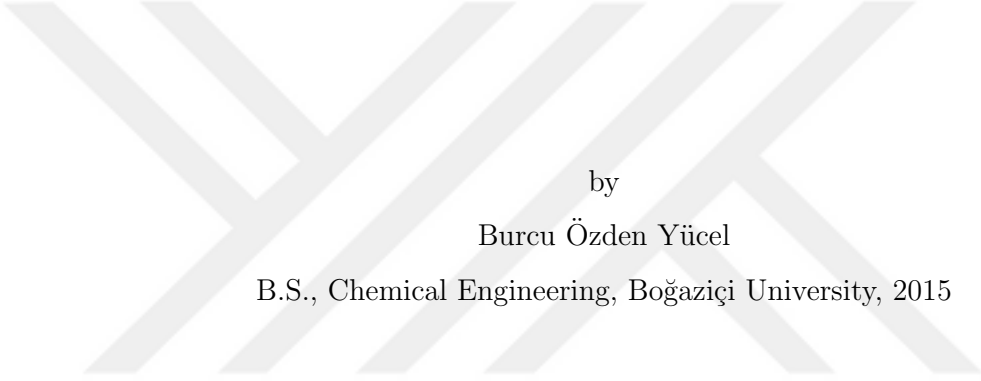


EXPLORING FUNCTIONAL MOTIONS OF MUSCARINIC ACETYLCHOLINE
M2 RECEPTOR



by
Burcu Özden Yücel
B.S., Chemical Engineering, Boğaziçi University, 2015

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Chemical Engineering
Boğaziçi University

2018

EXPLORING FUNCTIONAL MOTIONS OF MUSCARINIC ACETYLCHOLINE
M2 RECEPTOR

APPROVED BY:

Prof. Türkan Haliloğlu
(Thesis Supervisor)

Prof. Pemra Doruker

Assoc. Prof. Pemra Özbek

DATE OF APPROVAL: 01.06.2018

ACKNOWLEDGEMENTS

First and foremost, I would like to express my sincere thanks to my thesis supervisor, Prof. Türkan Halilođlu. It was a great pleasure to work with her. I would never achieve such success in my studies without her motivation, continuous support and understanding.

I'm grateful to the members of my thesis committee Prof. Pemra Doruker and Assoc. Prof. Pemra Özbek for valuable and insightful recommendations.

I would like to thank Elon Yariv, who provided me cavity calculation results of M2 receptor through transition pathway and contributed to my study generously.

I would like to give my deepest thanks Burçin Acar and Ayşe Burcu Aykaç Fas for their voluntary guidance throughout my research. I would like to thank everyone in my PRC family; Erge Akbař, Yiđit Kutlu, Gökçehan Kara for their support, advices and friendship. I am so grateful for all we shared together.

I am truly indebted to Burçin Camcı and Selin Erkan for their endless support and sharing all the memories with me. Special thanks to Fulya Bařađaç and Utku Can for being my second home in İstanbul and their invaluable friendship.

It was a great opportunity for me that my research has been supported by the projects TÜBİTAK 1001 (15M418), Bođaziçi University B.A.P. (11160/17A05M3) and Betil Fund.

And last but not least, I am thankful to my family for always being there for me with endless love, encouragement and support. Special thanks to my father for being an inspiration for my study and my husband for his invaluable support and endless love. I'm grateful to my mum being a perfect role model in my life and my sister for being next to me whenever I needed. Special thanks to my second family for their

unconditional support and amazing generosity.

I would like to dedicate my thesis to my father and husband.



ABSTRACT

EXPLORING FUNCTIONAL MOTIONS OF MUSCARINIC ACETYLCHOLINE M2 RECEPTOR

G-Protein Coupled Receptors (GPCRs) are the largest family of signaling proteins and a better understanding of dynamics underlying functional mechanism is essential in drug design. Although many studies have been performed to date, characterizing conformational changes and dynamics in-between inactive and active states remains elusive. Here, the activation mechanism of Muscarinic acetylcholine receptor M2, which belongs to the GPCRs family and responsible from decreasing heart rate to normal rhythm by inhibiting cAMP (cyclic adenosine mono phosphate), has been explored by ANM-LD computational methodology that combines anisotropic network model (ANM) with all-atom Langevin dynamics (LD) simulations. The predicted physically plausible multiple conformational transition pathways from the inactive (3UON) towards the active states (4MQS, 4MQT) disclosed the dynamic determinants underlying the M2's activation process. It was observed that certain collective ANM modes are essential for the activation and the hinge sites that coordinate the motion defined by these modes of motion aligns with the known functional important sites acting as molecular switches such as (DRY and NPxxY motifs, TM3-TM6 distance, salt bridge between R121 and E382). Furthermore, the two hydrophobic layers in TM domains display the coupling of the conformational changes with these switches and the continuous water pathway from extracellular side through intracellular side providing a continuous path from the ligand binding site to the G-protein binding site in the activation.

ÖZET

MUSKARİNİK ASETİLKOLİN M2 RESEPTÖRÜNÜN FONKSİYONEL HAREKETLERİNİN İNCELENMESİ

G-protein bağılı reseptörler (GPRs) en büyük sinyal protein ailesidir ve bu yüzden işleyiş mekanizmalarının altında yatan dinamik etkenlerin anlaşılması ilaç tasarımında büyük bir öneme sahiptir. Bu güne kadar yapılan pek çok çalışma olmasına rağmen, inaktif ve aktif yapılar arasında gerçekleşen kolektif konformasyonel değişimlerin özellikleri ve dinamikleri hala belirsizdir. Bu çalışmada, G-protein bağılı reseptörler ailesine ait olan Muskarinik asetilkolin M2 reseptörünün aktivasyon mekanizması ANM-LD metodu kullanılarak aydınlatılmaya çalışılmıştır. ANM-LD metodu, Anizotropik ağyapı modelinin (ANM) yavaş modlarının Langevin Dinamiği (LD) simülasyonları ile yönlendirilmesiyle elde edilen hibrit bir metoddur. M2 reseptörünün aktivasyon mekanizmasının anlaşılması için inaktif yapı (3UON), başlangıç yapısı olarak kullanılarak aktif yapıya (4MQS, 4MQT) ulaşılması hedeflenmiştir ve iki yapı arasında oluşturulan geçiş patikaları proteinin iç dinamiği kullanılarak incelenmiştir. Çalışmada belirli kolektif modların aktivasyon mekanizmasında etkin rol aldıkları ve bu modların hareketi hedef yapı doğrultusunda yönlendirdiği; aynı zamanda bu modlara ait menteşe rezidülerinin bilinen fonksiyonel rezidüler (DRY ve NPxxY motifleri, R121 ve E382 arasındaki tuz köprüsü, R121 ve T386 rezidüleri arasındaki mesafe (TM3-TM6 uzaklığı)) ile örtüştüğü sonucuna varılmıştır. Ayrıca transmembran domainlerin arasında yer alan hidrofobik bölgelerin, fonksiyonel öneme sahip rezidülerin yaptığı konformasyonel geçişlerle örtüştüğü ve su girişine izin verdiği; böylece aktivasyon için önemli olan sürekli su patikasının oluşabildiği görülmüştür.

TABLE OF CONTENTS

| | |
|---|------|
| ACKNOWLEDGEMENTS | iii |
| ABSTRACT | v |
| ÖZET | vi |
| LIST OF FIGURES | viii |
| LIST OF TABLES | xiv |
| LIST OF SYMBOLS | xv |
| LIST OF ACRONYMS/ABBREVIATIONS | xvii |
| 1. INTRODUCTION | 1 |
| 2. G PROTEIN COUPLED RECEPTORS (GPCRs) | 4 |
| 3. MATERIALS AND METHODS | 7 |
| 3.1. Normal Mode Analysis (NMA) And Elastic Network Models (ENMs) | 7 |
| 3.1.1. Gaussian Network Model (GNM) | 7 |
| 3.1.2. Anisotropic Network Model (ANM) | 8 |
| 3.2. Molecular Dynamic Simulations And Langevin Dynamics | 9 |
| 3.3. ANM-LD Methodology | 12 |
| 4. RESULTS AND DISCUSSION | 15 |
| 4.1. Activation of M2 Receptor | 15 |
| 4.1.1. ANM-LD Simulation Results from Antagonist Bound Structure to Agonist Bound Structure | 19 |
| 4.1.2. ANM-LD Simulation Results from Antagonist Bound Structure to Agonist and Allosteric Bound Structure | 54 |
| 5. CONCLUSION AND RECOMMENDATIONS | 62 |
| REFERENCES | 64 |
| APPENDIX A: ANM-LD SIMULATION RESULT SUMMARIES | 71 |

LIST OF FIGURES

| | | |
|-------------|--|----|
| Figure 2.1. | Cartoon representation of rhodopsin of bovine. | 4 |
| Figure 2.2. | Regulation process of M2 receptor coupled with Gi. | 5 |
| Figure 3.1. | The procedure of methodology of ANM-LD Simulations. | 14 |
| Figure 4.1. | The structure of M2 Receptor and naming terminology. | 16 |
| Figure 4.2. | M2 Receptor domain naming and corresponding residues. | 16 |
| Figure 4.3. | Active states of M2 Receptor. | 18 |
| Figure 4.4. | Active (4MQS) and inactive (3UON) crystal structures of M2 receptors. (a) Red arrows show displacement of TM6 through inward and outward on extracellular side and intracellular sides respectively. (b) Intracellular side view of structures. | 18 |
| Figure 4.5. | Alignment of generated final transition conformation of simulation 7 with the target structure, final RMSD value is 1.54Å. | 21 |
| Figure 4.6. | RMSD difference between generated transition conformations and target structure in simulations of M2 receptor from 3UON to 4MQS. | 22 |
| Figure 4.7. | Selected mode, RMSD, overlap and collectivity values plots of simulations from 3UON to 4MQS. RMSD and selected modes plots are given in right side. Selected modes and corresponding collectivity and overlap values plots are given in left side. | 23 |

| | | |
|--------------|---|----|
| Figure 4.8. | Functionally important distances through transition pathway (Left side) and RMSD of selected important residues between transition conformations and target state (Right side) in simulations of M2 receptor from 3UON to 4MQS. | 25 |
| Figure 4.9. | (a) Selected modes distributions. (b) Overlap values of selected modes of M2 simulations from 3UON to 4MQS according to the initial structure's first 30 modes. | 28 |
| Figure 4.10. | The most selected ANM mode shapes of initial state and corresponding pathway selected modes in M2 simulations from 3UON to 4MQS. | 29 |
| Figure 4.11. | Selected modes similarities with each other in M2 simulations from 3UON to 4MQS. | 30 |
| Figure 4.12. | (a) 5 th ANM Mode shape of the initial state. (b) Demonstration of minimum points of 5 th ANM mode on structure. | 33 |
| Figure 4.13. | Superimposition of 2 nd , 4 th and 6 th ANM modes with highlighted minimum regions. | 34 |
| Figure 4.14. | 2 nd ANM mode motion representation from ANM server and its ANM mode shape. | 34 |
| Figure 4.15. | 4 th ANM mode motion representation from ANM server and its ANM mode shape. | 35 |
| Figure 4.16. | 5 th ANM mode motion representation from ANM server and its ANM mode shape. | 35 |

| | |
|--|----|
| Figure 4.17. 6 th ANM mode motion representation from ANM server and its ANM mode shape. | 36 |
| Figure 4.18. Hinge residues of most selected ANM modes in Simulation 7. . . . | 36 |
| Figure 4.19. Hinge residues of most selected ANM modes in Simulation 24. . . | 37 |
| Figure 4.20. Hinge residues of most selected ANM modes in Simulation 30. . . | 37 |
| Figure 4.21. Hinge residues of most selected ANM modes in Simulation 33. . . | 38 |
| Figure 4.22. The first-half (Left) and full-pathway (Right) cross-correlation maps for ANM-LD simulations of M2 receptor from 3UON to 4MQS. . . | 40 |
| Figure 4.23. RMSD difference between generated transition and target conformations by using modified structures in simulations of M2 receptor from 3UON to 4MQS. | 42 |
| Figure 4.24. Mostly selected ANM mode shapes superimposition of modified and unmodified structures simulations. (a) 2 nd modes of three initial structures. (b) 6 th mode of unmodified structure, 7 th mode of only terminals modified structure, 6th mode of modeled and terminals modified structure. (c) 2 nd mode of unmodified structure, 1 st mode of only terminals modified structure. | 43 |
| Figure 4.25. RMSD difference between generated transition conformations and target conformation in simulations of M2 receptor from 3UON to 4MQS with 5 th mode restriction. | 45 |

| | |
|---|----|
| Figure 4.26. (a) Selected modes distributions (b) Overlap values of selected modes of M2 simulations from 3UON to 4MQS with 5 th ANM mode restriction according to the initial structure's first 30 modes. . . . | 45 |
| Figure 4.27. Functionally important distances through transition pathway (Left side) and RMSD of selected residues between transition conformations and target conformation (Right side) in M2 simulations from 3UON to 4MQS with 5 th mode restriction. | 47 |
| Figure 4.28. The first-half (Left) and full-pathway (Right) cross-correlation maps for ANM-LD simulations of M2 receptor from 3UON to 4MQS with 5 th mode restriction. | 48 |
| Figure 4.29. Selected mode, RMSD, overlap and collectivity values plots of simulations from 3UON to 4MQS with 4 th mode restriction. RMSD and selected modes plot is given in right side. Selected modes and corresponding collectivity and overlap values plot is given in left side. | 49 |
| Figure 4.30. (a) Selected modes distribution. (b) Overlap values of selected modes of M2 simulation from 3UON to 4MQS with 4 th ANM mode restriction according to the initial structure's first 30 modes. . . . | 50 |
| Figure 4.31. (a) Functionally important distances through transition pathway. (b) RMSD of selected residues between transition conformations and target conformation in M2 simulations from 3UON to 4MQS with 4 th mode restriction. | 50 |
| Figure 4.32. The first-half (Left) and full-pathway (Right) cross-correlation maps for ANM-LD simulations of M2 receptor from 3UON to 4MQS with 4 th mode restriction. | 51 |

| | |
|---|----|
| Figure 4.33. Calculated available channels in generated conformations, initial and final states by using Hollow. | 52 |
| Figure 4.34. Superimposition of calculated available channels of 3UON, Cycle 9 and 4MQS to demonstrate the formed continuous water pathway. DRY and NPxxY motifs are colored with magenta and cyan respectively. | 53 |
| Figure 4.35. Superimposition of calculated available channels of 3UON, Cycle 16 and 4MQS to demonstrate the formed continuous water pathway. DRY and NPxxY motifs are colored with magenta and cyan respectively. | 53 |
| Figure 4.36. Conformation change of residue W422 in active structures of M2 receptor. | 54 |
| Figure 4.37. Selected mode, RMSD, overlap and collectivity values plots of simulations from 3UON to 4MQS. RMSD and selected modes plots are given in right side. Selected modes and corresponding collectivity and overlap values plots are given in left side. | 56 |
| Figure 4.38. (a) Selected modes distributions. (b) Overlap values of selected modes of M2 simulations from 3UON to 4MQT according to the initial structure's first 30 modes. | 57 |
| Figure 4.39. (a) Mostly selected ANM modes in simulations from 3UON to 4MQT. (b) Minimum residues correspond to mostly selected ANM modes 2 (blue) and 3 (green). | 58 |

Figure 4.40. (a) Difference vectors between initial and final states. (b) Superimposition of 2nd ANM Mode of 4MQT to 4th ANM Mode of 4MQS. (c) Superimposition of 3rd ANM Mode of 4MQT to 5th ANM Mode of 4MQS. 60

Figure 4.41. Functionally important distances through transition pathway (Left side) and RMSD of selected residues between transition conformations and target conformation (Right side) in M2 simulations from 3UON to 4MQS. 61

LIST OF TABLES

| | | |
|------------|--|----|
| Table 4.1. | ANM-LD Simulation result summaries from 3UON to 4MQS. . . . | 21 |
| Table 4.2. | ANM-LD Simulation result summaries of modified structures. . . . | 41 |
| Table 4.3. | ANM-LD simulation result summaries from 3UON to 4MQS with 5 th mode restriction. | 44 |
| Table 4.4. | ANM-LD Simulation result summaries from 3UON to 4MQT. . . . | 55 |
| Table A.1. | ANM-LD Simulation result summaries of M2 Receptor activation pathway (Simulation No:1-30). | 71 |
| Table A.2. | ANM-LD Simulation result summaries of M2 Receptor activation pathway (Simulation No:31-70). | 72 |

LIST OF SYMBOLS

| | |
|----------------|--|
| a_i | Acceleration of particle i |
| \mathbf{G}_i | Adenylyl Cyclase inhibitor |
| C_{ij} | Cross-correlation value between the fluctuations of residues i and j |
| D_F | Deformation factor |
| d_i | Distance between the coordinates of atom i in two structures |
| \mathbf{D}_i | Distance difference vector |
| f | Force |
| F_i | Force exerted on particle i |
| \mathbf{H} | Hessian matrix |
| \mathbf{I}_i | Initial conformation coordinates in cycle i |
| k | Mode |
| k_B | Boltzmann constant |
| m_i | Mass of particle i |
| $modemax$ | Maximum number of modes to choose among |
| N | Number of atoms |
| $O_{i,j}$ | Overlap value of j th eigenvector, \mathbf{u}_j , with the difference vector, \mathbf{D}_i , for cycle i |
| $O_{i,max}$ | Maximum overlap value for cycle i |
| R_{cut} | Cut-off radius |
| \mathbf{R}_i | Position vector of residue i |
| R_{ij} | Instantaneous distance between residues i and j |
| R_{ij}^0 | Equilibrium distance between residues i and j |
| $stepmin$ | Number of minimization steps |
| $stepsim$ | Number of simulation steps |
| t | Time |
| T | Absolute temperature |
| \mathbf{T} | Target conformation coordinates |
| \mathbf{u}_k | k th eigenvector |

| | |
|----------------------|---|
| \mathbf{U} | Matrix of eigenvectors |
| V | Potential function |
| V_{ANM} | ANM potential function |
| $\alpha\mathbf{C}$ | Alpha carbon coordinates |
| \AA | Angstrom |
| $\Delta\mathbf{R}_i$ | Fluctuation of residue i |
| γ_f | Force constant for harmonic spring |
| γ | Damping constant of Langevin dynamics simulations |
| ζ | Friction (damping) coefficient |
| κ | The degree of collectivity |
| κ_m | The degree of collectivity of the mostly overlapping ANM mode m |
| λ_k | k th eigenvalue |
| ν_i | Velocity of atom i |
| $\mathbf{\Gamma}$ | Kirchhoff (or connectivity) matrix |
| Λ | Diagonal matrix of eigenvalues |

LIST OF ACRONYMS/ABBREVIATIONS

| | |
|--------|--|
| 1D | One dimensional |
| 3D | Three dimensional |
| AC | Adenylyl Cyclase |
| Amber | Assisted Model Building with Energy Refinement |
| ANM | Anisotropic Network Model |
| ANM-LD | Anisotropic Network Model Guided Langevin Dynamics |
| cAMP | Cyclic adenosine monophosphate |
| eV | Eigenvalue |
| ENM | Elastic Network Model |
| ECL | Extracellular Loop |
| GNM | Gaussian Network Model |
| GPCRs | Guanine Binding Protein Coupled Receptors |
| ICL | Intracellular Loop |
| LD | Langevin Dynamics |
| MD | Molecular Dynamics |
| MSF | Mean Square Fluctuations |
| NM | Normal Mode |
| PCA | Principal Component Analysis |
| PDB | Protein Data Bank |
| RMSD | Root-Mean-Square Deviation |
| Sander | Simulated Annealing with NMR-Derived Energy Restraints |
| ter | Terminal |
| TM | Transmembrane |

1. INTRODUCTION

Proteins play a crucial role in many biological processes like molecular recognition, enzymatic activity, and allosteric regulation. They may exist in many different conformations in order to conduct their functions in cells. Thus, it is important to discover the conformational changes of proteins in order to understand the molecular mechanism underlying their functions. Conformations at equilibrium near the minimum energy are stable conformations and they are long-lived. Conformations between these stable states are unstable and they are short-lived [1]. Although stable conformations can be discovered by experimental techniques, transition pathway unstable conformations are difficult to discover by experimental techniques due to their short-lived. However, also transition pathways are important to interfere protein dynamics and remedy a disease. Computational methods are so important to reveal these transition pathways and their processes in order to understand the functioning mechanism of cells.

There are many different methods to investigate the transition pathways and sample plausible conformations. Molecular Dynamic (MD) Simulation is one of the most realistic computational methods. They can provide very detailed simulation of every particle motions as a function of time. Desired specific contributions can be determined either removing or adding that specific contribution to the simulation, where all parameters can be controlled by the user [2]. Although its advantages, sampling can be limited due to computational cost of applications. Rough energy landscapes with many local minimums separated by high-energy barriers, huge systems that contain thousands of residues/atoms, existence of solvents and etc. are the parameters that can affect the sampling time [3]. Due to their timescales and/or large size, they are not feasible and many hybrid methods have been developed to reach enhanced sampling of the thermally accessible conformational space including computational and experimental techniques.

The realization of proteins undergo non-random changes had triggered coarse-grained methods and with the realization of experimentally observed functional motions of proteins can be predicted by coarse-grained normal mode analysis, coarse-grained methods began to used widely in structural biology [4, 5]. These non-random changes can usually be determined by using simple models like Elastic Network Models (ENMs). ENMs can provide an exact solution for the unique dynamics of each structure and these models can be applied to large biomolecular complexes and assemblies [6]. ENMs considers cooperative nature of biomolecular dynamics and it was observed that only a few principal modes of motion, also called soft modes, conduct intermolecular interactions and these modes can be used to discover unknown conformations [5]. These modes are insensitive to structural and energetic details.

Gaussian Network Model (GNM) and Anisotropic Network Model (ANM) are the most well-known simplified Elastic Network models. In the network α -carbons are noted nodes and it is assumed that these nodes are connected with elastic springs if α -carbon pairs are closer to each other than a specified cutoff distance. In GNM, these nodes undergo isotropic (Gaussian) fluctuations about their equilibrium positions. However in ANM, the residue fluctuations are considered as anisotropic that incorporate among components of the position vector independently. The main advantage of ANM is that it can generate alternative conformations in the close neighborhood of a given structure upon deforming the original structures along the dominant (lowest frequency) modes [5, 7–10].

Combining Molecular Dynamic simulation protocols and computational methods, many hybrid methods have been developed to investigate protein dynamics. ANM-LD is a simple and efficient hybrid method that implements the predicted intrinsic functional dynamics by ANM into an all-atom LD simulation protocol. The methodology disposes large conformational transitions and multiple transition pathways between given protein functional states. The cooperative changes underlying the functional transitions of proteins can be observed the generation of the alternative pathways by the restriction of certain modes can be controlled and differences in the pathways further by GNM can be assessed. The methodology requires two structures that are initial

and target structures. The procedure is iterative and its convergence is measured with RMSD value between target and discovered structure [11].

G protein-coupled receptors (GPCRs) are the largest family of mammalian membrane proteins and also the biggest target, nearly one-third of all drugs, in the drug market. Their versatility and ability to change their structural shape through transition among different conformations has increased their importance [12]. It has not been completely understood how GPCRs change their shapes through activation over time [13]. Motivation of this thesis is to explore the activation mechanism of muscarinic receptor M2 that belongs to GPCRs, in the light of generated transition pathways by Anisotropic Network Model Guided Langevin Dynamics (ANM-LD) methodology performing specificities in simulations. Dynamic behavior of structure was tried to reveal by analyzing collective ANM modes and generated conformations in terms of functionally known important distances.

2. G PROTEIN COUPLED RECEPTORS (GPCRs)

G (Guanine binding) protein coupled receptors (GPCRs) are the super-family of sensory proteins of eukaryotes. They are physiologically important membrane proteins that responsible from transmitting signals to from exterior to the interior of the cells and physiologic responses to neurotransmitters and hormones. GPCRs share three common regions which are an intracellular region (the C-terminus), a middle segment containing seven transmembrane domains and an extracellular region (the N-terminus) given in Figure 2.1.

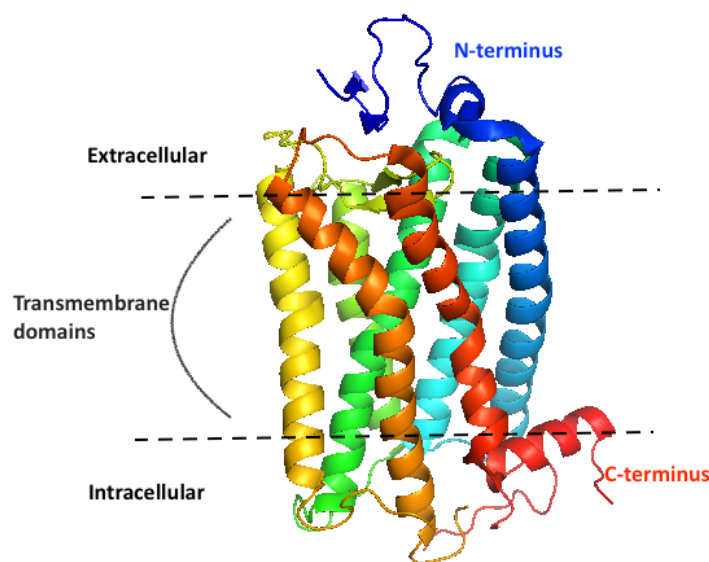


Figure 2.1. Cartoon representation of rhodopsin of bovine.

GPCRs are huge super-family and include five families that are Glutamate, Rhodopsin, Adhesion, Frizzled/Taste2, and Secretin, according to the GRAFS classification system [14]. There are more than 800 different human seven-transmembrane proteins and they are involved in many diseases which are related with signaling pathways including cancer, obesity, immunological, cardiovascular, inflammatory, viral infections,

senses disorders, cardiovascular and etc. Their versatility increases their importance in drug market and it is estimated that they are targets of drug market about 40%, so they have been the most popular target class for drug discovery [15–17].

Rhodopsin-like family constitutes the largest group of GPCRs, almost 85% of G-proteins, including light receptors, hormones and neurotransmitters. Muscarinic receptors belong to Rhodopsin-like family (Class A) that include five subtypes from M1 to M5 responsible from regulating many vital functions of the central and peripheral nervous systems and regulating a variety of physiological functions like intestinal airway, eye smooth muscle contraction, glandular secretions and heart rate [18,19].

M2 receptor, the subtype of the muscarinic acetylcholine receptors, is widely expressed in both central and peripheral nervous systems regulating the parasympathetic control of the heart, presynaptic inhibition of neurotransmitter release [20]. M2 receptor couple to $G_{i/o}$ and leads to inhibitory-type effect like inhibiting adenylyl cyclase (AC) activity the cell [21–23]. Inhibition mechanism of adenylyl cyclase (AC) activity is given in Figure 2.2.

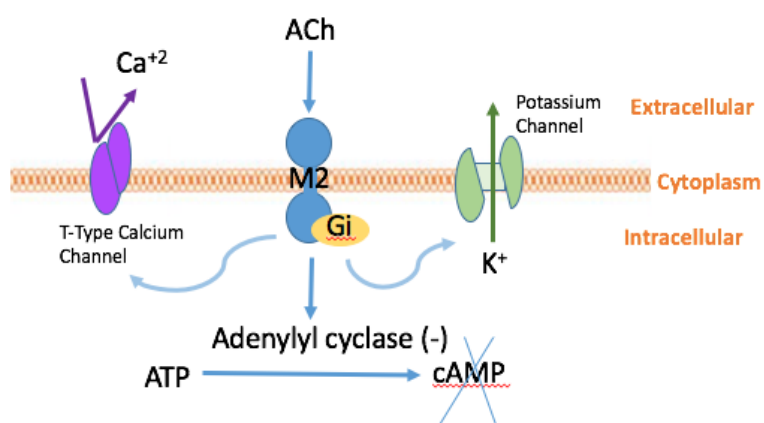


Figure 2.2. Regulation process of M2 receptor coupled with Gi.

Although many studies have been focused on muscarinic acetylcholine receptors due to their importance in both central and peripheral nervous systems, the activation mechanism of these receptors have not been completely understood yet as it is in other GPCRs [24,25]. However, it is known that conformational flexibility of GPCRs provide diverse ability to engage in many signaling pathways, although their structural similarity and they can be represented by an ensemble of conformations [26,27].



3. MATERIALS AND METHODS

3.1. Normal Mode Analysis (NMA) And Elastic Network Models (ENMs)

Native states of proteins are not rigid and proteins can undergo random fluctuations at their dynamic equilibrium (global minimum) and may sample ensemble of conformations with these fluctuations around their dynamic equilibrium. These fluctuations are related with the collective/cooperative/global motions that disclose the functions of proteins. Hence, NMA can predict the functionally important motions by solving the topology of modes of motions at equilibrium. These collective/global motions are generally related with the low frequency modes that are slow modes (soft modes). Elastic Network Models (ENMs) are simplified coarse-grained normal mode analysis, where amino acids are shown as nodes and linked with the neighbors by springs. In ENM, one dimensional uniform harmonic motions between interacting atoms is accepted rather than detailed atomic potential and this provides computational efficiency and unique exact solution to describe protein conformational changes and explore intrinsic dynamic behavior that is necessary for proteins to conduct their functions [4, 6, 28, 29].

Gaussian Network Model (GNM) and Anisotropic Network Model (ANM) [10, 30] are most well-known EN models [4].

3.1.1. Gaussian Network Model (GNM)

Gaussian Network Model (GNM) is one-dimensional, simple EN model. In the network α -carbons are noted as nodes and it is assumed that these nodes are connected with elastic springs if α -carbon pairs are closer to each other than a specified cutoff (\mathbf{R}_{cut}) distance and these nodes undergo isotropic (Gaussian) fluctuations about their equilibrium positions [8, 31].

The interaction network of residues is defined by Kirchhoff (connectivity) matrix ($\mathbf{\Gamma}$). Kirchhoff matrix is symmetric and sparse, so inverse of Kirchhoff matrix is calculated by using linear superimposition of $N - 1$ nonzero eigenvectors as given in Equation 3.1, where \mathbf{N} is total atom number, \mathbf{u}_k is the k th eigenvector of eigenvector matrix \mathbf{U} . k th eigenvector direct the motion along the k th mode.

$$\left[\mathbf{\Gamma}^{-1}\right]_{ij} = \left[\mathbf{U}\mathbf{\Lambda}^{-1}\mathbf{U}^T\right]_{ij} = \left[\lambda_k^{-1}\mathbf{u}_k\mathbf{u}_k^T\right]_{ij} \quad (3.1)$$

The Correlation between fluctuations of residue pairs is calculated by using Equation 3.2, where γ_f is the force constant of the potential function, T is the absolute temperature in Kelvin, k_B is the Boltzmann constant, $\Delta\mathbf{R}_i$ and $\Delta\mathbf{R}_j$ are the fluctuations of residue i and j .

$\Delta\mathbf{R}_i$ and $\Delta\mathbf{R}_j$ is given in Equation 3.1,

$$\langle\Delta\mathbf{R}_i.\Delta\mathbf{R}_j\rangle = \frac{3k_B T}{\gamma_f} \left[\mathbf{\Gamma}^{-1}\right]_{ij} \quad (3.2)$$

3.1.2. Anisotropic Network Model (ANM)

Anisotropic Network Model (ANM) is three dimensional Elastic Network Model, in which instead of isotropic fluctuations, directional preferences are taken into account considering anisotropic fluctuations of proteins in real life [10,30,32]. These anisotropic fluctuations can incorporate among components of the position vector independently and that means $3N-6$ modes instead of $N-1$ in GNM. The main advantage of ANM is that by using fluctuation vectors it can be generated alternative conformations in the close neighborhood of a given structure upon deforming the original structures along the collective (lowest frequency) modes [4, 5]. Overall potential of the system is calculated by using Equation 3.3, where R_{ij}^0 instantaneous distance and R_{ij} are the

equilibrium distance between residues i and j and γ_f is the force constant.

$$V_{ANM} = \frac{1}{2} \sum_{ij} \gamma_{f,ij} (R_{ij} - R_{ij}^0) \quad (3.3)$$

The correlation between fluctuations of residue pair i and j from their equilibrium position, $\langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j \rangle$, is calculated decomposition and reconstruction of $3N - 6$ modes of Hessian Matrix (\mathbf{H}) as distorted by the displacement of mode k is calculated as given in Equation 3.4 in terms of eigenvalue (λ_k) of Hessian matrix.

$$\langle \Delta \mathbf{R}_i \cdot \Delta \mathbf{R}_j \rangle = \frac{3k_B T}{\gamma_f} \text{tr} [\mathbf{H}^{-1}] = \frac{3k_B T}{\gamma_f} \sum_{k=1}^N \text{tr} [\lambda_k^{-1} \mathbf{u}_k \mathbf{u}_k^T]_{ij} \quad (3.4)$$

3.2. Molecular Dynamic Simulations And Langevin Dynamics

Molecular dynamic (MD) simulation is one of the most realistic computational methods. It depends on the solution of Newton's equation of motion as function of time iteratively. Energy function and position vectors are the only required items of MD Simulations. Initial positions of proteins can be obtained from NMR spectroscopy or X-Ray Crystallography. Then dividing time into discrete time steps, acting force on each atom is computed using a molecular force field. This causes change in atom positions in a given time and new positions can be calculated by solving Newton's equation of motion. This iterative procedure proceeds until the desired state is obtained if it is computationally possible [33, 34].

Newton's equation of motion is given in Equation 3.5, where (\mathbf{F}_i) is the force on atom i , (\mathbf{m}_i) and (\mathbf{a}_i) are the mass and acceleration of atom i respectively.

$$F_i = M_i a_i \quad (3.5)$$

Acceleration can be calculated by using second derivative of position vector (Equation 3.6) and force acting on a particle can be written in terms of gradient of potential energy (Equation 3.7).

$$\frac{d^2 r_i}{dt^2} = \frac{F_i}{M_i} \quad (3.6)$$

$$F_i = -gradV_i = -\nabla V_i \quad (3.7)$$

Newton's equation of motion can be rearranged as a function of time in terms of gradient of potential energy by replacing force acting on a particle with the derivative of potential energy given in Equation 3.8.

$$\frac{d^2 r_i}{dt^2} = -\frac{1}{M_i} \frac{\partial V(r_1, r_2, \dots, r_N)}{\partial r_i} = a_i \quad (3.8)$$

By using Maxwell-Boltzmann distribution (Eq.3.9) to calculate initial velocities and gradient of potential energy, acceleration distribution of initial atoms can be computed by using Equation 3.10. T , is the absolute temperature in Kelvin, k_B is the Boltzmann constant, v_{ix} is the velocity of atom i in x direction and m_i is the mass of atom i .

$$p(v_{ix}) = \left(\frac{M_i}{2\pi k_B T} \right)^{1/2} \exp \left[-\frac{1}{2} \frac{M_i v_{ix}^2}{k_B T} \right] \quad (3.9)$$

$$p = \sum_{i=1}^N M_i v_i = 0 \quad (3.10)$$

The absolute temperature and kinetic energy of the system are related and it can be calculated as given in Equation 3.11, where $\langle v_i^2 \rangle$ is the square of average velocity of atom i and N is the number of atoms.

$$T = \frac{2}{3Nk_B} \sum_{i=1}^N \frac{1}{2} M_i \langle v_i^2 \rangle \quad (3.11)$$

The Langevin equation is a stochastic differential equation obtained by adding two new terms to the Newton's second law to approximate the effects of neglected degrees of freedom (Equation 3.12). $\nabla_i V$ is the friction coefficient that is taken into account for viscous (dissipation) terms that represents the drag force due to friction of solvent and $R_i(t)$ is the random force acting on the system due collision between solvent particles and the molecule particles with zero mean. The random force is generally accepted as uncorrelated in time. When friction coefficient is taken as zero, the Langevin equation equals to classical MD equation. Langevin Dynamics may improve the simulations considering outside effects on the molecule in terms of random forces acting of particles [35, 36].

$$M_i \frac{d^2 r_i}{dt^2} = -\frac{\partial V}{\partial r_i} - M_i \gamma \frac{dr_i}{dt} + R_i(t) \quad (3.12)$$

The motion can be written in terms of gradient of potential (∇_V) and friction factor (ζ_i) and assuming random force on atom i (R_i) is independent from time; where M_i is mass of atom i , r is the position vector.

$$M_i \ddot{r}_i = -\nabla_i V - \zeta_i \dot{r}_i + R_i \quad (3.13)$$

3.3. ANM-LD Methodology

ANM Driven Langevin Dynamics (ANM-LD) simulation is a simple and efficient hybrid methodology that implements the ANM predicted intrinsic functional dynamics with LD simulations to generate alternative conformations between given two functional states [11]. Cooperative/global motions underlying the functional transitions can be observed by restricting certain modes and/or changing ANM and/or simulation parameters.

The methodology requires two functional states, which are used as initial ($\mathbf{I} = \mathbf{I}_0$) and target ($\mathbf{T} = \mathbf{T}_0$) structures. The procedure is iterative and the structure is distorted with the selected ANM mode in every iteration step i . Convergence to the target structure is measured with RMSD value between target and generated structure. RMSD calculation is given in Equation 3.14, where \mathbf{d}_i is the distance between atoms i of generated conformation and target structure and \mathbf{N} is the number of atoms in the structure. While selecting ANM mode, overlap value ($O_{i,j}$) between global mode \mathbf{u}_k and difference vector (\mathbf{D}_i) composed by target and instantaneous conformation is used. Overlap value calculation is given in Equation 3.15.

$$RMSD = \sqrt{\frac{\sum_{i=1}^N d_i^2}{N}} \quad (3.14)$$

$$O_{i,j} = \frac{|\mathbf{u}_j \cdot \mathbf{D}_i|}{|\mathbf{u}_j| \cdot |\mathbf{D}_i|} \quad (3.15)$$

$$\mathbf{D}_i = \mathbf{T}_0 - \mathbf{I}_i \quad (3.16)$$

Overlap value is calculated for all collective modes and then the mode with maximum overlap value is chosen as selected ANM mode. Utilizing selected ANM mode, the structure is distorted in that motion at given deformation factor (D_F) rate. The generated conformation will be the initial structure of next iteration step. In every iteration, the initial structure is energetically minimized and simulated with all-atom Langevin Dynamic simulation methodology. The procedure continuous as RMSD between the generated conformation and target state converges. The procedure of ANM-LD methodology is given in as flowchart in Figure 3.1. Simulation parameters are predefined before starting simulations, which are cut-off distance (R_{cut}), maximum number of modes ($modemax$), the number of modes available to choose a mode among them, deformation factor (D_F) to distort the structure along the chosen mode direction, energy minimization step ($stepmin$), absolute temperature (T), damping constant (γ) and LD simulation step ($stepsim$). First three parameters R_{cut} , $modemax$ and D_F belong to ANM part of simulation. LD simulations are conducted by using the Sander program of Amber 11 and the force field parameter is set by Amber ff03 and ff10 [37,38]. Protonation states of charged residues in Initial and final structures are calculated by PDB2PQR [39].

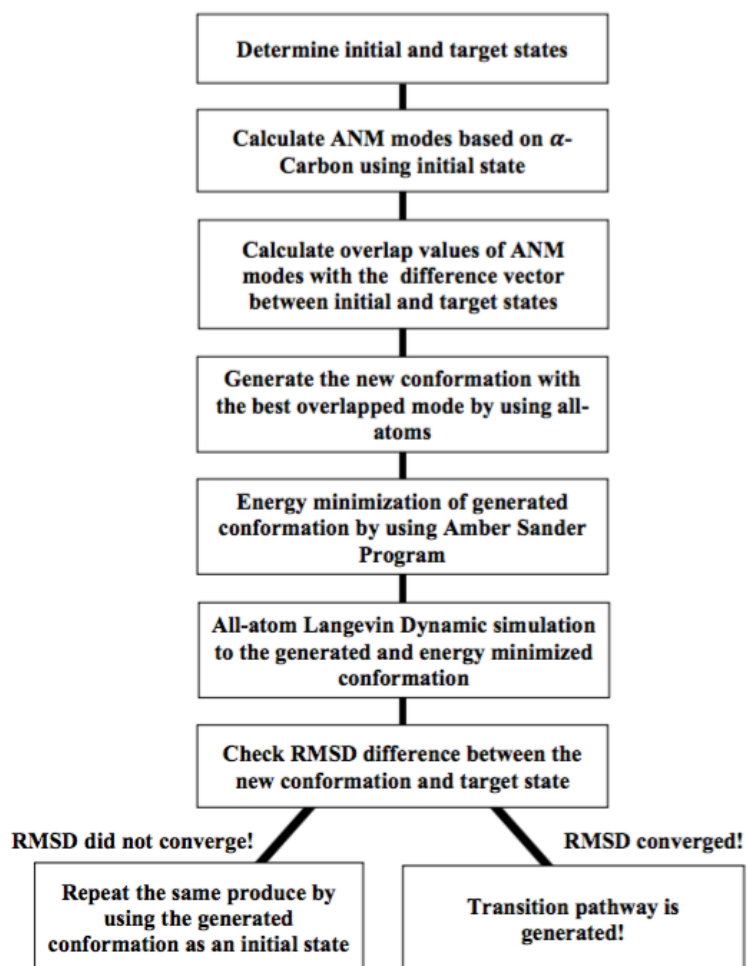


Figure 3.1. The procedure of methodology of ANM-LD Simulations.

4. RESULTS AND DISCUSSION

4.1. Activation of M2 Receptor

Muscarinic acetylcholine M2 receptor is responsible from physiologic control of cardiovascular function, that decreases heart rate, through activation of G protein coupled inwardly-rectifying potassium-ion membrane channels [40–42].

M2 receptor has been target for both allosteric and orthosteric ligands research studies due to its pharmacological specialties [41]. In this thesis, ANM-LD (Anisotropic Network Model is guided with all-atom Langevin Dynamic (LD) simulations) simulations have been explored in order to explore the activation mechanism of M2 receptor through generating conformational transition pathways between inactive and active states, and analyze the underlying dynamic behavior.

M2 receptor consists of a bundle of seven alpha helices, 3 intracellular loops (ICL) and 3 extracellular loops (ECL) as it is in other GPCRs and contains 276 residues. Alpha helices are located in transmembrane region and they are named TM domain numbers, loops located at the cytoplasmic side are named as intracellular loops (ICL) and loops located at the extracellular side are named as extracellular loops (ECL). The structure residue numbering starts 20 and ends up with 456, however residues located in ICL3 loop between 216 and 377 are missing. The structure of M2 receptor including their naming terminology that is same in all GPCR family is given in Figure 4.1 and residue numbers and their corresponding regions are summarized in Figure 4.2.

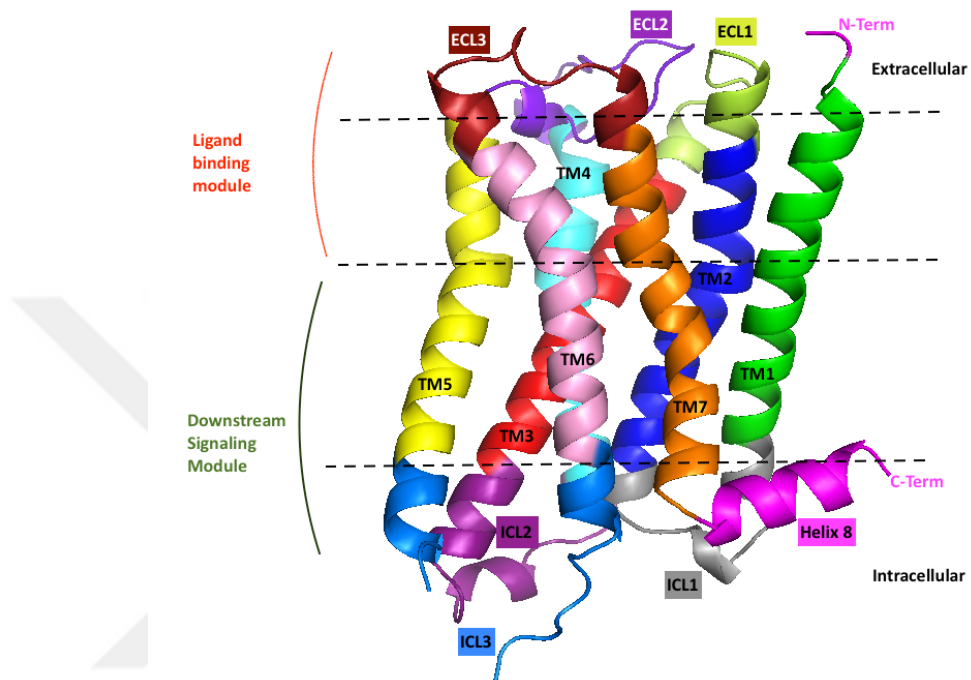


Figure 4.1. The structure of M2 Receptor and naming terminology.

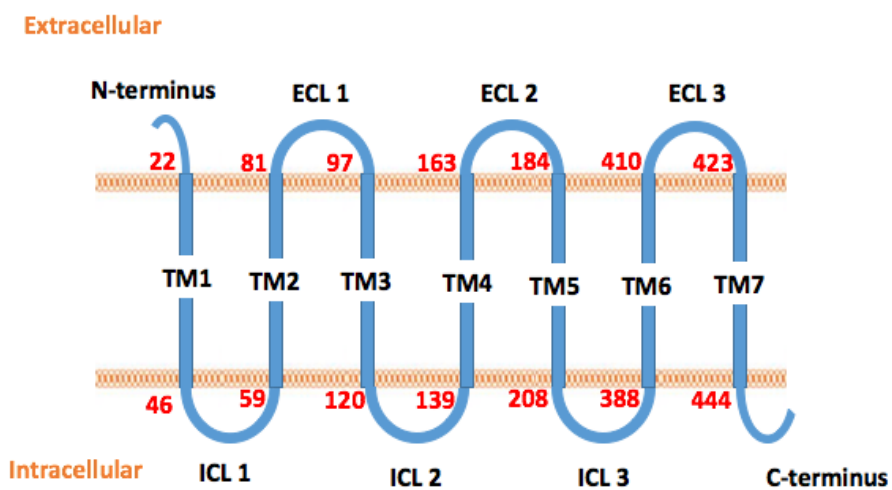


Figure 4.2. M2 Receptor domain naming and corresponding residues.

The inactive structure that is bound to an antagonist and two active structures were obtained from Protein Data Bank (PDB) with PDB IDs 3UON, 4MQS and 4MQT, respectively [40, 41]. Although 4MQT is bound to the agonist iperoxo and allosteric modulator, 4MQS is only bound to the agonist iperoxo. The structure of 4MQT is quite similar to the structure of 4MQS except 3 missing residues in ICL3 loop and slight changes around the allosteric binding site, which means the allosteric site could be preformed when agonist is bound the structure without any modulator [40]. Although, the RMSD between two active states is 0.99\AA , it has been explored here that whether small variations at certain regions may reflect in the transition pathway given the resolution of the present method. The aligned structures are given in Figure 4.3 in order to demonstrate the slight difference between these two active states.

During activation of M2 receptor, key outward displacement of the intracellular side of TM6 creates a cage for G protein to bind the receptor with slight outward displacement of TM5 and rearrangements around NPxxY (N436, P437, Y440) and DRY (D120, R121, Y122) motifs in TM7 and TM3, respectively. TM6 outward displacement, NPxxY and DRY motifs are highly conserved common features of G-protein-coupled receptors' activation mechanism [40, 43]. Displacement of transmembrane helices are shown in Figure 4.4 In addition to the displacements of transmembrane helices, water penetration is also required for activation of GPCRs. Water penetrated into the core of TM domains and the binding cage during transition. The region near the residue V105 below the orthosteric ligand binding cavity is hydrated and that may also have a functional role in the allosteric binding. The other hydrated regions are NPxxY motif region and the middle of TM2 near D69 residue. Water penetration may be important for biogenic amine receptor activation with the role of D69 residue and for highly conserved motif NPxxY in activation [13].

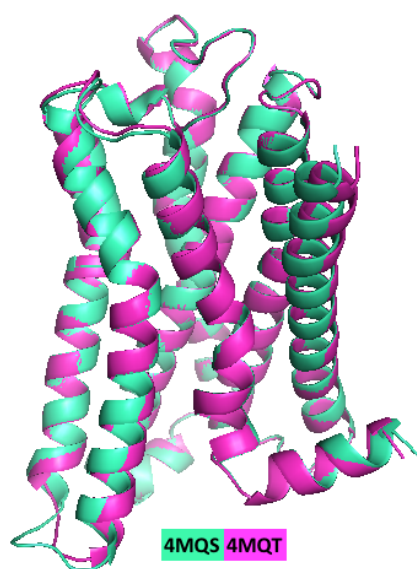


Figure 4.3. Active states of M2 Receptor.

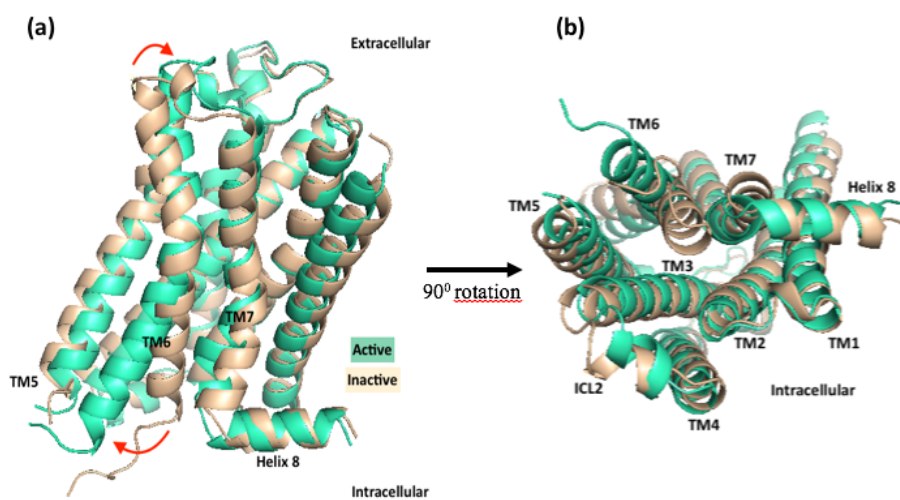


Figure 4.4. Active (4MQS) and inactive (3UON) crystal structures of M2 receptors.

(a) Red arrows show displacement of TM6 through inward and outward on extracellular side and intracellular sides respectively. (b) Intracellular side view of structures.

Where agonist iperexo binds to M2 receptor, tyrosine lid forms among Y104, Y403 and Y426 as a result of H-bonding of Y104 with Y403 and Y426. DRY (D120, R121, Y122) and NPxxY (N436, P437, Y440) motif regions also show structurally important displacements. In DRY motif region; N58 forms H-bond with R121 and D120, hence they are stabilized. Similarly, in NPxxY motif region; Y440 forms water-mediated H-bond with Y206. Allosteric modulator is placed just above the agonist iperexo and interacts with W422, N410, E172, N419, Y177, Y80, Y426.

Many studies have been focused on M2 receptor due to its importance and this study may contribute to the understanding of the activation mechanism through generating alternative transition pathways between initial and final states by utilizing collective ANM modes. For this purpose, many different cases have been taken into consideration like changing simulation parameters such as D_F , $modemax$ and R_{cut} , restricting different collective mode/modes, modelling and/or modifying initial and final structures after analyzing forward runs between 3UON and 4MQS in terms of selected ANM modes and their mean square fluctuations as well as running with the same simulation parameters, so as to disclose main determinants of the transition. Summary of performed simulations are given in Appendix.A and major results are represented in Section 4.1.1 and 4.1.2.

4.1.1. ANM-LD Simulation Results from Antagonist Bound Structure to Agonist Bound Structure

In this section of thesis, it is aimed to drive M2 receptor activation transition between initial-inactive (3UON) and target-active (4MQS) structures by utilizing the intrinsically accessible collective modes of 3UON by using ANM-LD simulations. For this purpose, parallel ANM-LD simulations were performed with D_F values 0.35, 0.4, 0.6, adaptive between 1 and 0.1; R_{cut} 10Å, 13Å, 18Å; $modemax$ 30, 100, all modes; restricting mostly selected ANM modes; modeling 3 and 23 residues from missing residues in ICL3 and deleting 4 and 2 residues from the beginning and end of the chain to predict conformational transition (activation) pathway(s).

Initially, the transition pathway was generated without making any changes on initial and target structures without modeling any missing residue in ICL3 loop. After analyzing the pathway results in terms of mostly selected ANM modes and mean square fluctuations, which direct the motion, it was observed that there are some unrealistic picks in the missing residue region, N-terminus and C-terminus of structures. Thus, we decided to model 3 missing residues (R216, I217, K218) in ICL3 by using MOD-LOOP server [44, 45] and delete 4 residues (K19, T20, F21, E22) from the beginning of structures and 2 residues (L455, M456) from the end of structures due to unrealistic picks in these regions in mostly selected ANM mode shapes and to avoid computational artifacts. Finally, by restricting envisaged most important modes individually and together, their effects on the transition pathway tried to be explored.

In order to find efficient simulation parameters R_{cut} , $modemax$ and D_F have been changed systematically. After trying different R_{cut} ($R_{cut}=10\text{\AA}$, 13\AA , 18\AA), $modemax$ ($modemax=30$, 100 , all modes) and D_F ($D_F=0.35$, 0.4 , 0.60 , $0.1-1$) values; R_{cut} value was set to 13\AA and it was decided to use all modes and adoptive D_F value between $0.1-1$ according to selected mode, which give the better final RMSD convergence to the target structure, to be able to compare results without concerning about whether different parameters cause a difference in M2 Receptor activation simulations or not.

Final RMSD between the predicted ANM-LD conformation at the end of the transition pathway towards the target conformation, the number of ANM-LD cycles that is the number of the predicted conformations defining the transition and the most selected ANM modes in the transition are summarized in Table 4.1.

RMSD is one of the most prevalent features to measure the protein similarities. RMSD difference between initial and target structures is 2.6\AA . RMSD plots belong to simulations are given in Figure 4.6 and for further analysis RMSD plots with selected modes, overlap and degree of collectivity values are also given in Figures 4.7. In the simulations, transition conformations are converged around 35 cycles; final RMSD values vary in the same range and near to 1.5\AA . Alignment of the predicted conformation at the end of the transition towards the target conformation and the target conforma-

tion are given in Figure 4.5 to demonstrate the success of simulations in terms of final RMSD. For the representation here, simulation 7 was used with the 1.54Å final RMSD value.

Table 4.1. ANM-LD Simulation result summaries from 3UON to 4MQS

| Simulation Number | Cycle Number | RMSD (Å) | The Most Selected ANM Modes |
|-------------------|--------------|----------|-----------------------------|
| 7 | 34 | 1.54 | 5, 4 |
| 24 | 30 | 1.46 | 5, 2, 6 |
| 30 | 30 | 1.49 | 1, 6 |
| 33 | 35 | 1.60 | 4, 5, 2 |

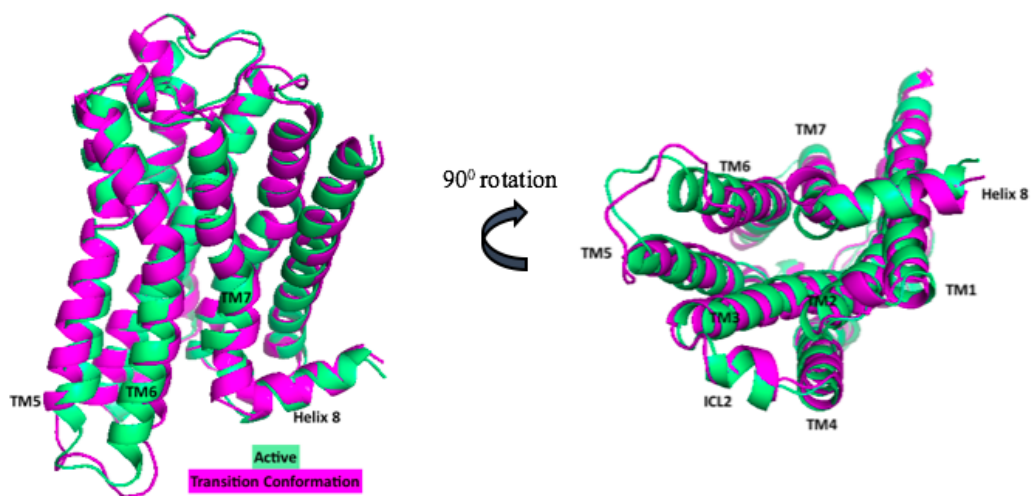


Figure 4.5. Alignment of generated final transition conformation of simulation 7 with the target structure, final RMSD value is 1.54Å.

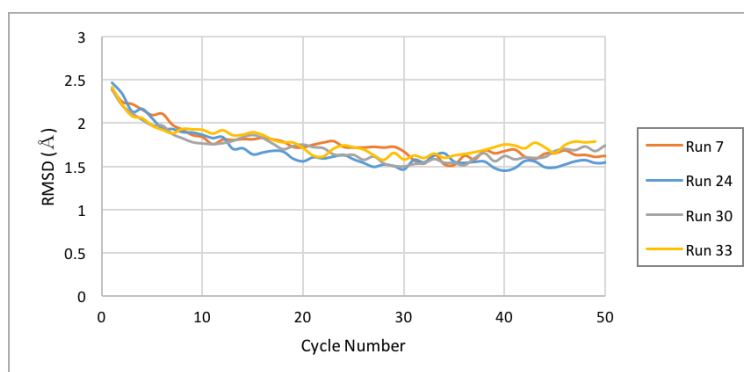
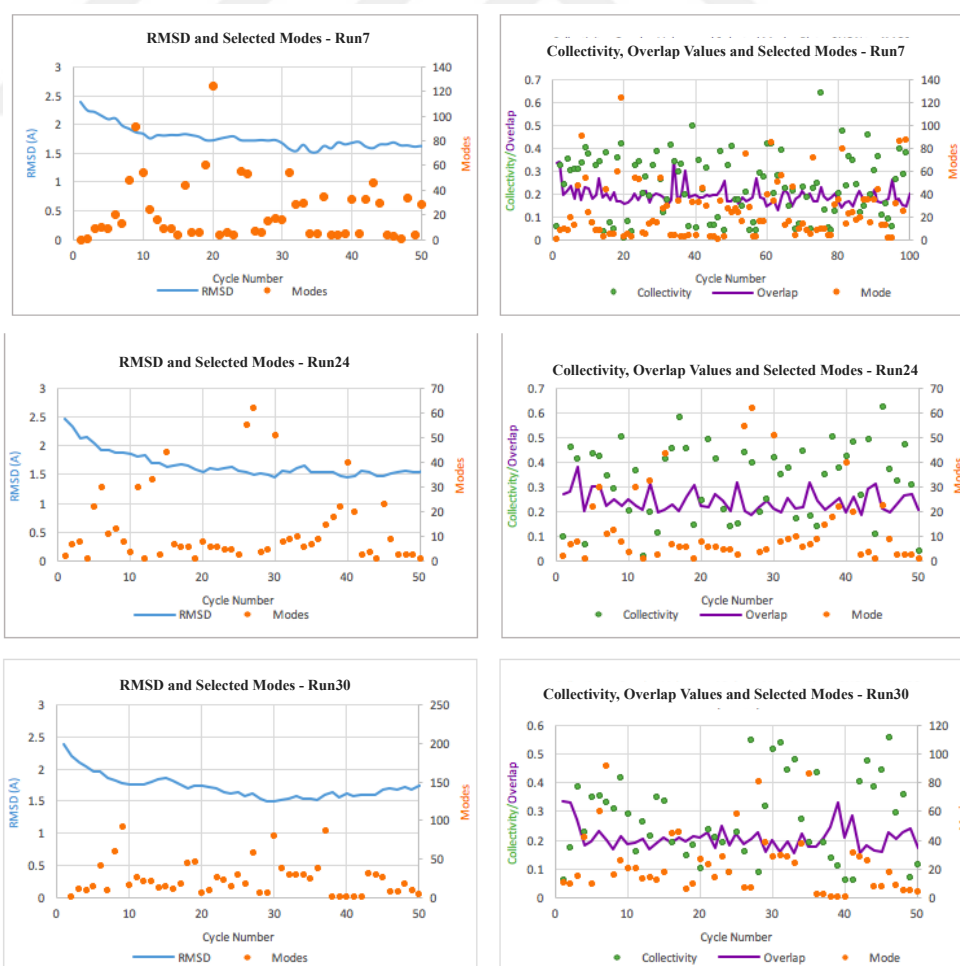


Figure 4.6. RMSD difference between generated transition conformations and target structure in simulations of M2 receptor from 3UON to 4MQS.



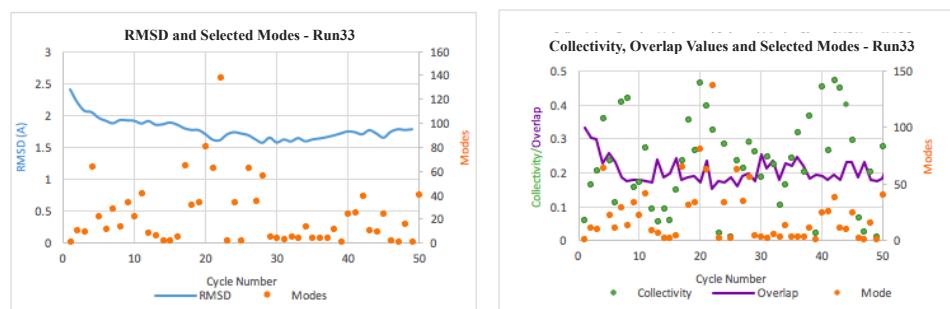


Figure 4.7. Selected mode, RMSD, overlap and collectivity values plots of simulations from 3UON to 4MQS. RMSD and selected modes plots are given in right side. Selected modes and corresponding collectivity and overlap values plots are given in left side.

For the activation of M2 receptor, the major displacements take place in TM6, TM5 and TM7 as it is mentioned in Section 4.1 Activation of M2 Receptor. The cage formation for ligand and G protein binding is important for activation of receptor. For this reason, major structural changes were designated by aligning inactive and active structures and calculating distances correspond to important regions like TM6 outward tilting region, middle of TM7 rearrangement region, DRY (D120, R121, Y122) and NPxxY (N436, P437, Y440) motifs. Beginning of TM5 (A184), ICL3 end of TM6 (T388), ECL3 end of TM6 (N410), region near the residue A438 in TM7 show big RMSD differences between initial-inactive structure (3UON) and final-target structure (4MQS). RMSD of ICL3 loop end of TM6 (T388) residues, ECL3 end of TM6 (N410) residues, residue A438 in TM7 and residue A184 at the beginning of TM5 are 7.5Å, 3.6Å, 3.1Å and 4.4Å respectively. Further than the structural changes, there are some known functionally important changes that take place during activation. They are hydrogen bond formation between Y206 located at ICL3 end of TM5 and Y440 located in TM7, breakage of salt bridge between R121 located at ICL2 end of TM3 and E382 located at ICL3 and the distance between residue R121 in TM3 and T386 in TM6 due to outward tilt of TM6 [41]. When the hydrogen bond is formed in active state, the distance between Y206 and Y440 decreases from 12.6Å to 3.7Å. The distance between TM3 and TM6 is calculated by using α -carbon atoms of R121 and T386, the values are

8.4Å and 13.7Å in inactive and active states respectively. Salt bridge between R121 and E382 is broken in the active state and the distance increases from 12.8Å to 17.4Å through activation.

In the simulations, these major distances and RMSD values were calculated and recorded in order to discover whether these major changes take place and required cage is formed for the ligand and G protein to bind the receptor is predicted to assess the physical and biological relevance of the pathway or not. When functionally important distances were analyzed, it was observed that H bond cannot be formed in the simulation. The distance between residues that form H bond decreases from 12.6Å to the minimum 7.2Å in simulation 24, which is the best result, instead of 3.7Å. Other H-bond distance results are 12.2Å, 11.6Å and 8.7Å for simulation 7, 30 and 33 respectively. The reason of unsuccessful H-bonding may be that H bond formation occurs in side chains and ANM-LD method depends on back bone calculation and this may not be enough to rotate side chain good enough to form H-bond. Other two important distances, which are breakage of salt bridge and TM3-TM6 distance, are closer to the target value in simulations. TM3-TM6 distance increases from 8.4Å to 11.0Å, 11.9Å, 9.8Å and 12Å in simulation 7, 24, 30 and 33 respectively, which the target value is 13.7Å. The distance between residues that breaks salt bridge is 19.6Å, 18Å, 18.8Å and 17.6Å in simulation 7, 24, 30 and 33 respectively, which the initial and target values are 12.8Å and 17.4Å. The RMSD values of selected residues are like this: RMSD of A184 decreases from 4.4Å to 1.9Å, 2.8Å, 1.7Å and 3.2Å, RMSD of T388 decreases from 7.5Å to 2.9Å, 2.4Å, 2.9Å and 2.1Å, RMSD of N410 decreases from 3.6Å to 1.4Å, 1.5Å, 1.7Å and 1.3Å, RMSD of A438 decreases from 3.1Å to 0.8Å, 1.5Å, 1.6Å and 1.7Å in Simulation 7, 24, 30 and 33 respectively. The distances and RMSD values through transition given in Figure 4.8 in detailed.

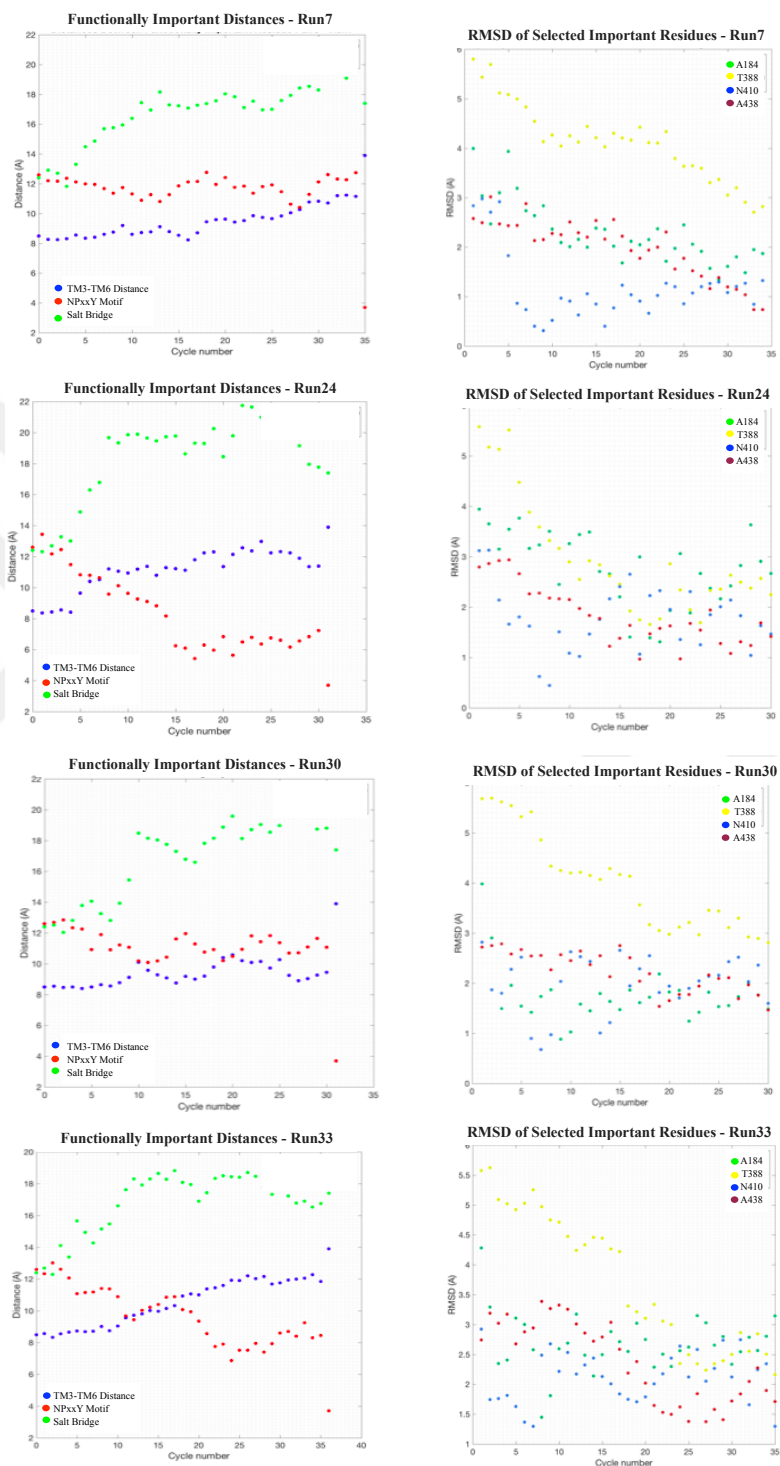


Figure 4.8. Functionally important distances through transition pathway (Left side) and RMSD of selected important residues between transition conformations and target state (Right side) in simulations of M2 receptor from 3UON to 4MQS.

Although overall RMSD converges around 1.5Å in the simulations, the final structures and transition pathways show differences as it is recognized from the distance graphs and RMSD graphs. These distances may also affect the water penetration through inside TM region, because it is known that functionally important sites like NPxxY motif are hydrated in the active state. These differences may be because of utilizing different collective ANM modes for transition in simulations. Different collective ANM modes direct the motion in different ways, thus affect the transition pathways and create different distances and RMSD values. The correlation between residue fluctuations were also examined to explore the network of dynamic interactions and also how this vary in the transition pathways generated. The cross-correlation maps are given in Figure 4.22.

Selected ANM modes during the transition pathway give an insight about how the inactive structure uses its intrinsic dynamics to undergo conformation changes to the active state. To this, the most selected modes were identified to disclose the dynamic determinants underlying the activation transition. ANM mode shapes (mean square fluctuations) were determined by calculating dot product of that ANM mode and then they were mapped to the initial structure's dynamic modes to be able to follow which modes of motion are repeatedly chosen and thus have the spectrum of chosen modes of motion on a common base, although ANM modes are calculated in every simulation step from generated conformation. During the transition pathway, with the conformational changes, the order of modes may shift or may slightly change with the topology. After mapping the selected modes to the initial structure's mode, the mode with the best overlap value was chosen as the selected mode. To be able to find the corresponding mode with the best overlap value, overlap value between selected mode and every mode of initial structure (among first slow 30 modes) is calculated by using Equation 4.1; where $O_{i,j}$ is overlap value between selected mode i in the simulation step and initial structure's mode j and \mathbf{u}_i and \mathbf{u}_j are ANM mode vectors of selected mode i in the simulation step and initial structure's mode j respectively. Then the mode with

the highest overlap value is chosen as corresponding mode in initial structure.

$$O_{i,j} = \frac{|\mathbf{u}_i \cdot \mathbf{u}_j|}{|\mathbf{u}_i| \cdot |\mathbf{u}_j|} \quad (4.1)$$

Overlap value below 0.5 was not regarded as that initial mode while generating selected mode distribution graphs, because the mode shapes become quite different below this overlap value. Selected mode distribution and overlapping plots are given in Figure 4.9. After determining the most selected modes, actual selected mode shapes that correspond to the most selected mode shape were also drawn in Figure 4.10 to display 0.5 is a good overlapping value and the selected modes are similar to the initial structure's modes. As it is seen in Figure 4.10, selected modes are quite similar to the one in the initial structure, so it is reliable to use the most selected modes to understand the transition mechanism. Selected mode similarities were also calculated by using overlap values between each selected ANM mode vector to the other selected ANM mode vectors. This similarity graphs give information about at which cycle the structure shows similar transition motions. It was observed that in the first-half of the transition, selected modes are much more similar (Figure 4.11), which is expected, since the structure makes more global motions by using slow modes in the first-half. In the second-half, however, the majority of the transition is mostly completed and RMSD begins to converge where mainly small (local) motions take the structure to the target state. Selected mode similarity graphs also give information about whether consecutive cycles' motions are similar and/or in the same direction or not. It was observed that the structure sometimes requires more than one ANM-LD simulation step to complete transition motion in the corresponding region.

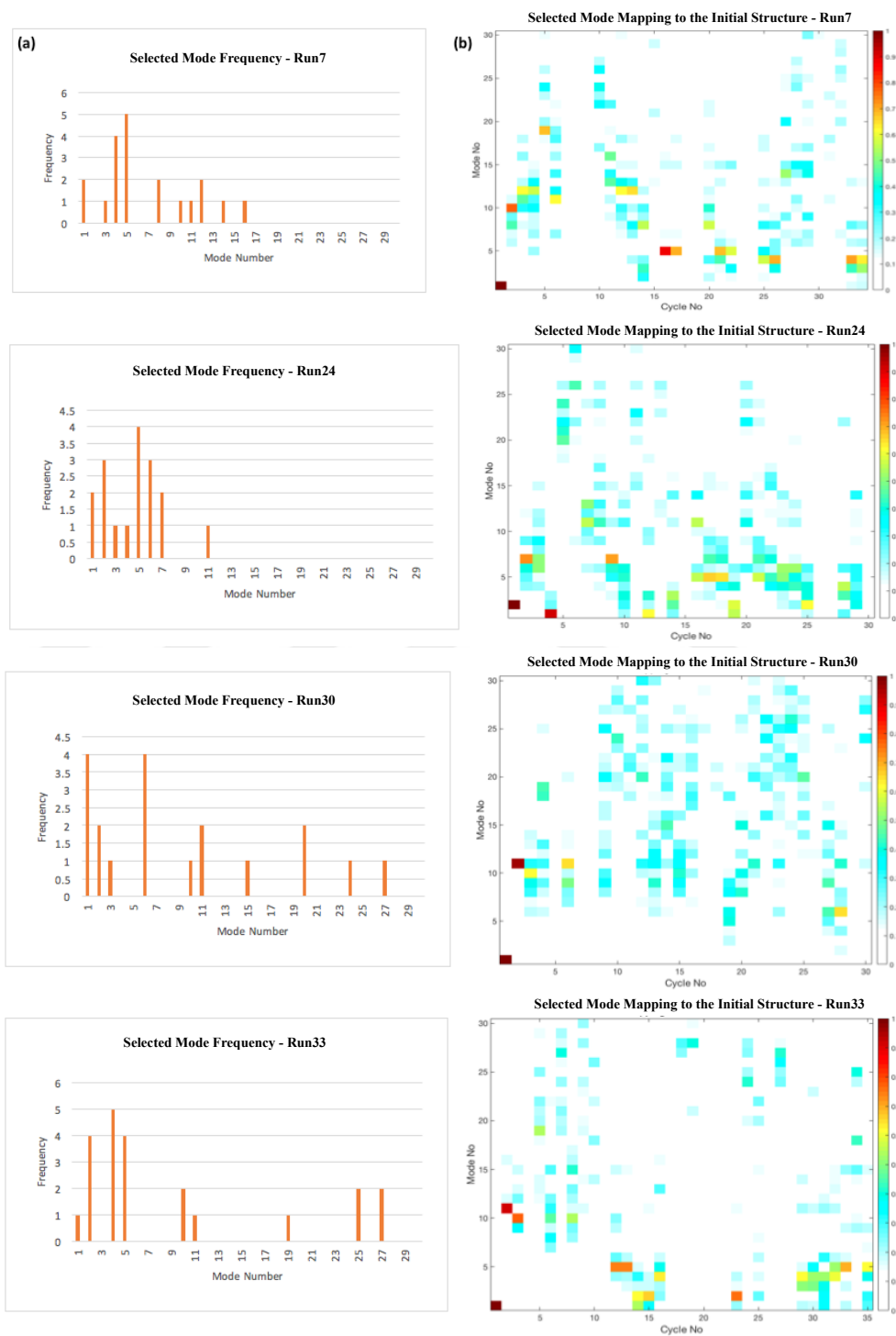


Figure 4.9. (a) Selected modes distributions. (b) Overlap values of selected modes of M2 simulations from 3UON to 4MQS according to the initial structure's first 30 modes.

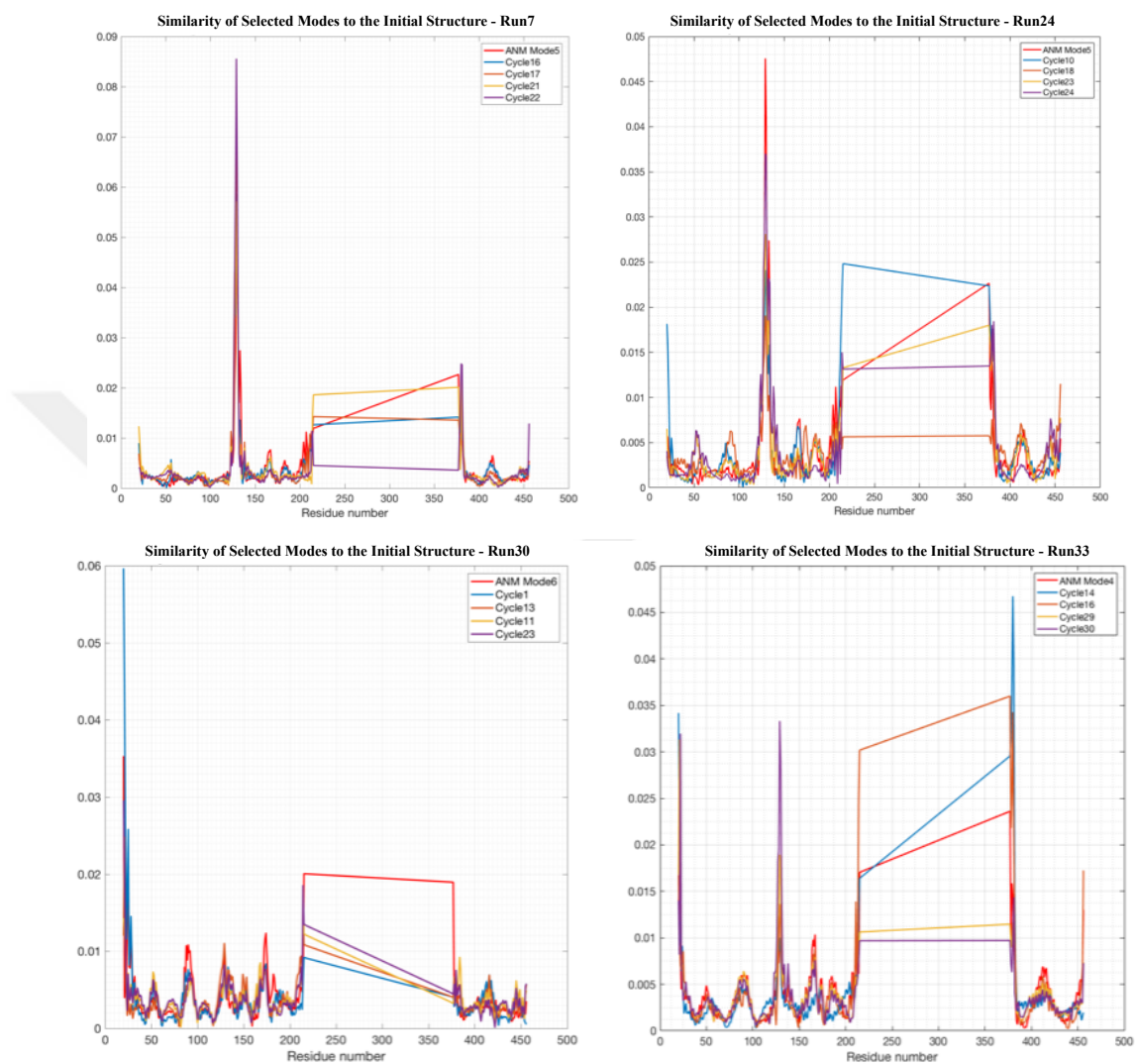


Figure 4.10. The most selected ANM mode shapes of initial state and corresponding pathway selected modes in M2 simulations from 3UON to 4MQS.

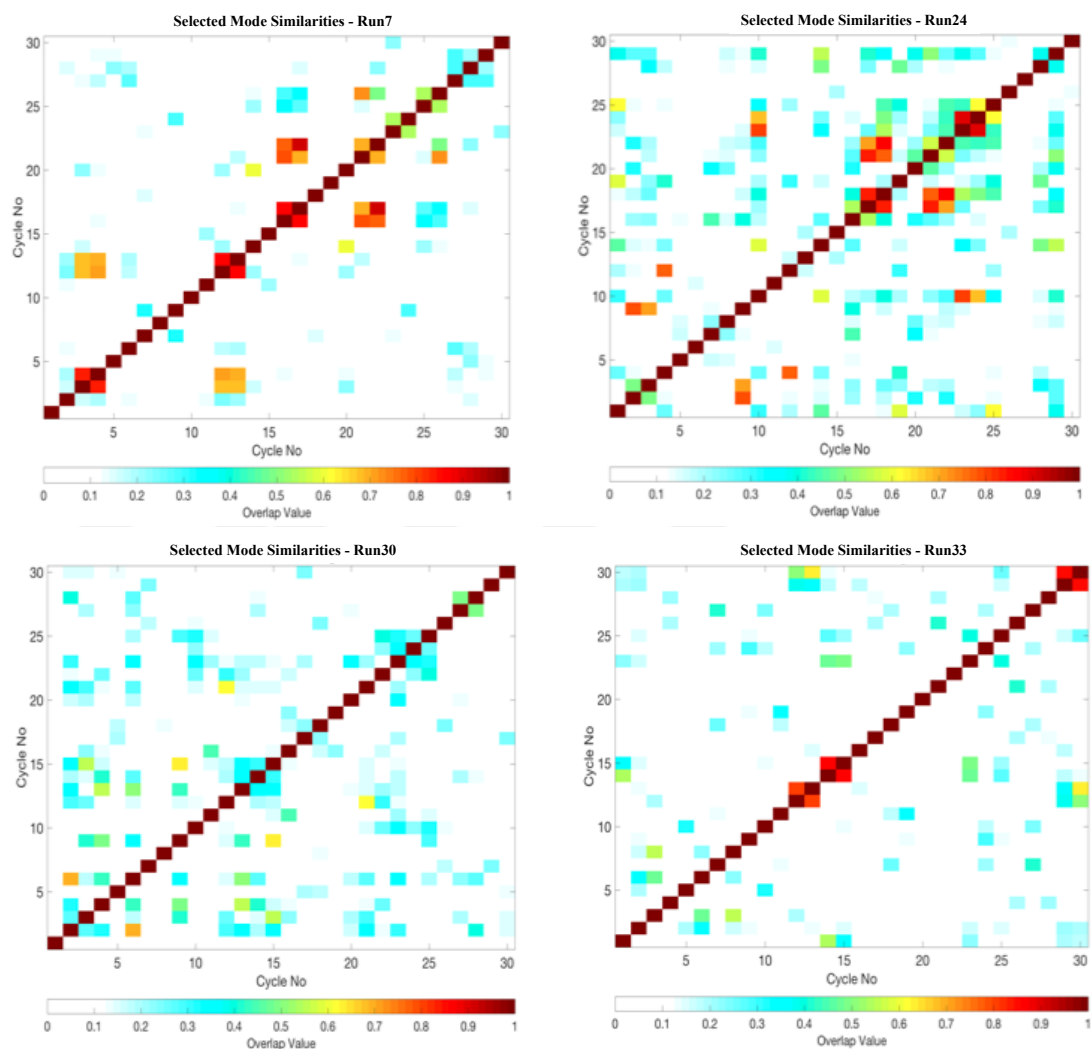


Figure 4.11. Selected modes similarities with each other in M2 simulations from 3UON to 4MQS.

Slowest 2^{nd} , 4^{th} , 5^{th} and 6^{th} ANM modes are the most selected modes considering ANM-LD simulations 7, 24, 30 and 33. Slowest 5^{th} mode is the most selected one; 4^{th} , 6^{th} and 2^{nd} modes appear interchangeably. Minimum sites in the profile of mean square fluctuations of a given mode corresponds to the hinge residues that coordinate the motion defined by this mode. The minimum points of 5^{th} ANM mode are V24, F25, I48-Y60, P90-V105, D120-R121, R135-A140, I153-W155, E172-C176, T190-F195, V385, N404-M406, V421-Y426, I435-Y440, A445-F451. It has been known that Y406 contacts with Y440 (related with NPxxY Motif), N58 contacts with R121 and D120

(related with DRY Motif) and Y104, Y403 and Y26 form tyrosine lid [40, 41]. The hinges of 5th ANM mode overlap with the functionally important residues, i.e., implying that while this modes of motion is coupled with DRY and NPxxY motifs. Mean square fluctuations of 5th ANM mode is given in Figure 4.12.a and the highlighted hinge residues of this mode on the structure by using PyMOL [46] is given in Figure 4.12.b. Highlighted residue regions correspond to the ligand binding site, where ligand is shown as sphere, and functionally important residues that are located on TM5, TM6 and TM7. When ANM 2nd, 4th and 6th modes are superimposed, their local minimum points are almost at the same sites and this may be the reason why these three modes appear interchangeably. Superimposition of 2nd, 4th and 6th ANM mode shapes are given in Figure 4.13. The minimum points of these modes correspond to residues G40-M45, Q55-Y80, L100-S110, W148-W155, S380-W400 and T420-Y440. Residues S380-W400 in TM6 and residues T420-Y440 in TM7 regions are functionally important, because the structure shows big conformational change in ICL3 end of TM6, which corresponds to residues S380-W400 and NPxxY motif's residues N436, P437 and Y440 are in TM7 region. Residues Q55-M80 in TM2 may be important for formation of water pathway, because there is a hypothesis about that residue D69 could be important for important for biogenic amine receptor activation with the formed water pathway [47].

Mode shapes give information about the motion in 1D by calculating the dot product of mode vector, however mode motions could be visualized in 3D by submitting the initial minimized structure and selecting corresponding mode from the ANM server [48]. Mostly selected ANM modes' motions, which are 2nd, 4th, 5th and 6th modes, according to initial structure by using ANM server in 3D are given in Figures 4.14-4.17 with the corresponding ANM mode shapes in 1D. The motions shown in the figures correspond to major/global motions of transition, where they are colored with red and demonstrated with bigger arrows. ICL3 end of TM6 shows the biggest conformational change during the transition. Beside than this region; ICL end of TM5, ECL end of TM1, TM2 and TM6 show slight conformational changes. When the mostly selected modes motions are examined, they are overlap with these changes. However, there are some picks in the mean square fluctuations graphs and these picks could dominate the motion in 3D and in ANM, local minimum points are much more important than the

picks to describe the motion, because minimum points behave like hinge points and direct the motions from these points. Due to this reason, it is also important to highlight hinge residues of mostly selected ANM modes on the structure to understand the transition pathway. The Figures 4.18-4.21 show the hinge residues of mostly selected modes in Simulation 7, 24, 30 and 33 respectively.

In ANM-LD methodology, transition pathway is generated by utilizing slow ANM modes and these modes cause global/major changes in transition as it was mentioned just above. It is expected that the corporation of most selected collective modes direct the transition motion from the corresponding hinge residue regions of most selected modes. As an exemplary case, Simulation 7, in which 4th and 5th ANM modes are selected mostly, was studied and results are presented in Figure 4.18. The minimum points of 4th and 5th mode shape graphs were colored on initial state with red and blue respectively and then initial state was aligned with target state. When the hinge residues of ANM mode 4 and 5 were examined, it was observed that 4th mode probably has effect on rearrangement of TM5 and middle of TM7, whereas 5th mode has effect on extracellular site of TM3, TM6 and TM7 and intracellular site of TM7. Hinge residues of mostly selected ANM modes in other simulations are given in Figures 4.19-4.21 for Simulation 24, 30 and 33 respectively.

When functionally important distances and RMSD values are taken into consideration, Simulation 30 has not good overlap with the target state. Although the overall RMSD converges to the target state in the same range as it is in other simulations, functionally important distances could not reach the target structure's values (4.8). This may be due to location of hinge residues of mostly selected ANM modes, since there isn't any hinge residue in the middle of TM7, ECL1 end of TM1 and the middle of TM6. These regions show important conformational changes especially TM6 and TM7. Other 3 simulations (Simulation 7, 24 and 33) could approach to the target values better than the simulation 30. The common feature of these simulations is selecting ANM mode 5 during the transition as one of the most selected modes. Mostly selected modes of Simulation 30 are 1 and 6. Although 6th mode's hinge residues contain the NPxxY motif and D69 residue in TM2 region,

it may not be enough to direct the required motion during transition with the 1st mode. Especially TM3-TM6 distance may not be obtained due to missing hinge residue in middle of TM6 (Figure 4.20). Thus, it can be said that the driven motions from hinge points are critical for activation of M2 Receptor, where they overlap with major displacements take place. This information is induced from ANM modes, so ANM modes harbor intrinsic dynamic of proteins that is directly with function.

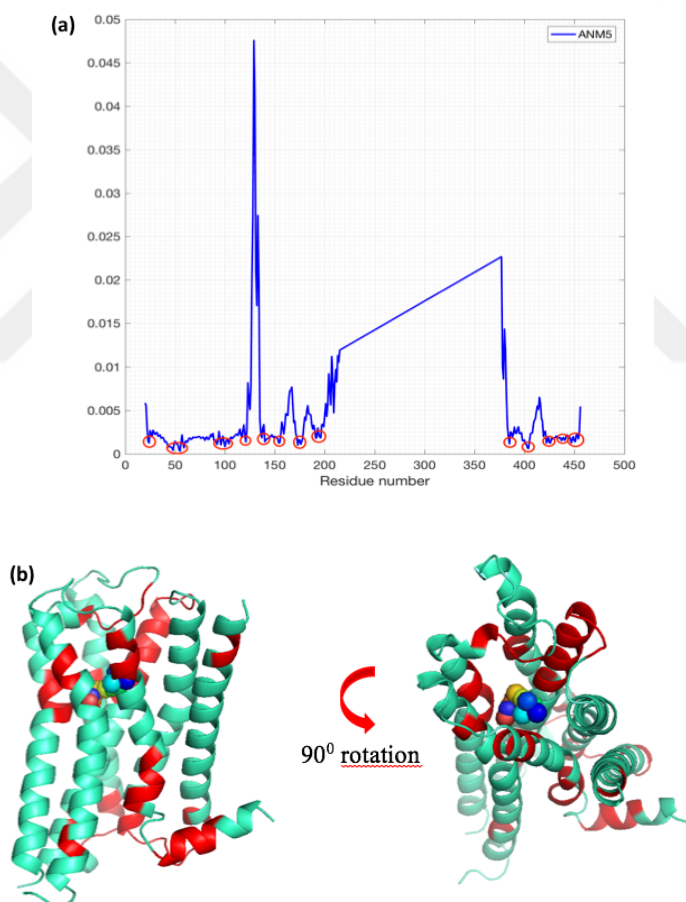


Figure 4.12. (a) 5th ANM Mode shape of the initial state. (b) Demonstration of minimum points of 5th ANM mode on structure.

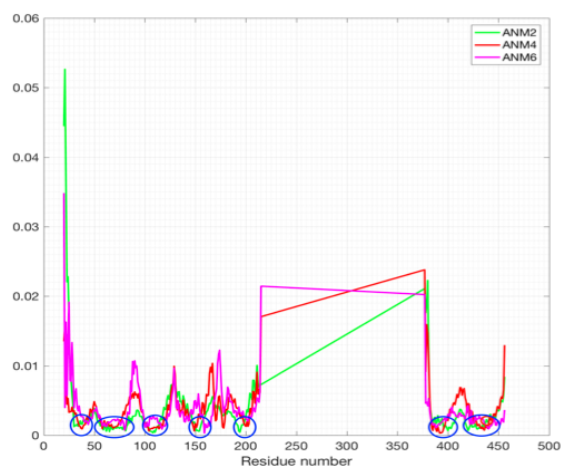


Figure 4.13. Superimposition of 2nd, 4th and 6th ANM modes with highlighted minimum regions.

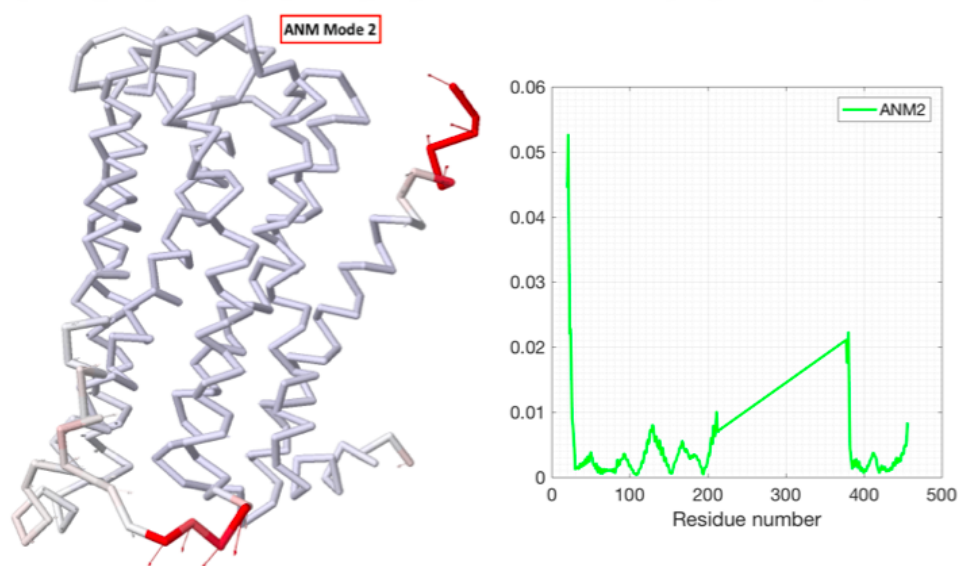


Figure 4.14. 2nd ANM mode motion representation from ANM server and its ANM mode shape.

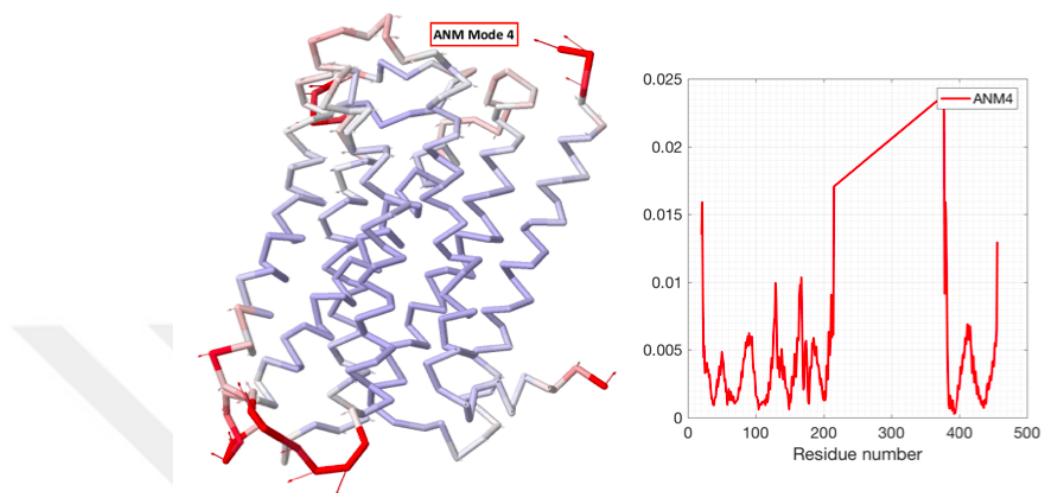


Figure 4.15. 4th ANM mode motion representation from ANM server and its ANM mode shape.

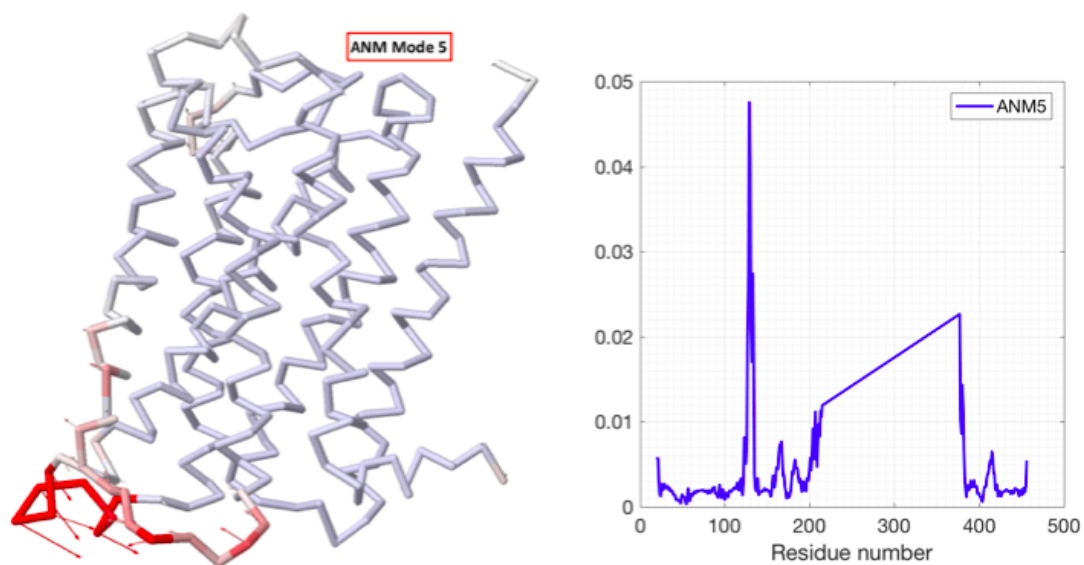


Figure 4.16. 5th ANM mode motion representation from ANM server and its ANM mode shape.

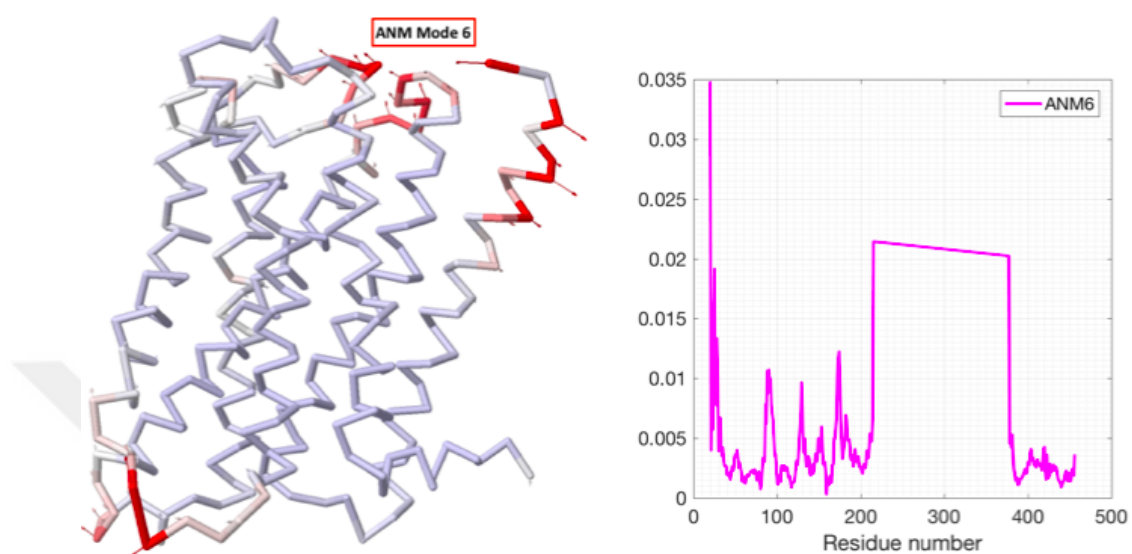


Figure 4.17. 6th ANM mode motion representation from ANM server and its ANM mode shape.

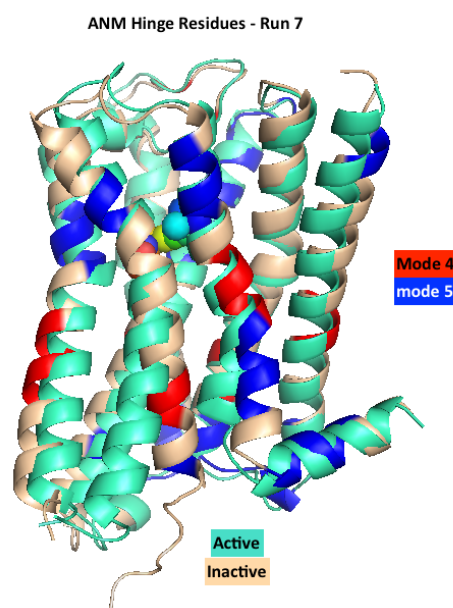


Figure 4.18. Hinge residues of most selected ANM modes in Simulation 7.

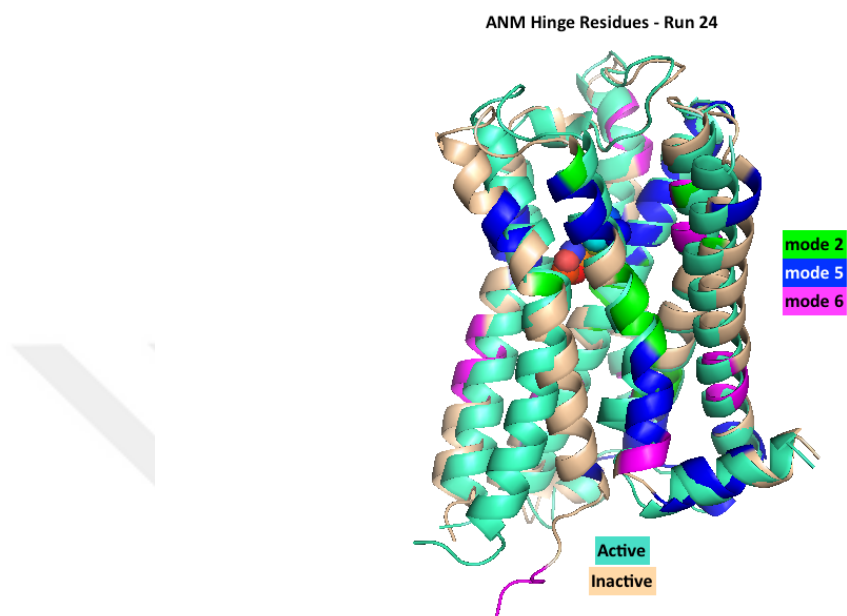


Figure 4.19. Hinge residues of most selected ANM modes in Simulation 24.

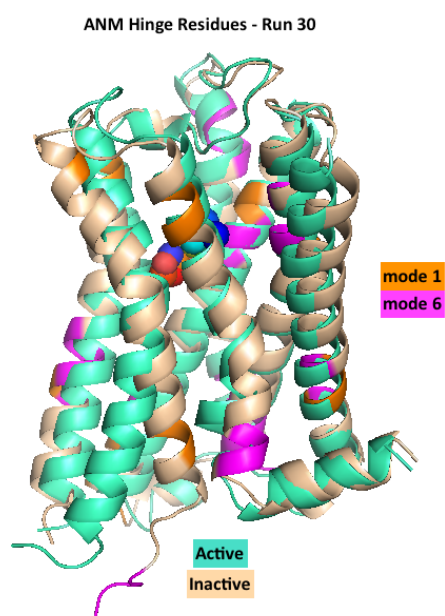


Figure 4.20. Hinge residues of most selected ANM modes in Simulation 30.

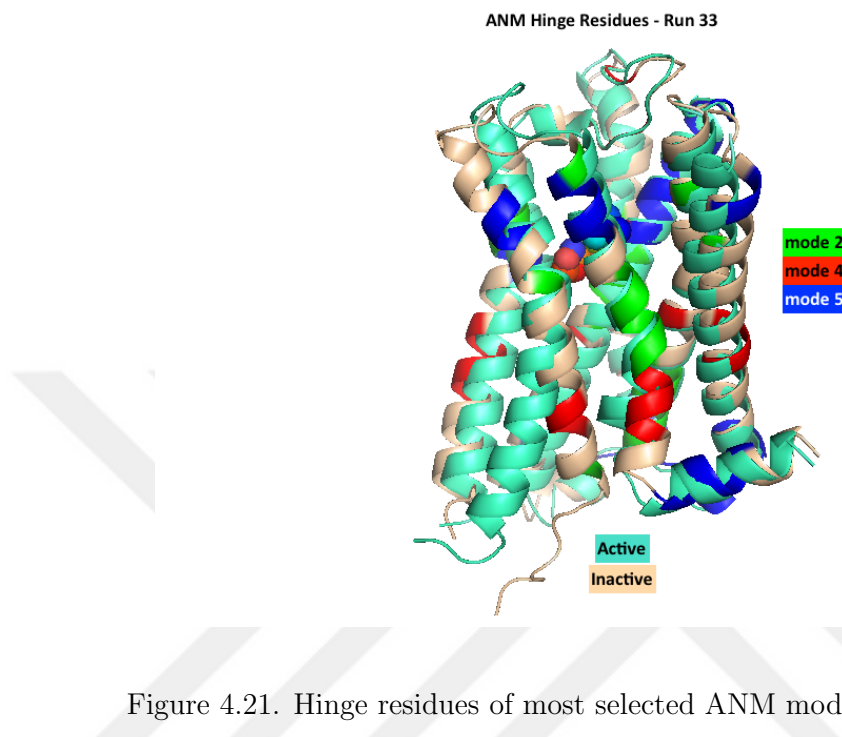
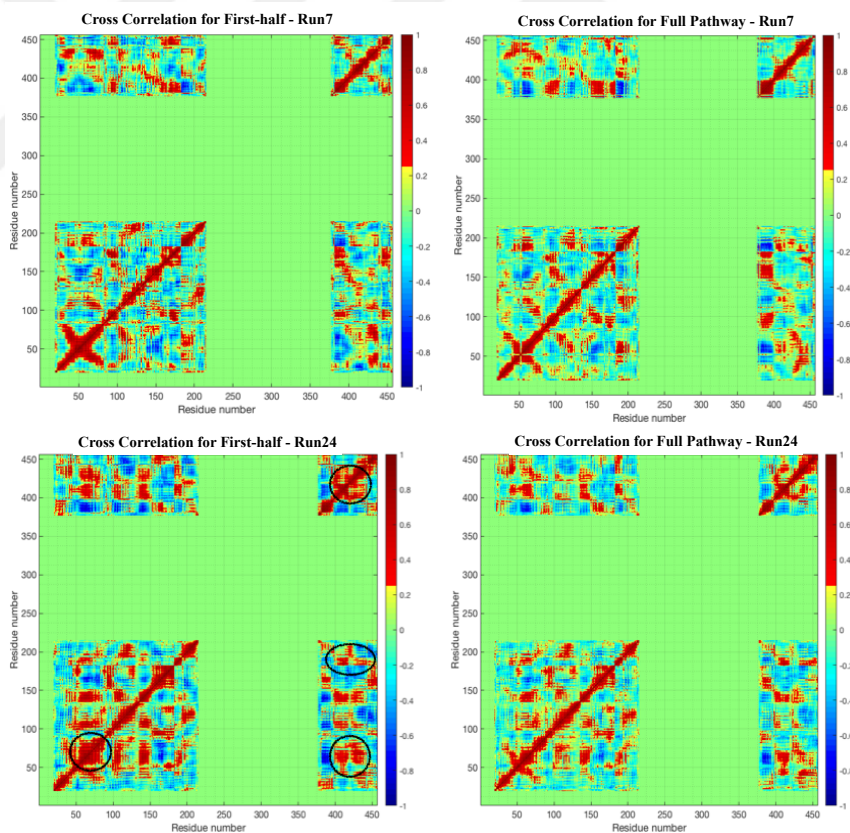


Figure 4.21. Hinge residues of most selected ANM modes in Simulation 33.

The correlation between residue fluctuations are given in normalized cross-correlation maps. Cross-correlations are normalized in the range $[-1, 1]$, the values below zero means negatively correlated residue pairs, and above zero means positively correlated residue pairs during the transition pathway. Correlation value near to the 1 means the most positively correlated motion and these regions may be functionally important regions. Cross-correlations maps were calculated for the full pathways and for the first-half of the pathways, where most of the transition is completed. It is important to note that these correlations are based on the transition pathway, and thus they are based on where the window is taken along the transitions. There are some differences in cross-correlations maps, although all the simulations were performed with the same parameters. Utilizing different collective ANM modes may be the reason of differences in cross-correlation maps. We focused on conserved regions in all maps to be able to reach a conclusion. Cross-correlation maps are presented in Figure 4.22.

When the first-half pathway cross-correlations were compared with the full pathway cross-correlations, it was observed that residues between M45 and Q55 in TM1 are

more strongly correlated with T430-Y440 residues in TM7, where TM7 is rearranged and become closer to the TM1 residues in the first-half. Middle of TM7 is rearranged during transition and this rearrangement may affect the correlation with the TM1 residues. Similarly, residues between W400-N410 in TM6 are more correlated with residues between A185-F195 in TM5 and less correlated with T420-Y430 residues in TM7 in the first-half. These correlations may be due to inward tilting of TM6 residues between W400-N410 during transition. The correlation between residues V50-F75 and W400-Y430 are more correlated with themselves in the first-half. The structure is quite small, which contains 276 residues, and this may also cause highly correlated regions in the structure.



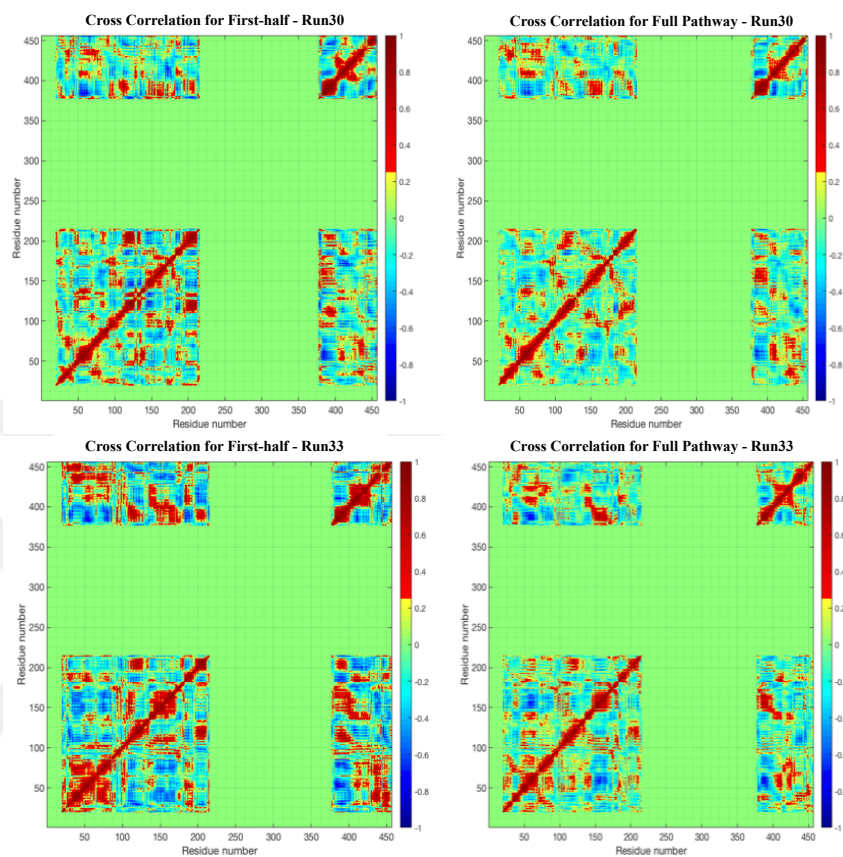


Figure 4.22. The first-half (Left) and full-pathway (Right) cross-correlation maps for ANM-LD simulations of M2 receptor from 3UON to 4MQS.

After analyzing ANM mode motions, mean square fluctuations, hinge residues and functionally important distances, it was decided to make some modifications on structures, due to unrealistic picks in mostly selected modes and then restrict certain modes; like mode 4 and 5 to reveal their effect on transition pathway. Firstly, initial and target structures were modified in two ways to hinder the unrealistic picks in ANM mode shapes. Then after deciding the functionally important ANM modes, restricted mode simulations were performed and results are presented here respectively.

In ANM mode shape graphs (Figure 4.14-4.17), there are some unrealistic picks around ICL3 loop, where the residues P216-S377 are missing, and/or near the C-terminal and/or N-terminal. These picks may effect the transition hindering other

motions due to their magnitude. For this reason, new ANM-LD simulations were performed by modifying initial and target structures in two ways. In the first way, it was excluded the residues cause pick formation which are located at the beginning and end of the structure. The first 4 (K19, T20, F21, E22) and last 2 (L455, M456) residues were deleted from the initial and target structures. In the second way, addition to the deleting the terminal residues, some known residues in ICL3 loop were modeled. After deleting residues, by using MODLOOP [44,45] server 3 residues were modeled in ICL3 loop, which are R216, I217 and K218. The modeled residues are known residues, which are given in their PDB files, although they were not located in the experiment. A quite small loop was modeled to avoid creating nonexistent correlations with other residues and/or hindering transition, but also enough length to bound two ends of ICL3 loop. The simulation results are summarized in Table 4.2.

Table 4.2. ANM-LD Simulation result summaries of modified structures.

| Simulation Number | Modification Explanation | Cycle Number | RMSD (Å) | The Most Selected ANM Modes |
|--------------------------|---------------------------------|---------------------|-----------------|------------------------------------|
| 56 | Terminals Modified | 50 | 1.55 | 6 |
| 57 | Terminals Modified | 40 | 1.44 | 2 |
| 58 | Terminals and ICL3 Modified | 50 | 1.50 | 2,7 |
| 59 | Terminals and ICL3 Modified | 35 | 1.50 | 1,2 |

Deleting first 4 residues (K19, T20, F21, E22) from the beginning and last 2 residues (L455, M456) at end of the structures reduces the unrealistic picks at the beginning and end of the selected modes. Modeling 3 residues in ICL3 loop decreases the pick height in ICL3 region considerably. However, when modified structures' simulations are compared with the unmodified structures' simulations, the most selected modes shapes are still almost the same and RMSD convergence is not so different than as it is in the unmodified simulations. The new simulations also converged around 1.5\AA , but in slightly longer time in the two cases that are 56^{th} and 58^{th} simulations. The final achievement in terms of RMSD value is not different and the final predicted conformations at the end of the transition towards the target state resemble to each other. RMSD graph of modified structures' simulations are given in Figure 4.23.

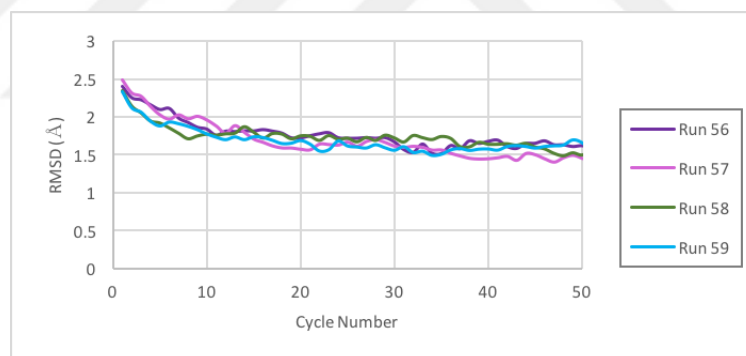


Figure 4.23. RMSD difference between generated transition and target conformations by using modified structures in simulations of M2 receptor from 3UON to 4MQS.

When the only terminals modified structure's mostly selected ANM mode shapes and both terminals and loop modified structure's mostly selected ANM mode shapes were superimposed to compare their similarities, it was observed that 6^{th} ANM mode of only terminals modified structure overlaps with the 7^{th} ANM mode of both terminal and loop modified structure and their 2^{nd} mode shapes also overlap in terms of minimum points. Superimposed ANM mode shapes are given in Figure 4.24, the difference occurs especially in ICL3 loop region in these selected modes, the modeled structure's ANM mode has lower pick in ICL3 loop region as expected, but with the same shape. Then,

modified structures and unmodified structures' most selected modes were superimposed to reveal modification effect on structures. 6th mode of unmodified structure is in good agreement with 7th mode of modeled and terminal modified structure and 6th mode of only terminal modified structure. The difference occurs around residues E175-A185, but in the 2nd mode shape there is also a minimum in this region and that may cover the change. 2nd mode shape of unmodified structure is in perfect agreement with new simulations' 2nd modes shapes and also good agreement with 1st mode shape of terminal modified structure with slight dislocations residues around L150 and F180.

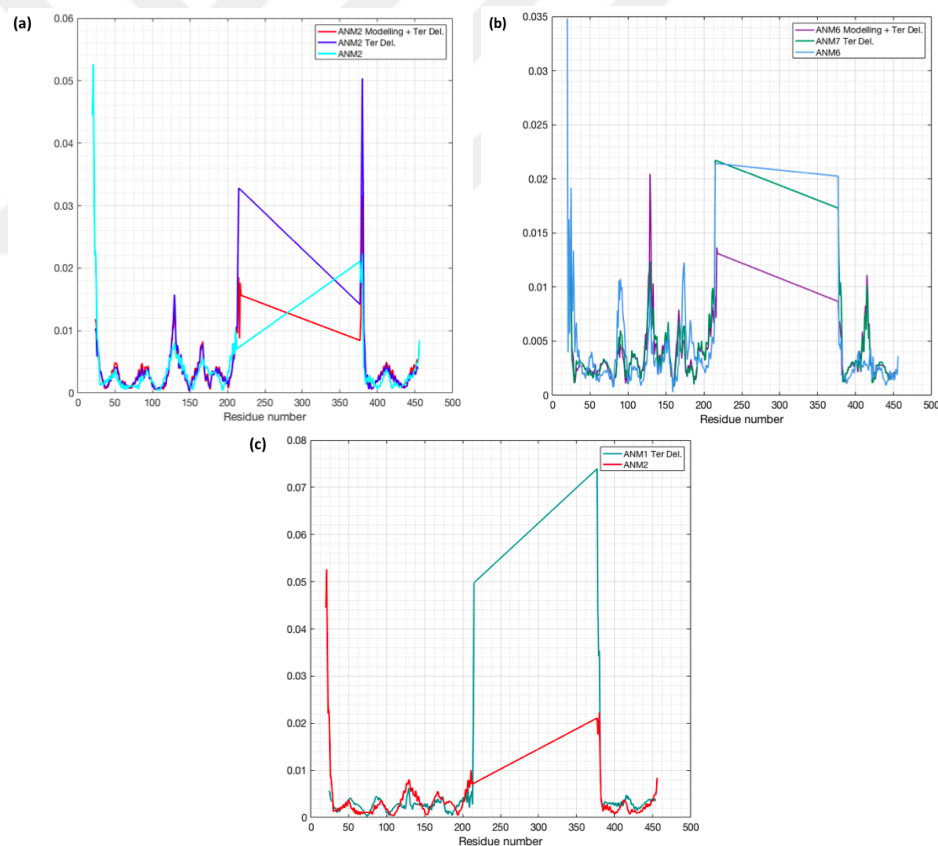


Figure 4.24. Mostly selected ANM mode shapes superimposition of modified and unmodified structures simulations. (a) 2nd modes of three initial structures. (b) 6th mode of unmodified structure, 7th mode of only terminals modified structure, 6th mode of modeled and terminals modified structure. (c) 2nd mode of unmodified structure, 1st mode of only terminals modified structure.

5^{th} and 4^{th} ANM modes are the other most selected modes in unmodified structures' simulations, however these two modes do not appear in modified structures' simulations. 4^{th} and 6^{th} mode appear interchangeably in unmodified simulations, maybe due to this interchangeability of modes, 4^{th} mode is not selected in the modified simulations. Probably instead of 5^{th} mode, 1^{st} mode appears as one of the most selected modes. It seems that 1^{st} mode substitutes with 5^{th} mode and utilizes required motion as collaborated with other modes, but in some cases the transition is completed in longer time. Interestingly, this result is also consistent with simulations given with unmodified structure and restriction of 5^{th} mode, which is the most selected mode.

The transition without selecting 5^{th} mode was simulated to reveal the effect of 5^{th} mode on the transition whether the absence of 5^{th} mode could hinder the transition or other modes could be replaced to make the required transition. In these simulations; 1^{st} mode appears as one of the most selected modes in all restricted simulations. 4^{th} and 15^{th} ANM modes are the other most selected modes. Overall RMSD converges around 1.3\AA and that is lower than the unrestricted values. Although overall RMSD decreases to 1.3\AA and 1.21\AA in simulation 22 and 28, the distance between Y440 and Y206 that forms H-bond does not show decreasing trend in simulation 22 and the distance between TM3 and TM6 does not approach to the target value in neither of simulations. These results show us there are some critical motions for the activation, and these motions can be revealed by utilizing certain ANM modes. Run parameters, RMSD convergence graph and selected mode shapes results for 5^{th} mode restricted simulations are presented in Table 4.3, Figure 4.25 and 4.26 respectively.

Table 4.3. ANM-LD simulation result summaries from 3UON to 4MQS with 5^{th} mode restriction.

| Simulation Number | Restricted Mode | Cycle Number | RMSD (\AA) | The Most Selected ANM Modes |
|-------------------|-----------------|--------------|-----------------------|-----------------------------|
| 22 | 5 | 32 | 1.30 | 1 |
| 28 | 5 | 62 | 1.21 | 4, 15, 1 |

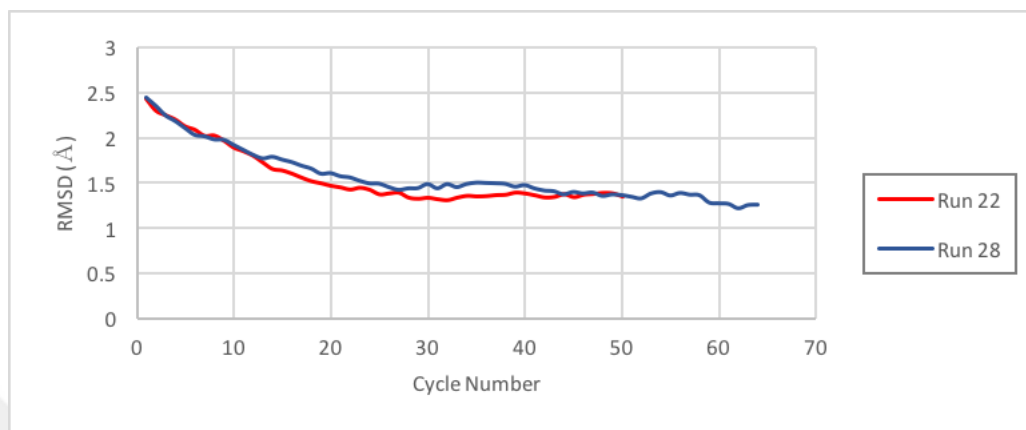


Figure 4.25. RMSD difference between generated transition conformations and target conformation in simulations of M2 receptor from 3UON to 4MQS with 5th mode restriction.

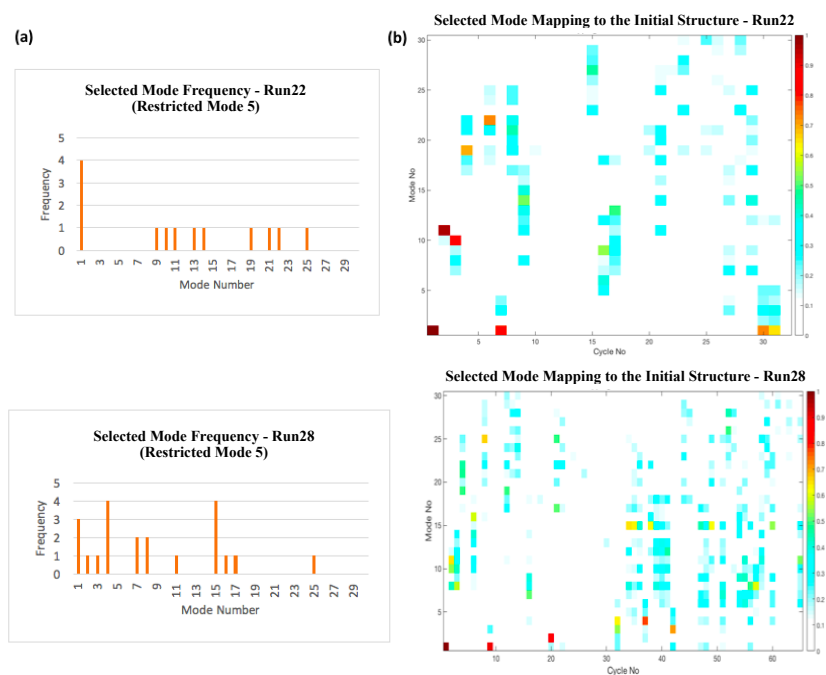
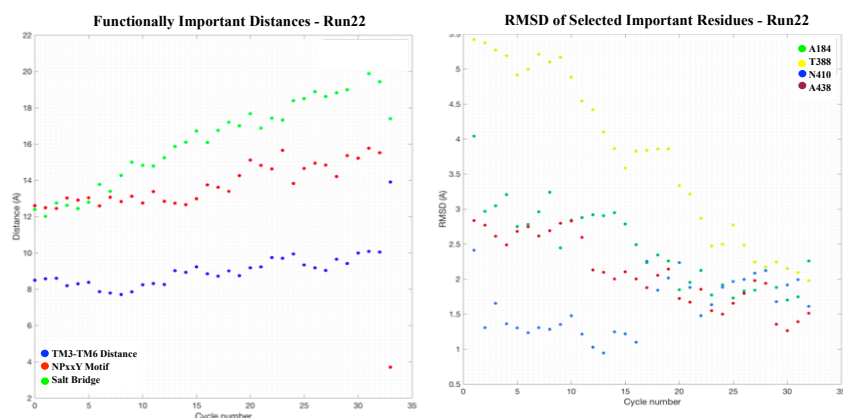


Figure 4.26. (a) Selected modes distributions (b) Overlap values of selected modes of M2 simulations from 3UON to 4MQS with 5th ANM mode restriction according to the initial structure's first 30 modes.

Functionally important distances and RMSD of superimposed residues between transition and target conformations of 5th mode restricted runs are given in Figure 4.27. In simulation 22, neither of functionally important distances could approach to the target values. In simulation 28, the distance between salt bridge atoms could achieve and although H-bond could not be formed the distance between H-bonding atoms decreases. TM3-TM6 distance increase from 8.4Å to 10Å instead of 13.7Å. When simulation 22 and 28 are compared in terms of functionally important distances, simulation 28 gives better results; since it could approach target values and/or show the appropriate trends to target values. The difference between simulation 22 and 28 may be due to selection of 4th mode in simulation 28. It can be said that the selection of 4th mode in simulation 28 affects the transition/activation of M2 receptor positively in terms of functionally important distances given in Figure 4.27.

When unrestricted run results given in Figure 4.8 were compared with the restricted run results given in Figure 4.27 in terms of functionally important distances and RMSD values, it was observed that functionally important distances could not approach as good as it is in unrestricted runs, although overall RMSD converged at lower values in the restricted simulations than the unrestricted simulations. This shows that 5th ANM mode's motion may be crucial for activation of M2 Receptor, because these distances have major effect on creating a cage that is necessary for ligand binding and/or activation.



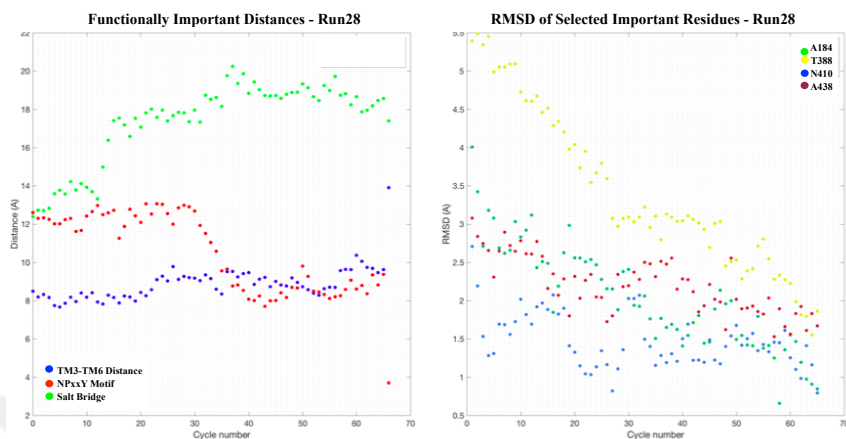


Figure 4.27. Functionally important distances through transition pathway (Left side) and RMSD of selected residues between transition conformations and target conformation (Right side) in M2 simulations from 3UON to 4MQS with 5th mode restriction.

Cross-correlations between residue pairs also examined for 5th mode restricted simulations in Figure 4.28. When full pathway cross-correlation maps were compared with the first-half cross-correlation maps, it was observed that residues between V50-L70 and Y430-T450 loose correlations with themselves. Residues between L100-D120 loose correlations with residues between V50-L100 and increase correlations with Y430-T450. These regions are functionally important regions, residues between Y430 and T450 include NPxxY motif and between L100 and D120 include DRY motif with the relation of N58 residue in TM2. Possibly these lost correlations affect to obtain functionally important distances and the convergence time, which is doubled in Run 28. Then cross-correlation maps were compared with unrestricted simulations' results (Figure 4.22, it was observed that correlation between L100-D120 and V50-L70 residues; T420-Y430 and L390-W400 residues increases, whereas correlation between P90-S110 and S380-W400 residues, Y80-L100 and S380-W400 residues and correlation of T420-Y430 residues with itself decreases. Correlation decreases in the region of Y80-L100, S380-W400 and T420-Y430 may affect the tyrosine lid formation between Y104, Y403 and Y426 and water pathway formation in the region of TM6. Outward tilting of TM6 maybe could not approach to the target due to lost correlation between S380-W400

in TM6 and P90-L100 in TM3. Further, the lost correlation of T420-Y430th residues with itself may affect the rearrangement of TM7. These cross-correlation maps show that although the initial state converges to the target state, it may use an alternative pathway by utilizing different slow ANM modes.

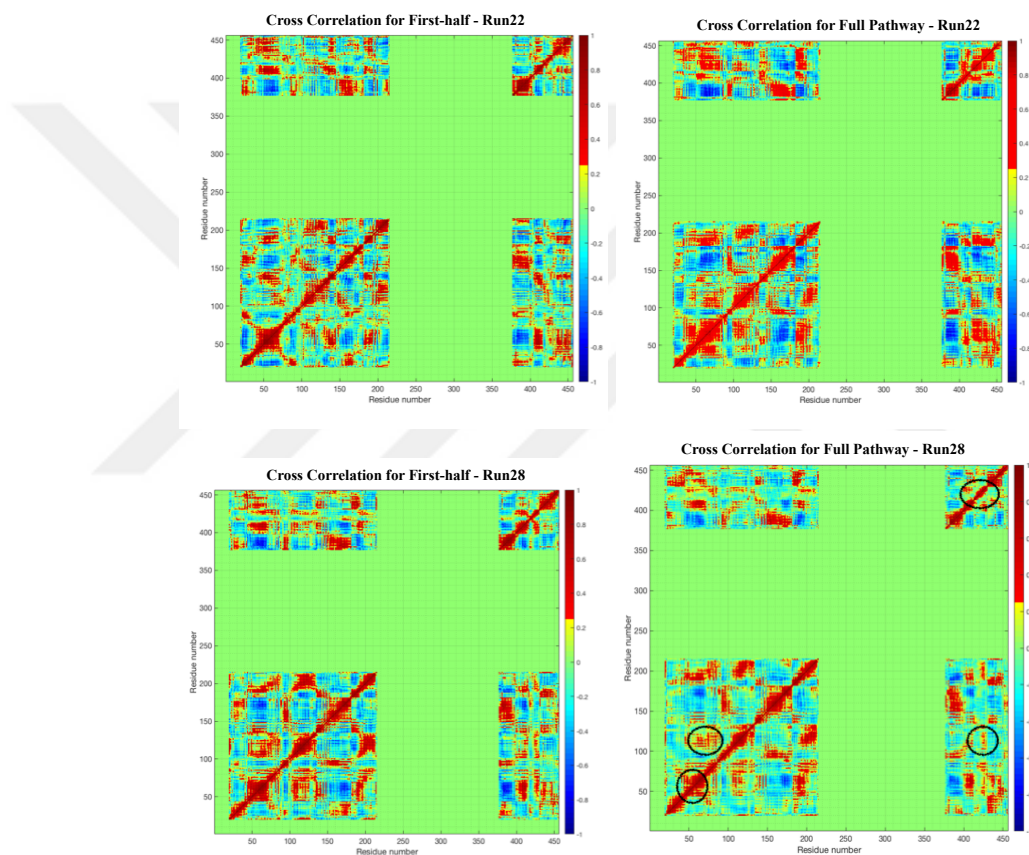


Figure 4.28. The first-half (Left) and full-pathway (Right) cross-correlation maps for ANM-LD simulations of M2 receptor from 3UON to 4MQS with 5^{th} mode restriction.

After analyzing 5^{th} mode restricted simulations' results, the importance of 4^{th} mode stands out and 4^{th} mode restricted runs were simulated to investigate its effect on the activation. RMSD, selected mode, collectivity and overlap value graphs are given in Figure 16 and selected mode distribution graph is given in Figure 4.29. RMSD decreases to 1.2\AA , but the convergence time more than double when it is compared with the unrestricted simulations' results. The most selected ANM modes are 1, 2

and 5. Functionally important distances quite approach to the target state given in Figure 4.31. Selection of 5th ANM mode could direct the motion in the correct direction in terms of functionally important distances with the help of 1st and 2nd modes, but the activation is decelerated. As further analysis, when cross-correlation map given in Figure 4.32 was compared with cross-correlation maps of unrestricted simulations (Figure 4.22), residues loose some correlations, like the residues between Y80-L100 and T420-T450 loose correlation with themselves and residues between F180-I200 loose correlation with residues between Y430-T450. As it is in the 5th mode restricted simulations, correlation lost between Y80-L100 and T420-T450 may affect the NPxxY motif. These regions are important regions for activation of M2 receptor and these lost correlations may also extend the simulation time. Hence, it can be said that 4th ANM mode is important for activation of M2 Receptor especially in terms of transition/activation time.

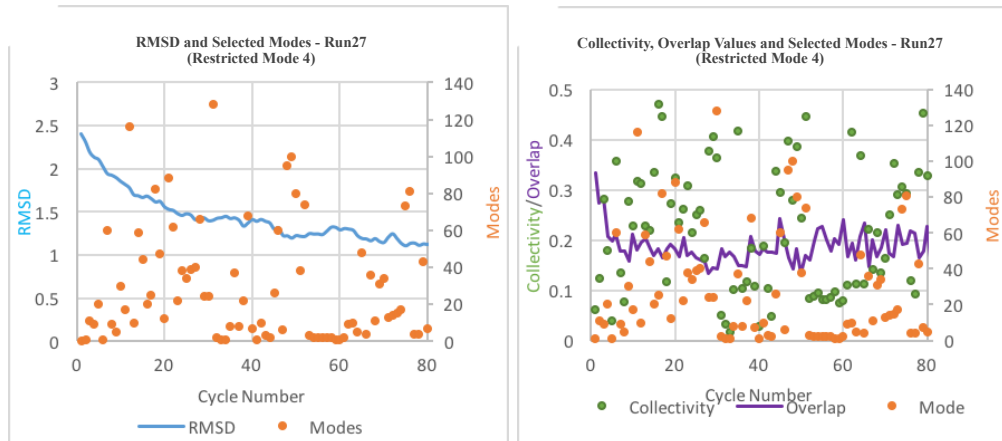


Figure 4.29. Selected mode, RMSD, overlap and collectivity values plots of simulations from 3UON to 4MQS with 4th mode restriction. RMSD and selected modes plot is given in right side. Selected modes and corresponding collectivity and overlap values plot is given in left side.

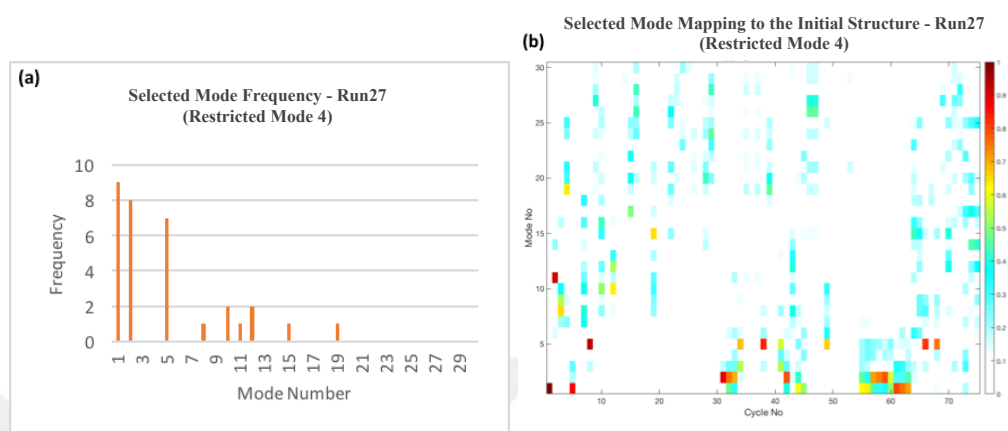


Figure 4.30. (a) Selected modes distribution. (b) Overlap values of selected modes of M2 simulation from 3UON to 4MQS with 4th ANM mode restriction according to the initial structure's first 30 modes.

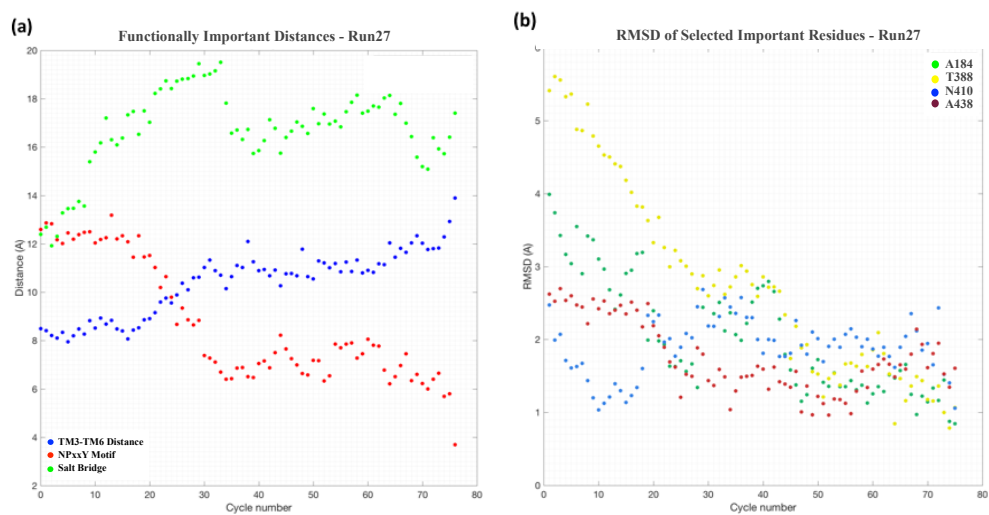


Figure 4.31. (a) Functionally important distances through transition pathway. (b) RMSD of selected residues between transition conformations and target conformation in M2 simulations from 3UON to 4MQS with 4th mode restriction.

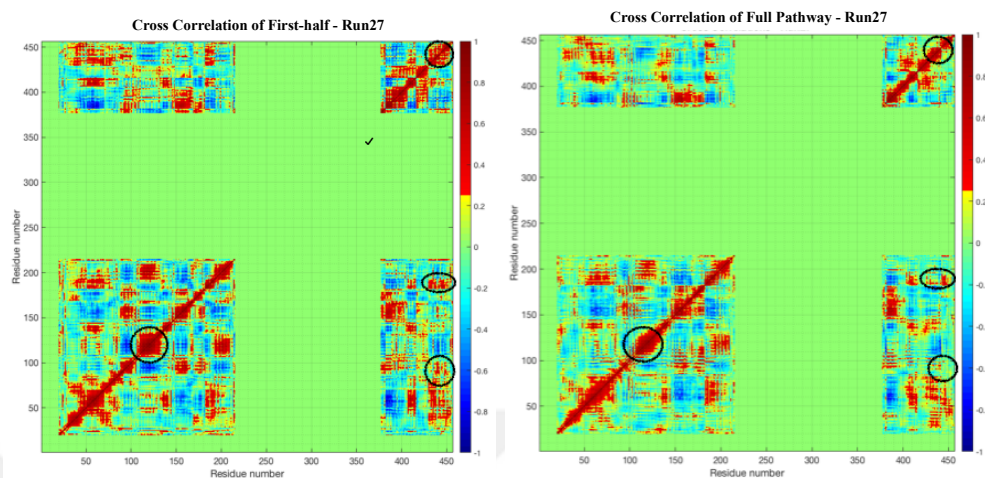


Figure 4.32. The first-half (Left) and full-pathway (Right) cross-correlation maps for ANM-LD simulations of M2 receptor from 3UON to 4MQS with 4th mode restriction.

Internal water molecules have crucial importance for activation of receptors in all GPCRs. It is known that intracellular end of domains and ligand binding cavity near the extracellular side show high water occupancy (hydration) and these two regions are disconnected because of the tight packing of hydrophobic residues in the middle of the TM regions. These hydrophobic residue clusters hinder the continuous “water wire” in inactive states of GPCRs [13, 47, 49]. Previous studies showed that the hydrolytic release of water molecules from intracellular region to the retinal by light activates the rhodopsin [50]. Hence, it may be novel to reveal this continuous water wire in transition conformations of M2 receptor in our case. Hollow [51] is a program that calculates the molecular surface of interior volume as dummy atoms. Thus, vestibules in the channels and/or geometric properties of the channel opening can be visualized and the information related to mechanism of proteins can be inferred [51]. In our case, we analyzed the ANM-LD predicted conformations on the activation pathway by using Hollow placing 1.3Å probes in the receptor to calculate the volume of cavities, hence to reveal the evolved channel inside the conformations through transition. For analysis, simulation 24 that gives the best results in terms of functionally important distances (Hydrogen bond formation between Y206 and Y440, broken of salt bridge between R121 and E382, TM3-TM6 distance between T386 and R121) among simulations from 3UON to 4MQS is selected. Figure 4.33.a and 4.33.h show the available channel in

initial and final states and Figure 4.33.b-g shows the channel formed in M2 receptor through the activation transition. Figure 4.34 and Figure 4.35 show the continuous water pathway through transition by superimposing cavities of initial, final and cycle 9 and cycle 16 transition conformations respectively. Hydrophobicity of middle of the TM regions decreases through middle of transition according to initial-inactive state, and two hydrated sites, which are ligand binding site and intracellular end of TM domains, have become almost connected during transition as it is seen in Figure 4.33,4.34 and Figure 4.35 as well. In the 9th cycle conformation, channel that linked two hydrated sites is formed closed to residues S64-C67, A109-V111, A146-L150, V199-W207, T386-A391 and T434-Y440. When functionally important distances' graph of Simulation 24 given in Figure 4.8 is examined, the major change in H-bond distance and TM3-TM6 distance are seen in Cycle 9 and Cycle 16 and interestingly continuous water pathway is formed in these cycles. Thus, it can be said that continuous water pathway formation may be crucial for activation of M2 receptor by playing a crucial role in functionally important changes. When functionally important residues are taken into account, which are Y206, T386, N436, P437 and Y440, this channel may be functionally important for activation of M2 receptor and thus the transition of ions for the inhibition of adenylyl cyclase activity.

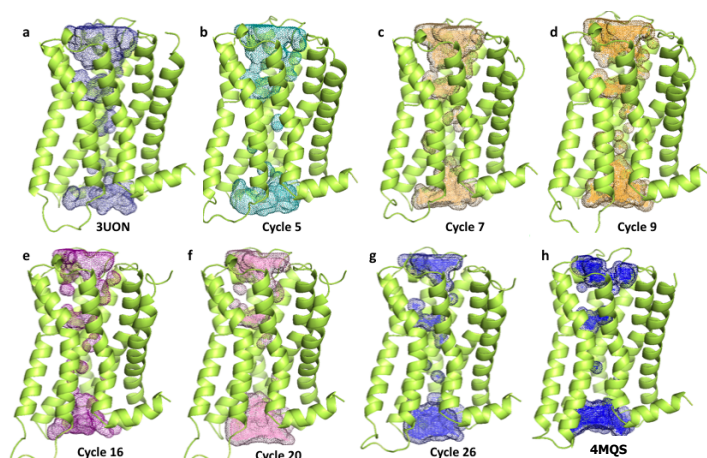


Figure 4.33. Calculated available channels in generated conformations, initial and final states by using Hollow.

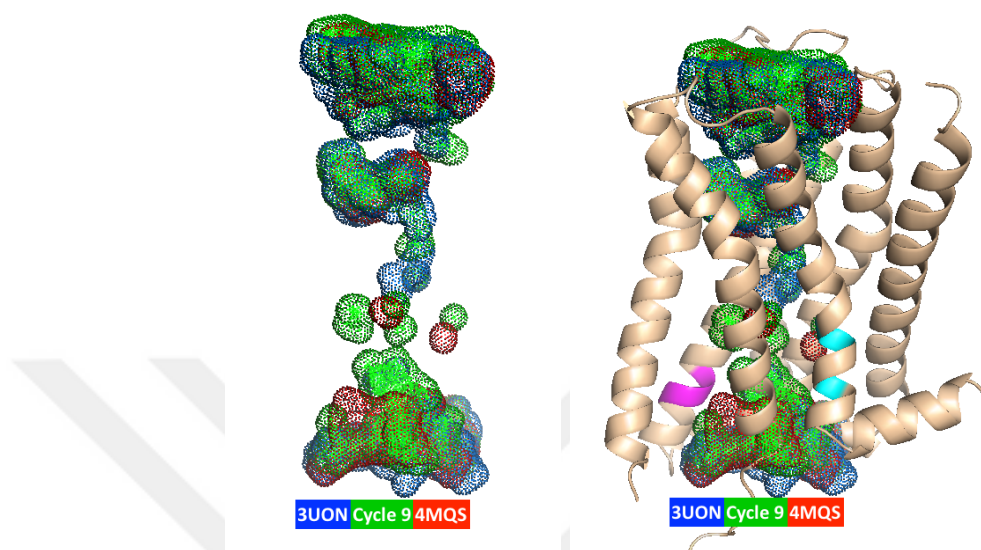


Figure 4.34. Superimposition of calculated available channels of 3UON, Cycle 9 and 4MQS to demonstrate the formed continuous water pathway. DRY and NPxxY motifs are colored with magenta and cyan respectively.

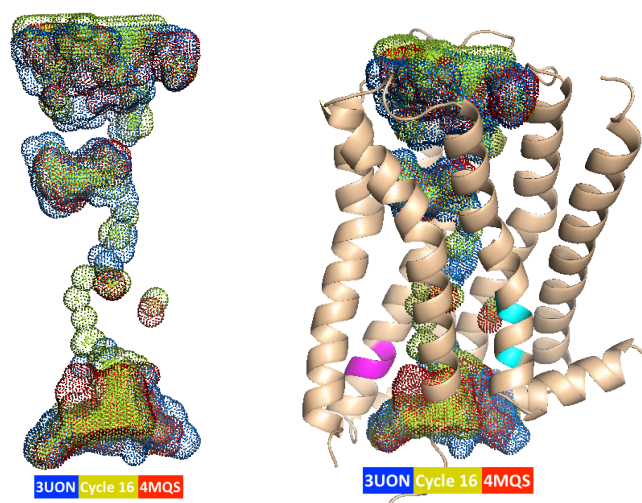


Figure 4.35. Superimposition of calculated available channels of 3UON, Cycle 16 and 4MQS to demonstrate the formed continuous water pathway. DRY and NPxxY motifs are colored with magenta and cyan respectively.

4.1.2. ANM-LD Simulation Results from Antagonist Bound Structure to Agonist and Allosteric Bound Structure

4MQT is the active state of M2 receptor that is bound to agonist iperoxo and allosteric modulator LY2119620. Allosteric modulators have such an importance of drug improvements through changing the protein allostery and dynamic with small ligands, so generating alternative pathways for the structure with the allosteric modulator may have an outstanding importance. Allosteric modulator and agonist bound structure show large similarity with only agonist bound structure (4MQS) except allosteric ligand region. Extracellular site of allosteric ligand region shows conformational change, especially residue W422 changes its conformation through vertical position that is shown in Figure 4.36 [41].

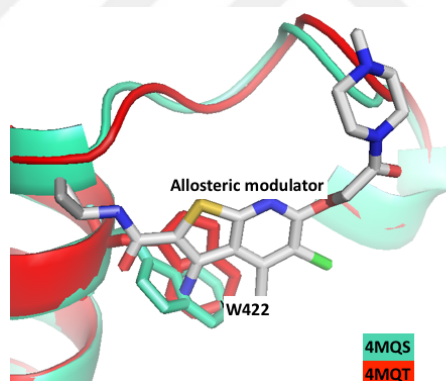


Figure 4.36. Conformation change of residue W422 in active structures of M2 receptor.

In this section of thesis, it is aimed to generate alternative transition pathways for activation of M2 receptor exploring whether allosteric modulator affects transition pathway utilizing different collaborative ANM modes or not. For this purpose, ANM-LD simulations were performed with using agonist iperoxo and allosteric modulator LY2119620 bound active state (4MQT) as a target structure; starting with the inactive state 3UON as initial structure is simulated by using ANM-LD methodology and results were compared with the simulations 4MQS as a target structure.

Parameters were selected as follow in simulations: Cut-off radius (R_{cut}) for ANM calculations is 13\AA , all modes are taken into account for transition and deformation factor (D_F) is adaptive between 0.1 and 1 and as they were in other M2 receptor activation simulations. The simulation results are summarized in Table 4.4 and detailed RMSD, selected mode, collectivity and overlap graphs are given in Figure 4.37. RMSD decreases from 2.4\AA to around 1.5\AA in 30 iterations. RMSD convergence is not so different than the simulations given using 4MQS as a target structure except Simulation 18.

Table 4.4. ANM-LD Simulation result summaries from 3UON to 4MQT.

| Simulation Number | Cycle Number | RMSD (\AA) | The Most Selected ANM Modes |
|-------------------|--------------|-----------------------|-----------------------------|
| 13 | 30 | 1.46 | 3 |
| 14 | 34 | 1.54 | 15, 3, 13, 17 |
| 18 | 55 | 1.01 | 8, 2, 10 |
| 70 | 20 | 1.51 | 2 |



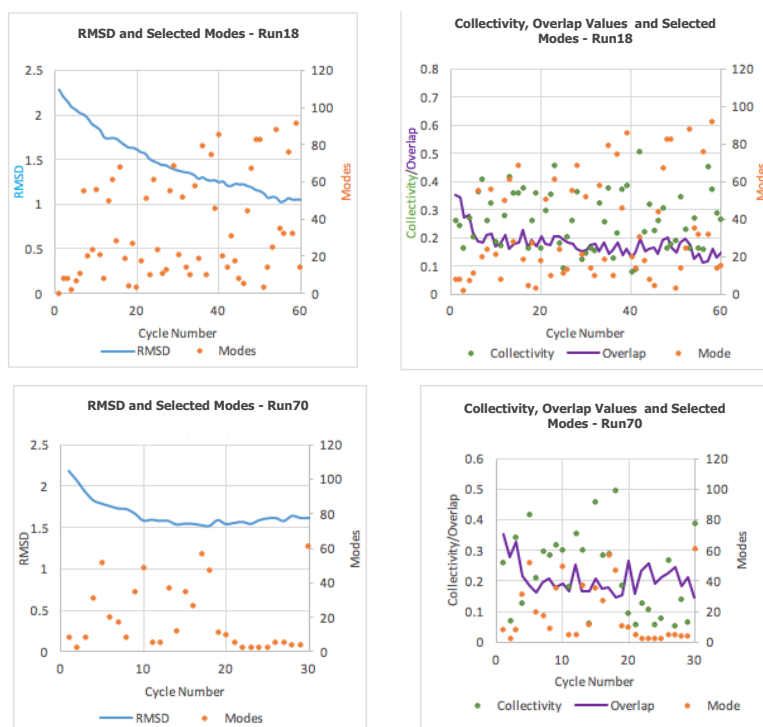
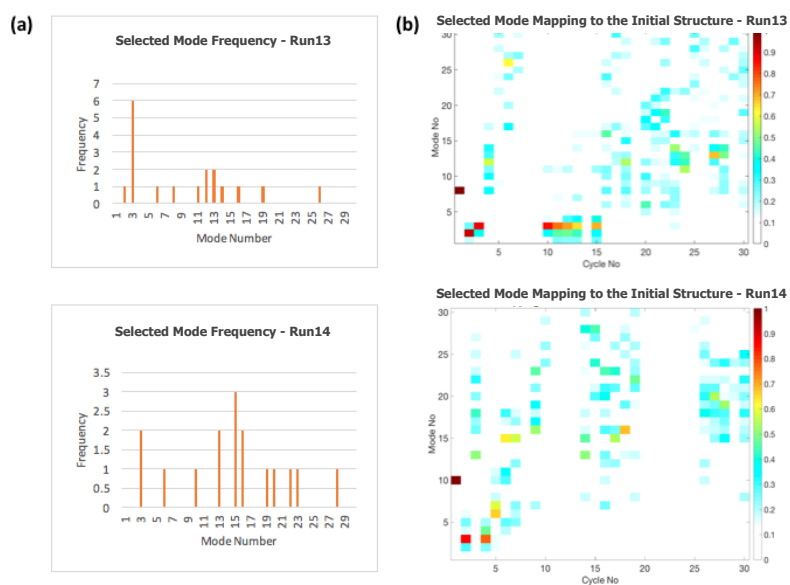


Figure 4.37. Selected mode, RMSD, overlap and collectivity values plots of simulations from 3UON to 4MQS. RMSD and selected modes plots are given in right side. Selected modes and corresponding collectivity and overlap values plots are given in left side.



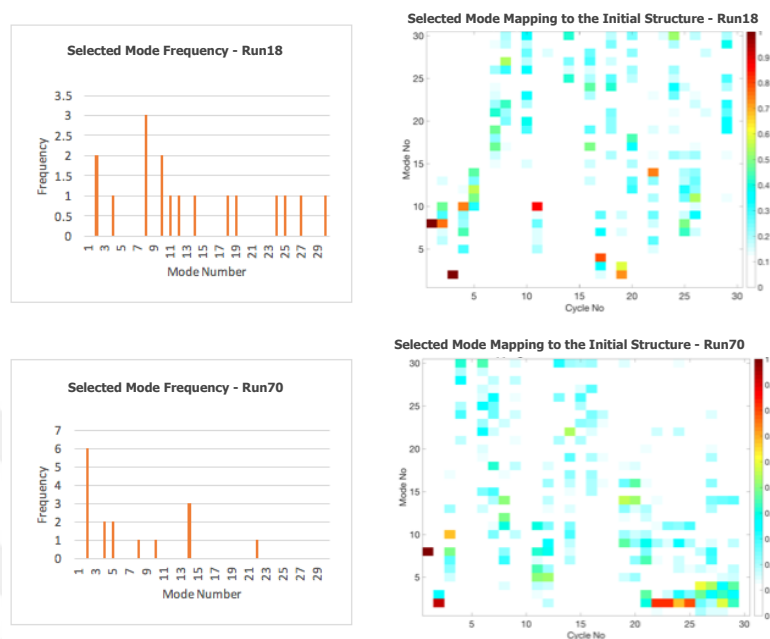


Figure 4.38. (a) Selected modes distributions. (b) Overlap values of selected modes of M2 simulations from 3UON to 4MQT according to the initial structure's first 30 modes.

Selected mode distribution graphs, given in Figure 4.38, show that 3^{rd} and 2^{nd} ANM modes are the most selected, common ANM modes in simulations from 3UON to 4MQT. The superimposed mode shape graphs of these modes and corresponding hinge residues are given in Figure 4.39. Hinges residues of 2^{nd} mode are L34-I38, Q55-N58, L70-G73, L103-V105, L115-D120, R135-A140, E175-S182, I200-S210, S380-L390, L428-S433 and hinge residues of 3^{rd} mode are F25-L28, T37-G40, Q55-L62, L103-S108, V149, L150, T170-F180, P198-I200, V385-A395, C443-K448. Hinge residues are located mostly at functionally important sites including DRY and NPxxY motifs, rearrangement of TM5 and outward tilting of TM6. Only 3^{rd} mode does not have a hinge residue in the middle of TM7, where TM7 is rearranged during transition, but other collective modes may contribute the motion in that direction.

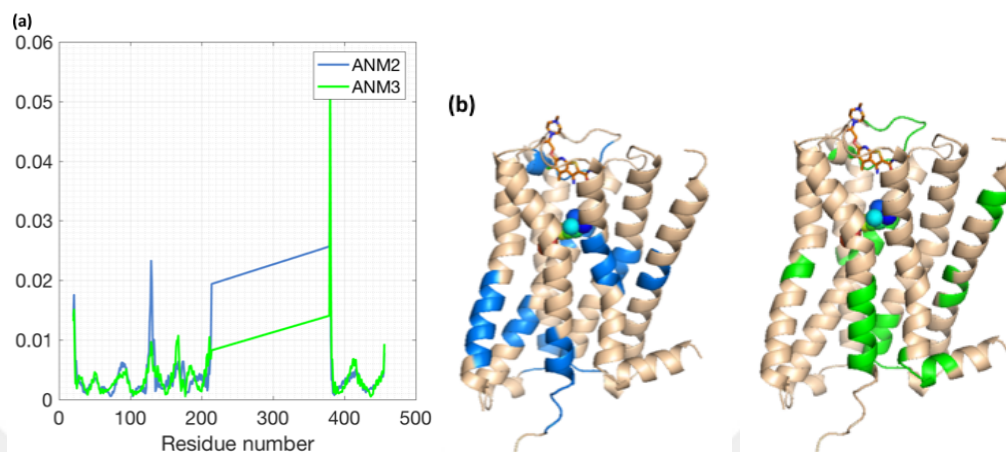


Figure 4.39. (a) Mostly selected ANM modes in simulations from 3UON to 4MQT. (b) Minimum residues correspond to mostly selected ANM modes 2 (blue) and 3 (green).

The structures of 4MQT and 4MQS are quite similar except missing 3 residues in ICL3 loop in 4MQT, there is only 0.99\AA RMSD difference between structures and major differences occur in N-terminus of TM1 region and ICL3 loop. Residues S215, P377 and P378 are missing in 4MQT structure. Different from the 4MQS, allosteric modulator in 4MQT interacts with the residues W422, N410, E172, N419, Y177, Y80, Y426. The simulations targeted to only agonist bound structure show that ANM modes 4 and 5 are quite important for the transition and restriction of these modes affect the transition. Due to similarity between two active states, it is expected to choose similar modes except differences in allosteric modulator binding region, N-terminus and ICL3 loop, to be able to say that certain collective modes are important for transition of M2 receptor. In order to find differences between these two structures more precisely, difference vectors between initial-inactive state (3UON) and target-active state (4MQS and 4MQT) were compared. In order to compare difference vectors, they were reduced to 1D by calculating their dot products, hence their difference vector shapes were obtained and compared in Figure 4.40.a. Although difference vector shapes are quite similar, residues between T20-A30, L70-Y80, S380-W400 show small differences. These different motions could be directed with selecting other modes and/or having slight differences in mode shapes. When mostly selected modes of simulations are

compared, 4th ANM mode of 4MQS overlaps with 2nd mode of 4MQT and 5th mode of 4MQS overlaps with 3rd mode of 4MQT with very slight differences given in Figure 25.b and 25.c. Mode shape graphs show that ANM modes are shifted, because of deleting missing three residues from initial structure. 4th mode replaced with the 2nd mode and 5th mode replaced with the 3rd mode. The similarity of selected modes in transition verifies the importance of mostly selected modes for activation of M2 receptor. The slight differences in target structures affected the transition in a manner of slight changes in modes including mode shifts. Also other mostly selected modes (2, 6) in simulations from 3UON to 4MQS were not selected in simulations from 3UON to 4MQT. Not utilizing these modes and slight changes in selected modes in transition may cause change in cross-correlation maps and affect the success of simulations in terms of functionally important distances.

When distance between functionally important residues, which are H-bond formation between Y206 located at ICL3 end of TM5 and Y440 located in TM7, breakage of salt bridge between R121 located at ICL2 end of TM3 and E382 located at ICL3 and the distance between residue R121 in TM3 and T386 in TM6 due to outward tilt of TM6, and RMSD between selected residues (A184, T388, N410, A438) of transition and target structure (Figure 4.41) were examined, it was observed that rather than the simulations targeted to agonist and modulator bound structure, the simulations targeted to only agonist bound structure approach to the target structure more successfully. This may be because of only selecting 3rd or 2nd mode individually in simulations. Different than the 4MQS structure, allosteric modulator in 4MQT structure changes the position of residue W422, where allosteric modulator is bound. W422 changes its conformation along vertical position. However, the RMSD difference of this residue between initial and target structure is quite small, it is difficult to measure its position through pathway.

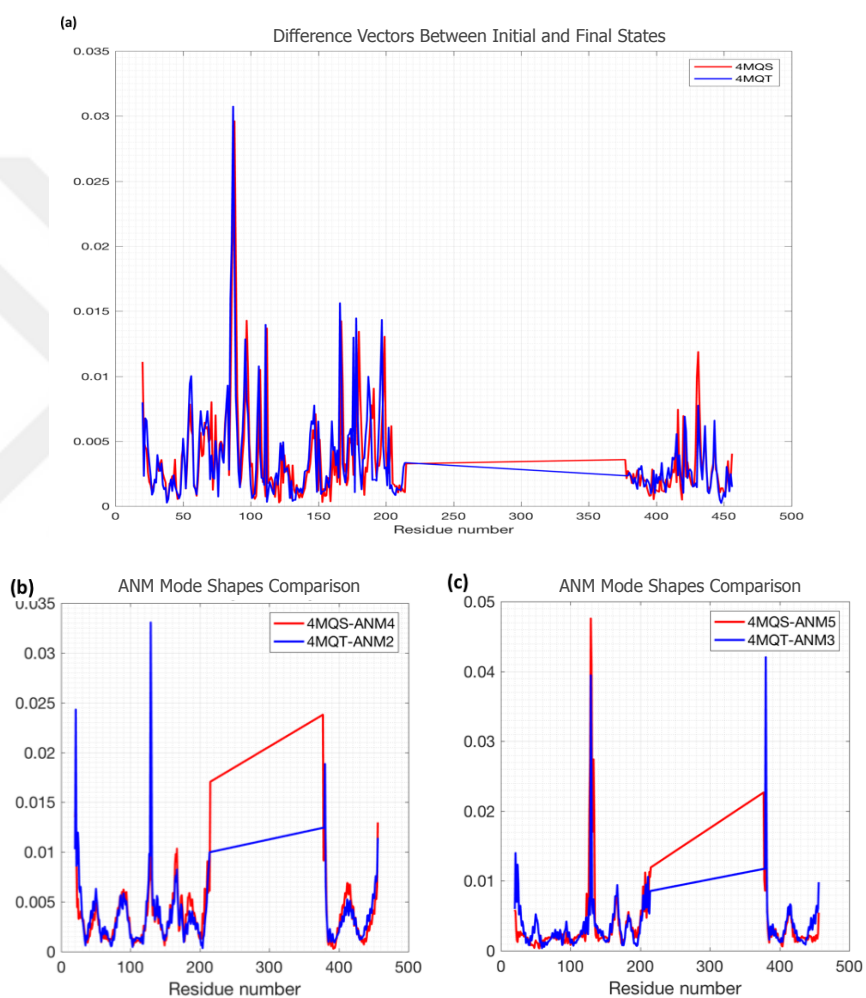


Figure 4.40. (a) Difference vectors between initial and final states. (b) Superimposition of 2nd ANM Mode of 4MQT to 4th ANM Mode of 4MQS. (c) Superimposition of 3rd ANM Mode of 4MQT to 5th ANM Mode of 4MQS.

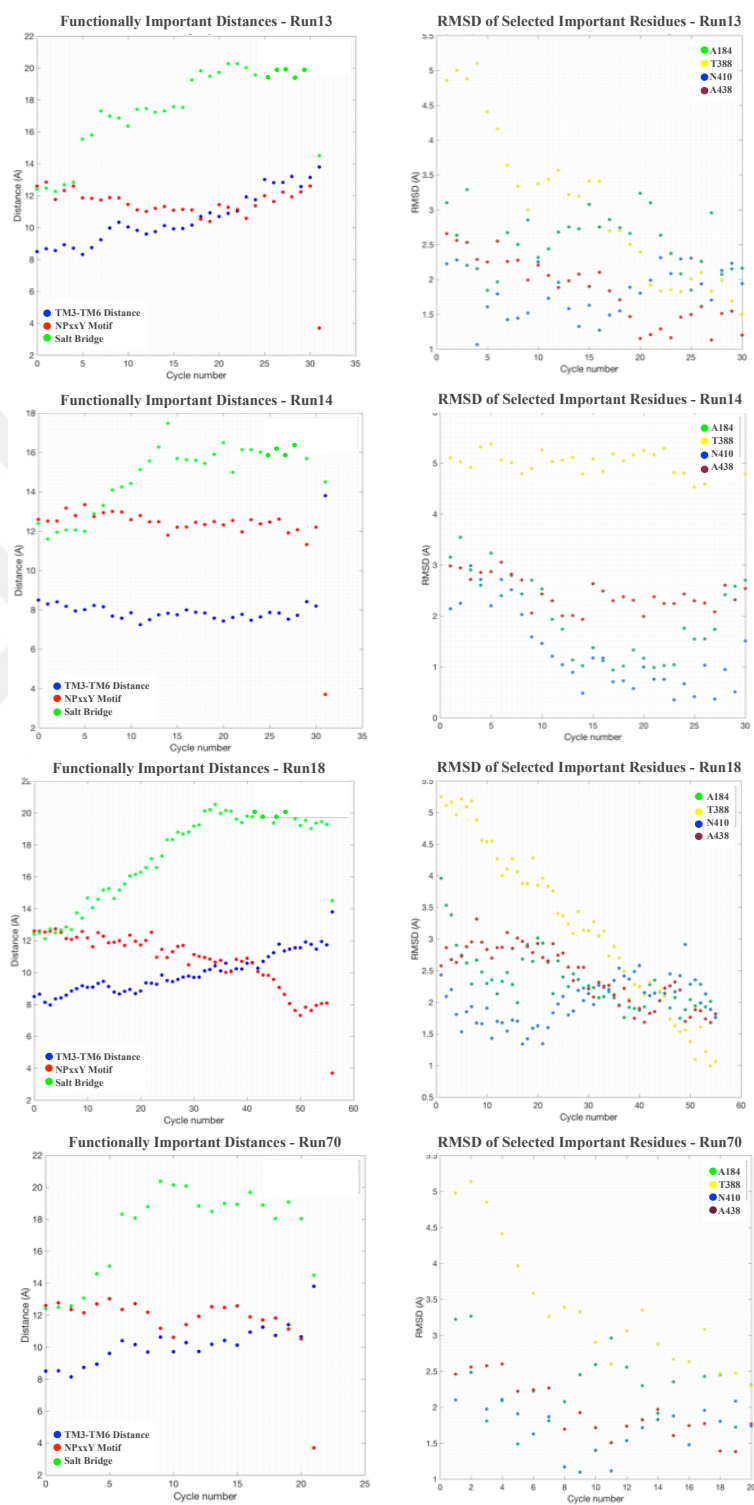


Figure 4.41. Functionally important distances through transition pathway (Left side) and RMSD of selected residues between transition conformations and target conformation (Right side) in M2 simulations from 3UON to 4MQS.

5. CONCLUSION AND RECOMMENDATIONS

In this thesis, the major aim is to explore activation mechanism of muscarinic M2 receptor in the light of generated transition pathways between initial-inactive state (3UON) and target-active states (4MQS, 4MQT) by implementing ANM-LD methodology. The conformational transition pathway(s) coupled to the activation with mostly selected intrinsic dynamic modes and the key sites that coordinate the motion(s) described by these modes of motion with respect to the functionally known important distances decoded as underlying the activation are the points of the focus in the analysis.

Activation of M2 receptor requires specific conformational changes through transition. Major changes are formation of Hydrogen bond between Y206 located at ICL3 end of TM5 and Y440 located in TM7, broken of salt bridge between R121 located at ICL 2 end of TM3 and E382 located at ICL3 and the distance between residue R121 in TM3 and T386 in TM6 due to outward tilt of TM6. Hydrogen bond formation between Y206 and Y440 is related to NPxxY (N436, P437, Y440) motif, which is highly conserved motif in all GPCRs. Another important common motif of GPCRs is DRY (D120, R121, Y122) motif. In DRY motif region; N58 forms H-bond with R121 and D120, hence these residues are stabilized in active state. Formation of tyrosine lid among Y104, Y403 and Y426 as a result of H-bonding of Y104 with Y403 and Y426 is another important change through transition. Structurally major displacements take place ICL3 loop end of TM6, in the middle of TM7 and TM5. [6,8].

The ANM-LD simulations of M2 receptor activation by using initial-inactive (3UON) and target-active (4MQS) states were performed first without restricting any certain ANM mode and modification on structures. After analyzing mostly selected modes, hinge residues and functionally important distances, simulations were repeated with making some modifications on structures to normalize unrealistic picks in ANM modes and then restricting ANM 4th and/or 5th modes. Finally, The ANM-LD simulations were performed by using target-active state 4MQT (agonist and allosteric

modulator bound) instead of 4MQS (only agonist bound). M2 receptor activation simulations show that 4th and 5th modes are important to direct the motion as target state. Hinge residues of these modes overlap with the functionally important residues. Hinge residues of 5th mode include DRY (D120, R121, Y122) motif and N58 residue, where D120 form H-bond with residue N58, NPxxY (N436, P437, Y440) motif and residues W22 and T386 and hinge residues of 4th mode include middle of TM7 rearrangement (residues between N431 and N435) and TM5 displacement (residues between I201 and L205). Hence, it is important to note that ANM could reveal the functionally important conformational changes successfully by using intrinsic dynamics of modes and there is a strong relationship between hinge residues of certain (mostly selected) ANM modes and functionally important residues. In addition to mostly selected modes overlapping with the functionally important switches, these switches and hinge residues of selected modes also align with the hydrophobic layers in TM domains and the motion targeted by these selected modes could provide the formation of continuous water pathway during the transition, which is also crucial for activation of M2 receptor.

In the analysis, some major side chain rotations and side chain bond formations like positional change of W422 and hydrogen bond formation between side chain of Y206 and Y440 could not reach the target value, although they show the trend in towards target conformation. This is because ANM-LD methodology is based on the backbone calculation and Side chains are rearranged during the all-atom Langevin Dynamic simulation, but this may be not enough to rotate side chains precisely. In future, ANM-LD methodology can be improved to include side chains rotations as well.

This study can be enhanced through other GPCRs, because all GPCRs have similar structures especially in their inactive states. Mostly selected ANM modes of other GPCRs through activation pathways can be compared with the mostly selected ANM modes of M2 receptor activation pathway to demonstrate whether there is a selective binding in terms of ANM modes or not.

REFERENCES

1. Das, A., M. Gur, M. H. Cheng, S. Jo, I. Bahar and B. Roux, “Exploring the conformational transitions of biomolecular systems using a simple two-state anisotropic network model”, *PLoS computational biology*, Vol. 10, No. 4, p. e1003521, 2014.
2. Karplus, M. and J. A. McCammon, “Molecular dynamics simulations of biomolecules”, *Nature Structural and Molecular Biology*, Vol. 9, No. 9, p. 646, 2002.
3. Bernardi, R. C., M. C. Melo and K. Schulten, “Enhanced sampling techniques in molecular dynamics simulations of biological systems”, *Biochimica et Biophysica Acta (BBA)-General Subjects*, Vol. 1850, No. 5, pp. 872–877, 2015.
4. Bahar, I. and A. Rader, “Coarse-grained normal mode analysis in structural biology”, *Current opinion in structural biology*, Vol. 15, No. 5, pp. 586–592, 2005.
5. Haliloglu, T. and I. Bahar, “Adaptability of protein structures to enable functional interactions and evolutionary implications”, *Current opinion in structural biology*, Vol. 35, pp. 17–23, 2015.
6. Bahar, I., T. R. Lezon, A. Bakan and I. H. Shrivastava, “Normal mode analysis of biomolecular structures: functional mechanisms of membrane proteins”, *Chemical reviews*, Vol. 110, No. 3, pp. 1463–1497, 2009.
7. Haliloglu, T. and I. Bahar, “Structure-based analysis of protein dynamics: Comparison of theoretical results for hen lysozyme with X-ray diffraction and NMR relaxation data”, *Proteins: Structure, Function, and Bioinformatics*, Vol. 37, No. 4, pp. 654–667, 1999.
8. Haliloglu, T., I. Bahar and B. Erman, “Gaussian dynamics of folded proteins”, *Physical review letters*, Vol. 79, No. 16, p. 3090, 1997.

9. Tozzini, V., “Coarse-grained models for proteins”, *Current opinion in structural biology*, Vol. 15, No. 2, pp. 144–150, 2005.
10. Atilgan, A. R., S. Durell, R. L. Jernigan, M. Demirel, O. Keskin and I. Bahar, “Anisotropy of fluctuation dynamics of proteins with an elastic network model”, *Biophysical journal*, Vol. 80, No. 1, pp. 505–515, 2001.
11. Fas, A. B. A., *Computational and experimental investigations into the mechanism of functional motion in biomolecular systems*, Ph.D. Thesis, Boğaziçi University, İstanbul, Turkey, 2016.
12. Latorraca, N. R., A. Venkatakrisnan and R. O. Dror, “GPCR dynamics: structures in motion”, *Chemical reviews*, Vol. 117, No. 1, pp. 139–155, 2016.
13. Bhattacharya, S., R. Salomon-Ferrer, S. Lee and N. Vaidehi, “Conserved mechanism of conformational stability and dynamics in G-protein-coupled receptors”, *Journal of chemical theory and computation*, Vol. 12, No. 11, pp. 5575–5584, 2016.
14. Schiöth, H. B. and R. Fredriksson, “The GRAFS classification system of G-protein coupled receptors in comparative perspective”, *General and comparative endocrinology*, Vol. 142, No. 1-2, pp. 94–101, 2005.
15. Hu, G.-M., T.-L. Mai and C.-M. Chen, “Visualizing the GPCR Network: Classification and Evolution”, *Scientific reports*, Vol. 7, No. 1, p. 15495, 2017.
16. Trzaskowski, B., D. Latek, S. Yuan, U. Ghoshdastider, A. Debinski and S. Filipek, “Action of molecular switches in GPCRs-theoretical and experimental studies”, *Current medicinal chemistry*, Vol. 19, No. 8, pp. 1090–1109, 2012.
17. Geppetti, P., N. A. Veldhuis, T. Lieu and N. W. Bunnett, “G protein-coupled receptors: dynamic machines for signaling pain and itch”, *Neuron*, Vol. 88, No. 4, pp. 635–649, 2015.

18. Thal, D. M., B. Sun, D. Feng, V. Nawaratne, K. Leach, C. C. Felder, M. G. Bures, D. A. Evans, W. I. Weis, P. Bachhawat *et al.*, “Crystal structures of the M1 and M4 muscarinic acetylcholine receptors”, *Nature*, Vol. 531, No. 7594, p. 335, 2016.
19. Eglen, R. M., “Muscarinic receptor subtype pharmacology and physiology”, *Progress in medicinal chemistry*, Vol. 43, pp. 105–136, 2005.
20. Brown, D. A., “Muscarinic acetylcholine receptors (mAChRs) in the nervous system: some functions and mechanisms”, *Journal of molecular neuroscience*, Vol. 41, No. 3, pp. 340–346, 2010.
21. Uustare, A., J. Näsman, K. E. Åkerman and A. Rincken, “Characterization of M2 muscarinic receptor activation of different G protein subtypes”, *Neurochemistry international*, Vol. 44, No. 2, pp. 119–124, 2004.
22. Hirshman, C. A., B. Lande and T. L. Croxton, “Role of M2 muscarinic receptors in airway smooth muscle contraction”, *Life sciences*, Vol. 64, No. 6-7, pp. 443–448, 1999.
23. Migeon, J. C., S. L. Thomas and N. M. Nathanson, “Differential coupling of m2 and m4 muscarinic receptors to inhibition of adenylyl cyclase by $G_{i\alpha}$ and $G_{o\alpha}$ subunits”, *Journal of Biological Chemistry*, Vol. 270, No. 27, pp. 16070–16074, 1995.
24. Moreau, C. J., J. Revilloud, L. N. Caro, J. P. Dupuis, A. Trouchet, A. Estrada-Mondragón, K. Nieścierowicz, N. Sapay, S. Crouzy and M. Vivaudou, “Tuning the allosteric regulation of artificial muscarinic and dopaminergic ligand-gated potassium channels by protein engineering of G protein-coupled receptors”, *Scientific reports*, Vol. 7, p. 41154, 2017.
25. Ye, L., C. Neale, A. Sljoka, B. Lyda, D. Pichugin, N. Tsuchimura, S. T. Larda, R. Pomès, A. E. García, O. P. Ernst *et al.*, “Mechanistic insights into allosteric regulation of the A_{2A} adenosine G protein-coupled receptor by physiological cations”,

- Nature communications*, Vol. 9, No. 1, p. 1372, 2018.
26. Ponzoni, L., G. Rossetti, L. Maggi, A. Giorgetti, P. Carloni and C. Micheletti, “Unifying view of mechanical and functional hotspots across class A GPCRs”, *PLoS computational biology*, Vol. 13, No. 2, p. e1005381, 2017.
 27. Gumbart, J., F. Khalili-Araghi, M. Sotomayor and B. Roux, “Constant electric field simulations of the membrane potential illustrated with simple systems”, *Biochimica et Biophysica Acta (BBA)-Biomembranes*, Vol. 1818, No. 2, pp. 294–302, 2012.
 28. Zheng, W., B. R. Brooks and G. Hummer, “Protein conformational transitions explored by mixed elastic network models”, *Proteins: Structure, Function, and Bioinformatics*, Vol. 69, No. 1, pp. 43–57, 2007.
 29. Ma, J., “Usefulness and limitations of normal mode analysis in modeling dynamics of biomolecular complexes”, *Structure*, Vol. 13, No. 3, pp. 373–380, 2005.
 30. Doruker, P., A. R. Atilgan and I. Bahar, “Dynamics of proteins predicted by molecular dynamics simulations and analytical approaches: Application to α -amylase inhibitor”, *Proteins: Structure, Function, and Bioinformatics*, Vol. 40, No. 3, pp. 512–524, 2000.
 31. Bahar, I., A. R. Atilgan and B. Erman, “Direct evaluation of thermal fluctuations in proteins using a single-parameter harmonic potential”, *Folding and Design*, Vol. 2, No. 3, pp. 173–181, 1997.
 32. Kuriyan, J., G. A. Petsko, R. M. Levy and M. Karplus, “Effect of anisotropy and anharmonicity on protein crystallographic refinement: an evaluation by molecular dynamics”, *Journal of molecular biology*, Vol. 190, No. 2, pp. 227–254, 1986.
 33. Karplus, M. and G. A. Petsko, “Molecular dynamics simulations in biology”, *Nature*, Vol. 347, No. 6294, p. 631, 1990.

34. Salsbury Jr, F. R., “Molecular dynamics simulations of protein dynamics and their relevance to drug discovery”, *Current opinion in pharmacology*, Vol. 10, No. 6, pp. 738–744, 2010.
35. Brünger, A., C. L. Brooks III and M. Karplus, “Stochastic boundary conditions for molecular dynamics simulations of ST2 water”, *Chemical physics letters*, Vol. 105, No. 5, pp. 495–500, 1984.
36. Adcock, S. A. and J. A. McCammon, “Molecular dynamics: survey of methods for simulating the activity of proteins”, *Chemical reviews*, Vol. 106, No. 5, pp. 1589–1615, 2006.
37. Duan, Y., C. Wu, S. Chowdhury, M. C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee *et al.*, “A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations”, *Journal of computational chemistry*, Vol. 24, No. 16, pp. 1999–2012, 2003.
38. Case, D., T. Darden, T. Cheatham III, C. Simmerling, J. Wang, R. Duke, R. Luo, R. Walker, W. Zhang, K. Merz *et al.*, “AMBER 11,(2010) University of California”, *San Francisco*.
39. Dolinsky, T. J., P. Czodrowski, H. Li, J. E. Nielsen, J. H. Jensen, G. Klebe and N. A. Baker, “PDB2PQR: expanding and upgrading automated preparation of biomolecular structures for molecular simulations”, *Nucleic acids research*, Vol. 35, No. suppl.2, pp. W522–W525, 2007.
40. Kruse, A. C., A. M. Ring, A. Manglik, J. Hu, K. Hu, K. Eitel, H. Hübner, E. Pardon, C. Valant, P. M. Sexton *et al.*, “Activation and allosteric modulation of a muscarinic acetylcholine receptor”, *Nature*, Vol. 504, No. 7478, p. 101, 2013.
41. Haga, K., A. C. Kruse, H. Asada, T. Yurugi-Kobayashi, M. Shiroishi, C. Zhang, W. I. Weis, T. Okada, B. K. Kobilka, T. Haga *et al.*, “Structure of the human M2

- muscarinic acetylcholine receptor bound to an antagonist”, *Nature*, Vol. 482, No. 7386, p. 547, 2012.
42. Kow, R. L. and N. M. Nathanson, “Structural biology: Muscarinic receptors become crystal clear”, *Nature*, Vol. 482, No. 7386, p. 480, 2012.
43. Miao, Y., S. E. Nichols, P. M. Gasper, V. T. Metzger and J. A. McCammon, “Activation and dynamic network of the M2 muscarinic receptor”, *Proceedings of the National Academy of Sciences*, Vol. 110, No. 27, pp. 10982–10987, 2013.
44. Fiser, A. and A. Sali, “ModLoop: automated modeling of loops in protein structures”, *Bioinformatics*, Vol. 19, No. 18, pp. 2500–2501, 2003.
45. Fiser, A., R. K. G. Do *et al.*, “Modeling of loops in protein structures”, *Protein science*, Vol. 9, No. 9, pp. 1753–1773, 2000.
46. DeLano, W. L., “The PyMOL Molecular Graphics System”, *DeLano Scientific*, Palo Alto, CA, USA, 2002.
47. Yuan, S., S. Filipek, K. Palczewski and H. Vogel, “Activation of G-protein-coupled receptors correlates with the formation of a continuous internal water pathway”, *Nature communications*, Vol. 5, p. 4733, 2014.
48. Eyal, E., G. Lum and I. Bahar, “The anisotropic network model web server at 2015 (ANM 2.0)”, *Bioinformatics*, Vol. 31, No. 9, pp. 1487–1489, 2015.
49. Yuan, S., H. Vogel and S. Filipek, “The Role of Water and Sodium Ions in the Activation of the μ -Opioid Receptor”, *Angewandte Chemie International Edition*, Vol. 52, No. 38, pp. 10112–10115, 2013.
50. Jastrzebska, B., K. Palczewski and M. Golczak, “Role of bulk water in hydrolysis of the rhodopsin chromophore”, *Journal of Biological Chemistry*, Vol. 286, No. 21, pp. 18930–18937, 2011.

51. Ho, B. K. and F. Gruswitz, "HOLLOW: generating accurate representations of channel and interior surfaces in molecular structures", *BMC structural biology*, Vol. 8, No. 1, p. 49, 2008.



APPENDIX A: ANM-LD SIMULATION RESULT SUMMARIES

Table A.1. ANM-LD Simulation result summaries of M2 Receptor activation pathway
(Simulation No:1-30).

| Run No | States | Modification | Restriction | DF | Modemax | Rcut | Cycle No | RMSD |
|--------|--------------|--------------|-------------|-------|-----------|------|----------|------|
| 1 | 3uon to 4mqs | | | 0.4 | 30 | 13 | 70 | 1.7 |
| 4 | 3uon to 4mqs | | | 0.35 | 100 | 18 | 70 | 1.6 |
| 5 | 3uon to 4mqs | | | 0.4 | 100 | 18 | 70 | 1.6 |
| 6 | 3uon to 4mqs | | | 0.35 | 100 | 13 | 70 | 1.2 |
| 7 | 3uon to 4mqs | | | 0.1-1 | all modes | 13 | 35 | 1.54 |
| 8 | 3uon to 4mqs | | 5 | 0.35 | 30 | 13 | 150 | 1.5 |
| 9 | 3uon to 4mqs | | 4 | 0.35 | 30 | 13 | 100 | 1.5 |
| 10 | 3uon to 4mqs | | 26 | 0.35 | 30 | 105 | 70 | 1.5 |
| 11 | 3uon to 4mqs | | 1 | 0.35 | 30 | 13 | 80 | 1.72 |
| 12 | 3uon to 4mqs | | 12 | 0.35 | 30 | 13 | 90 | 1.6 |
| 13 | 3uon to 4mqt | | | 0.1-1 | all modes | 13 | 30 | 1.46 |
| 14 | 3uon to 4mqt | | | 0.1-1 | all modes | 18 | 34 | 1.54 |
| 15 | 3uon to 4mqt | | | 0.4 | 100 | 13 | 100 | 1 |
| 16 | 3uon to 4mqt | | | 0.6 | 30 | 18 | 50 | 2 |
| 17 | 3uon to 4mqt | | | 0.6 | 30 | 13 | 80 | 1.3 |
| 18 | 3uon to 4mqt | | | 0.35 | 100 | 13 | 55 | 1.01 |
| 19 | 3uon to 4mqt | | | 0.4 | 100 | 10 | 50 | 1.15 |
| 20 | 3uon to 4mqt | | | 0.35 | 30 | 10 | 65 | 1.2 |
| 21 | 3uon to 4mqs | | 4 | 0.6 | 100 | 13 | 50 | 1.3 |
| 22 | 3uon to 4mqs | | 5 | 0.6 | 100 | 13 | 32 | 1.3 |
| 23 | 3uon to 4mqs | | | 0.5 | 100 | 13 | 40 | 1.21 |
| 24 | 3uon to 4mqs | | | 0.1-1 | all modes | 13 | 30 | 1.46 |
| 25 | 3uon to 4mqs | | 26 | 0.6 | 100 | 13 | 50 | 1.25 |
| 26 | 3uon to 4mqs | | | 0.6 | 100 | 13 | 60 | 1.3 |
| 27 | 3uon to 4mqs | | 4 | 0.1-1 | all modes | 13 | 80 | 1.15 |
| 28 | 3uon to 4mqs | | 5 | 0.1-1 | all modes | 13 | 62 | 1.25 |
| 29 | 3uon to 4mqs | | 26 | 0.1-1 | all modes | 13 | 65 | 1.27 |
| 30 | 3uon to 4mqs | | | 0.1-1 | all modes | 13 | 30 | 1.49 |

Table A.2. ANM-LD Simulation result summaries of M2 Receptor activation pathway (Simulation No:31-70).

| Run No | States | Modification | Restriction | DF | Modemax | Rcut | Cycle No | RMSD |
|--------|--------------|---------------------------------|-------------|---------|-----------|------|----------|------|
| 31 | 3uon to 4mqs | | 1 | 0.6 | 100 | 13 | 70 | 1.28 |
| 32 | 3uon to 4mqs | | 1 | 0.1-1 | all modes | 13 | 45 | 1.5 |
| 33 | 3uon to 4mqs | | | 0.1-1 | all modes | 13 | 35 | 1.6 |
| 34 | 3uon to 4mqs | | 19 | 0.6 | 100 | 13 | 85 | 1.14 |
| 35 | 3uon to 4mqs | | 19 | 0.35 | 30 | 13 | 95 | 1.5 |
| 36 | 3uon to 4mqs | | 19 | 0.1-1 | all modes | 13 | 100 | 1.48 |
| 37 | 3uon to 4mqs | | 19 | 0.6 | 100 | 13 | 70 | 1.3 |
| 38 | 3uon to 4mqs | | 2 | 0.35 | 30 | 13 | 70 | 1.6 |
| 39 | 3uon to 4mqs | | 2 | 0.1-1 | all modes | 13 | 70 | 1.46 |
| 40 | 3uon to 4mqs | | 2, 5 | 0.6 | 100 | 13 | 70 | 1.2 |
| 41 | 3uon to 4mqs | | 1, 2, 4 | 0.1-1 | all modes | 13 | 65 | 1.45 |
| 42 | 3uon to 4mqs | | 1, 2, 4 | 0.1-1 | all modes | 13 | 95 | 1.48 |
| 43 | 3uon to 4mqs | | 1, 2, 4 | 0.1-1 | all modes | 13 | 60 | 1.7 |
| 44 | 3uon to 4mqs | 23 res. modeled | | 0.1-1 | all modes | 13 | 60 | 1.7 |
| 45 | 3uon to 4mqs | 23 res. modeled | | 0.1-1 | all modes | 13 | 80 | 1.7 |
| 46 | 3uon to 4mqs | 23 res. modeled | | 0.6 | 100 | 13 | 40 | 1.7 |
| 47 | 3uon to 4mqs | | 6 | 0.1-1 | all modes | 13 | 55 | 1.5 |
| 48 | 3uon to 4mqs | | 6 | 0.1-1 | all modes | 13 | 40 | 1.5 |
| 49 | 3uon to 4mqs | | 6 | 0.6 | 100 | 13 | 40 | 1.3 |
| 50 | 3uon to 4mqs | 3 res. modeled | | 0.6 | 100 | 13 | 60 | 1.35 |
| 51 | 3uon to 4mqs | | 1, 2, 4, 6 | 0.6 | 100 | 13 | 40 | 1.5 |
| 52 | 3uon to 4mqs | | 1, 2, 4, 6 | 0.1-1 | all modes | 13 | 50 | 1.5 |
| 53 | 3uon to 4mqs | 3 res. modeled | | 0.1-1 | all modes | 13 | 25 | 1.56 |
| 54 | 3uon to 4mqs | 3 res. modeled | | 0.1-1 | all modes | 13 | 45 | 1.4 |
| 55 | 3uon to 4mqs | 3 res. modeled | | 0.1-1 | all modes | 13 | 50 | 1.5 |
| 56 | 3uon to 4mqs | Ter. deleted and 3 res. modeled | | 0.1-1 | all modes | 13 | 50 | 1.55 |
| 57 | 3uon to 4mqs | Ter. deleted and 3 res. modeled | | 0.1-1 | all modes | 13 | 40 | 1.44 |
| 58 | 3uon to 4mqs | Ter. deleted | | 0.1-1 | all modes | 13 | 50 | 1.5 |
| 59 | 3uon to 4mqs | Ter. deleted | | 0.1-1 | all modes | 13 | 35 | 1.5 |
| 60 | 3uon to 4mqs | Ter. deleted | | 0.1-1 | all modes | 13 | 60 | 1.55 |
| 61 | 3uon to 4mqs | Ter command in ICL3 | | 0.1-1 | all modes | 13 | 70 | 1.4 |
| 62 | 3uon to 4mqs | Ter command in ICL3 | | 0.1-1 | all modes | 13 | 60 | 1.52 |
| 63 | 3uon to 4mqs | Ter command in ICL3 | | 0.6 | 100 | 13 | 50 | 1.2 |
| 64 | 3uon to 4mqs | | 4 | 0.1-1 | all modes | 13 | 52 | 1.52 |
| 65 | 3uon to 4mqs | | 2, 5 | 0.1-1 | all modes | 13 | 20 | 1.56 |
| 66 | 3uon to 4mqs | | | 0.1-0.4 | all modes | 13 | 60 | 1.3 |
| 67 | 3uon to 4mqs | | 4, 5 | 0.1-1 | all modes | 13 | 80 | 1.7 |
| 69 | 3uon to 4mqs | | | 0.1-0.4 | all modes | 13 | 100 | 1.1 |
| 70 | 3uon to 4mqs | | | 0.1-1 | all modes | 13 | 20 | 1.51 |