
### 5.3.2 Binding modes of known M<sub>1</sub> selective ligands

The docking of the six known M<sub>1</sub> selective ligands revealed that these ligands did indeed favour the M<sub>1</sub> receptor through slightly better GScores compared to the M<sub>2</sub> and M<sub>3</sub> receptors (Table 5.5). Among the six ligands, GSK1034702 docked to the orthosteric site, while TBPB, 77-LH-281, VU0364527 and AC260584 portrayed bitopic characteristics, being bound to both the orthosteric and secondary/allosteric site. Brucine, the first reported M<sub>1</sub> PAM, unexpectedly bound only to the M<sub>1</sub> receptor, on top of the orthosteric site. Being structurally similar to TBPB, GSK1034702 is thought to be a potential bitopic ligand, however, it fitted well into the orthosteric site in the current docking study (Figure 5.1). TBPB, 77-LH-281, VU0364527 and AC260584 have been proposed to act on both the orthosteric and allosteric sites for their mode of actions (Valant, Sexton, & Christopoulos, 2009; Digby et al., 2012; Keov et al., 2013). This docking study is in agreement with this suggestion and confirmed that these allosteric agonists bind to both sites, acting as bitopic ligands (Figure 5.1). Based on the docking results and the fact that brucine is an M<sub>1</sub> PAM that selectively increases the affinity of M<sub>1</sub> for ACh, brucine is therefore inferred to bind at a putative site - cavity on top of the orthosteric site and between the ECL2 and ECL3 (Figure 5.1). For the bitopic ligands, besides preserving the interactions involving the orthosteric site, additional sets of interactions were established in the cavity formed by residues such as Y179<sup>ECL2</sup>, I180<sup>ECL2</sup>, L183<sup>ECL2</sup>, S184<sup>ECL2</sup>, V385<sup>6.55</sup> and W400<sup>7.35</sup>. For brucine, while there is an overlapping with the residues observed for making interactions with the bitopic ligands, it makes extra contacts with Y82<sup>2.61</sup>, Y85<sup>2.64</sup>, L86<sup>2.65</sup>, and E401<sup>7.36</sup>, and is separated from the orthosteric site by a tyrosine lid.

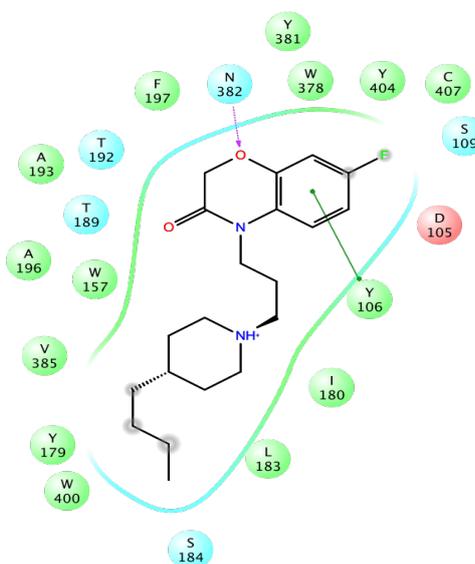
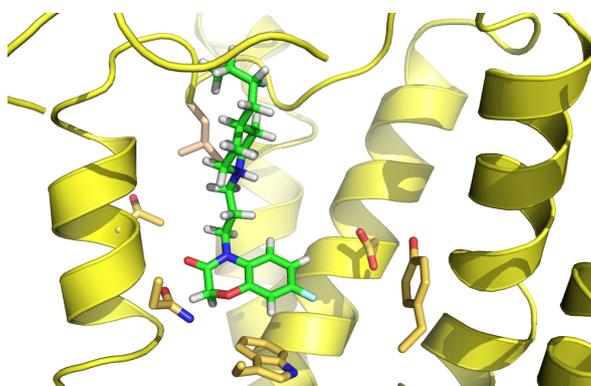
**Table 5.5.** GScore values of the known selective M<sub>1</sub> ligands showing their preferential docking to the M<sub>1</sub> model compared to the M<sub>2</sub> and M<sub>3</sub> models and Prime MM-GBSA estimations of the free energy changes for binding to the preferred M<sub>1</sub> model.

Compound	GScore			Prime MM-GBSA			
	M <sub>1</sub>	M <sub>2</sub>	M <sub>3</sub>	$\Delta G_{\text{Bind}}$	$\Delta G_{\text{Coul}}$	$\Delta G_{\text{Solv GB}}$	$\Delta G_{\text{vdW}}$
 GSK1034702	-10.24	-8.76	-8.36	-82.13	-14.00	25.32	-48.51
 TBPB	-9.51	-7.24	-8.25	-83.91	-9.55	29.97	-55.03
 77-LH-281	-8.57	-7.83	-6.89	-91.62	4.22	10.10	-51.74
 VU0364527	-8.50	-7.19	-7.07	-87.14	-0.39	16.91	-56.47
 AC260584	-7.62	-7.28	-7.08	-74.67	-12.85	21.10	-45.28
 Brucine	-6.268	nb	nb	-72.20	-23.93	27.34	-40.44

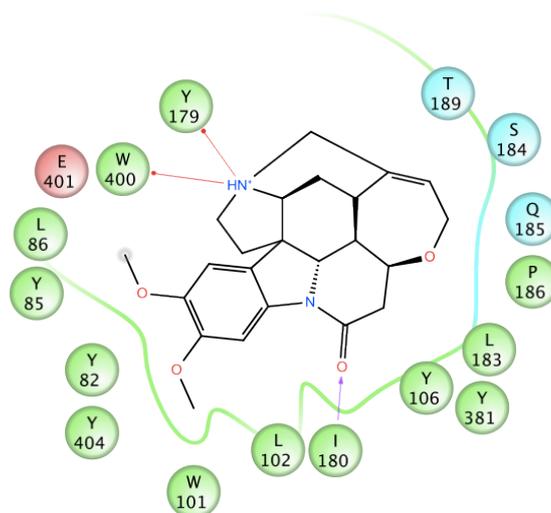
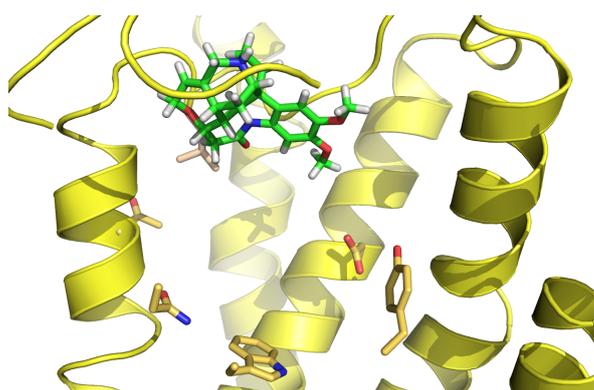
Un-shaded region bound to the primary/orthosteric binding site; purple shaded region bound to the secondary/allosteric binding site; and blue shaded region bound to the secondary binding site, on top of the orthosteric site. nb: not bound. Energies are in kcal/mol.  $\Delta G_{\text{bind}}$ : the estimated agonist-receptor binding free energies;  $\Delta G_{\text{Coul}}$ : the columbic binding free energy;  $\Delta G_{\text{vdW}}$ : the van der Waals binding free energy;  $\Delta G_{\text{Solv GB}}$ : the generalized Born solvation binding free energy.



77-LH-281

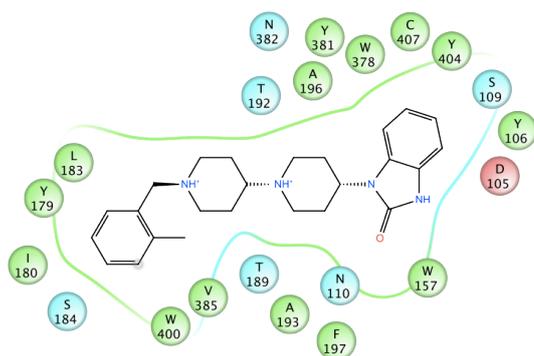
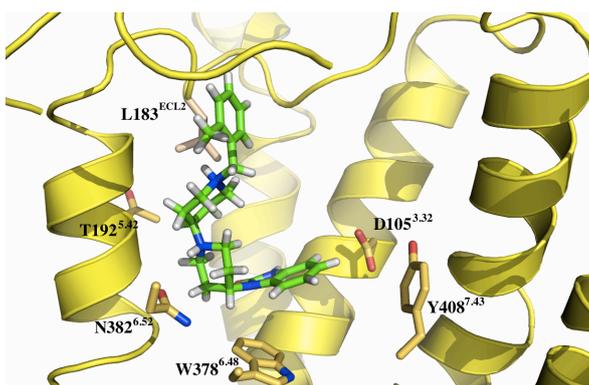


AC260584

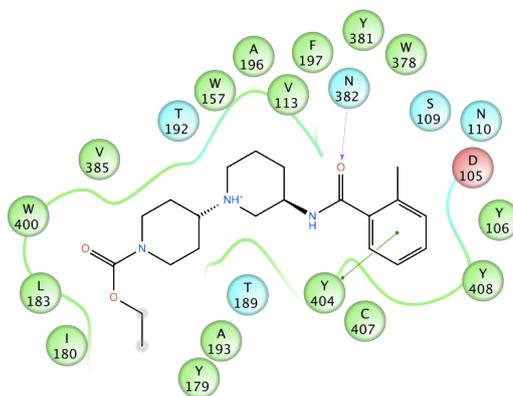
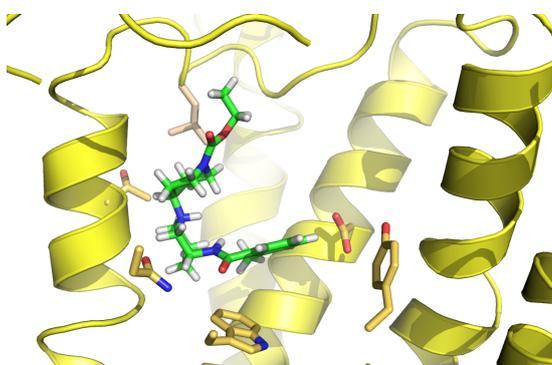


Brucine

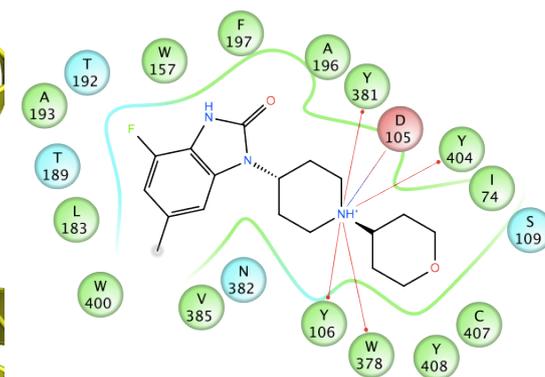
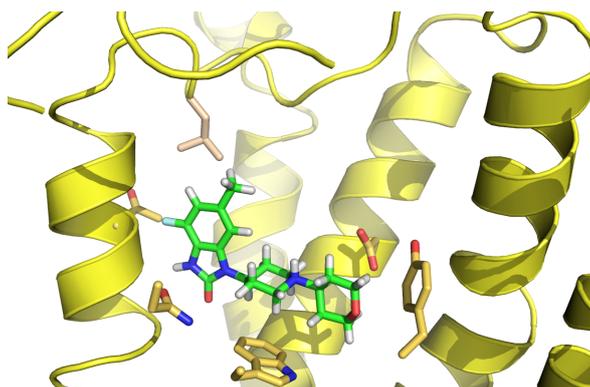
**Figure 5.1.** Binding modes and 2-D interactions map of the known  $M_1$  selective ligands. Red circle: -ve charged, light blue circle: polar, green circle: hydrophobic, green line: pi-pi stacking, red line: pi-cation, purple arrow: hydrogen bond (backbone), purple dash arrow: hydrogen bond (side-chain), red-blue line: salt bridge.



**TBPB**



**VU036452**



**GSK1034702**

**Figure 5.1.** Continued.

### 5.3.3 Identification of virtual screening hits

To identify new potential muscarinic ligands preferring the M<sub>1</sub> receptor and to evaluate the fitness of the models, a small-scale virtual screening against combinations of experimental structures and homology models of the M<sub>1</sub>, M<sub>2</sub> and M<sub>3</sub> receptors were carried out. The virtual screening successfully identified 19 of the ligands with a large difference in GScore among the receptors, favoring the M<sub>1</sub> receptor (Table 5.6). Thirteen of the ligands were retrieved from model 1, which was built based on the M<sub>3</sub> structure without any modifications, 4 ligands were retrieved from model 1A2, and 1 ligand from model 1A1 and model 2, respectively. These models (1A1, 1A2, 2) possess smaller binding pockets, representing the contracted binding pocket in the activated structure. The ligand binding modes were visually inspected and compared. The ligands either fitted poorly or were not able to dock in the binding sites of M<sub>2</sub> and M<sub>3</sub> mAChRs. However, molecular properties assessment showed that 8 of the hits (compounds 1 to 4, 7, 9, 10, and 19) possess properties which are outside the CNS drugs-likeness criteria (Table 5.7, shaded). The molecular weights for ligands 1 to 4, and 9 are more than 500 and hence these ligands were not investigated further. However, ligands 7, 10, and 19 possess only slightly higher polar surface area (and molecular weight, in the case of compound 7), and so they can still be considered as interesting hits as the M<sub>2</sub> allosteric modulator LY2119620 has a molecular weight of 450.05 and polar surface area of 131.55. The CNS drug-likeness concept is a modification of the original drug-likeness criteria based on the Lipinski rule of five taking into account the requirement of being able to cross the blood-brain barrier. Medicinal chemical properties of successful CNS drugs (with the ability to penetrate the blood-brain barrier affinity and exhibit CNS activity) are defined by the following molecular properties: molecular weight < 450 Da, polar surface area < 60-70 Å<sup>2</sup> to 90 Å<sup>2</sup>, calculated logP < 3, number of hydrogen bond donors < 4 and number of hydrogen bond acceptors < 8 (Pajouhesh & Lenz, 2005;

Reichel, 2006). A novelty check in the ChEMBL bioactivity database (Gaulton et al., 2012) through similarity searching at the lowest cutoff of 70% revealed that of the 14 selected hits, 9 compounds and their analogs have not been tested on mAChRs while compounds 6, 10, 11, 12, and 13 analogs have been previously tested on mAChRs. They are either active (compound 6 analog tested active as an antagonist for the M<sub>1</sub> mAChR, assay id: ChEMBL1613991), not active (compound 12 analog), or inconclusive (compound 11 and 13 analogs). However, it is worth noting that an analog of compound 10 has been tested active as M<sub>4</sub> mAChR selective allosteric modulator. The compound binds to an allosteric site on the M<sub>4</sub> mAChR and increase affinity for ACh and G proteins, while having no activity recorded for other muscarinic subtypes (Shirey et al., 2008).

**Table 5.6.** GScore values of the virtual screening hits showing their preferential docking to the M<sub>1</sub> models compared to the M<sub>2</sub> and M<sub>3</sub> structures and Prime MM-GBSA estimations of the free energy changes for binding to the preferred M<sub>1</sub> model.

Hits	Model	GScore			Prime MM-GBSA			
		M <sub>1</sub>	M <sub>2</sub>	M <sub>3</sub>	$\Delta G_{\text{Bind}}$	$\Delta G_{\text{Coul}}$	$\Delta G_{\text{Solv GB}}$	$\Delta G_{\text{vdW}}$
1	1	-11.18	-7.22	-5.00				
2	1	-10.44	-7.22	nb				
3	1	-10.30	-8.63	nb				
4	1	-10.21	-8.63	nb				
5	1	-10.19	-6.78	-7.98	-97.73	-13.10	16.25	-50.37
6	1	-10.18	-5.27	nb	-97.16	-17.58	25.82	-53.51
7	1	-9.99	nb	nb	-90.42	-16.25	37.30	-68.15
8	1	-9.67	-7.56	-8.09	-91.96	31.85	-34.39	-48.80
9	1	-9.44	-6.86	-7.78				
10	1	-9.43	-7.22	-7.11	-87.36	-18.33	20.99	-51.75
11	1	-9.08	-7.61	-8.06	-73.79	-7.48	12.42	-52.06
12	1	-9.05	-8.10	-7.83	-81.83	-1.53	-13.15	-35.37
13	1	-8.81	-6.30	-7.22	-68.17	-32.97	32.86	-37.56
14	1A2	-9.25	-7.94	-6.83	-68.92	-18.36	1.20	-16.18
15	2	-9.04	-7.84	-7.17	-63.86	8.24	-10.94	-27.65
16	1A2	-8.96	-7.64	-7.22	-70.70	-22.31	12.94	-26.38
17	1A2	-8.49	-7.25	-6.76	-80.88	8.14	-21.76	-40.80
18	1A2	-8.27	-7.08	-6.95	-71.04	-25.71	9.33	-25.85
19	1A1	-8.06	nb	nb	-82.45	-5.04	13.27	-49.01

nb: not bound. Energies are in kcal/mol.  $\Delta G_{\text{bind}}$ : the estimated agonist-receptor binding free energies;  $\Delta G_{\text{Coul}}$ : the columbic binding free energy;  $\Delta G_{\text{vdW}}$ : the van der Waals binding free energy;  $\Delta G_{\text{Solv GB}}$ : the generalized Born solvation binding free energy. Row shaded in grey indicated compounds with properties out of CNS drug-likeness range (see Table 5.7). MM-GBSA was not computed for compounds 1 to 4 and 9 as they were excluded on the basis of having molecular weight greater than 500.

**Table 5.7.** Molecular properties of known M<sub>1</sub> selective ligands and virtual screening hits.

M <sub>1</sub> selective ligands					
Compound	MW	AlogP	PSA	HBA	HBD
GSK1034702	334.417	1.132	51.460	2	2
TBPB	406.576	1.219	46.670	1	3
77LH281	329.510	2.660	24.750	1	1
VU0364527	374.507	1.479	63.080	3	2
AC260584	349.473	2.209	33.980	2	1
Brucine	395.483	-0.320	52.440	4	1
M <sub>1</sub> selective Hits					
Hits	MW	AlogP	PSA	HBA	HBD
1	551.694	6.519	139.380	5	1
2	576.119	7.130	111.660	5	1
3	551.581	2.140	195.410	7	2
4	509.613	0.073	87.630	3	3
5	325.434	1.659	51.000	2	3
6	407.962	2.915	53.210	1	2
7	488.554	3.399	90.770	5	1
8	300.430	-0.286	46.320	1	3
9	511.425	6.318	84.580	3	1
10	390.467	2.874	118.370	4	3
11	283.289	1.485	80.200	3	2
12	238.741	2.801	40.570	1	1
13	235.309	-0.320	61.860	2	4
14	204.080	0.537	51.610	0	4
15	161.225	1.443	30.200	0	2
16	161.291	0.846	76.910	1	4
17	317.131	0.087	40.770	2	1
18	205.282	1.226	33.540	1	2
19	433.514	2.434	102.480	4	3

Row shaded in grey indicated compounds with properties out of CNS drug-likeness range. CNS drug-likeness properties are defined by the following molecular properties: molecular weight < 450 Da, polar surface area < 60-70 Å<sup>2</sup> to 90 Å<sup>2</sup>, calculated logP < 3, number of hydrogen bond donors < 4 and number of hydrogen bond acceptors < 8. MW: molecular weight, PSA: polar surface area, HBA: hydrogen bond acceptor, HBD: hydrogen bond donor.

#### 5.3.4 Binding modes of selected virtual screening hits

The 14 selected virtual screening hits all bound to the orthosteric site, except compound 7, 10, and 19. Compound 7 and 10, bound to both the orthosteric and allosteric sites, similar to most of docked known M<sub>1</sub> selective ligands, while compound 19 resembled the docked brucine in M<sub>1</sub> mAChR, bound on top of the orthosteric site. The majority of the virtual hits were basic amines, which preserved the ionic, or hydrogen bonding interactions with the highly conserve D105<sup>3.32</sup>. Several others residues that are known to play key roles in mAChR recognition were also observed to establish interactions, including Y106<sup>3.33</sup>, S109<sup>3.36</sup>, W378<sup>6.48</sup>, Y381<sup>6.51</sup>, N382<sup>6.52</sup>, Y404<sup>7.39</sup> and Y408<sup>7.43</sup>. Although the ligands were bound to the orthosteric site, L183<sup>ECL2</sup> was found within 4 Å of the docked compounds 5, 6, 8, 11, and 12, similar to GSK1034702 (Figure 5.1). Compound 7 and 10 exhibited bitopic characteristics, making interactions that resembled most of the docked known M<sub>1</sub> selective ligands, where Y179<sup>ECL2</sup>, I180<sup>ECL2</sup>, L183<sup>ECL2</sup>, V385<sup>6.55</sup> and W400<sup>7.35</sup> form part of the binding pocket in addition to the orthosteric site. Compound 19 bound to the same cavity as brucine and established interactions similar to that of brucine. In the previous chapter, 4 residues (Y/F<sup>2.61</sup>, L/T/I<sup>2.65</sup>, S/N<sup>6.58</sup>, E/T/N<sup>7.36</sup>) at the extracellular vestibule with their side chains positioned facing towards the internal channel have been identified. These residues might be important for subtype selectivity, apart from those of the non-conserved ECL2 and ECL3. The docking of brucine and compound 19 revealed that Y82<sup>2.61</sup>, L86<sup>2.65</sup>, and E401<sup>7.36</sup> form part of the binding pocket, with S388<sup>6.58</sup> was also found forming part of the binding pocket for compound 19.

The M<sub>2</sub> crystal structure revealed that QNB binds within a deeply buried pocket defined by the side chains of TM3, 4, 5, 6 and 7. An aromatic cage consisting of three-conserved tyrosines – Y<sup>3.33</sup>, Y<sup>6.51</sup> and Y<sup>7.39</sup> covers the amine and forms a lid over the

ligand, separating the orthosteric site from the extracellular vestibule. D<sup>3.32</sup> and N<sup>6.52</sup> serve to orient the ligand in the largely hydrophobic binding cavity, where D<sup>3.32</sup> is responsible for the polar interactions with a positively charged amine moiety of the ligand, and N<sup>6.52</sup> forms a hydrogen bond pair with the hydroxyl and carbonyl groups in QNB. The TM residues that form the binding pocket are identical across all five mAChR subtypes and this explains the remarkable similarity in the binding mode observed in the M<sub>3</sub> crystal structure, in complex with tiotropium. Unexpectedly, in the agonist bound M<sub>2</sub> crystal structure, iperoxo bound to M<sub>2</sub> in the same binding cavity where the polar contacts resemble those in the binding of QNB to the inactive M<sub>2</sub> - D<sup>3.32</sup> interact with the ligand amine and N<sup>6.52</sup> engages in hydrogen bonding. The M<sub>2</sub> crystal structure with LY2119620 (M<sub>2</sub> PAM) revealed that the modulator is positioned directly above the orthosteric agonist, iperoxo, separated only by the tyrosine lid, with Y426<sup>7.39</sup> interacting with both ligands. The modulator establishes extensive interactions with the extracellular vestibule, specifically, Y177<sup>ECL2</sup> and W422<sup>7.35</sup> in aromatic stacking, and polar contacts involving Y80<sup>2.61</sup>, N410<sup>6.58</sup>, N419<sup>ECL3</sup> and E172<sup>ECL2</sup>. Our docking study showed that the ligand binding modes, both from known M<sub>1</sub> selective ligands and virtual screening share the key binding features that has been observed in the crystal structures. The ligands bound and established interactions as seen in the crystal structures while brucine binding mode was similar to LY2119620. However for the virtual screening ligands, Y381<sup>6.51</sup> and S109<sup>3.36</sup> were more dominant in engaging in hydrogen bonds compared to N382<sup>6.52</sup>. As the interactions were largely ligand dependent, different ligands may induce different interaction pairs. The fact that Y381<sup>6.51</sup> plays a more crucial role in the binding of agonists (through hydrogen bonding and cation- $\pi$  interactions) than N382<sup>6.52</sup> as shown by the site-directed mutagenesis may indicate that these ligands are possibly agonists (Huang, Nagy, Williams, Peseckis, & Messer, 1999; Ward, Curtis, & Hulme, 1999). Besides that the ligands were all docked

poorly in the M<sub>2</sub> and M<sub>3</sub> receptors, the residues possibly accounting for the subtype selectivity corresponding to Y179<sup>ECL2</sup> (Y177<sup>ECL2</sup> in M<sub>2</sub>, F221<sup>ECL2</sup> in M<sub>3</sub>), L183<sup>ECL2</sup> (F181<sup>ECL2</sup> in M<sub>2</sub>, L225<sup>ECL2</sup> in M<sub>3</sub>), Y82<sup>2.61</sup> (Y in M<sub>2</sub>, and F in M<sub>3</sub>), L86<sup>2.65</sup> (T in M<sub>2</sub> and I in M<sub>3</sub>), S388<sup>6.58</sup> (N in M<sub>2</sub> and M<sub>3</sub>), and E401<sup>7.36</sup> (T in M<sub>2</sub> and N in M<sub>3</sub>).

In general, antagonists are retrieved by virtual screening against inactive GPCR structures and agonists are discovered by virtual screening against active receptor structures. However, specific agonist discovery has been achieved using an inactive  $\kappa$ -opioid receptor structure (Negri et al., 2013) and one of the hits identified in the M<sub>2</sub>/M<sub>3</sub> virtual screening has shown agonist activity at the M<sub>3</sub> receptor (Kruse, Weiss, et al., 2013). This particular compound was able to activate the M<sub>3</sub> receptor but not the M<sub>2</sub> receptor and form hydrogen bond with Y<sup>3.33</sup>, unlike others ligands that form a pair of hydrogen bonds with N<sup>6.52</sup>, as seen in M<sub>2</sub> and M<sub>3</sub> crystal structures. Interestingly, the ligands identified in this study also showed that a hydrogen bond pair with N<sup>6.52</sup> is not mandatory. However, further studies are required to definitively determine the response elicited by these ligands.

### 5.3.5 Prime MM-GBSA

The six known M<sub>1</sub> selective ligands and the selected virtual screening hits were rescored by the Molecular Mechanics-Generalized Born Surface Area (MM-GBSA) approach in Prime (Prime MM-GBSA), with residues in the binding pocket of the receptor treated flexibly. Prime MM-GBSA has been widely used in the course of virtual screening, to provide a more quantitative ranking than the straight GScore (Mulakala & Viswanadhan, 2013; Shen et al., 2013; Arya & Coumar, 2014; Banavath, Sharma, Kumar, & Baskaran, 2014).

The binding free energy estimation for the known M<sub>1</sub> selective ligands ranged from 70-90 kcal/mol, while the virtual screening hits formed complexes with a binding free energy in the range of 60-90 kcal/mol (Table 5.5 & 5.6). Compounds 13, 14 and 15 performed worse than the known M<sub>1</sub> selective ligands, but the other compounds were comparable or better. The ligand-receptor interactions were largely driven by the vdW energy component in all the selected virtual screening hits, similar to the known M<sub>1</sub> selective ligands. However, four of them (compounds 13, 14, 16, and 18) bound with comparable vdW and columbic energy components. These ligands are among the least favorable with lower binding free energies. The favorable contributions of vdW interactions are not surprising since the binding cavity is largely hydrophobic, allowing ligands to participate in extensive hydrophobic contacts with the receptor. It is also apparent that hydrogen bonding is lacking for most of the known M<sub>1</sub> selective ligands and selected virtual screening hits.

#### **5.4 Conclusion**

In an effort to identify M<sub>1</sub> selective ligands, virtual screening of a small subset of a compound library against combinations of homology models and experimental structures of M<sub>1</sub>, M<sub>2</sub> and M<sub>3</sub> receptor were performed. Although model 1 was previously postulated to recognize antagonists better than agonists (Chapter 3, sections 3.2.3.2 and 3.2.3.3), the docking of six known M<sub>1</sub> selective ligands showed that, these compounds only bound to model 1, occupying the orthosteric and allosteric sites, and were not able to fit into the models with contracted binding pockets (models 1A1, 1A2 and 2). Muscarinic agonists are typically less complex and have smaller molecular weights than antagonists. It is assumed that the difference in size determines the recognition of agonists vs. antagonists by the models, where models modified or based according to the agonist bound structure might not be able to accommodate larger

ligands, because of steric clashes between the ligand and the binding pockets. Although the pharmacological properties of the ligands have yet to be tested, the current virtual screening study highlights the role of model 1 in retrieving ligands that have a preference for the M<sub>1</sub> receptor, and emphasizing the need of using more than a single model in virtual screening, to sufficiently represent a diverse range of “ligandable” and “dockable” conformations. The use of more than one structure for a single receptor ensures more ligand binding pocket variations are included in the virtual screening, representing receptor flexibility when bound to different ligands, and allowing the selection of the best docked GScore against all the models or crystal structures for a single ligand. Although selectivity determination between the subtypes might not be as straightforward due to the admittedly small difference between the orthosteric sites, the identification of ligands demonstrated by this current study illustrates the potential of this approach. The insights obtained from the current study will be tested experimentally to confirm the findings, which may then lead to the expansion of virtual screening for a larger number of molecules or further optimization of the pharmacological profile of any active ligands.

## CHAPTER 6

### CONCLUSION & FUTURE WORK

Despite the extensive progress that has been made in recent years in the field of GPCR structural biology, modeling of GPCRs is still considered to be a challenging task. Although a number of different strategies have been developed, such as the use of multiple/consensus templates (Yarnitzky, Levit, & Niv, 2010), non closest templates (Rataj, Witek, Mordalski, Kosciolk, & Bojarski, 2014), or refined models from molecular dynamics simulations (Tarcsey et al., 2013), the same strategies may not be applicable to all GPCRs, and the choice of strategy is probably case/receptor dependent, especially for the construction of models to recognize agonists. For example, it has been shown in this work that modeling based on the crystal structure of an inverse agonist bound  $\beta_2$ AR and consensus templates did not produce good M<sub>1</sub> mAChR models to recognize agonists even after refinement with MD simulations, but the same  $\beta_2$ AR structure has proved to be useful as a template in developing antagonist bound mAChR models (Thomas et al., 2014).

The lack of predictability may be due to the high diversity found in ligands that bind to GPCRs and the distinctive way that each receptor seems to respond to a ligand and change its conformation and binding pocket as the ligand binds (induced-fit effect). Furthermore, subtle changes such as in the orientation of a side chain, a slight rotation or a minor kink in the TM helix are crucial for ligand docking and screening and are the key to distinguishing the binding of agonists and antagonists (Roumen et al., 2011).

Hence it seems that the method used for modeling different GPCR family members may require customization, by making use all the experimental information that is available regarding the specific receptor.

Taking advantage of different resources, such as increasing experimental structures and structural information for GPCRs, pharmacological studies, mutagenesis data, computational techniques (such as induced-fit docking), and information from the community-wide GPCR dock assessments (Michino et al., 2009; Kufareva, Rueda, Katritch, Stevens, & Abagyan, 2011; Kufareva, Katritch, Stevens, & Abagyan, 2014), this work showcases the potential of computational approaches in studying complex ligand-receptor interactions at atomic level in the absence of an experimental structure. It has also shown that a single model may not be able to capture all the interactions for a set of ligands and using more than one model in ligand profiling may be an advantage, as it increases the dockability. In addition, the M<sub>1</sub> mAChR models generated in this study have been used to interpret experimental binding data of flavonoids at the M<sub>1</sub> mAChR (Swaminathan et al., 2014). However, it is clear that more experimentally solved GPCR structures bound to different ligands and G proteins and other signaling counterparts are required, together with improved algorithms and methods for structure predictions (especially in non conserved regions) and docking (i.e. better search methods to correctly position ligands in a relatively larger binding pocket) to realise the full potential of structure based drug discovery for GPCRs.

In this study, the models of M<sub>1</sub> mAChR have been developed and evaluated thoroughly for their capability to separate agonists from decoys and antagonists and to predict the binding modes of the bound ligands. MD simulations were used to reveal the dynamics behaviors of the receptors when bound to different ligands and in the apo form,

compared to other mAChR subtypes. Finally, the models were proved useful in retrieving M<sub>1</sub> mAChR preferring ligands in the virtual screening involving also M<sub>2</sub> and M<sub>3</sub> mAChRs. Future work to follow on from this study will include experimental testing of the virtual screening hits to confirm their pharmacological profile against the different mAChR subtypes. In addition, further virtual screening could be carried out on a compound library constructed based on known M<sub>1</sub> mAChR selective allosteric modulators and using multiple MD simulations snapshots for the identification of potential allosteric modulators. For this purpose, it may be helpful to use the MD snapshots extracted from the holo systems investigated in this study, as the binding affinities of allosteric modulators have been shown to be dependent on orthosteric site occupancy in MD simulations described elsewhere (Dror et al., 2013).

Last year's Nobel Prize in chemistry was awarded to Martin Karplus, Michael Levitt, and Arieh Warshel for their contributions to molecular modeling of complex chemical systems, indicating that computational methods have matured, are well recognized and are essential tools for imitating and simulating biological processes. Information on how GPCRs work at the molecular level is still lacking and there is much that remains to be explored, such as the interactions between GPCRs and intracellular effectors, the characterization of orphan GPCRs, the roles of dimerization and oligomerization, the impact of protein-membrane interactions on function and organization, the activation mechanism, ligand recognition profiling, biased signaling and allosteric control (Salon, Lodowski, & Palczewski, 2011). Computer modeling is therefore expected to continue to play a vital role in connecting structure and function for GPCRs (simulations), filling in the remaining gaps in the coverage of GPCR superfamily (homology modeling) and drug design and discovery (virtual screening).

## REFERENCES

- Abdul-Ridha, A., Lane, J. R., Mistry, S. N., Lopez, L., Sexton, P. M., Scammells, P. J., Christopoulos, A., & Canals, M. (2014). Mechanistic insights into allosteric structure-function relationships at the M1 muscarinic acetylcholine receptor. *Journal of Biological Chemistry*, 289(48), 33701-33711.
- Abdul-Ridha, A., Lopez, L., Keov, P., Thal, D. M., Mistry, S. N., Sexton, P. M., Lane, J. R., Canals, M., & Christopoulos, A. (2014). Molecular determinants of allosteric modulation at the M1 muscarinic acetylcholine receptor. *Journal of Biological Chemistry*, 289(9), 6067-6079.
- Allen, W. J., Lemkul, J. A., & Bevan, D. R. (2009). GridMAT-MD: a grid-based membrane analysis tool for use with molecular dynamics. *J Comput Chem*, 30(12), 1952-1958.
- Allman, K., Page, K. M., Curtis, C. A., & Hulme, E. C. (2000). Scanning mutagenesis identifies amino acid side chains in transmembrane domain 5 of the M(1) muscarinic receptor that participate in binding the acetyl methyl group of acetylcholine. *Molecular Pharmacology*, 58(1), 175-184.
- Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, 25(17), 3389-3402.
- Angel, T. E., Chance, M. R., & Palczewski, K. (2009). Conserved waters mediate structural and functional activation of family A (rhodopsin-like) G protein-coupled receptors. *Proceedings of the National Academy of Sciences of the United States of America*, 106(21), 8555-8560.
- Armougom, F., Moretti, S., Poirot, O., Audic, S., Dumas, P., Schaeli, B., Keduas, V., & Notredame, C. (2006). Espresso: automatic incorporation of structural information in multiple sequence alignments using 3D-Coffee. *Nucleic Acids Research*, 34(Web Server issue), W604-608.
- Aruksakunwong, O., Malaisree, M., Decha, P., Sompornpisut, P., Parasuk, V., Pianwanit, S., & Hannongbua, S. (2007). On the lower susceptibility of oseltamivir to influenza neuraminidase subtype N1 than those in N2 and N9. *Biophysical Journal*, 92(3), 798-807.
- Arya, H., & Coumar, M. S. (2014). Virtual screening of traditional Chinese medicine (TCM) database: identification of fragment-like lead molecules for filariasis target asparaginyl-tRNA synthetase. *Journal of Molecular Modeling*, 20(6), 2266.

- Avlani, V. A., Gregory, K. J., Morton, C. J., Parker, M. W., Sexton, P. M., & Christopoulos, A. (2007). Critical role for the second extracellular loop in the binding of both orthosteric and allosteric G protein-coupled receptor ligands. *The Journal of biological chemistry*, 282(35), 25677-25686.
- Avlani, V. A., Langmead, C. J., Guida, E., Wood, M. D., Tehan, B. G., Herdon, H. J., Watson, J. M., Sexton, P. M., & Christopoulos, A. (2010). Orthosteric and allosteric modes of interaction of novel selective agonists of the M1 muscarinic acetylcholine receptor. *Molecular Pharmacology*, 78(1), 94-104.
- Bakan, A., Meireles, L. M., & Bahar, I. (2011). ProDy: protein dynamics inferred from theory and experiments. *Bioinformatics (Oxford, England)*, 27(11), 1575-1577.
- Baker, N. A., Sept, D., Joseph, S., Holst, M. J., & McCammon, J. A. (2001). Electrostatics of nanosystems: application to microtubules and the ribosome. *Proceedings of the National Academy of Sciences of the United States of America*, 98(18), 10037-10041.
- Ballesteros, J. A., & Weinstein, H. (1995). Integrated methods for the construction of three dimensional models and computational probing of structure-function relations in G protein coupled receptors. In S. C. Sealfon & P. M. Conn (Eds.), *Methods in Neurosciences* (Vol. 25, pp. 366-428). San Diego, CA: Academic Press.
- Balupuri, A., & Sobhia, M. E. (2014). Exploring a model of human chemokine receptor CCR2 in presence of TAK779: A membrane based molecular dynamics study. *Journal of Molecular Structure*, 1063(0), 131-138.
- Banavath, H. N., Sharma, O. P., Kumar, M. S., & Baskaran, R. (2014). Identification of novel tyrosine kinase inhibitors for drug resistant T315I mutant BCR-ABL: a virtual screening and molecular dynamics simulations study. *Sci Rep*, 4, 6948.
- Beach, T. G., Walker, D. G., Potter, P. E., Sue, L. I., & Fisher, A. (2001). Reduction of cerebrospinal fluid amyloid beta after systemic administration of M1 muscarinic agonists. *Brain Research*, 905(1-2), 220-223.
- Bee, M. S., & Hulme, E. C. (2007). Functional analysis of transmembrane domain 2 of the M1 muscarinic acetylcholine receptor. *The Journal of biological chemistry*, 282(44), 32471-32479.
- Benz, R. W., Castro-Roman, F., Tobias, D. J., & White, S. H. (2005). Experimental validation of molecular dynamics simulations of lipid bilayers: a new approach. *Biophysical Journal*, 88(2), 805-817.
- Bigl, V., & Rossner, S. (2003). Amyloid precursor protein processing in vivo--insights from a chemically-induced constitutive overactivation of protein kinase C in Guinea pig brain. *Current Medicinal Chemistry*, 10(10), 871-882.

- Bissantz, C., Bernard, P., Hibert, M., & Rognan, D. (2003). Protein-based virtual screening of chemical databases. II. Are homology models of G-Protein Coupled Receptors suitable targets? *Proteins*, 50(1), 5-25.
- Blüml, K., Mutschler, E., & Wess, J. (1994). Functional role in ligand binding and receptor activation of an asparagine residue present in the sixth transmembrane domain of all muscarinic acetylcholine receptors. *The Journal of biological chemistry*, 269(29), 18870-18876.
- Boeckmann, B., Bairoch, A., Apweiler, R., Blatter, M.-C., Estreicher, A., Gasteiger, E., Martin, M. J., Michoud, K., O'Donovan, C., Phan, I., Pilbout, S., & Schneider, M. (2003). The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Research*, 31(1), 365-370.
- Bolton, E. E., Wang, Y., Thiessen, P. A., & Bryant, S. H. (2008). Chapter 12 PubChem: Integrated Platform of Small Molecules and Biological Activities. In A. W. Ralph & C. S. David (Eds.), *Annual Reports in Computational Chemistry* (Vol. Volume 4, pp. 217-241): Elsevier.
- Broadley, K., & Kelly, D. (2001). Muscarinic Receptor Agonists and Antagonists. *Molecules*, 6(3), 142-193.
- Budzik, B., Garzya, V., Shi, D., Foley, J. J., Rivero, R. A., Langmead, C. J., Watson, J., Wu, Z., Forbes, I. T., & Jin, J. (2010). 2' biaryl amides as novel and subtype selective M1 agonists. Part I: Identification, synthesis, and initial SAR. *Bioorganic & Medicinal Chemistry Letters*, 20(12), 3540-3544.
- Budzik, B., Garzya, V., Shi, D., Walker, G., Lauchart, Y., Lucas, A. J., Rivero, R. A., Langmead, C. J., Watson, J., Wu, Z., Forbes, I. T., & Jin, J. (2010). 2' biaryl amides as novel and subtype selective M1 agonists. Part II: Further optimization and profiling. *Bioorganic & Medicinal Chemistry Letters*, 20(12), 3545-3549.
- Budzik, B., Garzya, V., Shi, D., Walker, G., Woolley-Roberts, M., Pardoe, J., Lucas, A., Tehan, B., Rivero, R. A., Langmead, C. J., Watson, J., Wu, Z., Forbes, I. T., & Jin, J. (2010). Novel N-Substituted Benzimidazolones as Potent, Selective, CNS-Penetrant, and Orally Active M1 mAChR Agonists. *ACS Medicinal Chemistry Letters*, 1(6), 244-248.
- Bymaster, F. P., Carter, P. A., Yamada, M., Gomeza, J., Wess, J., Hamilton, S. E., Nathanson, N. M., McKinzie, D. L., & Felder, C. C. (2003). Role of specific muscarinic receptor subtypes in cholinergic parasympathomimetic responses, in vivo phosphoinositide hydrolysis, and pilocarpine-induced seizure activity. *The European journal of neuroscience*, 17(7), 1403-1410.
- Caccamo, A., Oddo, S., Billings, L. M., Green, K. N., Martinez-Coria, H., Fisher, A., & LaFerla, F. M. (2006). M1 receptors play a central role in modulating AD-like pathology in transgenic mice. *Neuron*, 49(5), 671-682.

- Canals, M., Lane, J. R., Wen, A., Scammells, P. J., Sexton, P. M., & Christopoulos, A. (2012). A Monod-Wyman-Changeux mechanism can explain G protein-coupled receptor (GPCR) allosteric modulation. *The Journal of Biological Chemistry*, 287(1), 650-659.
- Cang, X., Spomer, J., & Cheatham, T. E., 3rd. (2011). Explaining the varied glycosidic conformational, G-tract length and sequence preferences for anti-parallel G-quadruplexes. *Nucleic Acids Res*, 39(10), 4499-4512.
- Case, D. A., Darden, T. A., Cheatham, I., T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Crowley, F., Walker, R. C., Zhang, W., Merz, K. M., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Kolossváry, I., Wong, K. F., Paesani, F., Vanicek, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Yang, L., Tan, C., Mongan, J., Hornak, V., Cui, G., Mathews, D. H., Seetin, M. G., Sagui, C., Babin, V., & Kollman, P. A. (2008). AMBER 10, University of California, San Francisco.
- Case, D. A., Darden, T. A., Cheatham, I., T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Walker, R. C., Zhang, W., Merz, K. M., Roberts, B., Wang, B., Hayik, S., Roitberg, A., Seabra, G., Kolossvai, I., Wong, K. F., Paesani, F., Vanicek, J., Liu, J., Wu, X., Brozell, S. R., Steinbrecher, T., Gohlke, H., Cai, Q., Ye, X., Hsieh, M., Cui, G., Roe, D. R., Mathews, D. H., Seetin, M. G., Sagui, C., Babin, V., Luchko, T., Gusarov, S., Kovalenko, A., & Kollman, P. A. (2010). AMBER 11, University of California, San Francisco.
- Chen, S. F., Cao, Y., Chen, J. J., & Chen, J. Z. (2013). Binding selectivity studies of PKB $\alpha$  using molecular dynamics simulation and free energy calculations. *Journal of Molecular Modeling*, 19(11), 5097-5112.
- Cherezov, V., Rosenbaum, D. M., Hanson, M. A., Rasmussen, S. G. F., Thian, F. S., Kobilka, T. S., Choi, H.-J., Kuhn, P., Weis, W. I., Kobilka, B. K., & Stevens, R. C. (2007). High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. *Science (New York, N.Y.)*, 318(5854), 1258-1265.
- Chien, E. Y. T., Liu, W., Zhao, Q., Katritch, V., Han, G. W., Hanson, M. A., Shi, L., Newman, A. H., Javitch, J. A., Cherezov, V., & Stevens, R. C. (2010). Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. *Science (New York, N.Y.)*, 330(6007), 1091-1095.
- Cisse, M., Braun, U., Leitges, M., Fisher, A., Pages, G., Checler, F., & Vincent, B. (2011). ERK1-independent alpha-secretase cut of beta-amyloid precursor protein via M1 muscarinic receptors and PKC $\alpha$ /epsilon. *Molecular and Cellular Neuroscience*, 47(3), 223-232.
- Colovos, C., & Yeates, T. O. (1993). Verification of protein structures: patterns of nonbonded atomic interactions. *Protein science : a publication of the Protein Society*, 2(9), 1511-1519.

- Congreve, M., Langmead, C. J., Mason, J. S., & Marshall, F. H. (2011). Progress in structure based drug design for G protein-coupled receptors. *Journal of Medicinal Chemistry*, 54(13), 4283-4311.
- Conn, P. J., Christopoulos, A., & Lindsley, C. W. (2009). Allosteric modulators of GPCRs: a novel approach for the treatment of CNS disorders. *Nature reviews. Drug discovery*, 8(1), 41-54.
- Conn, P. J., Jones, C. K., & Lindsley, C. W. (2009). Subtype-selective allosteric modulators of muscarinic receptors for the treatment of CNS disorders. *Trends in Pharmacological Sciences*, 30(3), 148-155.
- Costanzi, S. (2012). Homology modeling of class a G protein-coupled receptors. *Methods in Molecular Biology*, 857, 259-279.
- Crowley, M. F., Williamson, M. J., & Walker, R. C. (2009). CHAMBER: Comprehensive support for CHARMM force fields within the AMBER software. *International Journal of Quantum Chemistry*, 109(15), 3767-3772.
- Darden, T., York, D., & Pedersen, L. (1993). Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. *The Journal of Chemical Physics*, 98(12), 10089-10092.
- Daval, S. B., Kellenberger, E., Bonnet, D., Utard, V., Galzi, J. L., & Ilien, B. (2013). Exploration of the orthosteric/allosteric interface in human M1 muscarinic receptors by bitopic fluorescent ligands. *Molecular Pharmacology*, 84(1), 71-85.
- Daval, S. B., Valant, C., Bonnet, D., Kellenberger, E., Hibert, M., Galzi, J. L., & Ilien, B. (2012). Fluorescent derivatives of AC-42 to probe bitopic orthosteric/allosteric binding mechanisms on muscarinic M1 receptors. *Journal of Medicinal Chemistry*, 55(5), 2125-2143.
- Davie, B. J., Christopoulos, A., & Scammells, P. J. (2013). Development of M1 mAChR allosteric and bitopic ligands: prospective therapeutics for the treatment of cognitive deficits. *ACS Chem Neurosci*, 4(7), 1026-1048.
- Davis, A. A., Fritz, J. J., Wess, J., Lah, J. J., & Levey, A. I. (2010). Deletion of M1 muscarinic acetylcholine receptors increases amyloid pathology in vitro and in vivo. *Journal of Neuroscience*, 30(12), 4190-4196.
- de Graaf, C., & Rognan, D. (2008). Selective structure-based virtual screening for full and partial agonists of the beta2 adrenergic receptor. *Journal of Medicinal Chemistry*, 51(16), 4978-4985.
- Deupi, X., Edwards, P., Singhal, A., Nickle, B., Oprian, D., Schertler, G., & Standfuss, J. (2012). Stabilized G protein binding site in the structure of constitutively active

metarhodopsin-II. *Proceedings of the National Academy of Sciences of the United States of America*, 109(1), 119-124.

- Deupi, X., & Standfuss, J. (2011). Structural insights into agonist-induced activation of G-protein-coupled receptors. *Current Opinion in Structural Biology*, 21(4), 541-551.
- Digby, G. J., Utley, T. J., Lamsal, A., Sevel, C., Sheffler, D. J., Lebois, E. P., Bridges, T. M., Wood, M. R., Niswender, C. M., Lindsley, C. W., & Conn, P. J. (2012). Chemical modification of the M(1) agonist VU0364572 reveals molecular switches in pharmacology and a bitopic binding mode. *ACS Chem Neurosci*, 3(12), 1025-1036.
- Dong, X., Zhao, Y., Huang, X., Lin, K., Chen, J., Wei, E., Liu, T., & Hu, Y. (2013). Structure-based drug design using GPCR homology modeling: toward the discovery of novel selective CysLT2 antagonists. *European journal of medicinal chemistry*, 62, 754-763.
- Dore, A. S., Robertson, N., Errey, J. C., Ng, I., Hollenstein, K., Tehan, B., Hurrell, E., Bennett, K., Congreve, M., Magnani, F., Tate, C. G., Weir, M., & Marshall, F. H. (2011). Structure of the adenosine A(2A) receptor in complex with ZM241385 and the xanthines XAC and caffeine. *Structure*, 19(9), 1283-1293.
- Dror, R. O., Green, H. F., Valant, C., Borhani, D. W., Valcourt, J. R., Pan, A. C., Arlow, D. H., Canals, M., Lane, J. R., Rahmani, R., Baell, J. B., Sexton, P. M., Christopoulos, A., & Shaw, D. E. (2013). Structural basis for modulation of a G-protein-coupled receptor by allosteric drugs. *Nature*, 503(7475), 295-299.
- Dror, R. O., Pan, A. C., Arlow, D. H., Borhani, D. W., Maragakis, P., Shan, Y., Xu, H., & Shaw, D. E. (2011). Pathway and mechanism of drug binding to G-protein-coupled receptors. *Proceedings of the National Academy of Sciences of the United States of America*, 108(32), 13118-13123.
- Espinoza-Fonseca, L. M., Pedretti, A., & Vistoli, G. (2008). Structure and dynamics of the full-length M1 muscarinic acetylcholine receptor studied by molecular dynamics simulations. *Archives of Biochemistry and Biophysics*, 469(1), 142-150.
- Evers, A., Hessler, G., Matter, H., & Klabunde, T. (2005). Virtual screening of biogenic amine-binding G-protein coupled receptors: comparative evaluation of protein- and ligand-based virtual screening protocols. *Journal of Medicinal Chemistry*, 48(17), 5448-5465.
- Fan, H., Irwin, J. J., Webb, B. M., Klebe, G., Shoichet, B. K., & Sali, A. (2009). Molecular docking screens using comparative models of proteins. *J Chem Inf Model*, 49(11), 2512-2527.
- Farce, A., Chugunov, A. O., Logé, C., Sabaouni, A., Yous, S., Dilly, S., Renault, N., Vergoten, G., Efremov, R. G., Lesieur, D., & Chavatte, P. (2008). Homology modeling of MT1 and MT2 receptors. *European journal of medicinal chemistry*, 43(9), 1926-1944.

- Farias, G. G., Godoy, J. A., Hernandez, F., Avila, J., Fisher, A., & Inestrosa, N. C. (2004). M1 muscarinic receptor activation protects neurons from beta-amyloid toxicity. A role for Wnt signaling pathway. *Neurobiology of Disease*, 17(2), 337-348.
- Filizola, M., Wang, S. X., & Weinstein, H. (2006). Dynamic models of G-protein coupled receptor dimers: indications of asymmetry in the rhodopsin dimer from molecular dynamics simulations in a POPC bilayer. *Journal of Computer-Aided Molecular Design*, 20(7-8), 405-416.
- Fisher, Michaelson, D. M., Brandeis, R., Haring, R., Chapman, S., & Pittel, Z. (2000). M1 muscarinic agonists as potential disease-modifying agents in Alzheimer's disease. Rationale and perspectives. *Annals of the New York Academy of Sciences*, 920, 315-320.
- Fisher, A. (2008). M1 muscarinic agonists target major hallmarks of Alzheimer's disease-the pivotal role of brain M1 receptors. *Neuro-degenerative diseases*, 5(3-4), 237-240.
- Flight, M. H. (2013). Drug discovery: Structure-led design. *Nature*, 502(7471), S50-52.
- Forlenza, O. V., Spink, J. M., Dayanandan, R., Anderton, B. H., Olesen, O. F., & Lovestone, S. (2000). Muscarinic agonists reduce tau phosphorylation in non-neuronal cells via GSK-3beta inhibition and in neurons. *Journal of Neural Transmission*, 107(10), 1201-1212.
- Foster, D. J., Choi, D. L., Conn, P. J., & Rook, J. M. (2014). Activation of M1 and M4 muscarinic receptors as potential treatments for Alzheimer's disease and schizophrenia. *Neuropsychiatric disease and treatment*, 10, 183-191.
- 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. (2004). Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of Medicinal Chemistry*, 47(7), 1739-1749.
- Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Montgomery, J. J. A., Vreven, T., Kudin, K. N., Burant, J. C., Millam, J. M., Iyengar, S. S., Tomasi, J., Barone, V., Mennucci, B., Cossi, M., Scalmani, G., Rega, N., Petersson, G. A., Nakatsuji, H., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Klene, M., Li, X., Knox, J. E., Hratchian, H. P., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Ayala, P. Y., Morokuma, K., Voth, G. A., Salvador, P., Dannenberg, J. J., Zakrzewski, V. G., Dapprich, S., Daniels, A. D., Strain, M. C., Farkas, O., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Ortiz, J. V., Cui, Q., Baboul, A. G., Clifford, S., Cioslowski, J., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Challacombe, M., Gill, P. M. W., Johnson, B., Chen, W., Wong, M. W., Gonzalez, C., & Pople, J. A. (2004). Gaussian 03, Gaussian, Inc., Wallingford CT.

- Gaulton, A., Bellis, L. J., Bento, A. P., Chambers, J., Davies, M., Hersey, A., Light, Y., McGlinchey, S., Michalovich, D., Al-Lazikani, B., & Overington, J. P. (2012). ChEMBL: a large-scale bioactivity database for drug discovery. *Nucleic Acids Res*, *40*(Database issue), D1100-1107.
- Gil-Mast, S., Kortagere, S., Kota, K., & Kuzhikandathil, E. V. (2013). An amino acid residue in the second extracellular loop determines the agonist-dependent tolerance property of the human D3 dopamine receptor. *ACS Chem Neurosci*, *4*(6), 940-951.
- Gohlke, H., & Case, D. A. (2004). Converging free energy estimates: MM-PB(GB)SA studies on the protein-protein complex Ras-Raf. *J Comput Chem*, *25*(2), 238-250.
- González, A., Perez-Acle, T., Pardo, L., & Deupi, X. (2011). Molecular basis of ligand dissociation in  $\beta$ -adrenergic receptors. *PLoS One*, *6*(9), e23815.
- Goodwin, J. A., Hulme, E. C., Langmead, C. J., & Tehan, B. G. (2007). Roof and floor of the muscarinic binding pocket: variations in the binding modes of orthosteric ligands. *Molecular Pharmacology*, *72*(6), 1484-1496.
- Gouet, P., Courcelle, E., Stuart, D. I., & Métoz, F. (1999). ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics (Oxford, England)*, *15*(4), 305-308.
- Govaerts, C., Lefort, A., Costagliola, S., Wodak, S. J., Ballesteros, J. A., Van Sande, J., Pardo, L., & Vassart, G. (2001). A conserved Asn in transmembrane helix 7 is an on/off switch in the activation of the thyrotropin receptor. *Journal of Biological Chemistry*, *276*(25), 22991-22999.
- Greenlee, W., Clader, J., Asberom, T., McCombie, S., Ford, J., Guzik, H., Kozlowski, J., Li, S., Liu, C., Lowe, D., Vice, S., Zhao, H., Zhou, G., Billard, W., Binch, H., Crosby, R., Duffy, R., Lachowicz, J., Coffin, V., Watkins, R., Ruperto, V., Strader, C., Taylor, L., & Cox, K. (2001). Muscarinic agonists and antagonists in the treatment of Alzheimer's disease. *Farmacologia*, *56*(4), 247-250.
- Gregory, K. J., Hall, N. E., Tobin, A. B., Sexton, P. M., & Christopoulos, A. (2010). Identification of orthosteric and allosteric site mutations in M2 muscarinic acetylcholine receptors that contribute to ligand-selective signaling bias. *Journal of Biological Chemistry*, *285*(10), 7459-7474.
- Grubmueller, H. (1996). SOLVATE v. 1.0, Theoretical Biophysics Group, Institute for Medical Optics, Ludwig-Maximilians University, Munich.
- Guedes, I., de Magalhães, C., & Dardenne, L. (2014). Receptor–ligand molecular docking. *Biophysical Reviews*, *6*(1), 75-87.
- Gurevich, V. V., & Gurevich, E. V. (2006). The structural basis of arrestin-mediated regulation of G-protein-coupled receptors. *Pharmacology & Therapeutics*, *110*(3), 465-502.

- Haga, K., Kruse, A. C., Asada, H., Yurugi-Kobayashi, T., Shiroishi, M., Zhang, C., Weis, W. I., Okada, T., Kobilka, B. K., Haga, T., & Kobayashi, T. (2012). Structure of the human M2 muscarinic acetylcholine receptor bound to an antagonist. *Nature*, *482*(7386), 547-551.
- Halgren, T. A., Murphy, R. B., Friesner, R. A., Beard, H. S., Frye, L. L., Pollard, W. T., & Banks, J. L. (2004). Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *Journal of Medicinal Chemistry*, *47*(7), 1750-1759.
- Haring, R., Gurwitz, D., Barg, J., Pinkas-Kramarski, R., Heldman, E., Pittel, Z., Danenberg, H. D., Wengier, A., Meshulam, H., Marciano, D., & et al. (1995). NGF promotes amyloid precursor protein secretion via muscarinic receptor activation. *Biochemical and Biophysical Research Communications*, *213*(1), 15-23.
- Heifetz, A., Mazanetz, M. P., James, T., Pal, S., Law, R. J., Slack, M., & Biggin, P. C. (2013). From Receptors to Ligands: Fragment-assisted Drug Design for GPCRs Applied to the Discovery of H3 and H4 Receptor Antagonist *Med chem*, *4*, 313-321.
- Helal, M. A., Darwish, K. M., & Hammad, M. A. (2014). Homology modeling and explicit membrane molecular dynamics simulation to delineate the mode of binding of thiazolidinediones into FFAR1 and the mechanism of receptor activation. *Bioorganic and Medicinal Chemistry Letters*, *24*(22), 5330-5336.
- Heng, B. C., Auel, D., & Fussenegger, M. (2013). An overview of the diverse roles of G-protein coupled receptors (GPCRs) in the pathophysiology of various human diseases. *Biotechnology Advances*, *31*(8), 1676-1694.
- Heo, J., Vaidehi, N., Wendel, J., & Goddard, W. A. (2007). Prediction of the 3-D structure of rat MrgA G protein-coupled receptor and identification of its binding site. *Journal of Molecular Graphics and Modelling*, *26*(4), 800-812.
- Hock, C., Maddalena, A., Raschig, A., Muller-Spahn, F., Eschweiler, G., Hager, K., Heuser, I., Hampel, H., Muller-Thomsen, T., Oertel, W., Wienrich, M., Signorell, A., Gonzalez-Agosti, C., & Nitsch, R. M. (2003). Treatment with the selective muscarinic m1 agonist talsaclidine decreases cerebrospinal fluid levels of A beta 42 in patients with Alzheimer's disease. *Amyloid*, *10*(1), 1-6.
- Homeyer, N., & Gohlke, H. (2012). Free Energy Calculations by the Molecular Mechanics Poisson-Boltzmann Surface Area Method. *Molecular Informatics*, *31*(2), 114-122.
- Hoofst, R. W., Vriend, G., Sander, C., & Abola, E. E. (1996). Errors in protein structures. *Nature*, *381*(6580), 272.
- Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., & Simmerling, C. (2006). Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins*, *65*(3), 712-725.

- Hou, T., Wang, J., Li, Y., & Wang, W. (2011). Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J Chem Inf Model*, 51(1), 69-82.
- Huang, S. Y., & Zou, X. (2007). Ensemble docking of multiple protein structures: considering protein structural variations in molecular docking. *Proteins*, 66(2), 399-421.
- Huang, X., Zheng, G., & Zhan, C.-G. (2012). Microscopic binding of M5 muscarinic acetylcholine receptor with antagonists by homology modeling, molecular docking, and molecular dynamics simulation. *The journal of physical chemistry. B*, 116(1), 532-541.
- Huang, X. P., Nagy, P. I., Williams, F. E., Peseckis, S. M., & Messer, W. S., Jr. (1999). Roles of threonine 192 and asparagine 382 in agonist and antagonist interactions with M1 muscarinic receptors. *British Journal of Pharmacology*, 126(3), 735-745.
- Hulme, E. C., Lu, Z. L., Saldanha, J. W., & Bee, M. S. (2003). Structure and activation of muscarinic acetylcholine receptors. *Biochemical Society Transactions*, 31(Pt 1), 29-34.
- Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: visual molecular dynamics. *Journal of Molecular Graphics*, 14(1), 33-38, 27.
- Irwin, J. J., & Shoichet, B. K. (2005). ZINC--a free database of commercially available compounds for virtual screening. *J Chem Inf Model*, 45(1), 177-182.
- Jakubik, J., Randakova, A., & Dolezal, V. (2013). On homology modeling of the M(2) muscarinic acetylcholine receptor subtype. *Journal of Computer-Aided Molecular Design*, 27(6), 525-538.
- Jiang, F., & Kim, S. H. (1991). "Soft docking": matching of molecular surface cubes. *Journal of Molecular Biology*, 219(1), 79-102.
- Jiang, S., Li, Y., Zhang, C., Zhao, Y., Bu, G., Xu, H., & Zhang, Y. W. (2014). M1 muscarinic acetylcholine receptor in Alzheimer's disease. *Neurosci Bull*, 30(2), 295-307.
- Jo, S., Kim, T., & Im, W. (2007). Automated builder and database of protein/membrane complexes for molecular dynamics simulations. *PLoS One*, 2(9), e880.
- Jo, S., Kim, T., Iyer, V. G., & Im, W. (2008). CHARMM-GUI: a web-based graphical user interface for CHARMM. *Journal of Computational Chemistry*, 29(11), 1859-1865.
- Johnson, D. J., Forbes, I. T., Watson, S. P., Garzya, V., Stevenson, G. I., Walker, G. R., Mudhar, H. S., Flynn, S. T., Wyman, P. A., Smith, P. W., Murkitt, G. S., Lucas, A. J., Mookherjee, C. R., Watson, J. M., Gartlon, J. E., Bradford, A. M., & Brown, F. (2010). The discovery of a series of N-substituted 3-(4-piperidinyl)-1,3-benzoxazolinones and

oxindoles as highly brain penetrant, selective muscarinic M1 agonists. *Bioorganic and Medicinal Chemistry Letters*, 20(18), 5434-5438.

- Jöhren, K., & Höltje, H.-D. (2005). Different environments for a realistic simulation of GPCRs-application to the M2 muscarinic receptor. *Archiv der Pharmazie*, 338(5-6), 260-267.
- Jórárt, B., Balint, A. M., Balint, S., & Viskolcz, B. (2012). Homology Modeling and Validation of the Human M1 Muscarinic Acetylcholine Receptor. *Molecular Informatics*, 31(9), 635-638.
- Jórárt, B., & Martinek, T. A. (2007). Performance of the general amber force field in modeling aqueous POPC membrane bilayers. *J Comput Chem*, 28(12), 2051-2058.
- Jones, C. K., Brady, A. E., Davis, A. A., Xiang, Z., Bubser, M., Tantawy, M. N., Kane, A. S., Bridges, T. M., Kennedy, J. P., Bradley, S. R., Peterson, T. E., Ansari, M. S., Baldwin, R. M., Kessler, R. M., Deutch, A. Y., Lah, J. J., Levey, A. I., Lindsley, C. W., & Conn, P. J. (2008). Novel selective allosteric activator of the M1 muscarinic acetylcholine receptor regulates amyloid processing and produces antipsychotic-like activity in rats. *The Journal of neuroscience*, 28(41), 10422-10433.
- Katritch, V., Cherezov, V., & Stevens, R. C. (2013). Structure-function of the G protein-coupled receptor superfamily. *Annual Review of Pharmacology and Toxicology*, 53, 531-556.
- Katritch, V., Jaakola, V. P., Lane, J. R., Lin, J., Ijzerman, A. P., Yeager, M., Kufareva, I., Stevens, R. C., & Abagyan, R. (2010). Structure-based discovery of novel chemotypes for adenosine A(2A) receptor antagonists. *Journal of Medicinal Chemistry*, 53(4), 1799-1809.
- Katritch, V., Rueda, M., Lam, P. C.-H., Yeager, M., & Abagyan, R. (2010). GPCR 3D homology models for ligand screening: lessons learned from blind predictions of adenosine A2a receptor complex. *Proteins*, 78(1), 197-211.
- Kaye, R. G., Saldanha, J. W., Lu, Z. L., & Hulme, E. C. (2011). Helix 8 of the M1 muscarinic acetylcholine receptor: scanning mutagenesis delineates a G protein recognition site. *Molecular Pharmacology*, 79(4), 701-709.
- Kenakin, T. (2004). Principles: receptor theory in pharmacology. *Trends in Pharmacological Sciences*, 25(4), 186-192.
- Kenakin, T. P. (2012). Biased signalling and allosteric machines: new vistas and challenges for drug discovery. *British Journal of Pharmacology*, 165(6), 1659-1669.
- Keov, P., Valant, C., Devine, S. M., Lane, J. R., Scammells, P. J., Sexton, P. M., & Christopoulos, A. (2013). Reverse engineering of the selective agonist TBPB unveils

both orthosteric and allosteric modes of action at the M<sub>1</sub> muscarinic acetylcholine receptor. *Molecular Pharmacology*, 84(3), 425-437.

Kiss, R., Jójárt, B., Schmidt, É., Kiss, B., & Keserű, G. M. (2014). Identification of Novel Histamine H4 Ligands by Virtual Screening on Molecular Dynamics Ensembles. *Molecular Informatics*, 33(4), 264-268.

Klauda, J. B., Venable, R. M., Freites, J. A., O'Connor, J. W., Tobias, D. J., Mondragon-Ramirez, C., Vorobyov, I., MacKerell, A. D., & Pastor, R. W. (2010). Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *The journal of physical chemistry. B*, 114(23), 7830-7843.

Kobilka, B. (2013). The structural basis of G-protein-coupled receptor signaling (Nobel Lecture). *Angewandte Chemie. International Ed. In English*, 52(25), 6380-6388.

Kołaczkowski, M., Bucki, A., Feder, M., & Pawłowski, M. (2013). Ligand-Optimized Homology Models of D1 and D2 Dopamine Receptors: Application for Virtual Screening. *J Chem Inf Model*, 53(3), 638-648.

Kolb, P., Rosenbaum, D. M., Irwin, J. J., Fung, J. J., Kobilka, B. K., & Shoichet, B. K. (2009). Structure-based discovery of beta2-adrenergic receptor ligands. *Proceedings of the National Academy of Sciences of the United States of America*, 106(16), 6843-6848.

Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., Lee, M., Lee, T., Duan, Y., Wang, W., Donini, O., Cieplak, P., Srinivasan, J., Case, D. A., & Cheatham, T. E. (2000). Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Accounts of Chemical Research*, 33(12), 889-897.

Korb, O., Olsson, T. S., Bowden, S. J., Hall, R. J., Verdonk, M. L., Liebeschuetz, J. W., & Cole, J. C. (2012). Potential and limitations of ensemble docking. *J Chem Inf Model*, 52(5), 1262-1274.

Koshland, D. E. (1995). The Key-Lock Theory and the Induced Fit Theory. *Angewandte Chemie International Edition in English*, 33(23-24), 2375-2378.

Krivov, G. G., Shapovalov, M. V., & Dunbrack, R. L. (2009). Improved prediction of protein side-chain conformations with SCWRL4. *Proteins*, 77(4), 778-795.

Kruse, A. C., Hu, J., Pan, A. C., Arlow, D. H., Rosenbaum, D. M., Rosemond, E., Green, H. F., Liu, T., Chae, P. S., Dror, R. O., Shaw, D. E., Weis, W. I., Wess, J., & Kobilka, B. K. (2012). Structure and dynamics of the M3 muscarinic acetylcholine receptor. *Nature*, 482(7386), 552-556.

- Kruse, A. C., Kobilka, B. K., Gautam, D., Sexton, P. M., Christopoulos, A., & Wess, J. (2014). Muscarinic acetylcholine receptors: novel opportunities for drug development. *Nature Reviews: Drug Discovery*, 13(7), 549-560.
- Kruse, A. C., Ring, A. M., Manglik, A., Hu, J., Hu, K., Eitel, K., Hubner, H., Pardon, E., Valant, C., Sexton, P. M., Christopoulos, A., Felder, C. C., Gmeiner, P., Steyaert, J., Weis, W. I., Garcia, K. C., Wess, J., & Kobilka, B. K. (2013). Activation and allosteric modulation of a muscarinic acetylcholine receptor. *Nature*, 504(7478), 101-106.
- Kruse, A. C., Weiss, D. R., Rossi, M., Hu, J., Hu, K., Eitel, K., Gmeiner, P., Wess, J., Kobilka, B. K., & Shoichet, B. K. (2013). Muscarinic receptors as model targets and antitargets for structure-based ligand discovery. *Molecular Pharmacology*, 84(4), 528-540.
- Krystek, S. R., Jr., Kimura, S. R., & Tebben, A. J. (2006). Modeling and active site refinement for G protein-coupled receptors: application to the beta-2 adrenergic receptor. *Journal of Computer-Aided Molecular Design*, 20(7-8), 463-470.
- Kucerka, N., Tristram-Nagle, S., & Nagle, J. F. (2005). Structure of fully hydrated fluid phase lipid bilayers with monounsaturated chains. *Journal of Membrane Biology*, 208(3), 193-202.
- Kuduk, S. D., & Beshore, D. C. (2012). Novel M(1) allosteric ligands: a patent review. *Expert opinion on therapeutic patents*, 22(12), 1385-1398.
- Kuduk, S. D., Chang, R. K., Di Marco, C. N., Ray, W. J., Ma, L., Wittmann, M., Seager, M. A., Koeplinger, K. A., Thompson, C. D., Hartman, G. D., & Bilodeau, M. T. (2010). Quinolizidinone Carboxylic Acids as CNS Penetrant, Selective M1 Allosteric Muscarinic Receptor Modulators. *ACS Medicinal Chemistry Letters*, 1(6), 263-267.
- Kuduk, S. D., Chang, R. K., Di Marco, C. N., Ray, W. J., Ma, L., Wittmann, M., Seager, M. A., Koeplinger, K. A., Thompson, C. D., Hartman, G. D., & Bilodeau, M. T. (2011). Quinolizidinone carboxylic acid selective M1 allosteric modulators: SAR in the piperidine series. *Bioorganic & Medicinal Chemistry Letters*, 21(6), 1710-1715.
- Kuduk, S. D., Chang, R. K., Greshock, T. J., Ray, W. J., Ma, L., Wittmann, M., Seager, M. A., Koeplinger, K. A., Thompson, C. D., Hartman, G. D., & Bilodeau, M. T. (2012). Identification of Amides as Carboxylic Acid Surrogates for Quinolizidinone-Based M1 Positive Allosteric Modulators. *ACS Medicinal Chemistry Letters*, 3(12), 1070-1074.
- Kuduk, S. D., Di Marco, C. N., Chang, R. K., Ray, W. J., Ma, L., Wittmann, M., Seager, M. A., Koeplinger, K. A., Thompson, C. D., Hartman, G. D., & Bilodeau, M. T. (2010). Heterocyclic fused pyridone carboxylic acid M1 positive allosteric modulators. *Bioorganic & Medicinal Chemistry Letters*, 20(8), 2533-2537.
- Kuduk, S. D., Di Marco, C. N., Cofre, V., Pitts, D. R., Ray, W. J., Ma, L., Wittmann, M., Seager, M., Koeplinger, K., Thompson, C. D., Hartman, G. D., & Bilodeau, M. T.

- (2010). Pyridine containing M1 positive allosteric modulators with reduced plasma protein binding. *Bioorganic & Medicinal Chemistry Letters*, 20(2), 657-661.
- Kuduk, S. D., Di Marco, C. N., Cofre, V., Pitts, D. R., Ray, W. J., Ma, L., Wittmann, M., Veng, L., Seager, M. A., Koeplinger, K., Thompson, C. D., Hartman, G. D., & Bilodeau, M. T. (2010). N-heterocyclic derived M1 positive allosteric modulators. *Bioorganic & Medicinal Chemistry Letters*, 20(4), 1334-1337.
- Kuduk, S. D., Di Marco, C. N., Cofre, V., Ray, W. J., Ma, L., Wittmann, M., Seager, M. A., Koeplinger, K. A., Thompson, C. D., Hartman, G. D., & Bilodeau, M. T. (2011). Fused heterocyclic M1 positive allosteric modulators. *Bioorganic & Medicinal Chemistry Letters*, 21(9), 2769-2772.
- Kuduk, S. D., DiPardo, R. M., Beshore, D. C., Ray, W. J., Ma, L., Wittmann, M., Seager, M. A., Koeplinger, K. A., Thompson, C. D., Hartman, G. D., & Bilodeau, M. T. (2010). Hydroxy cycloalkyl fused pyridone carboxylic acid M1 positive allosteric modulators. *Bioorganic & Medicinal Chemistry Letters*, 20(8), 2538-2541.
- Kufareva, I., Katritch, V., Stevens, R. C., & Abagyan, R. (2014). Advances in GPCR Modeling Evaluated by the GPCR Dock 2013 Assessment: Meeting New Challenges. *Structure*.
- Kufareva, I., Rueda, M., Katritch, V., Stevens, R. C., & Abagyan, R. (2011). Status of GPCR modeling and docking as reflected by community-wide GPCR Dock 2010 assessment. *Structure*, 19(8), 1108-1126.
- Langmead, C. J., & Christopoulos, A. (2006). Allosteric agonists of 7TM receptors: expanding the pharmacological toolbox. *Trends in Pharmacological Sciences*, 27(9), 475-481.
- Langmead, C. J., Fry, V. A. H., Forbes, I. T., Branch, C. L., Christopoulos, A., Wood, M. D., & Herdon, H. J. (2006). Probing the molecular mechanism of interaction between 4-n-butyl-1-[4-(2-methylphenyl)-4-oxo-1-butyl]-piperidine (AC-42) and the muscarinic M(1) receptor: direct pharmacological evidence that AC-42 is an allosteric agonist. *Molecular Pharmacology*, 69(1), 236-246.
- Langmead, C. J., Watson, J., & Reavill, C. (2008). Muscarinic acetylcholine receptors as CNS drug targets. *Pharmacology & Therapeutics*, 117(2), 232-243.
- Larsson, P., & Kasson, P. M. (2013). Lipid tail protrusion in simulations predicts fusogenic activity of influenza fusion peptide mutants and conformational models. *PLoS Computational Biology*, 9(3), e1002950.
- Laskowski, R. A., Macarthur, M. W., Moss, D. S., & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *Journal of Applied Crystallography*, 26, 283-291.

- Lazareno, S., Gharagozloo, P., Kuonen, D., Popham, A., & Birdsall, N. J. (1998). Subtype-selective positive cooperative interactions between brucine analogues and acetylcholine at muscarinic receptors: radioligand binding studies. *Molecular Pharmacology*, *53*(3), 573-589.
- Le Guilloux, V., Schmidtke, P., & Tuffery, P. (2009). Fpocket: an open source platform for ligand pocket detection. *BMC Bioinformatics*, *10*, 168.
- Leach, A. R. (1994). Ligand docking to proteins with discrete side-chain flexibility. *Journal of Molecular Biology*, *235*(1), 345-356.
- Leach, A. R. (2001). *Molecular modeling: principles and application* (second ed.). Harlow: Pearson education.
- Lebois, E. P., Bridges, T. M., Lewis, L. M., Dawson, E. S., Kane, A. S., Xiang, Z., Jadhav, S. B., Yin, H., Kennedy, J. P., Meiler, J., Niswender, C. M., Jones, C. K., Conn, P. J., Weaver, C. D., & Lindsley, C. W. (2010). Discovery and characterization of novel subtype-selective allosteric agonists for the investigation of M(1) receptor function in the central nervous system. *ACS Chem Neurosci*, *1*(2), 104-121.
- Lebois, E. P., Digby, G. J., Sheffler, D. J., Melancon, B. J., Tarr, J. C., Cho, H. P., Miller, N. R., Morrison, R., Bridges, T. M., Xiang, Z., Daniels, J. S., Wood, M. R., Conn, P. J., & Lindsley, C. W. (2011). Development of a highly selective, orally bioavailable and CNS penetrant M1 agonist derived from the MLPCN probe ML071. *Bioorganic and Medicinal Chemistry Letters*, *21*(21), 6451-6455.
- Lebon, G., Langmead, C. J., Tehan, B. G., & Hulme, E. C. (2009). Mutagenic mapping suggests a novel binding mode for selective agonists of M1 muscarinic acetylcholine receptors. *Molecular Pharmacology*, *75*(2), 331-341.
- Lebon, G., Warne, T., & Tate, C. G. (2012). Agonist-bound structures of G protein-coupled receptors. *Current Opinion in Structural Biology*, *22*(4), 482-490.
- Lefkowitz, R. J., & Shenoy, S. K. (2005). Transduction of receptor signals by beta-arrestins. *Science*, *308*(5721), 512-517.
- Lounnas, V., Ritschel, T., Kelder, J., McGuire, R., Bywater, R. P., & Foloppe, N. (2013). Current progress in Structure-Based Rational Drug Design marks a new mindset in drug discovery. *Comput Struct Biotechnol J*, *5*, e201302011.
- Lu, Z. L., & Hulme, E. C. (1999). The functional topography of transmembrane domain 3 of the M1 muscarinic acetylcholine receptor, revealed by scanning mutagenesis. *The Journal of biological chemistry*, *274*(11), 7309-7315.

- Lu, Z. L., Saldanha, J. W., & Hulme, E. C. (2001). Transmembrane domains 4 and 7 of the M(1) muscarinic acetylcholine receptor are critical for ligand binding and the receptor activation switch. *The Journal of biological chemistry*, 276(36), 34098-34104.
- Ma, L., Seager, M. A., Seager, M., Wittmann, M., Jacobson, M., Bickel, D., Burno, M., Jones, K., Graufelds, V. K., Xu, G., Pearson, M., McCampbell, A., Gaspar, R., Shughrue, P., Danziger, A., Regan, C., Flick, R., Pascarella, D., Garson, S., Doran, S., Kreatsoulas, C., Veng, L., Lindsley, C. W., Shipe, W., Kuduk, S., Sur, C., Kinney, G., Seabrook, G. R., & Ray, W. J. (2009). Selective activation of the M1 muscarinic acetylcholine receptor achieved by allosteric potentiation. *Proceedings of the National Academy of Sciences of the United States of America*, 106(37), 15950-15955.
- MacKerell, A. D., Bashford, D., Bellott, M., Dunbrack, R. L., Evanseck, J. D., Field, M. J., Fischer, S., Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T., Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., Wiorkiewicz-Kuczera, J., Yin, D., & Karplus, M. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B*, 102(18), 3586-3616.
- Marrink, S. J., & Mark, A. E. (2001). Effect of Undulations on Surface Tension in Simulated Bilayers. *The Journal of Physical Chemistry B*, 105(26), 6122-6127.
- Martínez-Archundia, M., & Correa-Basurto, J. (2013). Molecular dynamics simulations reveal initial structural and dynamic features for the A2AR as a result of ligand binding. *Molecular Simulation*, 40(13), 996-1014.
- Martinez-Archundia, M., Cordomi, A., Garriga, P., & Perez, J. J. (2012). Molecular modeling of the M3 acetylcholine muscarinic receptor and its binding site. *Journal of biomedicine & biotechnology*, 2012, 789741.
- McMillin, S. M., Heusel, M., Liu, T., Costanzi, S., & Wess, J. (2011). Structural basis of M3 muscarinic receptor dimer/oligomer formation. *Journal of Biological Chemistry*, 286(32), 28584-28598.
- McRobb, F. M., Capuano, B., Crosby, I. T., Chalmers, D. K., & Yuriev, E. (2010). Homology modeling and docking evaluation of aminergic G protein-coupled receptors. *J Chem Inf Model*, 50(4), 626-637.
- Mezei, M., & Filizola, M. (2006). TRAJELIX: a computational tool for the geometric characterization of protein helices during molecular dynamics simulations. *Journal of Computer-Aided Molecular Design*, 20(2), 97-107.
- Miao, Y., Nichols, S. E., Gasper, P. M., Metzger, V. T., & McCammon, J. A. (2013). Activation and dynamic network of the M2 muscarinic receptor. *Proceedings of the National Academy of Sciences of the United States of America*, 110(27), 10982-10987.

- Michino, M., Abola, E., Brooks, C. L., 3rd, Dixon, J. S., Moulton, J., & Stevens, R. C. (2009). Community-wide assessment of GPCR structure modelling and ligand docking: GPCR Dock 2008. *Nature Reviews: Drug Discovery*, 8(6), 455-463.
- Miller, B. R., McGee, T. D., Swails, J. M., Homeyer, N., Gohlke, H., & Roitberg, A. E. (2012). MMPBSA.py: An Efficient Program for End-State Free Energy Calculations. *Journal of Chemical Theory and Computation*, 8(9), 3314-3321.
- Mirza, N. R., Peters, D., & Sparks, R. G. (2003). Xanomeline and the antipsychotic potential of muscarinic receptor subtype selective agonists. *CNS drug reviews*, 9(2), 159-186.
- Mirzadegan, T., Benko, G., Filipek, S., & Palczewski, K. (2003). Sequence analyses of G-protein-coupled receptors: similarities to rhodopsin. *Biochemistry*, 42(10), 2759-2767.
- Mohr, K., Schmitz, J., Schrage, R., Trankle, C., & Holzgrabe, U. (2013). Molecular alliance—from orthosteric and allosteric ligands to dualsteric/bitopic agonists at G protein coupled receptors. *Angewandte Chemie. International Ed. In English*, 52(2), 508-516.
- Mordalski, S., Kosciolk, T., Kristiansen, K., Sylte, I., & Bojarski, A. J. (2011). Protein binding site analysis by means of structural interaction fingerprint patterns. *Bioorganic & Medicinal Chemistry Letters*, 21(22), 6816-6819.
- Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., & Olson, A. J. (1998). Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *Journal of Computational Chemistry*, 19(14), 1639-1662.
- Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem*, 30(16), 2785-2791.
- Mulakala, C., & Viswanadhan, V. N. (2013). Could MM-GBSA be accurate enough for calculation of absolute protein/ligand binding free energies? *Journal of Molecular Graphics and Modelling*, 46, 41-51.
- Mysinger, M. M., Weiss, D. R., Ziarek, J. J., Gravel, S., Doak, A. K., Karpiak, J., Heveker, N., Shoichet, B. K., & Volkman, B. F. (2012). Structure-based ligand discovery for the protein-protein interface of chemokine receptor CXCR4. *Proceedings of the National Academy of Sciences of the United States of America*, 109(14), 5517-5522.
- Nathan, P. J., Watson, J., Lund, J., Davies, C. H., Peters, G., Dodds, C. M., Swirski, B., Lawrence, P., Bentley, G. D., O'Neill, B. V., Robertson, J., Watson, S., Jones, G. A., Maruff, P., Croft, R. J., Laruelle, M., & Bullmore, E. T. (2013). The potent M1 receptor allosteric agonist GSK1034702 improves episodic memory in humans in the nicotine abstinence model of cognitive dysfunction. *The international journal of neuropsychopharmacology / official scientific journal of the Collegium Internationale Neuropsychopharmacologicum (CINP)*, 16(4), 721-731.

- Negri, A., Rives, M. L., Caspers, M. J., Prisinzano, T. E., Javitch, J. A., & Filizola, M. (2013). Discovery of a novel selective kappa-opioid receptor agonist using crystal structure-based virtual screening. *J Chem Inf Model*, 53(3), 521-526.
- Nickols, H. H., & Conn, P. J. (2014a). Development of allosteric modulators of GPCRs for treatment of CNS disorders. *Neurobiology of Disease*, 61, 55-71.
- Novikov, G. V., Sivozhelezov, V. S., Kolesnikov, S. S., & Shaitan, K. V. (2014). Investigation of the influence of external factors on the conformational dynamics of rhodopsin-like receptors by means of molecular dynamics simulation. *Journal of Receptor and Signal Transduction Research*, 34(2), 104-118.
- Novikov, G. V., Sivozhelezov, V. S., & Shaitan, K. V. (2013). Study of structural dynamics of ligand-activated membrane receptors by means of principal component analysis. *Biochemistry*, 78(4), 403-411.
- Novoa, E. M., Pouplana, L. R. d., Barril, X., & Orozco, M. (2010). Ensemble Docking from Homology Models. *Journal of Chemical Theory and Computation*, 6(8), 2547-2557.
- Nurisso, A., Daina, A., & Walker, R. (2012). A Practical Introduction to Molecular Dynamics Simulations: Applications to Homology Modeling. In A. J. W. Orry & R. Abagyan (Eds.), *Homology Modeling* (Vol. 857, pp. 137-173): Humana Press.
- Okuno, Y., Tamon, A., Yabuuchi, H., Nijima, S., Minowa, Y., Tonomura, K., Kunimoto, R., & Feng, C. (2008). GLIDA: GPCR--ligand database for chemical genomics drug discovery--database and tools update. *Nucleic Acids Research*, 36(Database issue), D907-912.
- Orry, A. J., & Abagyan, R. (2012). Preparation and refinement of model protein-ligand complexes. *Methods in Molecular Biology*, 857, 351-373.
- Overk, C. R., Felder, C. C., Tu, Y., Schober, D. A., Bales, K. R., Wu, J., & Mufson, E. J. (2010). Cortical M1 receptor concentration increases without a concomitant change in function in Alzheimer's disease. *Journal of Chemical Neuroanatomy*, 40(1), 63-70.
- Pajouhesh, H., & Lenz, G. R. (2005). Medicinal chemical properties of successful central nervous system drugs. *NeuroRx*, 2(4), 541-553.
- Pala, D., Beuming, T., Sherman, W., Lodola, A., Rivara, S., & Mor, M. (2013). Structure-based virtual screening of MT2 melatonin receptor: influence of template choice and structural refinement. *J Chem Inf Model*, 53(4), 821-835.
- Palczewski, K., Kumasaka, T., Hori, T., Behnke, C. A., Motoshima, H., Fox, B. A., Le Trong, I., Teller, D. C., Okada, T., Stenkamp, R. E., Yamamoto, M., & Miyano, M. (2000). Crystal structure of rhodopsin: A G protein-coupled receptor. *Science (New York, N.Y.)*, 289(5480), 739-745.

- Pardo, L., Deupi, X., Dolker, N., Lopez-Rodriguez, M. L., & Campillo, M. (2007). The role of internal water molecules in the structure and function of the rhodopsin family of G protein-coupled receptors. *ChemBioChem*, 8(1), 19-24.
- Patra, M. C., Maharana, J., Dehury, B., & De, S. (2014). Computational insights into the binding mechanism of antagonists with neuropeptide B/W receptor 1. *Mol Biosyst*, 10(8), 2236-2246.
- Peeters, M. C., van Westen, G. J., Li, Q., & AP, I. J. (2011). Importance of the extracellular loops in G protein-coupled receptors for ligand recognition and receptor activation. *Trends in Pharmacological Sciences*, 32(1), 35-42.
- Peng, J. Y.-c., Vaidehi, N., Hall, S. E., & Goddard, W. A. (2006). The predicted 3D structures of the human M1 muscarinic acetylcholine receptor with agonist or antagonist bound. *ChemMedChem*, 1(8), 878-890.
- Pittel, Z., Heldman, E., Barg, J., Haring, R., & Fisher, A. (1996). Muscarinic control of amyloid precursor protein secretion in rat cerebral cortex and cerebellum. *Brain Research*, 742(1-2), 299-304.
- Plesnar, E., Subczynski, W. K., & Pasenkiewicz-Gierula, M. (2012). Saturation with cholesterol increases vertical order and smoothes the surface of the phosphatidylcholine bilayer: a molecular simulation study. *Biochimica et Biophysica Acta*, 1818(3), 520-529.
- Rasmussen, S. G., Choi, H. J., Fung, J. J., Pardon, E., Casarosa, P., Chae, P. S., Devree, B. T., Rosenbaum, D. M., Thian, F. S., Kobilka, T. S., Schnapp, A., Konetzki, I., Sunahara, R. K., Gellman, S. H., Pautsch, A., Steyaert, J., Weis, W. I., & Kobilka, B. K. (2011). Structure of a nanobody-stabilized active state of the beta(2) adrenoceptor. *Nature*, 469(7329), 175-180.
- Rasmussen, S. G., Choi, H. J., Rosenbaum, D. M., Kobilka, T. S., Thian, F. S., Edwards, P. C., Burghammer, M., Ratnala, V. R., Sanishvili, R., Fischetti, R. F., Schertler, G. F., Weis, W. I., & Kobilka, B. K. (2007). Crystal structure of the human beta2 adrenergic G-protein-coupled receptor. *Nature*, 450(7168), 383-387.
- 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. (2011). Crystal structure of the beta2 adrenergic receptor-Gs protein complex. *Nature*, 477(7366), 549-555.
- Rataj, K., Witek, J., Mordalski, S., Kosciolk, T., & Bojarski, A. J. (2014). Impact of template choice on homology model efficiency in virtual screening. *J Chem Inf Model*, 54(6), 1661-1668.
- Reichel, A. (2006). The role of blood-brain barrier studies in the pharmaceutical industry. *Current Drug Metabolism*, 7(2), 183-203.

- Roe, D. R., & Cheatham, T. E. (2013). PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *Journal of Chemical Theory and Computation*, 9(7), 3084-3095.
- Rosenbaum, D. M., Cherezov, V., Hanson, M. A., Rasmussen, S. G. F., Thian, F. S., Kobilka, T. S., Choi, H.-J., Yao, X.-J., Weis, W. I., Stevens, R. C., & Kobilka, B. K. (2007). GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function. *Science (New York, N.Y.)*, 318(5854), 1266-1273.
- Roumen, L., Sanders, M. P. A., Vroiling, B., De Esch, I. J. P., De Vlieg, J., Leurs, R., Klomp, J. P. G., Nabuurs, S. B., & De Graaf, C. (2011). In Silico Veritas: The Pitfalls and Challenges of Predicting GPCR-Ligand Interactions. *Pharmaceuticals*, 4(9), 1196-1215.
- Ryckaert, J.-P., Ciccotti, G., & Berendsen, H. J. C. (1977). Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics*, 23(3), 327-341.
- Sadot, E., Gurwitz, D., Barg, J., Behar, L., Ginzburg, I., & Fisher, A. (1996). Activation of m1 muscarinic acetylcholine receptor regulates tau phosphorylation in transfected PC12 cells. *Journal of Neurochemistry*, 66(2), 877-880.
- Salon, J. A., Lodowski, D. T., & Palczewski, K. (2011). The significance of G protein-coupled receptor crystallography for drug discovery. *Pharmacological Reviews*, 63(4), 901-937.
- Sanders, J. M., Wampole, M. E., Thakur, M. L., & Wickstrom, E. (2013). Molecular determinants of epidermal growth factor binding: a molecular dynamics study. *PLoS One*, 8(1), e54136.
- Sansom, M. S., & Weinstein, H. (2000). Hinges, swivels and switches: the role of prolines in signalling via transmembrane alpha-helices. *Trends in Pharmacological Sciences*, 21(11), 445-451.
- Schiöth, H. B., & Fredriksson, R. (2005). The GRAFS classification system of G-protein coupled receptors in comparative perspective. *General and Comparative Endocrinology*, 142(1-2), 94-101.
- Schmidtke, P., Bidon-Chanal, A., Luque, F. J., & Barril, X. (2011). MDpocket: open-source cavity detection and characterization on molecular dynamics trajectories. *Bioinformatics (Oxford, England)*, 27(23), 3276-3285.
- Seibt, B. F., Schiedel, A. C., Thimm, D., Hinz, S., Sherbiny, F. F., & Muller, C. E. (2013). The second extracellular loop of GPCRs determines subtype-selectivity and controls efficacy as evidenced by loop exchange study at A2 adenosine receptors. *Biochemical Pharmacology*, 85(9), 1317-1329.

- Selkoe, D. J. (1991). The molecular pathology of Alzheimer's disease. *Neuron*, 6(4), 487-498.
- Sharman, J. L., Mpamhanga, C. P., Spedding, M., Germain, P., Staels, B., Dacquet, C., Laudet, V., Harmar, A. J., & Ne, I. (2011). IUPHAR-DB: new receptors and tools for easy searching and visualization of pharmacological data. *Nucleic Acids Research*, 39(Database issue), D534-538.
- Shen, M., Zhou, S., Li, Y., Pan, P., Zhang, L., & Hou, T. (2013). Discovery and optimization of triazine derivatives as ROCK1 inhibitors: molecular docking, molecular dynamics simulations and free energy calculations. *Mol Biosyst*, 9(3), 361-374.
- Sherman, W., Beard, H. S., & Farid, R. (2006). Use of an induced fit receptor structure in virtual screening. *Chem Biol Drug Des*, 67(1), 83-84.
- Sherman, W., Day, T., Jacobson, M. P., Friesner, R. A., & Farid, R. (2005). Novel Procedure for Modeling Ligand/Receptor Induced Fit Effects. *Journal of medicinal chemistry*, 49(2), 534-553.
- Sherman, W., Day, T., Jacobson, M. P., Friesner, R. A., & Farid, R. (2006). Novel procedure for modeling ligand/receptor induced fit effects. *Journal of Medicinal Chemistry*, 49(2), 534-553.
- Shirey, J. K., Brady, A. E., Jones, P. J., Davis, A. A., Bridges, T. M., Kennedy, J. P., Jadhav, S. B., Menon, U. N., Xiang, Z., Watson, M. L., Christian, E. P., Doherty, J. J., Quirk, M. C., Snyder, D. H., Lah, J. J., Levey, A. I., Nicolle, M. M., Lindsley, C. W., & Conn, P. J. (2009). A selective allosteric potentiator of the M1 muscarinic acetylcholine receptor increases activity of medial prefrontal cortical neurons and restores impairments in reversal learning. *The Journal of neuroscience*, 29(45), 14271-14286.
- Shirey, J. K., Xiang, Z., Orton, D., Brady, A. E., Johnson, K. A., Williams, R., Ayala, J. E., Rodriguez, A. L., Wess, J., Weaver, D., Niswender, C. M., & Conn, P. J. (2008). An allosteric potentiator of M4 mAChR modulates hippocampal synaptic transmission. *Nature Chemical Biology*, 4(1), 42-50.
- Shoichet, B. K., & Kobilka, B. K. (2012). Structure-based drug screening for G-protein-coupled receptors. *Trends in Pharmacological Sciences*, 33(5), 268-272.
- Spalding, T. A., Burstein, E. S., Henderson, S. C., Ducote, K. R., & Brann, M. R. (1998). Identification of a ligand-dependent switch within a muscarinic receptor. *Journal of Biological Chemistry*, 273(34), 21563-21568.
- Spalding, T. A., Ma, J. N., Ott, T. R., Friberg, M., Bajpai, A., Bradley, S. R., Davis, R. E., Brann, M. R., & Burstein, E. S. (2006a). Structural requirements of transmembrane domain 3 for activation by the M1 muscarinic receptor agonists AC-42, AC-260584, clozapine, and N-desmethyleclozapine: evidence for three distinct modes of receptor activation. *Molecular Pharmacology*, 70(6), 1974-1983.

- Spalding, T. A., Trotter, C., Skjaerbaek, N., Messier, T. L., Currier, E. A., Burstein, E. S., Li, D., Hacksell, U., & Brann, M. R. (2002). Discovery of an ectopic activation site on the M(1) muscarinic receptor. *Molecular Pharmacology*, *61*(6), 1297-1302.
- Srinivasan, J., Cheatham, T. E., Cieplak, P., Kollman, P. A., & Case, D. A. (1998). Continuum Solvent Studies of the Stability of DNA, RNA, and Phosphoramidate–DNA Helices. *Journal of the American Chemical Society*, *120*(37), 9401-9409.
- Standfuss, J., Edwards, P. C., D'Antona, A., Fransen, M., Xie, G., Oprian, D. D., & Schertler, G. F. (2011). The structural basis of agonist-induced activation in constitutively active rhodopsin. *Nature*, *471*(7340), 656-660.
- Strader, C. D., Fong, T. M., Tota, M. R., Underwood, D., & Dixon, R. A. (1994). Structure and function of G protein-coupled receptors. *Annual Review of Biochemistry*, *63*, 101-132.
- Sur, C., Mallorga, P. J., Wittmann, M., Jacobson, M. A., Pascarella, D., Williams, J. B., Brandish, P. E., Pettibone, D. J., Scolnick, E. M., & Conn, P. J. (2003). N-desmethylclozapine, an allosteric agonist at muscarinic 1 receptor, potentiates N-methyl-D-aspartate receptor activity. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(23), 13674-13679.
- Swaminathan, M., Chee, C. F., Chin, S. P., Buckle, M. J., Rahman, N. A., Doughty, S. W., & Chung, L. Y. (2014). Flavonoids with m1 muscarinic acetylcholine receptor binding activity. *Molecules*, *19*(7), 8933-8948.
- Swets, J. A. (1988). Measuring the accuracy of diagnostic systems. *Science (New York, N.Y.)*, *240*(4857), 1285-1293.
- Taddese, B., Simpson, L. M., Wall, I. D., Blaney, F. E., & Reynolds, C. A. (2013). Modeling active GPCR conformations. *Methods in Enzymology*, *522*, 21-35.
- Tarcsay, A., Paragi, G., Vass, M., Jojart, B., Bogar, F., & Keseru, G. M. (2013). The impact of molecular dynamics sampling on the performance of virtual screening against GPCRs. *J Chem Inf Model*, *53*(11), 2990-2999.
- Tate, C. G. (2012). A crystal clear solution for determining G-protein-coupled receptor structures. *Trends in Biochemical Sciences*, *37*(9), 343-352.
- Taylor, P., & Brown, J. H. (2006). Acetylcholine. In G. J. Siegel, R. W. Albers, S. T. Brady & D. L. Price (Eds.), *Basic neurochemistry: molecular, cellular, and medical aspects*. (7th ed., pp. 185-209). Burlington, Massachusetts: Elsevier Academic Press.
- Tehan, B. G., Bortolato, A., Blaney, F. E., Weir, M. P., & Mason, J. S. (2014). Unifying family A GPCR theories of activation. *Pharmacology & Therapeutics*, *143*(1), 51-60.

- Thomas, T., McLean, K. C., McRobb, F. M., Manallack, D. T., Chalmers, D. K., & Yuriev, E. (2014). Homology modeling of human muscarinic acetylcholine receptors. *J Chem Inf Model*, 54(1), 243-253.
- Topiol, S., & Sabio, M. (2009). X-ray structure breakthroughs in the GPCR transmembrane region. *Biochemical Pharmacology*, 78(1), 11-20.
- Triballeau, N., Acher, F., Brabet, I., Pin, J.-P., & Bertrand, H.-O. (2005). Virtual screening workflow development guided by the "receiver operating characteristic" curve approach. Application to high-throughput docking on metabotropic glutamate receptor subtype 4. *Journal of Medicinal Chemistry*, 48(7), 2534-2547.
- Truchon, J.-F., & Bayly, C. I. (2007). Evaluating virtual screening methods: good and bad metrics for the "early recognition" problem. *J Chem Inf Model*, 47(2), 488-508.
- Trzaskowski, B., Latek, D., Yuan, S., Ghoshdastider, U., Debinski, A., & Filipek, S. (2012). Action of molecular switches in GPCRs--theoretical and experimental studies. *Current Medicinal Chemistry*, 19(8), 1090-1109.
- Urizar, E., Claeysen, S., Deupi, X., Govaerts, C., Costagliola, S., Vassart, G., & Pardo, L. (2005). An activation switch in the rhodopsin family of G protein-coupled receptors: the thyrotropin receptor. *Journal of Biological Chemistry*, 280(17), 17135-17141.
- Uslaner, J. M., Eddins, D., Puri, V., Cannon, C. E., Sutcliffe, J., Chew, C. S., Pearson, M., Vivian, J. A., Chang, R. K., Ray, W. J., Kuduk, S. D., & Wittmann, M. (2013). The muscarinic M1 receptor positive allosteric modulator PQCA improves cognitive measures in rat, cynomolgus macaque, and rhesus macaque. *Psychopharmacology*, 225(1), 21-30.
- Vaidehi, N., Bhattacharya, S., & Larsen, A. B. (2014). Structure and dynamics of G-protein coupled receptors. *Advances in Experimental Medicine and Biology*, 796, 37-54.
- Vaidehi, N., Floriano, W. B., Trabanino, R., Hall, S. E., Freddolino, P., Choi, E. J., Zamanakos, G., & Goddard, W. A., 3rd. (2002). Prediction of structure and function of G protein-coupled receptors. *Proceedings of the National Academy of Sciences of the United States of America*, 99(20), 12622-12627.
- Valant, C., Gregory, K. J., Hall, N. E., Scammells, P. J., Lew, M. J., Sexton, P. M., & Christopoulos, A. (2008). A novel mechanism of G protein-coupled receptor functional selectivity. Muscarinic partial agonist McN-A-343 as a bitopic orthosteric/allosteric ligand. *Journal of Biological Chemistry*, 283(43), 29312-29321.
- Valant, C., Sexton, P. M., & Christopoulos, A. (2009). Orthosteric/allosteric bitopic ligands: going hybrid at GPCRs. *Mol Interv*, 9(3), 125-135.

- Vass, M., Schmidt, E., Horti, F., & Keseru, G. M. (2014). Virtual fragment screening on GPCRs: a case study on dopamine D3 and histamine H4 receptors. *Eur J Med Chem*, 77, 38-46.
- Venclovas, C. (2012). Methods for sequence-structure alignment. *Methods in Molecular Biology*, 857, 55-82.
- Venkatakrishnan, A. J., Deupi, X., Lebon, G., Tate, C. G., Schertler, G. F., & Babu, M. M. (2013). Molecular signatures of G-protein-coupled receptors. *Nature*, 494(7436), 185-194.
- Vilar, S., Ferino, G., Phatak, S. S., Berk, B., Cavasotto, C. N., & Costanzi, S. (2011). Docking-based virtual screening for ligands of G protein-coupled receptors: not only crystal structures but also in silico models. *Journal of Molecular Graphics and Modelling*, 29(5), 614-623.
- Vinh, N. B., Simpson, J. S., Scammells, P. J., & Chalmers, D. K. (2012). Virtual screening using a conformationally flexible target protein: models for ligand binding to p38alpha MAPK. *Journal of Computer-Aided Molecular Design*, 26(4), 409-423.
- Violin, J. D., & Lefkowitz, R. J. (2007). Beta-arrestin-biased ligands at seven-transmembrane receptors. *Trends in Pharmacological Sciences*, 28(8), 416-422.
- Vistoli, G., Pedretti, A., Testa, B., & Maticucci, R. (2007). The conformational and property space of acetylcholine bound to muscarinic receptors: an entropy component accounts for the subtype selectivity of acetylcholine. *Archives of Biochemistry and Biophysics*, 464(1), 112-121.
- Ward, S. D., Curtis, C. A., & Hulme, E. C. (1999). Alanine-scanning mutagenesis of transmembrane domain 6 of the M(1) muscarinic acetylcholine receptor suggests that Tyr381 plays key roles in receptor function. *Molecular Pharmacology*, 56(5), 1031-1041.
- Warne, T., Moukhametzianov, R., Baker, J. G., Nehme, R., Edwards, P. C., Leslie, A. G., Schertler, G. F., & Tate, C. G. (2011). The structural basis for agonist and partial agonist action on a beta(1)-adrenergic receptor. *Nature*, 469(7329), 241-244.
- Weiss, D. R., Ahn, S., Sassano, M. F., Kleist, A., Zhu, X., Strachan, R., Roth, B. L., Lefkowitz, R. J., & Shoichet, B. K. (2013). Conformation Guides Molecular Efficacy in Docking Screens of Activated  $\beta_2$  Adrenergic G Protein Coupled Receptor. *ACS chemical biology*, 8(5), 1018-1026.
- Wess, J., Han, S. J., Kim, S. K., Jacobson, K. A., & Li, J. H. (2008). Conformational changes involved in G-protein-coupled-receptor activation. *Trends in Pharmacological Sciences*, 29(12), 616-625.

- Woolf, N. J. (1996). The critical role of cholinergic basal forebrain neurons in morphological change and memory encoding: a hypothesis. *Neurobiology of Learning and Memory*, 66(3), 258-266.
- Xhaard, H., Nyrönen, T., Rantanen, V.-V., Ruuskanen, J. O., Laurila, J., Salminen, T., Scheinin, M., & Johnson, M. S. (2005). Model structures of alpha-2 adrenoceptors in complex with automatically docked antagonist ligands raise the possibility of interactions dissimilar from agonist ligands. *Journal of Structural Biology*, 150(2), 126-143.
- Xu, H., Greengard, P., & Gandy, S. (1995). Regulated formation of Golgi secretory vesicles containing Alzheimer beta-amyloid precursor protein. *Journal of Biological Chemistry*, 270(40), 23243-23245.
- Xu, L., Li, Y., Sun, H., Li, D., & Hou, T. (2013). Structural basis of the interactions between CXCR4 and CXCL12/SDF-1 revealed by theoretical approaches. *Mol Biosyst*, 9(8), 2107-2117.
- Xu, M., & Lill, M. A. (2013). Induced fit docking, and the use of QM/MM methods in docking. *Drug discovery today. Technologies*, 10(3), e411-418.
- Yarnitzky, T., Levit, A., & Niv, M. Y. (2010). Homology modeling of G-protein-coupled receptors with X-ray structures on the rise. *Curr Opin Drug Discov Devel*, 13(3), 317-325.
- Yuriev, E., Agostino, M., & Ramsland, P. A. (2011). Challenges and advances in computational docking: 2009 in review. *Journal of molecular recognition : JMR*, 24(2), 149-164.
- Yuriev, E., & Ramsland, P. A. (2013). Latest developments in molecular docking: 2010-2011 in review. *Journal of molecular recognition : JMR*, 26(5), 215-239.
- Zemek, F., Drtinova, L., Nepovimova, E., Sepsova, V., Korabecny, J., Klimes, J., & Kuca, K. (2014). Outcomes of Alzheimer's disease therapy with acetylcholinesterase inhibitors and memantine. *Expert Opinion on Drug Safety*, 13(6), 759-774.
- Zhang, L., & Hermans, J. (1996). Hydrophilicity of cavities in proteins. *Proteins*, 24(4), 433-438.
- Zhang, Y., Devries, M. E., & Skolnick, J. (2006). Structure modeling of all identified G protein-coupled receptors in the human genome. *PLoS Comput Biol*, 2(2), e13.
- Zhu, X. L., Yu, N. X., Hao, G. F., Yang, W. C., & Yang, G. F. (2013). Structural basis of femtomolar inhibitors for acetylcholinesterase subtype selectivity: insights from computational simulations. *Journal of Molecular Graphics and Modelling*, 41, 55-60.