# GATING MECHANISM IN NICOTINIC ACETYLCHOLINE RECEPTOR CHANNEL

by

Serkan Karadağ

BS. in Ch.E., Yıldız Teknik Üniversitesi, 2003

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

Graduate Program in Chemical Engineering

Boğaziçi University

2006

GATING MECHANISM IN NICOTINIC ACETYLCHOLINE RECEPTOR CHANNEL

APPROVED BY:

| | |
|---|---|
| Prof. Dr. Türkan Haliloğlu (Thesis Supervisor) | ..... |
| Asst. Prof. Şefika Banu Özkan | ..... |
| Prof. Dr. Viktorya Aviyente | ..... |

DATE OF APPROVAL: 13.06.2006

# ACKNOWLEDGEMENTS

I'd like to thank Türkan Haliloğlu for her continious support, Levent Tuna for helping me to get through the perilous fortran codes and discussing nAChRs throughout the programme.

# ABSTRACT

## GATING MECHANISM IN NICOTINIC ACETYLCHOLINE RECEPTOR CHANNEL

Nicotinic acetylcholine receptor (nAChR) is a ligand gated ion channel. Aberrations of nAChR is the primary reason of many congenital myasthenic syndromes. The mechanism behind the function of nAChRs is still unclear. In the present study nAChR from *Torpedo marmorata* was analyzed using the 3D structures of trasmembrane domain (TMD) (Protein Data Bank (PDB) Code: 1oed) and TMD with ligand binding domain (LBD) and five helices from the cyptoplasmic domain (PDB Code: 2bg9). Elastic Network Models (ENM), namely Gaussian Network Model (GNM) and Anisotropic Network Model (ANM), were employed to assess the underlying motions of nAChR. A program called MCPOOL was designed to analyze the correlated fluctuations by GNM to study the communication pathways between the binding domain and the gate. Conservation and correlated mutation analysis was performed to interpret the results of the ENMs from an evolutionary standpoint.

The results of GNM analysis suggested that TMD domain is largely affected by LBD and cytoplasmic domain. The highly fluctuating motion of M2-M3 loop is suppressed by the presence of LBD. The cross correlations elaborated by MCPOOL revealed the pathways between binding and gating region. The most visited regions are the initial residues of M1, M2-M3 loop, and the last residues of M4. They are on the pathways and overlap the minimum fluctuating regions and highly associate with the conserved residues and the residue pairs of correlated mutations. The results agree in the suggestion that these most visited regions are key passageways on the signal transmission pathway, and thus affect the gating mechanism.

# ÖZET

## ASETİLKOLİN RESEPTÖRLERİNDE KANAL AÇILMA MEKANİZMASI

Nikotinic asetilkolin reseptörü (nAChR) ligand kontrollü (ligand gated) iyon kanalıdır. Bu reseptördeki bozukluklar birçok kas hastalığına neden olur. Bu reseptörün çalışma mekanizması tam anlamıyla anlaşılamamıştır. Bu çalışmada *Torpedo marmorata*'nın nAChR'ı protein veri bankasındaki sadece hücre duvarının içindeki bölgeyi içeren 1oed ve hücre dışı ve hücre içi bölgeri de içeren 2bg9 kodlu yapılar incelenmiştir. Elastik ağ modelleri olan Gaussian ve anizotropik (Anisotropik) ağ yapı modelleri reseptörün hareketlerini incelemek için kullanımıştır. Çapraz korelasyonları inceleyip yapı içindeki haberleşme yollarını bulmak için MCPOOL adlı bir program tasarlanmıştır. Ayrıca korunmuş (conserved) aminoasitler ve beraber değişime uğramış aminoasit çiftleri (correlated mutation) belirlenerek aminoasitlerin evrimsel değişimi ile üç boyutlu yapıdaki dinamiğin nasıl örtüştüğü irdelenmiştir.

Dinamik sonuçlar hücre duvarındaki M1, M2, M3 ve M4 sarmal yapılarından oluşan, bölgenin diğer bölgelerin varlığından etkilendiğini göstermektedir. Hücre dışı bölge ile hücre duvarı içinde olan bölge arasındaki M2-M3 bölgesinin hareketlerinin hücre dışı bölge tarafından bastırıldığı görülmektedir. MCPOOL tarafından ortaya çıkarılan haberleşme yolu üzerinde M1'deki ilk M4'teki son ve M2-M3 ara bölgesindeki aminoasitler yer almaktadır. En kısa haberleşme yolu üzerinde en çok yer alan aminoasitlerin yapının mekanistik olarak önemli noktalarında yer aldığı ve korunan ve birlikte değişen aminoasit çiftleri ile örtüştüğü görülmektedir. Bu sonuçlar ışığında adı geçen bölgelerin reseptör kanalının açılma mekanizması için önemli oldukları belirlenmiştir.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# LIST OF SYMBOLS/ABBREVIATIONS

| | |
|---|---|
| $C(i,j)$ | Normalized cross correlation function |
| **crosscorr** | Matrix containing the cross correlation values |
| $gn$ | Minimum separation between residues |
| $H$ | Internal hamiltonian |
| **H** | Hessian matrix |
| **h** | Elements of Hessian matrix |
| $i$ | Residue index |
| $j$ | Residue index |
| $k_B$ | Boltzmann constant |
| $mincor$ | Minimum correlation value accepted |
| $N$ | Number of residues |
| **PC** | Matrix of possible pairs |
| **R** | Matrix containing positions of residues |
| $R_c$ | Cut off distance |
| **s** | Matrix containing positions of residues |
| **SR** | Vector of initial region |
| $T$ | Absolute temperature |
| **u** | Matrix containing eigenvectors |
| $V$ | Potential function |
| $W$ | Gaussian distribution function |
| $Z$ | Sum of cross correlation values for possible pairs |
| $Z_N$ | Configurational integral |
| | |
| $\alpha$ | Alpha subunit |
| $\alpha_\gamma$ | Alpha gamma subunit |
| $\alpha_\delta$ | Alpha delta subunit |
| $\beta$ | Beta subunit |
| $\gamma$ | Gamma subunit, and Hookean force constant |
| $\mathbf{\Gamma}$ | Kirchoff matrix |

| | |
|---|---|
| $\delta$ | Delta subunit |
| $\varepsilon$ | Epsilon subunit |
| $\lambda$ | Eigenvalue |
| | |
| 3D | Three dimensional |
| 5HT3 | Subtype 3 of the 5-hydroxytryptamine |
| A | Alanine |
| ACh | Acetylcholine |
| AChBP | Acetylcholine binding protein |
| Ala | Alanine |
| ANM | Anisotropic network model |
| Arg | Arginine |
| Asn | Asparagine |
| Asp | Aspartic acid (Aspartate) |
| B10 | Beta 10 |
| C | Cysteine |
| CMA | Correlated mutation analysis |
| CMS | Congenital myasthenic syndrome |
| Cys | Cysteine |
| D | Aspartic acid (Aspartate) |
| DPPC | dipalmitoylphosphatidylcholine |
| E | Glutamic acid |
| F | Phenylalanine |
| G | Glycine |
| $GABA_A$ | Gamma-amino butyric acid receptor type A |
| Gln | Glutamine acid (Glutamate) |
| Glu | Glutamic acid |
| Gly | Glycine |
| GNM | Gaussian network model |
| H | Histidine |
| His | Histidine |

| | |
|---|---|
| I | Isoleucine |
| Ile | Isoleucine |
| K | Lysine |
| L | Leucine |
| LBD | Ligand binding domain |
| Leu | Leucine |
| LGIC | Ligand gated ion channel |
| Lys | Lysine |
| M | Methionine |
| MD | Molecular dynamics |
| Met | Methionine |
| N | Asparagine |
| nAChR | Nicotinic Acetylcholine Receptor |
| P | Proline |
| PDB | Protein Data Bank |
| Phe | Phenylalanine |
| Pro | Proline |
| Q | Glutamine acid (Glutamate) |
| R | Arginine |
| S | Serine |
| Ser | Serine |
| T | Threonine |
| Thr | Threonine |
| TMD | Transmembrane domain |
| Trp | Tryptophan |
| Tyr | Tyrosine |
| V | Valine |
| Val | Valine |
| W | Tryptophan |
| Y | Tyrosine |

# 1. INTRODUCTION

## 1.1. Background Information

Nicotinic acetylcholine receptor (nAChR) is one of the most studied protein in the superfamily of ligand gated ion channels (LGIC). LGIC superfamily comprises several families of evolutionarily related proteins. These proteins can be referred as the Cys-Loop receptors, due to the fact that extracellular halves of all of the family members contain a conserved pair of disulfide bonded cysteine residues separated by only thirteen residues [1]. Serotonin (subtype 3 of the 5-hydroxytryptamine $5HT_3$), Gamma-Amino Butyric Acid receptor type A ($GABA_A$) and glycine receptors are all members of LGIC superfamily. The families of LGIC superfamily can be divided into cation an anion selective channels. Among these members, nAChR and serotonin belong to the cation selective group while $GABA_A$ and glycine receptors belong to the anion selective group of ion selective channels [2].

nAChRs reside at the neuromuscular junctions and at a number of different sites in the central nervous system [3]. nAChRs are neurotransmitter receptors, their main function is transducing chemical signal into an electrical signal [4, 5]. They mediate fast chemical transmission of electrical signals by opening a narrow pore across the cell membrane, where they reside, in response to neurotransmitters released from the nerve terminal into the synaptic cleft [6]. nAChRs have three distinct regions, namely ligand binding domain (LBD) (extracellular domain), transmembrane domain (TMD), and cytoplasmic domain (CYP). Figure 1.1 displays these regions (See Figure 1.1) on the structure of nAChR.

nAChRs are formed by five homologous subunits ($\alpha_2$, $\beta$, $\delta$, $\gamma$, or $\varepsilon$) These large subunits are almost 150 $\mathring{A}$ long, and formed by approximately 370 residues. In this work, nAChR from the electric organ of *Torpedo marmorata*, which is the marbled electric ray fish which is highly enriched in acetylcholine (ACh) receptor containing

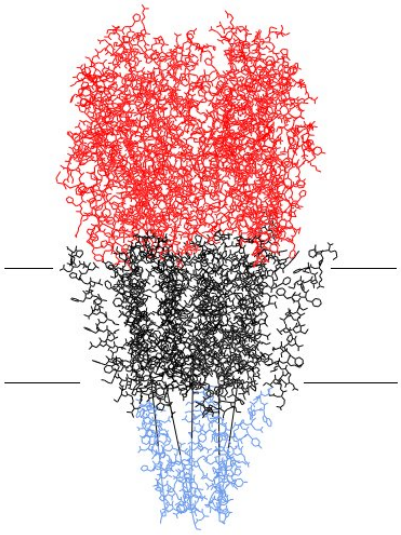

Figure 1.1. nAChR from *Torpedo marmorata* as seen from membrane surface. LBD is red, TMD is black and CYP is blue [10]. Black lines represent cell membrane

membranes [6, 7, 8], is analyzed, because of the available three dimensional (3D) structures (Protein Data Bank (PBD) code: 2bg9 and 1oed).

Until the release of the recent 3D structure of nAChR (PDB code: 2bg9), ligand binding and transmembrane domains did not exist in a single 3D structure. In fact the ligand binding domain was not resolved for *Torpedo marmorata*. During this period computational analyses were done either separately for ligand binding and transmembrane domains, using the 3D structure of transmembrane domain of *Torpedo marmorata* (PDB code: 1oed) [7] and the 3D structure of Acetylcholine Binding protein (AChBP), highly homologous to the ligand binding domain of nAChR (PDB Code: 1i9b [9]), or using a combined structure made up of 3D structures of TMD of nAChR from *Torpedo marmorata* and AChBP. The difference between AChBP and ligand binding domain of nAChR is that AChBP consists of only five identical subunits, it's a homopentamer. AChBP does not have an ion channel (a transmembrane domain). It's a soluble and has five binding pockets and is released to the synaptic cleft to modulate the synaptic transmission [9].

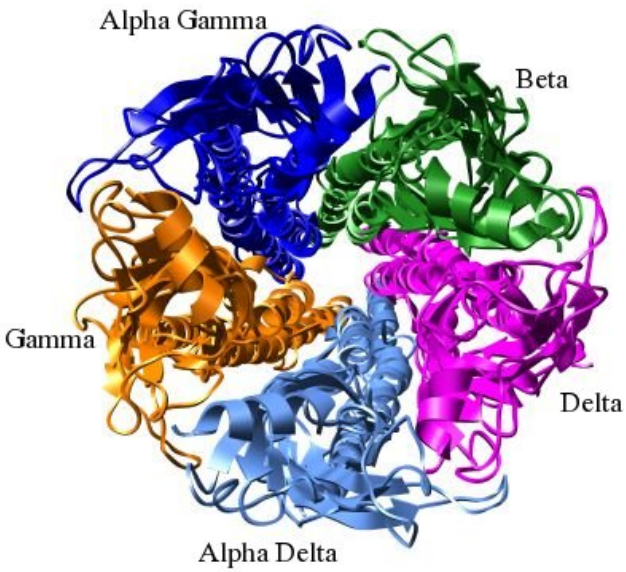


Figure 1.2. Subunits of nAChR viewed from the extracellular side. Alpha gamma, beta, delta, alpha delta, gamma subunits are shown (blue, green, magenta, light blue, orange respectively).

## 1.2. Structure

The structure of nAChR (PDB code: 2bg9) from the *Torpedo marmorata* is the closed structure of the receptor [10]. It is formed by five homologous subunits; these subunits are alpha gamma ($\alpha_\gamma$), beta ($\beta$), delta ($\delta$), alpha delta ($\alpha_\gamma$), and gamma ($\gamma$) ($\alpha_\gamma$, $\beta$, $\delta$, $\alpha_\delta, \gamma$) (See Figure 1.2) (See Appendix 1 for additional information about the sequences of each subunit). Each subunit, except for alphas which are identical, is at most 49 per cent identical to each other (delta and gamma) and at least 33 per cent identical (alpha and delta) (See Appexdix C, Pairwise alignment) according to the pairwise alignment (PIR SSearch, [11]). Figure 1.2 shows these subunits from the extracellular side of the membrane. Two alpha subunits are identical and are named alpha gamma and alpha delta. Alpha subunits are named with the neighbour with which they form the binding pocket. Each of the subunits is formed by three distinct regions, which are named according to their positions relative to the cell membrane (See Figure 1.1).

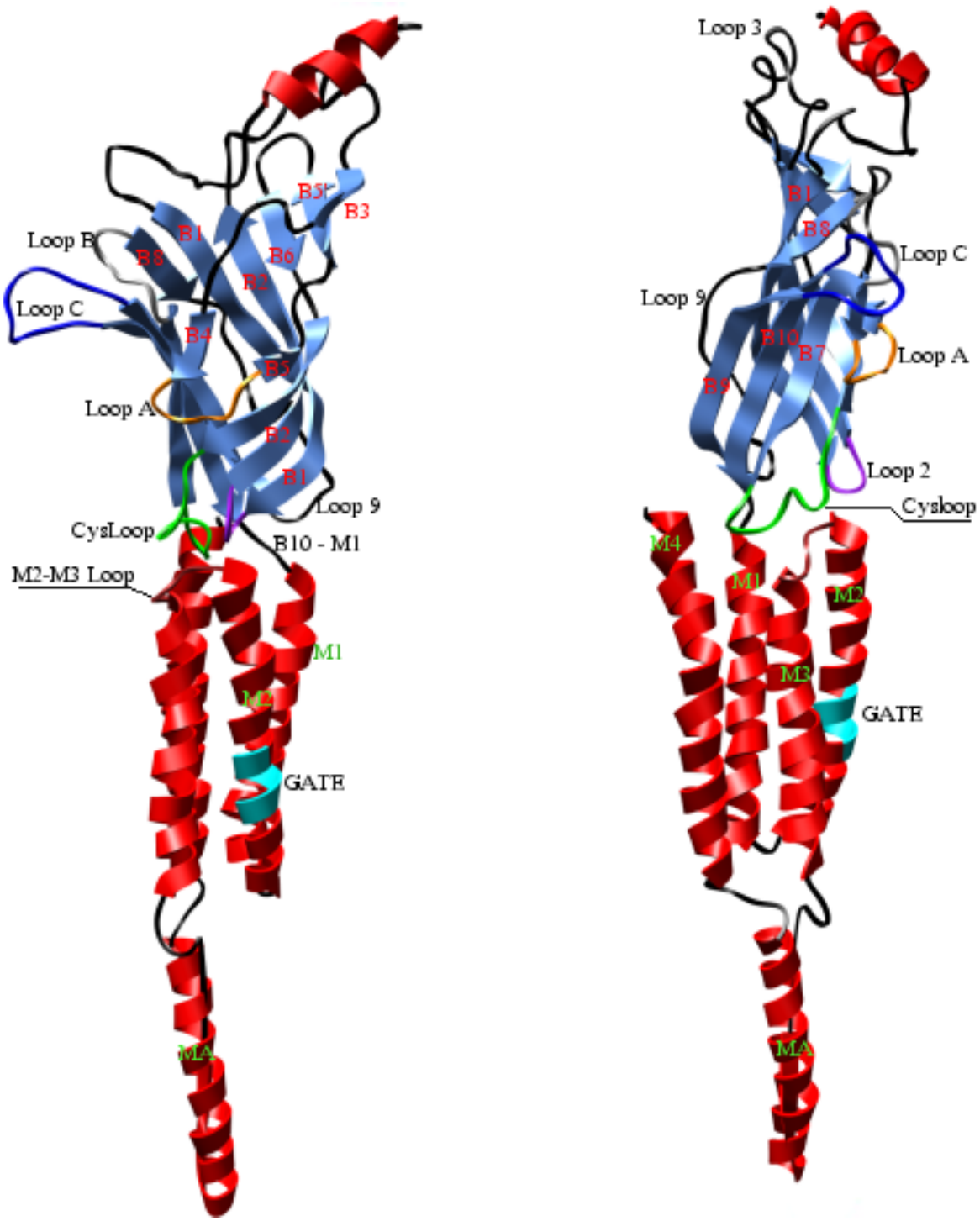


Figure 1.3. Secondary structure of alpha gamma subunit of nAChR. Residue which line the narrowest section of the pore is cyan [7, 10].

Each subunit is formed by mostly beta sheets and loops at the LBD and helices at the TMD. The LBD consist of one helix segment at the N terminal, twelve beta sheets and twelve loops between beta sheets. Each subunit has four helices at the transmembrane domain which named M1, M2, M3, and M4. Ligand binding domain and the transmembrane domains are connected covalently by beta 10 (B10) sheet and M1 helix. The interface of ligand binding and transmembrane domains are formed by Cys loop, loop two (between beta one (B1) and beta two (B2)), and loop nine, which connects beta eight (B8) to beta nine (B9) [12]. There are two ACh binding sites on nAChR [10]. These are located between the alpha subunits and their neighbours (alpha gamma-gamma, alpha delta-delta). Each of these binding pockets is formed by several loops and beta sheets contributed by an alpha subunit and the neighbouring subunit. Binding pocket is shaped by loop A (loop five), loop B (loop eight), and loop C (loop 10) from the alpha subunits and beta five (B5) and beta six (B6) of the adjacent subunit (delta or gamma subunit) [10] (See Figure 1.3). The residues which connect B10 strand to M1 helix are called B10-M1.

Transmembrane domain is formed by four alpha-helical segments of five subunits. These segments line around a central pore in a tapering manner (See Figure 1.1). Each TMD helix is divided into an upper half and a lower half for naming purposes. The upper half of a transmembrane helix is the half near the extracellular side, the lower half is the half near the intracellular side (See Figure 1.4.a). Among these alpha-helical segments are five M2 helices, which form the lumen of the pore (or the inner ring) [2, 13]; M1 and M2 helices form the inner ring, and M4 forms the outer ring (See Figure 1.4.b). Among these rings only outer ring have extensive lipid contacts, while middle ring makes much less lipid contacts [13, 14, 15]. M2 and M3 helices are connected with the M2-M3 loop which is thought to be one of the regions on the signal pathway from binding region to gate region [7, 10, 12, 16]. M1-M2 loop connects the M1 and M2 helices at the cytoplasmic half of the transmembrane domain (See Figure1.4.a).

The proposed gate, in the transmembrane domain, is thought to be formed by two constricting rings, which lie at the $\alpha_\gamma$Leucine251 ($\alpha_\gamma$Leu251 - Leucine Ring) and

$\alpha_\gamma$Valine255 ($\alpha_\gamma$Val255 - Valine Ring) [7] (See Figure 1.3). These two rings form the narrowest sections of the transmembrane domain, where upon activation widens by several Angstroms (freeze trapping experiments [17]). At these narrowest sections of the pore the M2 helices have very small separation and have bulky hydrophobic side chains. The symmetric nature of the pore leads to equal side to side hydrophobic interactions between equivalent residues at each subunit. Both small separation and the symmetric hydrophobic interactions keep the pore together [7, 10].

The cytoplasmic segment of the nAChR structure (PDB code: 2bg9) has missing loops. The only cytoplasmic segments present are the MA helices which are connected to the M4 helices. MA helices form an inverted pentagonal cone, which has five open spaces on the side. These open spaces are the only possible path for the ion flow [10]. Earlier kinetic studies established that nAChR can exist in at least three conformations with different functional properties, closed, open and desensitized [18].

### 1.3. Literature Survey

Nicotinic Acetylcholine receptors reside at the cell membrane facing the synaptic cleft. When ACh molecules bind to these receptors a fast conformational change occurs in the channel structure. This change leads to a change in the ionic permeability of the pore. The gating mechanism involves allosteric signalling caused by a brief exposure to ACh, and a conformational change at the alpha subunits. This conformational change is transmitted through the ligand binding domain to the transmembrane domain as a twist like motion [7], which in turn breaks the symmetric forces keeping the residues at the gate region together, allowing the pore to expand. Mutations of the genes coding the nAChR may alter this basic mechanism to an extend that renders nAChR impaired.

In LGIC super-family proper gating of the channel is the most important function. Improper functioning of nAChRs, due to their abundance at the neuro-muscular junction, is the reason of many congenital myasthenic syndromes (CMS), which are inherited disorders that cause muscle weakness (myasthenia). The loss of proper function damages the ability of post-synaptic cells to receive signals from the nervous system.

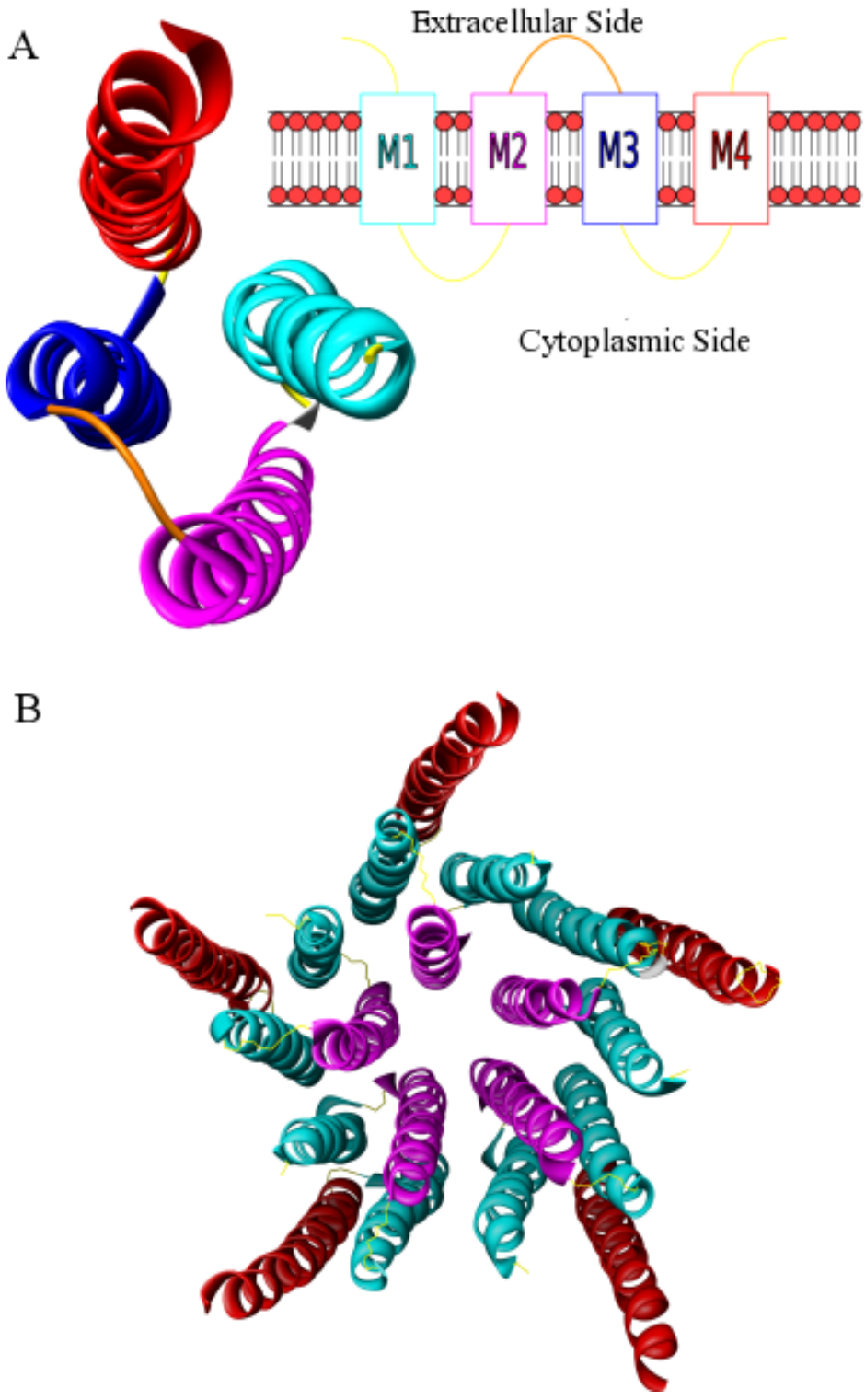


Figure 1.4. a) Transmembrane segment of nAChR and connecting loops. b) Presentation of inner ring (magenta), middlering (cyan), and outer ring (red).

Previous studies have revealed such mutations. The mutations of Valine 285 in the M3 helix of alpha subunit to Isoleucine ($\alpha_\gamma$V285I), and a mutation at the cytoplasmic loop between M3 and M4 helices of the $\beta$ subunit cause CMS [19, 20].

Nicotinic acetylcholine receptors are also found at the post-synaptic membranes of cells in brain. Mutations in these residues are thought to be involved in many diseases which includes, but are not limited to, autosomal dominant nocturnal frontal lobe epilepsy, Alzheimer's disease, schizophrenia, Parkinson's disease, Chrohn's disease [16, 21].

Nicotinic acetylcholine receptors have been a topic of experimental studies since the late sixties [22]. In late seventies it was found that delipidation and solubilization of nAChR irreversibly changed nAChR to a low affinity state [15]. Electron spin resonance studies in the late seventies of nAChRs in its native membrane environment disclosed a lipid belt region where lipids directly interacted with nAChR. This lipid belt region was shown to be qualitatively different than the rest of the membrane in muscle cells, regarding molecular mobility and lipid composition [2]. This association of lipid belt region and nAChR, supports the idea that nAChRs are, indeed, effected by their lipid surroundings, but the exact nature is still to be discovered. Also, there are evidence that nAChR interacts with the cholesterol molecules in the lipid membrane, which modulates nAChR for proper gating [15].

Experiments on nAChR using photoreactive phospholipids have shown that acetylcholine interacts with cholesterol and phospholipids [23]. Other experiments on nAChR label M4 transmembrane helices extensively with membrane partition photoactivatable probes. M3 and M1 transmembrane helices were also labelled [15, 24] In addition, photoaffinity experiments showed that gamma F292, L296, M295, and N300 faces lipid interface [25]. Experimental mutation studies on the residues of M3 helices, which are exposed to lipid environment, were conducted to disclose the functional contributions of the lipid exposed residues. The results have shown that an increase in either volume or hydrophobicity of these residues affect the function of nAChR. Also, experiments down

on the nAChR from the adult muscle cells indicate that 8' position (alpha Phe284) of the M3 domains of all subunits contribute to channel gating [26].

The M4 helix is the least conserved among the transmembrane helices and the most hydrophobic of all. Results of hydrophobic probe experiments have shown that it's the outermost among the transmembrane helices [24]. Experiments demonstrated that alpha T422I mutation, a residue on M4 helix, impedes this residues ability to make hydrogen bond, therefore influences the opening and closing rates of gating and that it's essential for proper gating [24, 27, 28, 29]. Another mutation experiment have shown that L416W and I419W mutations decreased the expression level to below 20 per cent of the wild type. Also an experimental study conducted by Tamamizu et. al. suggested that I425 at the M4 helix is important for proper folding and oligomerization of nAChR. A study on the contribution of M4 helices on gating demonstrated that the upper half of M4 helix have a larger effect when compared to the residues at the lower half. In addition, this study showed that mutations at positions L410, M415, C418, T422, and F426 of alpha M4 helix affects gating [29]. All of these studies prove that M4 is an essential part of the gating process [24].

In a study about the activation of nAChRs, it's suggested that activation involves a conformation change in the alpha subunits and Cys and M2-M3 loops are important for the transfer of signal from LBD to TMD [6]. A study on the coupling of ligand binding and gating on the $GABA_A$ receptors support the importance of Cys and M2-M3 loops [35]. Also, a study about the gating of $5HT3_A$ receptor supports that Cys and M2-M3 loops are important and also, adds that loop two is involved in signal transfer along with latter loops [30].

A previous study about the coupling of binding to gating proposed that loop nine from the neighbouring subunit contacts loop two, and loop two along with Cys loop straddles M2-M3 loop, suggesting a twisting motion that breaks apart the pore girdle and allows the flow of ions [12]. In a recent experimental study, the principal pathway is thought to begin with the binding pocket forming alpha loop C at the LBD and continue through the covalently connected B10 strand to the M1 helix region. In

between B10 and M1 alpha Arg209 is adjacent to alpha Glu45 of loop two at the LBD region. Glu45 along with Val46 couple with Pro272 at the M2-M3 loop [31]. This study also revealed that mutations to Arg209 impairs channel gating, and supports the importance of Cys loop and loop nine for the coupling of binding to gating as previously suggested [12, 31].

Regarding alpha subunit, a natural mutation of alpha Ser269Ile, last residue on M2 helix, causes a CMS, but mutations at the same position in non-alpha subunits do not cause a CMS [7], this evidence supports the idea that gating is mainly driven by alpha subunits.

The experimental data provides a foundation for computational studies to reach an understanding of the gating mechanism and delicacies of gating modulating interactions. Recently nAChRs have been the topic of several molecular dynamics (MD) simulations. MD simulations performed by Xu et. al. using the transmembrane domain of inserted into a dipalmitoylphosphatidylcholine (DPPC) bilayer, showed that M4 helices transmit the effects exerted by lipid surroundings into pore lining M2 helices [32]. Another MD simulation of the transmembrane domain in bilayer octane, mimicking cell membrane, pointed out that there is a hinge point near the gate, where all M2 helices showed bending motions. Also, this study underlines a correlation between M2-M3 loop and the gate region on the M2 helices, and that alpha M2-M3 loops fluctuates more than M2-M3 loops of non-alpha subunits [33]. MD simulations of ligand binding along with transmembrane domain state that LBD is anti-correlated with LBD. This study suggests that B9 and B10 form a lever that connects the binding site with the motions of TMD. Their simulation also supports the importance of loop two and M2-M3 and Cys loop. They propose that Cys loop pivots M2-M3 loop and loop two influences this motion [34].

Normal mode analysis (NMA) conducted by Cheng et. al. demonstrated that M2-M3 loop is correlated with Cys loop and loop two, which underlines that Cys loop and loop two undergo highly concerted motions with M2-M3 linker [18, 35]. Another NMA study states that while first mode motions are dominated with twisting, second

and third modes show symmetric expansion and asymmetric expansion respectively [34]. Also, one of the recent NMA study, suggests that only the first mode motions result in a structural change, which results in widening of the pore [36].

Correlated mutation analysis the alpha subunit of nAChR from *Torpedo californica* (99.13 per cent identical to *Torpedo marmorata*) has been the topic of a more recent publication Chen et. al. [37]. They identified clusters of correlated mutations at M1-M2 loop, M2-M3 loop. There are other correlated sites at M1 (Leu212, Ile219, Val232), M3 (Phe280, Thr281, Val293), and M4 (Val405, Leu411, Gly421, Thr422). On the LBD correlated residues were at sites on loop two (Gln48), loop nine (Trp176), and Cys loop (Val132, Phe137). They proposed that there is a network of correlated amino acids from the binding site to the gating region.

In this work, several different approaches were employed to identify important residues and elucidate the mechanism underlying the process which ends in an open pore. NMA was employed using a Gaussian Network Model (GNM) approach, and a custom program, MCPOOL, was written to analyze the cross correlation results of GNM. MCPOOL generates paths of given length biased by the cross correlation values and analyzes for communication pathways. GNM results do not include any information regarding the direction of motions (isotropic) [11]. A similar but anistropic method is the Anisotropic Network Model (ANM) [38]. ANM was also used to analyze the motions described by the lowest modes.

Conservation analysis was performed using the "Consurf" [39, 40] server at Bioinformatics Unit of Tel Aviv Unversity. Consurf provides data about conservation of each residue in a protein. There are some residues which are important for proper functioning of protein, but are not conserved during evolution. Evolution overcomes this problem with the coupled or correlated mutations, meaning that mutation a functionally or structurally important residue is kept, and passed to the next generation, if another mutation compensates for the loss of function caused by the first mutation. This analysis is called correlated mutation analysis (CMA). CMA was performed using the "Corrmut" [41] server at the Bioinformatics Unit of Tel Aviv University.

In the first half of the preparation period of this thesis only the 3D structure of the transmembrane domain of nAChR from *Torpedo marmorata* was present and all of the analyses were done using either the sequence or 3D structure (PDB code: 1oed) of this region. However a refined 3D structure was released in 2006, which include the ligand bindin domain and transmembrane domain together with a portion of cytoplasmic domain (PBD code: 2bg9). This new structure gave a new opportunity to assess the gating mechanism in a more complete manner.

This thesis will present the results of analysis for both of the structures and will compare the effect of the presence of LBD. Conservation and corelated mutation analyses were done using sequences homologous to the sequences of subunits of 2bg9, a set of homologous sequences was used for each subunit.

# 2. METHODS

## 2.1. Gaussian Network Model (GNM)

Gaussian network model (GNM) is a coarse grained one dimensional (isotropic) elastic network mode [11] where residues are thought to be connected to each other with harmonic springs. GNM assumes that fuctuations are isotropic, so motions are described with their amplitudes on a one dimensional scale and that the protein is in its native equilibrium state (in its minima). Interactions of individual atoms are coarse grained using a virtual bond model where individual residues are represented by their backbone carbon atoms ($C^{\alpha}$) and the interaction between different amino acid are not differentiated. A single-parameter potential approximation is adopted for all interacting backbone atoms and all backbone atoms are assumed to under undergo Gaussian fluctuations about their equilibrium values [42]. Backbone atoms are considered as interacting if the distance between two backbone atoms are less than a certain cut-off distance. The interaction map is described in a connectivity matrix, Kirchhoff matrix.

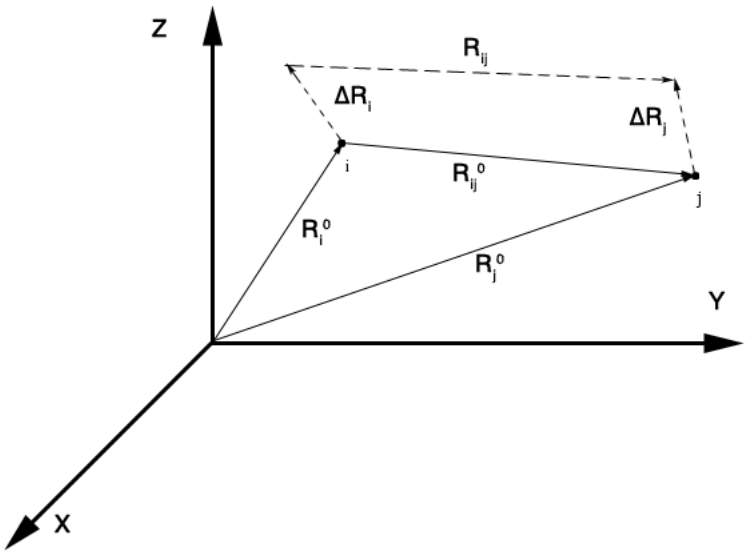


Figure 2.1. Representation of equilibrium positions $\mathbf{R}_i^0$ and $\mathbf{R}_j^0$ of the backbone atoms of $i^{\text{th}}$ and the $j^{\text{th}}$ residues. Each residue is represented by a circle.

The fluctuations of backbone atoms assumed to obey a Gaussian distribution (W($\Delta\mathbf{R}_i$)) is represented as

$$W(\Delta\mathbf{R}_i) = exp(-3(\Delta\mathbf{R}_i)^2/2 < (\Delta\mathbf{R}_i)^2 >) \tag{2.1}$$

where $< (\Delta\mathbf{R}_i)^2 >$ is the mean square fluctuation of the $i^{\text{th}}$ residue, $N$ is the total number of residues. The total potentil $V_{tot}$ can be written as the pairwise interactions of the molecular system of $N$ residues as

$$V_{tot} = \sum_i \sum_j V(\mathbf{R}_i, \mathbf{R}_j) = \sum_i \sum_j (1/2)\gamma(\Delta\mathbf{R}_{ij} \cdot \Delta\mathbf{R}_{ij}) \tag{2.2}$$

Here, the summation includes all $i$ and $j$ pairs ($1 \leq i < j \leq N$) for which equilibrium separation is less than a cut-off distance $R_c$, $\gamma$ represents the Hookean force constant assumed for all interacting pairs regardless of the chemical or physical nature of the interacting pair. $\Delta\mathbf{R}_{ij}$ can be described as

$$\Delta\mathbf{R}_{ij} = \mathbf{R}_{ij} - \mathbf{R}^0_{ij} \tag{2.3}$$

Here, $\Delta\mathbf{R}_{ij}$ is the fluctuation in the distance vector relative to the equilibrium separation $\mathbf{R}^0_{ij}$. $\mathbf{R}_{ij}$, which is the separation between sites $i$ and $j$, is

$$\mathbf{R}_{ij} = \mathbf{R}_i - \mathbf{R}_j \tag{2.4}$$

$\mathbf{R}_i$ ad $\mathbf{R}_j$ is the positions of $i$ and $j$ sites, respectively [11]. The fluctuations of the distance separating the residues are represented by $\Delta\mathbf{R}_{ij}$ can be rewritten as (See Figure 2.1)

$$\Delta\mathbf{R}_{ij} = \Delta\mathbf{R}_i - \Delta\mathbf{R}_j \tag{2.5}$$

where $\Delta\mathbf{R}_i$ represents the instantaneous fluctuations of the $i^{\text{th}}$ residue (See Figure 2.1).

$$\mathbf{\Gamma}_{ij} = \begin{cases} -1 & i \neq j \quad \mathbf{R}_{ij} \leq R_c \\ 0 & i \neq j \quad \mathbf{R}_{ij} \leq R_c \\ -\sum_{i \neq j} \mathbf{\Gamma}_{ij} & i = j \end{cases} \tag{2.6}$$

Here $R_c$ is the cut off distance, which is the maximum distance allowed for $i^{\text{th}}$ and $j^{\text{th}}$ residues ($10\mathring{A}$ ) to be treated as interacting. Using $\mathbf{\Gamma}$, the total potential function can be rewritten as

$$V_{tot} = H = (\gamma/2) tr[\Delta\mathbf{R}^T \mathbf{\Gamma} \Delta\mathbf{R}] \tag{2.7}$$

The $\Delta\mathbf{R}$ represents the matrix of instantaneous fluctuations of $\Delta\mathbf{R}_i$ of all residues, $(1 \leq i \leq N)$. $\Delta\mathbf{R}$ is a $N \times 3$ matrix whose $i^{\text{th}}$ row is the transpose of $\Delta\mathbf{R}_i$. $H$ is the internal Hamiltonian and is used to represent the total potential of a biomolecular system [11]. The average fluctuations of interacting residues are found from the inverse of Kirchoff matrix,

$$\langle \Delta\mathbf{R}_i \cdot \Delta\mathbf{R}_i \rangle = \frac{1}{Z_N} \int (\Delta\mathbf{R}_i \cdot \Delta\mathbf{R}_i) exp\{-H/k_B T\} d\{\Delta\mathbf{R}\} = (3k_B T/\gamma)[\mathbf{\Gamma}^{-1}]_{ij} \tag{2.8}$$

$Z_N$ represents the configurational integral which can be written as

$$Z_N = \int exp\{H k_B T\} d\{\Delta\mathbf{R}\} \tag{2.9}$$

where $k_B$ is the Boltzmann constant, $T$ is the absolute temperature and $d\{\Delta\mathbf{R}\}$ represents the integration over all residue fluctuations. Equation 2.8 provides a link between the connectivity matrix and the average fluctuation of interacting pairs and correlation between. This equation can further be simplified using the eigenvectors and eigenvalues obtained by decomposition of Kirchoff matrix along with equations 2.8 and 2.9 by

$$[\Delta\mathbf{R}_i \cdot \Delta\mathbf{R}_j]_{n,ni} = \frac{\sum_{k=n,ni}^{ni} \sum_{ni>n} [\mathbf{u}_k]_i [\mathbf{u}_k]_j \lambda_k^{-1}}{\sum_{k=n,ni}^{ni} \sum_{ni>n} \lambda_k^{-1}} \tag{2.10}$$

Here the correlation between the fluctuations of $i^{\text{th}}$ and $j^{\text{th}}$ residues in the modes between $n$ and $ni$ are presented (the contribution of several modes). $\lambda_k$ is the $k^{\text{th}}$ eigenvalue. The term $\mathbf{u}^k$ represents the $k^{\text{th}}$ eigenvector. The individual modes ($k$) of correlations can be calculated using the eigenvalue and the corresponding eigenvector instead of a range of modes. The smallest non-zero eigenvalue can be used to calculate the lowest mode and individual modes can be calculated using the corresponding eigenvalue. Autocorrelations can be calculated assuming $i$ and $j$ is equal. The normalized cross correlations of $i^{\text{th}}$ and $j^{\text{th}}$ ($C(i,j)$) nodes can be calculated using

$$C(i,j) = \frac{\langle \Delta\mathbf{R}_i \cdot \Delta\mathbf{R}_j \rangle}{[\langle \mathbf{R}_i \cdot \mathbf{R}_i \rangle \langle \mathbf{R}_j \cdot \mathbf{R}_j \rangle]^{1/2}} \tag{2.11}$$

Here cross correlations are normalized using the individual auto correlations (Eq. 2.10). Cross correlations represents the extend of correlation between the motions of $i^{\text{th}}$ and $j^{\text{th}}$ residues. Cross correlation values range from negative one to one. Positive values describe that $i^{\text{th}}$ and $j^{\text{th}}$ residues move in the same direction, while negative values mean that they move in opposite directions. Small absolute values of cross correlations represent a weak correlation between residues. Equation 2.11 can be used to calculate the contributions of individual or several modes together by

$$C(i,j) = \frac{[\Delta\mathbf{R}_i \cdot \Delta\mathbf{R}_j]_{n,ni}}{[[\mathbf{R}_i \cdot \mathbf{R}_i]_{n,ni} [\mathbf{R}_j \cdot \mathbf{R}_j]_{n,ni}]^{1/2}} \tag{2.12}$$

which is essentially the same as equation 8 without the ensemble averages.

The smallest non-zero eigenvalues define the slowest modes of motions while the highest eigenvalues define the fastest modes of motions. Slow modes represent a global

motion, while fast modes represent the isolated motions of individual backbone atoms [42]. The work done on GNM identified the residues, which are active in the fast modes, as key players upholding the stability of the folded molecule [42]. The local minima in the slow mode fluctuations are thought to be stable regions, with minimum global fluctuation [11]. These minimum fluctuating regions act as hinges, a static point, in between two more fluctuating (mobile). Therefore they are usually referred to as hinges. In this work autocorrelations at slow modes (small eigenvalue) and fast modes (large eigenvalue) was calculated using equation 2.10 and cross correlations was calculated using equation 2.12. The results of GNM analysis is used to understand the individual and correlated motions of nAChR. Results of slow and fast mode fluctuations were graphed along with other information to provide a base for observations on cross correlation maps. Cross correlation results were further analyzed with MCPOOL (explained later) to extract important information about each correlation, otherwise would be lost.

## 2.2. Anisotropic Network Model (ANM)

Anisotropic network model (ANM) is an extension of GNM where the fluctuations are anisotropic (affected by direction), thus incorporates the $X$, $Y$, and $Z$ components independently. Therefore, the overall potential for the ANM calculations includes the fluctuations for all components. The harmonic potential can be expressed by

$$\begin{aligned} V &= (1/2)\gamma(\mathbf{s}_{ij} - \mathbf{s}^0_{ij}) \\ &= (1/2)\gamma([(X_j - X_i)^2 + (Y_j - Y_i)^2 + (Z_j - Z_i)^2]^{1/2} - \mathbf{s}^0_{ij})^2 \end{aligned} \tag{2.13}$$

Here $\gamma$ is the harmonic force constant, and $\mathbf{s}^0_{ij}$ is the separation vector of $i^{\text{th}}$ and $j^{\text{th}}$ residues, $\mathbf{s}^0_{ij}$ is similar to $\mathbf{R}^0_{ij}$ in GNM (See Figure 2.1). In equation 2.13 $X_i$, $Y_i$, $Z_i$, are the components of vector $\mathbf{R}_i$. In ANM a similar matrix to Kirchoff matrix in GNM, the Hessian matrix, is employed. This matrix can be derived from the second derivative of equation 2.13. Both ANM and GNM assume that the structure is at equilibrium, therefore the first derivative of the potential is zero.

$$\frac{\partial V}{\partial X_i} = 0 \tag{2.14}$$

The second derivative for the potential function with respect to $X$ component becomes

$$\frac{\partial^2 V}{\partial X_i^2} = \frac{\gamma (X_j - X_i)^2}{\mathbf{s}_{ij}^2} \tag{2.15}$$

and the cross derivatives are

$$\frac{\partial^2 V}{\partial X_i \partial Y_j} = -\frac{\partial^2 V}{\partial X_j \partial Y_i} = -\gamma (X_j - X_i)(Y_j - Y_i)/\mathbf{s}_{ij}^2 \tag{2.16}$$

For the nodes surrounding $i^{\text{th}}$ node (where $\mathbf{s}_{ij}^0 < R_c$) equations 2.15 and 2.16 becomes

$$\frac{\partial^2 V}{\partial X_i^2} = \gamma \sum_j (X_j - X_i)^2/\mathbf{s}_{ij}^2 \tag{2.17}$$

$$\frac{\partial^2 V}{\partial X_i \partial Y_i} = \gamma \sum_j (X_j - X_i)(Y_j - Y_i)/\mathbf{s}_{ij}^2 \tag{2.18}$$

Here summations are over each $j$ node surrounding the $i^{\text{th}}$ node (where $\mathbf{s}_{ij}^0 < R_c$). The Hessian matrix is constructed as:

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}_{11} & \mathbf{h}_{12} & \dots & \mathbf{h}_{1N} \\ \mathbf{h}_{21} & \mathbf{h}_{21} & \dots & \mathbf{h}_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{h}_{N1} & \mathbf{h}_{N2} & \dots & \mathbf{h}_{NN} \end{bmatrix} \tag{2.19}$$

where the $\mathbf{h}_{ij}^{\text{th}}$ element is shown by

$$\mathbf{h}_{ij} = \begin{bmatrix} \partial^2 V/\partial X_i \partial X_j & \partial^2 V/\partial X_i \partial Y_j & \partial^2 V/\partial X_i \partial Z_j \\ \partial^2 V/\partial Y_i \partial X_j & \partial^2 V/\partial Y_i \partial Y_j & \partial^2 V/\partial Y_i \partial Z_j \\ \partial^2 V/\partial Z_i \partial X_j & \partial^2 V/\partial Z_i \partial Y_j & \partial^2 V/\partial Z_i \partial Z_j \end{bmatrix} \tag{2.20}$$

the elements of $\mathbf{h}_{ij}$ in the Hessian matrix are given by equation 2.16. When $i = j$, the diagonal elements of $\mathbf{h}_{ii}$ is given by equation 2.17 and the off-diagonal elements are given by equation 2.18. The decomposition of the Hessian matrix results in $3N - 6$ eigenvalues and $3N-6$ eigenvectors. In the present application, the conformations that describe the fluctuations from the average in the lowest ten nonzero modes were elaborated. Three dimensional (3D) structures were constructed using these vectors and the original node coordinates. A plus and a minus conformations were generated and the resulting 3D structure was animated using Visual Molecular Dynamics package (VMD) [43]. 'Molecular graphics images were produced using the Chimera package from the Computer Graphics Laboratory, University of California, San Francisco (supported by NIH P41 RR-01081)" [44]. The radius profile of the resulting structures pore were analyzed using HOLE [45] program suite.

The results of ANM analysis is important to visualize the isotropic results of GNM. ANM analysis yields the actual motions of each slow mode. ANM and GNM modes might not coincide perfectly, but together they provide information on the essential modes of motion. ANM result were used to qualitatively assess the dimension of the pore. The radius profile was used to interpret the occurring motions.

### 2.3. Multiple Sequence Alignment (MSA)

Multiple sequence alignment (MSA) is the alignment of more than two sequences [46]. Homologous proteins was searched for using the blastp web interface of "SIB BLAST Network Service" [47] with the sequence of nAChR from Torpedo marmorata as in the PDB file 2bg9. The sequences each subunit of nAChR from Torpedo marmorata (PDB code: 2bg9) was used as input, and sequences were searched "with within the

Swiss-Prot curated sequences" and "exclude the fragment sequences" options. Only the sequences with similarity above 35 per cent were retrieved, this set is called the standard sequence set.

The actual MSA calculations were performed by the web interface of ClustalW (version 1.83) [49] using the default values. The resulting outputs of MSA calculations were used as input files for conservation and correlated mutation analysis.

## 2.4. Conservation Analysis

Conservation is the resistance of residue to withstand the evolutionarily mutations [39]. Mutations at conserved sites either render the molecule impaired or seriously affect the molecules ability to perform its function. Therefore mutations at these sites are rarely passed onto the next generation. The conserved residues are often structurally or functionally important residues [40]

Conservation analysis uses sequences of a protein family or super-family to search against a base sequence for conserved residues, which are conserved at the same position on different proteins. In this work conservation analysis is performed using the third version of "Consurf" server at the Bioinformatics Unit of Tel Aviv University [39, 40]. To conform with other analysis the standard sequence set and its MSA was used as input files and the sequence of nAChR was analyzed for conserved sites.

The output of Consurf has two score types and two confidence intervals for each type. The first score is the conservation score, which shows the rate of evolutionarily change. Its ranges from negative real numbers to positive real numbers. Low scores show more conservation. Conservation score is normalized to yield an average of zero and a standard deviation of one [39]. Therefore, conservation scores can not be used to compare different structures' (or different subunits, or results from different runs) conservation status. The second score type is derived from the conservation score, this score divides the range of conservation scores into nine bins, first being the least conserved and nineth being the most conserved. Both type of scores have confidence

intervals which show the validity of the predicted value. Confidence information is vital against the possibility that too few sequences (containing a specific site) or too little diversity among the sequences might interfere with the results, and produce unreliable information. Therefore only the residue that fit into nineth bin and have confidence interval of "9,9" is assumed to be a highly conserved residue in the present analysis. The number of these highly conserved residues depends on the range of conservation score and the distribution of scores. Therefore a residue is only highly conserved relative to other residues in the particular structure or domain being analyzed, and a highly conserved residue in a domain might not be as conserved as another highly conserved residue in another domain.

## 2.5. Correlated Mutations

Correlated mutations refer to the coupled mutation of two residues, for which the mutation of the second residue compensates for the loss of function caused by the first mutation [48, 50]. These couples could be formed near neighbours or distant neighbours. These distant couples may reflect an allosteric regulation mechanism [51].

Correlated mutation analysis was performed using the web interface of "CorrMut" server at the Bioinformatics Unit of Tel Aviv University [41]. The default parameter was used along with the MSA from the ClustalW server. The results were filtered using a minimum correlation value of 0.5. The results were graphed using the Dot and Dotty programs [52].

## 2.6. Mcpool

MCPOOL is a program written to analyze the cross correlation outputs of GNM results. Cross correlation data includes essential information about the coupled motions of molecular regions. It's possible to analyze the relations between distant and close regions using a cross correlation map; but the vast amount of information present about the motions, makes it difficult to disclose a network of allosteric signals between remote residues. MCPOOL searches through this vast network for the most probable

paths of communication. This search begins with the generation of a pool of paths of given length and minimum cross correlation value (mincor), using an implementation similar to the Monte Carlo chain generation method, starting from a point or region (**SR**) on the cross correlation map. The generation of any given path start with the identification of possible pairs (**PC**), a possible pair represents a possible step that could be taken at a certain point on the path, using

$$\mathbf{PC}_i = (\mathbf{SR}_n, \mathbf{R}_k) \quad \text{where} \quad \begin{cases} \mathbf{crosscorr}(\mathbf{SR}_n, \mathbf{R}_k) > mincor \\ |\mathbf{SR}_n - \mathbf{R}_k| > gn \end{cases} \tag{2.21}$$

where **R** is the vector containing all of the residues present on the cross correlation map (with $Nr$ elements, generally total number of residues) and **crosscorr** is the matrix containing all of the cross correlation values of residue pairs. $gn$ limits the minimum number of residues separating $\mathbf{SR}_n$ (which is the residue number of $n^{\text{th}}$ residue in the starting region) and $\mathbf{R}_k$ (which is again a residue number). In this work $gn$ is taken as three. The reason for assuming $gn$ three is :

- it is not desirable to identify **PC** in the close proximity of starting region, this ends up in local paths
- it is not desirable to give a greater value for $gn$, as this may results in skipping neighbours with high cross correlation values with distant residues.

The minimum cross correlation value, a step in a random path may have, is $mincor$. The value of $mincor$ is taken as 0.5, because 0.5 is low enough to include the lower correlation values while it does not over saturate **PC** with low correlated pairs. Using equation 2.21 all of the possible pairs are identified, then the total cross correlation value ($Z$) of all $\mathbf{PC}_i$ pairs are calculated using

$$Z = \sum_{i=1}^{Npc} \mathbf{crosscorr}(\mathbf{PC}_i) \tag{2.22}$$

in which $Npc$ represents the number of $\mathbf{PC}_i$ identified using equation 2.21. Using $Z$ each cross correlation value of $\mathbf{PC}_i$ is normalized using

$$\mathbf{PoC}_i = \mathbf{crosscorr}(\mathbf{PC}_i)/Z \tag{2.23}$$

where, $\mathbf{PoC}_i$ represents the possibility of $i^{\text{th}}$ pair. Using $\mathbf{PoC}$ a range is setup for each $\mathbf{PC}_i$. The total range of the all $\mathbf{PC}_i$ is between zero and one. The range for a $\mathbf{PC}_i$ is setup using

$$\mathbf{RPoC}_{i,1} = \begin{cases} 0 & \text{for} \quad i = 1 \\ \mathbf{PR}_{i-1} & \text{for} \quad i \neq 1 \end{cases} \tag{2.24}$$

$\mathbf{RPoC}$ is the matrix which stores the range for each $\mathbf{PC}_i$, the first element ($\mathbf{RPoC}_{i,1}$) for $\mathbf{PC}_i$ is the starting point of the range, the second element is calculated using

$$\mathbf{RPoC}_{i,2} = \begin{cases} \mathbf{PR}_i & \text{for} \quad i \neq Npc \\ 1 & \text{for} \quad i = Npc \end{cases} \tag{2.25}$$

where $\mathbf{PR}_i$ term in equations 2.24 and 2.25 can be defined by

$$\mathbf{PR}_i = \sum_{j=1}^{i} \mathbf{PoC}_j \tag{2.26}$$

This method biases the possibility of each $\mathbf{PC}_i$ according to the cross correlation values. Then a uniform random number generator with a uniform distribution is used to select one element $\mathbf{PC}$. This basic scheme is used to create a pool of paths with given lengths.

The Monte Carlo path generation method of MCPOOL was checked against a true random path generator in which possibility of a residue appearing is not biased using the cross correlation values. In the true random run the possibility of any residue to appear at a given point on a path should be $1/Nr$. The possibility calculated from

the results of the true random run was approximately equal to $1/Nr$. However, the possibility of a given residue appearing at a certain point on a path, calculated from the results of MCPOOL significantly differs from latter results. The possibilities calculated from the results of MCPOOL closely resemble its source, the cross correlation map.

This pool of generated paths is searched for "connected paths". These connected paths are a subset of the generated paths which connect the starting region with the preselected target region. Then these connected paths are analyzed, on a residue basis, for important residues. This analysis results in two outputs:

1. A file containing the occurrence frequency of each residue on connected paths.
2. A file containing the connected paths in PAJEK [53] network format.

The first file, which contains the occurrence frequency of each residue (on the paths with defined initial and target regions) on the connected paths is used to analyze the important residues which are key to allosteric signalling. The occurrence frequency of a residue represents the connectedness of a residue to its neighbours. A residue with high occurrence frequency represents a crucial point on connected paths.

The second file is used to visualize and analyze these connected paths (paths with defined initial and target regions extracted from the initial pool) using the PAJEK program [67.9] which is a software for analyzing large networks. This file contains all of the nodes (residue) connected with arcs (steps). Each arc is given a value according to is occurrence frequency in the connected paths. The network, described in this file, is analyzed using the occurrence frequency of an arc. If an arc has a high occurrence frequency it may be referred as an essential step on the pathway from the binding site to the gate region. In PAJEK these frequencies are normalized using the maximum frequency value and normalized occurrence frequencies smaller that 0.1 is removed. For presentation purposes the values are multiplied by three. The resulting network is a representation of major steps on the pathway from binding region to gate.

Other useful information can be retrieved from the analysis of connected paths, such as the preference to use an external connected path, which is a connected path that includes at least one residue not on the subunit of starting region, or an internal connected path, which is a connected path that includes only residues that are on the subunit of starting region, for targets that are on the subunit of starting region.

# 3. RESULTS AND DISCUSSION

## 3.1. Conservation Analysis

Conservation analysis of nAChR from *Torpedo marmorata* was done using the sequence of 2bg9 (PDB code: 2bg9). This sequence was search for homologues with the web interface of BLASTP [47]. The resulting sequences were aligned by ClustalW using the default values [49]. Conservation analysis produced two types of results. The first is the normalized conservation score. The second score type is derived from the first results. The second score type represents the normalized conservation scores in nine distinct bins. First bin contains the least conserved residues and ninth bin contains the most conserved residues. Each residue in a bin has a confidence interval. This confidence interval represents the actual range of bins a residue can reside in. An example would be that, a residue with eight as a second type score has a confidence interval of five to nine (5,9). This means that there is not enough information to pinpoint the exact score of this residue. Therefore in this work only residues with a confidence interval of nine to nine (9,9) is considered as a highly conserved residue. In figure 3.1, a type of averaging called the "10 period moving average" (10 pe Mov. Ave.) (See figure 3.1. This type of averaging appear after the tenth point in the graph with the average value of the first ten points (such as normalized conservation scores), the second point on the "10 per. Mov. Avg". line is the average of ten normalized conservation scores starting from the value of the second point on the normalized conservation scores. In figure 3.1 only highly conserved residues are labeled. Stars represent the residues which are highly conserved in the other three non-alpha subunits. Rhombus represent the residues which are highly conserved in two of three other non-alpha subunits.

Conservation analysis showed that, except three sites, highly conserved sites in the alpha subunit are also conserved at the equivalent positions in beta, delta, and gamma subunits (See Table 3.1). However, those three sites on alpha subunit (Glu241, Ser268, Val405) are highly conserved in two of three non alpha subunits. Figure 3.1 shows the conservation distribution and highly conserved residues of alpha subunit of

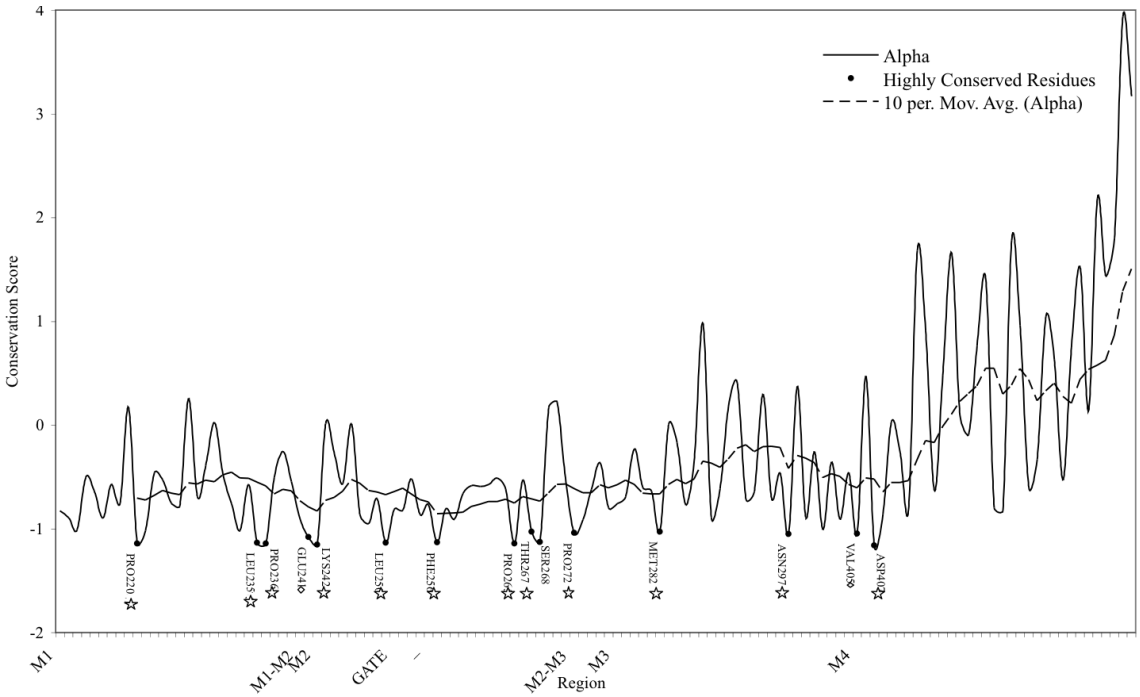


Figure 3.1. Conservation scores (normalized) of alpha TMD (solid line) along with the moving average of these scores (dashed line).

nAChR. In all the subunits M2 helices are the most conserved, while M2 helices and M3 helices are the second and third most conserved among the four transmembrane helices (See Figure 3.1, and figures B.1, B.2, B.3 in Appendix B). M4 transmembrane helices are the least conserved among all transmembrane helices which is consistent results of similar studies [24]. Most of the highly conserved resides, which line up with equivalent positions in other subunits, are hydrophobic (See Appendix C for MSA of nAChR subunits). However, there are residue which are not hydrophobic, such as aspartic acid (Asp), asparagine (Asn), lysine (Lys). Among these residues aspartic acids are on the M4 helix, Asparagines are on the M3 helix and lycines are the first residues on M2 helices. All of these residues are near the cytoplasmic domain. When plotted on the 3D structure, the highly conserved residues, which are highly conserved on all subunits, form four rings (See Figure 3.2). These four rings are respectively (from LBD domain to cytoplasmic domain) at the interface of TMD and LBD on M2-M3 loop, one residue before the gate section, one residue after the gate section, and at the interface of cytoplasmic domain and TMD. Rings adjacent to the narrowest section of the pore are formed by hydrophobic residues. Polar residues are only on the first and the last rings. Also the third ring approximately marks the middle of M2 helices.

Table 3.1. Highly conserved residues in TMD of nAChR, bold residues are conserved in all subunits, while italicized residues are conserved in three. The last column shows the ring which the residues on the same row contribute to.

| **Alpha** | **Beta** | **Delta** | **Gamma** | **Ring** |
|---|---|---|---|---|
| - | - | - | Phe220 | - |
| - | *Tyr220* | *Tyr228* | *Tyr221* | - |
| **Pro221** | **Pro227** | **Pro235** | **Pro228** | 2 |
| - | *Cys228* | *Cys236* | *Cys229* | - |
| **Leu235** | **Leu241** | **Leu249** | **Leu242** | 4 |
| **Pro236** | **Pro242** | **Pro250** | **Pro243** | 4 |
| - | - | Gly254 | Gln249 | - |
| *Glu241* | *Glu247* | *Glu255* | - | - |
| **Lys242** | **Lys248** | **Lys256** | **Lys250** | - |
| - | - | - | Ile255 | - |
| - | - | - | Ser256 | - |
| **Leu250** | **Leu256** | **Leu264** | **Leu258** | 3 |
| **Phe256** | **Phe262** | **Phe270** | **Phe264** | 2 |
| - | - | Leu272 | Leu265 | - |
| - | - | - | Phe266 | - |
| **Pro265** | **Pro271** | **Pro279** | **Pro273** | 1 |
| **Thr267** | **Thr273** | **Thr281** | **Thr275** | 1 |
| *Ser268* | *Ser274* | Ala282 | *Ser268* | - |
| **Pro272** | **Pro278** | **Pro286** | **Pro280** | 1 |
| - | Ile279 | - | Leu281 | - |
| **Met282** | **Met288** | **Met296** | **Met290** | 2 |
| - | Ser294 | Val302 | - | - |
| **Asn297** | **Asn303** | **Asn311** | **Asn305** | 4 |
| - | *Arg307* | *Arg315* | *Arg309* | - |
| - | - | Gly450 | - | - |
| *Val405* | *Val434* | Thr452 | *Val445* | - |
| **Asp407** | **Asp436** | **Asp454** | **Asp447** | 4 |
| - | - | - | Gly461 | - |

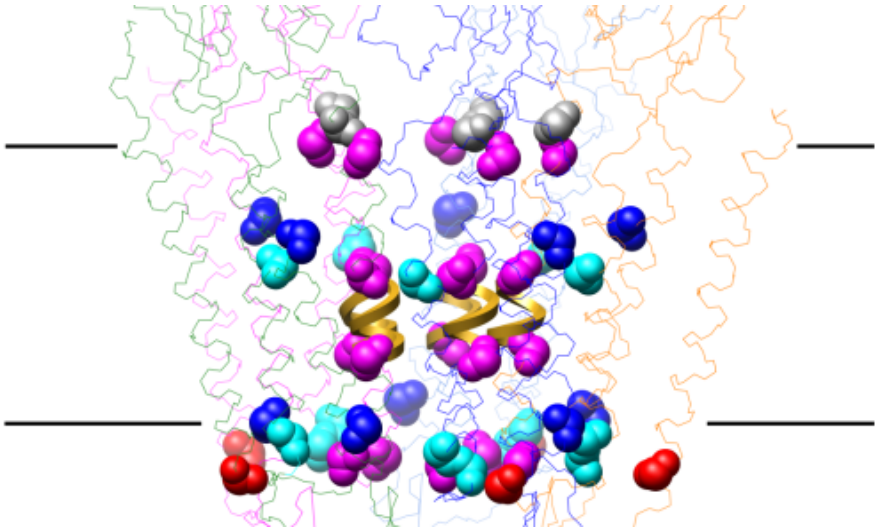

Figure 3.2. Highly conserved residues are displayed with a sphere. The golden ribbon is the gate. Spheres are on M1 (cyan), M2 (magenta), M3 (blue), M4 (red). Black lines represent the cell wall.

### 3.2. Correlated Mutations

Correlated mutations are changes with compensating effects. The mutation of a residue at one site causes a disturbance which is compensated for by the mutation of another residue at another site. These types of mutations can be identified computationally by analyzing the sequences of a family of homologous sequences.

CorrMut [41] has been employed here to search through the sequences for correlated mutations. CorrMut analyzed nAChR from *Torpedo marmorata* using the sequence of 2bg9 in five subunits. This sequence was search for homologues with the web interface of BLASTP [47]. The resulting sequences were aligned by ClustalW using the default values [49]. This alignment was used as the input file of CorrMut. A minimum correlation coefficient of 0.5 was chosen. Figure 3.3 shows the results of correlated mutations at the TMD of alpha subunit.

The helix with minimum correlated mutations is the M2 helix. This is explained by the fact that M2 is more conserved than other helices. Serine at the 287 position (See Figure 3.4, green sphere) on the M3 helix is the most connected residue. Figure 3.3 shows the network of correlated mutations. As Figure 3.3 suggests, apart from the

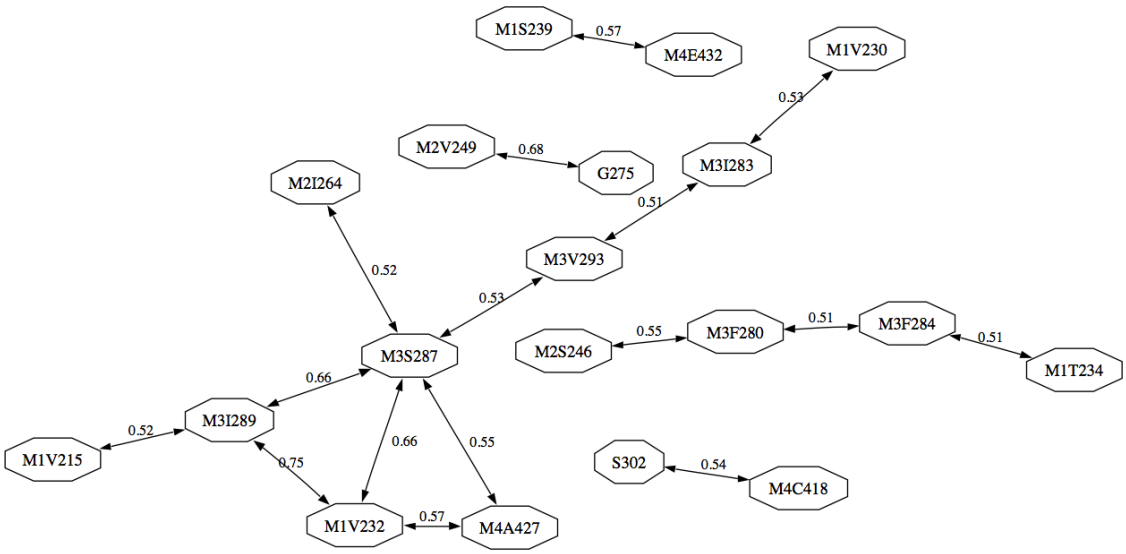


Figure 3.3. The Network of correlated mutations for alpha subunit of nAChR with correlation coefficients (above lines).

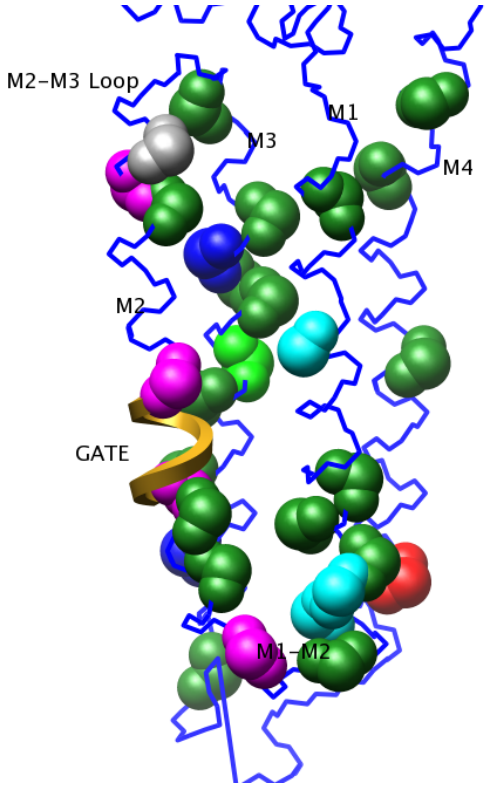


Figure 3.4. Correlated mutations (dark green) and conserved residues (not green) on the alpha subunit of nAChR.

conserved residues, alpha subunit of nAChR has a network of correlated mutations. Ser287 of M3 helix is correlated with residues from M1, M2, and M4 helices. Ser287 forms a closed loop of correlated residues with I289 on M3 helix, V232 on M1 helix, and A427 on M4 helix. A closed loop may represent a cluster of residues which is important for the function. It's noteworthy that residues on M1 and M4 helices do not have a direct correlated residue on M2 helix. Residues on M1 and M4 correlated with residues on M2 helix via residues on M3. A recent study showed that mutation of Phe426 on M4 helix, which is adjacent to A427 on M4 helix, affects the gating [29]. In addition a recent MD study proposed that M4 acts as means to transfer the lipids modulating effects to the channel [32]. It's also interesting that Ser287 on M3 helix is at the junction in between I264 on M2 helix and residues from the other subunits. This may suggest that Ser287 is a conduit in between M1 and M4 and M2 helices. Figure 3.4 show the residues identified by conservation and correlated mutation analysis (CMA) together for alpha subunit. Conserved and correlated sites form a network from the M2-M3 loop and from lipid facing helices to the gating region.

Correlated mutation sites can be seen on the M1-M2 and M2-M3 loops (See Figure 3.4). These include the coupling of the Gly275 on the M2-M3 helix with the Val249 near the gate region on M2, and the coupling of Ser239 on M1-M2 with the Glu432 on M4. M2-M3 region has been noted to be important to the transfer of allosteric signal from the LBD region to the TMD [7, 12, 29, 30, 31]. Also, several studies point out that M1-M2 loop is on the signaling pathway [16].

Correlated mutation analysis of Beta subunit of nAChR differed from that of alpha subunit. In Beta subunit M1 and M4 helices are coupled with the residues on the M2 helix (See Figure B.4). In addition, Beta subunit has a similar closed loop of correlated residues. These residues are Leu456 on M4, Val238 on M1, Phe293 on M3, V295 on M3. This closed loop is connected to M2 helix at Ser252 via Phe286 on M3. It's also noteworthy that, M4 helix has several correlated mutations with residues on M2 helix. The positions of the correlated mutations can be seen on Figure 4. M2-M3 loop is correlated with gate region on the M2 helix by the correlation of a residue juxtaposed to M2-M3 helix, Iso281. This residue is correlated with the Ala255, which

is near the proposed gateway. It's also notable that one end of M2-M3 loop is the site of a correlated mutation (Ile281) while the other end is conserved (Thr273). The conserved proline (Pro278) is also notable.

The results of correlated mutation analysis for delta subunit are represented schematically on figure B.6. Same correlated mutation pairs, present in all subunits, exists in Delta subunit of nAChR. A residue on the M2-M3 linker is conserved and this residue (Pro286) is flanked by (See Figure B.7) the two residues identified by the correlated mutation analysis (Leu283 and Gly289). Both of these residues are correlated with residues on M2 helix near the gate region. Also, M1-M2 loop is flanked by a conserved residue (Lys256) and a residue identified by CMA (Ser253). Figure B.7 show the positions of residues identified by CMA on the 3D structure of nAChR (PDB Code: 2bg9). M4 and M1 helices of delta subunit also have correlated residues with residues from M2 helix, there is direct coupling between M4, M1 and M2 helices as in beta subunit.

The CMA analysis of gamma subunit from the nAChR showed no direct correlations mutations between the residues of M4 and M1 with M2. As in alpha subunit, M1 correlates with M2 through M3 helix (See Figure B.8). M4 helix has a single correlated mutation with the loop after M3 helix. Similar to all subunits M2-M3 helix has a conserved residue flanked by two residues identified by CMA (See Figure B.9). These residues are Leu277 correlated with Leu294 and Val257 correlated with Gly283.

Correlated mutation analysis was performed on the nAChR and the correlated residues with a correlation coefficient higher than 0.5 were analyzed. In all four different subunits of nAChR alpha has the highest number of correlated couples. Residues on M1 and M4 helices of alpha and gamma subunits correlated with residues on M2 helix through a correlation with M3 helix. In beta and delta subunits, residues on M4 helices have direct coupling with the residues on M2. On delta subunit residues on M1 helix have pairs identified by correlated mutation analysis, with M2 residues.

## 3.3. GNM Results

### 3.3.1. Contact Map

Contact map of the TMD of nAChR from the *Torpedo marmorata* was calculated using the 3D structure from 1oed (See Figure D.1). Each subunit has extensive intra-subunit contacts. Inter-subunit contacts are made with neighboring subunits. Most of the inter-subunits contacts are contacts of M2 helices of neighboring subunits (See Figure D.1, parallelogram). Inter-subunit contacts include the contacts between M3 helices with the M1 helices of subunit at the counter clockwise position, when viewed from the extracellular side (marked by an oval on Figure D.1). M2 helices of each subunit have a number of contacts with the M1 of neighboring subunits (marked by an octagon on Figure D.1). The number of contacts between individual helices of TMD of nAChR can be seen on figure 3.5.

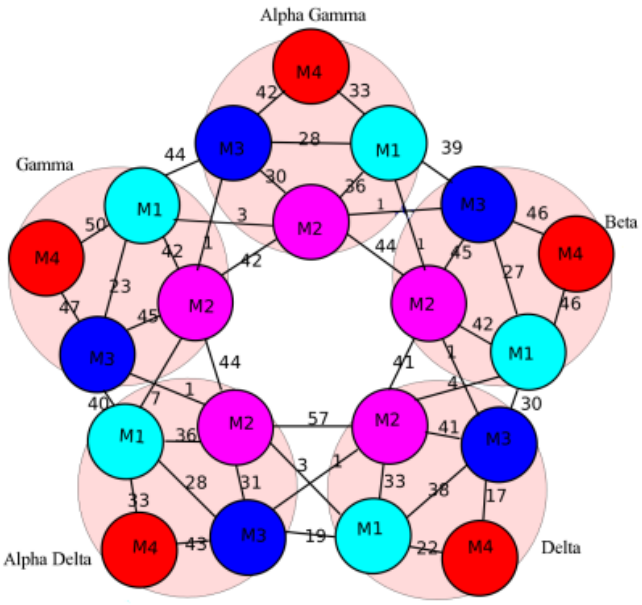


Figure 3.5. Contact network and number of contacts for TMD of nAChR.

The contact network represented by Figure 3.5 show that M4 helices have no contacts with M2 helices, but M2 subunits are connected to M4 subunits via the extensive contact network. M1 and M2 helices have extensive contacts with M2 helices in their subunit, and inter-subunit contacts are dominant between M2 helices. M4 helix of delta subunit has the least number of contacts. It has less than half of number contacts made

by any helix in the TMD domain. Also, delta subunit has the least number of contacts with the neighboring subunits.

### 3.3.2. Isolated M2 Helices

M2 helices (inner ring) of nAChR have been analyzed without the other TMD helices. This analysis was done with GNM. The resulting slow and fast mode fluctuations were analyzed for intrinsic flexibility. The slow mode fluctuation of the first two slowest modes suggested that the upper halves (the half near the extracellular side) of all M2 helices are more mobile than the lower halves (the half near the cytoplasmic side) (See Figures 3.6).

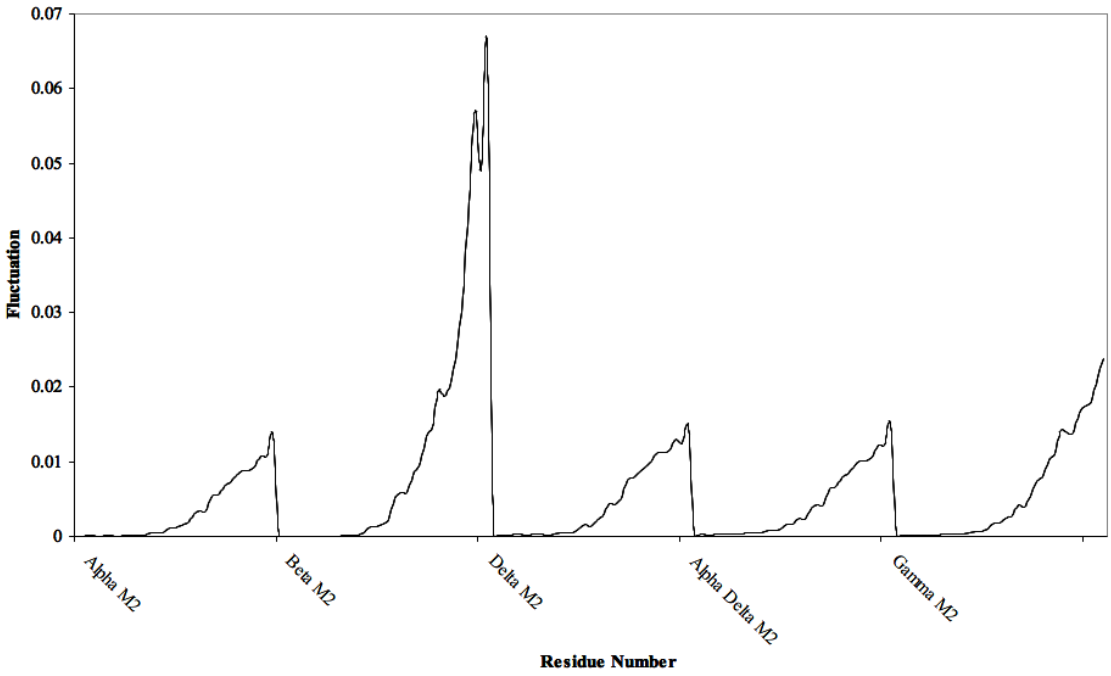


Figure 3.6. Mean square fluctuations of the average of the lowest two modes for M2 helices of nAChR.

This stationary lower half and mobile upper half may be the results of the tapering structure. The tapering structure of the TMD domain of nAChR lead to fewer contacts at the upper half. M2 helix of the beta subunit is the most mobile among the M2 helices at the upper half. Also, beta M2 helix has the smallest number of contacts with the other M2 helices, but it has only one less contact than alpha M2 and Gamma M2

helices (alpha gamma M2 helix has 86, beta M2 helix has, 85, delta M2 helix has 98, alpha delta helix has 101, and gamma M2 helix has 86 contacts).

The cross correlations of M2 helices also show signs of the tapering structure of TMD domain (See Figure D.2). Figure 3.7 shows the key for interpreting cross correlation maps.

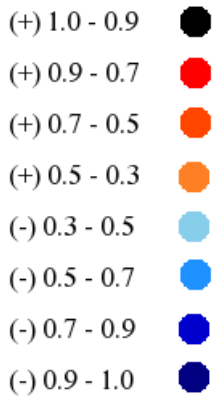


Figure 3.7. Key for correlation coefficients ranges and the representative colors.

The lower halves of all M2 helices are positively correlated, while the upper halves of neighboring subunits do not show correlation and non neighboring subunits have a weak negative correlation. The tapering structure leads to less contacts in the upper half and this in turn leads to very weak correlations (white spaces). The lack of intra-subunit correlation between the upper half and the lower half represents the difference in motions of the two parts of the subunit.

### 3.3.3. M1, M2, and M3 helices

M1, M2, and M3 helices make up the inner and middle rings of the TMD of nAChR, their structure should contain the key elements to allow for channel opening, since M4 helices do not have a single contact with M2 helices, M4 helices can not alter the structural mechanism dramatically. The mean square fluctuations of the slowest two modes of show a similar fluctuation scheme to the fluctuations of isolated M2 helices, the lower half of each helix is more stationary than the upper halves (See Figure 3.8).

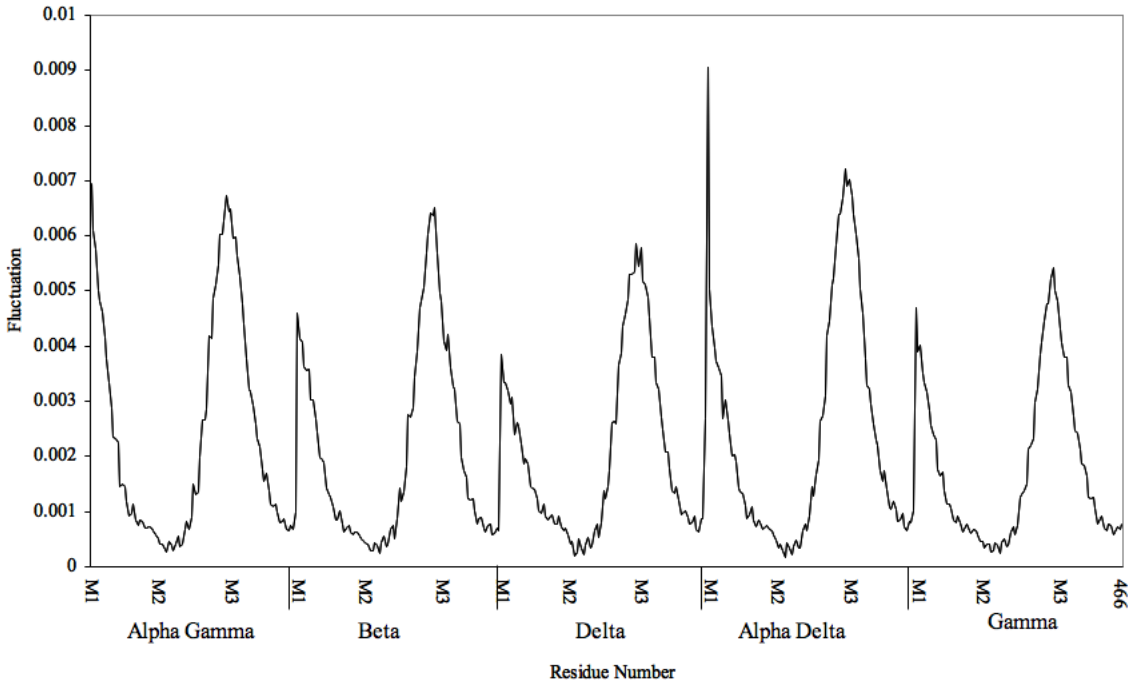


Figure 3.8. The mean square fluctuations of M1, M2, and M3 helices of the TMD in the average slowest two modes.

The high fluctuations of starting residues of each M1 helix is the results of missing connection (a covalent bond) with the LBD. M2-M3 loop (between the end of M2 and starting position of M3) shows high mobility in each subunit. The results of MD simulation of TMD in a membrane mimicking environment also found M2-M3 loop to be region with highest amplitude fluctuations [33], they suggested that this may be a requirement for transmitting structural changes to the TMD from LBD. Also, their MD simulations exhibited a correlation between the fluctuations of M2-M3 loop and the twisting bending motion of M2 helices, this correlation is present in the cross correlation map of M1, M2, and M3 loops (See Figure D.3, region inside the oval), suggesting that these helices have the structural requirements of mechanics of gating.

The M1, M2, and M3 helices (inner and middle rings) have been analyzed using GNM to identify the effects of M4 helices (outer ring). The correlation of M2 helices of M2 subunit show results similar to the analysis of isolated M2 helices, the correlation displays the tapering character of the TMD of nAChR. Upper half of M2 helices show no correlation (See Figure D.3 and Figure 3.9). The result of the tapering structure can be seen on the correlations of M1 helices. The effects of a tapering structure can

be seen within the positive correlations between subunits with no contacts. Positive correlations exist between the last residues of M3 helix (which are near the intracellular domain) and the M1-M2 loop. The lack of correlation between M1-M2 loop and M2-M3 loop can, also, be attributed to the tapering structure of TMD.

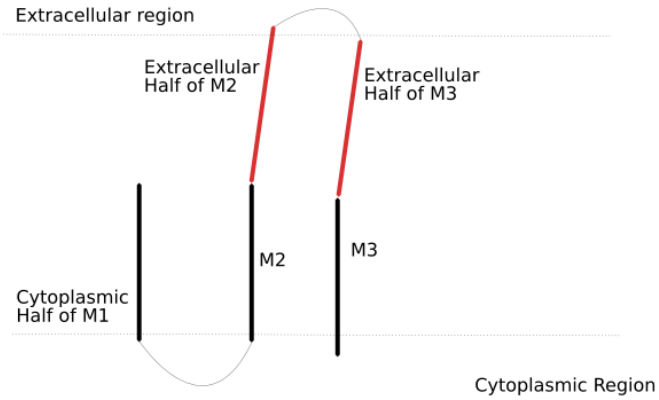


Figure 3.9. Results of tapering structure, the helices of a single subunit with little or no correlation.

Unwin et. al. [7] proposed that the twisting motion at the TMD domain weakens the symmetric forces acting on the pore lining residues, and pore widens when M2 helices collapse onto the middle and outer rings [17]. This suggests that there may be a relatively flexible hinge at or near the gate region. Once the interacting forces are broken by the twisting motion, this hinge may protrude toward middle ring and widen the pore. The tapering structure and the lack of correlation between the upper half and lower half support this idea of twisting motion opening the pore.

### 3.3.4. Transmembrane Domain of nAChR

Normal mode analysis was performed on the TMD of nAChR. The 3D structure of TMD of nAChR was taken from the PDB item 1oed. In its native environment the TMD of nAChR is embedded in the lipid bilayer of cell membrane [7]. To assess the changes imposed by the lipid bilayer two systems were constructed such that both systems have approximately same number of contacts per residue. In the first system each residue on the lipid facing side of nAChR was added a contact. This resulted in a shape similar to a five point star (See Figure 3.10).

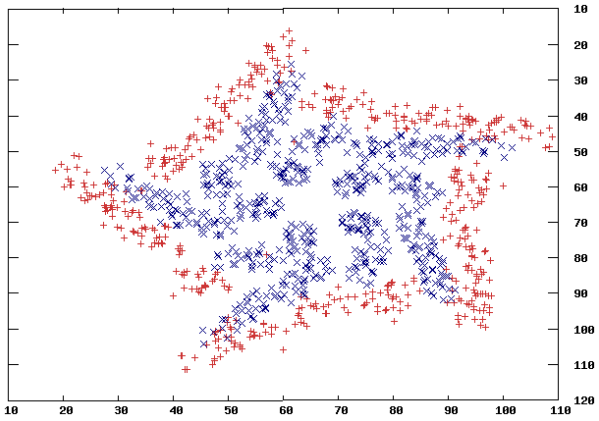


Figure 3.10. TMD of nAChR with additional contacts for every lipid facing residue to represent the lipid bilayer. Blue xs are the backbone atoms of nAChR.

The second system used a periodic repetition of equally spaced atoms to form a cube, then TMD of nAChR was embedded into this cube. The overlapping atoms of cube was removed (See Figure 3.11). This system has more additional atoms than the first system.

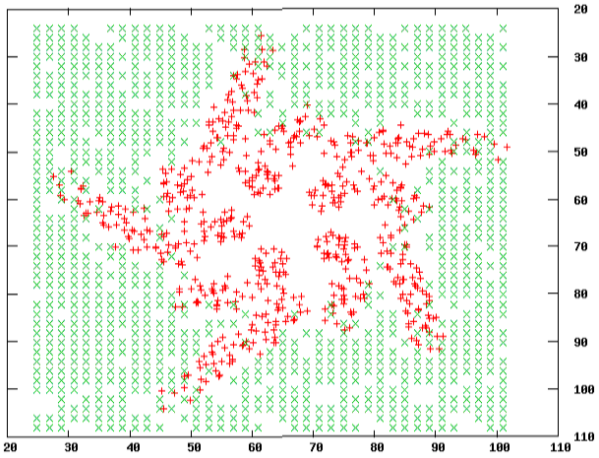


Figure 3.11. TMD of nAChR embedded in a cube of generic atoms to represent the lipid bilayer. Red crosses are the backbone atoms of nAChR.

The representations of the membrane environment use the fact that GNM does not differentiate atom types. Only one force constant is used for the interaction between individual atoms. Therefore, only the distribution and locations of atoms in the membrane is important to incorporate a membrane into the GNM analysis. These two systems were constructed to assess the general effect of nAChR in its native environment. Slow mode fluctuation of all systems are similar in a general sense, but the

amplitudes of fluctuations decreases as the number of residues increase (second system has the highest number of additional atoms) (See Figure 3.12).

The first difference between isolated TMD of nAChR and the embedded TMD nAChR is that the fluctuations of most fluctuating regions diminish, because of the additional contacts. Also, the hinge on the lower half of M3 helix (near M4) diminishes (increases its relative fluctuation) from isolated TMD of nAChR to the second system. The fluctuations of the loose end of M4, caused by the tapering structure, diminishes in the presence of lipid mimicking atoms to point that it fluctuates only as much as M2-M3 loop. Another important difference is that the smooth fluctuation difference between the lower end of M1 and lower half of M2 changes to an abrupt decrease. The mobility of the M2-M3 loop relative to the rest of the TMD does not change. The relative fluctuations of the lower half of M3 helix increases and relative fluctuations of M4 decrease to a similar level. This behavior may suggest that M4 helices transfer the effect of lipid bilayer to the channel through the M3 helices.

The cross correlation map of TMD of nAChR (See figure D.4) shows the results of slowest ten modes of nAChR, which corresponds to 23 per cent of the total dynamics. The presence of M4 helices increases the intra-subunit correlations. The presence of M4 helices increases the correlation of the region around M1-M2 loop (lower part of M1 and M2) to the region around M2-M3 loop (upper part of M2 and M3), but the individual loops still have limited correlation (See Figure D.3). A recent experimental study [31] points out that the Arg209 on the end of B10 is correlated to the M2-M3 loop by a residue on the loop two of LBD. This suggests that initial residues of M1 should also be correlated with M2-M3, since they are covalently bonded to Arg209 in the whole structure of nAChR. This correlation of initial residues of M1 and M2-M3 helix was missing in the absence of M4 helices but it is present in the presence of M4 helices.

The inter-subunit correlations include the correlations of neighboring and non neighboring subunit. Each neighboring subunit has extensive correlations at the lower end of M2 and M3 helices. Non neighboring subunits have high positive correlations

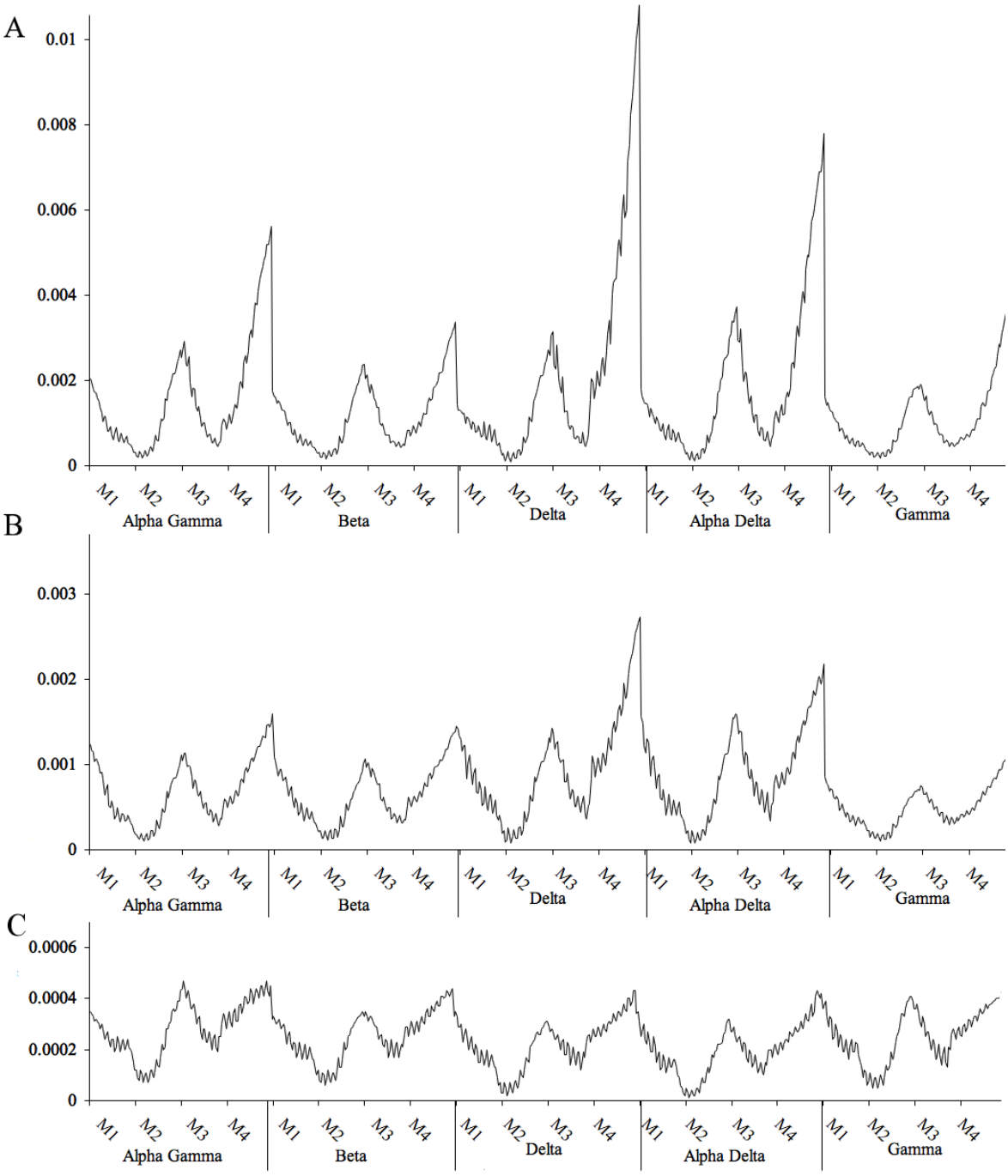


Figure 3.12. The mean square fluctuations of the average of the first two modes of TMD of nAChR isolated (A), in the first type of environment (B), and in the second type environment (C).

between the lower halves of M1 and M2 and lower half of M2 helix and the lower half of M3 helix. The high correlation between the subunits at the lower half of TMD of nAChR may suggest that the M1-M2 loop plays an important role in the gating.

The additional atoms result in a more correlated intra-subunit correlations for all subunit except for beta and gamma subunits, in which the correlations decrease in the first system. In the second type system all of the intra-subunit positive correlations increase. Also, in this system the inter-subunit positive correlations increase to great extend, while in the first type system the negative correlations increase.

These observations suggest that even if the lipid contacts may change correlations between regions of TMD of nAChR to a certain extent the basic intrinsic flexibility remains.

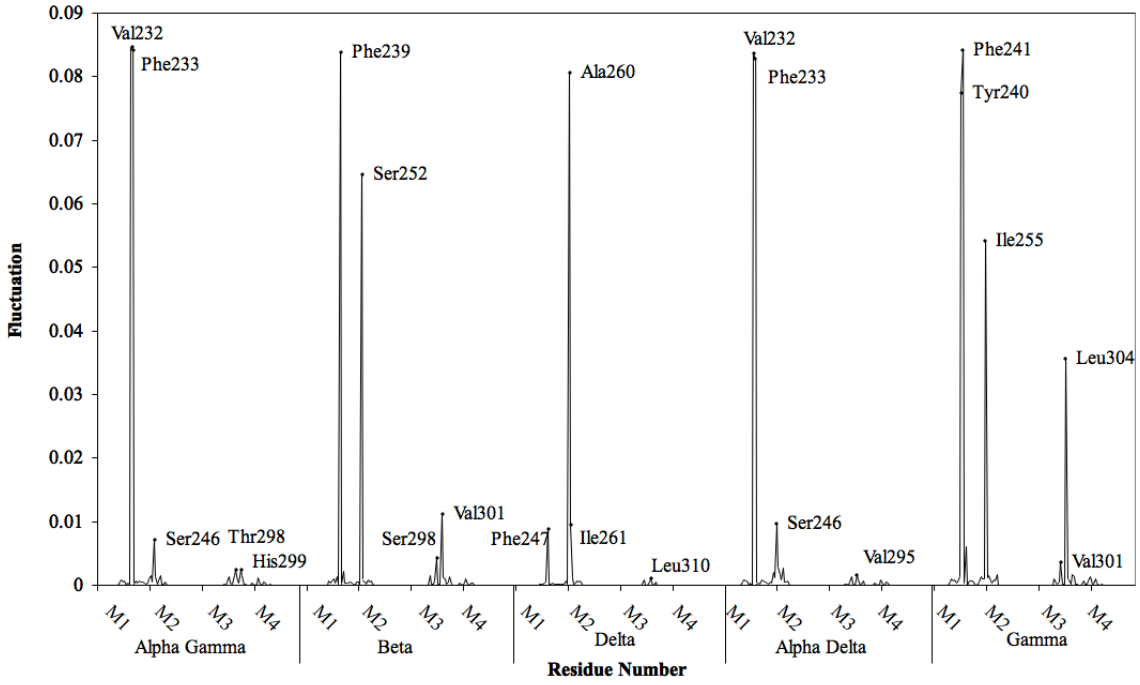


Figure 3.13. Average of fasted 10 fast mode mean square fluctuations of the TMD of nAChR.

The residues which are active in the fast modes may be referred as the hot spots of the protein, which are important for the structure [11]. The identified residues, active in the fast modes, are labeled on figure 3.13. These residues cluster at the lower half of the TMD under the level of the gate (See Figure 3.14). The positions of these residues

and the physical meaning attributed to these active residues lead to an idea that the lower half of the TMD is structurally important while the upper half of the TMD is functionally important. This idea may be supported by the fact that the binding region is at the extracellular side and that the signal from binding site is transferred through the upper half of the TMD.

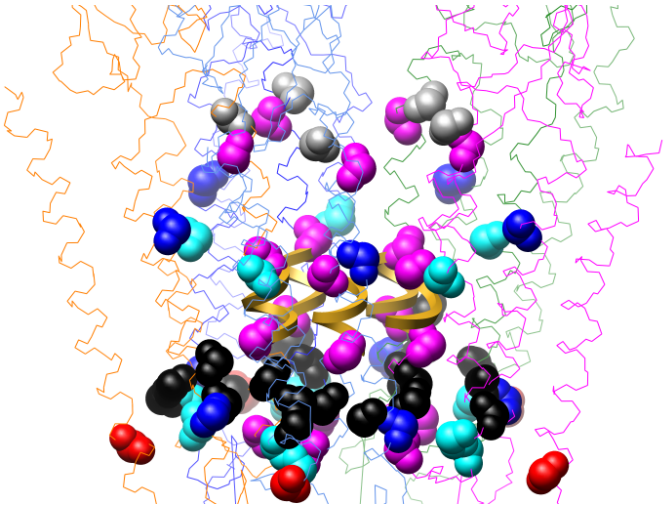

Figure 3.14. The conserved residues (colored spheres) and the residues active in the fast modes (black spheres). The gate region is shown in golden ribbons.

### 3.3.5. Transmembrane Domain results from 2bg9

Recently Unwin et. al released a more extensive 3D structure of the nAChR from *Torpedo Marmorata* [10] (PDB code: 2bg9). This new structure has the LBD, TMD and some residues of the cytoplasmic section together. GNM analysis had been performed on this new and more complete structure to elucidate the effect of LBD and the cytoplasmic domain on the TMD of nAChR. The analyses were performed using 2bg9, and TMD was extracted from extracted from the results. The mean square fluctuations of the average of the slowest two modes of TMD of nAChR as 2bg9, showed different results than the slow mode results of TMD of nAChR isolated from LBD and TMD (See Figure 3.15). The most distinct difference is the absence of a hinge at M1-M2 loop, this loop was relatively stationary in the isolated analysis of TMD of nAChR. The main reason behind this change the cytoplasmic helices (MA) present in the structure in 2bg9. These helices have few contacts and can freely move in the structure of 2bg9. They are covalently bonded to M4 helices and make contacts

with M1-M2 loop and M3-M4 region. These contacts might transfer the exaggerated fluctuations of MA to the M1-M2 loop and M3-M4 region (See Figure 3.15).

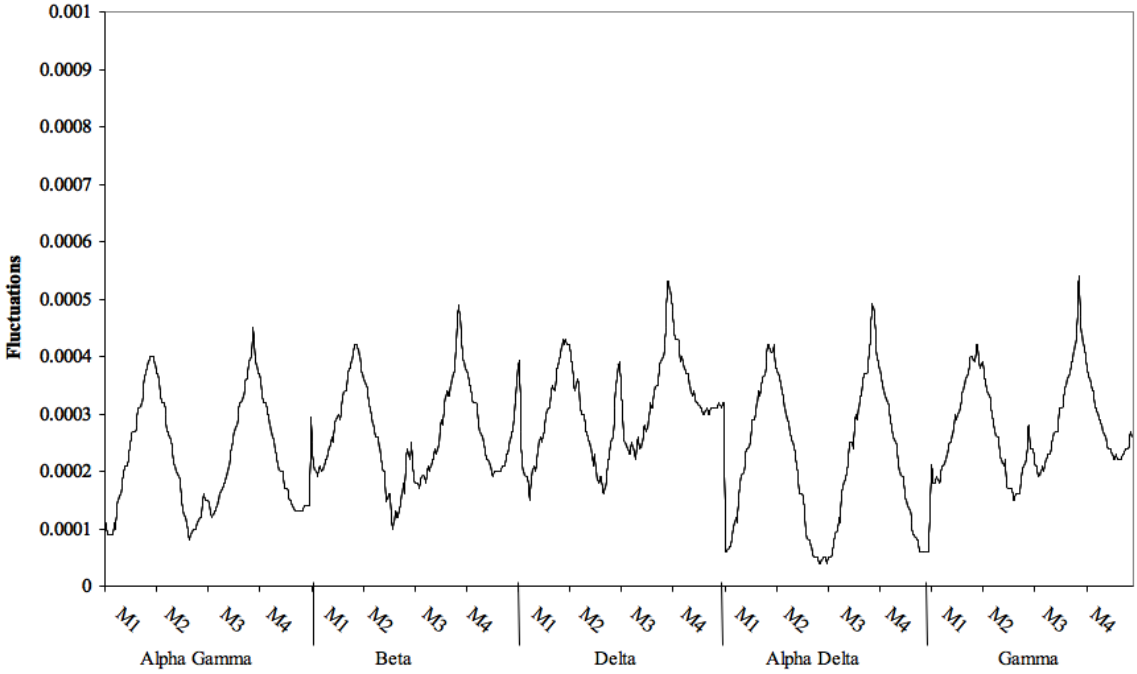


Figure 3.15. Slow mode mean square fluctuations (average of slowest two modes) of TMD of nAChR extracted from the GNM analysis of 2bg9.

Another noticeable difference is that the hinge at the region between M3-M4 has become a highly fluctuating region (presence of MA) and the flexible M2-M3 loop lost the flexibility to a great extend in alpha subunits and to a lesser extend in non-alpha subunits. These differences may be the result of the coupling between the cytoplasmic helices and TMD and LBD. Some studies suggest that the Cys loop and loop two together torques the M2-M3 loop and that Cys loops acts as the pivot point in between LBD and TMD [12, 16, 31, 35]. These suggestions support the results of GNM analysis of 2bg9. M2-M3 is stationary relative to the regions between M1-M2 and M3-M4, because it is interacting with the Cys loop and loop two. The M2-M3 loop in alpha delta subunit lies on a plateau of lower than any part in the TMD. This plateau extends from the gate on the M2 helix to the M2-M3 Loop. In alpha gamma subunit M2-M3 is more flexible than the M2-M3 loop in alpha delta subunit, but both are stationary when compared to the rest of the TMD. Several studies suggest that the activation of the channel is triggered with separate contributions from alpha subunits [6, 7]. This suggestion might explain the relative stationary structure of the M2-M3 loop to the

other M2-M3 loops. A rigid body is more effective in transmitting changes, than a flexible body.

Noticeably the last residues of M4 helices in all subunits show more restricted fluctuations than the analysis of isolated TMD of nAChR depicted, and in both alpha subunits these last residue of M4 forms stationary regions. This clearly shows that M4 subunits have contacts with the LBD, and might suggest that M4 interacts with LBD. Several studies, both experimental and computational, suggest that M4 conveys the effects of environment to the channel [27, 29, 32, 54].

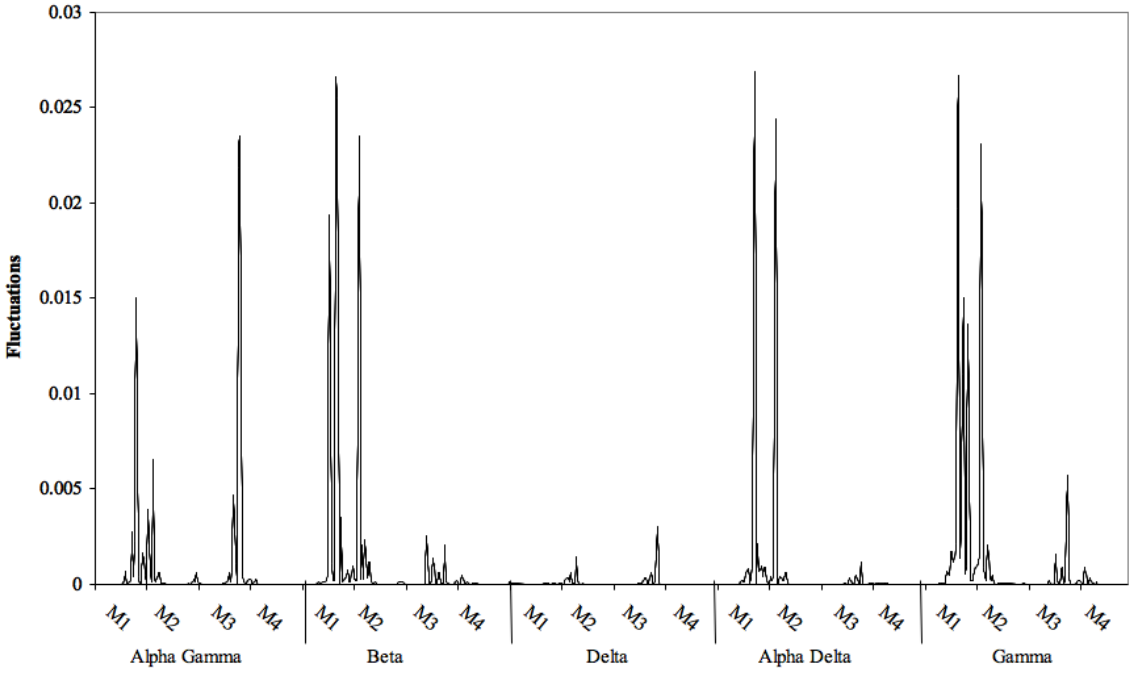


Figure 3.16. The fast mode fluctuations of the fastest 30 modes of TMD of nAChR extracted from the analysis of 2bg9.

The fast mode fluctuations of the TMD from the analysis of 2bg9 shows similar results to the fast mode (See figures 3.13, 3.16) results of isolated TMD of nAChR. The residues active in these modes are clustered similarly. The residues resides mostly in the lower half of the TMD.

The cross correlation results of the first ten modes of TMD from the analysis of 2bg9 (represents the 20 per cent of dynamics) is dominated by positive correlations

of each transmembrane region. Only the inter-subunit correlations of beta and delta subunits with other subunits have negative correlations (See Figure D.5) M2-M3 loop of beta and delta subunits with most of the regions of alpha subunits and gamma subunit.

The correlation map suggests a tightly integrated TMD. Except for the delta subunit M4 of all subunits have high intra-subunit correlations of M4 with its' M2-M3 loop (grey boxes on Figure D.5). The highest inter-subunit correlation are mostly between the lower half of M1 and M2 with the lower half of M3 and M4.

The results of GNM analysis suggests that the structure of TMD of nAChR is largely affected by the presence of LBD and the cytoplasmic helix. The isolated TMD has the intrinsic flexibility and rigidity, but the neighboring domains modulate these structural properties. The interaction between TMD and LBD shapes the correlations important for proper gating. Therefore, the structural properties of the TMD of nAChR are required but are not the only requirements for the gating mechanism.

### 3.4. ANM Results

ANM analyses have been performed to elucidate the motions described by cross correlation maps and slow mode fluctuations. ANM analysis produces coordinate files which include the amplitude and directions of each residue (vectors). These vectors are used to produce a 3D image of the protein at various modes and conformations. Each mode has two conformations one positive, obtained by adding the vectors to the original conformation, and one negative, obtained by subtracting vectors. Also animations of the motion from negative conformation to positive conformation through the original structure were generated and used to analyze the motions. The resulting 3D representation of the protein at each mode was analyzed with HOLE [45]. HOLE produced outputs containing the radius profile of the channel along the Z axis (the axis pore runs along). The output vectors does not describe the absolute amplitude of motion along each axis, these vectors were multiplied by an arbitrary constant for a better presentation of the movement described. Therefore the results of HOLE are only infor-

mative of the state of the pore after these motions. It is possible to see if pore widens or gets narrower from the results of HOLE, but the radius values are not absolute. The analyses were performed on the 3D structure of nAChR from *Torpedo marmorata* using the structure file retrieved from Protein Data Bank (PBD) (PDB code: 2bg9). This structure contains the ligand binding domain along with the transmembrane domain and a helix from the cytoplasmic domain. After computational analyses the results for TMD of nAChR were extracted. The analyses were not performed on the isolated TMD domain of nAChR, because of the data provided by GNM analyses. GNM results suggested that the presence of LBD domain and cytoplasmic section produce an effect which can not be excluded from the ANM analysis. Therefore, isolated structure was not analyzed using ANM.

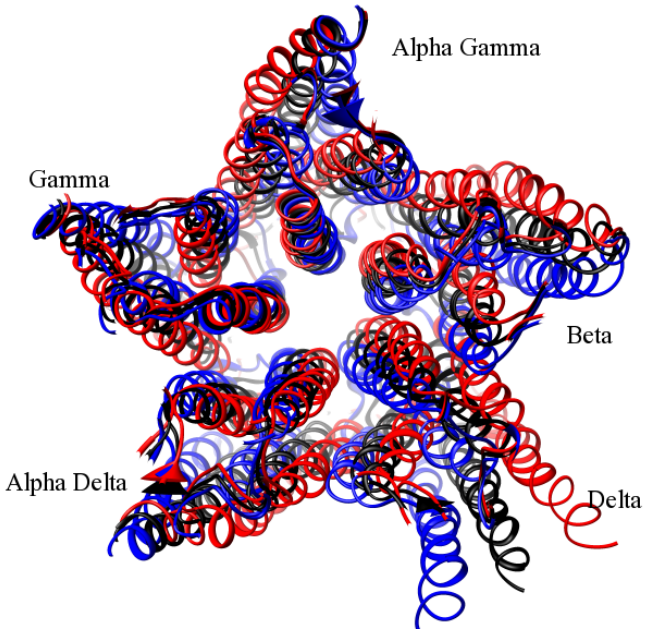


Figure 3.17. ANM results, first mode of nAChR (2bg9). Black ribbon represents the original structure. Negative and positive conformations are represented by blue and red ribbons, respectively.

The motion in the first mode of ANM analysis exhibits a dominating twisting motion. The lower half of the TMD twists to a greater extend than the upper half. The motion seems to start with the M2-M3 linkers motion and it's transferred to the lower end via M3 helices. M2 helices are stationary relative to the lower end. (See Figure 3.17). These motions produced a wider pore in the negative conformation and the positive conformation does not provide a wider pore (See Figure 3.18). The twisting

motion described in this mode was described by Cheng et. al. [18] and they suggest that twisting motion is important for the opening of gate.

The motions in the second mode of ANM analysis show a transitional movement with a small amount of rotation. The extent of motion is different in each subunit. Only the negative conformation leads to a wider channel, the positive conformation leads to a narrower pore (See Appendix E, Figure E.1).

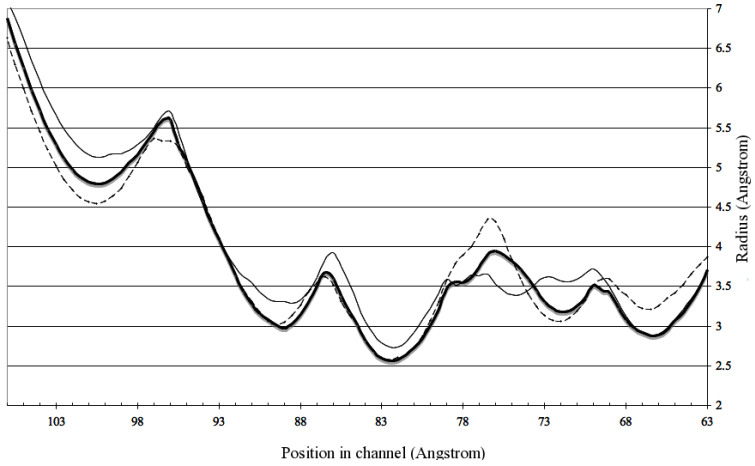


Figure 3.18. The pore profile calculated by HOLE in the first mode. Thick line represents the original structure. Negative and positive conformations are represented by thin and dashed lines, respectively.

The third mode of motions produces a transition, rotation motion at the gate level. The M1 helices and M2-M3 loop seems to bend the transmembrane helices in the middle. The M1-M2 loop is stationary relative to the gate section. The effects of these movements produce a wider pore for the negative conformation but a narrower pore for the positive conformation (See Figure E.2). The fourth mode does not produce a wider channel, therefore it is omitted.

The fifth mode produces a wider pore (See Figure E.3) for the negative conformation, but the positive conformation lead to a narrower pore. In fifth mode the structure seems to be fluctuating less than the other modes, but the all M2 helices, except the gamma M2, seem to be bending towards the M4 helices. The motion from the M2-M3 loop seems to drive the combined motions, leading to a wider gate. The lower halves of each subunit seem to move like a rigid body. The motion from M2-M3 seems to bend

the upper half and the bending diminishes near the gate region. The combined effect of each of these modes produces a wider pore in the fifth mode (See Figure 3.19).

The results from the ANM analysis up to fifth modes suggest that twisting motion is a dominating motion producing wider gate. The motions from M2-M3 loop and upper half of M1 helix should be underlined as the driving elements in the gating motions.

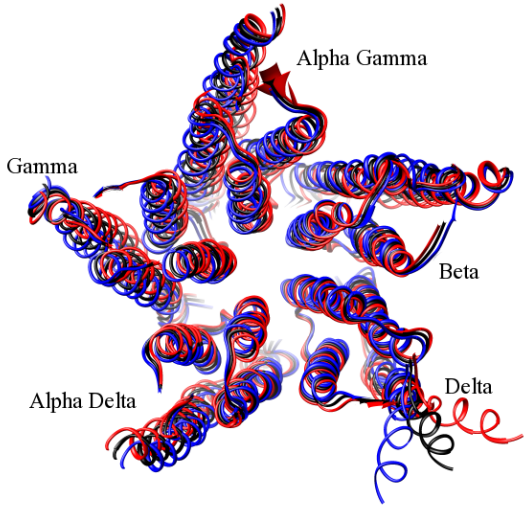


Figure 3.19. ANM results, fifth mode of nAChR (2bg9).

## 3.5. Mcpool

MCPOOL is program to analyze the cross correlation data produced by GNM analysis. The abundance of data in the cross correlation map makes it difficult to read and produce valuable information. MCPOOL analyzes the data via generating paths biased by cross correlation data, starting from a give set of points, and analyzes these paths (one million were generated) for connected paths. These connected paths are a subset within the random paths, which reach a set of target regions. This set of target regions were the gate regions defined by the Leu251 and Val255 [7, 10] rings and the residues falling in between these rings. The starting set of residues were selected are the binding region shaping residues of loop A (loop five), loop B (loop eight), loop C (loop 10). This analysis was performed on the cross correlation map (first ten slowest modes) of recently published structure of nAChR (PDB code: 2bg9)

[10]. This structure contains the ligand binding domain, transmembrane domain and a cytoplasmic domain consisting of five helices (named MA). After MCPOOL analysis results for TMD was extracted and presented with the findings of correlated mutation analysis and conservation analysis. The reason for using the cross correlation data of 2bg9 is that GNM results depicted TMD domain differentiates to a great extent in the presence of LMB and cytoplasmic domain (see section 3.3).

The lengths of paths were set at ten, twenty and thirty steps, but the results of only paths with ten steps were elaborated. The results of twenty and thirty steps were largely similar to the results of ten steps. The length of paths does not impose that all paths from binding region to gate region to be exactly ten steps. It only imposes a maximum length.

The first result of the MCPOOL analysis is the occurrence frequency of residues on connected paths. The importance of the occurrence frequency of residues is that residues with high occurrence frequencies form the key points in the structure for communication between two preselected regions. They may be interpreted as passes for efficient transmission of signal between any given point; i.e. the hub regions.

The second result of MCPOOL is the network formed by the connected paths. This network has a defined starting region and a defined ending region. The network was analyzed and visualized with a software for network analysis, PAJEK [67.9]. This analysis provides additional benefits to MCPOOL, which are visualization and realization of key points more effectively. The only drawback of network analysis is that representation of each residue is inconvenient. Therefore the residues of nAChR are grouped as the structural part they are on (residues on Cys loop is grouped into a group called Cys loop). This grouping results in 41 groups per subunit and 205 groups for TMD. Sometimes this grouping causes key residues to be left out, but it provides more information than it interferes with. Also, the two results of MCPOOL are complementary and should be analyzed together.

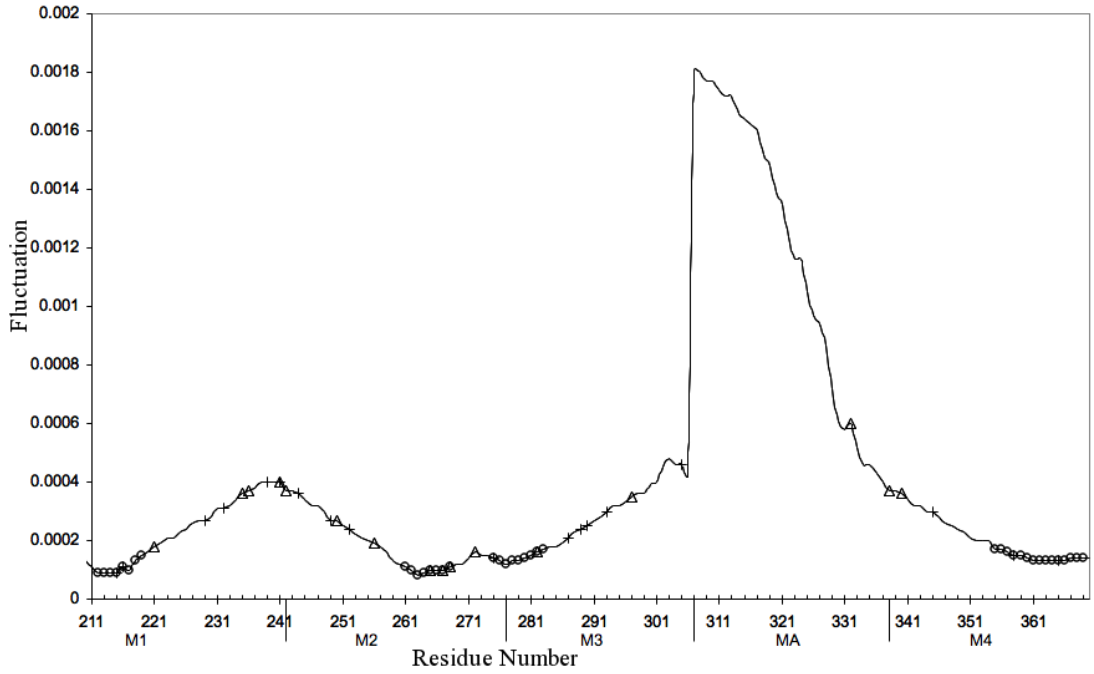


Figure 3.20. Most used residues of alpha gamma subunit plotted on slow mode fluctuations.

### 3.5.1. Occurrence Frequencies

Most occurring residues figures show forty residues with highest occurrence frequency (circles) on the paths from binding region to the gate region on alpha gamma subunit in TMD, plotted on the average fluctuations of first two modes (line), along with conserved residues (triangles), and correlated mutation sites (crosses). The most occurring residues cluster at four locations (See Figure 3.20). These are the initial residues of M1, lower half of M2, M2-M3 loop and initial residues of M3 and the upper half of M4 helices. It is also notable that most frequently occurring residues reside at hinges. The frequent occurrence of the initial residues of M1 helix support the importance of the covalent bond between Beta10 strand in LBD and M1 helix of alpha subunit [16, 31]. The important M2-M3 loop is also present, which is repeatedly mentioned in articles [1, 16, 31, 35, 55]. Residues on M4 helices are also important for regulating the gating, and mutation analyses underline the importance of M4 helix [15, 29, 55], but residues on M4 haven't been identified as residues important for the coupling of LBD and TMD. One of the frequently occurring residues Phe426 (359 on Figure 3.20, from now on the number between parentheses will show the number of the

residue on the corresponding figure if they differ) is found to be affecting the gating when mutated [29].

In addition, some of these frequently occurring residues are conserved residues and some of them are overlap with correlated mutations. The conserved residues which are frequently occurring are alpha gamma Pro265, Thr267 and Ser268 on the M2 helix, Met282 on M3 helix. The frequently occurring residues which are conserved through correlated mutations are alpha gamma Val215, Val216 on M1 helix, Gly275 on M2-M3 loop, and Val425 (358 on Figure 3.20) and Glu432 (365 on Figure 3.20).

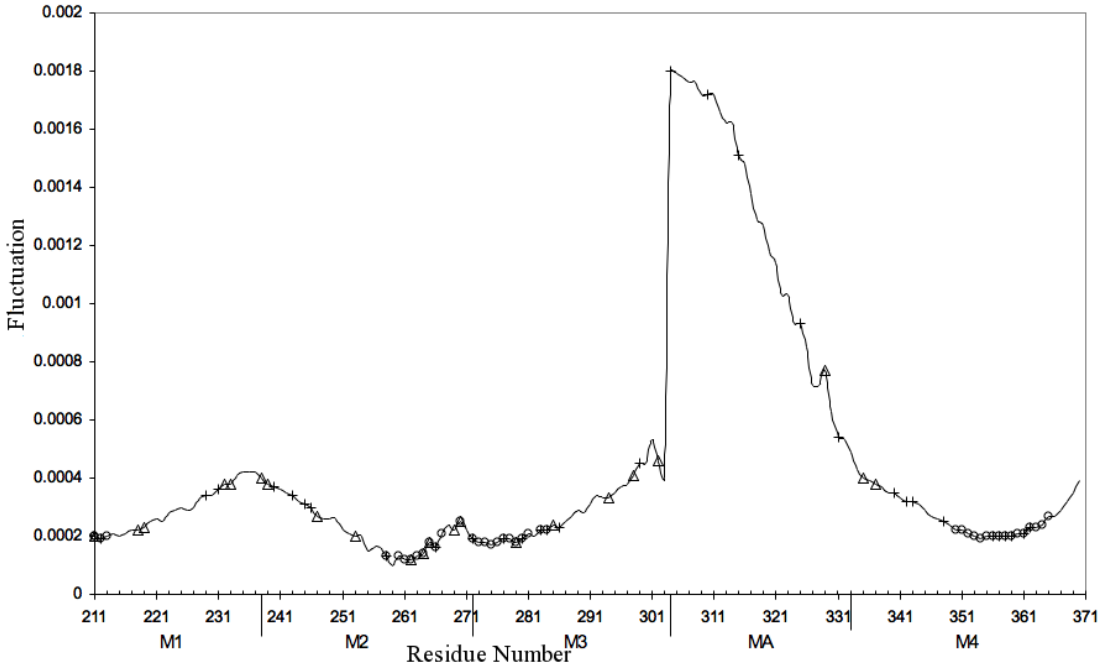


Figure 3.21. Most used residues of alpha gamma subunit plotted on slow mode fluctuations.

The results of MCPOOL analysis for beta subunit (See Figure 3.21) resemble the result of alpha subunit. The most noticeable difference between the results is the diminished cluster of frequently occurring residues on the M1 helix. Only two of such residues exits, Tyr220 (211 on Figure 3.21), which is also conserved, and Ile221 (212 on Figure 3.21), which is conserved by correlated mutations. The remaining results are similar. The frequently occurring residues cluster in the upper half of M2 and M3 and on the M2-M3 loop and the upper half of M4.

The results of delta subunit are similar to beta subunit with the difference that the cluster of most frequently occurring residues on the upper half of M4 is decreased (See Appendix F, figure F.1). MCPOOL analysis produced similar results for alpha delta and gamma subunits (See Appendix F, figures F.2, F.3).

### 3.5.2. Network of Pathways

The network of connected paths were analyzed for key regions of communication with the network analysis program PAJEK. The network output of MCPOOL was first normalized with respect to the maximum occurrence frequency of a step. Than the steps with normalized occurrence frequency less than 0.1 (0.15 for gamma subunit and 0.25 for delta subunit) was removed. This number is selected in order to present the 10 per cent of number of groups with the maximum occurrence frequencies, which led to around 20 groups. For representation purposes loop C was selected as the starting point of pathway, due to the recent findings about loop C [31, 55], but groups on the LBD were not interpreted in the present thesis.

The network representation of the pathway between alpha gamma loop C (loop 10) and alpha gamma gate group (residues Leu251 to Val255) uses the upper half of M4 subunit (See Figure 3.22) at the most. M2-M3 is on the pathway but it is not as underlined as the Upper half of M4 or the upper half (symbolized by "Ust") of M1 helix. Interestingly the pathway to the gate region passes from the neighboring gamma subunits M1 helix and the lower half (symbolized by "Alt") of M4 helix.

The network representation of the alpha delta, and gamma subunits are similar to the network representation of alpha gamma subunit with one difference that these subunits (alpha gamma, gamma) do not use the neighboring subunits in the communication (See Appendix F, Figure F.4, F.6, and F.7). The upper half of M4 helices of each subunit is a key passageway for all subunits.

The results for Delta subunit is similar to the other results, as the upper part of M4 helix is underlined, but delta subunit uses the TMD of beta subunit extensively

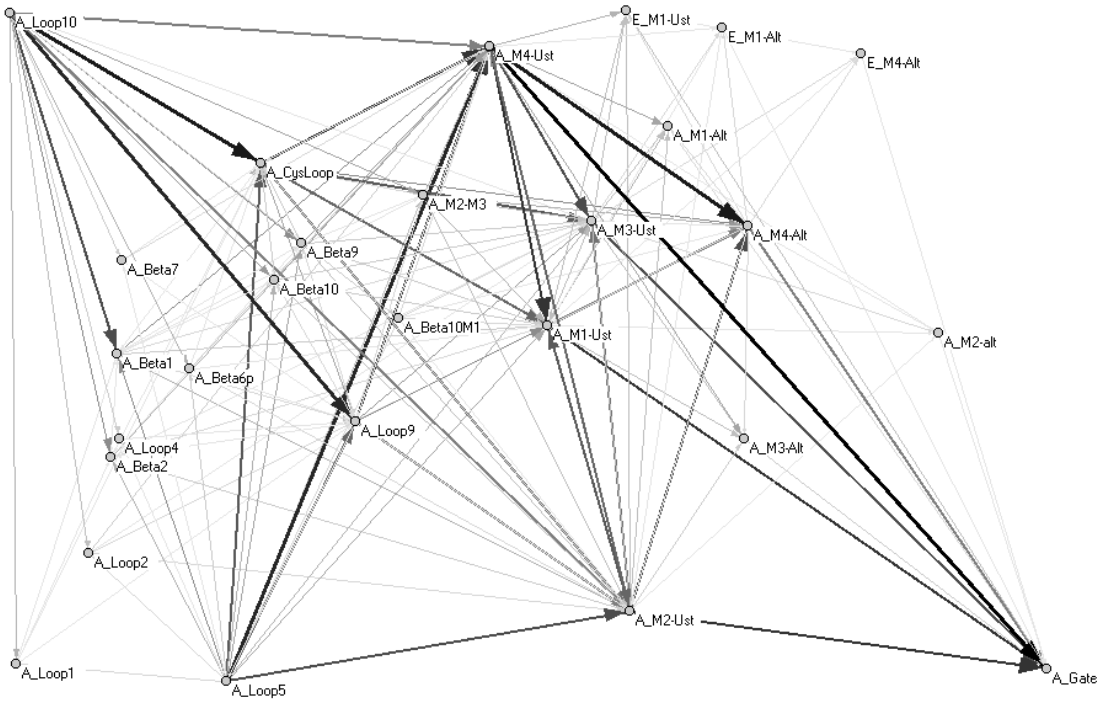


Figure 3.22. Network representation of the pathway network from alpha loop C (represented by A_Loop10) to the gate region of alpha subunit (A_Gate).

(See Figure F.5). The regions of beta subunit used by the paths between delta loop C and delta gate region includes upper halves of beta M1, M2, M3 and M4. This may be the result of extensive high correlations between beta and delta subunit present in the cross correlation map and the fact that M4 helix of delta subunit makes the smallest number of contacts with M1 and M3 helices of delta subunit.

The results of MCPOOL analysis suggest that upper half of M4 helices are key passageways on the pathway to the gate region from the binding region. In addition, the results suggest that the covalent bond between M1 helices and LBD, and M2-M3 Loop is important for signal transmission. In the present study there were no noticeable difference between alpha subunits.

# 4. CONCLUSIONS

Nicotinic acetylcholine receptors (nAChR) are a family of ligand gated ion channels (LGIC) super-family. They mediate the signals at the neuromuscular junctions and in the central nervous system. Aberrations in nAChR family is the cause behind many diseases such as Congenital Myasthenic Syndrome (CMS) (weakness of muscles) and Alzheimer's disease (a disease deteriorating a persons memory and ability to learn). Nicotinic acetylcholine receptors have been studied extensively since the late sixties and the most studied nAChR is from a ray fish, *Torpedo marmorata*, which utilizes electric shocks to stun its prey. In the present study the structure of of nAChR is studied.

GNM analysis was performed on the near complete 3D structure of nAChR and the result for TMD was extracted. The correlation between the fluctuations of residues by GNM analysis suggests that the tapering structure of TMD, a more tightly packed lower half and a loosely packed upper half, is important for the structural requirements of a twist to open type of movement suggested by Unwin et. al. [7]. The tapering structure of TMD leads to a wider upper portion and a narrower lower portion in TMD, which in turn leads to an increase in number contacts made by a residue from the extracellular side (upper) to cytoplasmic side (lower). The wider region of M2-M3 loop creates spaces for the interface regions of LBD, therefore a wider region of contact. The symmetric positions of residues [7] and increased number of contacts at the gate region makes an expansion difficult due to side to side interactions. From a logical stand point it requires more energy to expand the gate region without first weakening the side to side interactions of M2 helices. The abundance of highly correlated residues of M2 helices suggest that, these residues on M2 helices tend to move in a single direction (because the results of GNM are isotropic). Only the decreased correlations between non neighboring gate regions on M2 helices hint the rotational movement. This rotational movement could weaken the side to side interactions [7] and allow the expansion of gate.

The analysis of fluctuations in the most cooperative modes of TMD of the nAChR, with LBD and cytoplasmic domain, show that the fluctuations of M2-M3 loop is restricted relative to the rest of the regions. The flexible joints in a structures is more efficient in transmitting the structural information than a mobile segment, therefore the M2-M3 loop behaves as a hinge, with low amplitude motions, mediating the structural information between two domains. Also, an efficient way of transmitting signals from binding region would be more successful, in the way that it will allow sharper reflexes (faster signaling), and in the greater sense of interaction between living creatures and evolutionary prey and hunter relation. The initial residues of M1 helices are also displays restricted fluctuations. The fluctuations of M1 helix changes into unrestricted fluctuations from the upper half of TMD to the lower half. This mobility change should be attributed to the cytoplasmic helices which have much less contacts than the rest of the nAChR described in the 3D structure. This could be interpreted as a drawback for the signal transmitting capabilities of M1 but the cross correlations with the M2-M3 loop and the upper part of M2 may allow the signal to be directly transmitted to M2 helices. The slow mode fluctuations of M4 helices and the cross correlation of M4 helices with other regions in the TMD domain suggests that M4 should be involved in the gating mechanism as much as the M2-M3 loop. Even though mutagenesis studies of M4 suggests that M4 affects the gating, it is not discussed as on the pathway from binding site to the gate.

The analyses with the isolated structure of TMD of nAChR shows that the fluctuations in the most cooperative modes of TMD are affected to a great extent by the LBD. The TMD domain, in the isolated analyses, shows the intrinsic flexibilities required for the gating mechanism, but the interaction between LBD and TMD shapes the correlations. Therefore, individual interactions of LBD and TMD are essential for gating mechanism.

Conservation and correlated mutation analysis suggests that the residues shaping the M2-M3 helix and the gate region are important for structure and function of nAChR. It's noticeable that even most of the residues on M4 helices are not conserved, correlated mutations at these sites conserve the important residues.

The method MCPOOL, designed in this thesis, has shown to be an efficient and informative tool in the analysis of the correlated fluctuations. By MCPOOL, it has been possible to acquire a network of connected paths from a starting region to a preselected target, and identified the most visited residues on these paths. These most visited residues highly associate with the results of conservation and correlated mutation analyses. The results of MCPOOL suggests that the interaction between LBD and TMD is not dependent on any single region, but the interaction is transmitted by a set of regions. These regions include the M2-M3 loop, initial residues of M1 (which are covalently bonded to B10 at the LBD), and the upper half of M4. They contribute to the gating mechanism.

The gating mechanism involves the cooperative interactions of LBD interacting regions of TMD helices. The signal from the binding region is transmitted through M1, M2-M3 loop and M4 to the rest of the channel to allow for a cooperative twisting motion and a bending at the gate region which in turn leads to a wider pore.

# APPENDIX A: SEQUENCE OF nAChR

>2BG9:A - Alpha Gamma
SEHETRLVANLLENYNKVIRPVEHHTHFVDITVGLQLIQLINVDEVNQIVETNVRLRQQWIDVRLRWNPADYGGIKKIRL
PSDDVWLPDLVLYNNADGDFAIVHMTKLLLDYTGKIMWTPPAIFKSYCEIIVTHFPFDQQNCTMKLGIWTYDGTKVSISP
ESDRPDLSTFMESGEWVMKDYRGWKHWVYYTCCPDTPYLDITYHFIMQRIPLYFVVNVIIPCLLFSFLTVLVFYLPTDSG
EKMTLSISVLLSLTVFLLVIVELIPSTSSAVPLIGKYMLFTMIFVISSIIVTVVVINTHHRSPSTHSAIEGVKYIAEHMK
SDEESSNAAEEWKYVAMVIDHILLCVFMLICIIGTVSVFAGRLIELSQEG
>2BG9:B - Beta
SVMEDTLLSVLFENYNPKVRPSQTVGDKVTVRVGLTLTSLLILNEKNEEMTTSVFLNLAWTDYRLQWDPAAYEGIKDLSI
PSDDVWQPDIVLMNNNDGSFEITLHVNVLVQHTGAVSWHPSAIYRSSCTIKVMYFPFDWQNCTMVFKSYTYDTSEVILQH
ALDAMINQDAFTENGQWSIEHKPSRKNWRSDDPSYEDVTFYLIIQRKPLFYIVYTIVPCILISILAILVFYLPPDAGEKM
SLSISALLALTVFLLLLADKVPETSLSVPIIISYLMFIMILVAFSVILSVVVLNLHHRSPNTHEAVEAIKYIAEQLESAS
EFDDLKKDWQYVAMVADRLFLYIFITMCSIGTFSIFLDASHNVPPDNPFA
>2BG9:C - Delta
VNEEERLINDLLIVNKYNKHVRPVKHNNEVVNIALSLTLSNLISLKETDETLTTNVWMDHAWYDHRLTWNASEYSDISIL
RLRPELIWIPDIVLQNNNDGQYNVAYFCNVLVRPNGYVTWLPPAIFRSSCPINVLYFPFDWQNCSLKFTALNYNANEISM
DLIIDPEAFTENGEWEIIHKPAKKNIYGDKFPNGTNYQDVTFYLIIRRKPLFYVINFITPCVLISFLAALAFYLPAESGE
KMSTAICVLLAQAVFLLLTSQRLPETALAVPLIGKYLMFIMSLVTGVVVNCGIVLNFHFRTPSTHSGIDSTNYIVKQIKE
KNAYDEEVGNWNLVGQTIDRLSMFIITPVMVLGTIFIFVMGNFNRPPAK
>2BG9:D - Alpha Delta
SEHETRLVANLLENYNKVIRPVEHHTHFVDITVGLQLIQLINVDEVNQIVETNVRLRQQWIDVRLRWNPADYGGIKKIRL
PSDDVWLPDLVLYNNADGDFAIVHMTKLLLDYTGKIMWTPPAIFKSYCEIIVTHFPFDQQNCTMKLGIWTYDGTKVSISP
ESDRPDLSTFMESGEWVMKDYRGWKHWVYYTCCPDTPYLDITYHFIMQRIPLYFVVNVIIPCLLFSFLTVLVFYLPTDSG
EKMTLSISVLLSLTVFLLVIVELIPSTSSAVPLIGKYMLFTMIFVISSIIVTVVVINTHHRSPSTHSAIEGVKYIAEHMK
SDEESSNAAEEWKYVAMVIDHILLCVFMLICIIGTVSVFAGRLIELSQEG
>2BG9:E - Gamma
NEEGRLIEKLLGDYDKRIKPAKTLDHVIDVTLKLTLTNLISLNEKEEALTTNVWIEIQWNDYRLSWNTSEYEGIDLVRIP
SELLWLPDVVLENNVDGQFEVAYYANVLVYNDGSMYWLPPAIYRSTCPIAVTYFPFDWQNCSLVFRSQTYNAHEVNLQLS
AEEGIDPEDFTENGEWTIRHRPAKKNYNWQLTKDDIDFQEIIFFLIIQRKPLFYIINIIAPCVLISSLVVLVYFLPAQAG
GQKCTLSISVLLAQTIFLFLIAQKVPETSLNVPLIGKYLIFVMFVSLVIVTNCVIVLNVSLRTPNTHSCVEACNFIAKST
KEQNDSGSENENWVLIGKVIDKACFWIALLLFSLGTLAIFLTGHLNQVPE

Figure A.1. Sequence of nAcetylcholine receptor from *Torpedo marmorata*

# APPENDIX B: CONSERVATION AND CORRELATED MUTATION FIGURES

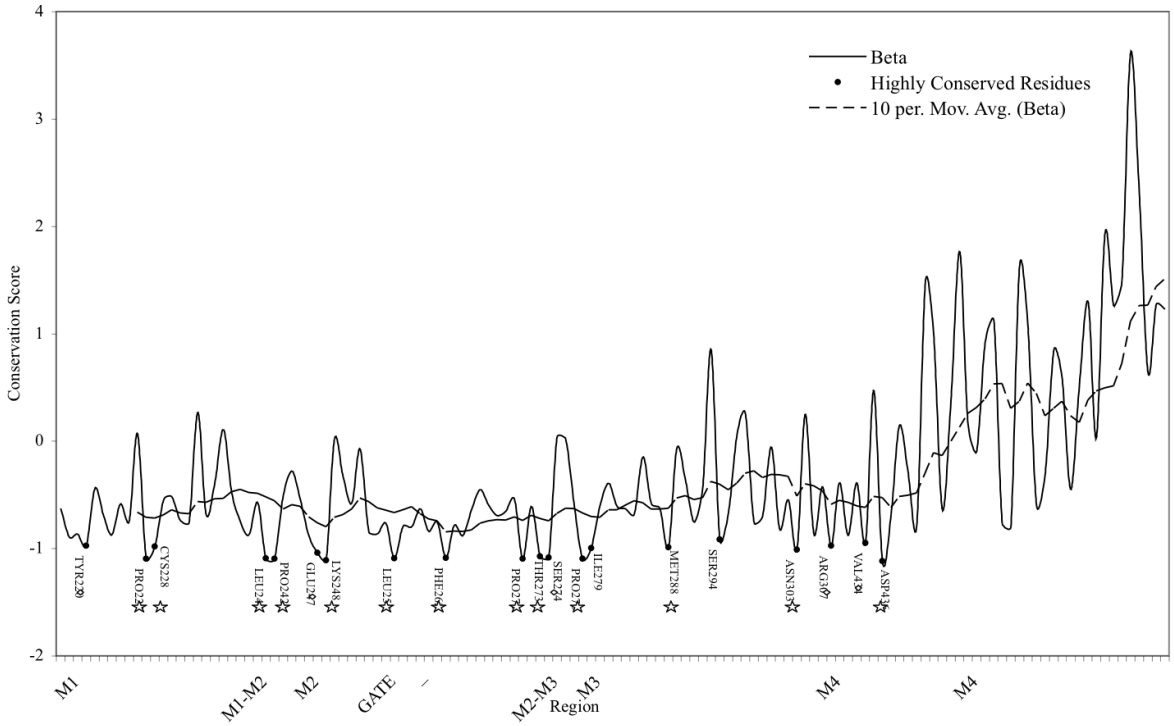


Figure B.1. Conservation scores (normalized) of beta TMD.

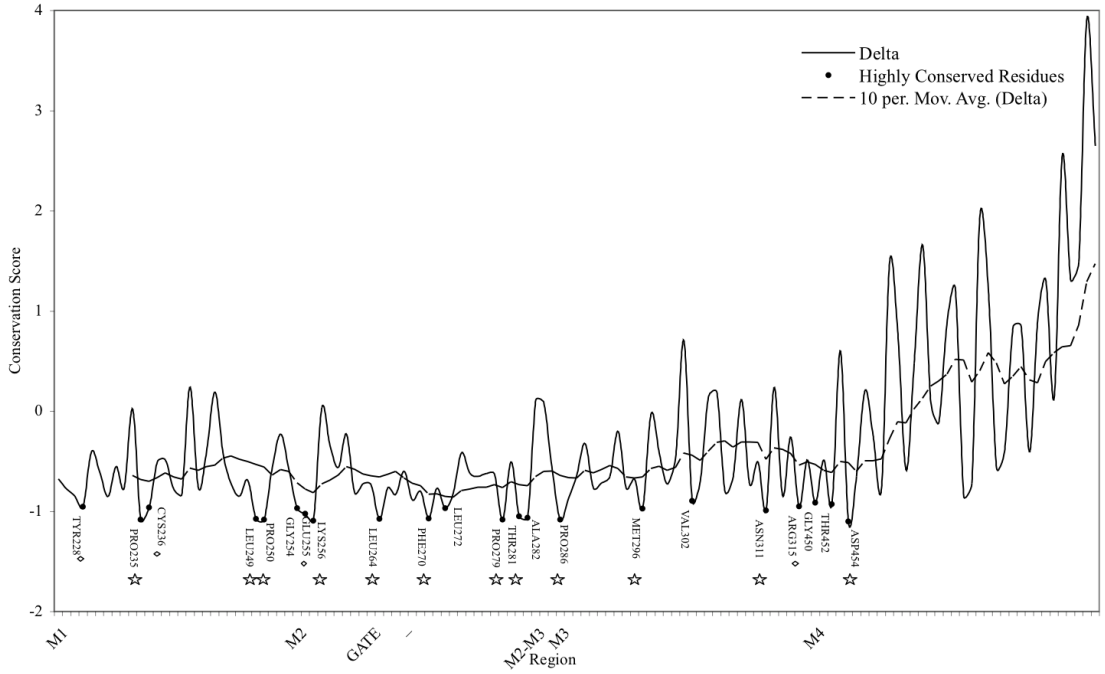


Figure B.2. Conservation scores (normalized) of delta TMD.

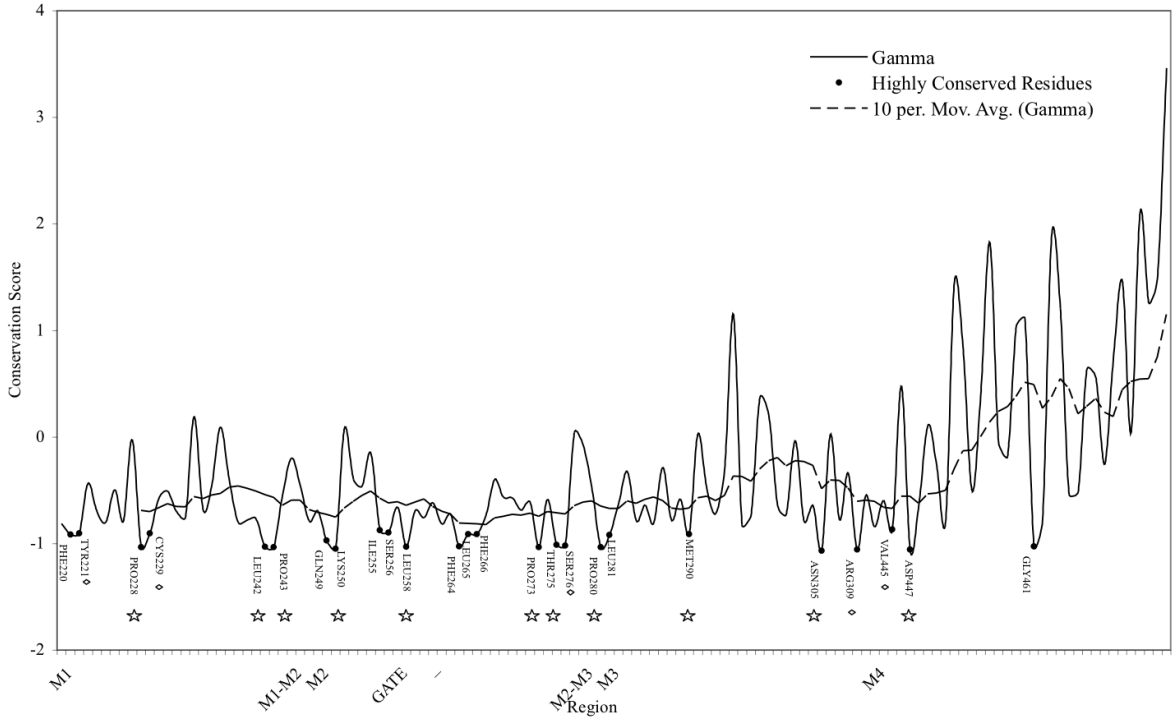


Figure B.3. Conservation scores (normalized) of gamma TMD.

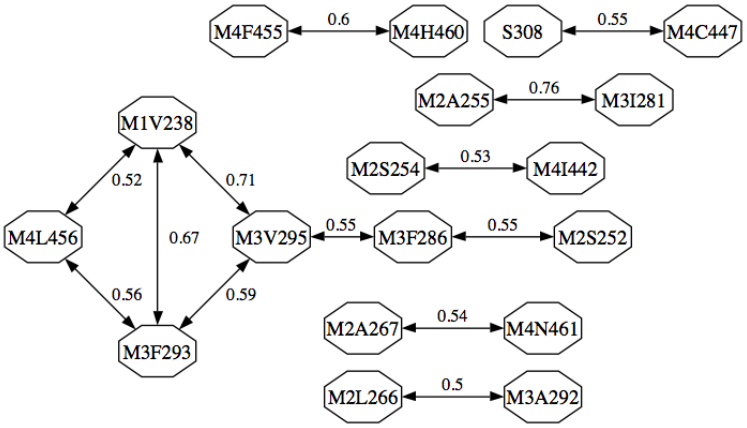


Figure B.4. The Network of correlated mutations for beta subunit of nAChR.

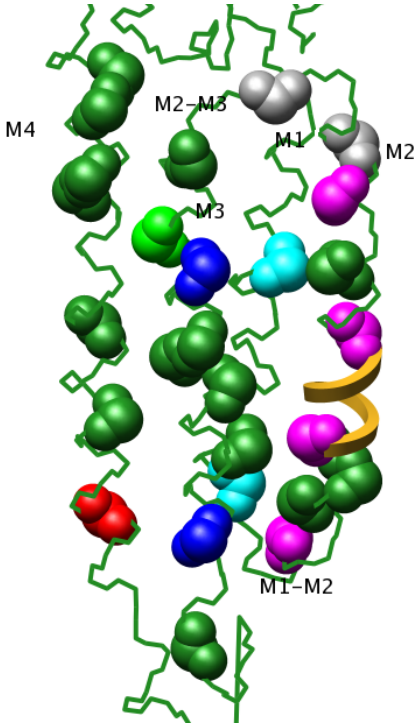


Figure B.5. Correlated mutations (dark green) and conserved residues (not green) on the beta subunit of nAChR.

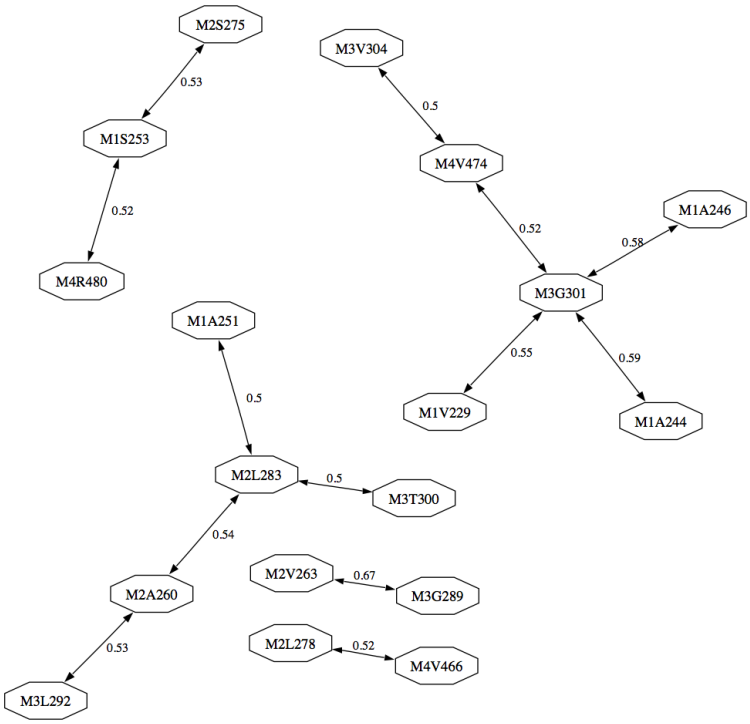


Figure B.6. The Network of correlated mutations for delta subunit of nAChR.

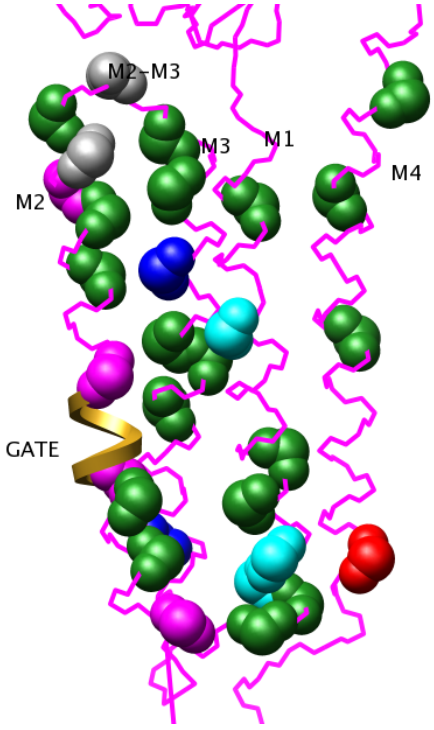


Figure B.7. Correlated mutations (dark green) and conserved residues (not green) on the delta subunit of nAChR.

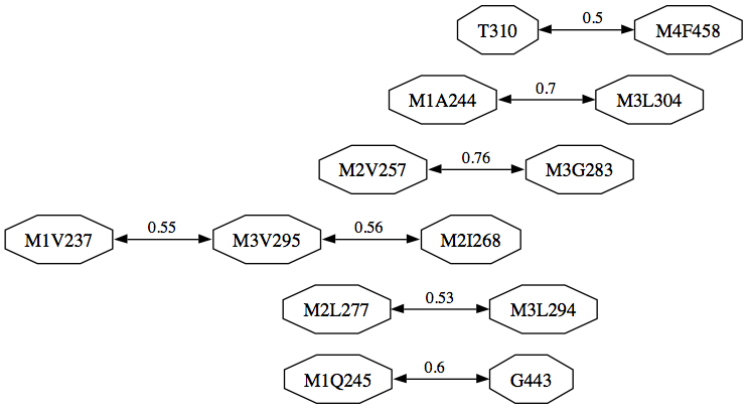


Figure B.8. The Network of correlated mutations for gamma subunit of nAChR.

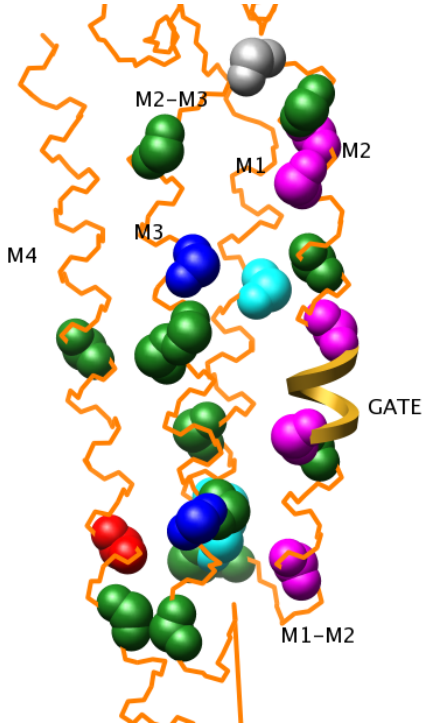


Figure B.9. Correlated mutations (dark green) and conserved residues (not green) on the gamma subunit of nAChR.

# APPENDIX C: MULTIPLE SEQUENCE ALIGNMENT OF nAChR SUBUNITS

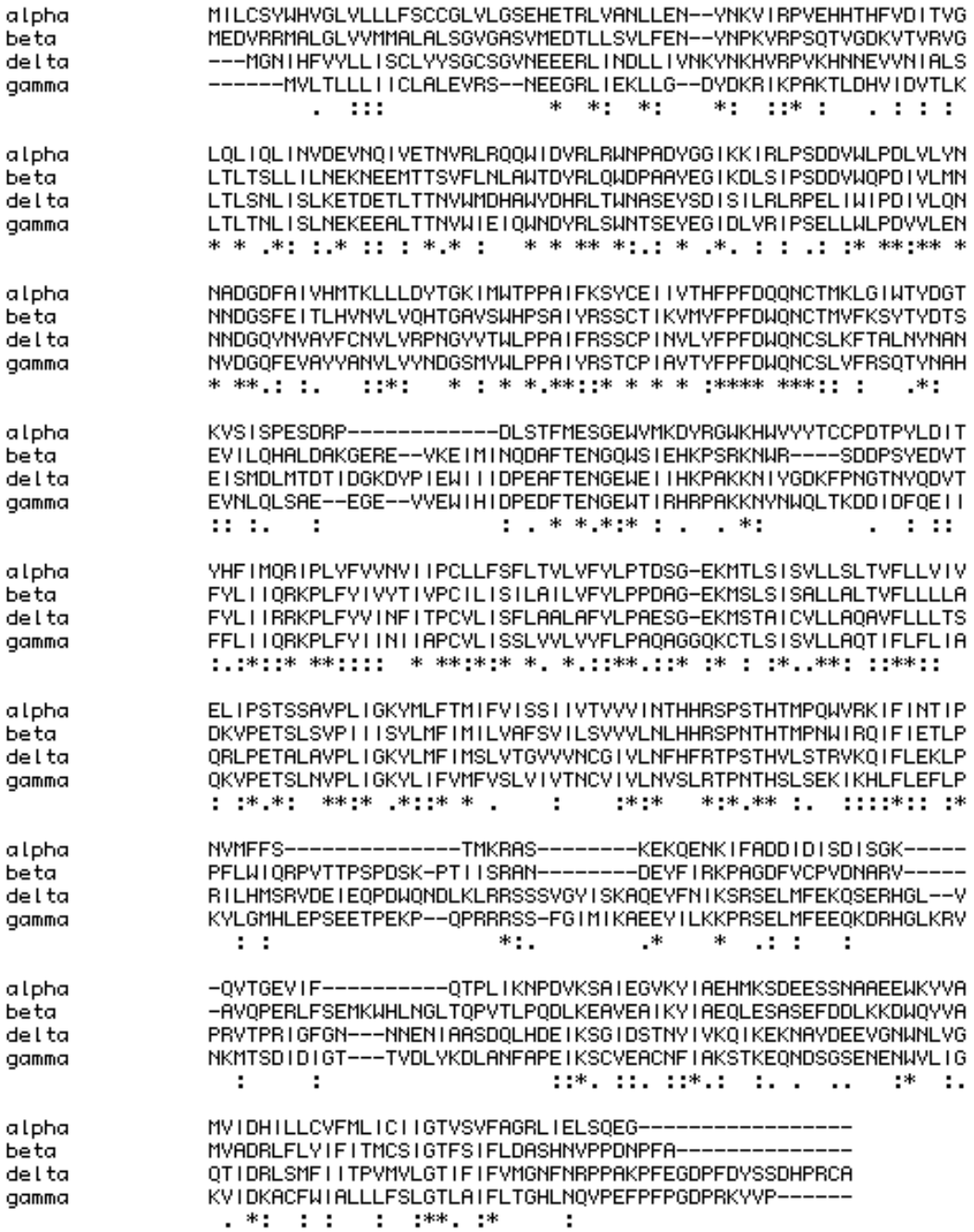


Figure C.1. Multiple sequence alignment of nAChR subunits

# APPENDIX D: GNM FIGURES

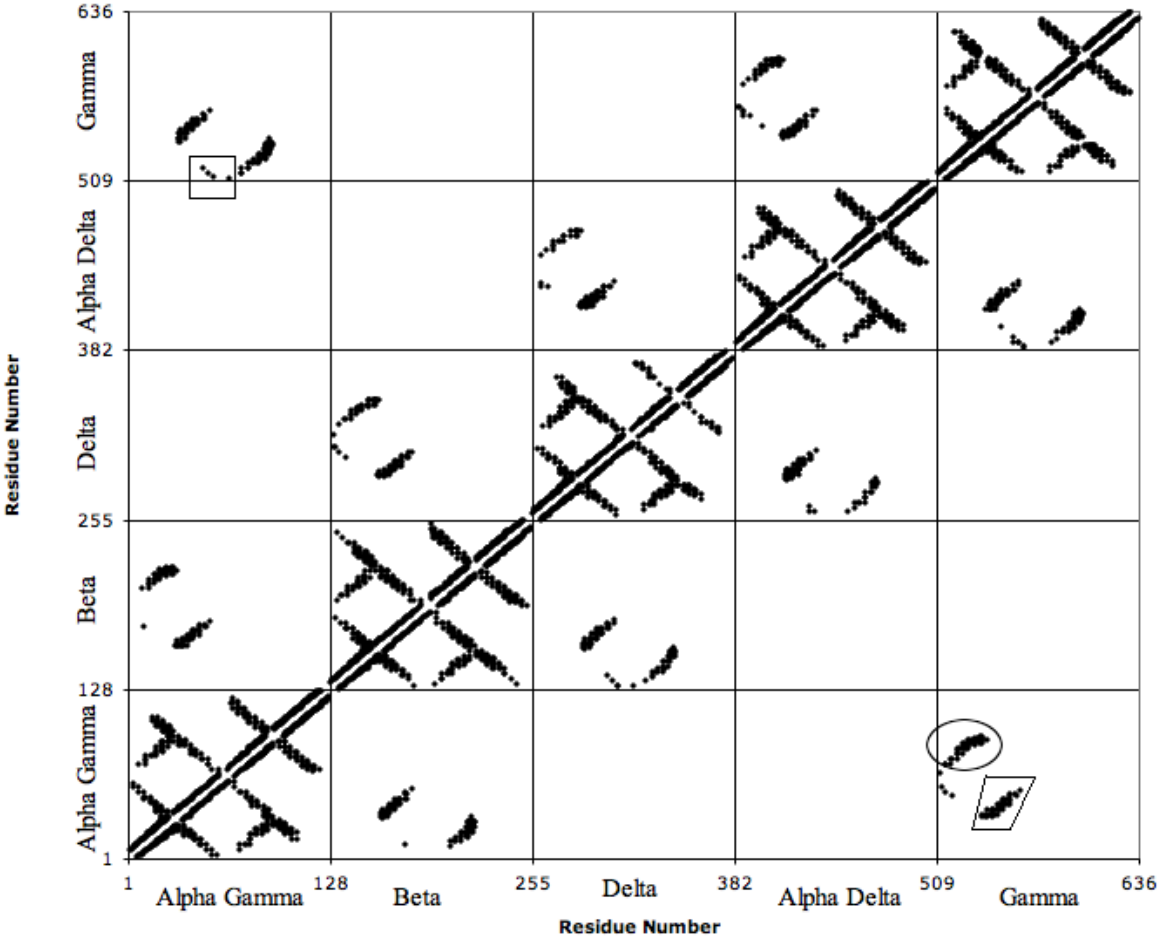


Figure D.1. Contact Map of 1oed with maximum backbone separation of 10$\mathring{A}$.

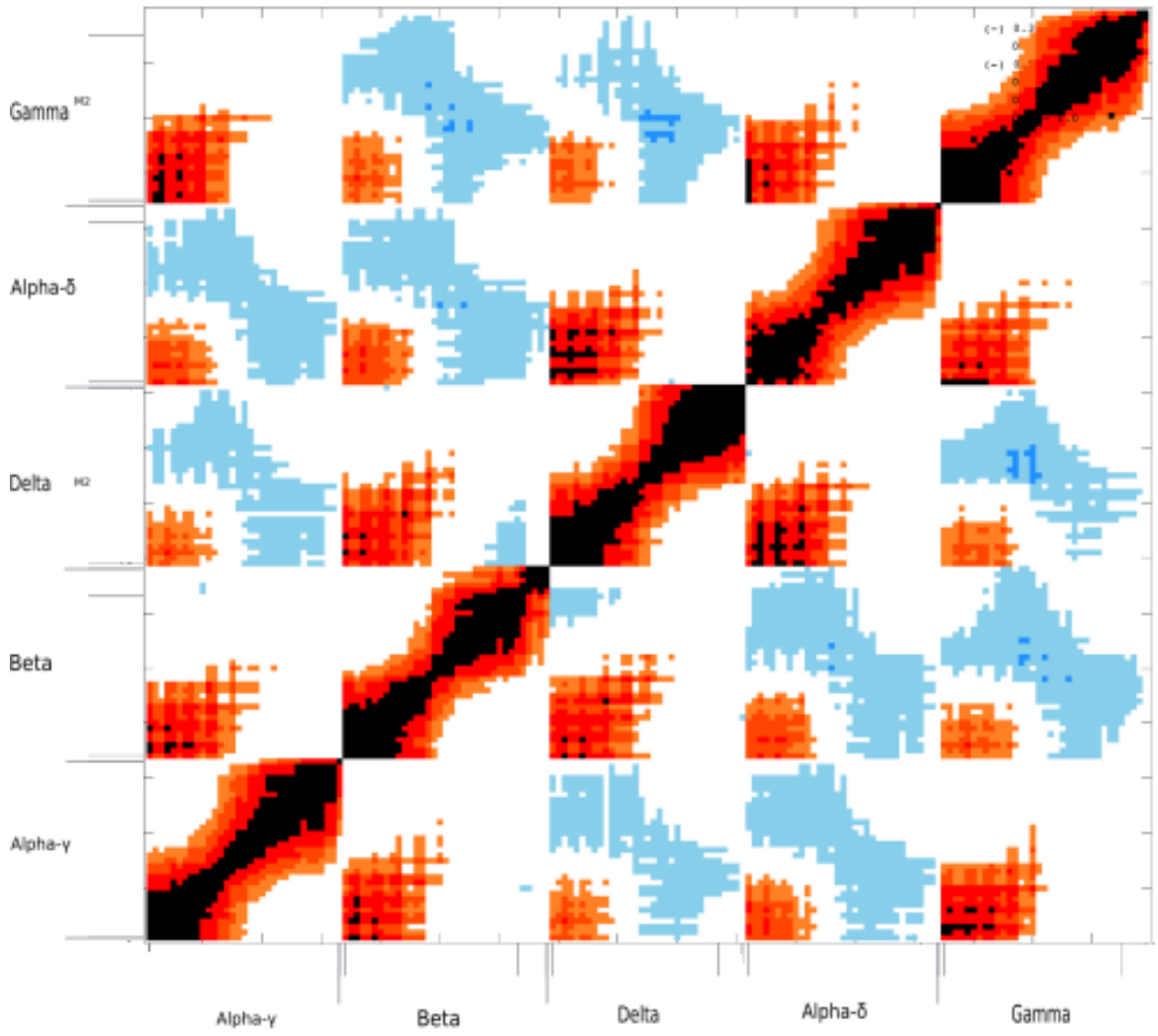


Figure D.2. Cross correlation map for slowest ten modes of M2 helices.

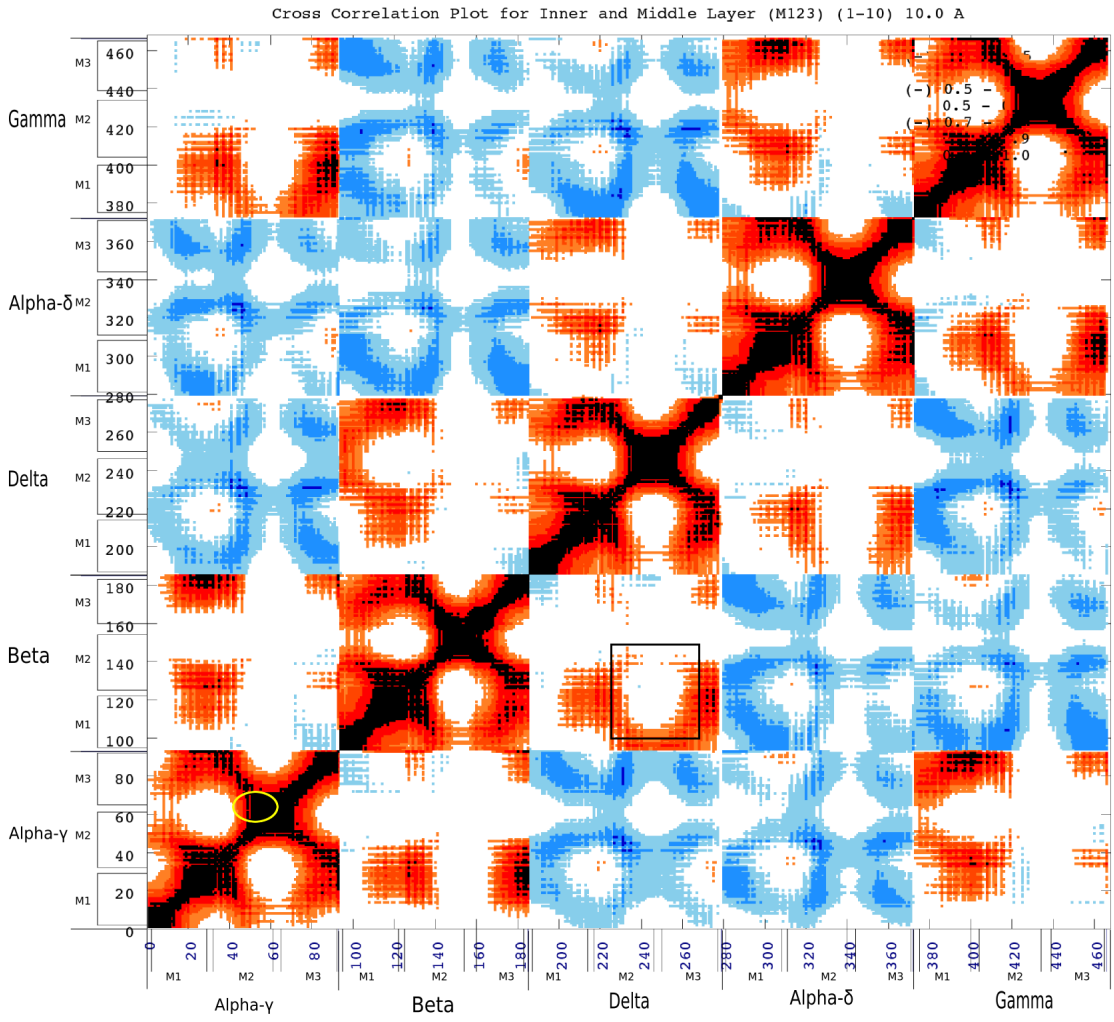


Figure D.3. The cross correlation map for the slowest ten modes of M1, M2, and M3 helices (inner and middle ring) of TMD of nAChR.

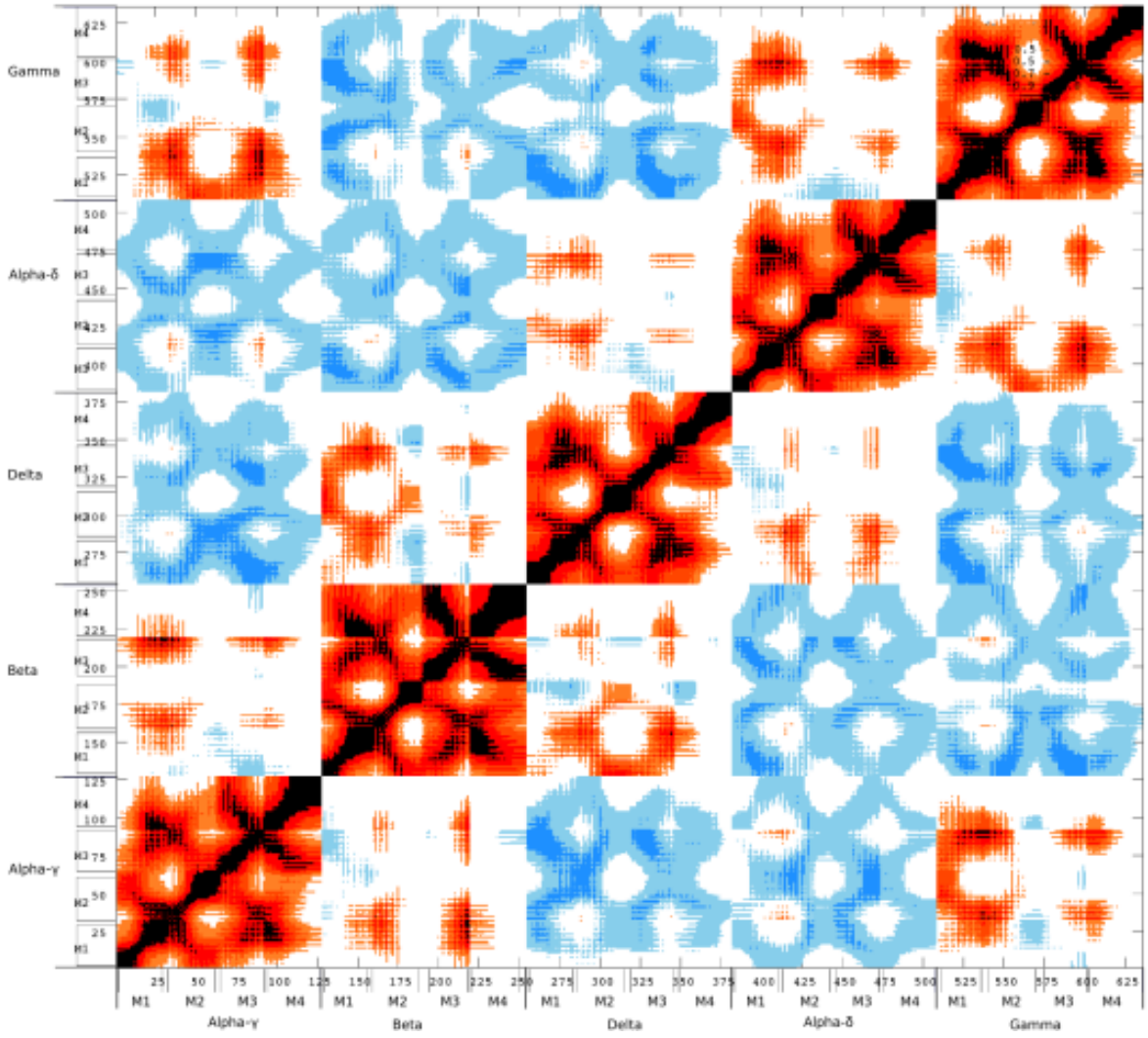


Figure D.4. The cross correlation map for the TMD of nAChR (isolated), the results of the slowest ten modes.

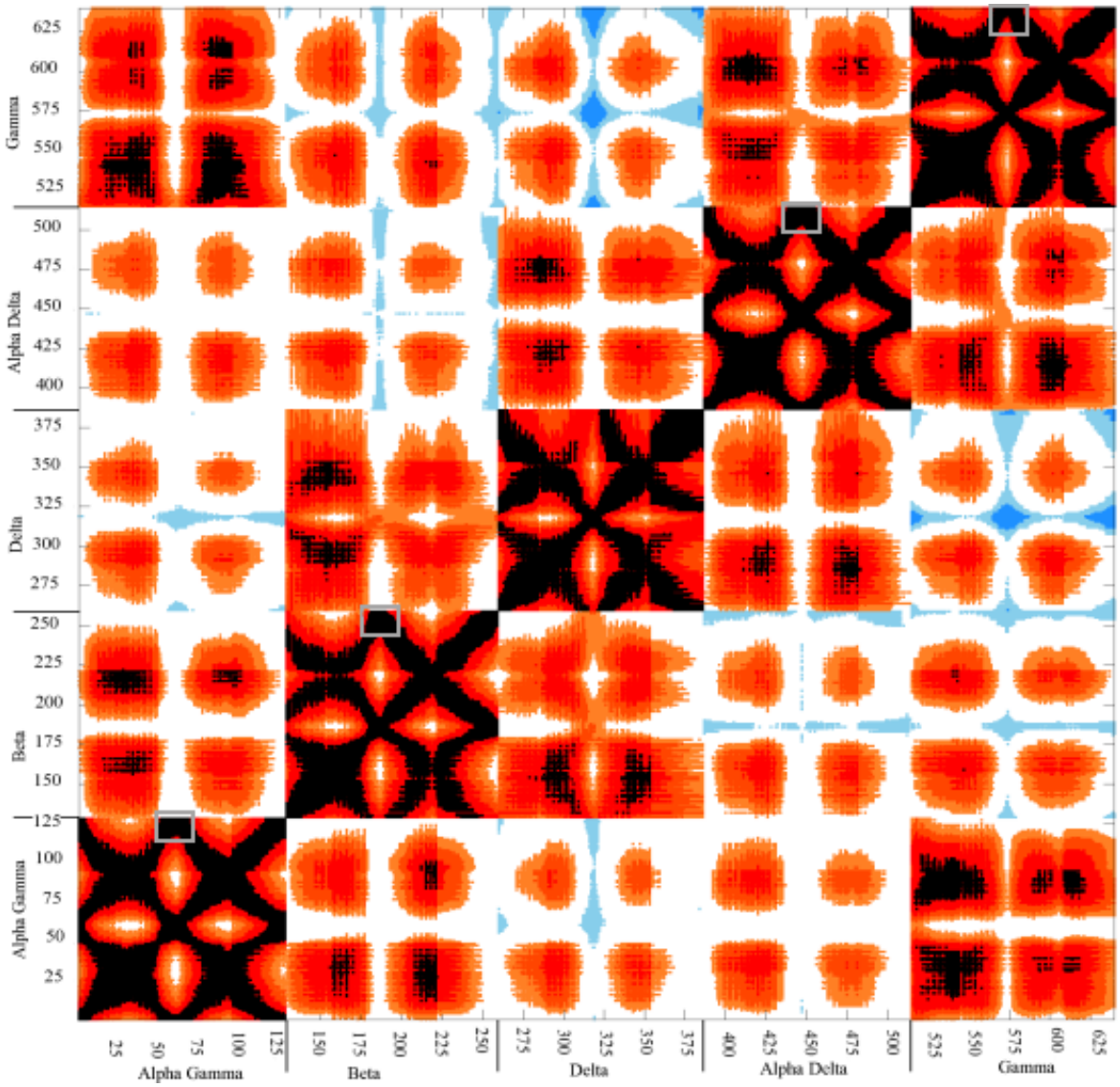


Figure D.5. Cross correlation map of the TMD region of nAChR extracted from the analysis of 2bg9.

# APPENDIX E: PORE RADIUS PROFILES

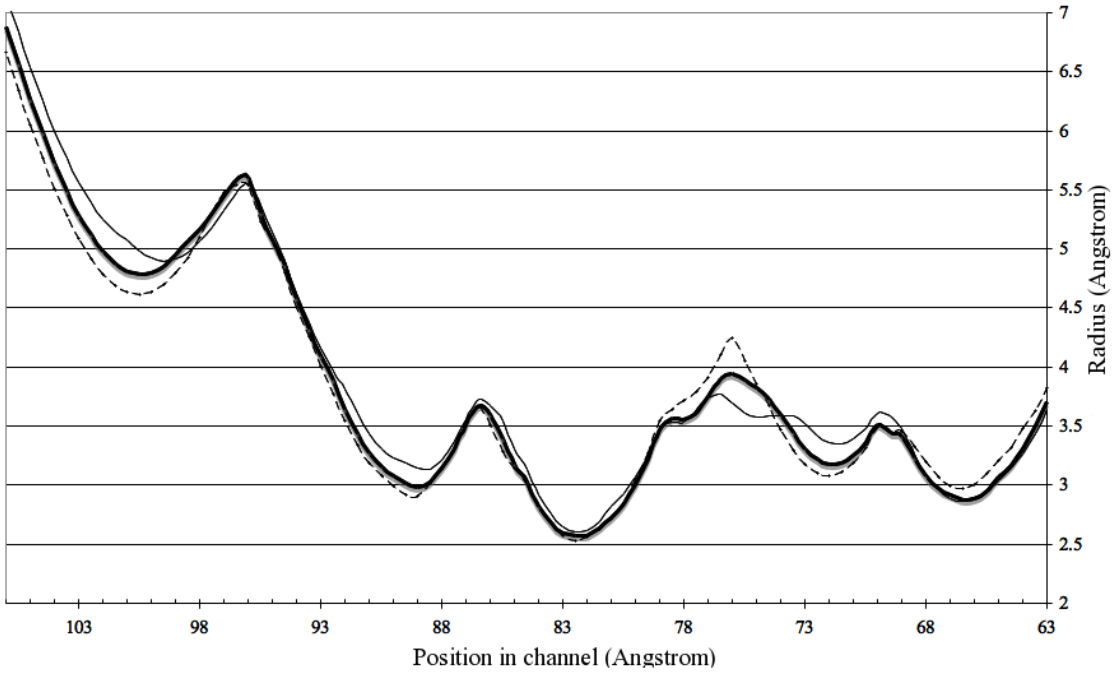


Figure E.1. The pore profile calculated by HOLE in the second slowest mode. Thick line represents the original structure. Negative and positive conformations are represented by thin and dashed lines, respectively.

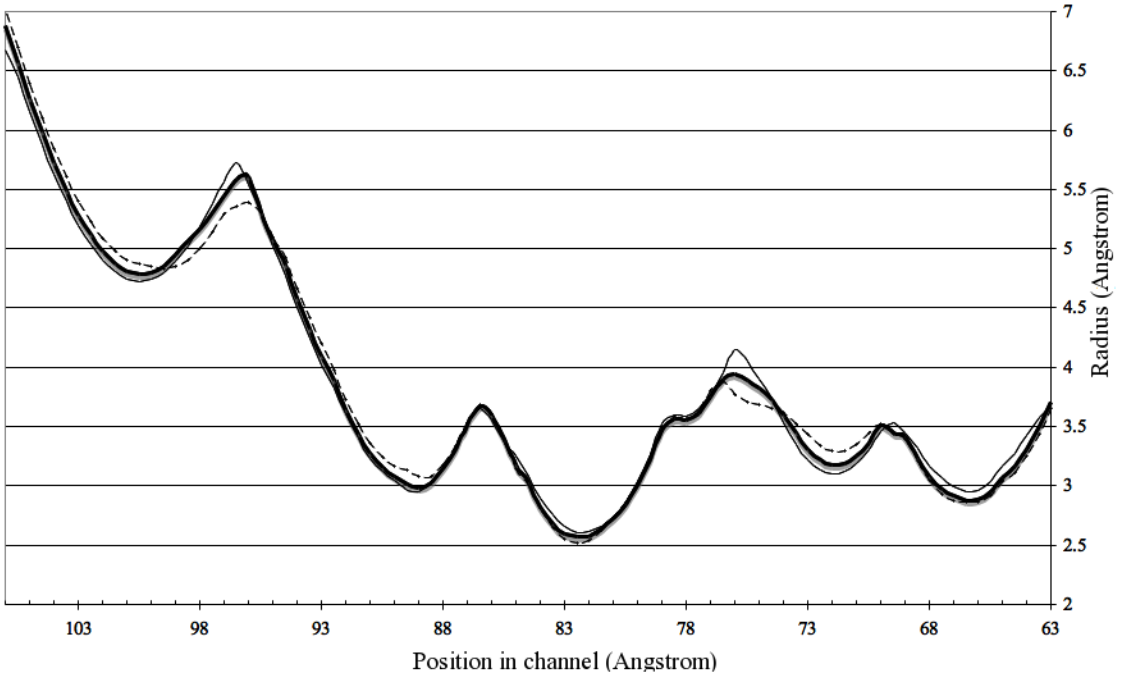


Figure E.2. The pore profile calculated by HOLE in the third slowest mode.

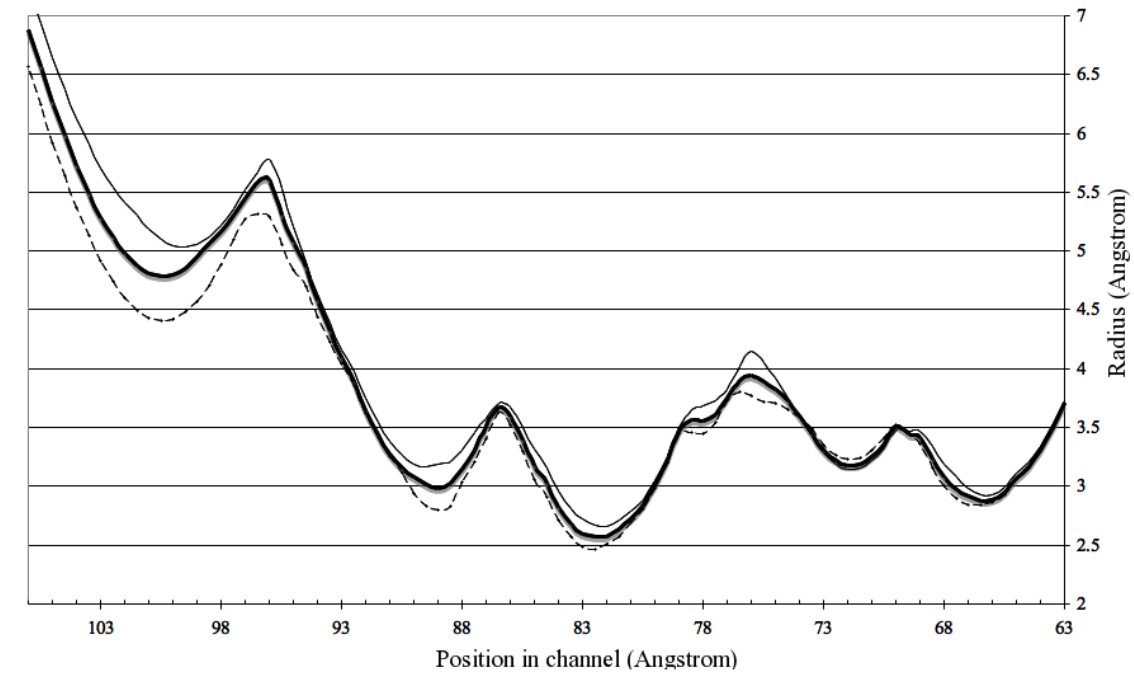


Figure E.3. The pore profile calculated by HOLE in the fifth mode.

# APPENDIX F: MCPOOL ANALYSIS RESULTS

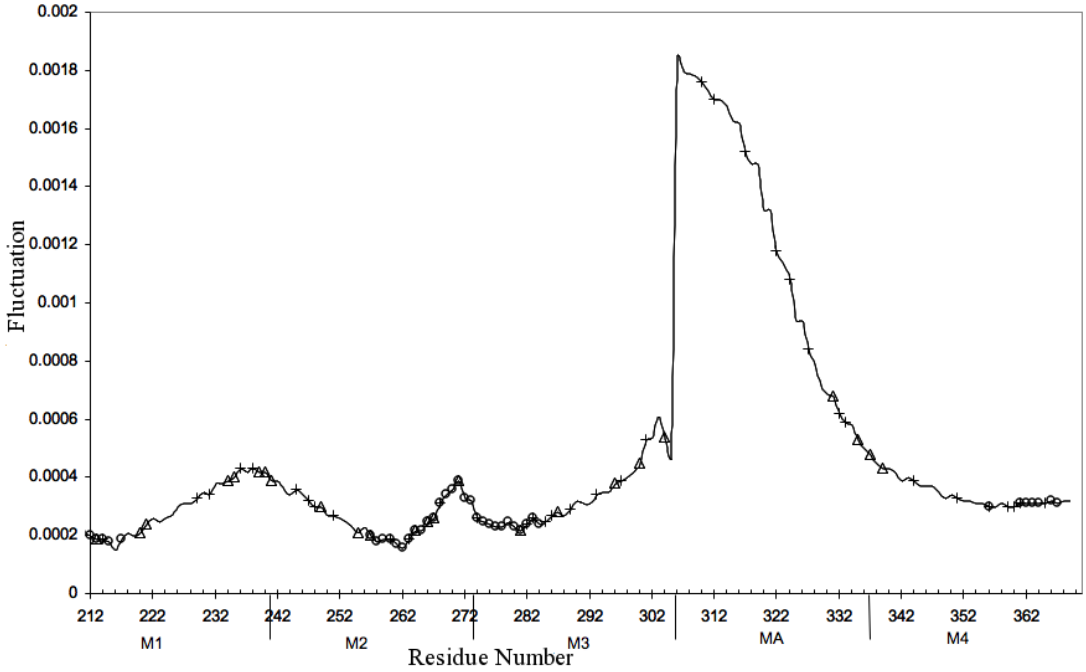


Figure F.1. Most used residues of delta subunit plotted on slow mode fluctuations.

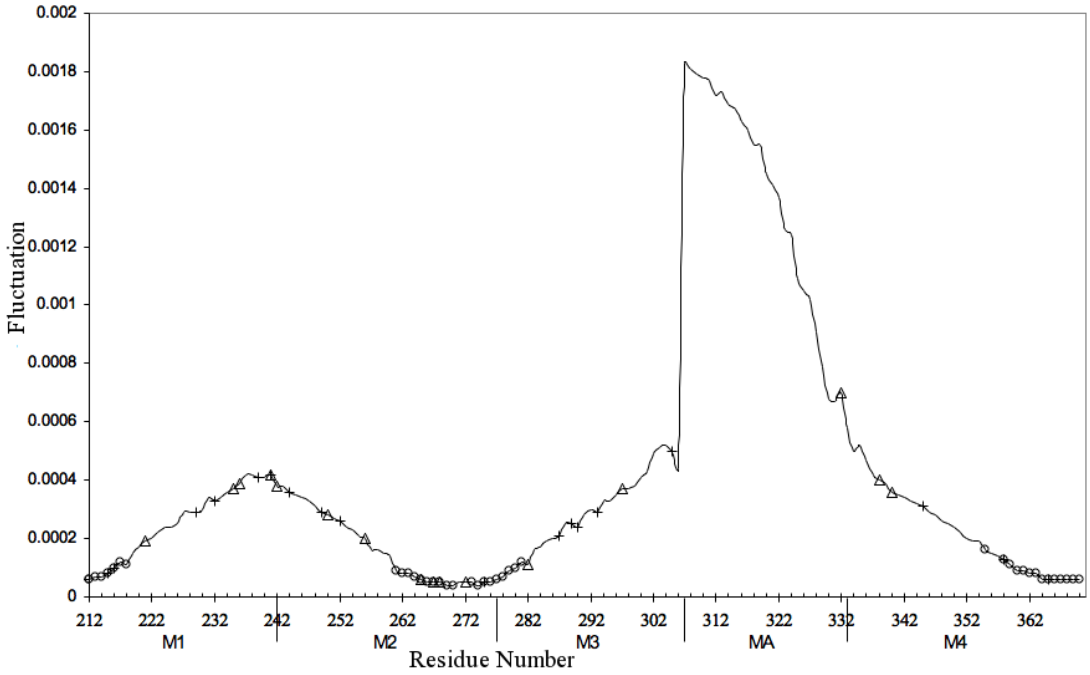


Figure F.2. Most used residues of alpha delta subunit plotted on slow mode fluctuations.

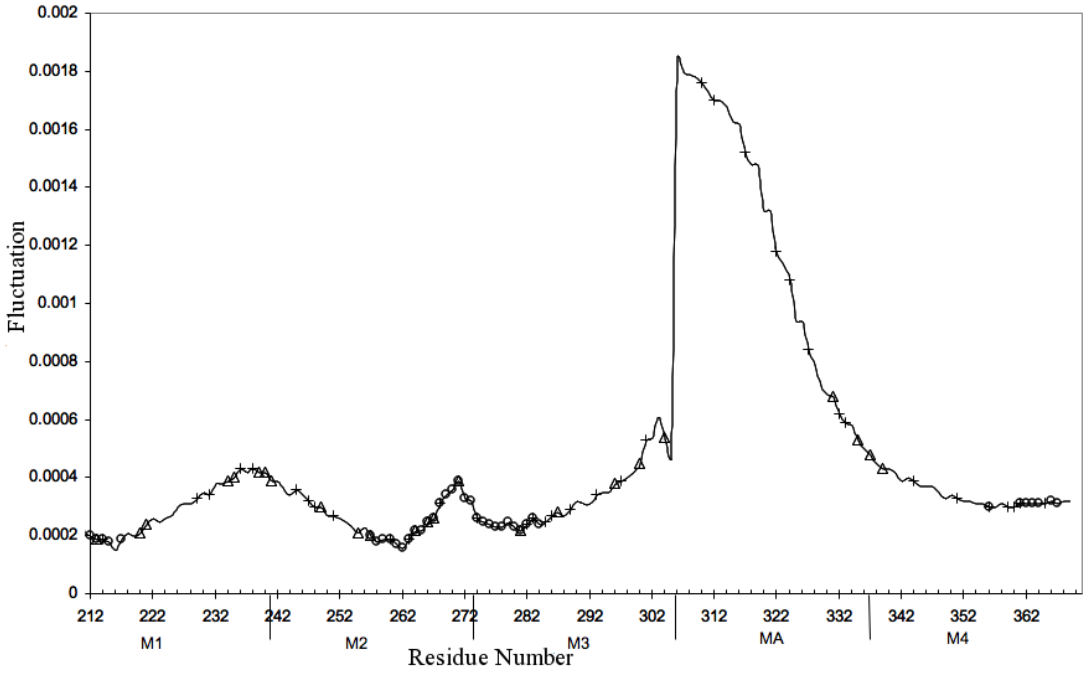


Figure F.3. Most used residues of gamma subunit plotted on slow mode fluctuations.

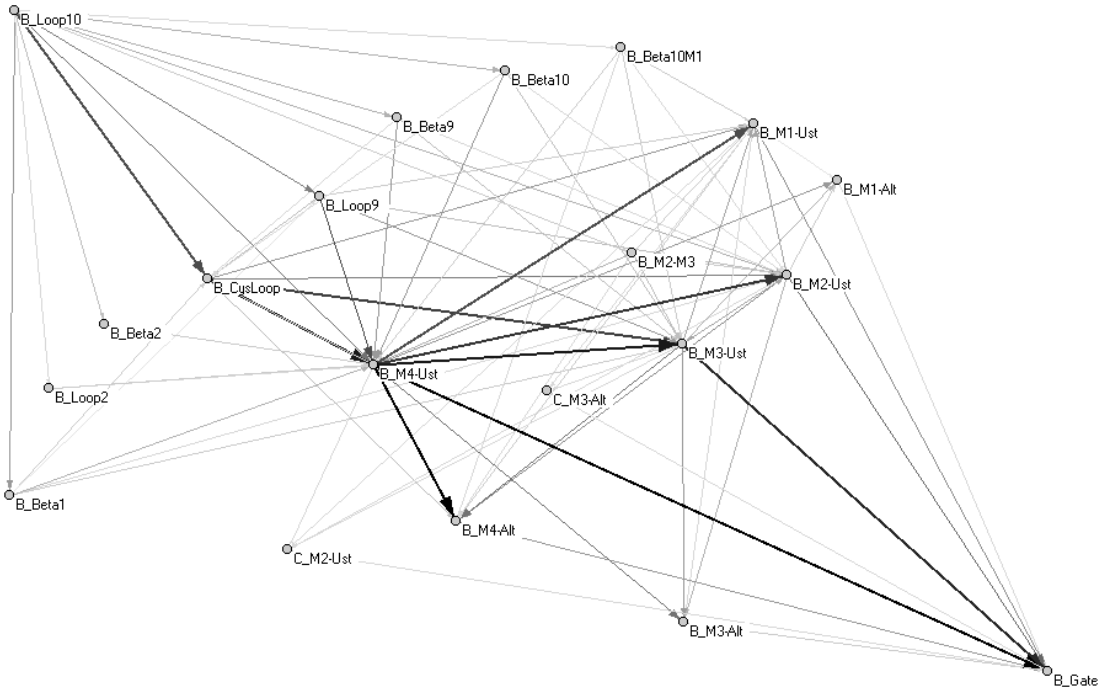


Figure F.4. Network representation of the pathway network from beta loop C (represented by B_Loop10) to the gate region of beta subunit (B_Gate).

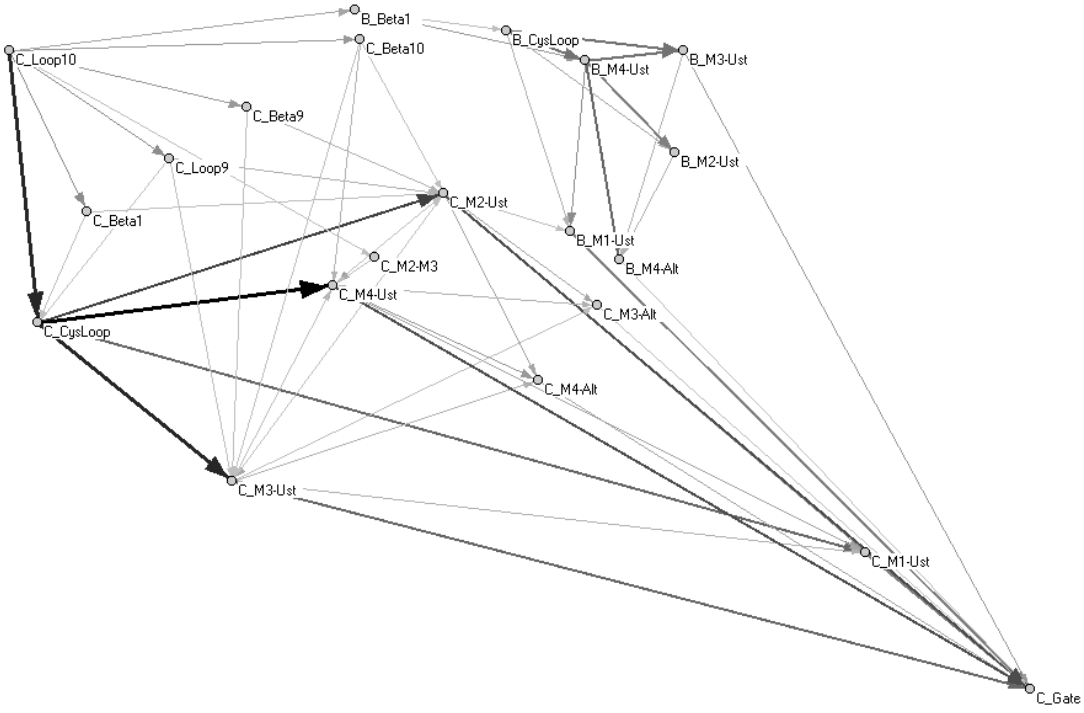


Figure F.5. Network representation of the pathway network from delta loop C (represented by C_Loop10) to the gate region of delta subunit (C_Gate).

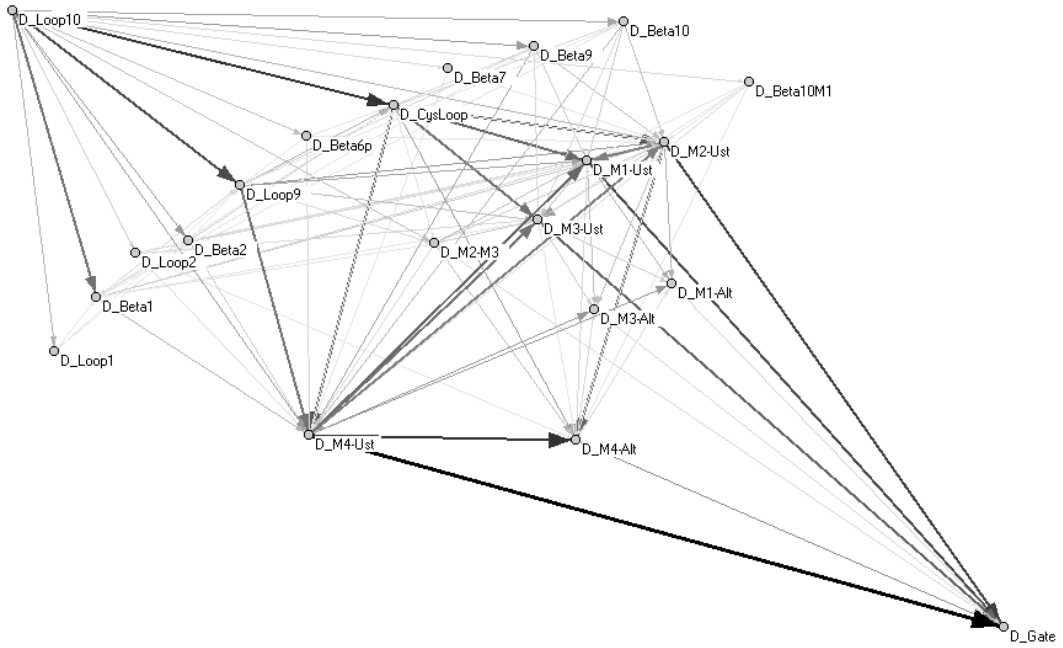


Figure F.6. Network representation of the pathway network from alpha delta loop C (represented by D_Loop10) to the gate region of alpha delta subunit (D_Gate).

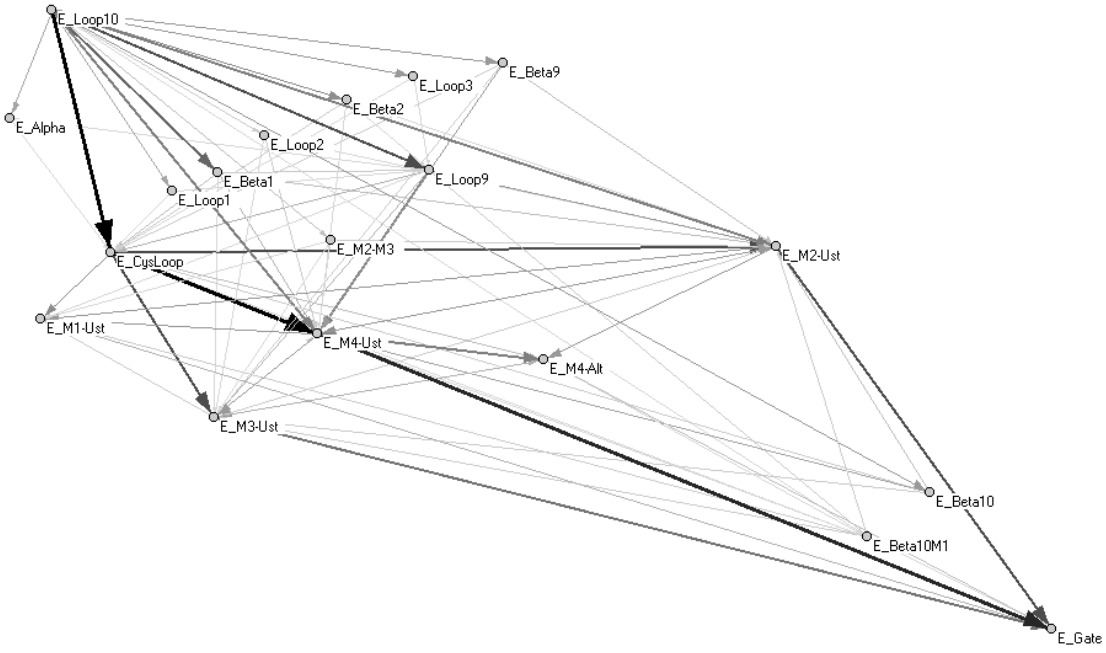


Figure F.7. Network representation of the pathway network from gamma loop C (represented by E_Loop10) to the gate region of gamma subunit (E_Gate).

# REFERENCES


1. Karlin, A., "Ion Channel Structure: Emerging structure of the nicotinic acetylcholine receptors',' *Nature Reviews Neuroscience*, Vol. 3, No. 2, pp. 102-14, February 2002.

2. Barrantes, F. J., "Structural studies of the acetylcholine receptor in the transmembrane environment", *Current Science*, Vol. 81, No. 8, pp. 983-91, October 2001.

3. Schapira M., R. Abagyan and M. Totrov, "Structural model of nicotinic acetylcholine receptor is isotypes bound to acetylcholine and nicotine", *BMC Structural Biology*, Vol. 2, No. 1, pp. 1, January 2002.

4. Barrantes, F. J., "Modulation of nicotinic acetylcholine receptor function through the outer and middle rings of transmembrane domain", *Current Opinion in Drug Discovery & Development*, Vol. 6, No. 5, pp. 620-32, September 2003.

5. Corry, B., "Theoretical conformation of the closed and open states of the acetylcholine receptor channel", *Biochimica et Biophysica Acta*, Vol. 1663, No. 1-2, pp. 2-5, February 2004.

6. Unwin, N., A. Miyazawa, J. Li and Y. Fujiyoshi, "Activation of the Nicotinic Acetylcholine Receptor Involves a Switch in Conformation of the Alpha Subunits", *Journal of Molecular Biology*, Vol. 339, No. 5, pp. 1165 - 76, June 2002.

7. Miyazawa, A., Y. Fujiyoshi and N. Unwin, "Structure and gating mechanism of the acetylcholine receptor pore", *Nature*, Vol. 423, pp. 949-55, June 2003.

8. Burger, K., G. Gimpl and F. Fahrenholz, "Regulation of receptor function by cholesterol", *Cellular and Molecular Life Sciences*, Vol. 57, pp. 1577-92, January 2002.

9. Brejc, K., W. J. V. Dijk, R. V. Klaassen, M. Schuurmans, J. von der Oost, A. B.

Smitt and T. K. Sixma, "Crystal structure of an Ach-binding protein reveals the ligand-binding domain of nicotinic receptors", *Nature*, Vol. 411, pp. 269-76, May 2001.

10. Unwin, N., "Refined Structure of the Nicotinic Acetylcholine Receptor at 4 $\mathring{A}$ Resolution", *Journal of Molecular Biology*, Vol. 346, No. 4, pp. 967-89, March 2005.

11. Bahar, I., "Dynamics of Proteins and Biomolecular Complexes: Inferring Functional Motions From Structure", *Reviews in Chemical Engineering*, Vol. 15, No. 4, pp. 319-47, 1999.

12. Bouzat, C., F. Gumilar, G. Spitsmaul, H. Wang, D. Reyes, S. B. Hansen, P. Taylor and S. M. Sine, "Coupling of agonist binding to channel gating in an Ach-binding protein linked to an ion channel", *Nature*, Vol. 430, pp. 896-900, August 2004.

13. Williamson, P. T. F., B.H. Meier and A. Watts, "Structural and functional studies of the Nicotinic acetylcholine receptor by solid-state NMR", *European Biophysics Journal*, Vol. 33, No. 3, pp. 247-54, May 2004.

14. De Planque M. R.R., D. T.S. Rijkers, J. I. Fletcher, R. M.J. Liskamp and F. Separovic, "The alphaM1 segment of the nicotinic acetylcholine receptor exhibits conformational flexibility in a membrane environment", *Biochimica et Biophysica Acta*, Vol. 1665, No. 1-2, pp. 40-7, October 2004.

15. De Almeida, R. F. M., L. M. S. Loura, M. Prieto, A. Watts, A. Fedorov and F. J. Barrantes, "Cholesterol Modulates the Organization of the gamma M4 Transmembrane Domain of the Muscle Nicotinic Acetylcholine Receptor", *Biophysical Journal*, Vol. 86, No. 4, pp.2261-72, April 2004.

16. Lester, H. A., M. I. Dibas, D. S. Dahan, J. F. Leite and D. A. Dougherty, "Cys-loop receptors: new twists and turns", *Trends in Neurosciences*, Vol. 27, No. 6, pp. 329-36, June 2004.

17. Unwin, N., "Structure and action of the nicotinic acetylcholine receptor explored by electron microscopy", *FEBS Letters*, Vol. 555, No. 1, pp. 91-5, November 2003.

18. Cheng, X., B. Lu, B. Grant, R. J. Law and J. A. McCammon, "Channel Opening Motion of alpha7 Nicotinic Acetylcholine Receptor as Suggested by Normal Mode Analysis", *Journal of Molecular Biology*, Vol. 355, No. 2, pp. 310-24, January 2006.

19. Cruz-Martin, A., J.L. Mercado, L.V. Rojas, M.G. McNamee and J.A. Lasalde-Dominicci "Tryptophan Substitutions at Lipid-exposed Positions of the Gamma M3 Transmembrane Domain Increase the Macroscopic Ionic Current Response of the Torpedo californica Nicotinic Acetylcholine Receptor", *The Journal of Membrane Biology*, Vol. 183, No. 1, pp. 61-70, September 2001.

20. Quiram, P. A., K. Ohno, M. Milone, M. C. Patterson, N. J. Pruitt, J. M. Brengman, S. M. Sine and A. G. Engel, "Mutation causing congenital myasthenia reveals acetylcholine receptor beta/delta subunit interaction essential for assembly", *The Journal of Clinical Investigation*, Vol. 104, No. 10, pp. 1403-10, November 1999.

21. Salamone, F. and M. Zhou, "Aberrations in Nicotinic Acetylcholine Receptor Structure, Function, and Expression: Implications in Disease", *McGill Journal of Medicine*, Vol. 5, No. 2, pp. 90-7, 2000.

22. Katz, B., "Nerve, Muscle and Synapse", *McGraw Hill*, New York, 1996.

23. Giraudat J., C. Montecucco, R. Bisson and J.P. Changeux , "Transmembrane topology of acetylcholine receptor subunits probed with photoreactive phospholips", *Biochemistry*, Vol. 24, No. 13, pp. 3121-27, June 1985.

24. Bouzat, C., F. J. Barrantes and S. M. Sine, "Nicotinic Receptor Fourth Transmembrane Domain Hydrogen Bonding by Conserved Threonine Contributes to Channel Gating Kinetics", *The Journel of General Physiology (JGP)*, Vol. 115, No. 5, pp. 663-71, May 2000.

25. Blanton, M.P. and J.B. Cohen, "Identifying the lipid-protein interface of the Torpedo nicotinic acetylcholine receptor: secondary structure implications", *Biochemistry*, Vol. 33, No. 10, pp. 2859-72, March 1994.

26. De Rosa, MJ, D Rayes, G Spitzmaul and C Bouzat, "Nicotinic receptor M3 transmembrane domain: Position 8' contributes to channel gating", *Molecular Pharmacology*, Vol. 62, No. 2, pp. 406-14, August 2002.

27. Bouzat, C., F. Gumilar, M. del Carmen Esandi and S. M. Sine, "Subunit-Selective Contribution to Channel Gating of the M4 Domain of the Nicotinic Receptor", *Biophysical Journal*, Vol. 82, No. 4, pp. 1920-29, April 2002.

28. Senes, A.I. Ubarretxena-Belandia and D.M. Engelman, "The Calpha —H-O hydrogen bond: a determinant of stability and specifity in transmembrane helix interactions", *Proceedings of the National Academy of Sciences USA*, Vol. 98, No. 16, pp. 9056-61, July 2001.

29. Auerbach, A. L., T. D. Bailey and A. Mitra, "Structural Dynamics of the M4 Transmembrane Segment during Acetylcholine Receptor Gating", *Structure*, Vol. 12, No. 10, pp. 1909-18, October 2004.

30. Castillo, M., J. Mulet, J. A. Bernal, M. Criado, F. Sala and S. Sala, "Improved gating of a chimeric alpha7-5HT3A receptor upon mutations at the M2-M3 extracellular loop", *FEBS Letters*, Vol. 580, No. 1, pp. 256-60, January 2006.

31. Lee, W. Y. and S. M. Sine, "Principal pathway couplingagonist binding to channel gating in nicotinic receptors", *Nature*, Vol. 438, No. 7065, pp. 243-7, November 2005.

32. Xu, Y., F. J. Barrantes, X. Luo, K. Chen, J. Shen and H. Jiang, "Conformational Dynamics of the Nicotinic Acetylcholine Receptor Channel: A 35-ns Molecular Dynamics Simulation Study", *Journal of the American Chemical Society*, Vol. 127, No. 4, 1291-9, February 2005.

33. Hung, A., K. Tai and M. S. P. Sansom, "Molecular Dynamics Simulation of the M2 helices within the Nicotinic Acetylcholine Receptor Transmembrane Domain: Structure and Collective Motions", *Biophysical Journal*, Vol. 88, No. 5, pp. 3321-33, May 2005.

34. Law, R. J., R. H. Henchman and J. A. McCammon, "A gating mechanism proposed from a simulation of a human alpha7 nicotinic acetylcholine receptor", *Proceedings of the National Academy of Sciences*, Vol. 102. No. 19, pp. 6813-8, May 2005.

35. Kash, T. L., A. Jenkins, J. C. Kelley, J. R. Trudell and N. L. Harrison, "Coupling of agonist binding to channel gating in the GABA(A) receptor", *Nature*, Vol. 421, No. 6920, pp. 272-5, January 2003.

36. Taly, A., M. Delarue, T. Grutter, M. Nilges, N. Le Novere and P.J. Corringer, "Normal Mode Analysis Suggests a Quaternary Twist Model for the Nicotinic Receptor Gating Mechanism", *Biophysical Journal*, Vol. 88, No. 6, pp. 3954-65, June 2005.

37. Chen, Y., K. Reilley and Y. Chang, "Evolutionarily conserved allosteric network in the cys-loop family of ligand-gated ion channels revealed by statistical covariance analyses", *Journal of Biological Chemistry*, BEFORE PRINT, April 2006.

38. Bahar, I., A.R. Atilgan, S.R. Durell, R.L. Jernigan, M.C. Demirel and O. Keskin, "Anisotropy of Fluctuation Dynamics of Proteins with and Elastic Network Model", *Biophysical Journal*, Vol. 80, No. 1, pp. 505-15, January 2001.

39. Landau, M., I. Mayrose, Y. Rosenberg, F. Glaser, E. Martz, T. Pupko and N. Ben-Tal, "ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information", *Bioinformatics*, Vol. 19, No. 1, pp. 163-4, January 2003.

40. Landau, M., I. Mayrose, Y. Rosenberg, F. Glaser, E. Martz, T. Pupko and N. Ben-Tal, "ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures", *Nucleic Acids Research*, Vol. 33 W, pp. 299-302, July 2005.

41. Fleishman, S.J., O. Yifrach and N. Ben-Tal, "An Evolutionarily Conserved Network of Amino Acids Mediates Gating in Voltage-dependent Potassium Channels", *Journal of Molecular Biology*, Vol. 340, No. 2, pp. 307-18, July 2004.

42. Bahar, I., "Vibrational Dynamics of Folded Proteins: Significance of Slow and Fast Mode Motions in Relation to Function and Stability", *Physical Review Letters*, Vol. 80, No. 12, pp. 2733-6, 1998.

43. Humphrey, W., A. Dalke and K. Schulten, "VMD - Visual Molecular Dynamics", *Journal of Molecular Graphics*, Vol. 14, No. 1, pp. 33-8, February 1996.

44. Huang, C.C., G.S. Couch, E.F. Pettersen and T.E. Ferrin, "Chimera: An Extensible Molecular Modeling Application Constructed Using Standard Components", *Pacific Symposium on Biocomputing*, Vol. 1, pp. 724, 1996.

45. Smart, O.S., J.M. Goodfellow and B. A. Wallace, "The Pore Dimensions of Gramicidin A", *Biophysical Journal*, Vol. 65, No. 6, pp. 2455-60, December 1993.

46. Leach, A. R., "Molecular Modelling: Principles and Applications", *Pearson / Prentice Hall*, Essex, 2001.

47. Gasteiger, E., A. Gattiker, C. Hoogland, I. Ivanyi, R. D. Appel and A. Bairoch, "ExPASy: the proteomics server for in-depth protein knowledge and analysis", *Nucleic Acids Research*, Vol. 31, No. 13, pp. 3784-8, July 2003.

48. Kass, I. and A. Horovitz, "Mapping Pathways of Allosteric Communication in GroEL by Analysis of Correlated Mutations", *Proteins: Structure, Function, and Gentics*, Vol. 48, No. 4, pp. 611-617, September 2002.

49. Chenna, R., H. Sugawara, T. Koike, R. Lopez, T. J. Gibson, D. G. Higgins and J. D. Thompson, "Multiple sequence alignment with the Clustal series of programs", *Nucleic Acids Research*, Vol. 31, No. 13, pp. 3497-3500, July 2003.

50. Pazos, F., M. Helmer-Citterich, G. Ausiello and A. Valencia, "Correlated muta-

tions contain information about protein-protein interaction", *Journal of Molecular Biology*, Vol. 271, No. 4, pp. 511-23, August 1997.

51. Perez-Jimenez, R., R. Godoy-Ruiz, A. Parody-Morreale, B. Ibarra-Molero and J. M. Sanchez-Ruiz, "A simple tool to explore the distance distribution of correlated mutations in proteins", *Biophysical Chemistry*, Vol. 119, No. 3, pp. 240-6, February 2006.

52. Gansner, E.R. and S. C. North, "An Open Graph Visualization System and Its Applications to Software Engineering", *Software-Practice And Experience*, Vol. 30, No. 11, pp. 1203-33, September 2000.

53. Batagelj, V. and A. Mrvar, "Pajek - Program for Large Network Analysis", *Connections*, Vol. 21, No. 2, pp. 47-57, 1998.

54. Barrantes, F. J., "Lipid matters: nicotinic acetylcholine receptor-lipid interactions (Review)", *Molecular Membrane Biology*, Vol. 19, No. 4, pp. 277-84, October 2002.

55. Sine S. M. and A. G. Engel, "Recent Advances in Cys-Loop receptor structure and function", *Nature*, Vol. 440, No. 7083, pp. 448-55, March 2006.