KADIR HAS UNIVERSITY
GRADUATE SCHOOL OF SCIENCE AND ENGINEERING

LOOP MODELING AND MOLECULAR DYNAMICS SIMULATIONS OF APO AND LIGAND-BOUND HUMAN GLUN1-GLUN2A NMDA TYPE RECEPTORS

MUHAMMED AKTOLUN

September 2017

LOOP MODELING AND MOLECULAR DYNAMICS SIMULATIONS OF APO AND LIGAND-BOUND HUMAN GLUN1-GLUN2A NMDA TYPE RECEPTORS

MUHAMMED AKTOLUN

Submitted to the Graduate School of Science and Engineering
in partial fulfillment of the requirements for the degree of
Master of Science
in
Computational Biology and Bioinformatics

KADIR HAS UNIVERSITY

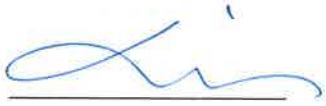
September 2017

KADIR HAS UNIVERSITY
GRADUATE SCHOOL OF SCIENCE AND ENGINEERING

LOOP MODELING AND MOLECULAR DYNAMICS SIMULATIONS OF APO AND LIGAND-BOUND HUMAN GLUN1-GLUN2A NMDA TYPE RECEPTORS

MUHAMMED AKTOLUN

APPROVED BY:

Asst. Prof. Dr. Şebnem Eşsiz (Thesis Supervisor)	(Kadir Has University)	
Prof. Dr. Mine Yurtsever	(Istanbul Technical University)	
Asst. Prof. Dr. Bahar Delibaş	(Kadir Has University)	

DATE OF APPROVAL: 15.09.2017

“I, Muhammed AKTOLUN, hereby declare that this master’s thesis is my own original work and that due references have been appropriately provided on all supporting literature and resources.”

MUHAMMED AKTOLUN

Istanbul, 2017

ABSTRACT

N-Methyl-D-Aspartate receptors (NMDARs) are glutamate-gated ion channels found in the nerve cell membranes. The functioning of the receptor is of crucial importance in consciousness and normal brain functions. As a result of overexcitation of NMDARs, neuronal death occurs and may lead to diseases such as epilepsy, stroke, Alzheimer's and Parkinson's. Understanding the molecular mechanism and structure function relationships of the receptor might lead to discovery of new drug target mechanisms.

Recently, there are two intact X-ray structures available, one is from *Xenopus laevis* and the other one is from *Rattus norvegicus*, for GluN1-GluN2B type NMDA receptor. First, both X-ray structures are examined and compared, for the ion channel especially, by taking the general problems into consideration which arise from crystallization conditions. Human GluN1-GluN2A type NMDAR structure is modeled based on the structure of *Xenopus laevis* template and missing loops are added by ab-initio loop modeling. Final structure is chosen according to the model assessment scoring function. NMDAR activation requires binding of two co-agonists, glycine and glutamate. To be able to observe the structural changes upon ligand binding, glycine and glutamate molecules are docked into the corresponding binding sites of the receptor. Subsequently, Molecular Dynamics (MD) simulations of 1 microsecond are performed for both apo and ligand-bound structures. 10 structural parameters, which have been considered as functionally important in previous NMDA studies, are developed to understand the dynamics of the conformational changes that is associated with the function of the protein throughout the simulations. Moreover, Principal Component Analysis is performed for the equilibrated part of the simulations to classify similar conformations together.

In the ligand-bound simulation, certain loop regions showed higher mobility. Upon ligand binding, closure in LBD clamshell, smaller ATD-LBD inter-domain distance and larger LBD-TMD linker distance is observed in specific subunits. Opening in the bottom TMD girdle is observed for a short time. Correlated motions of the receptor in the ligand-bound simulation increased. The structure showed rotation-like motion in the apo simulation whereas sliding-like motion within the neighboring heterodimers are observed.

ÖZET

N-Metil-D-Aspartat (NMDA) reseptörleri sinir hücrelerinin zarında bulunan glutamat-kapılı iyon kanallarıdır. Reseptörün fonksiyon göstermesi, bilinçlilik ve normal beyin işlevlerinde büyük öneme sahiptir. NMDA reseptörlerinin aşırı uyarılması sonucu, nöronal hücre ölümü oluşur ve bunun sonucunda epilepsi, inme, Alzheimer ve Parkinson gibi hastalıklar ortaya çıkar. Reseptörün moleküler mekanizmasını ve yapı-fonksiyonu ilişkilerini anlamak, yeni ilaç geliştirme mekanizmalarının keşfedilmesini sağlayacaktır.

Yakın zamanda, GluN1-GluN2B reseptörü için iki yeni bütün X-ışını yapısı, biri *Xenopus laevis*'ten, diğeri *Rattus norvegicus*'tan olmak üzere, elde edilmiştir. İki yapı zar proteinlerinin yapılarını kristalize ederken ortaya çıkan genel problemler göz önüne alınarak analiz edilmiş ve karşılaştırılmıştır. İnsan GluN1-GluN2A tipi NMDAR yapısı, *Xenopus laevis* şablon yapısına dayalı olarak modellenmiştir ve eksik döngüler ab-initio döngü modellemesi ile eklenmiştir. Son yapı, model değerlendirme ve puanlama fonksiyonuna göre seçilmiştir. NMDAR aktivasyonu için, glisin ve glutamat olmak üzere iki agonistin aynı anda yapıya bağlanması gereklidir. Ligand bağlandığında yapısal değişiklikleri gözlemleyebilmek için, glisin ve glutamat molekülleri reseptörün ilgili bağlama bölgelerine "dock" edilmiştir. Sonra hem apo hem de ligand bağlı yapılar için 1 mikrosaniyelik Moleküler Dinamik (MD) simülasyonları uygulanmıştır. Önceki NMDA çalışmalarında fonksiyonel olarak önemli olduğu tespit edilen 10 yapısal parametre, simülasyonlar boyunca yapısal değişikliklerin dinamiklerini anlamak için incelenmiştir. Ayrıca, benzer hareketleri birlikte sınıflandırmak amacıyla simülasyonların dengelenmiş kısmı için "Principal Component Analysis" gerçekleştirilmiştir.

Ligand bağlı simülasyonda, bazı döngü bölgeleri daha yüksek mobilite göstermiştir. Ligand bağlandığında, belli zincirlerde, LBD deniz kabuğunda kapanma, daha küçük ATD-LBD mesafesi ve daha büyük LBD-TMD "linker" mesafesi gözlemlenmiştir. Alttaki TMD girdle'ın açılması kısa bir süre için gözlemlenmiştir. Reseptörün korelasyonlu hareketleri ligand bağlı simülasyonda artmıştır. Yapı, apo simülasyonunda rotasyon benzeri bir hareket gösterirken yan yana duran heterodimerler arasında kayma benzeri hareket gözlemlenmiştir.

ACKNOWLEDGEMENTS

Foremost, I would like to express my sincere gratitude to my thesis advisor Assist. Prof. Dr. Şebnem Eşsiz for the continuous support of my master's study and research, for her patience, motivation, enthusiasm, and immense knowledge. Her guidance helped me in all the time of research and writing of this thesis.

Besides my advisor, I would like to express my appreciations to the rest of my thesis committee: Prof. Dr. Mine Yurtsever and Asst. Prof. Dr. Bahar Delibaş, for their encouragement and insightful comments.

My special thanks also go to my colleague Ayhan Demir at this project for his friendship and support, and for creating a cordial working environment.

Last but not least, I would like to thank to my family and friends for their endless supports.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS/ABBREVIATIONS

3D	3-Dimensional
5-HT3	5-hydroxytryptamine3
Å	Angstrom
ACPC	1-aminocyclopropane-1-carboxylic acid
ADT	AutoDock Tools
ALA, A	Alanine
AMPA	α-amino-3-hydroxyl-5-methyl-4-isoxazole-propionate
AMPAR	α-amino-3-hydroxyl-5-methyl-4-isoxazole-propionate receptor
APL	Area per lipid headgroup
ARG, R	Arginine
ASN, N	Asparagine
ASP, D	Aspartic acid
ATD	Amino Terminal Domain
C	Carbon
Cα	Carbon alpha
CNS	Central Nervous System
CTD	Carboxy Terminal Domain
CYS, C	Cysteine
DOPE	Discrete Optimized Protein Energy
fs	Femtosecond
GABAA	Gamma-aminobutyric acid type A
GLU, E	Glutamate
GLY, G	Glycine
H	Hydrogen
HIS, H	Histidine
iGluR	Ionotropic Glutamate Receptor
ILE, I	Isoleucine
K	Kelvin
KIC	Kinematic closure
LBD	Ligand Binding Domain

LEU, L	Leucine
LGA	Lamarckian genetic algorithm
LGIC	Ligand-gated ion channel
LYS, K	Lysine
MD	Molecular dynamics
mGluR	Metabotropic Glutamate Receptor
MK-801	Dizocilpine
N	Nitrogen
nAChR	Nicotinic acetylcholine receptor
NMDA	N-methyl-D-aspartate
NMDAR	N-methyl-D-aspartate receptor
NMR	Nuclear Magnetic Resonance
ns	Nanosecond
O	Oxygen
PCA	Principal component analysis
PCP	Phencyclidine
PDB	Protein Data Bank
pLGIC	Pentameric ligand-gated ion channels
POPC	Palmitoyl-oleoyl phosphatidylcholine
ps	Picosecond
RMSD	Root mean square deviation
RMSF	Root mean square fluctuation
t-ACBD	Trans-1-aminocyclobutane-1,3-dicarboxylic acid
THR, T	Threonine
TMD	Transmembrane Domain
VAL, V	Valine
vdW	van der Waals
z-DOPE	Normalized Discrete Optimized Protein Energy
φ	Phi
Ψ	Psi
ω	Omega

1. CHAPTER 1: INTRODUCTION

1.1. Ion Channels

Neurodegenerative diseases which occur in the central nervous system (CNS) are related with ion channels, since they play a fundamental role in neurotransmission of ions through the cell membrane. Ion channels are protein complexes which are embedded in the cell membrane and their function is to regulate the diffusion of ions across the membrane. Malfunction in the regulation mechanism of ion channels can cause common diseases such as Alzheimer's, Parkinson's, Huntington's, epilepsy and schizophrenia (Dworakowska & Dołowy, 2000). According to the World Alzheimer Report 2016, the number of people who live with dementia was 46.8 million in 2015 and it is expected to increase to 131.5 million in 2050 (World Alzheimer Report, 2016). Therefore, a better understanding of the relationship between the structure and the working mechanism of the ion channels could provide valuable information in order to design novel drugs and to find new treatment methods.

The cell membrane is a biological membrane that consists of phospholipid bilayer and separates interior of the cell from the outside environment. It has a vital importance for the living cells as it controls and maintains the internal composition of the cell. It is selectively permeable to some molecules where only small uncharged molecules can pass freely through the lipid bilayer while large uncharged polar molecules such as glucose and some amino acids cannot. Charged molecules such as potassium (K^+), calcium (Ca^{2+}), sodium (Na^+) and chloride (Cl^-) ions are also unable to pass through the lipid bilayer regardless of their sizes.

Differences in the concentrations of ions on the opposite sides of a cellular membrane causes a voltage called the membrane potential. Cell membranes maintain a small voltage or "potential" through the membrane in its normal or resting state. In the rest state, the inside of the nerve cell membrane is negative with respect to the extracellular environment. The voltage results from differences in concentration of the electrolyte ions K^+ and Na^+. For the nerve cell, this equilibrium is disturbed by the arrival of a suitable stimulus. The dynamic changes in the membrane potential in response to the stimulus is called an action potential. The action potential plays an important role in the communication between nerve cells. After the action

potential, the cell membrane comes back to its resting state by making the intracellular side relatively negative with respect to extracellular side of the membrane.

The lipid bilayer has ion channels on it in order to allow ions to flow across the membrane. An external physical or chemical perturbation determines if the channel is open so that ions can pass through, or if it is closed which prevents the ion flow (Bernard & Shevell, 2008). Transition of the channel between these states is called 'gating'. By the gating mechanism, ion channels regulate the formation of resting and action potentials mentioned above.

Ion channels are fundamental for the normal cellular functioning of the nervous system with three important features including very rapid response to ion flow, to electrical, chemical and mechanical stimuli and selective ion permeation (Hille, 2001).

Ion channels are classified in 5 different subclasses according to how their gating mechanisms work (Keiser & Utzinger, 2010):

- voltage-gated
- ligand-gated
- seconder messenger gated
- mechanosensitive
- gap junctions

Voltage-gated ion channels are the type of transmembrane proteins that is opened when the membrane potential exceeds a threshold value. There are several classes regarding the ion type that pass through the channel, such as Na^+, K^+, Ca^{2+} and Cl^-. When a potential difference exists over the membrane, it triggers the conformational changes in the channel. The conformational change distorts the shape of the channel in such a way that the channel opens to allow ion flow across the membrane. This gating is a result of certain mechanisms related to the structure in the protein. For instance, sodium and calcium channels have voltage sensing helices. These helices are positively charged due to their amino acid sequence. A high positive charge outside the membrane repels these helices keeping the channel in its closed state, or vice versa. The migration of ions balances their concentration in both sides of the membrane which produces an electric current to depolarize the cell membrane (Hille, 2001).

Second messenger-gated channels are only found in eukaryotes and they are activated by light-sensitive compounds such as odors and pheromones (Oldham & Hamm, 2008; McCudden et al., 2005). The mechanosensitive channels are found in all cell types and opened by mechanical stimuli such as touch, cardiovascular regulation and sensing of gravity (Ghazi, 1998). Gap junctions are intracellular channels and they allow diffusions of ions directly cell to cell (Goodenough & Paul, 2009).

Ligand-gated ion channels (LGICs) are opened upon ligand binding. LGICs produce rapid synaptic potentials which inhibits or excite the neuron. This channel family will be discussed further in the following section.

1.2. Ligand-Gated Ion Channels

LGICs are a class of membrane proteins whose functional state is modulated by the binding of a ligand. Upon binding of their ligand, conformational changes take place that open the ion pore and allow the passage of ions. In addition to open and closed conformational states, LGICs can be found in desensitized state where the receptor is unresponsive upon prolonged exposure to its neurotransmitter. This causes the channel to remain closed upon ligand binding to the receptor (Keramidas & Lynch, 2013).

Examples of LGICs are the nicotinic acetylcholine receptor (nAChR), glycine receptor, gamma-aminobutyric acid type A (GABAA) receptor, the 5-hydroxytryptamine3 (5-HT3) serotonin receptor, α-amino-3-hydroxyl-5-methyl-4-isoxazole-propionate (AMPA) receptor, kainate receptor and N-methyl-D-aspartate (NMDA) receptor (Barry & Lynch, 2005). First four receptors are also known as "Cys-loop" receptors due to a pair of conserved cysteines in one of the ligand binding domain loops. AMPA, kainate and NMDA receptors are also known as ionotropic glutamate receptors.

These proteins contain at least two different domains, the first one is a transmembrane domain (TMD) which forms the channel, and the second is an extracellular domain that includes the ligand binding domain (LBD). Ligand-gated ion channels convert chemical signals which come with neurotransmitter released from presynaptic neuron to the electrical signals for the

postsynaptic neuron. LGICs are modulated by allosteric ligands, by channel blockers, or by ions (Barry & Lynch, 2005).

Glutamate receptors are protein complexes that contribute to memory and learning functions, and play a crucial role in synaptic conduction and in the formation of neuronal connections during the development of the nervous system (Traynelis et al., 2010). Glutamate receptors are found in the postsynaptic neuron and divided into two categories such as metabotropic glutamate receptors (mGluRs) and ionotropic glutamate receptors (iGluRs).

mGluRs are G-protein-dependent receptors that participate in the modulation of synaptic transmission and neuronal excitability throughout the central nervous system. The mGluRs bind glutamate within a large extracellular domain and transmit signals through the receptor to intracellular signaling partners.

iGluRs are tetrameric ligand-gated ion channels that are activated by the neurotransmitter glutamate (Traynelis et al., 2010). Structurally, iGluRs consist of one or more subunit subtypes forming a pore in the membrane. Each subunit is comprised of several components that act in a semi-modular fashion: an extracellular amino-terminal domain (ATD), an extracellular ligand binding domain (LBD), a transmembrane domain (TMD), and an intracellular carboxy-terminal domain (CTD) (Laube et al., 1998; Rosenmund et al., 1998). The extracellular amino-terminal sequence of each iGluR subunit forms the clamshell-shaped ATD, composed of upper (R1) and lower (R2) lobes (Fig. 1.1). The ATD is followed by a sequence (S1) that forms much of the upper lobe (D1) of the similarly clamshell-shaped LBD. S1 links to the first (M1) of three transmembrane helices (M1, M3 and M4). A short intracellular linker connects M1 to M2 and a pore-lining reentrant loop (M2/p-loop), which forms the selectivity filter toward the intracellular aspect of the channel. Another short linker connects the M2/p-loop segment to M3 helix. The longer sequence (S2) that connects M3 to M4 comprises most of the lower lobe of the LBD (D2). The CTD, likely to be in an extended conformation, is formed by residues C-terminal to M4 (Ryan et al., 2008).

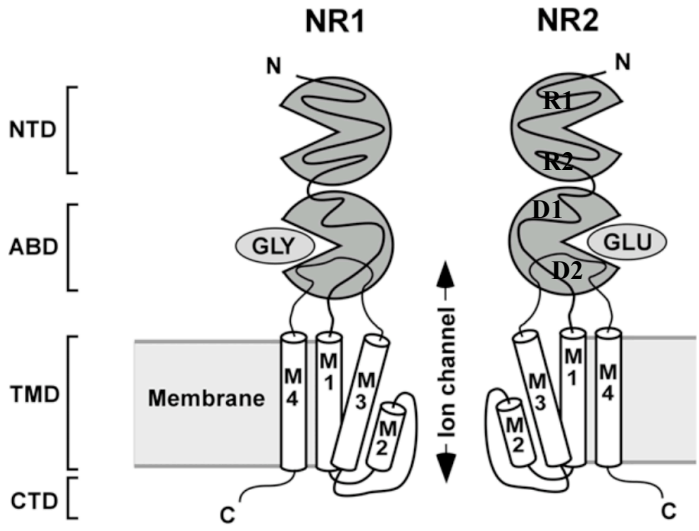


Figure 1.1. *NMDAR Topology. Cartoon diagram depicts a NMDAR with front and back subunits removed for clarity. Each NMDAR subunit contains an extracellular N-terminal domain (ATD), a ligand binding domain (LBD) and a transmembrane domain (TMD) with the helices M1, M2, M3 and M4 and an intracellular C-terminal domain. The ATD and LBD domains are clamshell-shaped with upper lobes termed R1 and D1, respectively, and lower lobes termed R2 and D2, respectively. Reprinted from "Mechanistic Basis of NMDA Receptor Channel Property Variation" by B. S. Retchless, 2011, University of Pittsburgh.*

There are three different members of the iGluR family including AMPA receptors, kainate receptors and NMDA receptors. They are named according to the agonists which activate them.

AMPA receptors (AMPARs) are heterotetrameric receptors that belong to the iGluR family. These receptors are primarily responsible for the rapid excitatory glutamatergic neurotransmission. In nature, the majority of AMPA receptors are heterotetrameric and made up of symmetric dimer of dimers of GluR1, GluR2, GluR3 and GluR4 subtype combinations (Mayer, 2005; Greger et al., 2007). There are four binding sites in AMPA receptors. Each binding site or subunit can bind exactly one agonist (Mayer, 2005). The active site is presumed to be uniquely formed by the amino-terminal and the extracellular loop which is between the transmembrane helices M3 and M4 (Sukumaran et al., 2011). In the case where the agonist binds, the pore opens after some of the two loops move towards one another. The channel opens only if the two binding sites are occupied thus, increasing its currents giving rise to more binding sites to be occupied (Platt, 2007; Rosenmund et al., 1998). When a channel opens, it may be subjected to a rapid desensitization state resulting the current to stop or reach a instantaneous state. The so-called desensitization mechanism is assumed to arise from a little

change in the angle of the regions of the binding site that lead to closing the channel pore (Armstrong et al., 2006; Sukumaran et al., 2011).

One of the two non-NMDA ionotropic glutamate receptors are kainate receptors functioning in the CNS. They are named according to their activator ligands, kainate. They are originally identified by their preferential response to kainate, namely a rapid desensitization in its presence that distinguished them from the other class of non-NMDA iGluRs; AMPA receptors (Davies & Watkins, 1979; Bettler & Mulle, 1995). Kainate receptors have five subunits including GluR5, GluR6, GluR7, Ka1 and Ka2 that are similar to AMPA and NMDA receptor subunits and can be arranged in different ways to form a tetramer (Hollmann & Heinemann, 1994). It is possible that GluR5-GluR7 are formed in homomeric or heteromeric arrangement, however Ka1 and Ka2 subunits are only combined with one of GluR5-GluR7 subunits (Traynelis et al., 2010; Sukumaran et al., 2011).

1.3. N-Methyl-D-Aspartate (NMDA) Receptors

NMDA receptors (NMDARs) are heterotetrameric members of iGluRs located in the membrane of the nerve cells and assembled from a pool of three different gene families: GluN1, GluN2 and GluN3 (Dingledine et al., 1999). The most common subtype combination that exist in human brain is GluN1-GluN2A. NMDA receptors include at least one GluN1 subunit and two or more GluN2 or GluN3 subunits (Ozawa et al., 1998). The GluN1 subunit is essential in all functional NMDARs. It possesses the binding site for regulatory glycine and D-serine and is necessary in the formation of tetrameric NMDAR channels (Kuryatov et al., 1994; McIlhinney et al., 2003). The four GluN2 subunits (GluN2A-D) possess the binding sites for the neurotransmitter glutamate (Anson et al., 2000). GluN3 subunits also contains a binding site for regulatory glycine and D-serine and exist in two different isoforms such as GluN3A-B (Yao & Mayer, 2006). Some important functional properties such as kinetics, open channel conductance and voltage dependence can be influenced by receptor composition with respect to these isoforms (Gillesen et al., 2002).

NMDARs characteristically bind the agonist glycine in LBD of GluN1 subunits and glutamate in LBD of GluN2 subunit. When glutamate and glycine binds to their corresponding binding

sites, lower lobe (D2) of the LBD clamshell comes close to the upper lobe (D1) and packs the ligand in the binding site. This closure motion pulls the pore-forming M3 helices in TMD up of the membrane which leads to the activation of the channel allowing the passage of small cations including Na^+, K^+ and Ca^{2+}. Influx of ions through these channels generates neuronal action potentials (Schneggenburger et al., 1993; Lee et al., 2014; Zheng et al., 2017).

There are several distinguishing properties of neuronal NMDARs that make them unique when compared to other ionotropic receptors: 1) in addition to the natural agonist glutamate, they require an additional co-agonist, either glycine or D-serine, to function and 2) they are normally blocked by Mg^{2+} ions located within the channel pore. Membrane depolarization occurs in order to repel Mg^{2+} from the channel pore (Zhu et al., 2013). Other substances which are known to interact with NMDARs and modulate their activity include zinc cations, protons, polyamines such as dizocilpine (MK-801) and the psychotropic drug phencyclidine (PCP) (Ransom & Stec, 1988).

The binding sites of NMDA receptors are of great pharmacological importance. NMDA receptors have 7 different binding sites (Kandel et al., 2014; Kayaalp, 2009).

- Agonist recognition regions; NMDAR reacts with glutamate and other agonists. It enables the opening of the ion channel in the receptor to allow the normal exciter effect to occur (Kandel et al., 2014).
- Glycine binding region; Glycine, normally acting as an inhibitor neurotransmitter, facilitates the activation of NMDA receptors; Physiologically, is called a co-agonist of glutamate (Kandel et al., 2014)
- The polyamine binding site; The natural polyamines, spermidine and spermine, increase glutamate activity as the glycine does (Kandel et al., 2014).
- Zinc binding site; The Zn^{2+} ion binds to its own specific binding site on the NMDA receptor, reducing the glycine effect (Kayaalp, 2009). A voltage-dependent blockade is associated with the binding of Zn^{2+} ion to this region. This region has an inhibitory effect.
- Antagonist binding site; Phencyclidine, ketamine, dizocilpine and benzomorphan derivatives are the sites to which the substances bind. When antagonists bind this region, channel is closed to flow ions. (Kandel et al., 2014). This region is suitable for binding antagonists which will prevent excitotoxicity.

- Cation binding site; This region is located in the channel, where Mg^{2+} binds and blocks the ion current through the membrane. The effect of Mg^{2+} is voltage dependent. In order for the cation channel to be able to get rid of the inhibitor effect of Mg^{2+}, the membrane voltage needs to be raised to a sufficient degree (Kayaalp, 2009).
- Allosteric modulation site REZ

1.4. Structural Studies of Glutamate Receptor

The main obstacle to understanding the gating mechanism of NMDAR is that the active, glutamate bound and intact receptor structure open to ion transport has not been experimentally resolved yet (Schmid & Hollmann, 2008).

There are studies that focused on the topics which are based on molecular dynamics such as the ligand selectivity for the LBD part (Lau & Roux, 2011), the semi-active agonist activity (Postila et al., 2011) and the closure of the two LBD shells triggered by the agonist (Dravid et al., 2010; Frydenvang et al., 2009). In addition to these studies, it has been also found that certain amino acids actively coordinated to close LBD clamshell upon glutamate binding to GluR2 LBD. In this study, 256 simulations of 2 ns MD are run, and it is observed that the binding of the glutamate molecule on six of them led to the closure of two halves of LBD onto ligand molecules (Okada et al., 2012). Furthermore, as a result of examining the results obtained by the two different MD simulations in terms of dominant protein motions, it is found that, in contrast to the AMPA receptors, the partial agonists are as effective as full-agonist molecules in the closure of the clamshell. Differences between the agonist and partial agonists are observed only in the part playing the hinge role in the closure of two flaps.

In another simulation study, it has been observed that, in addition to the glutamate molecule, water molecules in LBD play an important role in conformational changes caused by ligand binding, and that water molecules can regulate themselves in the nanosecond time interval (Sahai & Biggin, 2011; Vijayan et al., 2010).

Moreover, in the molecular dynamics simulations of the isolated ATD, the clamshell-like motion of LBD and the twisting motions of two lobes in ATD rotating against each other are observed as dominant global protein movements (Dutta et al., 2012; Zhu et al., 2013). These

studies have identified that the ATD of the NMDA receptors plays a role in ion channel activation as the main difference between AMPA and NMDA receptors.

The iGluR family exists as heterotetramer structures in the nature. This heterotetramer structure consists of two dimers of dimers in all iGluR members including NMDAR. Until 2014, the only resolved X-ray structure was AMPAR-GluR2, which is homotetramer in nature. In 2014, two valuable X-ray structures have been crystallized. Both of those are the GluN1-GluN2B structure in the heterotetramer form. The first is resolved by Lee and his colleagues (Lee et al., 2014) at 3.7 Å, and published in Nature in July, while the second is resolved at 4 Å resolution by Karakas and Furukawa (Karakas & Furukawa, 2014). The first structure is crystalized from the *Xenopus laevis* with the allosteric inhibitor Ro25-6981 in ATD dimer interface and partial agonists 1-aminocyclopropane-1-carboxylic acid (ACPC) in GluN1 LBD and trans-1-aminocyclobutane-1,3-dicarboxylic acid (t-ACBD) in GluN2B LBD. The second structure is crystalized from the *Rattus norvegicus* with the allosteric inhibitor ifenprodil in ATD dimer interface and the agonists glycine in GluN1 LBD and glutamate in GluN2B LBD. Both structures are found in physiologically inactive states. These proteins revealed that the structures of NMDARs are topologically the same as that of AMPAR published in 2009.

When the AMPA receptor structure is compared with the newly resolved X-ray structures, it has been determined that the domains outside of the ion channel are much more compact in the NMDA receptor (Karakas & Furukawa, 2014; Lee et al., 2014).

Finally, in August 2014, a work has been published in Cell which inspects the 'desensitized' AMPA receptor, insensitive to the ligand molecule but closed to ion flow, by obtaining low-resolution images via cryo-electron microscopy method. Conformational changes between the apo state (in the absence of ligand) and the desensitized state of the structure suggested in this study involves very large rotational and twisted movements of the ATDs and LBDs on the ion channel (Dürr et al., 2014). In such a large conformational change, `global` motions in protein domains rather than local atomic motions appear.

In this thesis work, homology model of human GluN1-GluN2A NMDAR will be simulated with computational approaches. The available crystal structures contain a number of mutations to reduce flexibility and to improve X-ray diffraction quality such as disulfide cross-links between subunits and removal of some flexible regions in the receptor. Moreover, the

structures are crystallized in the presence of antagonists and allosteric inhibitors, and the ion channel is closed. Hence, the gating mechanism of the receptor remains still controversial.

The aim of the study is to investigate the dynamics of the apo receptor in the absence of ligand molecules, to observe the open and closed states of the ligand-bound receptor upon glycine and glutamate binding using molecular dynamics (MD) simulations and to compare the conformational differences or correlations take place in whole structure and domains including ATD, LBD and TMD in both simulations. In addition, binding regions of ligand molecules that regulate gating mechanism and also how allosteric changes at the receptor are transmitted through the ion channel in the cell membrane will be examined. Allosteric changes are the result of structural changes in a local area of the protein that alter function of the protein by spreading to the other regions. The consequences of such a study would provide information to find molecules that have a controlled effect on the function of this protein. So, this will directly assist in the design of new drug molecules to be used in the treatment of neuronal death-related diseases such as Alzheimer's and Parkinson's.

2. CHAPTER 2: MATERIALS AND METHOD

2.1. Loop Modeling

It is possible that the alignment between model and template sequences contains gaps either in the model sequence or in the template sequence. If gaps in the model, residues are omitted from the template, creating a hole in the model that must be closed. If gaps in the template, the continuous backbone is taken from the template, and missing residues are inserted. Both cases imply a conformational change of the backbone. These changes in loop conformation make predict the structure difficult. There are two main approaches to loop modeling:

Knowledge based: The loop which is inserted to the structure is searched on the database for known loops with endpoint that match the residues. The database, typically contains thousands of loops, includes loop coordinates, sequences, number of residues in loop, Cα-Cα distances and many more. When consistent loops are found, placement and positioning of these loops are done via using superposition algorithm. Then loop fits are evaluated using root mean square deviation (RMSD) or energy function.

Ab initio: If any loop is not found in database, an algorithm is used to create loops using randomly generated φ/Ψ angles. Then an energy function is used to determine the quality of a loop. This loop is minimized, using Monte Carlo (Simons et al., 1999) or molecular dynamics techniques (Fiser et al., 2000) to attain the best loop conformation.

2.1.1. ROSETTA Kinematic Closure (KIC) with Fragments

The Kinematic Closure (KIC) which is component of ROSETTA's loop modeling package, is an algorithm to create closed loops. Kinematic Closure (KIC) with fragments is a significant method in structural biology since it is used for modeling loops such as missing loops in the crystal structure. KIC is also used to predict conformations of loops in homology modeling, predict conformational changes upon protein-protein interactions. Two main protocols of the algorithm are;

- Loop reconstruction: Ab-initio prediction of loop structure according to given amino acid sequence.
- Loop refinement: Searching low-energy conformations of constructed loops.

In the KIC with fragments protocol, φ/Ψ/ω torsion angles from sequential residues in fragments of peptide (default monomers, trimers or nonamers) are used to sample all without 6 torsional degrees of freedom. KIC solves 6 torsions analytically.

There are 2 stages, namely loop reconstruction and loop refinement, to model the loops with "Kinematic Closure with fragments" algorithm. The loop reconstruction aim is basically structure improvement by remodeling or rebuilding peptide fragments and is advantageous due to rapid procedure and tremendous number of backbone conformation generation. The very first step is to produce initial loop which consist assumption of bond lengths, angles and omega torsions. Later, by the help of Ramachandran, idealized loop φ/Ψ torsions are changed with random values which guarantees the dispose of all the information about native loop. φ/Ψ/ω torsions are modeled with the (information derived from fragment) fragment data, continued with kinematic closure of the loop. This step utilizes Rosetta low-resolution scoring function and a centroid image to produce backbone conformations of each loop within the loop definition file. Predicting low energy conformations of peptide segments according to presented structure is performed at refine step. In refinement stage, Rosetta's high-resolution scoring function Talaris (Shapovalov & Dunbrack, 2011; Song et al., 2011; Leaver-Fay et al., 2013) applied for evaluation of side-chains and backbone conformations (Mandell et al., 2009).

2.2. Molecular Docking

Docking is a computational method used to predict the preferred conformation and position of atoms in a molecule when forming a stable complex with another molecule. In molecular docking technique, there are two common approaches to dock ligand into protein. The first is geometric shape complementarity which is defined as an algorithm to match solvent-accessible surface area of macromolecule and molecular surface of the ligand according to their conformations. (Goldman et al., 2000; Meng et al., 1992; Morris et al., 1998) The second approach simulates the real docking procedure where the ligand-protein pairwise interaction energies are evaluated (MacKerell et al., 2004; Vogt & Cera, 2012).

AutoDock Tools (ADT) (Morris et al., 2009) is a suite of automated docking tools which is designed to predict how small molecules, such as substrates or drug candidates, bind to a receptor of known 3D structure. It contains two main programs: the first one; AutoGrid4 is used for pre-computation of the area of the 3D "grid" structure, i.e. the region to which the ligand is connected. The second is AutoDock4 (Morris et al., 2009), which allows the ligand molecule to be targeted to the binding site. AutoDock4 allows the most appropriate conformation of the complex formed by the protein and the ligand.

AutoGrid uses a set of AMBER force field packages to define three-dimensional structure and to calculate van der Waals and Coulomb's interactions (Huang et al., 2010). For each atom type in the ligand molecule to be docked, energy evaluation is achieved by pre-calculating atomic affinity potentials. In the AutoGrid procedure, the protein is embedded in a three-dimensional grid and a probe atom is located at every grid point. The interaction energy of this single atom with the protein is appointed to the grid point. AutoGrid affinity grids are calculated for each atom type in the ligand molecule, typically carbon, oxygen, nitrogen and hydrogen, as well as grids of electrostatic and desolvation potentials. Then, during the course of AutoDock calculation, the energetics of a certain ligand configuration is evaluated using the values from the grids.

AutoDock's most widely used conformational search algorithm is the Lamarckian genetic algorithm (LGA). It also includes simulated annealing and traditional genetic algorithms. AutoDock uses a semi-empirical force field to predict the free energy of binding of small molecules to macromolecules. AutoDock is run for several times to produce several docked ligand conformations. The program allows the entire ligand to be completely flexible, as well as some specific portions of the protein (Morris et al., 2009). Finally, the best conformations of the ligand in the desired binding site are obtained according to their energy scores.

2.3. Molecular Dynamics

The Molecular Dynamics (MD) simulation technique is a computational method aimed at investigating the motions of multi-particle systems over time using classical physics. MD first emerged in the 1950s and is now widely used in multi-atomic and complex systems, such as

biological structures. Since the experimental methods provide static images of the biomolecules, MD is an efficient tool to investigate the structure-function relationship of the biosystems by studying the dynamics of the system.

In molecular level, interactions on the atoms result from both interatomic bonds and their charges and positions. Basically, 5 different interactions can be mentioned about on an atom. Interatomic interaction terms can be called angle bending, bond stretching, and torsion, originating from the covalent bonds between the molecular atoms. These interactions are found only in atoms with covalent bonds between them. The atoms that are not connected to each other can also have Coulomb interaction or Lennard-Jones interaction due to their charges and the distribution of these charges. Although the Lennard-Jones interaction is between all atoms, there is no Coulomb interaction in atoms with a net charge of zero.

The potential energy function, or force field, that includes all the intra- and inter- molecular forces within the system mentioned above is used in MD simulation. This force field contains terms that describe how the energy changes as bonds are rotated and the terms that describe interaction between non-bonded parts of the system. A functional force field mainly consists of bonded and non-bonded interactions and is generally represented by the following:

$$E_{potential} = \sum E_{stretching} + \sum E_{bending} + \sum E_{torsion} + \sum E_{nonbonded\ interactions} \qquad (2.1)$$

$$\mathcal{V}(r^N) = \sum_{bonds} \frac{k_l}{2}(l - l_0)^2 + \sum_{angles} \frac{k_\theta}{2}(\theta - \theta_0)^2 + \sum_{torsions} \frac{V_n}{2}(1 + \cos(n\varphi - \gamma)) + \sum_{i=1}^{N} \sum_{j=1+1}^{N} (4\varepsilon_{ij}[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}] + \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}}) \qquad (2.2)$$

$\mathcal{V}(r^N)$ denotes the potential energy, which is a function of the positions (r) of N atoms. k_l and k_θ are constants for bond stretching and angle bending terms, respectively. The first term in Equation 2.1 and 2.2 models the interaction between pairs of bonded atoms, modelled here by a harmonic potential that gives the increase in energy as the bond length l_i deviates from the reference value $l_{i,0}$. The second term is a summation over all angles formed between three atoms in the molecule, again modelled using a harmonic potential. The third term in Equation 2.2 is a torsional potential that models how the energy changes as a bond rotates. The fourth contribution is the non-bonded term. This is calculated between all pairs of atoms (i and j) that are in different molecules or that are in the same molecule but separated by at least three bonds. In a simple force field, the non-bonded term is usually modelled using a Coulomb potential term for electrostatic interactions and a Lennard-Jones potential for van der Waals interactions.

The equations of motion of particles interacting with each other within a certain time period are numerically solved and trajectories are determined. Here the trajectory represents all of the particle coordinates at a certain time interval obtained in each time step. In the Newtonian mechanics, the potential energy on each particle is calculated, and from this potential energy, the net force on the particle is found. With the help of this net force, the function that determines the motion of the particle is solved numerically using the second law of Newton.

2.3.1. Potential Energy Velocity Verlet

Throughout the simulation, the force on each atom alters whenever the position of the atom changes, or whenever the positions of other interacting atoms changes. As the number of atoms increases, the number of interactions also increases, so the calculation of the potential energy becomes challenging, namely many-body problem. A finite difference method is used to handle such a situation. The fundamental idea in this method is to break down the integration into many small stages. One of the commonly used algorithms in finite difference methods is Velocity Verlet algorithm. In velocity verlet, positions and velocities are calculated at the same time value of the time variable by solving Taylor series expansion.

Fig. 2.1 shows the operation of a typical MD simulation. As can be seen from Fig. 2.1, the most important factor that must be known before starting the simulation is the initial position of the atoms. There are several experimental (e.g. Nuclear Magnetic Resonance (NMR) and X-ray

Crystallography) and theoretical (e.g. Homology Modeling) ways of obtaining these positions. The atomic coordinates of the structures used in this work are obtained from experimental studies. Although theoretical methods are very powerful today, the effective approach is to experimentally determine the structure of the molecule of interest.

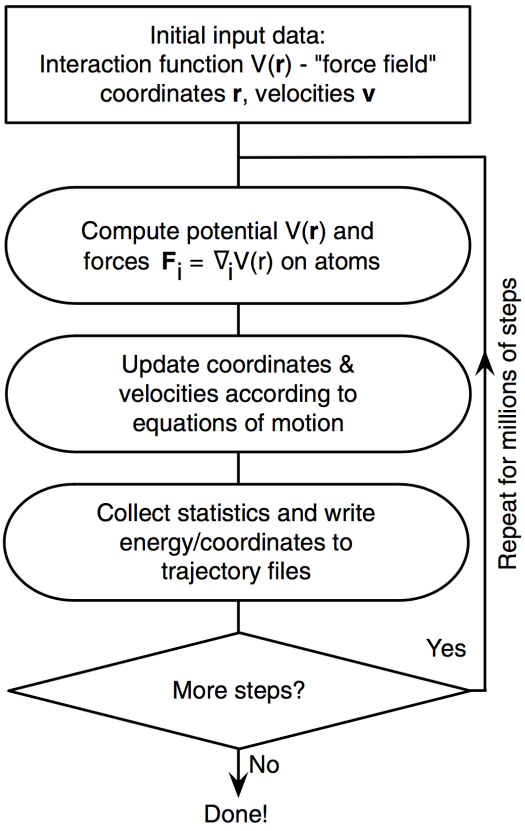


Figure 2.1. *Operation steps of a typical molecular dynamics (MD) simulation.*

Knowing the initial positions of atoms is required to start a MD simulation. In addition, one of the important things that must be known for the simulation to run is the initial speeds of the atoms. The initial speeds of the atoms are not known. Instead, a random initial velocity is distributed to each atom according to the Maxwell-Boltzmann distribution. In the next step, the forces on each atom in the system are calculated from the first derivative of potential energy. With the help of Newton's second law, the acceleration of the atoms is calculated from the forces acting on them. Once the acceleration is known, velocities are calculated and then, the positions of the atoms after certain time interval are determined according to their velocities. During the simulation, it is possible to monitor the change of many physical properties such as pressure, temperature and the energies related to the system.

The most time-consuming part of this integration scheme is calculating pairwise interactions. Especially calculation of the non-bonded interactions is challenging since the number of interactions dramatically increase when all of the atom pairs are included in potential energy calculation. In order to eliminate this difficulty, certain cutoff distance, generally 12-15 Å, is used to calculate the potential energy of the system by taking into account the interactions for only the atoms within that cutoff distance. Moreover, periodic boundary conditions are generally set up in MD simulations to deal with the boundary effects. Periodic boundary conditions enable a simulation to be performed using a relatively small number of particles in such a way that the particles experience forces as if they are in bulk fluid. For instance, in a cubic box simulation system, there are 26 imaginary neighbors around the center box in each dimension. The coordinates of the particles in the image boxes can be computed simply by adding or subtracting integral multiples of the box sides. In the case an atom leaves the center box during the simulation then it is replaced by an image atom that enters from the opposite side. Thus, the number of atoms within the central box remains constant.

2.4. MD Software

NAMD is a parallel Molecular Dynamics software designed for the high-performance simulation of biomolecular systems using force fields (Phillips et al., 2005). There are numerous empirical force fields designed for the simulations in various computational areas. Some of them, such as CHARMM, AMBER and GROMOS, have been intended specifically for the simulations of biomolecular systems. Either CHARMM or AMBER force fields can be used for parameterizations within NAMD.

Both CHARMM and AMBER are empirical, all-atom force fields that support proteins, nucleic acids, lipids and carbohydrates (Kukol, 2015). Although the general form of potential energy functions in both force fields are similar, significant philosophical and parameter optimization distinctions exist due to the additional improper (Equation 2.3) and Urey-Bradley (Equation 2.4) terms in CHARMM (Leach, 2001). Improper torsion term is basically calculated based on four atoms not successively bonded and used to select the correct geometry or chirality of atoms. Urey-Bradley is a cross-term accounting for angle bending using 1,3 nonbonded interactions and introduces a virtual bond between atoms 1 and 3. It is added into CHARMM

force field in order to apply a harmonic restraint on the distance between atoms 1,3 with respect to each other (Leach, 2001).

$$E_{improper} = \sum \frac{k_\omega}{2} (\omega - \omega_0)^2 \qquad (2.3)$$

$$E_{Urey-Bradley} = \sum \frac{k_u}{2} (u - u_0)^2 \qquad (2.4)$$

2.5. MD Trajectory Analysis Tools

2.5.1. Root Mean Square Deviation (RMSD)

The root-mean square deviation (RMSD) is a method that allows quantifying the average structural displacement of the protein atoms with respect to the selected reference (starting) structure. Often, instead of all atoms in RMSD calculations, only Cα atoms or backbone atoms (N, Cα, O) are involved. The RMSD is calculated according to the Equation 2.5:

$$RMSD = \sqrt{\frac{\sum_{i=1}^{N_{atoms}} (r_i(t_1) - r_i(t_2))^2}{N_{atoms}}} \qquad (2.5)$$

where N_{atoms} is the number of atoms whose positions are being compared and $r_i(t)$ is the position of atom i at time t after superimposing on the reference frame. The calculation is repeated for every frame throughout the simulation trajectory.

2.5.2. Root Mean Square Fluctuation (RMSF)

Root mean square fluctuation (RMSF) is a method that gives information about the local conformational changes in the protein structure throughout the simulation trajectory. RMSF is measured according to the Equation 2.6:

$$RMSF = \sqrt{\frac{\sum_{t_j=1}^{T} < (r_i(t_j) - r_i^{ref})^2 >}{T}} \qquad (2.6)$$

where T is the duration of the simulation (time steps), r_i^{ref} is the reference frame, $r_i(t_j)$ is the coordinate of the residue i at time t_j after superimposing on the reference, and the angle brackets indicate that the average of the square distance is taken over the selection of atoms in the residue.

2.5.3. Area Per Lipid (APL)

GridMAT-MD (Allen et al., 2009) is a simple program, designed to analyze two important parameters in simulations of lipid bilayers; area per lipid headgroup (APL) and bilayer thickness. The program is particularly useful in the case of an embedded membrane protein, as the program can calculate the APL, while still compensating for the lateral area occupied by the protein at the intersection with the membrane interface.

2.5.4. Principal Component Analysis (PCA)

Molecular dynamic simulations can produce a huge amount of data which must be processed and analyzed. It is crucial to select significant conformations from among smaller data that has the representative set of conformations for further analysis. This can be achieved using cluster analysis which classifies similar conformations together.

A principal component is basically the linear combination of the variables. Principle component analysis (PCA) is applied to the simulation trajectories in order to reduce dimensionality of the system and to acquire collective results from the simulation. The first

step in PCA is the superimposition of each conformation in the trajectory into a reference structure. The second step is to construct variance-covariance matrix by calculating the deviations of positions from the average structure. For a system consisting of N atoms, the dimension of the matrix is 3Nx3N. In the third step, the eigenvalues and the eigenvectors are obtained by matrix diagonalization where 3N eigenvalues are yielded. The eigenvectors represent the intrinsic collective motions of the protein and the corresponding eigenvalues, after removing 6 zero eigenvalues (3 for rotational and 3 for translational motions), represent the magnitudes of these motions (David & Jacobs, 2014).

PCA is performed for the structures taken from MD simulations, using the PTRAJ utility of AMBER software (Case, 2005, 2008).

The input is an *n* by *p* coordinate matrix, *X*, where *n* is the number of snapshots and *p* is three times the number of atoms. Each row in *X* represents the atom coordinates of each snapshot structure. The elements of the covariance matrix, *C*, is calculated from the Equation 2.7:

$$C_{ij} = <x_i - <x_i>>. <x_j - <x_j>> \qquad (2.7)$$

where averages are over the *n* snapshots. The covariance matrix, *C*, can be decomposed as

$$C = P\Delta P^T \qquad (2.8)$$

where the eigenvectors, *P*, represent the principal components (PCs) and the eigenvalues are the elements of the diagonal matrix, Δ. Each eigenvalue is directly proportional to the variance it captures in its corresponding PC.

Correlation analysis of the system is performed by the Equation 2.9:

$$C_{ij} = \frac{\sum_{l=1}^{10} \frac{U_{il}U_{jl}}{\Omega_{ll}}}{(\sum_{m=1}^{10} \frac{U_{im}U_{jm}}{\Omega_{mm}})^{\frac{1}{2}} (\sum_{n=1}^{10} \frac{U_{in}U_{jn}}{\Omega_{nn}})^{\frac{1}{2}}} \qquad (2.9)$$

where U is the matrix of eigenvectors, and Ω is the diagonal matrix of eigenvalues. The cross-correlations maps indicate correlation of motion between different parts of the protein.

2.6. Modeling and Simulation Setup

In this thesis study, two homology modeled structures of NMDA receptors have been used. The homology modeling of the human NMDAR GluN1-GluN2A based on the crystal structure of *Xenopus laevis* GluN1-GluN2B NMDAR (PDB code: 4TLM) has been carried in a separate study (Demir & Essiz, 2017). Missing loops in the crystal structure are modeled by ROSETTA (Rohl et al., 2004).

Since the initial crystal structure has agonists and allosteric inhibitors, these molecules in the model are removed in order to obtain ligand-free structure of the protein, and this model will be named as 'Apo' throughout the thesis. The second structure used in MD simulation is obtained by docking the agonists glycine and glutamate into the LBDs of the receptor using AutoDock and this model will be named as 'Ligand-bound' throughout the thesis. The docking procedure will be explained below in detail.

During the preparation of both models for MD simulations, the protonation states of histidine (HIS) and other charged residues (GLU, ASP, LYS and ARG) have been determined by PROPKA (Olsson et al., 2011; Søndergaard et al., 2011) web server by calculating the pKa values of these residues at pH 7.4. All of the disulfide bridges between the close cysteine residues recommended in the X-ray structure pdb file are applied to the system in VMD (Humphrey et al., 1996). Both models are inserted into an unequilibrated palmitoyl-oleoyl phosphatidylcholine (POPC) membrane. Then they are solvated with TIP3P water molecules and are neutralized by adding Na^+ and Cl^- ions into the system. At the end, a system consisting of nearly 230,000 atoms are obtained.

First, in order to equilibrate the POPC membrane, a simulation is performed in which all other atoms (water, ion, protein and lipid head groups) in the system are kept fixed except for lipid tails. This first step consists of 2 picoseconds (ps) (1,000 steps) minimization and 2 nanosecond (ns) (1,000,000 steps) MD runs with 1 fs time step at a constant temperature of 300 K and at constant volume (NVT). In the second step, both systems are minimized for 2 ps and simulated for 8 ns (four 2 ns simulation one after another with the spring constant 1, 0.75, 0.5 and 0.25 kcal/mol/$Å^2$, respectively) with positional constraints implemented on protein atoms at 300 K constant temperature and 1 atm constant pressure (NPT). Furthermore, by the use of "NAMD

Tcl Forces" it is prevented the penetration of water atoms into the hydrophobic region of the membrane and the membrane-protein interface. This step is important for the equilibration of other molecules around the protein. Then, in the third step, the positional constraints on the protein atoms are completely removed and simulations are run 5 ns (NPT, lipid area not constant) with constant system pressure (1 atm) and temperature (310 K) while all system atoms are released. In the last step, the lipid area is kept constant, and 200 simulations of 5 ns (1 microsecond in total) are performed in 1 atm, 310 K (NPT).

Table 2.1. *Protocols used in MD simulations.*

	Protocol 1	**Protocol 2**
Step 1 (Equilibration of Lipid Tails)	Minimization: 2 ps MD: 2 ns All atoms are fixed except lipid tails	Minimization: 2 ps MD: 2 ns All atoms are fixed except lipid tails
Step 2 (Protein Constrained)	1) Minimization: 2ps, MD: 1 ns Force constant on protein atoms: 1 kcal/mol/Å^2 2) MD: 0.5 ns Force constant on protein atoms: 0.5 kcal/mol/Å^2 3) MD: 0.5 ns Force constant on protein atoms: 0.25 kcal/mol/Å^2	1) Minimization: 2ps, MD: 2 ns Force constant on protein atoms: 1 kcal/mol/Å^2 2) MD: 2 ns Force constant on protein atoms: 0.75 kcal/mol/Å^2 3) MD: 2 ns Force constant on protein atoms: 0.5 kcal/mol/Å^2 4) MD: 2 ns Force constant on protein atoms: 0.25 kcal/mol/Å^2
Step 3 (Protein Released)	MD: 5 ns NPT system, Lipid Area Not Constant	MD: 5 ns NPT system, Lipid Area Not Constant
Step 4 (Protein Released)	MD: 5 ns x 200 = 1000 ns NPT system, Lipid Area Constant	MD: 5 ns x 200 = 1000 ns NPT system, Lipid Area Constant (In progress)

All MD simulations are run using the CHARMM force field in NAMD with a nonbonded vdW cutoff 12 Å. Constant pressure is maintained by a Langevin piston set at 1 atm and constant temperature is maintained using a Langevin temperature piston set at 310 K in all the production runs. Particle Mesh Ewald electrostatics settings are applied to both systems. All of the production runs are performed using 2 fs time step.

2.6.1. Docking

AutoDock Tools (ADT) is used to calculate the positioning of glycine and glutamate, which are the natural agonists of the receptor that is obtained from homology modeling.

First, the ligands to be positioned on the model with AutoDock4 are isolated from the crystal structure. Glycine molecules for GluN1 (1^{st} and 2^{nd} chains, will be named $GluN1_1$ and $GluN1_2$ in this document) and glutamate molecules for GluN2A (1^{st} and 2^{nd} chains, $GluN2A_1$ and $GluN2A_2$) are separately saved in pdb format. The ligand binding sites of each chain in the model and the ligands for those sites are matched according to the best conformation. In this process, the torsion numbers of each ligand are first determined, i.e. how many and which bonds in the molecule are allowed to be rotated, are given as inputs to the ADT. Thus, different conformations can be calculated for the molecule. This is done with default values for each ligand identified by the ADT. Ligands are saved in pdbqt format with this information.

During the docking process, ADT Grid Options Widget is used to determine the region to be searched where the ligands are located at the most appropriate location. This region is determined by selecting a center and then by setting a certain number of points in each direction at equal distance from each other. Based on the crystal structure, the center for each ligand is selected from the LBD regions of each chain, and a grid box dimension is determined as 50 * 50 * 50 in Å. This process is completed using AutoGrid4 via ADT.

After the preparation stage, ligands for the docking process and information on the areas where these ligands can be positioned on the macromolecule are obtained. Genetic algorithm is used as a search parameter for this positioning. With this algorithm, each ligand is evaluated 2,500,00 times for each run in the Grid box. Lamarckian genetic algorithm is used as the output. With this algorithm, the best 20 conformations are selected according to energy scores (Huey et al., 2007). AutoDock4 is run for each chain and ligand with these parameters, and the best conformation of ligands is placed on the model to be used for simulation.

3. CHAPTER 3: RESULTS AND DISCUSSION

3.1. Analysis of Two Template Structures

There are two X-ray structures available for homology modelling step. Both X-ray structures were resolved in 2014, first one was published in Science and the second in Nature, with resolutions 4 Å and 3.7 Å, respectively. The sequence identity of the human NMDA receptor and X-ray structure derived from the *Xenopus laevis* (Lee et al., 2014) is 92% and 81% with human GluN1 and GluN2A, respectively. The sequence identity of the human NMDA receptor and X-ray structure derived from *Rattus norvegicus* (Karakas & Furukawa, 2014) is about 98% for both chains.

Using *Rattus norvegicus* NMDA structure as a template for homology modeling is thought to be more acceptable because the sequence identity with human sequence is higher in *Rattus norvegicus* NMDAR. Unfortunately, a number of mutations have been applied to sequences in order to crystallize these structures. These structural changes can alter the characteristics of the conformational changes driving the function of the protein. For instance, approximately 15 neutralizations and 5 CYS cross-linking mutations are added to the *Rattus norvegicus* structure. Four of these CYS mutations are shown in Fig. 3.1(yellow beads indicate cross link locations). These mutations connect TMD M1 and M4 helices. The CYS mutations in TMD on each subunit might be essential in terms of transmembrane helices moving together when gating. Namely, when these mutations are recovered, the structure can gain freedom between the subunits (Karakas & Furukawa, 2014; Lee et al., 2014).

There is a K216C cross-link cysteine mutation which links ATDs in the structure of *Xenopus laevis* (Fig. 3.2A ATD site). In the absence of this mutation, ATD fragments are observed to be opened approximately 25 degrees (Fig. 3.2 B) (Lee et al., 2014).

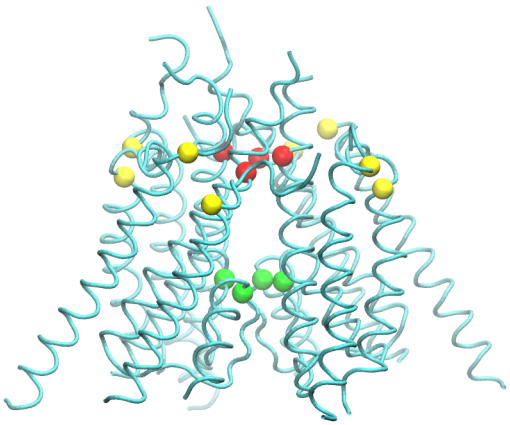


Figure 3.1. *Structure of ion channel (only TMD is shown). Red beads indicate TMD girdle (top), green beads TMD bottom, yellow beads cross link locations.*

As a result of the above-mentioned reasons, the primary focus is the comparison and stability of these two NMDAR X-ray structures as the initial step of this study. The structures of *Xenopus laevis* and *Rattus norvegicus* are compared in Fig 3.2A. TMD part is the most challenging step of crystallization and it is obtained by homology modeling using the potassium channel as the template (Lee et al., 2014). Fig. 3.2 A reveals that the M1-M2 linker helix, which have been shown to play important role in channel closure (Lee et al., 2014), in TMD of the *Rattus norvegicus* structure, is not completely resolved. Additionally, some other helices are missing (red colored regions in Fig. 3.2 A). Moreover, since obtaining a stable TMD structure is challenging, the different subunits of TMD are linked by CYS cross-links as depicted above (yellow balls in Fig. 3.1). Green ball in Fig. 3.2 A shows the location of the CYS cross-link between ATD of GluN2B dimers.

The structure from *Xenopus laevis* seems less modified, especially for TMD segment. Still the two loop structures in GluN2B are cut off. These loop structures consist of 9 or 10 amino acids. One of them links the M1 helix to M2 helix (red colored and labeled as missing pore helix in Fig. 3.2A) and the other connects M3 helix to the LBD region (green colored and labeled as missing linker in Fig. 3.2A) (Karakas & Furukawa, 2014; Lee et al., 2014). There are also missing ATD-LBD linkers in $GluN1_2$, $GluN2B_1$ and $GluN2B_2$ (Fig. 3.2A).

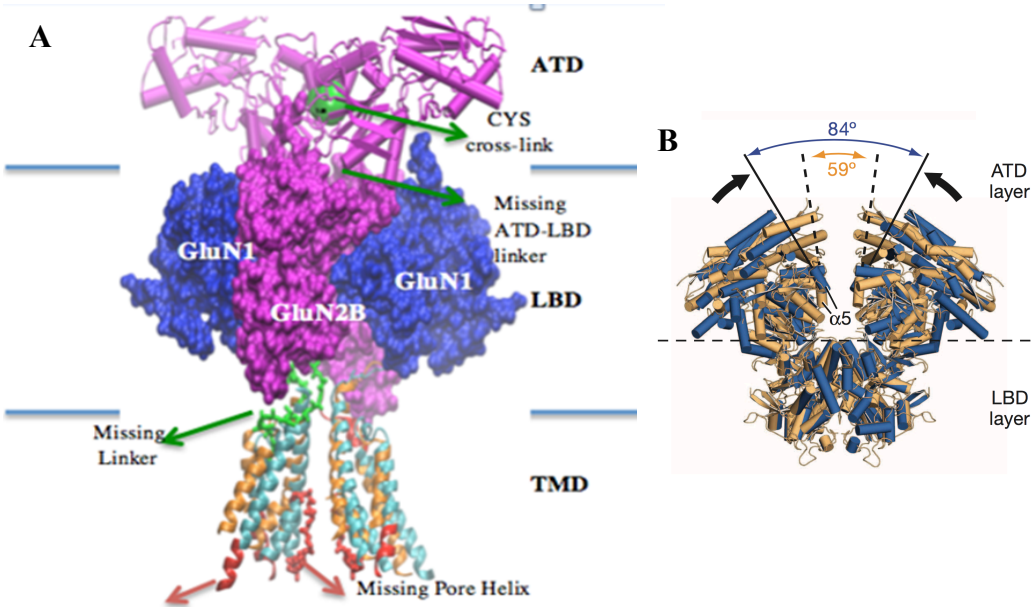


Figure 3.2. (A) *Structural differences between the structures of 4TLM and 4PE5.* ***(B)*** *Disulfide bridge mutation in the ATD and the observed increase in angle in the absence of this mutation.*

Additionally, the ion-channel radius profiles are generated using the HOLE program (Smart et al., 1996) in both structures (see Fig. 3.3C). Pore profiles along the channel of 4TLM (*Xenopus laevis*) and 4PE5 (*Rattus norvegicus*) structures are shown in Fig. 3.3A and B, respectively. The surface along the channel is also depicted using HOLE in Fig. 3.3A and B in the order of 4TLM and 4PE5. Two girdles appeared on the pore profiles of the structures. One of them, named as TMD top girdle, is observed at the top of the M3 helix near the extracellular side (between 30-40 Å in the Z axis in Fig. 3.3C). The other one, TMD bottom girdle, is observed at the lower end of the M2 helix. It is observed at 65 Å in the Z axis in Fig. 3.3C. The second girdle region in the X-ray structure of *Rattus norvegicus* is missing. The M2 pore lining helix in the channel has not been resolved in this structure and the TMD bottom girdle corresponds exactly to this point. Overall analysis showed that the TMD structure of *Xenopus laevis* is more detailed and complete compared to the structure of *Rattus norvegicus*.

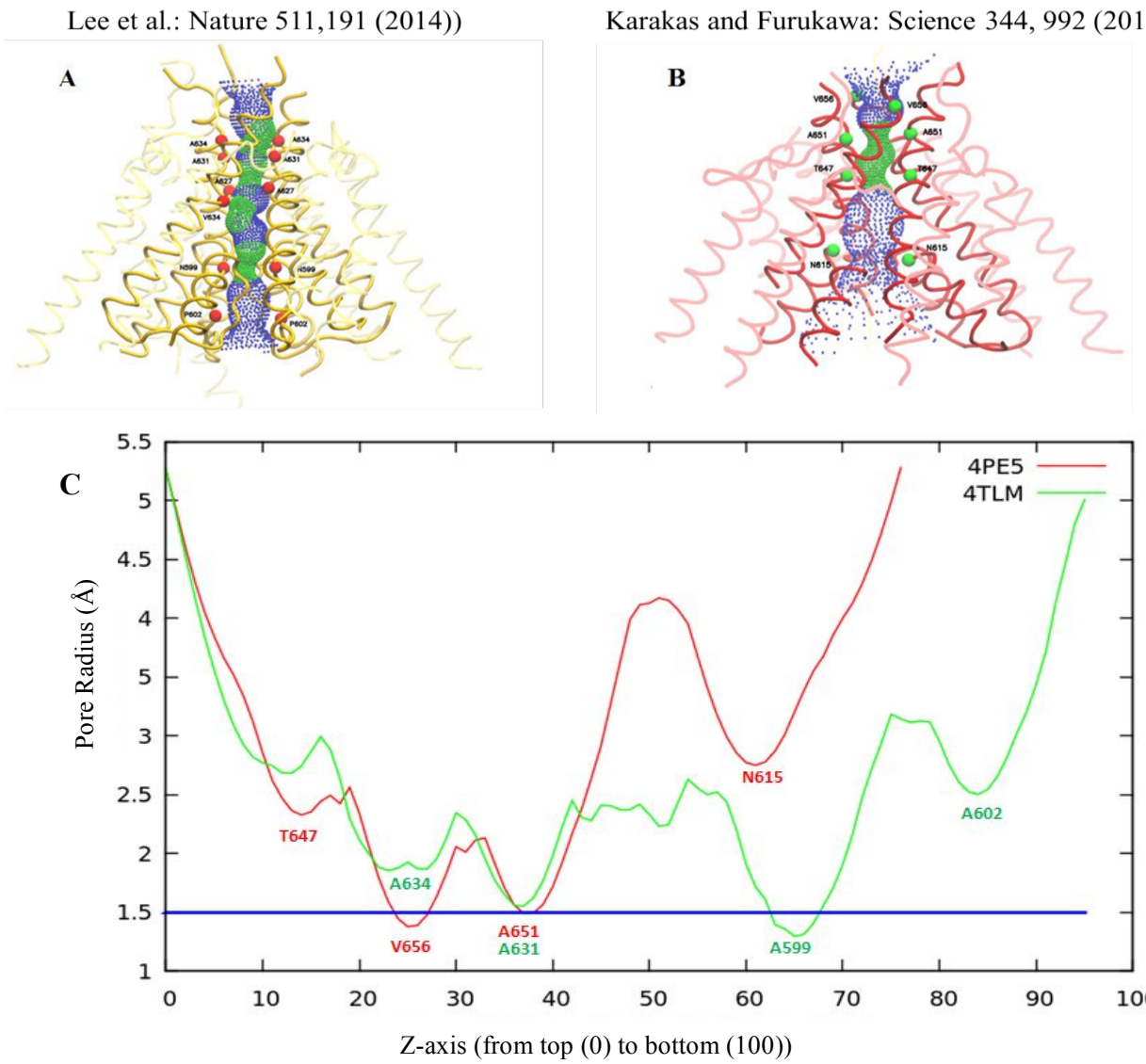


Figure 3.3. *Pore profiles of TMDs in structures* ***(A)*** *4TLM (yellow) and* ***(B)*** *4PE5 (red). The small green dots show the narrow regions of the channel and the blue dots show the wider regions. The helices forming the ion channel are shown in bold colors. Pore lining residues are shown as red beads for 4TLM and green beads for 4PE5.*

3.2. Loop Modeling

Homology model of the receptor has been completed by Demir and Essiz (Demir & Essiz, 2017). Two different modeling strategies were applied. The first one was modeling the protein independently by taking each individual subunit at a time, then combining these subunits to form the tetramer structure. The second one was modeling the protein as a tetramer structure from the tetramer template.

When chains are modeled individually, more accurate models in terms of structural quality can be obtained since the homology modeling is an optimization problem and when the problem becomes smaller better refined models can be obtained. However, in an intact structure, the chains in the structure are not independent from each other. Namely, when modeling the structure as a tetramer, neighboring residues in different subunits can interact in a more accurate way and they are arranged in the structure correctly during optimization.

In the work of Demir and Essiz, the orientation of the side chains in tetrameric structure modeled with loops has been tested for clashing residues via using PROCHECK (Laskowski et al., 1993) program. Fig. 3.4A and B shows PROCHECK Ramachandran plot of the template and the structure obtained by modeling all chains together with loops. In the Ramachandran plots, the percentage of residues found in the allowed regions was 77 % for the model, while it was 94 % for the template structure.

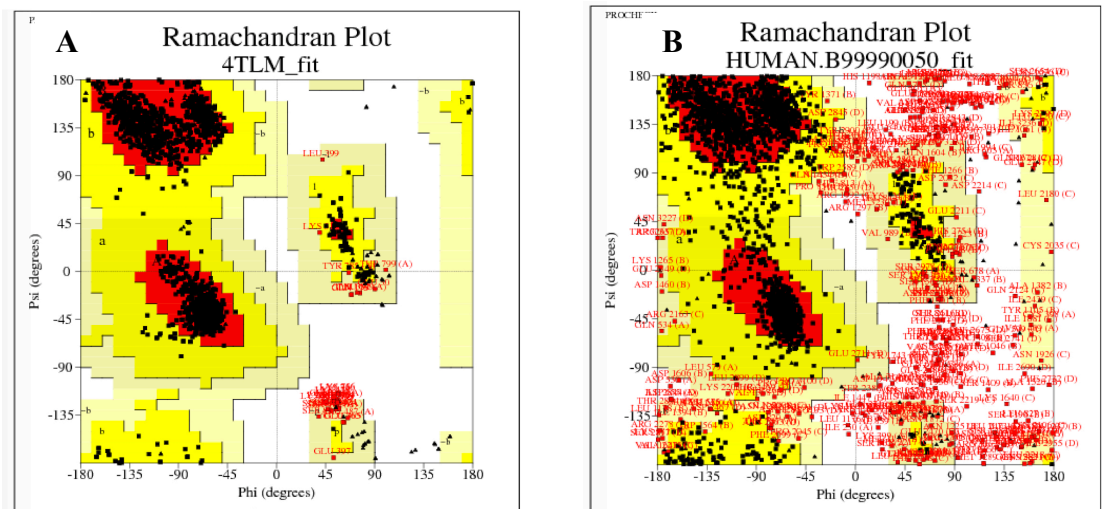


Figure 3.4. *Ramachandran plots.* ***(A)*** *The results of the template structure.* ***(B)*** *The results of the model obtained by modeling as tetramer with loops.*

Validating the template structure in addition to the model is also important since the template directly affects the accuracy of the model. In other words, if the template structure has problematic regions, those regions would be acquired from template as well. But in our case, model and template has some difference. That means model could have still been improved.

Moreover, there are several loop structures that are not experimentally resolved in the X-ray structure although these loop sequences are present in the protein sequence. In addition, some of the loops that exist in the human NMDAR sequence do not exist in the *Xenopus laevis* and *Rattus norvegicus* NMDAR sequences. These loops are located in important interface regions of the protein, i.e. loops connecting ATD to LBD or loops connecting LBD to TMD. Since these loops reduced the structural quality of the model when modeling with whole receptor, the new model is obtained by modeling without loop regions (Demir & Essiz, 2017).

Ramachandran graphic values are calculated for the model which does not have loops with using the PROCHECK (see Fig. 3.5) (Laskowski *et al.*, 1993). The percentage of residues found in the allowed regions of the structure modeled without loops has increased to 90 %. This model, which is more accurate in terms of geometric and secondary structures, was used in a different loop modeling program.

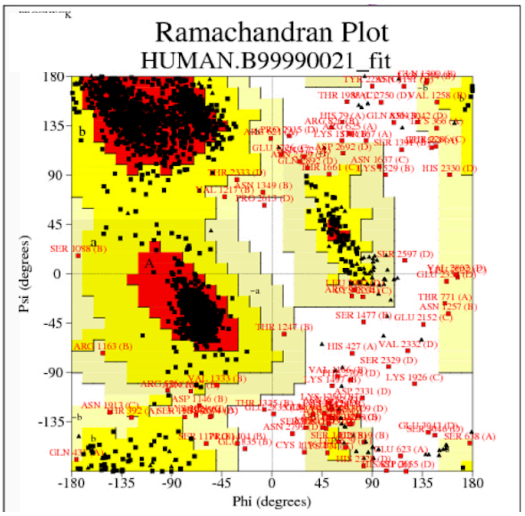


Figure 3.5. *Ramachandran plot of the structure modeled as tetramer without loops.*

Consequently, it is decided to model the protein as tetramer by removing the loops that do not exist in X-ray structure (Demir & Essiz, 2017).

Basically, one loop between transmembrane helices M1 and M2, two loops linking the ligand binding site to the transmembrane domain, and one loop between the ATD and the LBD are subjected to a remodeling step for each monomer (in Fig. 3.6, the missing loops in two subunits are displayed with blue and red colors). In this step, rest of the protein are fixed.

In the loop modeling part, the loop modeling program ROSETTA (Rohl et al., 2004) and MODELLER (Šali & Blundell, 1993) are used and compared for the quality of the obtained results.

There are 100 loop structures obtained from each program. DOPE scoring function of MODELER is used to evaluate loop structures. Results for the $GluN1_1$ subunit are shown in Fig. 3.7. This chain is picked since it is the most complicated chain in terms of the number of missing loops.

In Fig. 3.7A, the x-axis is the amino acid index and the y-axis is the DOPE score per amino acid. The graph is plotted for the loop structure modeled between M2 and M3 helices in $GluN1_1$. The difference only appears in this loop region and the lower scores represent the better models.

When the results of loop structures which are produced with ROSETTA and MODELLER are compared, ROSETTA is producing loops which have better scores (Fig. 3.7A). Please note that scoring functions in a particular program generally have better results for the structures that are generated in the same program. However here both ROSETTA and MODELLER scoring functions displayed better results for loops modeled in ROSETTA.

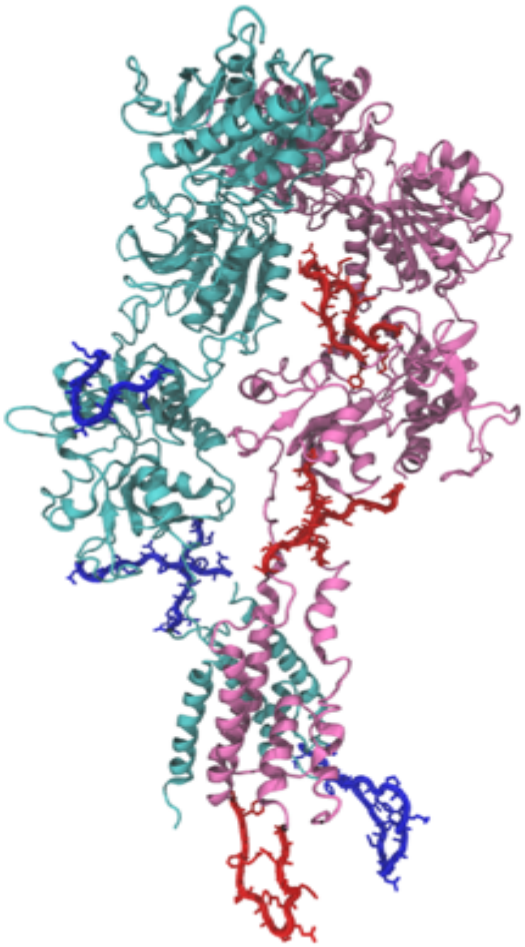


Figure 3.6. *Dimeric structure of NMDAR. GluN1 is shown as pink and GluN2A as light blue cartoon representation. Red and blue loops represent the missing loops to be modeled in these two chains.*

Then the missing loops in all the remaining subunits are completed by creating 100 models for each loop in Rosetta and choosing the best according to DOPE score.

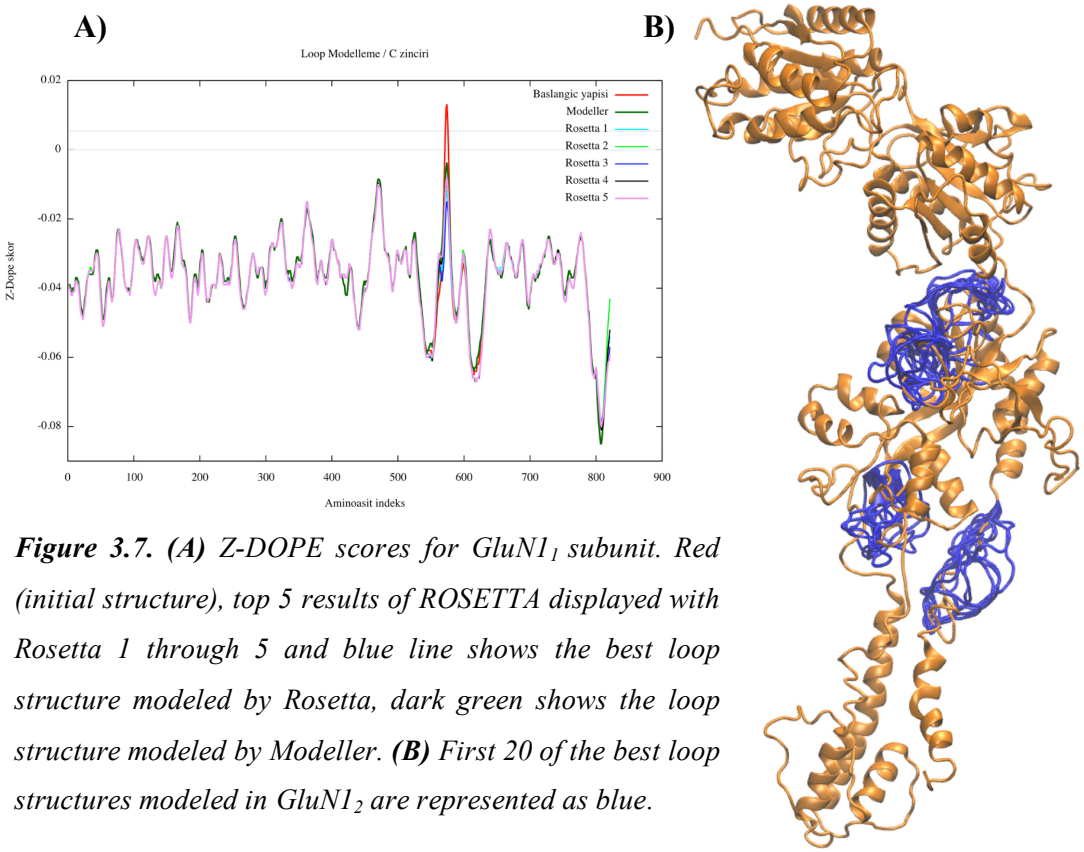


Figure 3.7. (A) *Z-DOPE scores for GluN1$_1$ subunit. Red (initial structure), top 5 results of ROSETTA displayed with Rosetta 1 through 5 and blue line shows the best loop structure modeled by Rosetta, dark green shows the loop structure modeled by Modeller.* ***(B)*** *First 20 of the best loop structures modeled in GluN1$_2$ are represented as blue.*

3.3. Molecular Dynamics Simulations

3.3.1. Equilibration Step

After obtaining ligand-free model structure from the loop modeling step, MD simulation of ligand-free (Apo) structure is initiated using the Protocol 1 shown in Table 2.1. Then, the simulation of ligand-bound structure is started using the Protocol 1 after the best scored ligand in AutoDock is docked into the binding site (see Section 2.6.1).

The structural stability of the models is further analyzed throughout molecular dynamics study. First, the root-mean-square deviation (RMSD) of Cα atoms from their starting positions is calculated for both simulation. The RMSD graphs presented a problem that we came across with the ligand-bound simulation. A jump is observed to ~8 Å in RMSD around 80 nanoseconds in the ligand-bound simulation run with Protocol 1 (shown with blue line in Fig. 3.10 and see Table 2.1 for the protocols used).

From the visual inspection of the simulation results, it has been observed that the ligand in $GluN2A_2$ started to move out from the binding pocket around 80 nanoseconds. Fig. 3.8 A shows how the ligand glutamate moves throughout the trajectory. In Fig. 3.8 A, $GluN2A_2$ is shown with green, the red vdW representation shows the first and last positions of the glutamate, and the black lines show the way it follows through the simulation. Additionally, change in the position of center of mass of glutamate throughout the simulation with respect to the X-axis (upper right), Y-axis (mid right) and Z-axis (lower right) is shown in Fig. 3.8 B, C and D respectively.

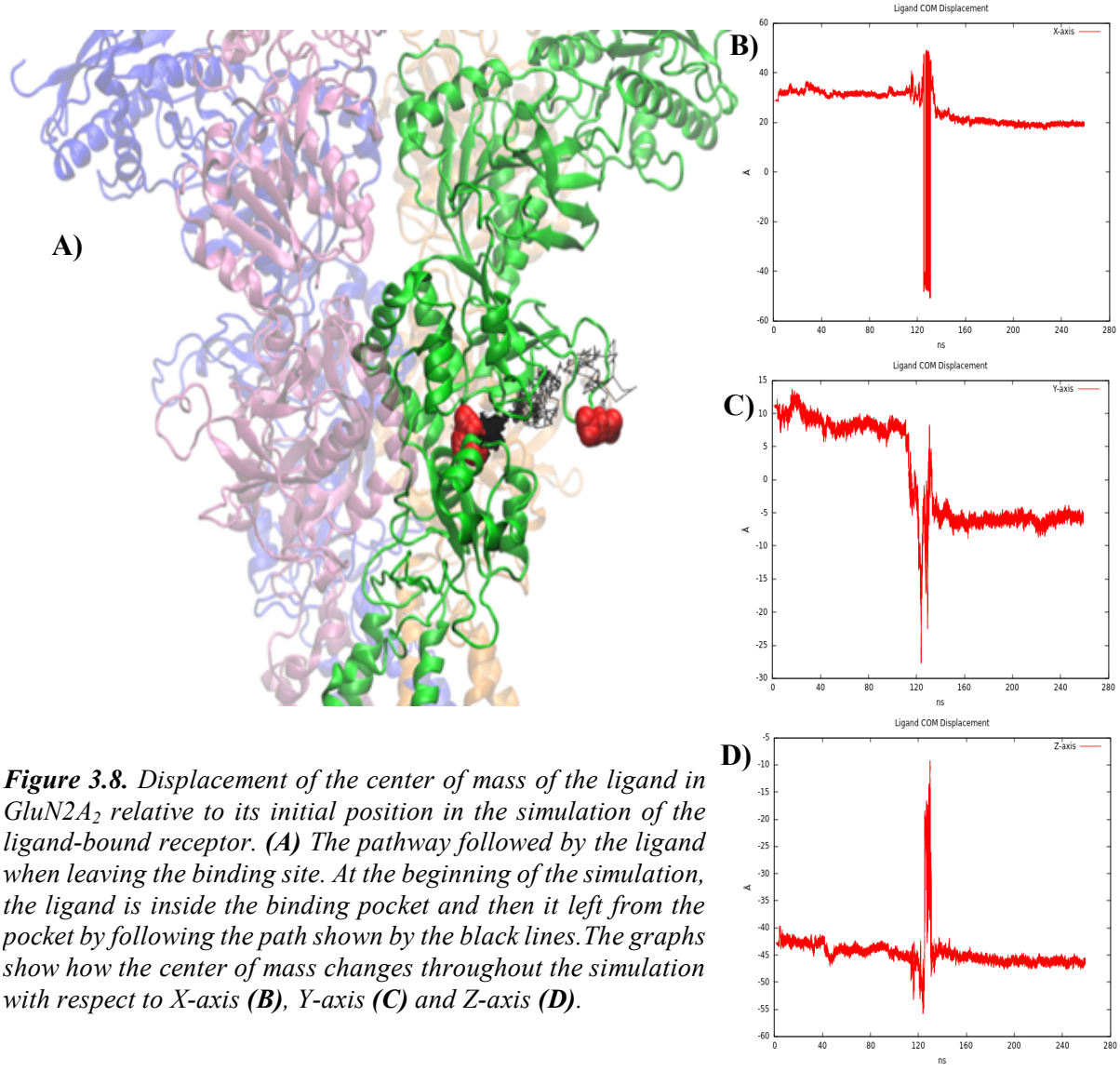


Figure 3.8. *Displacement of the center of mass of the ligand in GluN2A$_2$ relative to its initial position in the simulation of the ligand-bound receptor.* ***(A)*** *The pathway followed by the ligand when leaving the binding site. At the beginning of the simulation, the ligand is inside the binding pocket and then it left from the pocket by following the path shown by the black lines.The graphs show how the center of mass changes throughout the simulation with respect to X-axis* ***(B)****, Y-axis* ***(C)*** *and Z-axis* ***(D)****.*

To solve this problem, the poses of the ligand molecules sorted by AutoDock results are examined first. The best scored orientation of the glutamate molecule in GluN2A$_2$ and orientation of which in the X-ray structure of the NMDA receptor is shown in Fig. 3.9 A and B, respectively. It is observed that the best scored pose rotated 180 degrees in the binding pocket with respect to the pose in the X-ray structure. However, orientations of the best scored poses in other chains are similar to those observed in X-rays structure.

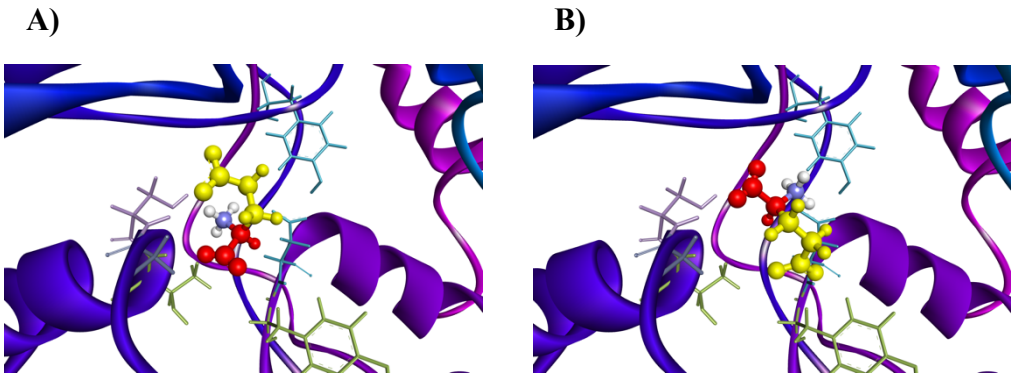


Figure 3.9. *Glutamate molecule in ligand binding domain.* ***(A)*** *Orientation of glutamate in LBD of the model.* ***(B)*** *Orientation of the glutamate in LBD of X-ray structure. Glutamate molecule is shown in yellow, red and light blue for the clarity of the orientation.*

Instead of using the best pose from AUTODOCK, a new ligand-bound simulation is run by docking the second-best pose, which has a similar orientation in the X-ray structure, into GluN2A$_2$. In the second simulation of ligand-bound structure, a new protocol (Protocol 2 shown in Table 2.1) is also used in order to ensure the ligands to stay in the binding pockets. Mainly, Protocol 2 has higher restrained simulation times compared to Protocol 1.

Finally, both apo and ligand-bound simulations have been updated to be run by Protocol 2 as a result of this change in the ligand-bound simulations.

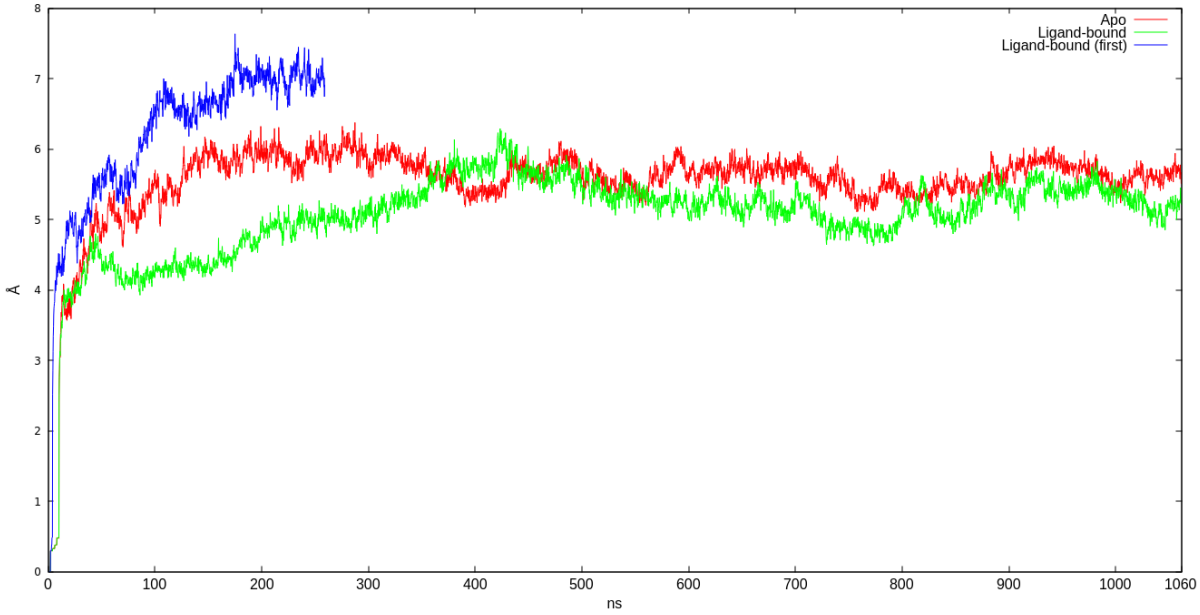


Figure 3.10. *Ca RMSD plot of protein in apo-Protocol 2 (red), ligand-bound-Protocol 2 (green) and ligand-bound-Protocol 1 (blue) simulations. In the ligand-bound-Protocol 1 simulation, ligand left the binding pocket.*

According to the RMSD graph, apo simulation seem to reach to a plateau after 120 ns (shown with red line in Fig. 3.10). For the ligand-bound simulation, RMSD slightly increased from 4 Å to 6 Å after 80th ns and then decreased to ~5.5 Å (shown with green line in Fig. 3.10). The apo simulation seems to reach to an equilibrated structure faster than the ligand-bound simulation.

Another parameter monitored during the equilibration step is the area per lipid head groups (APL). Since a pre-equilibrated POPC bilayer from VMD is used, the stability of the bilayer is also required to be monitored. This parameter is calculated by the GridMAT-MD (Allen et al., 2009) software. The experimental APL value for POPC is found to be 68.3 ± 1.5 Å^2 (Kučerka et al., 2006). Lipid molecules in this work have remained along the starting point throughout the simulation. Fig. 3.11 A shows how the APL changes throughout the melting of lipid molecules at 300 K. Similarly, Fig. 3.11 B shows the state of the lipid molecules before starting the simulation, Fig. 3.11 C shows the heated and collapsed state of the lipid molecules at the end of the first step of the Protocol 2 summarized in Table 2.1.

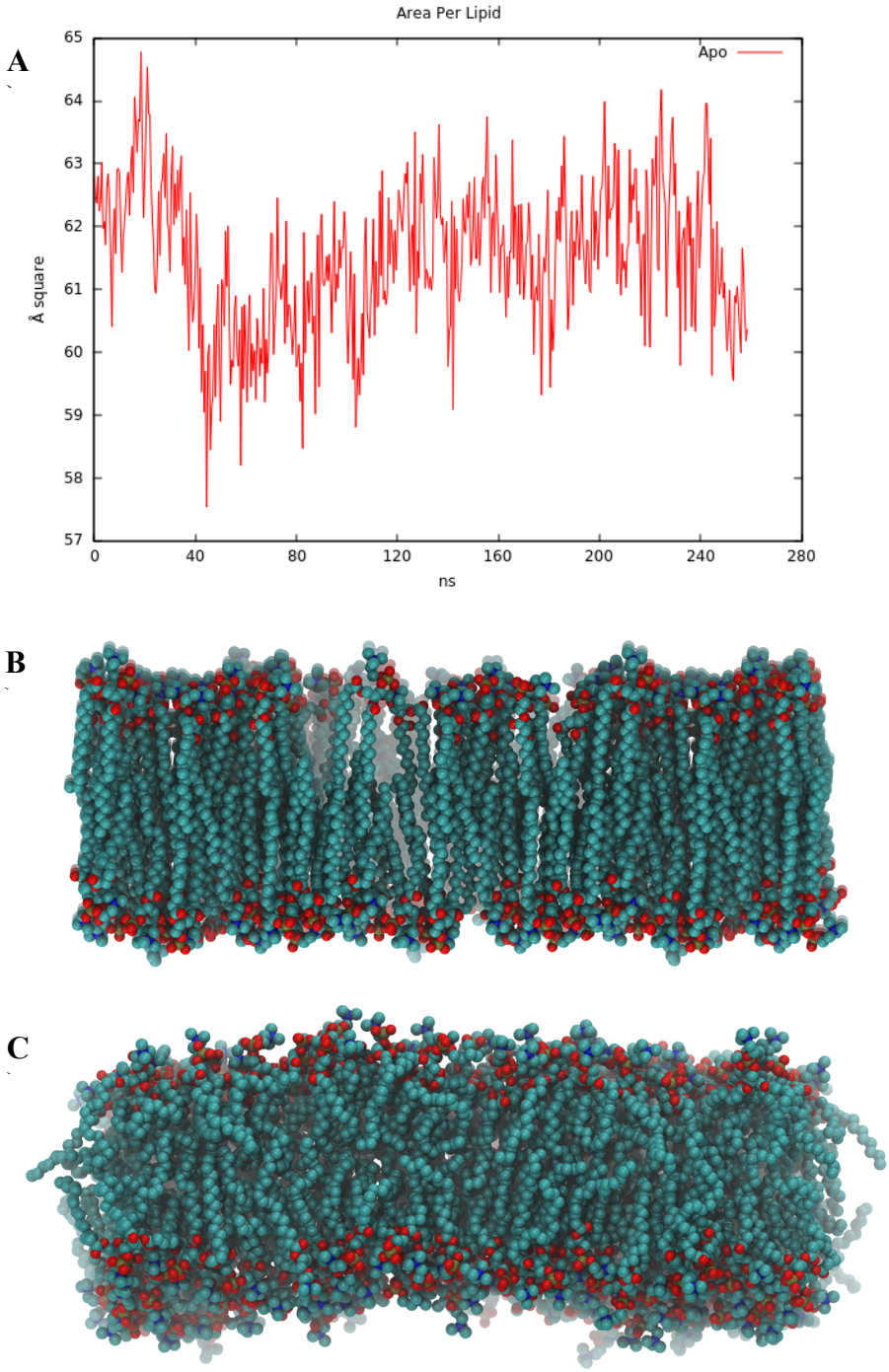


Figure 3.11. *Area per lipid graph and reaching to equilibrium of lipid bilayer.* ***(A)*** *Change of area per lipid throughout the apo simulation.* ***(B)*** *The representation of the unequilibrated lipid bilayer prior to simulation and* ***(C)*** *equilibrated bilayer at the end of the simulation. Carbon atoms in lipid tails (cyan), oxygen atoms in lipid head groups (red), nitrogen atoms (blue) and phosphorous atoms (tan) are shown as vdW representation. Hydrogen atoms are not shown in this figure for clarity.*

3.3.2. Analysis of the Equilibrated MD Trajectory

The simulations reached to equilibrium after 300 ns according to the RMSD plot. Therefore, the first 300 ns of the simulations has been discarded and the rest 700 ns simulations were used for the analysis part.

During apo and ligand-bound simulations, a number of structural parameters have been monitored. These parameters are mainly collected from the previous studies of the NMDA receptor and shown to be important on the mechanism of gating (Karakas & Furukawa, 2014; Lee et al., 2014). They consist of distances and angles between certain parts of the protein and they can be summarized as the following: (See Fig. 3.12 through Fig. 3.17)

- Distance between the center of masses of ATDs (Fig. 3.12 C).
- K216C Mutation Region in ATD: This mutation has implemented to reduce the structural flexibility of the protein in the crystallization step and important for restricting the ATD motion (Karakas & Furukawa, 2014; Lee et al., 2014) (Fig. 3.13A).
- TMD Girdle: The narrowest region and the gate of the channel near the extracellular side. The distance is calculated from Cα of the residues T638s on GluN1s and A631s on GluN2As (Fig. 3.12 A).
- TMD Bottom Girdle: This girdle is the second narrowest region in the channel toward intracellular side after the TMD Girdle. The distance is calculated from Cα of the residues N606s on GluN1s and N599s on GluN2As (Fig. 3.12 B).
- Distance between LBDs of GluN1 and GluN2A homodimers. The distance is calculated from Cα of the residues of ASP769 in GluN1 and ARG684 in GluN2A (Fig. 3.12 D).
- LBD Separation Angle: The angle is calculated between 2-fold axes passing through the $GluN1_1$-$GluN2A_2$ and $GluN1_2$-$GluN2A_1$ subunits (Fig. 3.13 B).
- Clamshell shaped D1-D2 Closure in Ligand Binding Domain (Fig. 3.14).
- Distance of the loop in LBD from the ATD helix (Fig. 3.15).
- Distance between the center of masses of ATD and LBD inter-domains (Fig. 3.16).
- TMD-M3 linker: This is the linkage between LBD and the ion channel. Distance is calculated from Cα of the residue pairs V646-I657 for GluN1 and I638-L649 for GluN2A (Fig. 3.17).

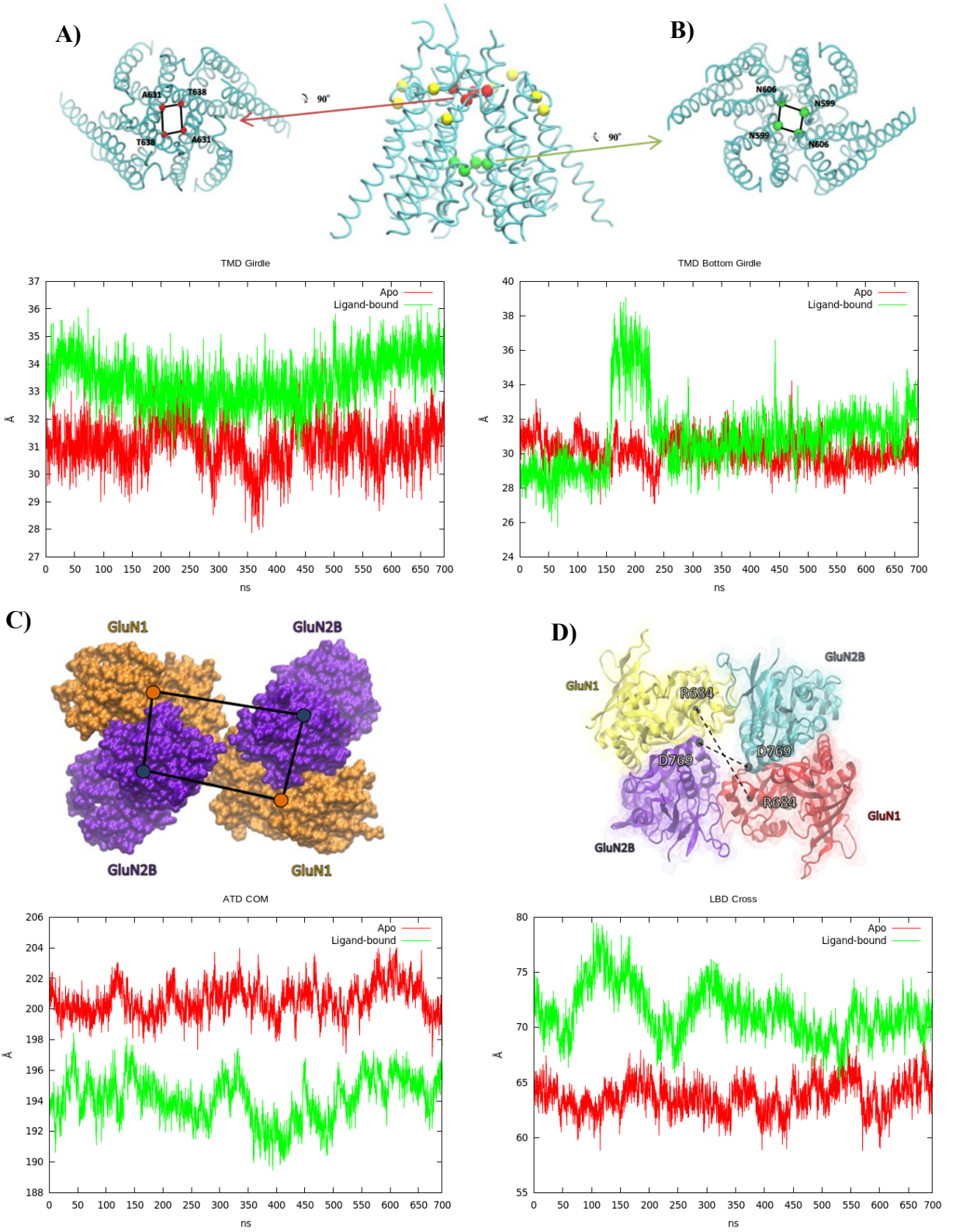


Figure 3.12. *Structural parameters measured for the simulations.* ***(A)*** *TMD Top Girdle. Selected residues are shown as bead representation in red and black rods represent the measured distances.* ***(B)*** *TMD Bottom Girdle. Selected residues are shown as bead representation in green and black rods represent the measured distances.* ***(C)*** *ATDs are shown as surface representation in orange for GluN1 and purple for GluN2A. Circles show the center of masses of each domain in same color codes and the black rods represent the measured distances.* ***(D)*** *LBDs are shown as cartoon representation with the color codes yellow and red for GluN1, cyan and purple for GluN2A. Residues selected are shown as black spheres and dashed lines show the distances measured. Distance graphs for the apo and ligand-bound trajectories are shown under each of the corresponding parameters.*

In order to see the conformational changes in the transmembrane (TMD) part of the receptor, the distances between the residues that are located in the two narrowest regions along the z-axis of the channel are measured. The upper narrow region consists of the residues T610, A1401, T2207 and A2992, and the lower narrow region consists of the residues N578, N1369, N2175 and N2960. The distances between the successive residues are measured and summed up for both upper and lower regions during the production runs. Summation of the distances between the residues throughout apo and ligand-bound simulations are shown in Fig. 3.12 A and B, respectively, together with the visual depiction of the residue locations on the TMD above the plots. In these graphs, there is no quite difference between the apo and ligand-bound simulations for the upper part of the TMD. However, the distance mostly remained smaller in apo simulation, which is around 33 Å while it is around 35 Å for the ligand-bound simulation. For the lower narrow region, the distance is 37 Å in apo simulation and 32 Å in ligand-bound simulation at the beginning of the trajectory. The distance remained almost constant for the apo throughout the trajectory. A huge jump is observed for the ligand-bound simulation at 160 ns from 32 Å to 42 Å which is an indicator of the opening in the lower girdle. Then, it decreased to the level of apo simulation (35 Å) at the end of the simulation.

The distance between the ATDs is calculated by choosing the center of mass of the ATD region of each chain. The constituent residues of ATDs are 1-371 in $GluN1_1$, 806-1173 in $GluN2A_1$, 1599-1969 in $GluN1_2$ and 2396-2752 in $GluN2A_2$. Then, these distances are summed up and plotted to observe how the center of masses moved during 700 ns part of the trajectory (Fig. 3.12 C). For the apo simulation (shown with red line in Fig. 3.12 C), the distance started from around 200 Å and remained stable for the most part of the trajectory except slight increases to 203 Å, around 120 ns to 220 ns. For the ligand-bound simulation (shown with green line in Fig. 3.12 C), the distance started from 194 Å, and ended at 192 Å. A larger distance is observed in the apo simulation compared to the ligand-bound simulation. The distances of the center of masses from each other (not the summation) in ATD are also shown in Supplementary Materials section (Fig. S1).

To monitor the changes in the LBD, a predetermined residue is selected from each chain and the cross distances between the residues in the same type of subunits (GluN1 and GluN2A subunits) are measured. Distances are calculated between R656 of the $GluN1_1$ and R2253 of the $GluN1_2$, between D1539 of the $GluN2A_1$ and D3130 of the $GluN2A_2$. Selected residues are shown as black spheres and calculated distances are shown in black dashed lines in Fig. 3.12D.

The graph showing the change of these distances is plotted for apo and ligand-bound simulation. LBD cross distance in apo simulation (shown with red line in Fig. 3.12 D) slightly fluctuated around 65 Å while in the ligand-bound simulation (shown with green line in Fig. 3.12 D), it started around 75 Å and raised up to 80 Å around 120 ns, and then decreased to 70 Å around 240 ns. This parameter in apo simulation looks less fluctuating compared to ligand-bound simulation in terms of the LBD cross distance and the distance mostly remained lower in apo meaning that LBDs have a more compact structure. The distances of the residues from each other (not the summation) in LBD are also shown in Supplementary Materials section (Fig. S2).

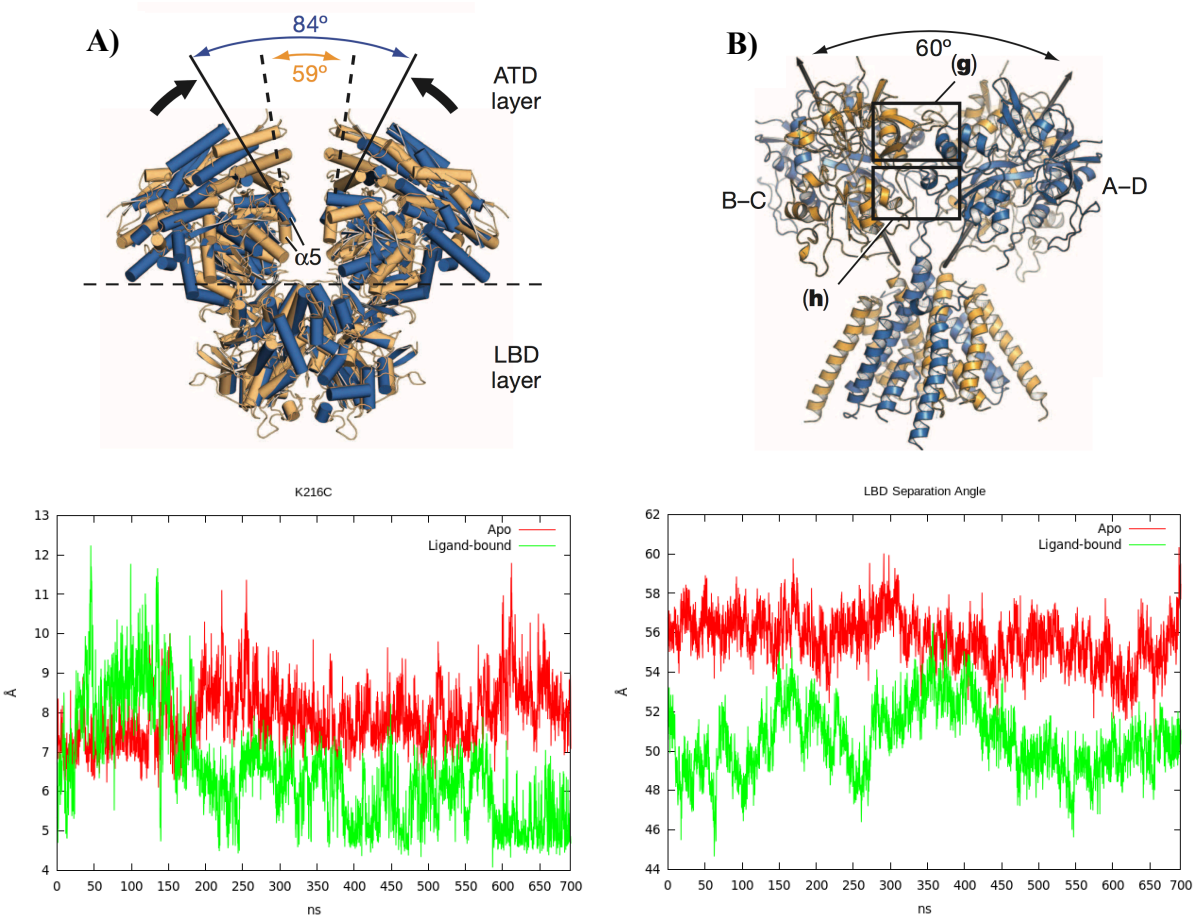


Figure 3.13. (A) *K216C mutation site. Cartoon representation of ATD and the distance between the lysine residues are plotted below.* ***(B)*** *LBD separation angle. Cartoon representation of LBD and TMD. Arrows represent the two-fold axis of LBD heterodimers. Change in the angle between the axes is plotted below the cartoon representations.*

Another parameter monitored is the distance between the lysine residues (997 in $GluN2A_1$ and 2586 in $GluN2A_2$) on ATD subunits. These two residues have been crosslinked to each other by substituting lysine residues with cysteine residues while crystallizing the structure (Lee et al., 2014). The distance between Cα atoms of those residues is plotted for the simulations (Fig. 3.13 A). For the ligand-bound (shown with green line in Fig. 3.13 A), the distance fluctuated up to 10 Å till150 ns and then decreased down to 5 Å for the rest of the trajectory. An overall larger distance (~9 Å) is observed in the apo simulation (shown with red line in Fig. 3.13A) after 150 ns. Namely, the distance between lysine residues, thus ATDs, are close to each other in ligand-bound simulation compared to apo.

Angle between the $GluN1_2$-$GluN2A_1$ and $GluN1_1$-$GluN2A_2$ LBD heterodimers are monitored throughout the equilibrated trajectory (Fig. 3.13 B). First two vectors are generated at the $GluN1_2$-$GluN2A_1$ and $GluN1_1$-$GluN2A_2$ heterodimer interfaces, which represent the twofold axes of LBD dimers, and then angle between two vectors are calculated. The angle fluctuated around 55-56 degrees for the apo simulation (shown with red line in Fig. 3.13 B) while it started from 52 degrees in the ligand-bound simulation (shown with green line in Fig. 3.13 B) and increased up to 55 at 150^{th} and 350^{th} nanoseconds, and then finally ended with 50 degrees. This angle is always smaller for the ligand-bound simulation and about 5 degrees closure is observed in ligand-bound simulations relative to the apo simulations.

D1-D2 Closure in Ligand Binding Domain

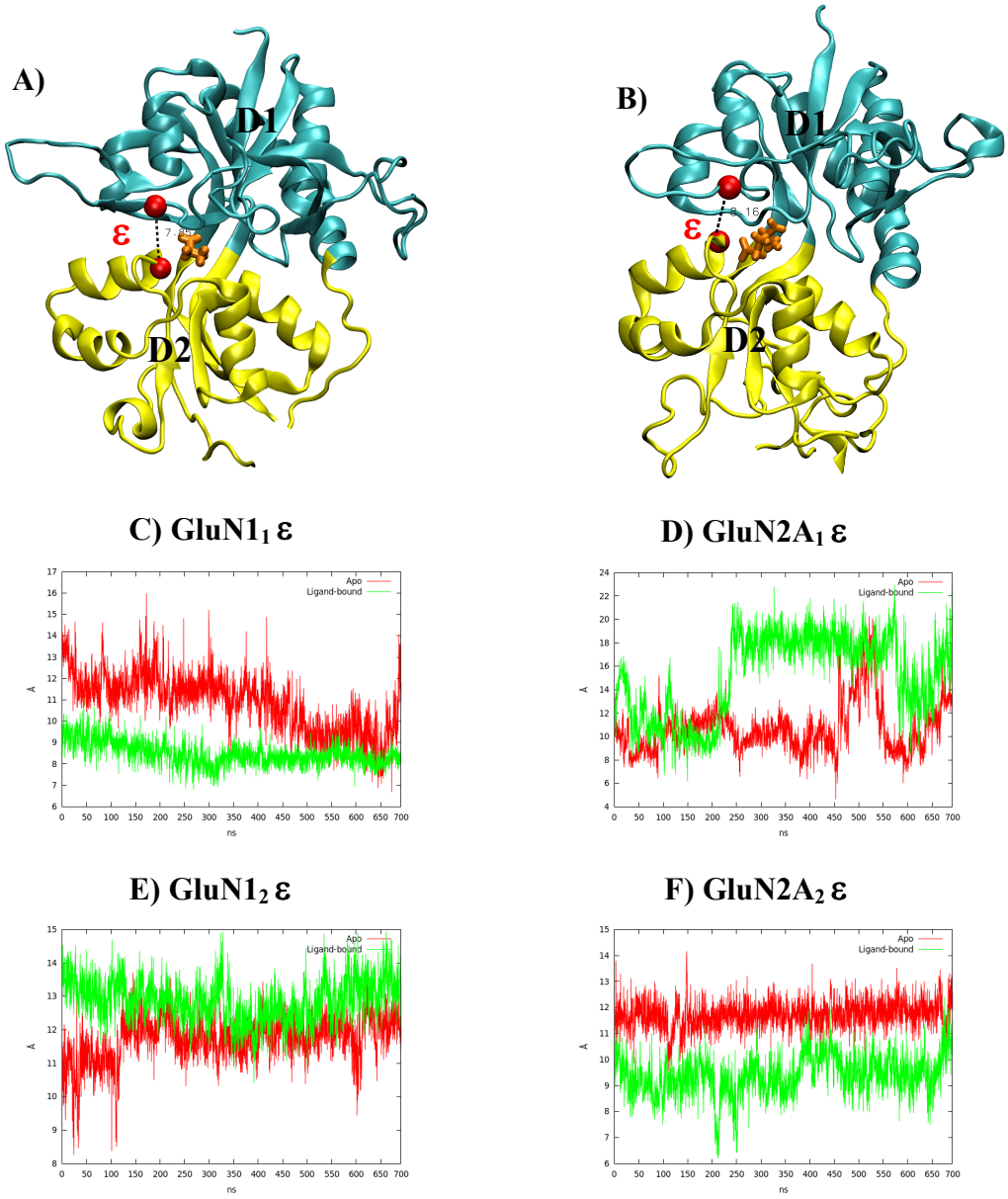


Figure 3.14. *D1-D2 closure measured in ligand binding domain.* ***(A)*** *Tube representation of GluN1$_1$ LBD and* ***(B)*** *GluN2A$_1$ LBD. Two sub-domains of LBD, D1 and D2 are colored as cyan and yellow on the crystal structure, respectively. Glycine in GluN1 and glutamate in GluN2A, both are represented as orange licorice model, bind to the cleft between D1 and D2. The center of mass pairs of the residues used in distance calculations are showed as red, blue and green beads for the distances ε_1, ε_2 and ε_3, respectively.* ***(C, D, E, F)*** *Plot of the ε_1, ε_2 and ε_3 distances in GluN1$_1$, GluN2A$_1$, GluN1$_2$ and GluN2A$_2$ throughout the trajectory, respectively. Apo simulation is showed as red and ligand-bound simulation is showed as green in the plots.*

As explained in the introduction section 1.2, D1 and D2 are the two sub-domains of the LBD, which closes like a clamshell when the ligand binds to the ligand binding sites. For different agonists, partial agonist and apo forms, different closure degrees have been observed (Armstrong & Gouaux, 2000). To analyze inter-domain motions of LBD for both apo and ligand-bound simulations, a distance between D1-D2 domains that have been developed in previous studies for determining the degree of closure of the clamshell (Mamonova et al., 2008) is also used. The parameter named as ε is the distance between Cα atoms of S652 and G451 (shown with red balls in Fig. 3.14 A and B). The distance is monitored for the equilibrated 700 ns trajectory for different chains in Fig. 3.14 C, D, E and F. Apparently, in $GluN1_1$ and $GluN2A_2$, larger ε are observed in the apo simulation. For $GluN2A_1$ subunit (Fig 3.14 D), larger distance is observed in ligand-bound simulation. However, apo simulation is also sampling the same large distance around 450 ns. Finally, in Fig. 3.14 E, the ligand-bound and apo simulations are sampling similar distances while the ligand-bound distance seems slightly larger.

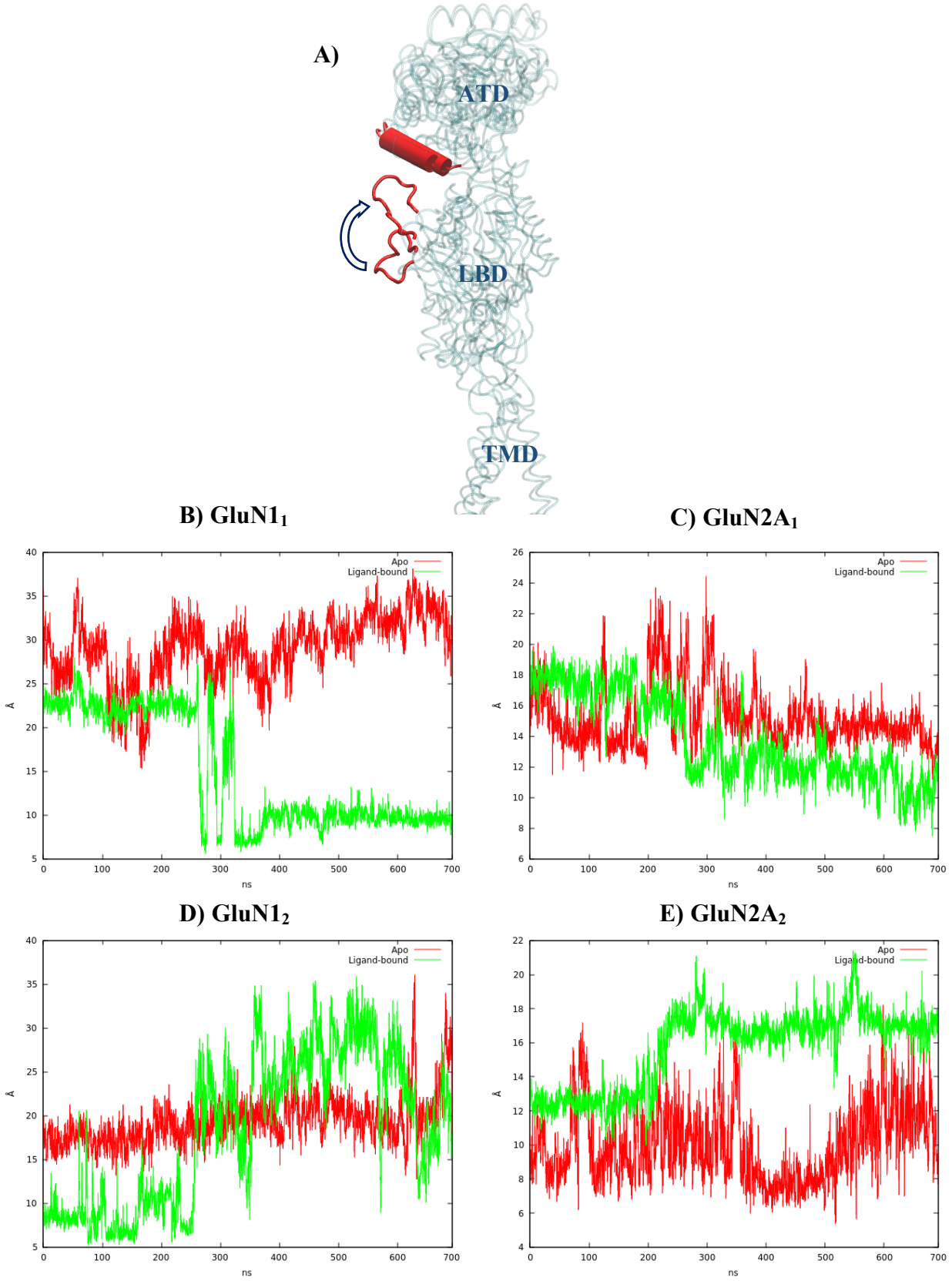


Figure 3.15. *Distance change comparison of the loops in LBD to ATD helix in apo and ligand-bound simulations.* ***(A)*** *Locations of the loop and helix are shown as cartoon representation on the top of the figure. Plots show the distances for each chain separately,* ***(B)*** *GluN1$_1$ on upper left,* ***(C)*** *GluN2A$_1$ on upper right,* ***(D)*** *GluN1$_2$ on lower left and* ***(E)*** *GluN2A$_2$ on lower right. Apo simulation is shown as red and ligand-bound simulation is shown as green.*

From the visual inspection of the trajectory, loop structure hanging from LBD has been observed to be very flexible in the ligand bound simulations when compared with apo. The graphs plotted above are produced by selecting the residues from that loop region in LBD and a helix in ATD. This flexibility might be important in the communication between ATD and LBD. The distances are calculated between the alpha carbons of the residues 187-422 in $GluN1_1$ (Fig. 3.15 B), 1167-1219 in $GluN2A_1$ (Fig. 3.15 C), 1785-2020 in $GluN1_2$ (Fig. 3.15 D) and 2756-2808 in $GluN2A_2$ (Fig. 3.15 E). In all of the chains, what is observed is that this loop structure is sampling two different conformations in the ligand-bound simulations when compared with apo simulations. In the $GluN1_1$, the loop gets closer to ATD and stays in contact with ATD at a distance of 6 Å. In the other chains also, the distance in the ligand-bound simulation is sampling at least two different conformations. In Fig. 3.15 A, two conformations of the loop observed in the ligand-bound simulation are displayed. Segments highlighted with red show the ATD helix and LBD loop region in $GluN1_1$.

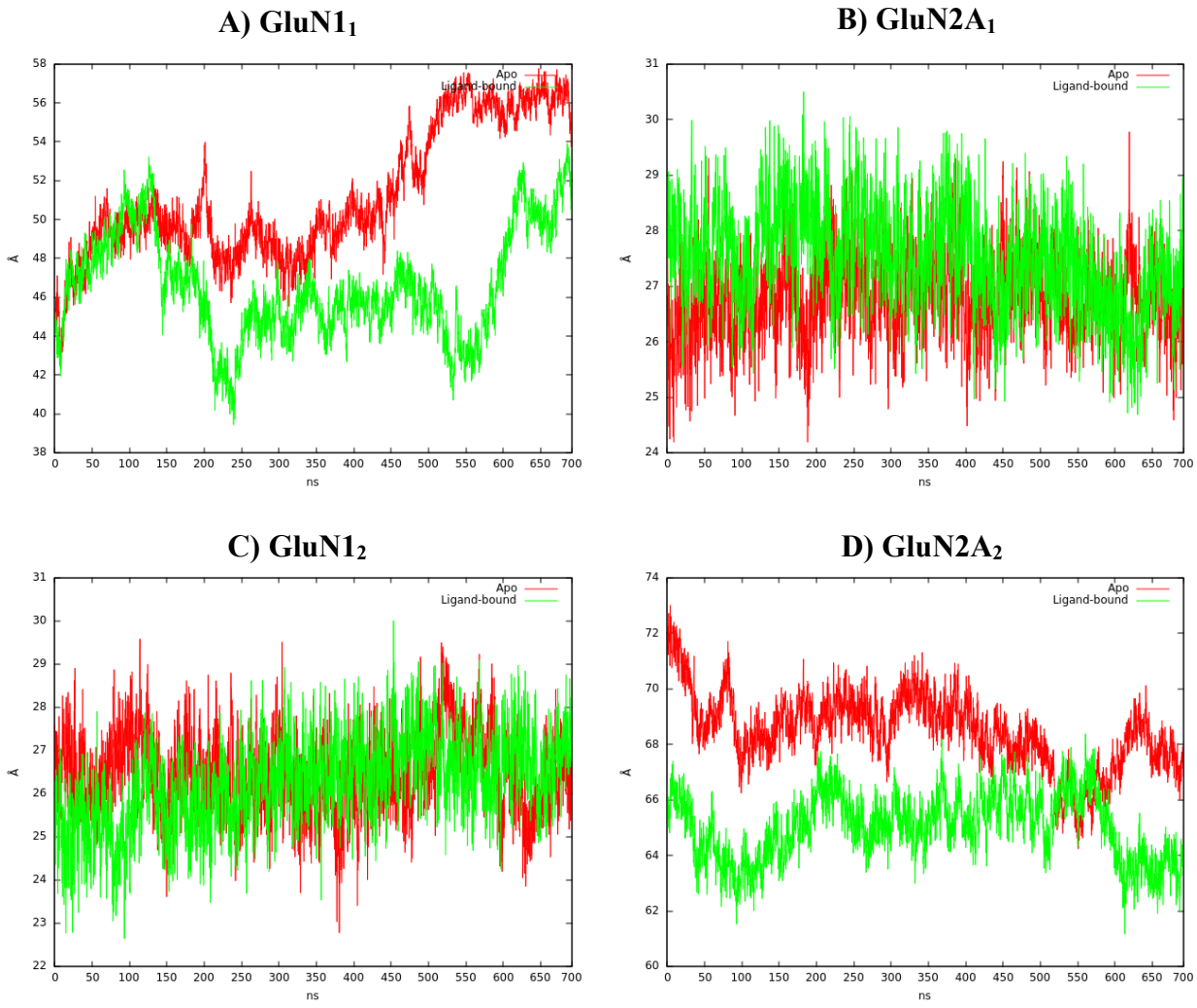


Figure 3.16. *Distance changes between ATD and LBD center of masses. Plots show the distances for each chain separately, (A) GluN1$_1$ on upper left, (B) GluN2A$_1$ on upper right, (C) GluN1$_2$ on lower left and (D) GluN2A$_2$ on lower right. Apo simulation is shown as red and ligand-bound simulation is shown as green.*

In addition to the specific loop region in LBD, the distance between ATD and LBD is monitored. To observe the distance change between ATD and LBD, center of masses of ATD and LBD are determined and then the distance between these centers is calculated for each subunit separately. In GluN1$_1$ and GluN2A$_2$ subunits, this distance is observed to be smaller in ligand-bound simulations (Fig. 3.16A and D) while it stayed almost same in GluN1$_2$ and GluN2A$_1$ (Fig. 3.16 B and C) subunits.

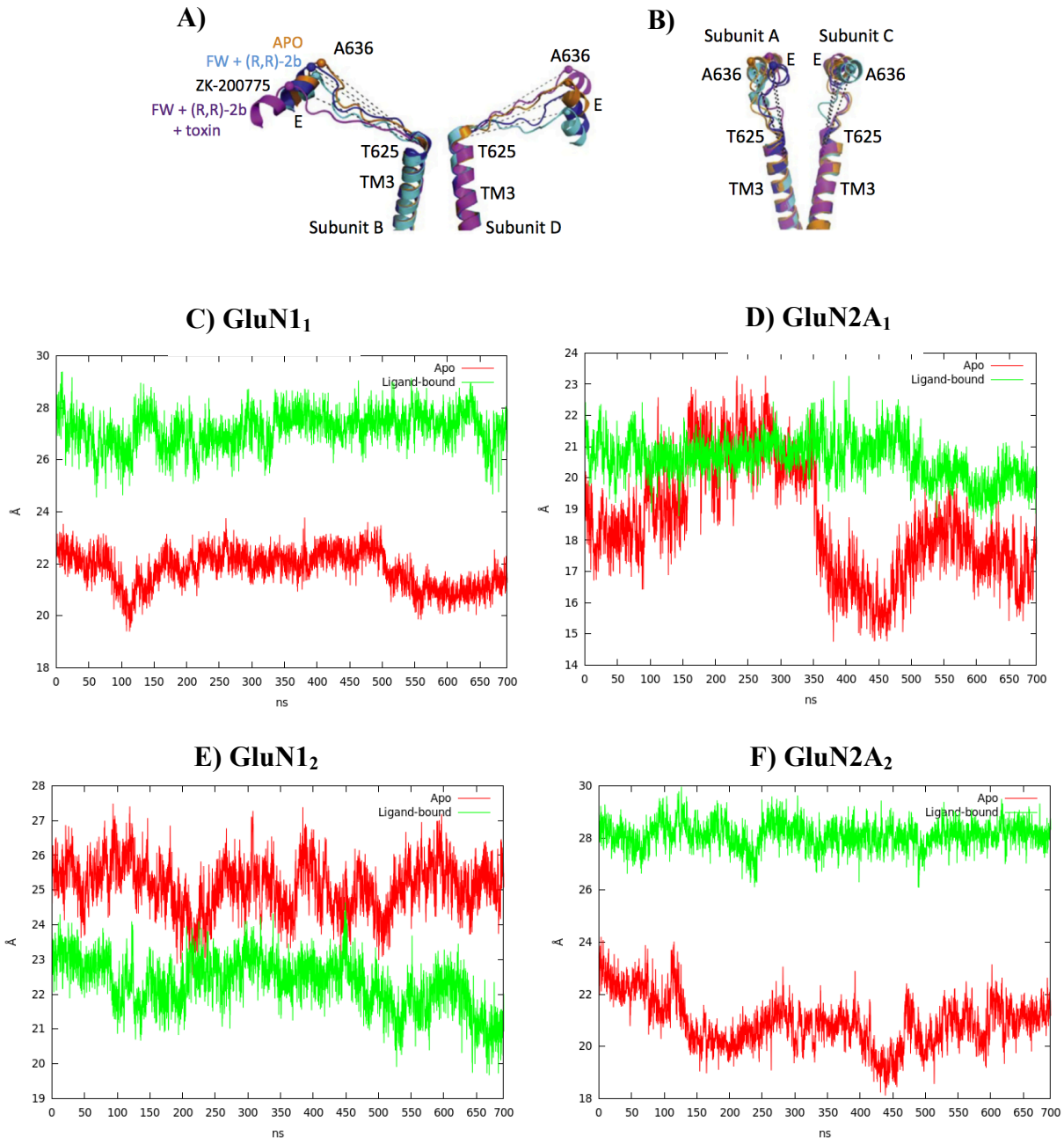


Figure 3.17. *Linker connecting the LBD and TMD M3 helices are showed as cartoon representation for each subunit, B and D subunits are on the left and A and C subunits are on the right. Distances of these linkers throughout the simulation are plotted below for each subunit.*

In Fig. 3.17, the linkers that connect LBD and TMD M3 helices and changes in the distances of these linkers are monitored throughout the trajectory. These linker distances are shown to be important in terms of pulling M3 helices up for opening the channel. The NMDAR channel structure resembles K^+ channel however it is upside down in terms of the channel axis. Namely, bundle crossing region which is linked to LBD domain in NMDAR is near intracellular end in K^+ channel. While K^+ channel is voltage sensitive, NMDAR is activated by ligand binding. The K^+ channel structure is much more studied in the literature and the pH sensor which is located at the intracellular domain of the channel is expected to be replaced by the movement of LBD upon ligand binding in NMDAR (Cuello et al., 2010). Namely, there is a pH sensitive helix in potassium channels which pulls the TM helices to open the helix bundling at the bottom of the channel. In the NMDAR case, this pulling motion is expected to be near the intracellular side of the channel.

The linker distances for the $GluN1_1$ and $GluN2A_2$ subunits are larger in ligand-bound simulations. The distances of these subunits in ligand-bound simulation (shown with green lines in Fig. 3.17) are similar and fluctuated around 28 Å whereas the distance of $GluN1_1$ in apo simulation (shown with red lines in Fig. 3.17 C) is a little bit higher (~22 Å) than that of $GluN2A_2$ in apo simulation (~20 Å) (Fig. 3.17 F). For the $GluN2A_1$ (Fig. 3.17 D), the linker distance in apo simulation started from the below of the ligand-bound simulation and then increased to the same level of ligand-bound. Then, a downward jump in apo simulation is observed from 21 Å to 15 Å at around 350 ns. In $GluN1_2$ (Fig. 3.17 E), the linker distance for apo simulation fluctuated at higher level (~25 Å) than that of ligand-bound simulation (~22 Å).

To summarize the results in this section, TMD girdle top which is expected to act the gate at the closed state of the channel remains larger in the ligand bound simulations (Fig. 3.12 A). Bottom girdle is staying same in both simulations however there is a brief opening for a length of approximately 100ns in the ligand bound simulations (Fig. 3.12 B). ATD subunits stays more compact while LBD subunits is more relaxed in the ligand bound simulations (Fig. 3.12 C-D). When the separation angles of ATD and LBD is considered, in both cases ligand is bringing these subunits closer to each other (Fig. 3.13 A-B).

When the individual ligand binding sites are considered, the distance between D1 and D2 sub-domains is staying smaller for only $GluN1_1$ and $GluN2A_2$ (Fig. 3.14). This distance is 5-8 Å when the clamshell is closed and 8-11 Å is accepted to be a semi-open and above 12 Å when

the clamshell is open (Mamonova et al., 2008; Okada et al., 2012). This is also reflected to the linker region in between TMD and LBD. In $GluN1_1$ and $GluN2A_2$, the linker distances are longer. That means when the ligand can close the clamshell type domains, this is reflected to the linker region (Fig. 3.14 and 3.17).

Again, just like LBD conformational change reflection to the TMD linker regions, when the distance between the centers of masses of ATD and LBD are considered ligand bound simulation display a shorter inter-domain distance, in other words more compact structure for $GluN1_1$ and $GluN2A_2$ (Fig. 3.16).

Both of the hydrophobic girdles stay bigger. The chains in which ligand induces a clamshell closure deduces a more compact ATD–LBD inter-domain distances along with longer TMD-LBD inter-domain linkers. This mechanism resembles the K^+ channel pH sensitive helix pulling mechanism for opening the hydrophobic girdle at helix bundle. Actually, it was not possible to observe a complete opening of the channel in the limited time of our MD simulations. However, the correlated motions in linker region, LBD clamshell closure and ATD and LBD inter-domain closure is observed in ligand-bound simulations only (Fig. 3.22 and 3.23).

Additionally, the separation angles intra LBD and intra ATD domains is smaller in ligand bound simulations and there is one loop region from LBD which is getting in very close contact with ATD in ligand-bound simulations (Fig. 3.15).

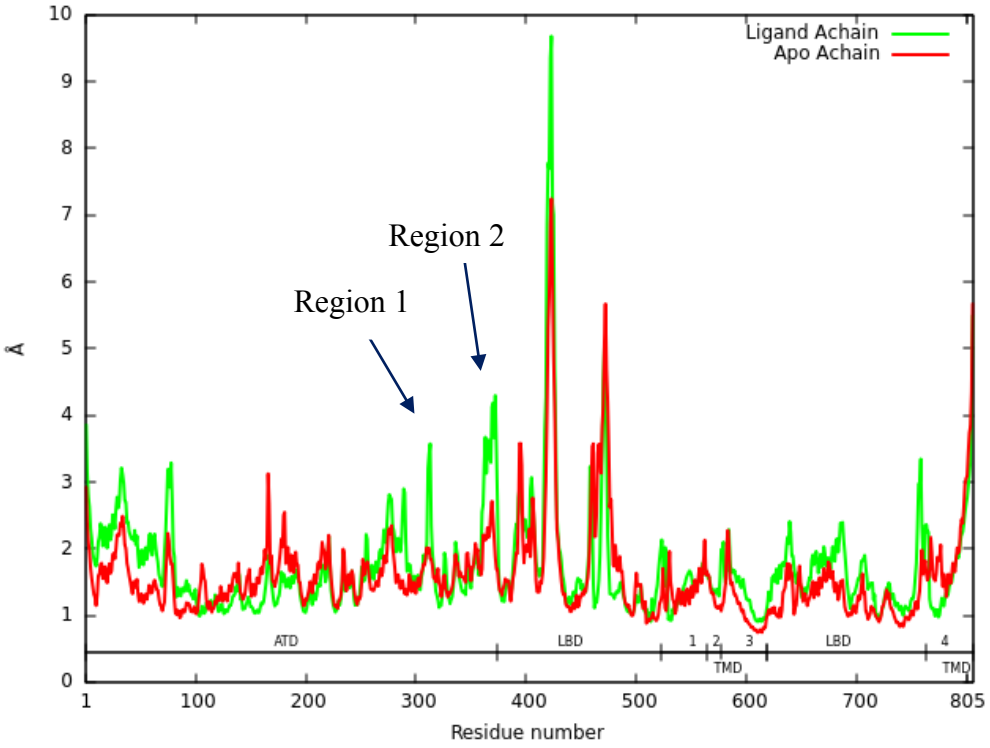


Figure 3.18. *Root-mean-square fluctuations (RMSF) of GluN1$_1$ Cα atoms from residues 1–805 for both apo (red) and ligand-bound (green) simulations.*

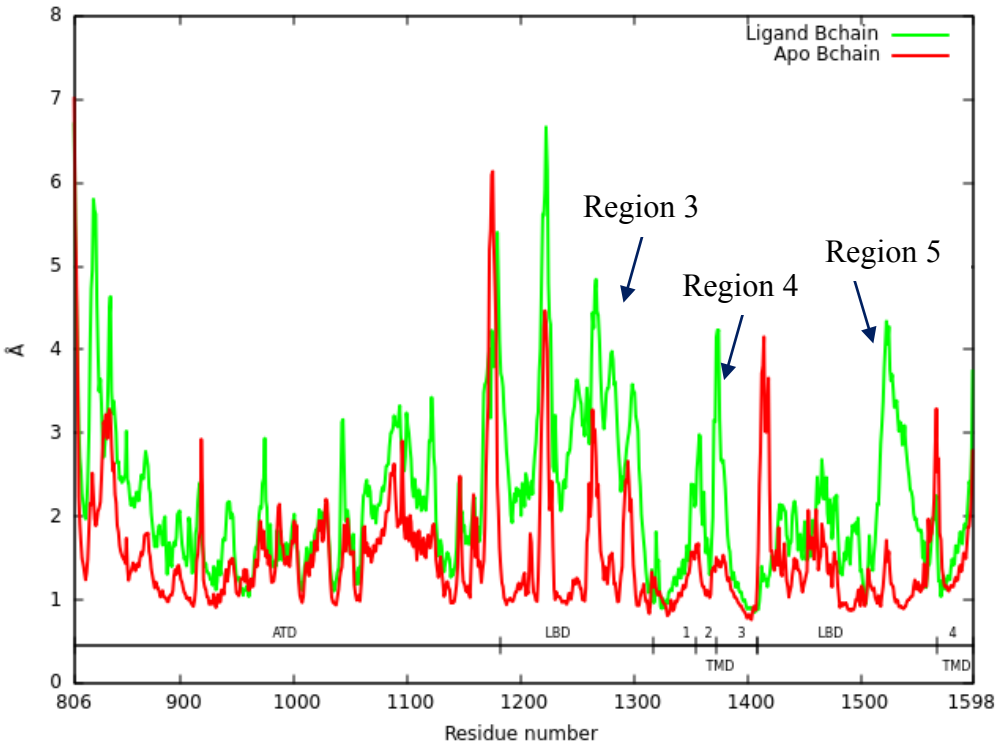


Figure 3.19*. Root-mean-square fluctuations (RMSF) of GluN2A$_1$ Cα atoms from residues 806–1598 for both apo (red) and ligand-bound (green) simulations.*

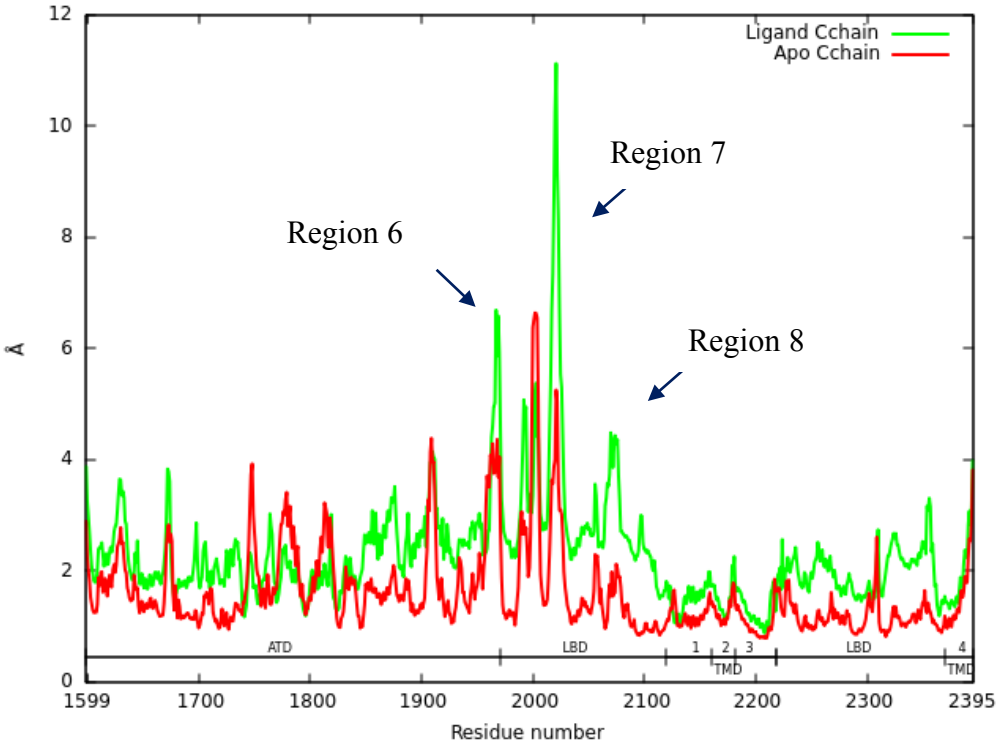


Figure 3.20. *Root-mean-square fluctuations (RMSF) of GluN1$_2$ Cα atoms from residues 1599 -2395 for both apo (red) and ligand-bound (green) simulations.*

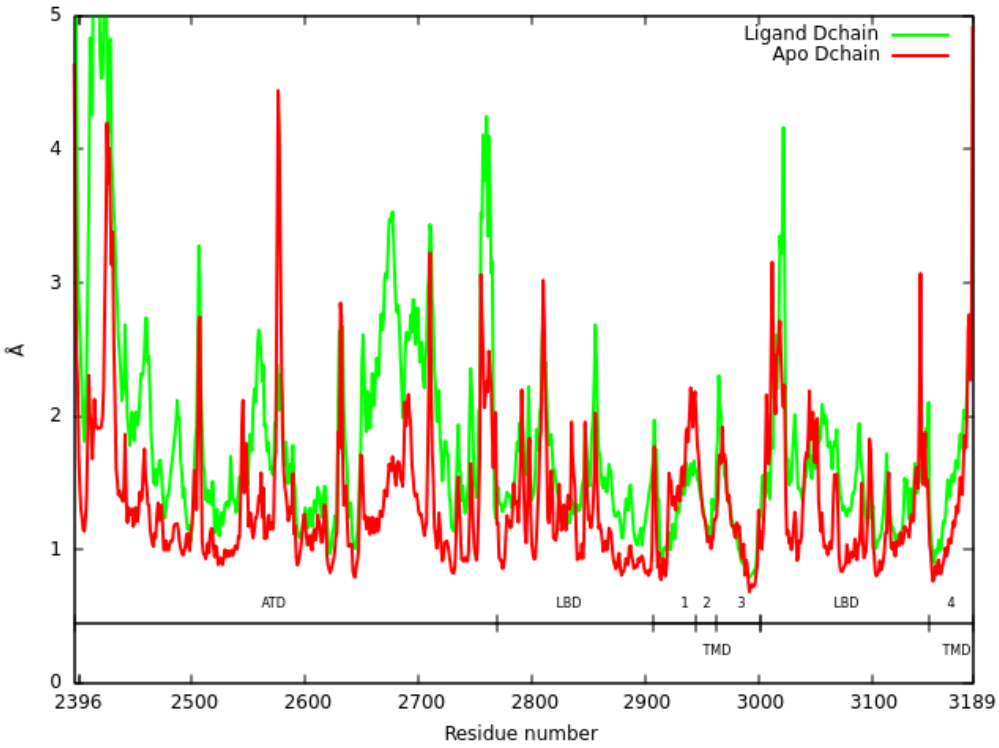


Figure 3.21. *Root-mean-square fluctuations (RMSF) of GluN2A$_2$ Cα atoms from residues 2396–3189 for both apo (red) and ligand-bound (green) simulations.*

In the RMSF graph of $GluN1_1$, a loop in the upper lobe of ATD (Region 1) and a loop in the lower lobe of ATD (Region2) displays a higher mobility in the ligand-bound simulation (Fig. 3.18). Region 2 corresponds to the linker connecting ATD to LBD.

According to RMSF plot of $GluN2A_1$, certain regions in the ligand-bound simulation show higher mobility compared to the apo simulation (Fig. 3.19). Especially, LBD of the ligand-bound simulation exhibits a noticeable flexibility in S1 region (Region 3). Also, a remarkable increase in S2 region of LBD near the M4 helix is observed in ligand-bound simulation (Region 5). Region 4 corresponds to two linkers M1-M2 and M2-M3.

Higher flexibilities are also observed in the $GluN1_2$ subunit of ligand-bound simulation, Region 6, 7 and 8 (Fig. 3.20). The first (Region 6) is the ATD-LBD linker region in the lower lobe of ATD that occurred in $GluN1_1$ as well. The second (Region 7) is the loop in the upper lobe of LBD, the distance of which to the ATD helix is also monitored separately in the previous section (Fig. 3.17). The third (Region 8) is a loop region that is located in the upper lobe of LBD and having a close contact with the lower lobe of ATD. Additionally, more flexible regions are also observed along all the S2 region of LBD that is between M3 and M4 helices.

According to RMSF plot of $GluN2A_2$, the whole ATD segments seems more mobile in the ligand-bound simulation, especially in the lower lobe of ATD and ATD-LBD linker regions (Fig.3.21).

3.4. Principal Component Analysis (PCA)

PCA is performed for the 700 ns equilibrated trajectories of apo and ligand-bound simulations after discarding the first 300 ns of the simulations. The correlations shown in Fig. 3.22 and 3.23 are the normalized covariance matrix values calculated by Equation 2.9. The system consists of 3189 residues and so, 9567 eigenvalues and eigenvectors are calculated with PTRAJ suite of AMBER MD package. The eigenvalues are sorted in descending order; thus, the first eigenvalue captures the largest amplitude motion of the protein. In the previous studies, it has been seen that the first 10 eigenvalues are enough to capture most of the fluctuations in the protein or the slowest motion of the protein (Hayward & de Groot, 2008; Issack et al., 2012).

In this study, first 10 eigenvalues cover 62 % of total fluctuations in apo simulation and 77 % of total fluctuations in ligand-bound simulation (See Table 3.1). Table 3.1 below is plotted in order to see the coverage of first 10 eigenvalues among all the eigenvalues calculated in the analysis. By using the first 10 eigenvectors and eigenvalues, the covariance matrix is calculated (Equation 2.9).

Table 3.1. *Percentage contribution of the first 10 eigenvalues among all the eigenvalues in PCA.*

Apo (700 ns)	**Ligand-bound (700 ns)**
2332.12	7255.37
730.86	1370.61
565.75	972.42
385.55	563.88
346.69	511.34
291.48	448.36
264.41	270.01
182.52	229.98
169.12	189.72
147.55	156.29
62.19 %	77.11 %

The results in Fig. 3.22 and 3.23, are normalized covariance matrix values based on Cα atoms and they represent how correlated motions are affected upon ligand binding into the LBD. In the correlation graphs, diagonal segments are for spatially close segments of the protein while off-diagonal segments show the correlations between distant segments. +1 in color bars on the right of the plots means the segments are moving in a fully correlated fashion while -1 means anti correlated movements. Figures are colored according to the degree of correlation where +1 is represented with red while -1 is represented with blue.

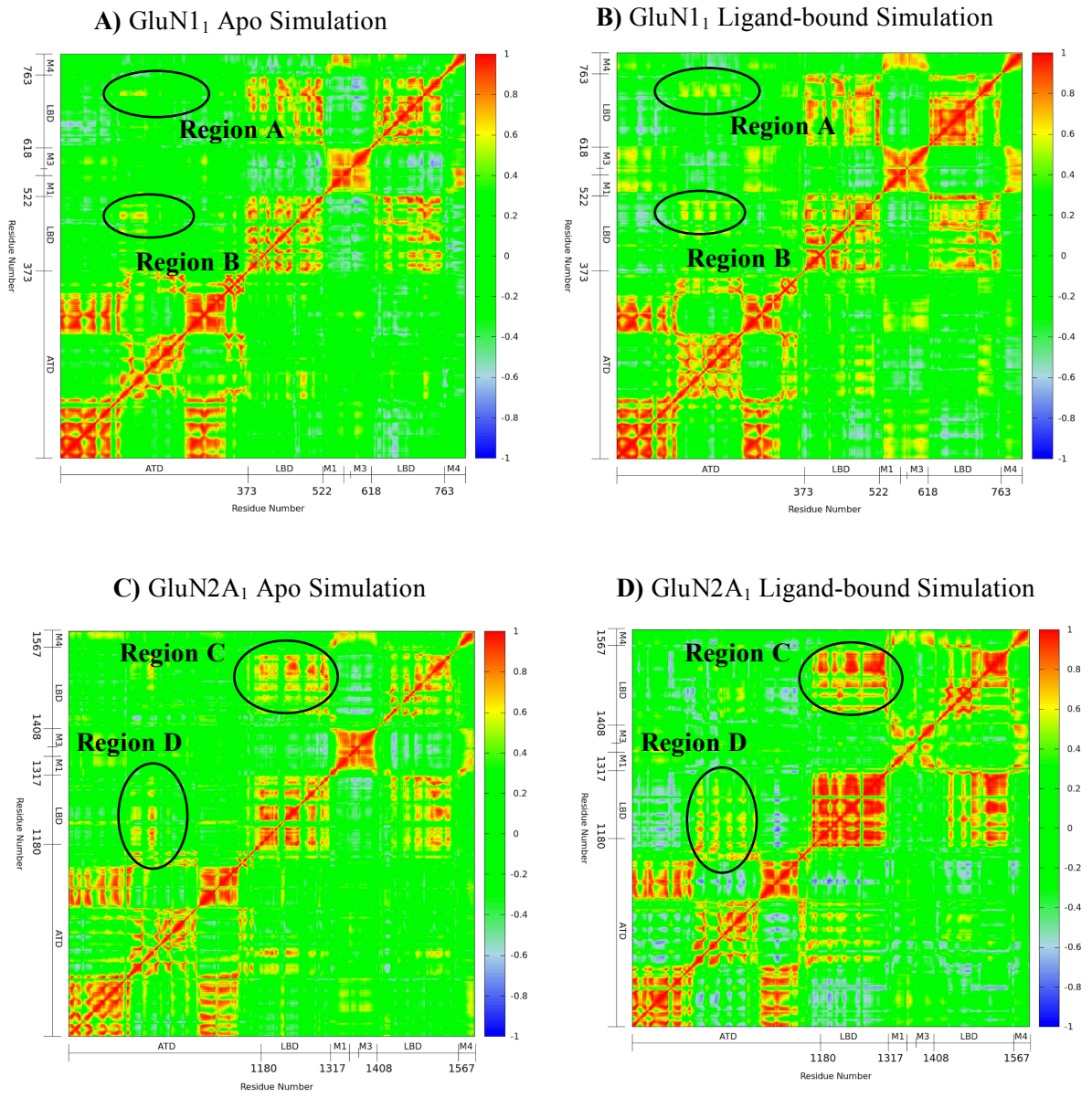


Figure 3.22. *Correlation plots of apo and ligand-bound simulations are represented for the GluN1$_1$ and GluN2A$_1$ subunits.* ***(A)*** *GluN1$_1$ Apo* ***(B)*** *GluN1$_1$ Ligand-bound* ***(C)*** *GluN2A$_1$ Apo and* ***(D)*** *GluN2A$_1$ Ligand-bound simulations are shown. Cross-correlation values range from -1 (blue) to 1 (red).*

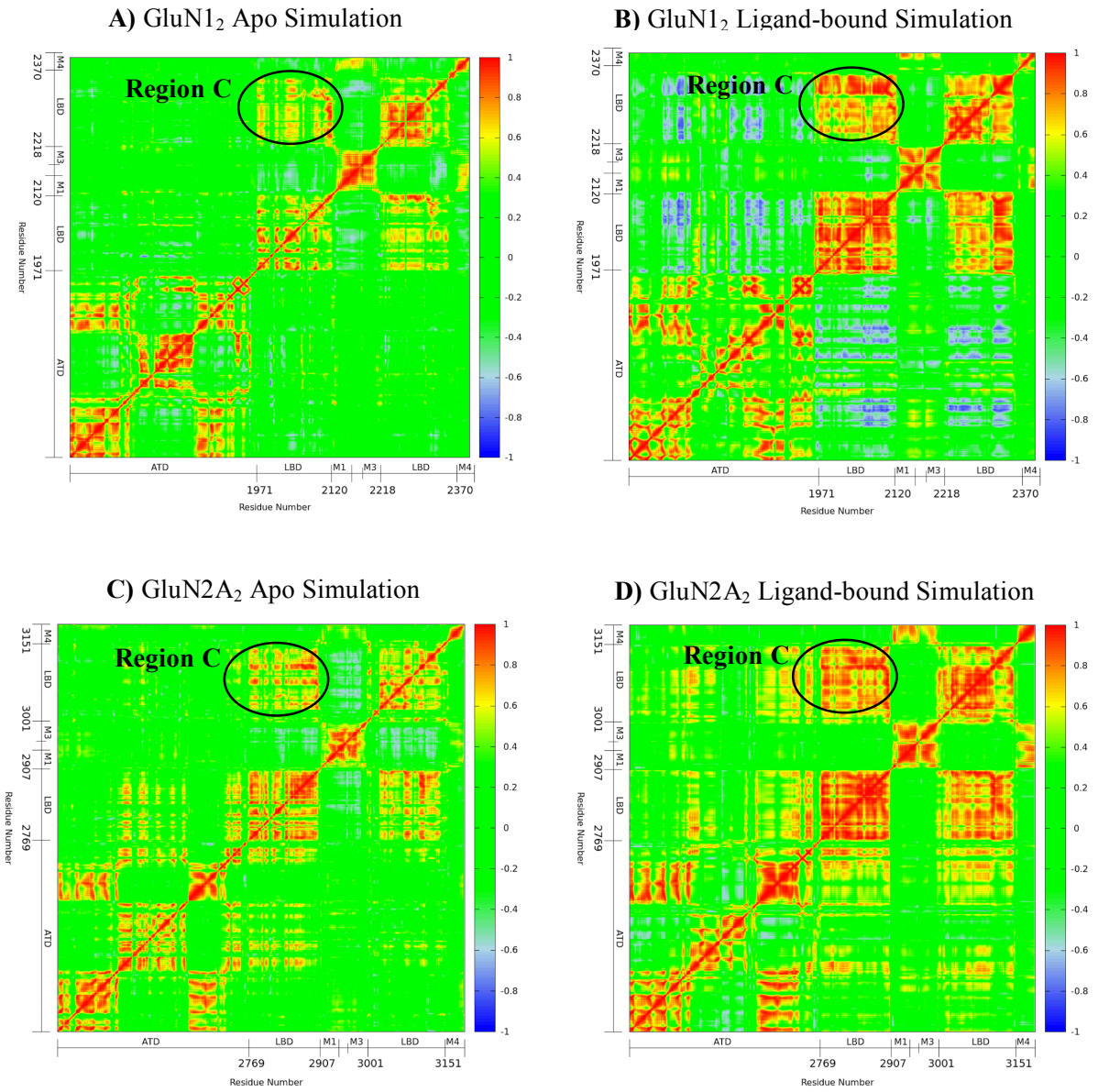


Figure 3.23. *Correlation plots of apo and ligand-bound simulations are represented for the GluN1$_2$ and GluN2A$_2$ subunits.* ***(A)*** *GluN1$_2$ Apo* ***(B)*** *GluN1$_2$ Ligand-bound* ***(C)*** *GluN2A$_2$ Apo and* ***(D)*** *GluN2A$_2$ Ligand-bound simulations are shown. Cross-correlation values range from -1 (blue) to 1 (red).*

According to the cross-correlation plots of $GluN1_1$, there are two remarkable increases in the correlated motions in the ligand-bound simulation. These regions correspond to the correlation of the same region of the ATD with two different regions of LBD, labeled as Region A and B in Fig. 3.22 A-B. The Region A shows the correlation between ATD and S2 of LBD and the correlated motion of this region increased in ligand-bound simulation. Region B shows an increase in the correlated motion of the same ATD segment in Region A and S1 of LBD in ligand-bound simulation. S1 is the loop connecting LBD to M1 and S2 is the loop connecting M3 to M4 helix in TMD.

Fig. 3.22 C and 3.22 D show the correlations in $GluN2A_1$ of apo and ligand-bound simulations, respectively. There are two noticeable regions in the ligand-bound simulation appearing upon ligand binding. The first one (Region C) shows the correlation between whole S1 and S2 regions of LBD. This correlation increased in ligand-bound simulation. The second one (Region D) corresponds to the regions ATD and S1 of LBD in the $GluN2A_1$. The movement of these regions are more correlated in ligand-bound simulation compared to apo.

In both of the subunits $GluN1_2$ and $GluN2A_2$, the same region observed in $GluN2A_1$ (Region C) showed higher correlation in ligand-bound simulation compared to apo. Additionally, in $GluN2A_2$, an increase in correlated motion between M1-M3 and M4 helices is also observed in the ligand-bound simulation. Moreover, as it can be seen from the increase of blue zones in correlation graph of $GluN1_2$ in ligand-bound simulation, the regions between ATD and both S1 and S2 regions of LBD (ATD-S1 and ATD-S2) showed less correlation in ligand-bound simulation.

3.5. PCA First Mode Comparison

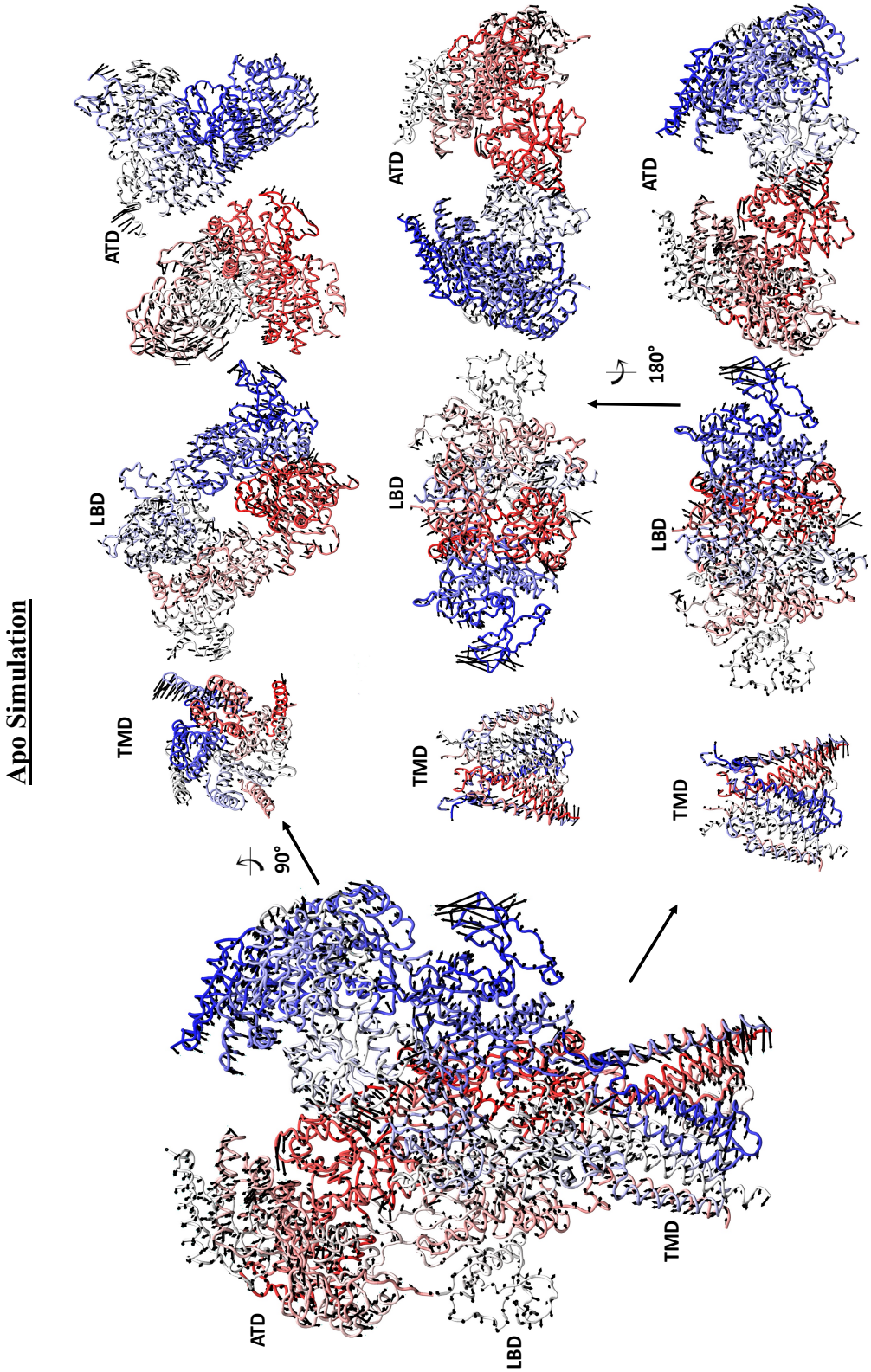


Figure 3.24. *First mode of apo simulation from PCA. Structure on the left represents the whole receptor. ATDs, LBDs and TMDs are also shown separately from different views on the right side of the figure. Black arrows on the structures show the movement of that region in the direction of that arrow.*

Ligand-bound Simulation

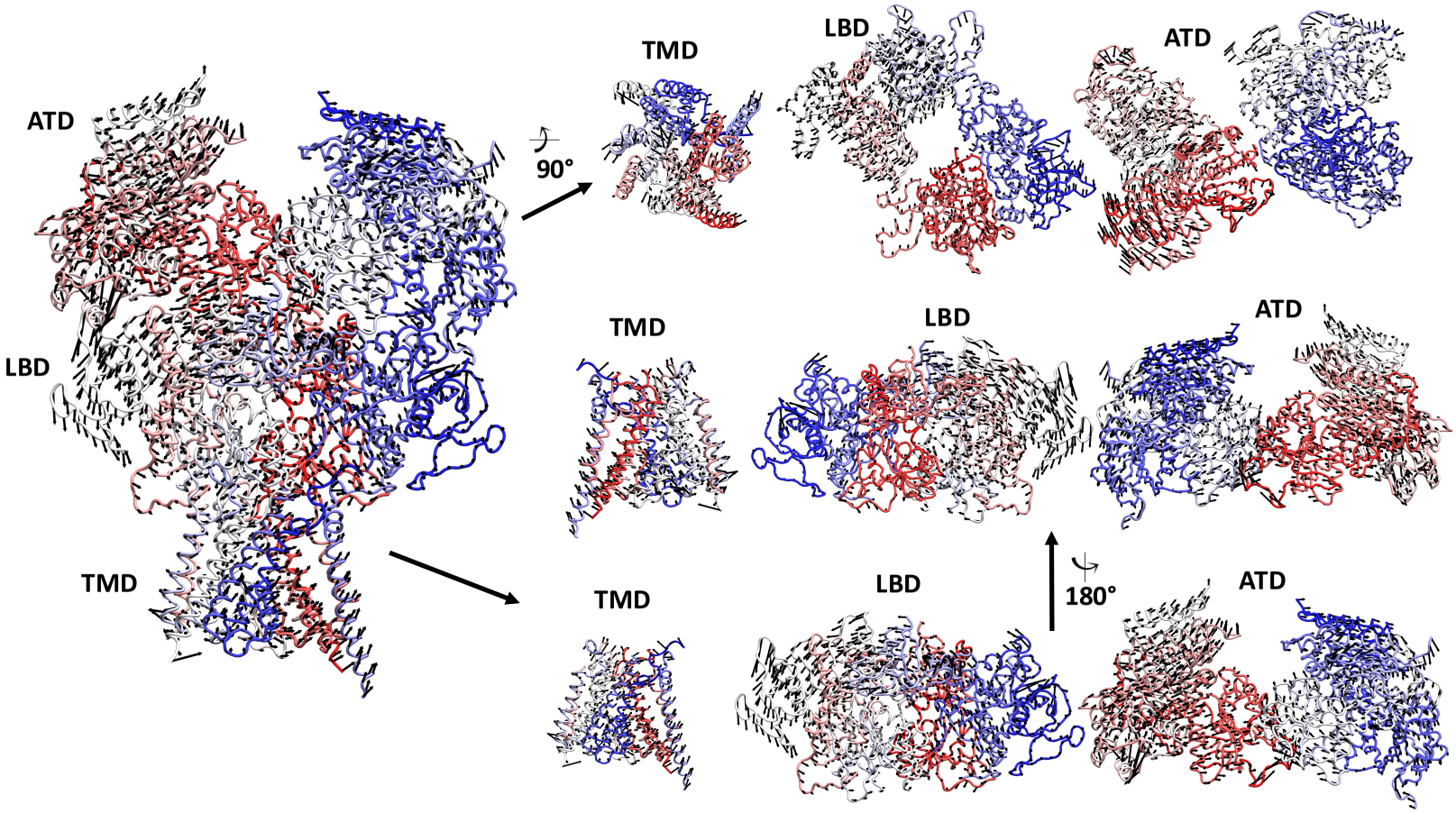


Figure 3.25. *First mode of ligand-bound simulation from PCA. Structure on the left represents the whole receptor. ATDs, LBDs and TMDs are also shown separately from different views on the right side of the figure. Black arrows on the structures show the movement of that region in the direction of that arrow.*

Fig 3.24 and 3.25 depicts the movement of the receptor in the first mode of PCA for apo and ligand-bound simulation, respectively. What is observed in the first modes is that the apo structure shows a rotation motion in the extracellular domains while the ligand-bound structure shows a sliding motion. Sliding-like motion is observed in the extracellular domains ATD and LBD of neighboring heterodimers $GluN1_1$-$GluN2A_2$ and $GluN1_2$-$GluN2A_1$ subunits.

NMDA receptors exhibit an overall two-fold axis. The GluN1 and GluN2A subunits are arranged in a 1-2-1-2 orientation with two-fold symmetry between the two GluN1-GluN2A heterodimers in the ATD and LBD. Additionally, swapping of dimer pairs between the ATD and LBD regions are observed in NMDAR. In the ATD region, heterodimer pairs assemble as $GluN1_1$-$GluN2A_1$ and $GluN1_2$-$GluN2A_2$ whereas in the LBD region, they assemble as $GluN1_1$-$GluN2A_2$ and $GluN1_2$-$GluN2A_1$. The results in LBD clamshell closure, ATD-LBD inter-domain closure and LBD-TMD linker regions represented changes in the neighboring heterodimers $GluN1_1$-$GluN2A_2$ subunits. This fashion is also seen in the first mode motions of PCA in ligand-bound simulation. As a result, LBD clamshell closure, distance decrease between ATD-LBD inter-domains and the increase in the distance of LBD-TMD linker might be related with the symmetry axis crossing between the $GluN1_1$-$GluN2A_2$ and $GluN1_2$-$GluN2A_1$ subunits.

4. CHAPTER 4: CONCLUSION

In this study, missing loops in the structure of human GluN1-GluN2A type NMDA receptor, which is homology modeled based on the *Xenopus laevis* GluN1-GluN2B type NMDAR template structure, are modeled. Since the NMDA receptor is a large heterotetrameric membrane protein complex, crystallizing the X-ray structure is challenging. NMDAR X-ray structures crystallized from *Xenopus laevis* and *Rattus norvegicus* are the first intact structures of NMDA receptor. For the human receptor, although higher sequence identity is observed in *Rattus norvegicus* structure, homology model was created based on *Xenopus laevis* because there were many mutations introduced in *Rattus norvegicus* structure. For such a large system with high sequence similarity, various challenges have been observed in the modelling section. Finally, the homology model is modeled as a tetramer after cutting the loops off since the loops reduced the structural quality of the model. Missing loops were modeled by two different modeling programs, MODELLER and ROSETTA. According to the structural assessment of the models, ROSETTA results were better compared to MODELLER. The loops were assessed by ROSETTA and MODELLER scoring functions.

Loops are the most mobile regions of the proteins. In general, it is very difficult to capture them in one unique conformation. This is also the reason in X-ray methods they are problematic. In this study, a special modeling step included a detailed evaluation of the loops generated by different algorithms to ensure the best starting structures to the simulation step.

After obtaining the final model, natural agonists of NMDA receptor, glycine and glutamate, are docked into the binding sites in LBDs by AutoDock. The challenge that we came across during the simulation was the escape of the ligand from the binding pocket in one of the four subunits. The orientation of the ligand in this subunit was problematic. We changed the protocol used in order to keep the ligand in the binding pocket by restraining the protein and ligand to their initial states for longer times in the simulation.

We used 10 parameters to monitor the structural changes in both apo and ligand-bound structures throughout the simulations. The first was the distance between the residues of TMD top girdle which is the narrowest region of the ion channel near the extracellular side. This distance remained larger in the ligand-bound simulation compared to apo. Bottom girdle, which

is the second narrowest region of the channel near intracellular side, remained generally same in both simulations.

For the closure of D1-D2 in clamshell shaped LBDs, we monitored a distance (ε) between two residues which have been studied in a previous NMDAR work. In that work, the closure of the clamshell was studied by using free energy method. We have seen that the distance stayed smaller for specific chains in the structure ($GluN1_1$ and $GluN2A_2$) in ligand-bound simulation meaning that the residues came closer to each other. This change in ligand-bound simulation is also reflected to other regions of the protein, such as the linker between LBD and TMD and also inter-domain distance between ATD and LBD. For the chains $GluN1_1$ and $GluN2A_2$, the linker distances remained longer in ligand-bound simulation. Shorter inter-domain distance between ATD and LBD is observed in ligand-bound simulation, representing a more compact ATD-LBD distance. These two mechanisms resemble the K^+ channel pH sensitive helix pulling mechanism for opening the hydrophobic girdle at helix bundle. In NMDAR, the channel is upside down of the K^+ channel. In K^+ channel, the hydrophobic helix bundle crossing is near the intracellular side while it is near extracellular side in NMDAR. Actually, it was not possible to observe a complete opening of the channel in the limited time of our MD simulations. However, the correlated motions in linker region, LBD clamshell closure and ATD-LBD inter-domain closure is observed in ligand-bound simulations only. That can be interpreted as if the ligand induces a closure in ligand binding D1 and D2 domains, it is reflected to ATD and linker region to TMD.

In TMD region, opening of the top girdle has not been observed while the bottom girdle, which is the selectivity filter of the ion channel and stabilizes the ion flow, opened for a short time. The simulation time of 1 microsecond is very long to run in terms of computer power but still might not be sufficient to observe channel opening which occurs in millisecond time scales.

5. REFERENCES

Allen, W. J., Lemkul, J. A., & Bevan, D. R. (2009). GridMAT-MD: A grid-based membrane analysis tool for use with molecular dynamics. *Journal of Computational Chemistry*, *30*(12), 1952–1958.

Anson, L. C., Schoepfer, R., Colquhoun, D., & Wyllie, D. J. (2000). Single-channel analysis of an NMDA receptor possessing a mutation in the region of the glutamate binding site. *The Journal of Physiology*, *527 Pt 2*(Pt 2), 225–37.

Armstrong, N., & Gouaux, E. (2000). Mechanisms for activation and antagonism of an AMPA-sensitive glutamate receptor: crystal structures of the GluR2 ligand binding core. *Neuron*, *28*(1), 165–81.

Armstrong, N., Jasti, J., Beich-Frandsen, M., & Gouaux, E. (2006). Measurement of Conformational Changes accompanying Desensitization in an Ionotropic Glutamate Receptor. *Cell*, *127*(1), 85–97.

Barry, P. H., & Lynch, J. W. (2005). Ligand-Gated Channels. *IEEE Transactions on Nanobioscience*, *4*(1), 70–80.

Bernard, G., & Shevell, M. I. (2008). Channelopathies: A Review. *Pediatr Neurol*, *38*, 73–85.

Bettler, B., & Mulle, C. (1995). AMPA and kainate receptors. *Neuropharmacology*.

Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., Woods, R. J. (2005). The Amber biomolecular simulation programs. *Journal of Computational Chemistry*, *26*(16), 1668–1688.

Case, Darden, T., Cheatham, Simmerling, C., Wang, J., Duke, R., Kollman, P. (2008). AMBER 10.

David, C. C., & Jacobs, D. J. (2014). Principal Component Analysis: A Method for Determining the Essential Dynamics of Proteins. In *Methods in molecular biology,* (Vol. 1084, pp. 193–226).

Davies, J., & Watkins, J. C. (1979). Selective antagonism of amino acid-induced and synaptic excitation in the cat spinal cord. *The Journal of Physiology*, *297*, 621–635.

Demir, A., & Essiz, S. (2017). Homology modeling and normal mode analysis of human NR1-NR2A NMDA type receptors. Master's thesis, Kadir Has University, Istanbul, Turkey.

Dingledine, R., Borges, K., Bowie, D., & Traynelis, S. F. (1999). The glutamate receptor ion channels. *Pharmacological Reviews*, *51*(1), 7–61.

Dravid, S. M., Burger, P. B., Prakash, A., Geballe, M. T., Yadav, R., Le, P., Traynelis, S. F. (2010). Structural Determinants of D-Cycloserine Efficacy at the NR1/NR2C NMDA Receptors. *Journal of Neuroscience*, *30*(7), 2741–2754.

Dürr, K. L., Chen, L., Stein, R. A., De Zorzi, R., Folea, I. M., Walz, T., Gouaux, E. (2014). Structure and dynamics of AMPA receptor GluA2 in resting, pre-open, and desensitized states. *Cell*, *158*(4), 778–792.

Dutta, A., Shrivastava, I. H., Sukumaran, M., Greger, I. H., & Bahar, I. (2012). Comparative dynamics of NMDA- and AMPA-glutamate receptor n-terminal domains. *Structure*.

Dworakowska, B., & Dołowy, K. (2000). Ion channels-related diseases. *Acta Biochimica Polonica*, *47*(3), 685–703.

Fiser, A., Do, R. K. G., & Šali, A. (2000). Modeling of loops in protein structures. *Protein Science*, *9*(9), 1753–1773.

Frydenvang, K., Lash, L. L., Naur, P., Postila, P. A., Pickering, D. S., Smith, C. M., Kastrup, J. S. (2009). Full Domain Closure of the Ligand-binding Core of the Ionotropic Glutamate Receptor iGluR5 Induced by the High Affinity Agonist Dysiherbaine and the Functional Antagonist 8,9-Dideoxyneodysiherbaine. *Journal of Biological Chemistry*, *284*(21), 14219–14229.

Ghazi, A., Berrier, C., Ajouz, B., & Besnard, M. (1998). Mechanosensitive ion channels and their mode of activation. *Biochimie*, *80*(5–6), 357–362.

Gillessen, T., Grasshoff, C., & Szinicz, L. (2002). Mitochondrial permeability transition can be directly monitored in living neurons. *Biomedicine and Pharmacotherapy*, *56*(4), 186–193.

Goodenough, D. A., & Paul, D. L. (2009). Gap Junctions. *Cold Spring Harbor Perspectives in Biology*.

Greger, I. H., Ziff, E. B., & Penn, A. C. (2007). Molecular determinants of AMPA receptor subunit assembly. *Trends in Neurosciences*.

Hayward, S., & Groot, B. L. (2008). Normal Modes and Essential Dynamics. In *Methods in molecular biology (Clifton, N.J.)* (Vol. 443, pp. 89–106).

Hille, B. (2001). *Ion channels of excitable membranes*. Sinauer.

Hollmann, M., Maron, C., & Heinemann, S. (1994). N-glycosylation site tagging suggests a three transmembrane domain topology for the glutamate receptor GluRl. *Neuron*, *13*(6), 1331–1343.

Huang, S.-Y., & Zou, X. (2010). Advances and Challenges in Protein-Ligand Docking. *International Journal of Molecular Sciences*, *11*(8), 3016–3034.

Huey, R., Morris, G. M., Olson, A. J., & Goodsell, D. S. (2007). A semiempirical free energy force field with charge-based desolvation. *Journal of Computational Chemistry*, *28*(6), 1145–1152.

Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: Visual molecular dynamics. *Journal of Molecular Graphics*, *14*(1), 33–38.

Issack, B. B., Berjanskii, M., Wishart, D. S., & Stepanova, M. (2012). Exploring the essential collective dynamics of interacting proteins: Application to prion protein dimers. *Proteins: Structure, Function, and Bioinformatics*, *80*(7).

Kandel, E. R., Dudai, Y., & Mayford, M. R. (2014). The Molecular and Systems Biology of Memory. *Cell*, *157*(1), 163–186.

Karakas, E., & Furukawa, H. (2014). Crystal structure of a heterotetrameric NMDA receptor ion channel. *Science*, *344*(6187), 992–997.

Keiser, J., & Utzinger, J. (2010). The Drugs We Have and the Drugs We Need Against Major Helminth Infections. In *Advances in parasitology* (Vol. 73, pp. 197–230).

Keramidas, A., & Lynch, J. W. (2013). An outline of desensitization in pentameric ligand-gated ion channel receptors. *Cellular and Molecular Life Sciences*, *70*(7), 1241–1253.

Kučerka, N., Tristram-Nagle, S., & Nagle, J. F. (2006). Structure of Fully Hydrated Fluid Phase Lipid Bilayers with Monounsaturated Chains. *Journal of Membrane Biology*, *208*(3), 193–202.

Kukol, A. (2015). *Molecular Modeling of Proteins*. *Methods (San Diego, Calif.)* (Vol. 1215).

Kuryatov, A., Laube, B., Betz, H., & Kuhse, J. (1994). Mutational analysis of the glycine-binding site of the NMDA receptor: structural similarity with bacterial amino acid-binding proteins. *Neuron*, *12*(6), 1291–300.

Lau, A. Y., & Roux, B. (2011). The hidden energetics of ligand binding and activation in a glutamate receptor. *Nature Structural & Molecular Biology*, *18*(3), 283–7.

Laube, B., Kuhse, J., & Betz, H. (1998). Evidence for a tetrameric structure of recombinant NMDA receptors. *The Journal of Neuroscience, 18*(8), 2954–61.

Leach, A. R. (2001). *Molecular modelling : principles and applications*. Prentice Hall. R

Leaver-Fay, A., O'Meara, M. J., Tyka, M., Jacak, R., Song, Y., Kellogg, E. H., Kuhlman, B. (2013). Scientific benchmarks for guiding macromolecular energy function improvement. *Methods in Enzymology*, *523*, 109–43.

Lee, C.-H., Lü, W., Michel, J. C., Goehring, A., Du, J., Song, X., & Gouaux, E. (2014). NMDA receptor structures reveal subunit arrangement and pore architecture. *Nature*, *511*.

MacKerell, A. D., Feig, M., & Brooks, C. L. (2004). Improved Treatment of the Protein Backbone in Empirical Force Fields. *Journal of the American Chemical Society*, *126*(3), 698–699.

Mamonova, T., Yonkunas, M. J., & Kurnikova, M. G. (2008). Energetics of the cleft closing transition and the role of electrostatic interactions in conformational rearrangements of the glutamate receptor ligand binding domain. *Biochemistry*, *47*(42), 11077–85.

Mandell, D. J., Coutsias, E. A., & Kortemme, T. (2009). Sub-angstrom accuracy in protein loop reconstruction by robotics-inspired conformational sampling. *Nature Methods*, *6*(8), 551–552.

Mayer, M. L. (2005). Glutamate receptor ion channels. *Current Opinion in Neurobiology*.

McCudden, C. R., Hains, M. D., Kimple, R. J., Siderovski, D. P., & Willard, F. S. (2005). G-protein signaling: back to the future. *Cellular and Molecular Life Sciences*, *62*(5), 551–577.

Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., & Olson, A. J. (1998). Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *Journal of Computational Chemistry*, *19*(14), 1639–1662.

Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *Journal of Computational Chemistry*, *30*(16), 2785–91.

Okada, O., Odai, K., Sugimoto, T., & Ito, E. (2012). Molecular dynamics simulations for glutamate-binding and cleft-closing processes of the ligand-binding domain of GluR2. *Biophysical Chemistry*, *162*, 35–44.

Oldham, W. M., & Hamm, H. E. (2008). Heterotrimeric G protein activation by G-protein-coupled receptors. *Nature Reviews Molecular Cell Biology*, *9*(1), 60–71.

Olsson, M. H. M., Søndergaard, C. R., Rostkowski, M., & Jensen, J. H. (2011). PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical pK_a Predictions. *Journal of Chemical Theory and Computation*, *7*(2), 525–537.

Ozawa, S., Kamiya, H., & Tsuzuki, K. (1998). Glutamate receptors in the mammalian central nervous system. *Prog Neurobiol*, *54*(5), 581–618.

Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Schulten, K. (2005). Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, *26*(16), 1781–1802.

Platt, S. R. (2007). The role of glutamate in central nervous system health and disease - A review. *Veterinary Journal*.

Postila, P. A., Ylilauri, M., & Pentikäinen, O. T. (2011). Full and Partial Agonism of Ionotropic Glutamate Receptors Indicated by Molecular Dynamics Simulations. *Journal of Chemical Information and Modeling*.

Prince, M., Comas-Herrera, A., Knapp, M., Guerchet, M., & Karagiannidou, M. (2016). World Alzheimer Report 2016 Improving healthcare for people living with dementia. Coverage, Quality and costs now and in the future, 1–140.

Ransom, R. W., & Stec, N. L. (1988). Cooperative modulation of [3H]MK-801 binding to the N-methyl-D-aspartate receptor-ion channel complex by L-glutamate, glycine, and polyamines. *Journal of Neurochemistry*, *51*(3), 830–6.

Retchless, B. S. (2011). Mechanistic basis of NMDA receptor channel property variation. Doctoral dissertation, University of Pittsburgh, PA, USA. Retrieved from http://d-scholarship.pitt.edu/10007.

Rohl, C. A., Strauss, C. E. M., Chivian, D., & Baker, D. (2004). Modeling Structurally Variable Regions in Homologous Proteins with Rosetta. *Proteins: Structure, Function and Genetics*, *55*(3), 656–677.

Rosenmund, C., Stern-Bach, Y., Stevens, C. F., Karlin, A., Kuusinen, A., Arvola, M., Stevens, C. F. (1998). The tetrameric structure of a glutamate receptor channel. *Science (New York, N.Y.)*, *280*(5369), 1596–9.

Ryan, T. J., Emes, R. D., Grant, S. G., & Komiyama, N. H. (2008). Evolution of NMDA receptor cytoplasmic interaction domains: implications for organisation of synaptic signalling complexes. *BMC Neuroscience*, *9*(1), 6.

Sahai, M. A., & Biggin, P. C. (2011). Quantifying Water-Mediated Protein–Ligand Interactions in a

Glutamate Receptor: A DFT Study. *The Journal of Physical Chemistry B*, *115*(21), 7085–7096.

Šali, A., & Blundell, T. L. (1993). Comparative Protein Modelling by Satisfaction of Spatial Restraints. *Journal of Molecular Biology*, *234*(3), 779–815.

Schmid, S. M., & Hollmann, M. (2008). To gate or not to gate: Are the delta subunits in the glutamate receptor family functional ion channels? *Molecular Neurobiology*.

Schneggenburger, R., Zhou, Z., Konnerth, A., & Neher, E. (1993). Fractional contribution of calcium to the cation current through glutamate receptor channels. *Neuron*, *11*(1), 133–143.

Shapovalov, M. V, & Dunbrack, R. L. (2011). A smoothed backbone-dependent rotamer library for proteins derived from adaptive kernel density estimates and regressions. *Structure*, *19*(6), 844–858.

Simons, K. T., Bonneau, R., Ruczinski, I., & Baker, D. (1999). Ab initio protein structure prediction of CASP III targets using ROSETTA. *Proteins: Structure, Function and Genetics*, *37*, 171–176.

Smart, O. S., Neduvelil, J. G., Wang, X., Wallace, B. A., & Sansom, M. S. P. (1996). HOLE: A program for the analysis of the pore dimensions of ion channel structural models. *Journal of Molecular Graphics*, *14*(6), 354–360.

Søndergaard, C. R., Olsson, M. H. M., Rostkowski, M., & Jensen, J. H. (2011). Improved Treatment of Ligands and Coupling Effects in Empirical Calculation and Rationalization of pK_a Values. *Journal of Chemical Theory and Computation*, *7*(7), 2284–2295.

Song, Y., Tyka, M., Leaver-Fay, A., Thompson, J., & Baker, D. (2011). Structure-guided forcefield optimization. *Proteins*, *79*(6), 1898–909.

Sukumaran, M., Rossmann, M., Shrivastava, I., Dutta, A., Bahar, I., Greger, I. H., Traynelis, S. (2011). Dynamics and allosteric potential of the AMPA receptor N-terminal domain. *The EMBO Journal*, *30*(5), 972–982.

Traynelis, S. F., Wollmuth, L. P., McBain, C. J., Menniti, F. S., Vance, K. M., Ogden, K. K., Dingledine, R. (2010). Glutamate receptor ion channels: structure, regulation, and function. *Pharmacological Reviews*, *62*(3), 405–96.

Vijayan, R., Sahai, M. A., Czajkowski, T., & Biggin, P. C. (2010). A comparative analysis of the role of water in the binding pockets of ionotropic glutamate receptors. *Physical Chemistry Chemical Physics*, *12*(42), 14057.

Vogt, A. D., & Di Cera, E. (2012). Conformational selection or induced fit? A critical appraisal of the kinetic mechanism. *Biochemistry*, *51*(30), 5894–5902.

Yao, Y., & Mayer, M. L. (2006). Characterization of a Soluble Ligand Binding Domain of the NMDA Receptor Regulatory Subunit NR3A. *Journal of Neuroscience*, *26*(17).

Zheng, W., Wen, H., Iacobucci, G. J., & Popescu, G. K. (2017). Probing the Structural Dynamics of the NMDA Receptor Activation by Coarse-Grained Modeling. *Biophysical Journal*, *112*(12), 2589–2601.

Zhu, S., Stroebel, D., Yao, C. A., Taly, A., & Paoletti, P. (2013). Allosteric signaling and dynamics of the clamshell-like NMDA receptor GluN1 N-terminal domain. *Nature Structural & Molecular Biology*, *20*(4), 477–85.

6. SUPPLEMENTARY MATERIALS

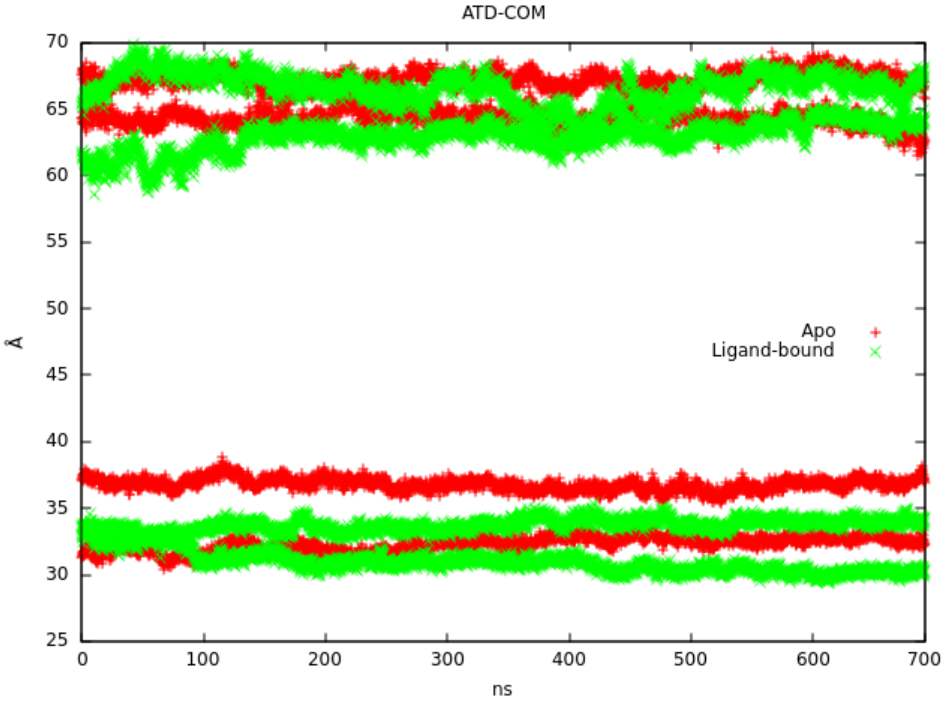


Figure S1. *Separate distances of ATD center of masses from each other for apo (red) and ligand-bound (green) simulations. The summation of the distances is shown in Fig. 3.12 C.*

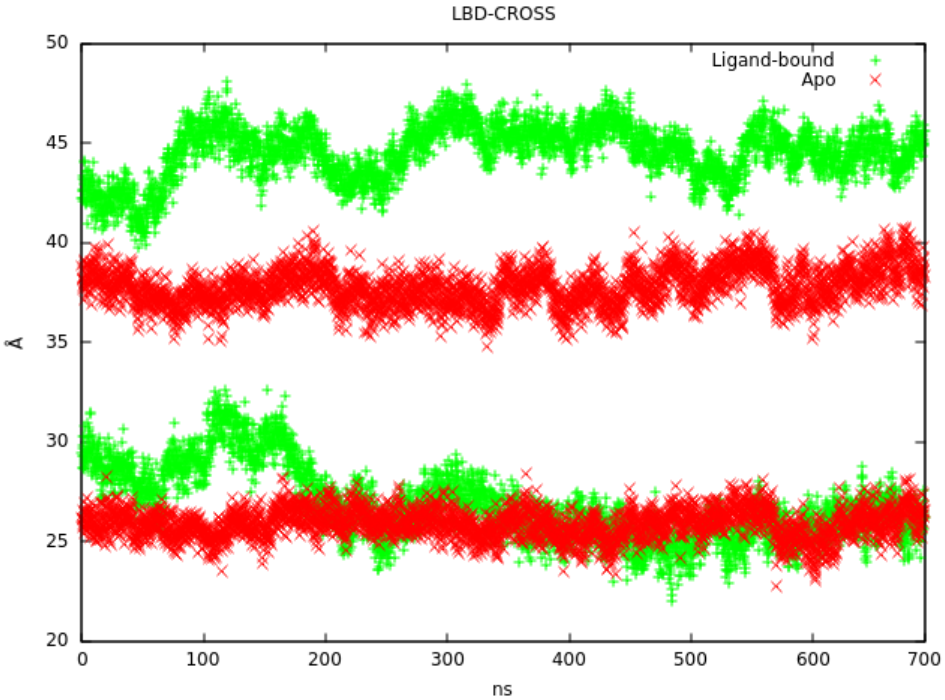


Figure S2. *Separate distances of LBD cross residues from each other for apo (red) and ligand-bound (green) simulations. The summation of the distances is shown in Fig. 3.12 D.*