

# Msc Thesis

by

Ebru Çetin

Investigation of the intracellular loop 3 and allosteric mechanisms of  $\beta_2$ -  
Adrenergic among G-protein coupled receptors

Master of Science

in

Computational Sciences and Engineering



**KOÇ**  
**UNIVERSITY**

August 17, 2017

MSc Thesis

Koç University

Graduate School of Sciences and Engineering

This is to certify that I have examined this copy of a master's thesis by

Ebru Çetin

and have found that have found that it is complete and satisfactory in all respects,  
and that any and all revisions required by the final  
examining committee have been made.

Committee Members:

---

Prof.Dr. Burak Erman

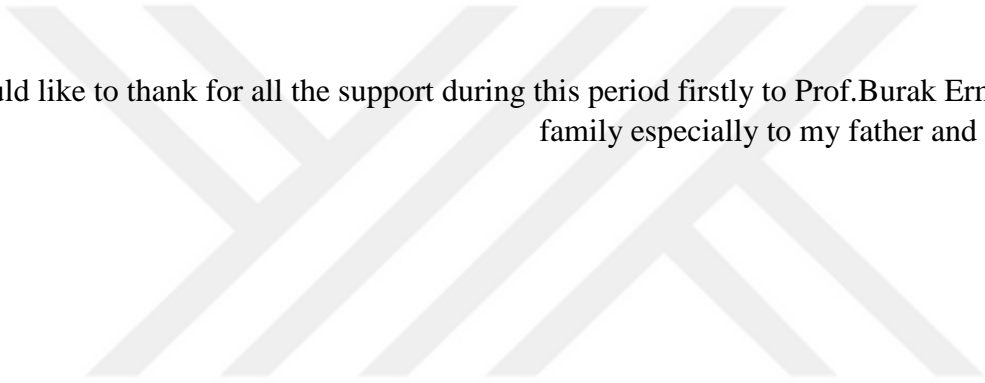
---

Prof.Dr. A. Levent Demirel

---

Prof. Dr. Türkan Haliloğlu

Date:



I would like to thank for all the support during this period firstly to Prof.Burak Erman and my family especially to my father and my mother.

## ABSTRACT

G-protein coupled receptors are encoded over 700 genes in human genome. They share a common seven transmembrane domain and differing intracellular and extracellular loops. As transmembrane domains, they function as gate-keepers of intercommunication of cells. Thus, their functioning mechanisms are one of most unique ongoing research area. In this study after the introduction of GPCR families, human  $\beta$ 2-adrenergic receptor and endogenous ligand epinephrine are investigated. First of all, usually omitted part of  $\beta$ 2-Ar intracellular loop 3 and its unbinding energy from inactive to the active state is estimated. Secondly, intracomunications between residues are investigated via correlation matrices.

In allostery analysis two trajectories of  $\beta$ 2-Ar embedded in lipid membrane are analyzed. One of the trajectories is gathered from 1  $\mu$ s molecular dynamics simulation and second trajectory is gathered from 500 ns molecular dynamics simulation where the ligand binding pocket is restrained to 8 Å in order to mimic ligand binding.

Specifically for  $\beta$ 2Ar, experimental measurements show that the distance range of 8Å - 10Å between Asp113 and Ser207 is sufficient for receptor activation.

In 1  $\mu$ s simulation, it is observed that ICL1 (residues 64 to 66) is in communication with intracellular end of TM6 (Cys265, Lys267). Thr66 of ICL1 is important in protein stabilization in lipid membrane and functions through Tyr123 (TM3) and Ile154 (TM4). Trp99 of ECL1 is correlated to ECL2 (Cys191) and membrane region of TM5 (Ile 214). Phe101 of ECL1 is correlated extracellular end of TM6 and intracellular end of TM5.

In 500 ns simulation, Met98 of ECL1 is correlated to intracellular end of TM1 (Arg53), ICL1 (Phe61), intracellular end of TM2 (Thr68) and cytoplasmic tail (Leu339). Gly102 of ECL1 is correlated to membrane region of TM2 (Glu82), intracellular end of TM3 (Arg131), ICL2 (Ser143) and intracellular end of TM5 (Phe223). Ile135 of TM3 at intracellular region is correlated to extracellular part of TM2 (Ala92), membrane region of TM4 (Val160, Leu163) and membrane region of TM5 (Ala202). These finding imply that ICL1 may be responsible for ligand recognition signals whereas ECL1 is responsible for protein stabilization and recognition of intracellular effectors.

In estimation of required energy for unbinding of intracellular loop 3 is found as 1629 kJ/mol and confirmed with G-protein coupling energy to  $\beta$ 2-Ar with the work of -1455 kJ/mol.

Finally, the endogenous ligand epinephrine of  $\beta$ -adrenoceptors is studied in order to estimate critical differences between  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$  ligand binding pockets for epinephrine and their selectivity.

## ÖZETÇE

G-protein kenetli reseptörler insan genomunda 700'den fazla gen ile kodlanmaktadır. G-protein kenetli reseptörler ortak 7 transmembran yapısı ve değişken hücre içi ve hücre dışı düğümlerden oluşur. Transmembran reseptörleri olarak, hücrelerin hücre dışı ile iletişimlerini sağladıkları birer kapı ve kapıcı görevi sürdürmektedirler. Bu nedenle G-protein kenetli reseptörlerin işleyiş mekanizmaları güncel bilimsel sorulardan biri olarak önemli bir yer kapsamaktadır. Bu çalışmada G-protein kenetli reseptörlerin takdimi, insan  $\beta$ 2-adrenergic proteinin karakteristikleri ve  $\beta$ -adrenergic reseptörlerin doğal ligandının bağlanma afinitesi araştırılmıştır. Literatürde genellikle  $\beta$ 2-Ar'ın üçüncü hücre içi ilmiği yapılan çalışmalarda yoksayılmaktadır. Çalışmanın ilk ayağı olarak  $\beta$ 2-Ar üçüncü hücre içi ilmiğinin proteinin inaktif halinden aktif halindeki konformasyona geçişi için gerekli olan enerji hesaplanmıştır.

ICL3 bölgesinin aktif olmayan konformasyondan aktif konformasyonuna getirilebilmesi için gereken enerji 1629 kJ/mol olarak hesaplanmış, G-protein bağlanmasından açığa çıkan enerjinin -1455 kJ/mol olduğu farklı bir moleküler simülasyon hesabı ile desteklendiği gözlemlenmiştir.

Çalışmanın ikinci ayağı olarak rezidüler arası iletişim ağları korelasyon matrislerinin analizi üzerinden tespit edilmesi hedeflenmiştir.

Allosteri analizinde  $\beta$ 3-Ar proteinin lipid hücre zari içine gömülü halde simüle edildiği iki farklı gidişiz kullanılmıştır. Gidişizlerden bir tanesi 1  $\mu$ s boyunca monomeric ve herhangi bir kısıtın bulunmadığı simülasyonun sonucunda, bir tanesi ise 500 ns boyunca reseptörün aktif durumundaki yapısı tanımlanmış olan Asp113 ve Ser207 arasında 8 Å kısıt uygulanarak ligand bağlanmasının mimic edildiği simülasyonun sonucunda elde edilmiştir.

1  $\mu$ s simülasyon sonucunda, birinci hücre içi ilmiğinin (ICL1), transmembrane 6'nın hücre içi sonlanma bölgesiyle (Cys265, Lys267) iletişim halinde olduğu gözlemlenmiştir. Birinci hücre içi ilmiğinde yer alan Thr66'nın proteinin lipid membrane içerisindeki stabilizasyonunda rol oynadığı ve Tyr123 (TM3) ve Ile154 (TM4) üzerinden işleyişini gerçekleştirdiği gözlemlenmiştir. Birinci hücre dışı ilmiğinde yer alan Trp99'un ikinci hücre dışı düğümünde yer alan Cys191 ve transmembrane 5'in membrane bölgesinde bulunan Ile214 ile yine birinci hücre dışı ilmiğinde yer alan Phe101'in transmembrane 6'nın hücre dışı bölgesi ile transmembran 5'in hücre içi kısmıyla korele olduğu gözlemlenmiştir.

500 ns süren ve ligand bağlanmasını mimic eden simülasyonda ise birinci hücre dışı ilmiğinde yer alan Met98'in Phe61 (ICL1), Thr68 (transmembrane 2'nin hücre içi sonu) ve Leu339 (sitoplazmik düğüm) ile iletişim halinde olduğu gözlemlenmiştir. Yine birinci hücre dışı ilmiğinde bulunan Gly102'nin ise Glu82 (TM2 membran bölgesi), Arg131 (TM3 hücre içi sonu), Ser143 (ikinci hücre içi ilmiği) ve Phe223 (TM5 hücre içi sonu) ile bir iletişim ağı kurduğu gözlemlenmiştir. Üçüncü iletişim ağı ise transmembrane 3'te bulunan Ile 135 ile Ala92 (TM hücre dışı bölgesi), Val160, Leu163 (TM4 membran bölgesi) ve Ala202 (TM5 membran bölgesi) arasında olduğu gözlemlenmiştir. Elde edilen bulgular birinci hücre içi ilmiğinin (ICL1) ligand tanıma sinyalleri ile ilintili iken ECL1'in protein stabilizasyonu ve hücre içi efektörlerin tanımlanmasında rol oynayabiliyor olacağını göstermektedir.

Son olarak, epinephrine'in  $\beta$ -adrenergic reseptörleri arasında seçiciliği  $\beta_2 > \beta_3 > \beta_1$  olarak gözlemlenmiştir.

## ACKNOWLEDGMENTS

All trajectories used in allostery analysis and initial conformation of intracellular loop 3 unbinding energy calculations is contributed by Assoc. Prof. Ebru Demet AKTEN, Kadir Has University.



# TABLE OF CONTENTS

Chapter 1	
1.1 Introduction .....	1
Chapter 2	
2.1 GPCR Structure and Function .....	2
2.1.1 G-Protein and Corresponding Signaling Pathways .....	2
2.1.2 Class A G-protein coupled receptors: Rhodopsin Family .....	5
2.1.3 Class B G-protein coupled receptors .....	8
2.1.4 Class C G-protein coupled receptors: Glutamate Family .....	8
2.1.5 Class F Frizzled / TAS2 .....	9
2.2 $\beta_1$ Adrenergic .....	10
2.3 $\beta_2$ Adrenergic .....	11
2.4 $\beta_3$ Adrenergic .....	15
Chapter 3	
3.1 Potential Mean Force Analysis of Beta-2-Adrenergic at G-Protein Binding Site .17	
3.1.1 Steered MD Simulations .....	19
3.1.1.1 $\beta_2$ -Adrenergic Intracellular Loop 3 Unbinding Energy Calculations .....	20
3.1.1.2 Force and Work Calculations .....	20
3.1.2 0.5 A/ns Pulling .....	20
3.1.3 2.4 A/ns Pulling .....	21
Chapter 4	
4.1 Allosteric Couplings of Beta-2-Adrenergic	
4.1.1 What is allostery? .....	25
4.1.2 Two-state Allosteric Models .....	27
4.1.3 Computational Approaches .....	30
4.1.3.1 Structural Surveys .....	30
4.1.3.1.1 Statistical Coupling Analysis .....	30
4.1.3.1.2 Principal Component Analysis .....	30
4.1.3.1.3 Elastic Network Models .....	30

4.1.3.2 Comparisons within conformational ensembles .....	31
4.1.3.2.1 Monte Carlo simulation .....	31
4.1.3.2.2 Coarse grained molecular simulations .....	31
4.1.4 Important Interactions .....	32
4.1.5 Method .....	33
4.1.6 1000 ns APO $\beta$ 2-AR .....	34
4.1.7 500 ns 8 Å Restrained $\beta$ 2-AR .....	41
4.1.7.1 0 – 100 ns .....	41
4.1.7.2 400 – 500 ns .....	44
<b>Chapter 5</b>	
5.1 Epinephrine Binding to $\beta$ -adrenoceptors .....	51
5.2 Method .....	51
5.3 Binding Pockets .....	51
5.4 SMD Simulations .....	52
<b>Chapter 6</b>	
6.1 Conclusion .....	56
<b>Appendix I</b> .....	57
<b>Appendix II</b> .....	58
<b>Appendix III</b> .....	60



## LIST OF TABLES

Table 3.1.1.1 .....	19
Table 3.1.2.1 .....	20
Table 3.1.3.1 .....	21



## LIST OF FIGURES

Figure 2.1 .....	3
Figure 2.2 .....	4
Figure 2.3.1 .....	11
Figure 3.1.2.1 .....	20
Figure 3.1.2.2 .....	21
Figure 3.1.3.1 .....	22
Figure 3.1.3.2 .....	22
Figure 3.1.3.3 .....	23
Figure 3.1.3.4 .....	23
Figure 3.1.3.5 .....	24
Figure 4.1.2.1 .....	28
Figure 4.1.6.1 .....	34
Figure 4.1.6.2 .....	35
Figure 4.1.6.3 .....	35
Figure 4.1.6.4 .....	36
Figure 4.1.6.5 .....	36
Figure 4.1.6.6 .....	37
Figure 4.1.6.7 .....	37
Figure 4.1.6.8 .....	38
Figure 4.1.6.9 .....	38
Figure 4.1.6.10 .....	39
Figure 4.1.6.11 .....	39
Figure 4.1.6.12 .....	40
Figure 4.1.6.13 .....	40
Figure 4.1.7.1.1 .....	41
Figure 4.1.7.1.2 .....	41
Figure 4.1.7.1.3 .....	42
Figure 4.1.7.1.4 .....	42

Figure 4.1.7.1.5 .....	43
Figure 4.1.7.1.6 .....	43
Figure 4.1.7.1.7 .....	44
Figure 4.1.7.2.1 .....	44
Figure 4.1.7.2.2 .....	45
Figure 4.1.7.2.3 .....	45
Figure 4.1.7.2.4 .....	46
Figure 4.1.7.2.5 .....	46
Figure 4.1.7.2.6 .....	47
Figure 4.1.7.2.7 .....	47
Figure 4.1.7.2.8 .....	48
Figure 4.1.7.2.9 .....	48
Figure 4.1.7.2.10 .....	49
Figure 4.1.7.2.11 .....	49
Figure 5.4.1 .....	52
Figure 5.4.2 .....	52
Figure 5.4.3 .....	53
Figure 5.4.4 .....	53
Figure 5.4.5 .....	54
Figure 5.4.6 .....	54
Figure 5.4.7 .....	55

## Chapter 1

### 1.1 Introduction

G-protein coupled receptors are one of the largest families which encode over 700 genes in the human genome and %40 of drug market targets GPCR proteins. Thus, their functioning mechanism is one of the utmost important topics of current scientific community. They mainly function in mediation of adenylyl cyclase dependent pathways and tissue specific pathways such as clathrin-cargo signalling events. They mainly couple to  $G_i$ ,  $G_q$ ,  $G_s$  and  $G_{12/13}$  G-alpha proteins. General characteristics of G-protein coupled receptors are defined in Chapter 1. In this study, mainly the role of intracellular loop 3 in G-protein coupled activation, inner mechanism of  $\beta_2$ -adrenergic protein and binding abilities of endogenous ligand of beta-adrenergic protein are investigated.  $\beta_2$ -adrenergic protein functions in smooth muscle relaxation and breathing. Dysfunctioning of  $\beta_2$ -adrenergic causes asthma. We tried to shed light into the role of intracellular loop 3 of  $\beta_2$ -ar which is usually omitted in structural investigation of  $\beta_2$ -ar activation. In Chapter 3, the intracellular loop 3 of inactive state of  $\beta_2$ -ar is taken as initial conformation and the energy required to bring the intracellular loop 3 into active-state conformation is calculated via Steered Molecular Dynamics. In Chapter 4, one of the current important subjects of scientific community, the subject of allostery, is investigated by comparing different points of views and perspectives. Two different trajectories where the protein is embedded in the membrane are generated. One trajectory, 1000 ns NVT, was generated in the absence of ligand (apo). The second was a 500 ns NVT trajectory in which the ligand binding pocket was fixed at 8 Å, thus mimicking the ligand bound state of the protein.

In Chapter 5, the endogenous ligand of beta-adrenergic receptor epinephrine binding to  $\beta_1$ Ar,  $\beta_2$ Ar and  $\beta_3$ Ar is investigated using Steered Molecular Dynamics Method.

## 2.1 GPCR Structure and Function

G-protein coupled receptors correspond to one of the largest protein families which covers over 700 genes of human genome. They share a common 7 transmembrane alpha helix structure with altering intracellular and extracellular coil formations. The GPCR family has 7 subclasses 2 of which (Class D and Class E) are not found in vertebrates. [1] Deciphered genes in vertebrates belonging to the GPCR family divides proteins of this family into 5 subclasses as Rhodopsin (class A); Secretin and Adhesion (Class B); Glutamate (Class C); and Frizzled/Taste Receptor 2. (Class F) [2] GPCR proteins are in charge of signal transfer to cytoplasmic region recognizing extracellular stimuli and transferring acquired information through photons, ions, ligands, peptides or proteins to intracellular agents or acts in formation, activation and deactivation of ion channels. Thus, they are important stations of intercellular communication and most targeted receptors of marketed drugs. [2] To understand these critical information stations, a collaborative effort was established as The GPCR Network in 2010 funded by NIGMS's Protein Structure Initiative based at The Scripps Research Institute. The GPCR Networks effort resolved 19 unique structures and deposited to the Protein Data Bank. The Protein Data Bank has 49 unique structures of GPCR receptors and 198 structures of GPCRs in different conformations were revealed of these structures, 133 belong to Homo sapiens as of 2017.

### 2.1.1 G-Protein and Corresponding Signaling Pathways

G protein is the next step of information transfer which comes from the corresponding G-protein coupled receptor and are activated by phosphorylation. They are heterotrimeric proteins which serve as molecular switches via association and dissociation of the trimeric sub-structure and composed of 3 subunits designated as  $\alpha$ ,  $\beta$  and  $\gamma$ . Activated G-protein coupled receptors cause dissociation of  $\alpha$  subunit of G-protein from  $G\beta\gamma$  dimer and initiates signal transferring cascades. After the signaling procedure is complete,  $G_\alpha$  subunits eventually hydrolyse the bound GTP into GDP and reforms trimeric ( $\alpha\beta\gamma$ ) structure of G-protein.

The most common signaling pathways associated can be classified as  $G_{\alpha(i)}$ ,  $G_{\alpha(q)}$ , and  $G_{\alpha(s)}$  pathways.

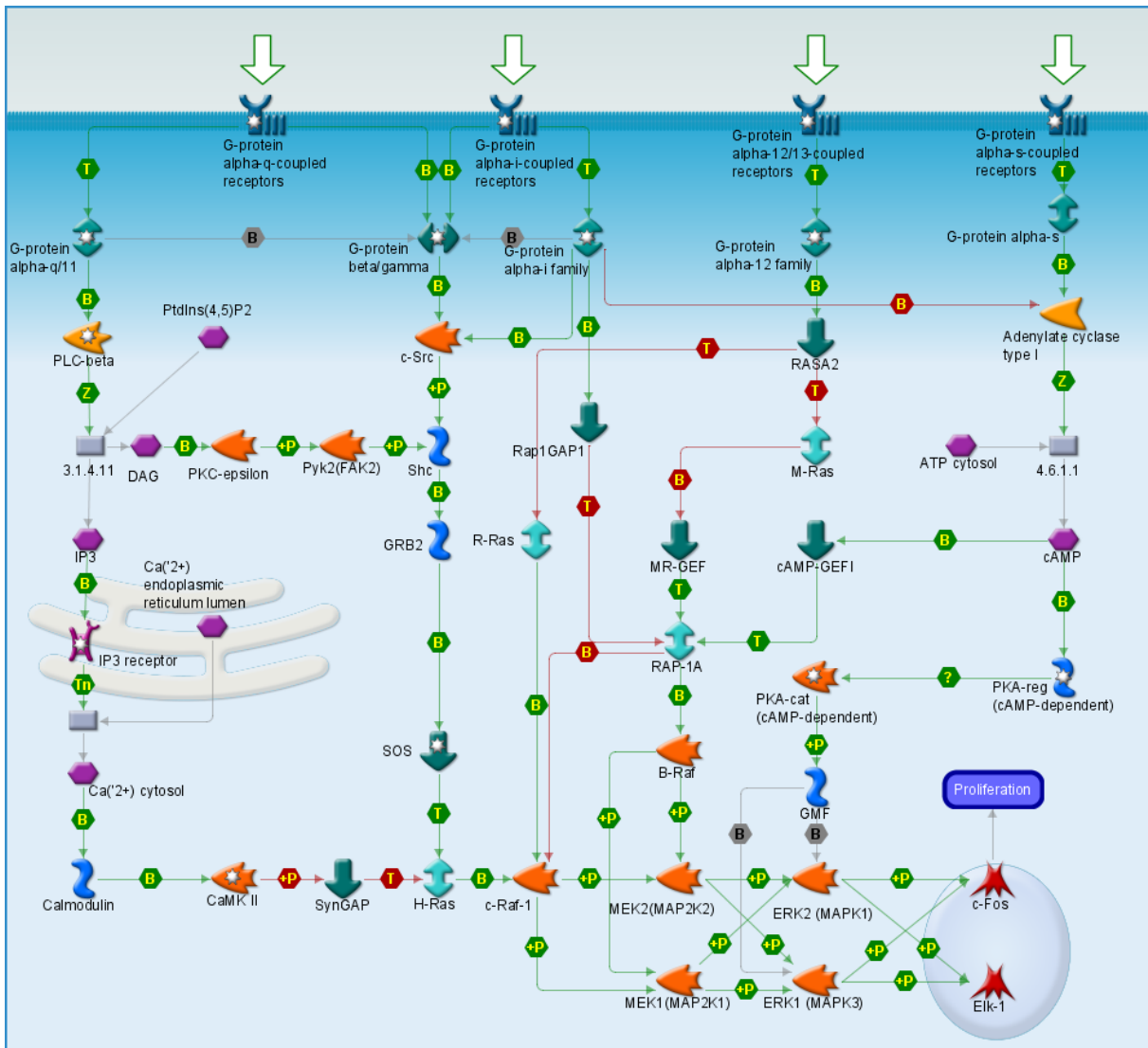


Figure 2.1. G-protein coupled receptors and activated G-protein signaling pathways

$G_{\alpha(i)}$  pathway also known as  $G_{i/o}$  pathway inhibits adenylyl cyclase activity thus inhibits the cAMP dependent pathway. The  $G_{\alpha(i)}$  pathway mainly works opposite to the cAMP-dependent protein kinase. Adenylyl cyclase has 6 classes. Their main activity is to transform ATP to 3'-5' cyclic AMP (cAMP) phosphatase. Produced cAMP will function as a signal for cAMP binding proteins.

$G_{\alpha(q)}$  ( $G_{q/11}$ ) signaling initiates the formation of inositol-1,4,5-triphosphate 3 (IP3) by phospholipase C and activates protein tyrosine phosphorylation. Phosphorylation of  $G_{\alpha(q11)}$  activates  $G_{q11}$  leading to release of  $Ca^{2+}$  from intracellular stores. [3]

The  $G_s$  pathway activates the adenylyl cyclase pathway leading to the activation of protein kinase A (pKA) dependent intracellular events.

$G_{12/13}$  signaling occurs a bit differently than other  $G_{\alpha}$  pathways. After activation of cAMP pathway  $G_{13}$  subunit binds to Rho guanine nucleotide exchange factors (RhoGEFs) therefore initiating regulation of cell specific pathways.

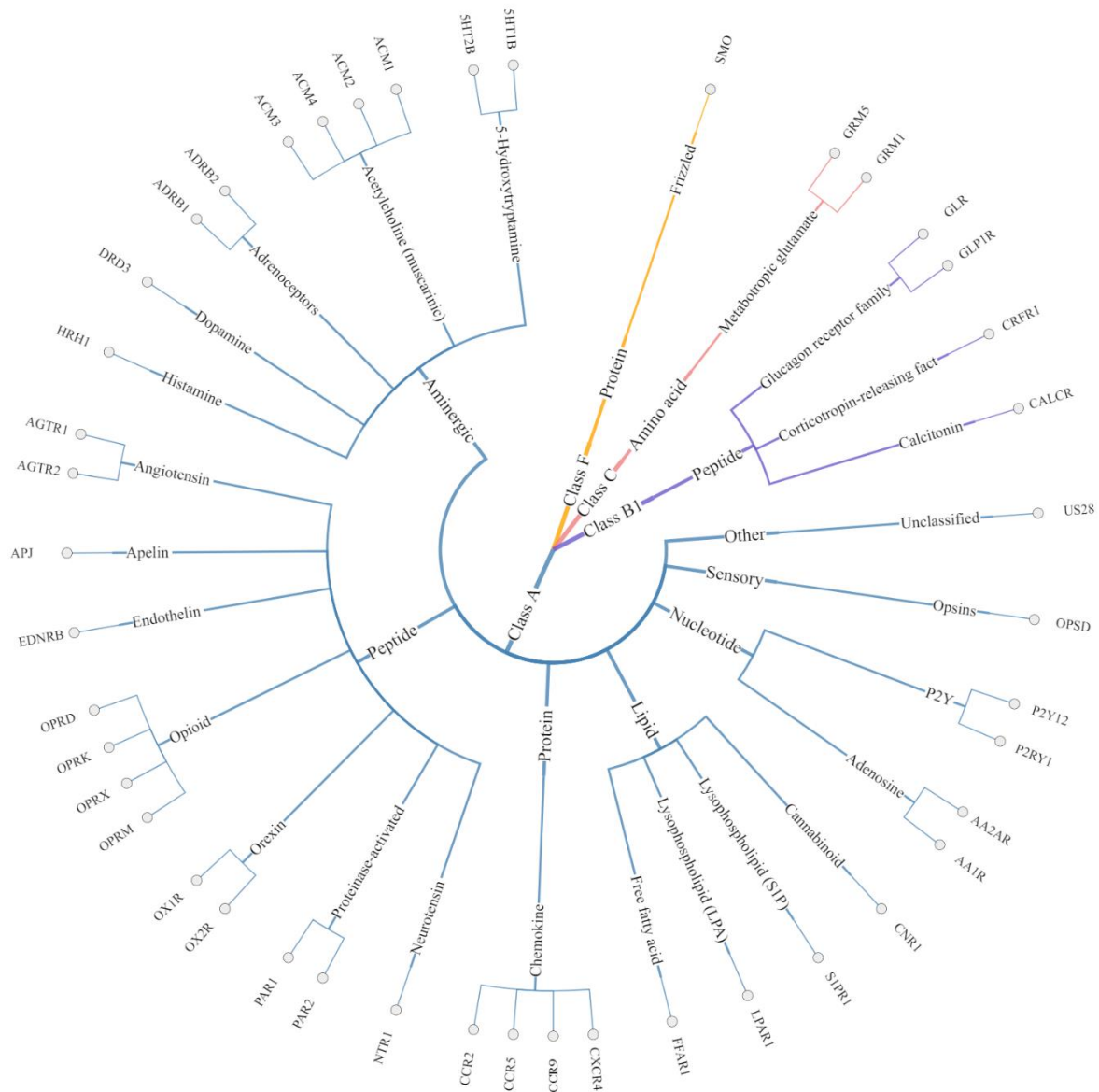


Figure 2.2 G-protein coupled receptors polygenetic tree

### 2.1.2 Class A G-protein coupled receptors: Rhodopsin Family

The Rhodopsin family consists of 719 genes functioning in a wide variety of mechanisms. The sub families of the Rhodopsin class divide into 4 main branches ( $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ ) which further divide into sub clusters according to their functional similarities. The Olfactory receptors of the  $\delta$ -group comprise the largest group of Rhodopsin family with 390 genes. Important sub clusters of Class A are listed as: Orphans, Olfactory receptors, Vision, Taste, Pheromone receptors and receptors with known ligands. Receptors with known structures are also divided into 18 receptor types.

#### *5-Hydroxytryptamine receptors:*

Thirteen receptors are found in this subfamily [1] but only 2 unique structures are resolved. Their main function is the regulation of serotonin, dopamine and acetylcholine uptake and downstream regulation of signaling. [4, 5] 5-Hydroxytryptamine receptors are also known for various ergot alkaloid derivatives and psychoactive substances. They affect neural activity, but their function in perception of pain and cerebral vasoconstriction is speculated. Two known structures belong to 5-Hydroxytryptamine 1B and 5-Hydroxytryptamine 2B. 5-Hydroxytryptamine 1B signaling inhibits adenylyl cyclase activity and 5-Hydroxytryptamine 2B signaling activates a phosphatidylinositol-calcium second messenger system that modulates the activity of phosphatidylinositol 3-kinase and down-stream signaling cascades and promotes the release of  $\text{Ca}^{2+}$  ions from intracellular stores.[6]

#### *Acetylcholine receptors (muscarinic) :*

Muscarinic receptors consist of 5 types as M1 to M5. Their main function is the inhibition of adenylyl cyclase, breakdown of phosphoinositide and modulation of the potassium channels. [7] Structures of M1 to M4 have been resolved so far, excluding M5.

#### *Adenosine receptors :*

Adenosine receptors are classified as A1, A2A, A2B and A3 receptors. They function in the activation and inhibition of adenylyl cyclase. Structures of A1 and A2A are deposited to Protein Data Bank with various ligands.

#### *Adrenoceptors :*

Adrenoceptors also called adrenergic receptors branch into two subgroups denoted as  $\alpha$  and  $\beta$ . The Alpha subgroup consists of 6 receptors ( $\alpha_{1A}$ ,  $\alpha_{1B}$ ,  $\alpha_{1D}$ ,  $\alpha_{2A}$ ,  $\alpha_{2B}$ ,  $\alpha_{2C}$ ) and beta subgroup consists of  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$  adrenoceptors.  $\alpha$ -adrenoceptors functions in  $G_{\alpha(i)}$  and  $G_{\alpha(q)}$  signaling pathways whereas  $\beta$ -adrenoceptors functions in  $G_{\alpha(s)}$  signaling pathway. [8]

#### *Angiotensin receptors :*

Angiotensin is a peptide hormone which causes vasoconstriction in control of blood pressure. Angiotensin type 1 and Angiotensin type 2 constitutes an angiotensin receptor subgroup, Angiotensin type 1 transduces signals trough the  $G_{\alpha(q)}$  signaling pathway in activation of phospholipase C (PLC) whereas Angiotensin type 2 signals trough the  $G_{\alpha(i)}$  pathway in inhibition of adenylyl cyclase.



### *Apelin receptor :*

The Apelin receptor inhibits adenylyl cyclase activity and functions in control of blood pressure, heart contractility and heart failure. In early heart development, Apelin receptor functions in gastrulation and morphogenesis via APELA hormone

### *Cannabinoid receptors:*

CB1 and CB2 receptors mainly function in the nervous system in the brain and the immune system. They mainly transduce signals in the  $G_{\alpha(i)}$  signaling pathway; however, it has been also reported that CB1 also plays a role in calcium, potassium ion channels and kinase pathways. Structure and signaling mechanisms of CB1 is more abundant in literature whereas mechanisms of CB2 have not been presented yet. [9]

### *Chemokine receptors:*

Chemokines constitute one of the largest sub families which have a main group coupled to G-proteins. Their main function is the regulation of leukocyte trafficking. Some other chemokines also function in non-G-protein activated processes. [10] Chemokines are also divided into 4 subgroups according to structure similarity and conservation of cysteine residues: CC chemokines, CXC chemokines, CX3C chemokines and C chemokines. G-protein coupled chemokines are CCR1 to CCR10, CXCR1 to CXCR6, CX<sub>3</sub>CR1, XCR1 and CCRL2. Their signaling pathways varies through leukocyte trafficking, nevertheless they also transduce signals through  $G_{\alpha(i)}$  pathway.

### *Dopamine receptors:*

Dopamine receptors are important in the central nervous system in regulation of the neurotransmitter dopamine. There are 5 types of dopamine receptors, of which D1 and D5 transduce signals through the  $G_{\alpha(s)}$  pathway and D2, D3, D4 through the  $G_{\alpha(i)}$  pathway. Only crystal structure of D3 has been resolved up to now. [11]

### *Endothelin receptors:*

They are receptors for 3 types of endothelin 1-3 and classified as ETA, ETB receptors. Only the Endothelin type B structure is known as a non-specific receptor for all types of endothelin functioning in the  $G_{\alpha(q)}$  pathway in activation of phosphatidylinositol-calcium, second messenger system and  $G_{\alpha(i)}$  pathway. Endothelin type A receptor function in  $G_{\alpha(q)}$  and  $G_{\alpha(s)}$  pathways. [12]

### *Free fatty acid receptors:*

Receptors activated by long-chain saturated and unsaturated fatty acids are classified as FFA1 to FFA4 and GPR42. [13] FFA1 and FFA4 are activated by long-chain fatty acids whereas FFA2 and FFA3 are activated by short fatty acids. FFA1 and FFA2 are important in homeostasis and act through  $G_{\alpha(i)}$  and  $G_{\alpha(q)}$  pathways.[14]

### *Histamine receptors:*

Histamine receptors have 4 types: H1, H2, H3 and H4. The structure of H1 is resolved and acts on the  $G_{\alpha(q)}$  pathway and plays a role in smooth muscle contraction, increase in capillary permeability, and neurotransmission in central nervous system and catecholamine release from adrenal medulla.

### *Lysophospholipid (LPA) receptors:*

Lysophospholipid receptors are activated through phospholipid metabolite LPA and consists of 6 types from LPA1 to LPA6. Besides LPA activation. They can be activated by some phosphorylated intermediates, endocannabinoids or heterogeneous LPA molecules. LPA1 crystal structure is resolved and LPA1 is known for signaling through  $G_{\alpha(i)}$  and  $G_{\alpha(q)}$  pathways and downstream MAP kinases. They are required for normal brain development and proliferation and plays important role in chemotaxis and wounding.

### *Neurotensin receptors:*

Neurotensin I (Nstr1) and Neurotensin II (Nstr2) are receptors activated via neurotensin. The activated receptor signals through G-protein coupled phosphatidylinositol-calcium second messenger system and leads to activation of downstream MAP kinases. [15]

### *Opioid receptors:*

Opioid receptors consist of 4 types with resolved structures and classified as  $\delta$ ,  $\kappa$ ,  $\mu$  and Nociceptin/Orphanin FQ (NOP) receptor. The NOP receptor is classified in some studies as the opioid-like receptor due to its distinct pharmacology. Delta, kappa and mu receptors are widely found in the brain and spinal cord. They are responsible for pain modulation and addiction, regulation of membrane ionic homeostasis, cell proliferation, emotional response, epileptic seizures, immune function, feeding, obesity, respiratory and cardiovascular control and some neurodegenerative disorders. [16] All of these receptors act in the  $G_{\alpha(i)}$  signaling pathway.  $\delta$ ,  $\kappa$ ,  $\mu$  receptors controls inhibition of adenylyl cyclase activity and regulation of calcium ion current and potassium ion conductance.  $\mu$  receptor differing from  $\delta$  and  $\kappa$  plays a role in regulation of mitogen-activated protein kinase (MAPK), phospholipase C (PLC), phosphoinositide/protein kinase (PKC), phosphoinositide 3-kinase (PI3K) and NF-kappa-B. [17] Additionally, agonist specific phosphorylation of the receptor leads to G-protein dependent or independent activation of the ERK pathway. NOP receptor mediates alternative signaling pathways in activation of MAP kinases. [18]

### *Orexin receptors:*

Orexins are receptors activated by either orexin-A or orexin-B and have 2 types as OX1 and OX2. OX1 has high affinity for orexin-A whereas OX2 is non-selective over orexin-A or orexin-B. They trigger an increase of cytoplasmic calcium concentration in presence of orexin-A binding. [19] They also signal through  $G_{\alpha(q)}$  pathway.

### *P2Y receptors:*

P2Y receptors consist of 8 types as: P2Y<sub>1</sub>, P2Y<sub>2</sub>, P2Y<sub>4</sub>, P2Y<sub>6</sub>, P2Y<sub>11</sub>, P2Y<sub>12</sub>, P2Y<sub>13</sub> and P2Y<sub>14</sub>. They are receptors for ATP, ADP, uridine triphosphate, uridine diphosphate and UDP-glucose. They processes signals through the  $G_{\alpha(i)}$ ,  $G_{\alpha(q)}$  and  $G_{\alpha(s)}$  pathways. Crystal structures of P2Y<sub>1</sub> and P2Y<sub>12</sub> are resolved. P2Y<sub>12</sub> is activated by ATP and ADP and inhibits the adenylyl cyclase

second messenger system. P2Y<sub>1</sub> functions in mobilization of intracellular calcium ions via activation of phospholipase C. [20]

### 2.1.3 Class B G-protein coupled receptors

Adhesion and Secretin family constitutes Class B GPCRs. Secretin family is encoded with 15 genes in human and adhesion family is phylogenetically related to class B family with 33 genes however they conceive a large N-termini with a GPCR auto proteolysis inducing domain (GAIN) over 320 residues. [21] Secretin family has 2 unique protein structures deposited to Protein Data Bank of Corticotrophin-releasing factor receptors and Glucagon receptor family.

*Corticotrophin-releasing factor receptors:*

Corticotrophin-releasing factor receptors have 2 types as CRF1 and CRF2 both of which are activated non-selectively by corticotrophin-releasing hormone and urocortin. Crystal structure of CRF1 is deposited to Protein Data Bank. CRF1 functions in activation of adenylyl cyclase through the G<sub>α(s)</sub> signaling pathway and inhibits the activity of the calcium channel CACNA1H.

*Glucagon receptor family:*

Glucagon receptor family consists of 6 proteins which are activated via the endogenous glucagon hormone. Only the structure of the glucagon receptor is known and it's functioning in control of glucose levels in blood through G<sub>α(s)</sub> signaling pathway. [22]

### 2.1.4 Class C G-protein coupled receptors: Glutamate Family

Class C includes metabotropic glutamate receptors, calcium sensing receptor (CaS) and GABA<sub>B</sub>, three taste type 1 receptors and V2 receptors which are encoded with 22 genes. Structures of GABA<sub>B</sub> and few of metabotropic glutamate receptors are resolved.

*Metabotropic glutamate receptors:*

mGluRs consist of 8 types. Only 2 of the 8 structures are resolved. They are receptors for glutamate and perform a variety of functions in the central nervous system. mGluR1 mediates signaling through the phosphatidylinositol-calcium second messenger system and participates in long-term potentiation in hippocampus and long term depression in cerebellum. mGluR5 plays an important role in synaptic plasticity and neural network. Both of them mediate through G<sub>α(q)</sub> pathway. [23]

*GABA<sub>B</sub>:*

GABA is a heterodimeric protein formed by GABBR1 and GABBR2. GABBR1 binds to ligand whereas GABBR2 mediates G-protein coupling. [24] GABA functions in inhibition of the adenylyl cyclase pathway, high voltage calcium channel leading to neurotransmitter inhibition and activation of potassium channels and phospholipase A2 and modulates inositol phospholipid hydrolysis. [25]

### **2.1.5 Class F Frizzled / TAS2**

Frizzled family is encoded with 10 genes and activated by secreted lipoglycoproteins of WNT family which passes signals into cell via membrane proteins. Class F also includes SMO gene for smoothed receptor.

*Smoothened Homolog:*

Smoothened homolog is thought to be transducing hedgehog's protein signal indirectly.



## 2.2 $\beta_1$ Adrenergic

$\beta_1$ -Adrenergic is located mainly in the heart and regulates cardiac function and  $\text{Ca}^{2+}$  homeostasis. [27] Thus it is one of the utmostly important  $\beta$ -adrenergic receptors. However, currently human  $\beta_1$ -adrenergic receptor has not yet been crystallized due to its flexible conformation. Engineered form of turkey (*Meleagris gallopavo*)  $\beta_1$ -adrenergic has been crystallized by Warne et. al. as mutant  $\beta_1\text{AR-m23}$  complex.[28]

$\beta_1, \beta_2, \beta_3$  of family A shares % 51 sequence identity except amino and carboxyl termini and cytoplasmic loop and also human  $\beta_1$  and  $\beta_2$  receptors share %67 sequence identity. Thus one of the main challenges about  $\beta$ -adrenergic receptors is determining ligand specificity to targeted  $\beta$  receptor. Sugimoto et. al. showed that antagonist CGP 20712A has selectivity over  $\beta_1\text{AR}$  via binding 500 times more strongly than to  $\beta_2\text{AR}$  and ICI-118,551 binds 550 times more strongly to  $\beta_2\text{AR}$  than  $\beta_1\text{AR}$ . [29]

One of well known antagonist of  $\beta_1\text{-Ar}$  is carvedilol and it is not sufficiently selective over  $\beta_1\text{AR}$ . However they are still used for treatments of cardiovascular diseases. In 1997, clinical trials conducted by Doughty and Sharpe showed that metoprolol and carvedilol targeting  $\beta_1\text{AR}$  blockage improve LV (Left Ventricle) function and hemodynamics similarly. However it has been stated that even though carvedilol has a weak  $\beta_1$ -selectivity, it has other potentially beneficial effects such as anti-proliferative and antioxidant activity. [30]

Radio ligand binding assay study conducted on  $\beta$ -blocker selectivity at  $\beta_1$  and  $\beta_2$  receptors by Smith and Teitler reports ligands that bisoprolol, betaxolol, metoprolol and atenolol are selective of  $\beta_2\text{AR}$  over  $\beta_1\text{AR}$  whereas carvedilol, propranolol and ICI-188,551 has no selectivity over either receptor. [31]

As structural specifications ionic lock between helices 3 and 6 observed in  $\beta_2\text{AR}$  is also in  $\beta_1\text{AR}$  and aromatic interaction between TM7 and helix 8 in NPxxYxF motif. Another interaction that attracts attention is the salt bridge network between extracellular loop 2 and extra cellular loop 3 at residues D184/D186 and R317, respectively. In the binding pocket it has been observed that only Phe325 in  $\beta_1\text{Ar}$  equivalent of Tyr308 behaves differently in  $\beta_1\text{Ar}$ . Upon ligand binding hydrophobic interactions lock ECL2 in a rigid conformation, in the presence of sodium ion, ECL2 gets coordinated with backbone carbonyl groups of residues Cys192, Asp195 and Cys198 and two water molecules. Hydrogen bonds between side chains of Ser<sup>(5.42)</sup>, Ser<sup>(5.41)</sup> and Tyr<sup>(5.38)</sup> is observed and two water molecules in the cavity between TM3, TM4 and TM5. Phe<sup>(5.32)</sup> could be playing an important part in the regulation entry to binding pocket. Phe325 cannot form hydrogen bonds with Asn<sup>(6.55)</sup> as observed in  $\beta_2\text{Ar}$ .

## 2.3 $\beta_2$ Adrenergic

$\beta_2$  adrenergic after rhodopsin has been one of the mostly studied GPCR. The structure of  $\beta_2$ Ar first revealed in 2010 and has been a model for abounding number of studies. Flexible structure of  $\beta_2$ Ar and signaling capabilities to different pathways makes GPCRs an important area to work on and makes  $\beta_2$ Ar special as on-research model of GPCR mechanisms. In literature ionic lock on the mostly conserved (D/E)R(W/Y) sequence at the at the cytoplasmic end of TM3 and acidic amino acid at the cytoplasmic end of TM6 [34], rotamer toggle switch observed at W286 and F290 modulating bend angle of TM6 upon activation [34,35], salt bridge between Asp192 and Lys305 [3] is reported as important constituents of protein. Ligand binding pocket characteristics and agonist/antagonist binding mechanisms, GPCR independent signaling mechanisms of  $\beta_2$ Ar (basal activity,  $\beta$ -arrestin dependent pathways), activation and G-protein selective activation mechanisms are mostly acknowledged and studied issues.

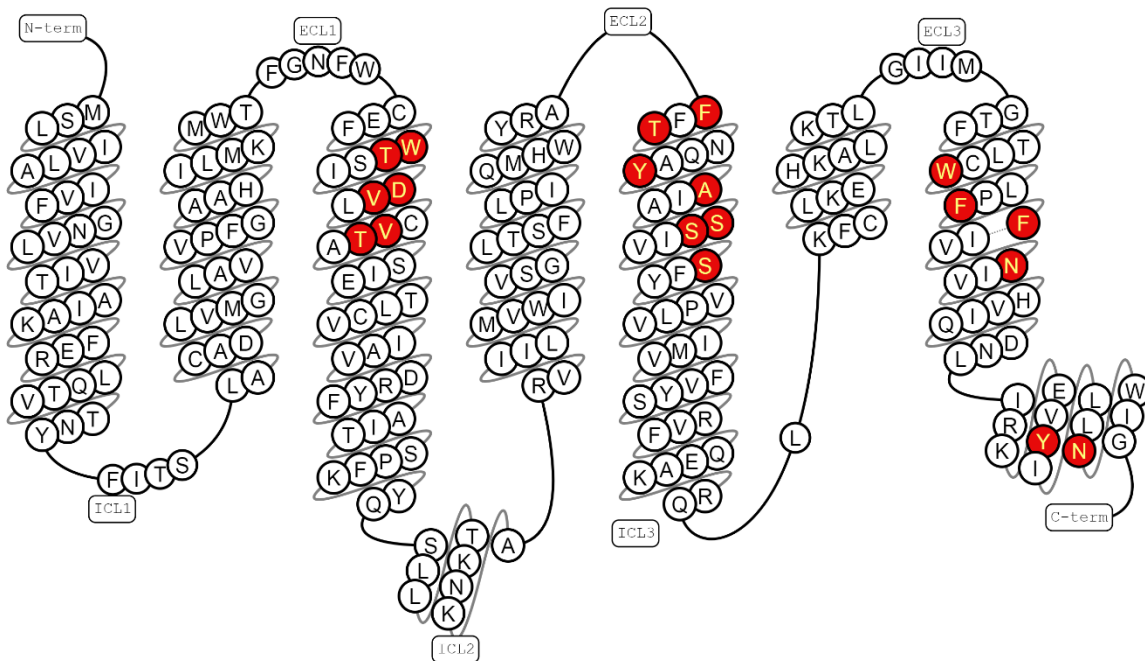


Figure 2.3.1 Chain representation of  $\beta_2$ -Ar (PDB code: 2RH1)

### *Toggle switch:*

CWxP motif in TM6 (toggle switch) and NPxxY motif is known as important in receptor activation. Kobilka and his group findings present that disruption of ionic lock and activation of toggle switch are not related events. [34]

### *Ionic Lock:*

Ionic lock between Arg131 and Glu268 is seen during the study carried out by Vanni et.al. [33] In ionic lock disruption side chain conformation of  $\beta_2$ Ar is asserted as an important indicator for signaling levels as in  $\beta_1$ Ar, Ser211 is observed in a distinctly different conformation. [33]

### *Salt Bridge:*

Salt bridge interactions involved in  $\beta_2$ Ar indicates that Asp192, Lys305 is important in salt bridge formation and disruption of salt bridge between Arg131 and Glu268 seems as a critical actor for shifting equilibrium of protein towards to active state. [33]

### *Hydrogen Network:*

Upon activation 2 Å inward shift of TM5 results in catechol 'head' re-orientation with Ser203 and Ser204 side chains and Asn293 stabilizes TM5 shift. [36]. Additionally Vanni et.al observed that ethanolamine groups of ligand and interaction with Asp113 and Asn313 is conserved for 0.5 microsecond.

### *Intracellular Loop 3 (ICL3):*

In crystal data of  $\beta_2$ Ar, intracellular loop 3 is removed and T4 lysosome is inserted. Due to this, most studies lacks the importance of ICL3 in  $\beta_2$ Ar signaling mechanism. In work of Özgür et.al.  $\beta_2$ Ar is investigated with ICL3 for around 1 microsecond. Also, HDX data suggests that ICL3 plays a role as molecular switch where upon antagonist and inverse agonist binding ICL3 covers the region at the ends of TM5 and TM6 G-protein binding site and shifts equilibrium toward inactive state. [46] In their work they observe that around 600 ns of MD 1 $\mu$ s, ICL3 packs underneath  $\beta_2$ Ar and keeps its stationary state until the rest of simulation. In second continuation run, 5 Å RMSD value towards active state is observed around 20 ns. After ICL3 packing a net of hydrogen bonds are formed and maintained till the end of simulation. ICL3 packing seemed to significantly correlated with lower end of TM6 causing 7.5 Å inward movement in the opposite direction to 14 Å outward movement of TM6.

It is also known that ICL3 is important in G-protein activation, the peptides synthesized with primary sequence 259-273 corresponding to ICL3 region G $\alpha$ s-coupled  $\beta_2$ Ar selectively bind with G-proteins stimulate their functional activity.[58 ,59] ICL3 derived peptides is crucial for interacting with negatively charged receptor binding site of G $\alpha$ (s) subunit.

### *G-protein signaling:*

In G-protein activation, G-protein interacts with TM5 and TM6, ICL2 and ICL3 but makes no important interactions with TM6 and TM7. Conformation change occurring in core region generally affects ICL2 and ICL3 intracellular loop (which are directly linked to TM3 and TM4 that constitutes one of the key sites for G-protein recognition and activation. [40]. $\beta_2$ Ar-G $\alpha\beta\gamma$  heterotrimeric complex shown that G $\alpha$  subunit of G-protein makes contact with TM5 and TM6 resulting 14 Å outwards movement of TM6 upon activation. [37] Position and conformation of

TM8 does not change upon activation. [38] Phosphorylation is also important in G-protein signaling, in the work of Daaka et.al it has been seen that  $\beta_2$ Ar efficiently coupled to G $\alpha_s$  but not G $\alpha_i$ . Once phosphorylated by PKA becomes efficiently coupled to G $\alpha_i$  but not G $\alpha_s$ . Moreover, peptide study conducted by Bockaert and Pin shows that a peptide corresponding to the TM6 of  $\beta_2$ Ar inhibits both receptor dimerization and activation suggesting dimerization may be important for G-protein activation. [40]

#### *$\beta$ -arrestins:*

In G-protein independent signaling, phosphorylated GPCRs recruit  $\beta$ -arrestins in competing with G-proteins and desensitize G-protein signaling. Binding of  $\beta$ -arrestins to agonist bound GPCRs promotes clathrin-dependent endocytosis. Arrestins binds to GPCRs phosphorylated by GRKs [38].

$\beta$ -arrestin mediated signaling includes mitogen-activated protein kinases (MAPKs), non-receptor protein kinases, glycogen synthase kinase 3, protein phosphatase 2A and transcription factors [41]

Schukla et.al. concluded that TM6 and TM7 are most probably involved in  $\beta$ -arrestin biased signaling.

#### *Downstream signaling:*

$\beta$ -arrestin and G-protein signaling pathways seems to be interconnected on several levels and may be involving other activation and rearrangement of cytoplasmic and periplasmic loops.[42] It has been observed that ligands for a given GPCR can show different efficacies for coupling to distinct signaling pathways. [43] Basal activity may be reflecting the inherent flexibility of a protein and its tendency to exist in several signaling pathways in the absence of ligands. It also could refer to a highly constrained state with a relatively high affinity for G-protein [34]. It is known that  $\beta_2$ Ar shows higher basal activity when pH is reduced to 6.5 from 7.5. pH increase probably disrupts the ionic lock and other intermolecular interactions as a result of protonation [44]. It is also reported that mutation of Leu272 to alanine results in increased basal activity [45].

#### *Ligand Interactions:*

In ligand interactions it has been seen that distinction between  $\beta_1$  and  $\beta_2$  receptors is based on relative affinities for catecholamine agonists and their distribution in cell tissue. [33] Specifically for  $\beta_2$ Ar, experimental measurements show that distance range of 8Å - 10Å between Asp113 and Ser207 is sufficient for receptor activation. [46] Full agonists upon binding establishes two major interactions in which amino groups form a salt bridge with Asp113 carboxyl, whereas the polar groups of catechol or similar moiety interacts with Ser204 side chains in TM5 [47, 48]. It has been observed that  $\beta$ -OH group in agonist forms a polar interaction with Asn293 in  $\beta_2$ Ar. [49]  $\beta$ -OH moiety of ethanolamine 'tail' of agonists has seen consistently anchored to the side between Asp113 and Asn312 as in inverse agonists and antagonists known to structure. [36] Additionally, in the fluorescence labeling studies conducted by Liu et.al shows that full agonist isoproterenol bound  $\beta_2$ Ar shows equal populations at Cys327 and Cys265. Partial agonists tulobuterol, clonbuterol, and norepinephrine have less effect on Cys265 than isoproterenol whereas full agonist farmoterol elicits a greater shift in conformational equilibrium state on Cys265. The inverse agonist



carazolol makes no apparent effect compared to apo form. Isoetharine and carvedilol cause large shift on Cys327 toward to active state. Carvedilol induce no effect on Cys327 whereas isoetharine procures a large shift on Cys265 toward active state. [37] The antagonist/weak partial agonist alprenolol can induce a small shift on Cys265 but minimal effect of Cys327 in agreement with previous findings. [49, 50] Large activation changes isoetharine and carvedilol performs on Cys327 equilibrium suggests that TM7 primarily effects on  $\beta$ -arrestin signaling pathway. [37]

Kobilka and his co-workers resulted that catechol does not disrupt the ionic lock and salbutamol does not activate rotamer toggle switch and concluded that norepinephrine occurs trough at least one conformational change which has an intermediate state similar to dopamine.

Carvedilol previously thought as an inverse agonist characterized as a weak  $\beta$ -arrestin biased ligand of  $\beta_2$ Ar. [51, 52, 53] Carvedilol bound  $\beta_2$ Ar preserves a closed conformation as ICL3 becomes more buried in the presence of carvedilol as reference to full agonist isoproterenol. [51] Carvedilol makes contact with Leu101 (TM2), Asp200 (ECL2), TRY207 (ECL2), TRP330 (TM7) and Phe201 (ECL2). [38]

In ligand based interactions it has been seen that in order to achieve an activation, around 2 Å inward shift of TM5 is required. [54] In the binding pocket Asp113, Asn313 and Ser202 are important in both agonist and antagonist binding. Ser203, Ser206 and Asn293 are not involved in antagonist binding. [55, 56, 57] Tyr 308 is proposed as an important residue responsible of agonist selectivity. Tyr 308 is located in close proximity to ligand interface and solvent and having a hydrogen bond contact with Asn293.

## 2.4 $\beta_3$ - Adrenergic

$\beta_3$ Ar of beta-adrenoceptors sub family mainly locates in adipose tissue and fewer extent in skeletal muscle.[60]  $\beta_3$ Ar acts upon enhancement of lipolysis and thermogenesis. [61]  $\beta_3$ -Ar when activated couples with Gs signaling pathway and induce/inhibit adenylyl cyclase pathway.

Encoding of murine (mouse)  $\beta_3$ Ar receptor translates to a 388 amino acid residue protein which shows %82 homology with  $\beta_3$ Ar. [62] Striking difference between murine  $\beta_3$ Ar and human  $\beta_3$ Ar is the carboxyl-terminal tail of the murine  $\beta_3$ Ar is 11 residues shorter than that of the human  $\beta_3$ Ar, and the high homology between these two proteins ends after the conserved cysteine residue (position 358 of murine sequence) which was shown to anchor the carboxy-terminus of the  $\beta_3$ Ar in the membrane through palmitoylation. Only 3 serines present in murine  $\beta_3$ Ar and they are not favorable in palmitoylation. Palmitoylation is a type of post-translational modification acts as lipid binding on cysteine and threonine residues. [62, 63, 64]

Another important difference noted between mouse and the human  $\beta_3$ Ar are found in the intracellular loops, close to the transmembrane regions where three arginine residues (at positions 61,150 and 289 of murine sequence) are substituted for a tryptophan, a cysteine and a second cysteine respectively in the human receptor. [62]

Known agonists of  $\beta_3$ Ar include Amibegron (SR-58611A) [65, 66], CL-316,243 [67], L-742,791 [68], L-786,568 [69], LY-368,842, Migabegron [70], RO40-2148, Solabegron [71]. Known antagonists are L-748,328, L-748,337 and SR 59230A. [68]

Additionally,  $\beta_3$ Ar is characterized with a low affinity toward antagonist lodocyanopindolol, and low efficiency toward propranolol, ICI-118,551 and CGP 20712A in inhibition of cAMP production induced by isoproterenol. [62] BRL 37344, a weak agonist for  $\beta_1$ Ar and  $\beta_2$ Ar behaves as potent activator of lipolysis and thermogenesis in rodent fat cells and strongly inhibits rat colon motility. [72, 73]

In the work of Ligett et.al beta-adrenoceptors is expressed in CHCW cells for measuring Gs coupling and stimulating adenylyl cyclase activity. The rank order of potency for agonists for stimulating adenylyl cyclase was isoproterenol = BRL 37344 > norepinephrine > epinephrine for  $\beta_3$ Ar. Along with agonist activity they have observed that  $\beta_3$ Ar did not display short-term agonist promoted functional desensitization or sequestration or long-term down regulation. [64]

In the work of Lewis et.al potent beta adrenergic antagonist (-)[<sup>3</sup>H]dihydroalprenolol was used to directly estimate the number and affinity of beta-adrenergic receptors in rat heart membranes from control and hyperthyroid rats. Their result showed that isoproterenol and (-)[<sup>3</sup>H]dihydroalprenolol affinity was identical and thyroid hormone is an important regulator in the number of cardiac beta adrenergic receptors. The increment in the number of beta-adrenergic receptors caused by thyroid hormone is highly statistically significant and is likely to be physiologically relevant. [74]

Over and above as mentioned in structure differences,  $\beta_3$ Ar does not demonstrate agonist-dependent phosphorylation. [64] After long term agonist exposure the number of beta adrenoceptors is decreased. [75] Measurements showed that after long-term agonist exposure

net cellular expression of  $\beta_2$ Ar becomes decreased regardless of localization of beta adrenoceptors.  $\beta_3$ Ar number is not affected via this exposure most probably due to thyroid hormone regulation on  $\beta_3$ Ar function, and lacking cysteine and threonine residues on cytoplasmic tail which causing agonist dependent post-translational modification and PKA phosphorylation.



### 3.1 Potential Mean Force Analysis of Beta-2-Adrenergic at the G-Protein Binding Site

In 1  $\mu$ s simulation of  $\beta_2$ Ar it has been observed that the intracellular loop 3 (ICL3) region creates numerous hydrogen bonds and transform into a closed form. Open and closed forms of ICL3 are one of main indicators of  $\beta_2$ Ar active and inactive states. In this study, we performed Steered MD simulations to estimate required energy for transition of ICL3 from closed form to open form.

Steered MD simulations are giving opportunity to understand the interaction of an area of interest both qualitatively and quantitatively. In steered MD simulations an atom is chosen for applying force and force is applied to this atom via a spring attaching a real and a dummy atom, then force is applied in a vectorial direction.

#### Steered MD

In steered MD simulations a force is applied on one or a group of atoms (SMD atoms) whereas one or a group of atoms are kept in a fixed position in cartesian space and the force vector is calculated accordingly to the vector between steered MD atom and fixed atom. Steered MD simulations can be performed either in constant force or constant velocity. SMD method merely mimics Atomic Force Microscopy technique and pulls smd atoms accordingly.

$$\vec{F} = -\nabla U$$

$\vec{F}$  : force

$\nabla U$  : potential gradient

$$U = \frac{1}{2}k[v t - (\vec{r} - \vec{r}_0) \cdot \vec{n}]^2$$

$k$  : spring constant

$\vec{n}$  : pulling direction vector

$\vec{r}_0$  : center of mass of smd atom at initial position

$\vec{r}$  : center of mass of smd atom after force applied

$v$  : velocity

$t$  : time

In AFM experiments springs attached to consol have spring constants of 1 pN/Å, in steered MD simulations due to time constraints attached springs have spring constants at 70 – 600 pN/Å. Due to fast pulling acquired work values are irreversible work values. In order to calculate reversible free energy value Jarzinsky inequality is used. [76]

### Jarzynski Inequality

Jarzynski inequality shows that free energy of the system from state A to state B can be calculated from finite-time work measurements. In order to be able to calculate Gibbs free energy hence reversible work from finite-time measurements, the transition from state A to state B should occur in one parameter space.

In reactions occurring reversibly, free energy of the system is equal to work from the transition of the system from state A to state B.

$$W = H^B - H^A$$

W: work;  $H^B$ : free energy at state B;  $H^A$ : free energy at state A

In finite-time transitions which occur irreversibly, work is greater or equal to the free energy of the system.

$$\bar{W} \geq \Delta H$$

Jarzynski inequality shows that the difference between reversible and irreversible work is the dissipated work and whether the dissipated work originated from irreversible work can be overcome by averaging finite-time work values over exponential. Jarzynski inequality states that the difference between work values should be in the range of a few  $k_B T$ . [76]

$$\overline{\exp(-\beta W)} = \exp(-\beta \Delta H)$$

If work values acquired are more than a few  $k_B T$ , PMF potential should be reorganized and Jarzynski inequality could be used for calculation of free energy.[77]

In this work, *Gaussian Drift Method* [77] is followed and a relatively higher force constant is used and with known force constant assumption and the dissipated work formed due to friction is subtracted from the force applied.

$$U(x) = U(0) + \int_0^x dx' (F - \gamma v)$$

Free energy values are calculated according to the new force values formed after subtraction of friction effect and using Jarzynski Inequality.

ICL3 of  $\beta_2$ AR is stretched via pulling steered MD atom to the distance of  $\beta_2$ Ar in active state with constant velocity method, in smd simulations last frame of 8 Å restrained 500 ns simulation is used as initial condition. Following simulations are performed and corresponding free energies are calculated.

1. 0.5 Å/ns 5.5 Å
2. 2.4 Å/ns 13 Å SMD

### 3.1.1 Steered MD Simulations

In the simulations performed, Particle Mesh Ewald periodic boundary conditions are used, 1 fs is chosen as timestep, spring constant is chosen as 417 pN/A and SMD output is collected at every 100 timestep.

In all simulations besides smd constraints an 8 Å constraint is subjected to binding site of  $\beta_2$ Ar in order to mimic agonist binding. Restrained residues are presented in Table 3.1.1.1.

<i>Residue</i>	<i>Atom</i>	<i>Residue</i>	<i>Atom</i>	<i>k</i>	<i>Constraint</i>
SER 203	H- $\beta_2$	ASP 113	H- $\beta_2$	50	8
SER 207	H- $\beta_2$	ASP 113	H- $\beta_2$	50	8
SER 204	H- $\beta_2$	ASP 113	H- $\beta_2$	50	10
ASN 293	H- $\alpha$	ASP 113	H- $\beta_2$	60	8
PHE 289	H- $\alpha$	ASP 113	H- $\beta_2$	60	8
ASN 312	H- $\alpha$	ASP 113	H- $\beta_2$	60	8
PHE 289	H- $\alpha$	ASN 312	H- $\alpha$	50	8

Table 3.1.1.1 Restrained atoms and their positions.

#### 3.1.1.1 $\beta_2$ -Adrenergic Intracellular Loop 3 Unbinding Energy Calculations

For 8 Angstrom Restrained 500 ns simulation system, last frame of simulation where protein is mimicking an inactive state is taken and equilibrated for 2 nanoseconds where protein is fixed onto simulation space through 50 GLY, 76 ALA and 127 ILE. System is prepared in this way to lower degree of freedom in energy space in order to be able to account for only the change of intracellular loop three (ILC3) conformational energy.

During Steered MD simulations an 8 Å constraint is performed at ligand binding pocket in order to overcome antagonist effects might happen during unbinding of ICL3 from inactive state.

Simulations are performed in 1 femtosecond time step with langevin dynamics on at constant volume. SMD vector is calculated as in NAMD Tutorial Steered MD Constant Velocity Pulling Section.

Trajectory file and restart files are collected in every 5000 steps. Each simulation is lasted for 5.4 ns equivalent of 13 Angstrom pulling.

Indication of steered MD atoms, fixed atoms and constraints are used in presented in Appendix I.

Force and distance trajectories are gathered from simulation data using VMD. Force and Work values are calculated as it is described in Appendix II. PMF values are calculated accordingly to Stretching Deca-Alanine Tutorial.

In the following section findings of 0.5 Å/ns simulations and 2.4 Å/ns simulations are summarized. Force and Work profiles and respective PMF calculations are presented. It is seen that when pulling speed is increased force exerted on smd atom also increases, however catching a Gaussian distribution of work values acquired is more prominent than lower force and work values. Due to this histogram analysis of work values acquired is calculated.

### 3.1.2 0.5 A/ns Pulling

Each simulation is lasted for 11 ns and 81.49 kcal/mol standard deviation of work values is acquired as shown in Table 3.1.2.1.

Table 3.1.2.1 Work values of 10 trajectories

Simulation Number	Work (kcal/mol)
1	232.28
2	27.73
3	264.11
4	150.01
5	43.11
6	218.35
7	198.09
8	174.29
9	140.73
10	249.42

Force vs extension relations are shown in Figure 3.1.2.1 where different colors correspond to the 10 different simulations.

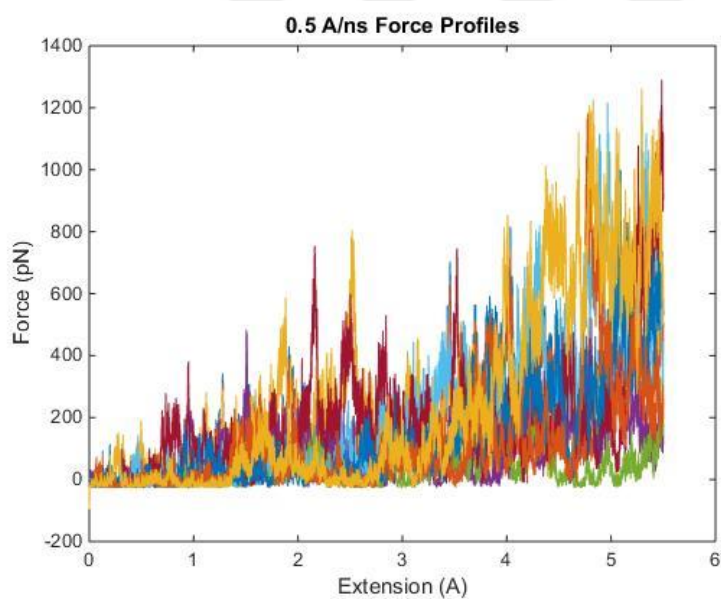


Figure 3.1.2.1 Force profiles of 10 trajectories

Work vs extension relations are shown in Figure 3.1.2.2 where different colors correspond to the 10 different simulations

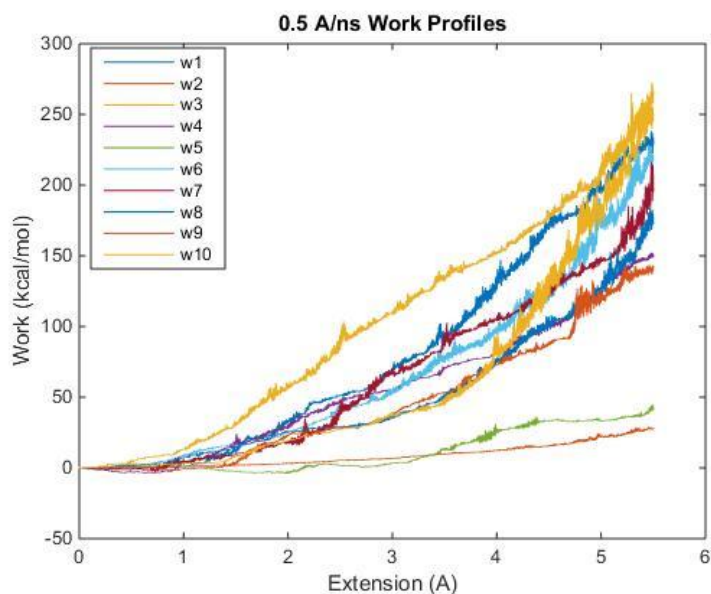


Figure 3.1.2.2 Work profiles of 10 trajectories

### 3.1.3 2.4 A/ns Pulling

Each simulation is lasted for 5.4 ns and 23.91 kcal/mol standard deviation is acquired.

Table 3.1.3.1 Work values of 20 trajectories

Simulation Number	Work (kcal/mol)
1	392.85
2	430.35
3	412.29
4	426.92
5	439.03
6	420.16
7	500.58
8	417.65
9	425.22
10	436.11
11	456.34
12	432.99
13	453.99
14	397.96
15	415.71
16	419.03
17	457.38
18	419.09
19	413.41
20	426.46

Force vs extension relations are shown in Figure 3.1.3.1 where different colors correspond to the 20 different simulations.



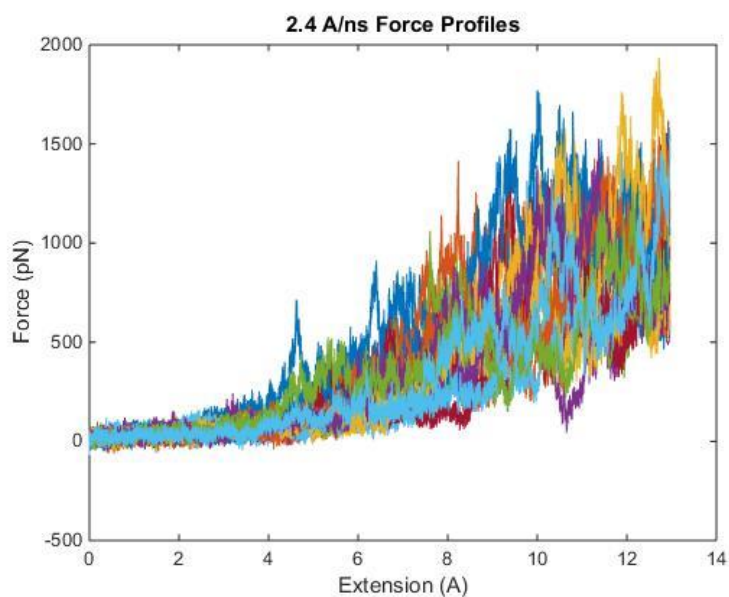


Figure 3.1.3.1 Force profiles of 20 trajectories

Work vs extension relations are shown in Figure 3.1.3.2 where different colors correspond to the 20 different simulations.

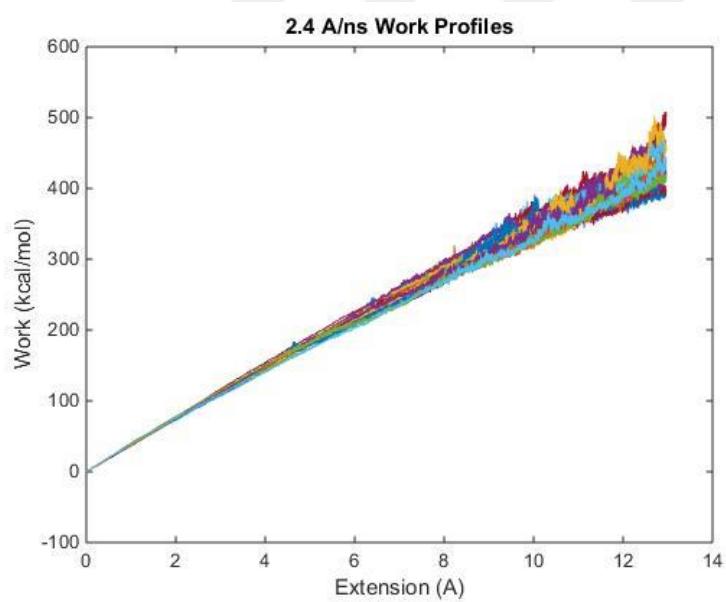


Figure 3.1.3.2 Work profiles of 20 trajectories

Potential mean force calculated is presented in Figure 3.1.3.3 out of 20 trajectories.

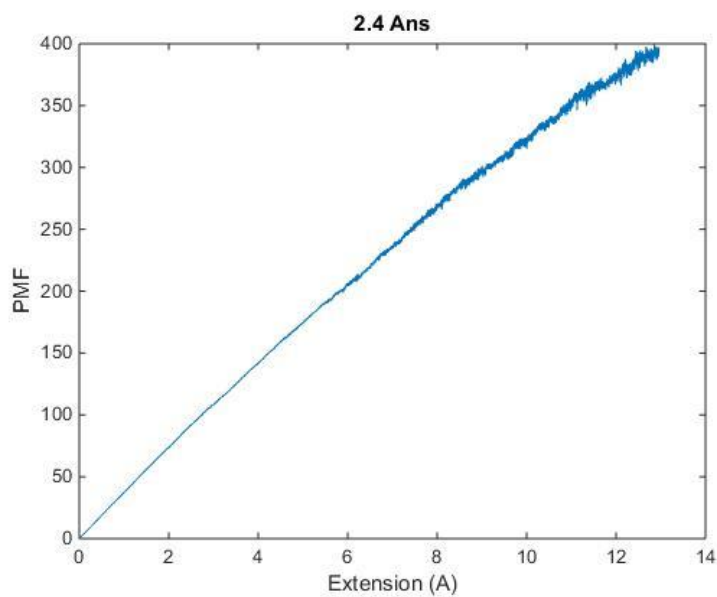


Figure 3.1.3.3 Potential Mean Force profiles of 2.4 Å/ns pulling

Work values acquired is subjected to a Gaussian fit whether chosen velocity is prominent for application of Jarzinsky inequality. In gaussian fit of 0.5 A/ns work values does not give a Gaussian distribution. In 2.4 A/ns velocity measurement required Gaussian fit is achieved.

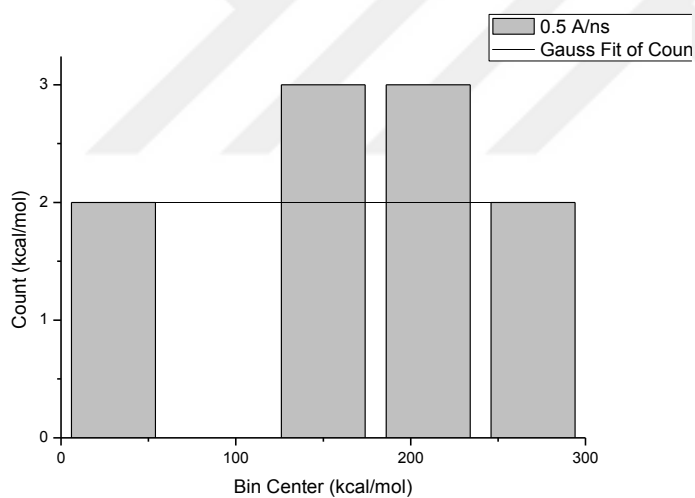


Figure 3.1.3.4 Work distribution of 0.5 Å/ns pulling over 10 trajectories

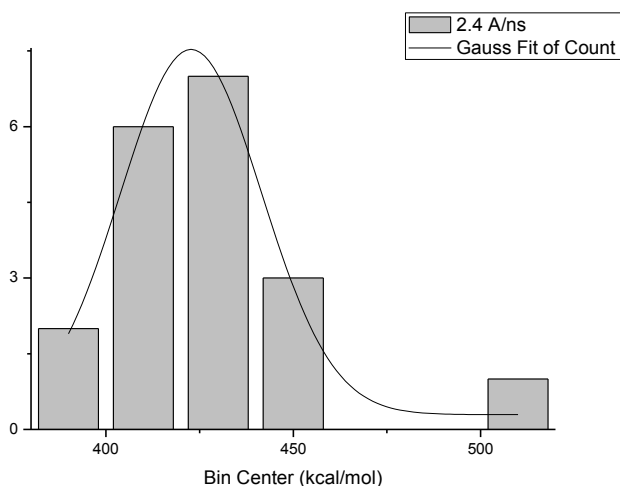


Figure 3.1.3.5 Work distribution of 2.4 Å/ns over 20 trajectories

Performed simulations show that right value of velocity has to be found via trial and error. Jarzinsky states that acquired work values should be in range of few  $k_B T$ 's and if not possible acquired work values should be in Gaussian distribution.

Jarzinsky states that acquired work values should be in range of  $k_B T$ 's regardless of the pulling speed. Second rule of thumb to follow Jarzinsky equality is that acquired work values should have a Gaussian distribution. Therefore right value of velocity has to be found via trial and error.

In deletion studies it has been seen that ICL3 region is an essential for Gs coupling. [78] Work required for unbinding of ICL3 region, hence gibbs free energy of G-protein binding is calculated as 389.3 kcal/mol, 1628.8312 kJ/mol. [79] In the work of Straßer and Whittman they have conducted Molecular Dynamics simulations and calculated  $\beta_2\text{Ar-G}\alpha$  interaction energy as -1455 kJ/mol. Our findings are in good agreement with their study.

## 4.1 Allosteric Couplings of Beta-2-Adrenergic

### 4.1.1 What is allostery?

The process by which signals originating at one site in a protein propagate reliably to affect distant functional sites is called allosteric communication. A small subset of residues forms physically connected networks that link distant functional sites in the tertiary structure. In this perspective, allostery represents information transfer to a distinct surface mediating downstream signaling. [80] There is a two-way transfer of information that communicates the state and needs of the cell to the extracellular space. They do this through a change in shape, specifically referred to as a change in conformation. [81] Throughout the years, allostery has come to encircle different yet in a way related mechanisms by which protein function is regulated. It is developed initially in the field of enzymology, is that many proteins possess more than one binding site as later characterized as orthosteric sites and allosteric sites. [82] One explanation is that multiple binding sites leading to multiple conformations. In which those conformations, the bottom of the energy wells is probably “rugged” allowing a range of nearly iso-energetic conformers; the more flexible is the protein, the larger ensemble of conformers it produces. [83] Allosteric sites are the sites conformationally linked such that the inhibition/activation could be transmitted from the effector site to receiving site. [84] In the work of Kenakin and Miller, allosterism is defined in terms of three interacting species [81]:

- The modulator : a ligand or protein that binds to a conduit
- A conduit that transduces the thermodynamic allosteric energy to a guest
- A guest which receives the influence of the modulator through a conduit.

Many events can serve as modulators: binding of a molecule, a change in pH, temperature, ionic strength or concentration, a covalent modification such as tethering glycosylation, phosphorylation and ubiquitination. [85] A key manifestation of allosteric interaction is an alteration in the affinity of an orthosteric ligand for its binding site. [82] Homo or heterodimeric ligand-binding domains of several nuclear receptors display different conformations in the agonist-bound versus antagonist bound states, with a few characteristic changes in helix structure within largely common protein fold structure. [86]

There must be a balance between thermodynamic stability to support specialty and flexibility to mediate conformational change to catalyze biochemical reaction pathways. [81] There are many allosteric theories and methods which try to explain allosteric modulation in network based view and ligand-activation based view. Two of main arguments having its roots from active and inactive states and referred as two-state models are conformational induction and conformational selection theories. Conformational induction is a concept proposed by Koshland (1958). [87] In this perspective molecule contributes energy to cause a change in the conformation of the receptor whereas conformational selection is rooted in the “population shift” model emanating from the MWC model of allostery. In conformational induction scenario; the ligand binds to the low energy inactive state of the receptor to cause in protein to the active state molecular simulations favor conformational shift models for the binding of small molecules. In conformational selection hypothesis receptor scans both active-like and

inactive-like conformations with different probabilities. In literature, it has been argued that receptor pre-existing in an active-like state without a ligand bound found as a scarce probability.

As another perspective, rather than two-state models as MWC and KNF models, there is energetic “hot wire” idea to understand allosteric couplings, in this scope the coupling between sites depends upon the intrinsic stabilities of the domains and the interactions between them which in turn depend upon probability distributions resulting from the conformational changes within the receptor protein.

Carminé and his coworkers measured the binding affinities of six catecholamines differing only in ethanolamine tail and performed double mutant cycle analysis at positions 203,204 and 207 of  $\beta_2\text{ar}$ . In their study, they have performed 33 double mutant cycles and estimated free energy coupling values and found out that replacements of serines (SER) with alanine (ALA) or cysteine (CYS) resulted in similar losses in energy whereas threonine replacement caused different effect relying on different positions. Threonine replacement in position 207 showed a conservative behavior as effecting a very small change whereas in position 204 they have measured equal losses of alanine and cysteine. Finally, with the ligand mutations they have concluded that during ligand binding to the aforesaid binding pocket is an interactive occasion rather than rigid on-off signal.

To depict allostery there are several experimental methods for protein dynamics. X-ray crystallography, Nuclear Magnetic Resonance, Fluorescence Microscopy, Hydrogen-Deuterium Exchange Mass Spectrometry and Atomic Force Microscopy are few of them. [85]

### 4.1.2 Two-state Allosteric Models

The main models that are used to describe the mechanism of two-state allostery is the Monod-Wyman-Changeux (MWC) model and the Koshland-Nemethy-Filmer (KNF) model. In MWC model, it is assumed that, in the absence of the effector, the protein samples both inactive/tense (T) and active/relaxed (R) states, but with different probabilities. The R state has a higher free-energy and therefore sampled less, but the appearance of the allosteric effector lowers the free energy of R state enough to lock protein in it. [86]

The key statements of MWC model are that allosteric proteins are oligomeric, they possess an axis of symmetry, they can exist in an equilibrium between (at least) two distinct states in the absence of ligand, and they possess multiple ligand recognition sites binding to which stabilizes a subset of conformational states at the expense of others. MWC predicts that it should be possible not only orthosteric but allosteric drugs that preferentially favor either active or inactive receptor states and can selectively modulate properties of co-bound ligands in a manner that correlates with the efficacy (positive, negative or neutral) of such ligands. [89]

Orthosteric site is a distinctive binding site for respective endogenous ligand in receptors. Binding and activation by an orthosteric allosteric agonist conclusively results in a structural rearrangement of the receptor that results in increased affinity for G-proteins. [90]

Allosteric sites are spatially and functionally distinct sites which modulate orthosteric ligand binding by shifting equilibrium increasing or decreasing affinity or efficacy for orthosteric ligand without activating the receptor. Such ligands are habitually called as positive allosteric modulators (PAMs) and negative allosteric modulators (NAMs). Many GPCRs possess allosteric binding sites. [83,91,92]

Ternary complex model (TCM) is an expansion of simple two-state model which takes into account the interaction between a ligand and its receptor by including G-protein role in activation and giving rise to a four point 2-D model, there are 16-point quaternary complex models as well. Still, such models take only active (R\*) and inactive (R) state into account. [91]

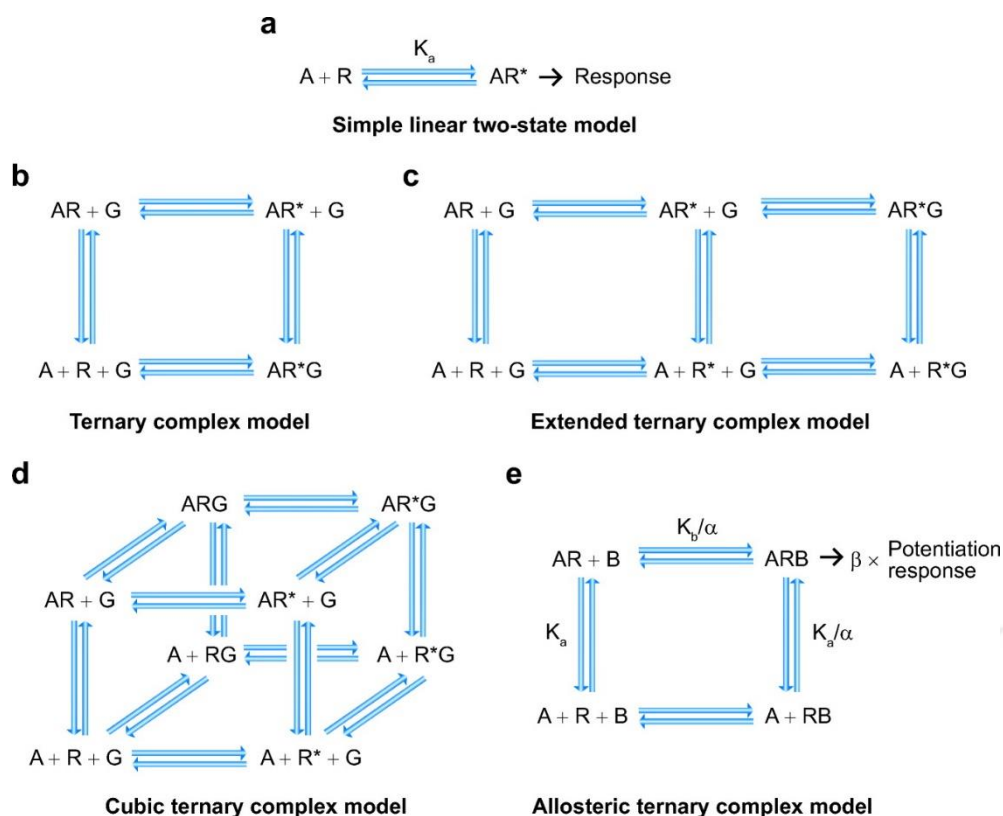


Figure 4.1.2.1 Models of GPCR ligand–receptor interactions. a) A simple linear two-state model employing a single dissociation constant governed by the law of mass action. b) The initial ternary complex model for ligand, receptor, and G-protein interactions. c) The extended ternary complex model, which includes the scenario of constitutive or agonist-independent receptor activation. d) The cubic ternary complex model, which adds the possibility of active receptor and G-protein association that does not causing signaling. e) The allosteric ternary complex model, a concise framework for modeling the interaction of two ligands (*e.g.*, an orthosteric agonist and an allosteric modulator) on a receptor, taking into account both cooperativity between the ligands ( $\alpha$ ) and the effect of one ligand on the other's efficacy ( $\beta$ ).

In Allosteric Ternary Complex Model (ATCM), orthosteric and allosteric ligands bind reversibly and saturably to their respective binding sites. [83]

TCM can result in shallow and biphasic binding curves if G-protein amount is limiting, whereas ATCM does not because exogenous allosteric modulator ligands are present in vast excess relatively to the concentration of the receptor. Additionally, the ATCM does not differentiate among ligands that do or do not possess efficacy whereas the cooperativity factor in the G-protein TCM is an index of orthosteric ligand efficacy. [94]

Allosteric two state model (ATSM) describes the interaction between an allosteric modulator with an orthosteric ligand on a receptor that is isomerize between inactive (R) and active (R\*) states. [95]

In KNF model, the protein only visits T state in the absence of the allosteric effector. It is the appearance of the effector and interaction with the protein what makes protein undergo a conformational change into R state.

MWC model involves conformational changes that require the collective motions of many atoms at the same time, while in KNF model transition from the T state to R state is more sequential. [86]

Fluorescence studies of beta-2-adrenergic activation kinetics suggests that conformational change occurs after ligand binding which may leading to an 'induced fit' mechanism. [96]

In addition to PAMs and NAMs, allosteric binding sites also allow other novel modes of receptor modulation. An ago-allosteric modulator is defined as allosteric ligand which displaying both allosteric agonism and an ability to allosterically modulate the binding and/or function of orthosteric ligands from allosteric ligands that display only agonism or modulation. [83]

Enhanced potency might arise from modulator-induced enhancement of agonist binding affinity. Stabilization of the agonist bound sensing module, for example, might lead to a reduced rate of agonist dissociation. Provided that the time spent by the effector module in its active conformation remains unchanged and that signal transmission from the sending unit is intact, the effect of the modulator would be confined to an effect on potency.

Enhanced efficacy implies a conformation state in which the effector unit interacts with a larger number of signaling molecules, e.g. by remaining active for longer upon receipt of the intramolecular message from the signal transmission unit or by enhancing the rate of access of signaling molecules to the effector site. [97]

“Pure” positive or negative allosteric modulators which are typically sought after in most allosteric drug discovery programs are unlikely to exist. The degree of modulation will depend on coupling efficiency, such that amplified responses will appear to be modulated to a greater extent than less amplified responses. [89]

A major controversy of GPCR agonist drugs that they require chronic administration and leading receptors to desensitization, down regulation and/or internalization over time. Efficiently turning receptor to active state via agonist administration is stated as unnatural and causing receptor to lose sensitivity. [91]



### 4.1.3 Computational Approaches

It is needed to note that there is no approach that has been achieved consensus in order to depict and understand allostery. Experimental approaches in this scope of the study is usually used for understanding two-state model allostery and remaining transition or not-recognized states and conformations is still under investigation. Some of the methods in allostery identification can be listed as non-equilibrium molecular Dynamics simulations, procedure of anisotropic thermal diffusion, evolutionary trace analysis and normal mode/principal component analysis. [98, 99] Used techniques can be divided into two categories as feature-based models and structural surveys. In feature-based methods the scope is mostly upon similar characteristic properties seen in similar functions. Thus, investigation of Hotspots in a protein, Atomic density, solvent accessibility, Hydrophobicity, Noncanonical hydrogen bonds and  $\Pi$ -interactions can be listed as examples of this type of analysis. [85]

#### 4.1.3.1 Structural Surveys

##### 4.1.3.1.1 Statistical Coupling Analysis

Statistical Coupling Analysis developed by Lockless and Raganathan mines large databases for homologous proteins and compares their sequences in order to identify evolutionary covarying residues. [100] Given its power on the availability of large databases it is unlikely that SCA will be able to identify novel allosteric networks between orthosteric and small-molecule allosteric site specific to individual GPCRs. [83]

##### 4.1.3.1.2 Principal Component Analysis

Principal Component Analysis, the covariance matrices are diagonalized to yield collective variables (eigenvectors) sorted according to their contribution to the total mean-squared fluctuation.

MD trajectory is projected onto principal components (collective variables) and it has been found out that coordinates constitute a low dimensional sub-space that represents a significant part of the macromolecular motion. [101, 102]

##### 4.1.3.1.3 Elastic Network Models

Elastic Network Models (ENM) where macromolecule is defined as a set of nodes connected by harmonic springs. Normal modes / collective coordinates are then derived from this simplistic harmonic approximation of the potential energy function near equilibrium. [103]

ENM based methods suffer from two major limitations.

- By either treating the macromolecule in isolation, or using implicit solvent schemes to screen interatomic interactions, the methods likely fail to account for configurational entropy of the surrounding environment.
- Given ruggedness of the potential energy surface and an-harmonic of the energy barriers for conformational transitions, harmonic approximations tend to lose critical information relevant for characterizing macromolecular conformational behavior. [104, 105]

### **4.1.3.2 Comparisons within conformational ensembles**

A different approach modeling allosteric signal propagation using a protein structure network has been introduced by Park and Kim in signal transmission through the network is modeled using a Markov random walk. This modeling approach is limited in its ability to address dynamic, large scale changes in structure. [106]

#### ***4.1.3.2.1 Monte-Carlo simulation***

Doruker et.al coupled MC simulation with anisotropic network modeling to study transition pathways in two-well known allosteric protein. E.coli adenyl cyclase and human hemoglobin. Their iterative approach is comprised of alternating cycles of normal mode analysis and MC simulation applied to pre-transition structures, resulting in post-transition conformations. In each cycle, normal modes are calculated for starting structures and the slow mode overlapping the desired transition is selected and a new structure is generated by deforming the protein along the slow mode. [107]

#### ***4.1.3.2.2 Coarse grained molecular dynamics***

It is also known as reduced representation approach. Okazaki and Takada have used the Gō-like single-bead approach to represent glutamate-binding protein in a comparative study to protein systems to see which are best represented. [108]

Among structural surveys if enough time and memory is present Molecular dynamics simulations, targeted MD simulations can also be used in the investigation of allostery or above-mentioned methods can be used in combination of MD simulations accordingly to the problem set up.

#### 4.1.4 Important Interactions

GPCRs fundamental nature requires extracellular ligand binding to result in a dynamic change in receptor conformation that is reflected in exposure of a signaling domain at the cytosolic surface, which interacts with the effector partner, a heterotrimeric G-protein. [81]

Allosteric ligands provide a new energy landscape for the receptor which can essentially create a new receptor. Allosteric ligands have unique properties in that they produce saturable and probe-dependent effects; these can lead to correspondingly special therapeutic properties [81] Thus it is important to understand how allosteric ligands effect the landscape of energy when receptor has come upon with a signaling event and internal signaling pathways of receptor.

Carmine et.al states that aromatic residues residing in TM6 may be involved in catechol ring of ligand during binding of catecholamines through  $\pi$ - $\pi$  stacking as it is observed in Phe 290 near toggle switch. Their another statement is that tilting motion at intracellular region of TM3 and TM6 which occurs during G-protein binding might have been causing upper segments of other helices to get closer towards each other and enhancing catecholamine binding with residues at binding site. [88]

In the work of Lakkaraju et.al. it was found that the apo form of  $\beta_2$ Ar and carazolol bound  $\beta_2$ Ar were assigned to distinct “inactive-like” conformations in which the cavity of Gas binding site is not formed. They have pointed out that agonists of  $\beta_2$ Ar preserved G-protein binding site cavity which is a little bit smaller than the cavity found out in crystal structures. In agonists, they have found out that BI-167107 bound  $\beta_2$ Ar is sampled active-like state more frequently than epinephrine and salbutamol. They have found out that Salbutamol is keener to inactive-states indicating another binding partner is required for required modulation. [109]

In experimental studies it has been observed that amino group of catecholamines form a salt bridge with Asp113 whereas polar groups of catecholamine interact with Ser203 and Ser204 in TM5. [48, 55]

Usually  $\beta$ -OH moiety of ligands anchor between Asp113 and Ans312. [36] Carvedilol known as an antagonist makes van der Waals in  $\beta_1$ Ar with Leu101 (TM2), Asp200 (ECL2), Tyr207 (ECL2), Trp330 (TM7) and Phe201 (ECL2). [40]

In the work of Lakkaraju et.al. where they have been compared different effectors they had found that in Apo versus carazolol bound simulations residues which have shown biggest conformational changes were Arg151, Phe166, Ile169, Gln170, Gln197, Leu275 and Cys327. In Apo- BI-167107 comparison they have reported Glu122, Arg151, Trp158, Thr164 and Tyr326. In epinephrine binding Gly35, Met36, Ile38, Leu42, Phe208, Phe217, Tyr219, Leu275, Ile277, Phe282, Asn318, Tyr326 and Arg328 were showing biggest conformational change. In salbutamol binding pointed out residues were Gly35, Phe49, Phe166, Gln170, Gln197, Val206, Phe217, Ile278, His296, Ile298, Glu306, Leu310, Ile325 and Arg328. [109]

Dimerization of GPCRs is another subject in discussion. There is increasing evidence that dimerization is a requisite for function in some GPCRs. [110] Changes in receptor sensitivity to agonist upon dimerization has been worked on and characterized for opioid receptor heterodimers upon now. [81]

## 4.1.5 METHOD

### Production of Correlational Deviation Matrices (Covariance Matrices)

In correlation analysis which are performed, average positions of alpha carbon atoms are represented accordingly to cumulative change that has been observed in the targeted time frame. Every position is divided to total station number in order to get mean positions of backbone.

$$\bar{X} = \frac{(\sum_{t=0}^N X_1(t))}{N} \quad \bar{Y} = \frac{(\sum_{t=0}^N Y_1(t))}{N} \quad \bar{Z} = \frac{(\sum_{t=0}^N Z_1(t))}{N}$$

For each carbon atom, deviation from mean value is accounted for each station and average correlational deviation values are calculated. This will be giving the average correlational deviation for each residue during whole simulation.

$$\langle \Delta R_i^2 \rangle = \left( \sum_{t=0}^N (\Delta X_{ij}^2 + \Delta Y_{ij}^2 + \Delta Z_{ij}^2) \right) / N$$

$$\Delta X_j = X_j - \bar{X}, \Delta Y_j = Y_j - \bar{Y}, \Delta Z_j = Z_j - \bar{Z}, i: \text{residue index}, j: \text{station index}$$

Average correlational deviation can be used to calculate temperature factors of each residue using the relation below, this will be resulting in detecting rigid and flexible regions of protein.

$$B - \text{factor} = \langle \Delta R_i^2 \rangle \left( \frac{298}{310} \right) \left( \frac{8\pi^2}{3} \right)$$

Beta factor represents the relative harmonical movements of atoms and it is defined in X-ray crystallography studies for every protein.

A correlational deviation of a residue pair can be obtained through scalar multiplication of each residues correlational deviations of that station.  $C_{ij}$ , correlational deviation of residue pair results in a value between +1 and -1 due to final division to average deviation which is performed to get a normalized value. If there is no correlation between residue pair  $C_{ij}$  value will be resulting as zero.

$$C_{ij} = \left\langle \frac{\langle \Delta R_i \rangle}{|\Delta R_i|} \cdot \frac{\langle \Delta R_j \rangle}{|\Delta R_j|} \right\rangle$$

Correlational deviation values will be useful to detect the structural difference of a protein in the presence of two different effectors. These could be allosteric or intrinsic effectors upon analyzer's choice.

Matlab code which is implemented for this method to calculate  $C_{ij}$  values for a definite time scale is presented in Appendix III.

After gathering  $C_{ij}$  values mostly correlated residues are depicted for distance analyses. Residue pairs are chosen according to their branching number as involving multiple sites of proteins.

## Important Correlations

### 4.1.6 1000 ns APO $\beta$ 2-AR

It is seen that ICL1 is correlated with lower end of TM6-ICL3 region residues Cys265, Glu268 and His269. In inactive conformation ICL1 and lower end of TM6 comes closer to lower end of TM2 as seen in Figure 4.1.6.1 Gln65-Cys265 distance becomes 10 Å narrower during 600 – 1000 ns. Glu268 is speculated to be involved in G-protein coupling along with Leu64 and Asp130 which are located on the intracellular side of receptor and. [111] These residues could be important in G-protein recognition as underneath of TM6 and TM2 becomes tighter in distance and thus signals an inactive state due to absence of G-protein.

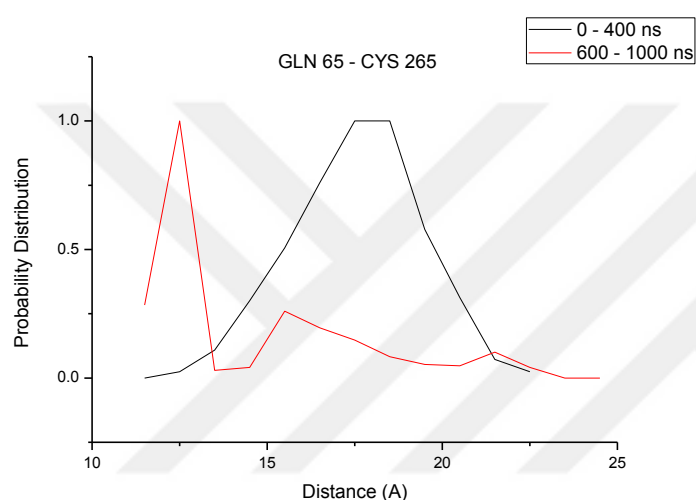


Figure 4.1.6.1 Probability distribution of distance between Gln65 - Cys265. The distance between Gln65 and Cys265 decreases in 600 – 1000 ns. Gln65 – Cys265 distance is an indicator of behavior of intracellular parts of TM2 and TM6.

Another residue of ICL1, Thr 66 has important correlations with Ile 154 of TM4 and Thr 123 of TM3. (Figure 4.1.6.2, Figure 4.1.6.3) In these correlations lower end of TM2 forms stable distance with TM4 and reaches further away from TM3. Ile154 is known for  $\beta$ <sub>2</sub>Ar interaction with cholesterol rings A and B [113]. Ile154 keeps its distance with Leu64 in a stable manner with Thr66 (Figure 4.1.6.2) whereas Tyr123-Thr66 distance increase about 8 Å. This may be indicating that  $\beta$ <sub>2</sub>Ar follows a two sided control mechanisms in order to position in fully active state. Packing of ICL3 underneath TM5 and TM6 makes protein to adopt a new conformation which locks some of the important residues in locked signaling states.

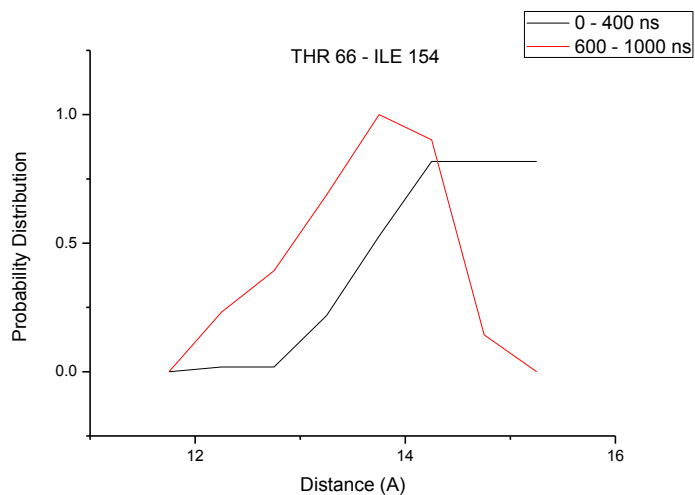


Figure 4.1.6.2 Probability distribution of distance between intracellular end of TM2 (Thr66) and membrane region of TM4 (Ile154)

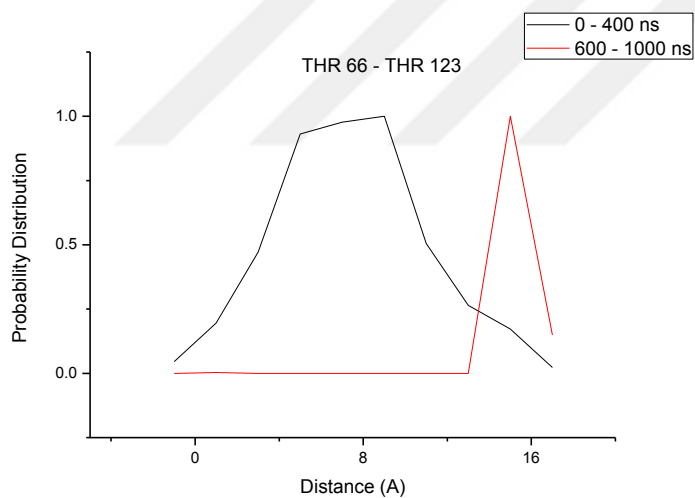


Figure 4.1.6.3 Probability distribution of distance between intracellular end of TM2 (Thr66) and membrane region of TM3 (Thr123)

Trp 99 of ECL1 locks into a specific distance in 600 – 1000 ns with middle portion of TM5 (Figure 4.1.6.4) and Trp99 - Cys191 of ECL2 gets 7 Å wider. Cys 191 is an important residue in folding of protein into its functional conformation. [114] This cluster controls probably rotational degree of transmembrane 5 which may lead to a specific binding cavity formation due to flexibility of two half's of transmembrane 5.

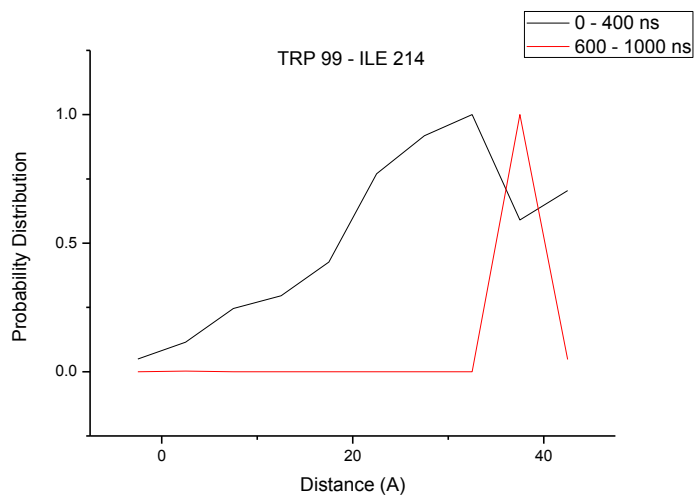


Figure 4.1.6.4 Probability distribution of distance between ECL1 (Trp99) and membrane region of TM5 (Ile214)

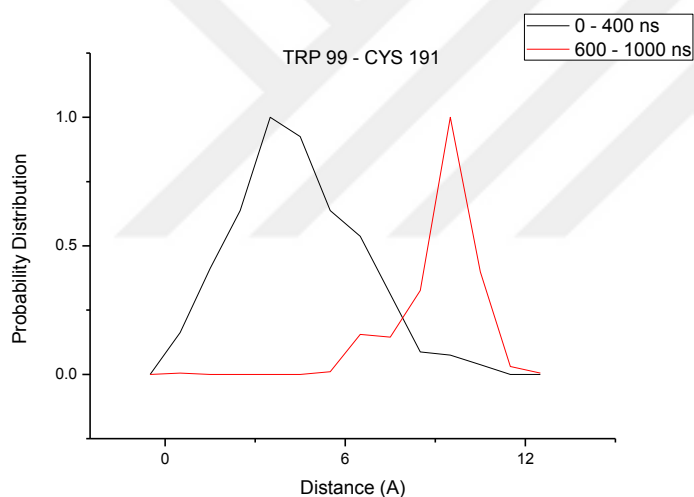


Figure 4.1.6.5 Probability distribution of distance between ECL1 (Trp99) and ECL2 (Cys191)

ECL1 moves apart from both extracellular part of TM6 around 20 Å (Figure 4.1.6.6) and intracellular of TM5 around 25 Å (Figure 4.1.6.7). Ile294 is known for its importance in JDTic receptor selectivity in kappa-opioid receptor. [115] Opioid receptors are functioning in regulation of membrane ionic homeostasis. Intracellular part of TM5 and distance between Phe101 is not stabilizing at 25 Å. (Figure 4.1.6.7) This seems like transmembranes 2 and 3 and transmembranes 5 and 6 is taking apart from each other where Phe101 of ECL1, TM5 and TM6 is opening a wider space on the center of 7 TM structure. Thus it can be speculated that Phe101 interactions are important for selective binding and activation of ionic channels.

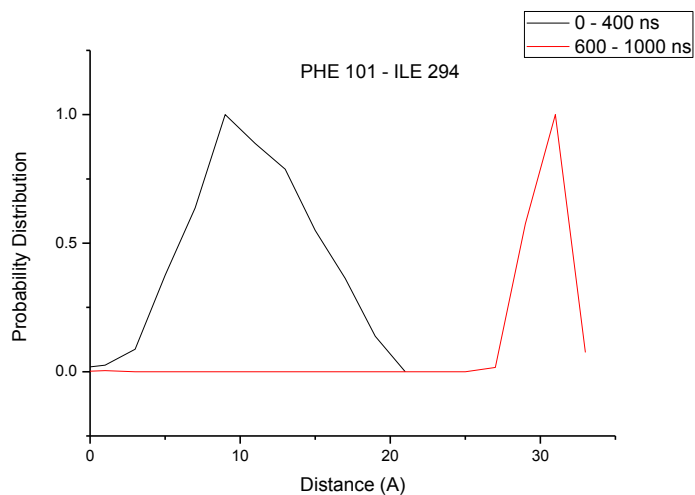


Figure 4.1.6.6 Probability distribution of distance between ECL1 (Phe101) and extracellular end of TM6 (Ile294)

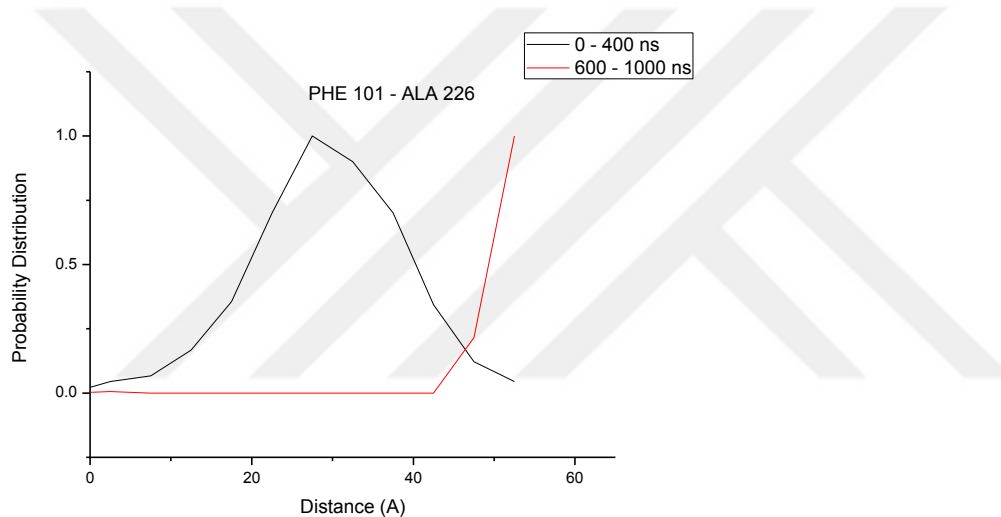


Figure 4.1.6.7 Probability distribution of distance between ECL1 (Phe101) and intracellular end of TM5 (Ala226)

Cys 106 of TM3 - Phe 89 at extracellular part of TM2 shifts into a conformation where they are apart for 10 Å in 600 – 1000 ns. In the work of Liu et.al it is seen that Phe89 is located in an important antagonist binding pocket of CCR1 protein. Usually benzene rings of antagonists forms strong  $\pi$ - $\pi$  stacking with the side chains of Phe89. [122]



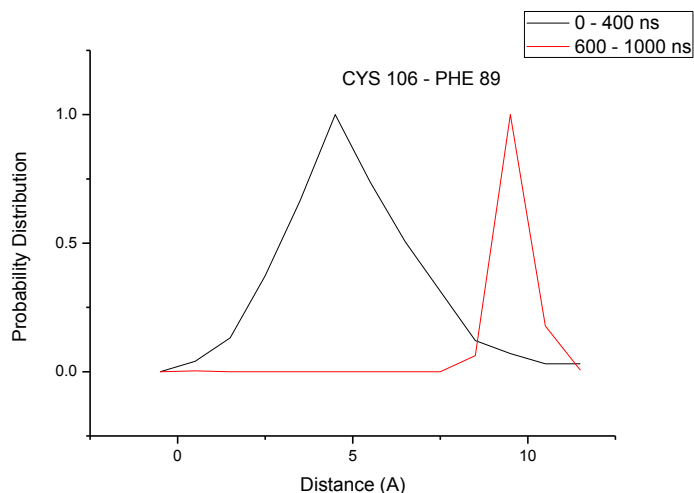


Figure 4.1.6.8 Probability distribution of distance between extracellular end of TM3 (Cys106) and extracellular end of TM2 (Phe89)

Cys 184 of ECL2 and Ile 291 of TM6 in active-like conformation scans around 20 Å area which locks into 30 Å distance in last 400 ns. In the work of Hogan et.al [128] it is postulated that Ile291 resides in a hydrophobic pocket on top of the DPxxY motif and thought to act as a rotamer switch.

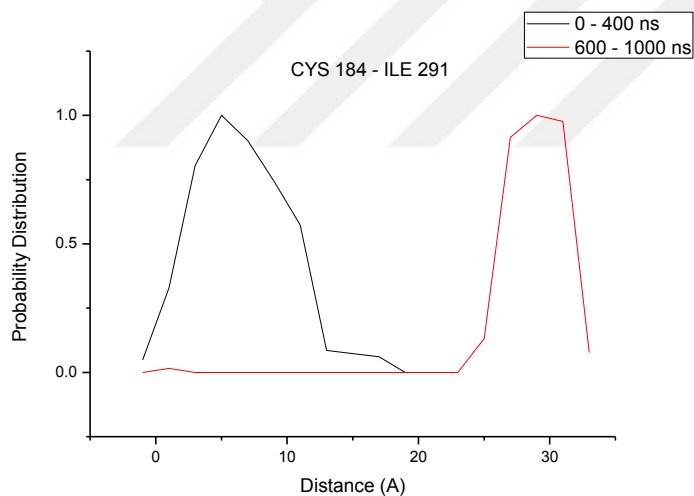


Figure 4.1.6.9 Probability distribution of distance between ECL2 (Cys184) and extracellular end of TM6 (Ile291)

Lower end of TM6 forms correlations with lower end of TM1 and ICL1. Lower end of TM2 and TM6 comes closer whereas TM1 is pushed away 10 Å (Figure 4.1.6.11). Lys267 as an important residue underneath TM5 and TM6 where ICL3 packs around, correlates mainly with ICL1 (Figures 4.1.6.10-12-13) via pushing intracellular end of TM1 which probably causing the energy barrier for activation.

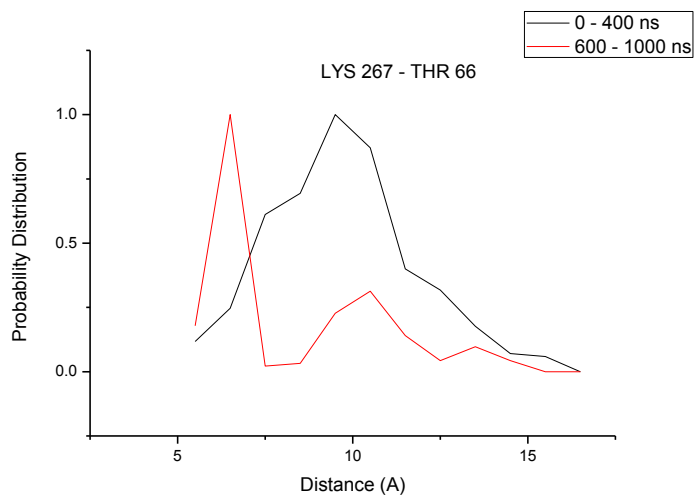


Figure 4.1.6.10 Probability distribution of distance between intracellular part of TM2 (Thr66) and intracellular end of TM6 (Lys267)

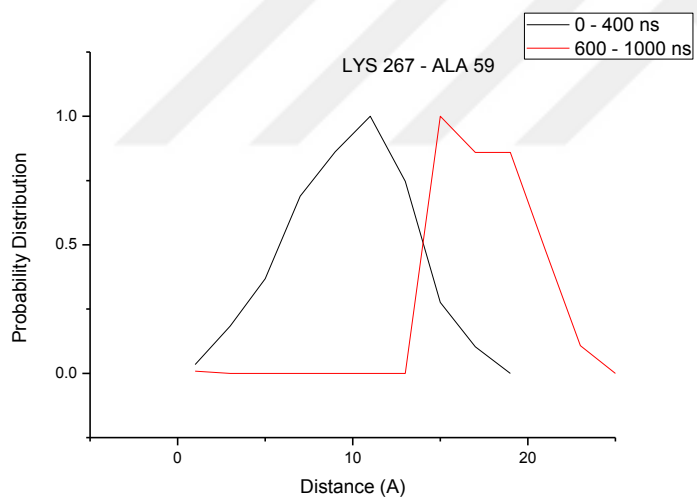


Figure 4.1.6.11 Probability distribution of distance between intracellular part of TM1 (Ala59) and intracellular end of TM6 (Lys267)

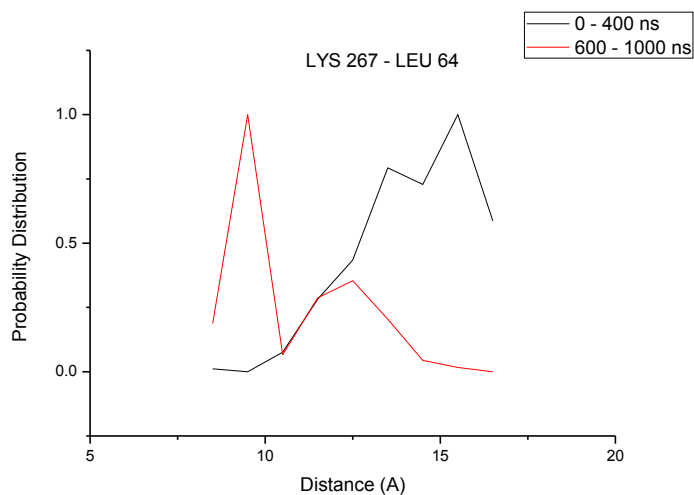


Figure 4.1.6.12 Probability distribution of distance between ICL1 (Leu64) and intracellular end of TM6 (Lys267)

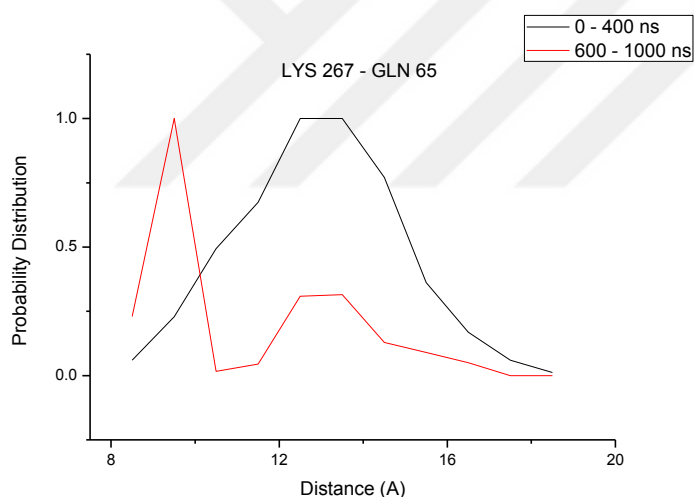


Figure 4.1.6.13 Probability distribution of distance between ICL1 (Gln65) and intracellular end of TM6 (Lys267)

During 1000 ns simulation in the last 400 ns three parts of  $\beta 2A_r$  adopts three distinct movements. Intracellular part ends of TM2, TM6 and ICL1 comes closer, lipid membrane region of TM4 and intracellular end of TM2 spans a 14 Å distance where most probably Ile154 is making stabilizing contacts with cholesterol, and membrane region of TM3 become distant to TM2 and TM4. In extracellular part the distance between ECL1, ECL2, extracellular part of TM6 and intracellular part of TM5 increased as well as the distance between extracellular ends of TM2 and TM3 enabling a wide space for ligand binding, clathrin-cargo signaling or ionic passage.

### 4.1.7 500 ns 8 Å Restrained $\beta$ 2-AR

#### 0 – 100 ns

Phe166 is located in binding pocket of  $\alpha_{2A}$ -adrenoceptor, this trend seen in 400-500 ns of 8 Å restrained simulation shows that TM4-TM2 domain is scanning for ligand binding from extracellular site .[116] (Figure 4.1.7.1.1). Asn69 is reported as N-glycosylation site in human vasoactive intestinal peptide 1 receptor a class II GPCR. [123]

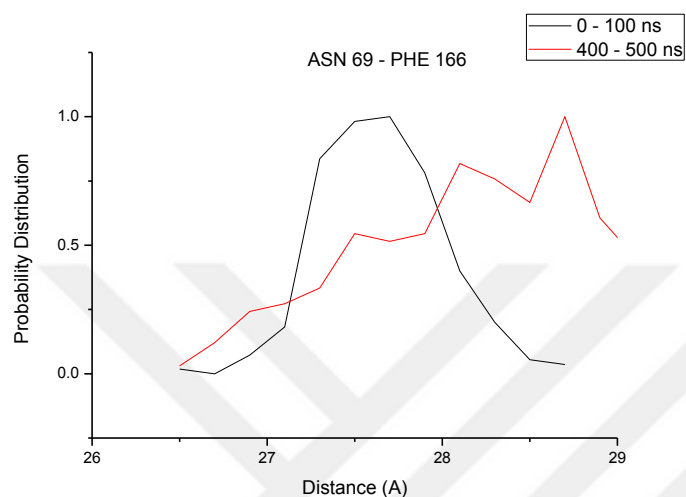


Figure 4.1.7.1.1 Probability distribution of distance between intracellular part of TM2 (Asn69) and extracellular end of TM4 (Phe166)

Lys97 is noted for its importance in cholesterol binding. [124] In last 100 ns the distance between Lys97 and Ala134 resonates between 49 Å and 52 Å.

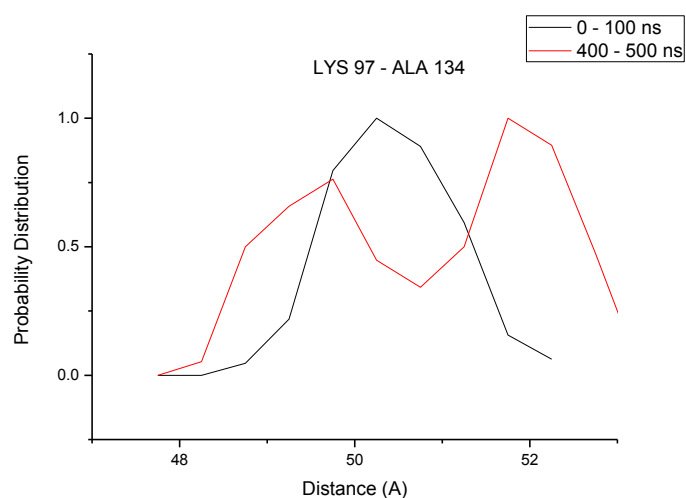


Figure 4.1.7.1.2 Probability distribution of distance between ECL1 (Lys97) and extracellular end of TM3 (Ala134)

Phe104 resides in binding pocket of P<sub>2</sub>Y<sub>12</sub> receptors. [125] P<sub>2</sub>Y<sub>12</sub> inhibits second messenger adenylyl cyclase activity upon activation. Probability distribution of Phe104 and Leu342 is stabilizes at 42 Å rather than scattered peaks observed during 0 – 100 ns. (Figure 4.1.7.1.3) This may indicate a lock which causes and G-protein independent signaling.

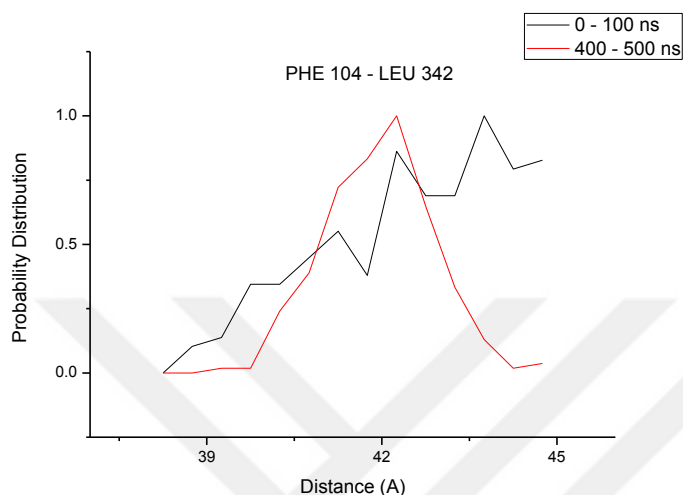


Figure 4.1.7.1.3 Probability distribution of distance between extracellular end of TM3 (Phe104) and cytoplasmic tail (Leu342)

In following correlations lower ends of TM3 and TM4 adopts a tighter distance in intracellular region and a wider distance with extracellular end of TM2. Ile135 is highly conserved among GPCRs whereas it has been substituted with valine in some GPCRs, correlations shows it forms van der Waals contacts with alanine at position 92 and 202.

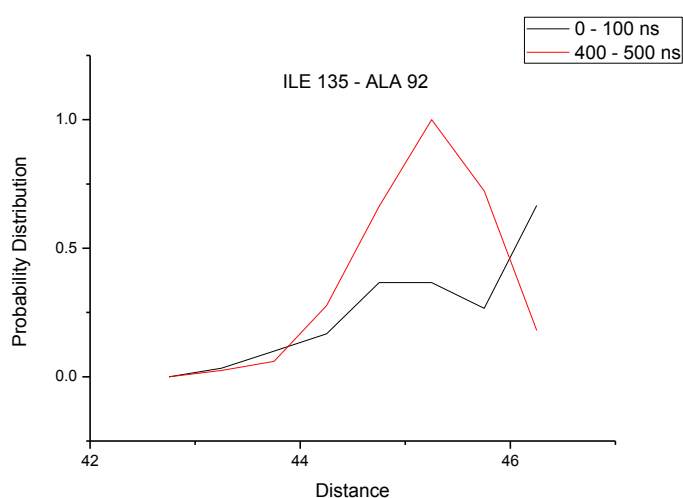


Figure 4.1.7.1.4 Probability distribution of distance between extracellular end of TM3 (Ile135) and extracellular end of TM2 (Ala92)

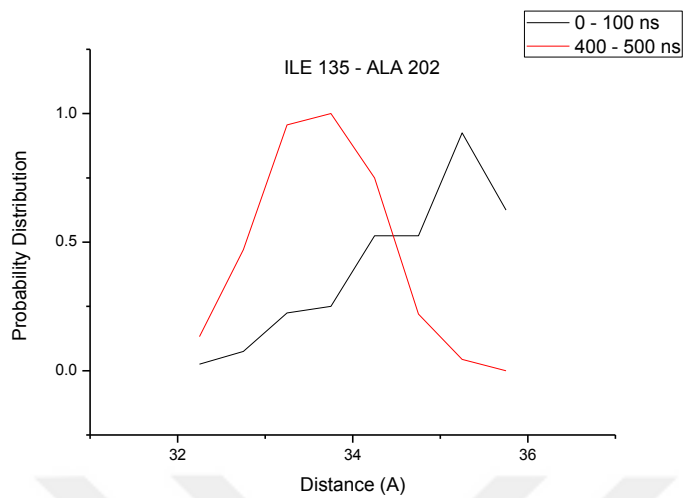


Figure 4.1.7.1.5 Probability distribution of distance between extracellular end of TM3 (Ile135) and extracellular part of TM5 (Ala202)

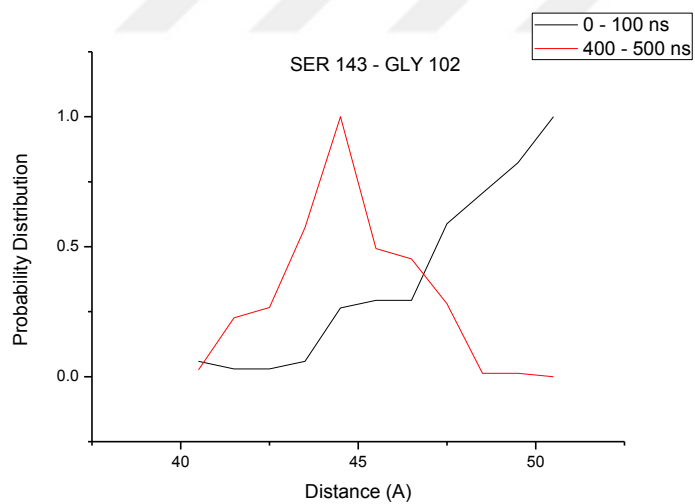


Figure 4.1.7.1.6 Probability distribution of distance between ICL2 (Ser143) and ECL1 (Gly102)

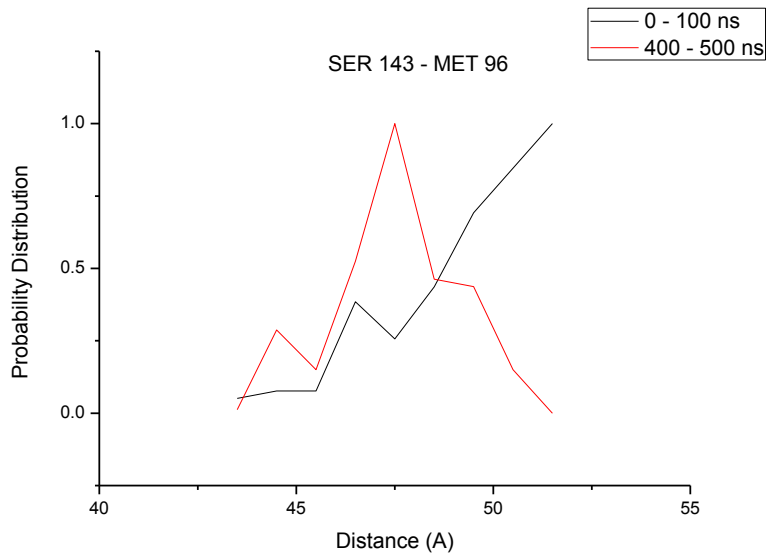


Figure 4.1.7.1.7 Probability distribution of distance between ICL2 (Ser143) and ECL1 (Met96)

#### 400 – 500 ns

It is observed Met98 is important for signal transduction from intracellular region to extracellular region. Met98 correlates with intracellular loop 1 and Leu339 of cytoplasmic tail where depicted as important in G $\alpha$ s signalling. [117]

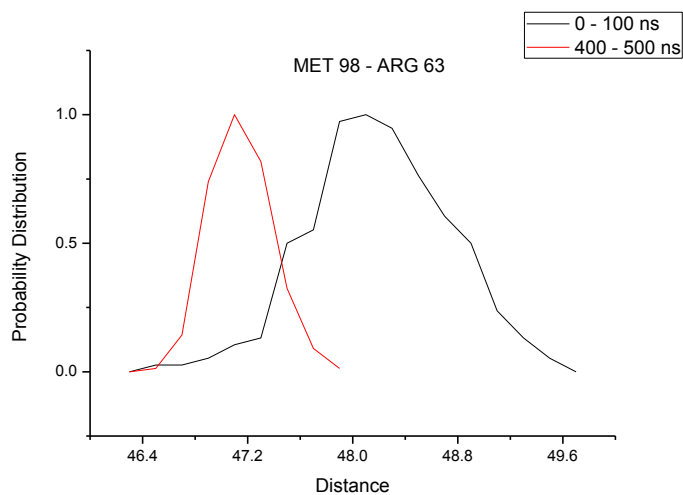


Figure 4.1.7.2.1 Probability distribution of distance between ECL1 (Met98) and ICL1 (Arg63)

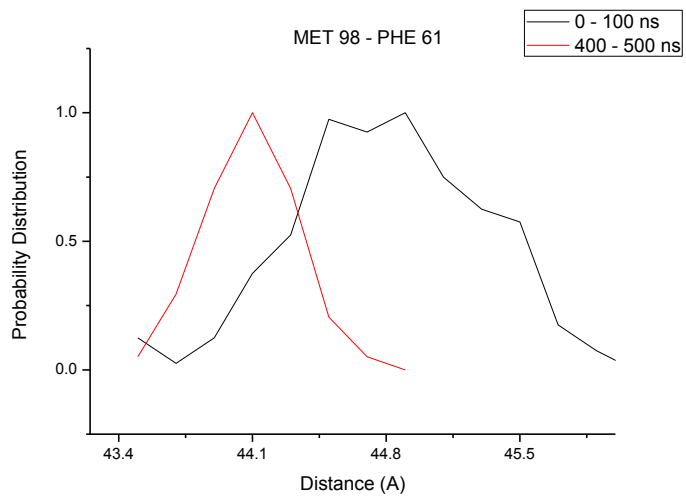


Figure 4.1.7.2.2 Probability distribution of distance between ECL1 (Met98) and ICL1 (Phe61)

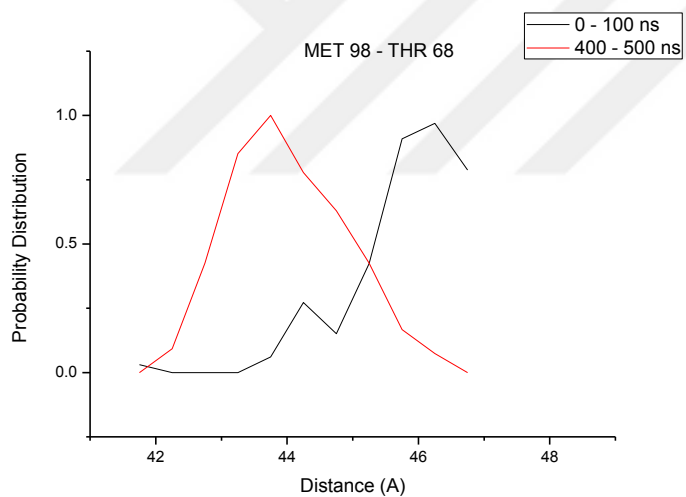


Figure 4.1.7.2.3 Probability distribution of distance between ECL1 (Met98) and intracellular end of TM2 (Phe68)



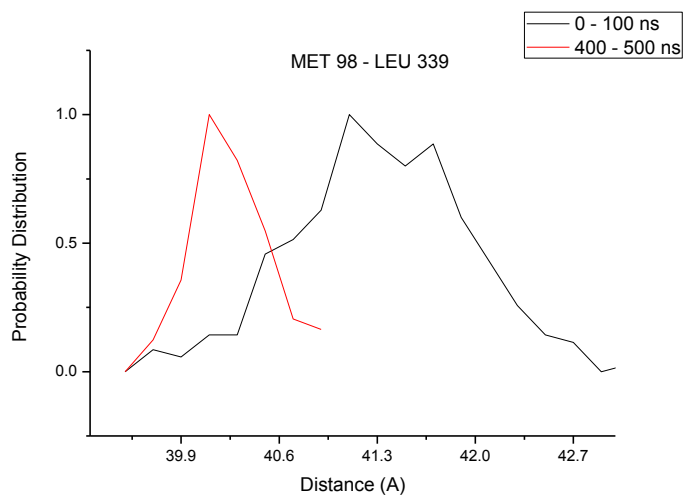


Figure 4.1.7.2.4 Probability distribution of distance between ECL1 (Met98) and cytoplasmic tail (Leu339)

Gly102 is another highly conserved residue in GPCRs and seems to be another important residue in signal transduction it increases the distance between binding pocket most probably to be able to hinder steric clashes which may occur during ligand binding.

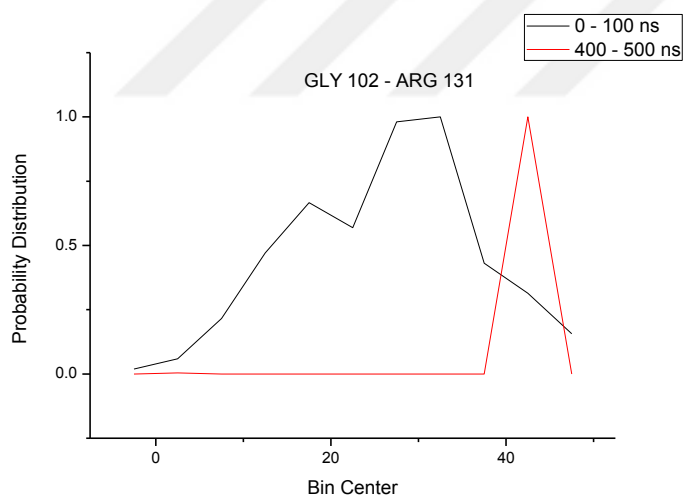


Figure 4.1.7.2.5 Probability distribution of distance between ECL1 (Gly102) and extracellular end of TM3 (Arg131)

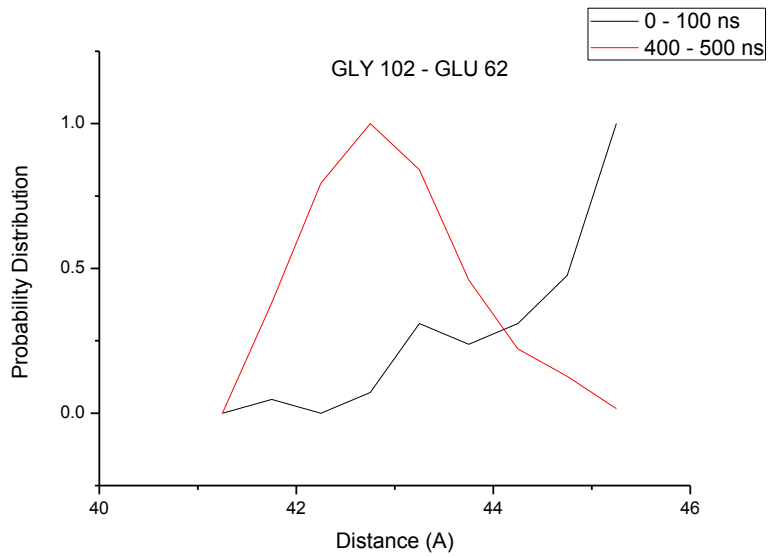


Figure 4.1.7.2.6 Probability distribution of distance between ECL1 (Gly102) and ICL1 (Glu62)

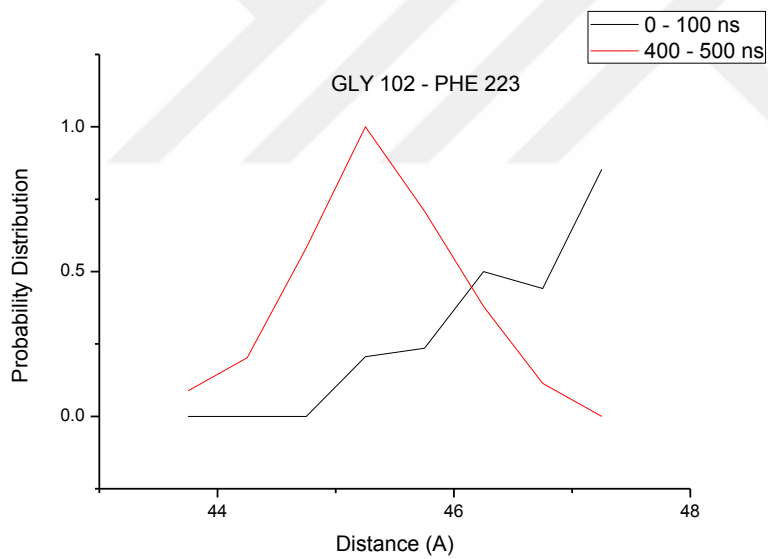


Figure 4.1.7.2.7 Probability distribution of distance between ECL1 (Gly102) and membrane region of TM5 (Phe223)

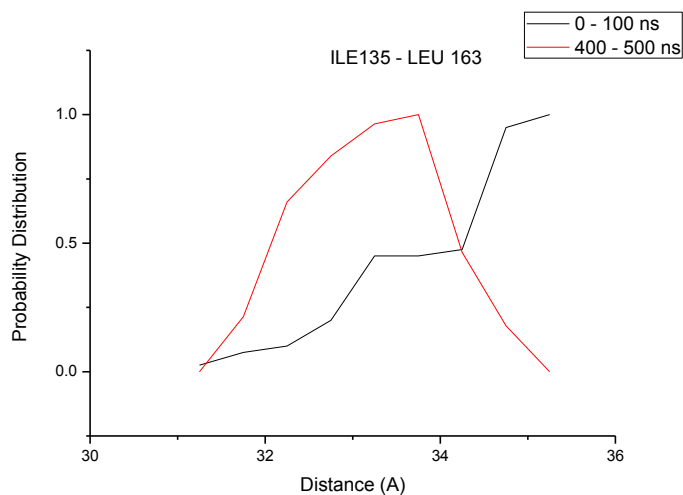


Figure 4.1.7.2.8 Probability distribution of distance between intracellular end of TM3 (Ile135) and membrane region of TM4 (Leu163)

Val160 is an important residue in ligand pocket formation of olfactory receptors. In olfactory receptors Val113 of MOR42-3, forms hydrophobic interactions with residues in TM3, TM4 and TM5 bringing these helices together. When Val113 is interacted with a polar residue, these interactions are disrupted and formation of a larger binding pocket is initiated. [126] In  $\beta$ 2Ar it is seen that these kind of interactions are governed by Gly102 and Ile 135.

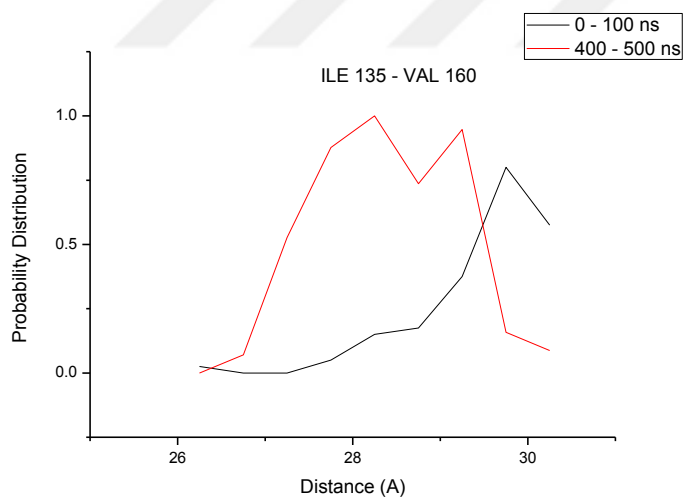


Figure 4.1.7.2.9 Probability distribution of distance between intracellular end of TM3 (Ile135) and membrane region of TM4 (Val160)

Phe193-Tyr141 are both involved in ligand binding, 4 Å change occurring in this region might be speculated as binding pocket reformation. [118, 119]

In receptor G-protein interaction in the patent of Kobilka et.al it is depicted Thr68, Asp130 interact with ICL2 via Tyr141 and enables docking of the receptor into a hydrophobic pocket on the G-protein surface. [127]

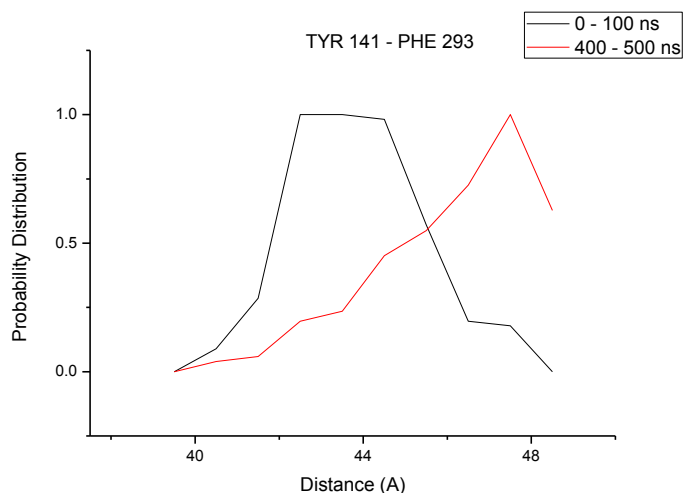


Figure 4.1.7.2.10 Probability distribution of distance between intracellular end of TM4 (Tyr141) and extracellular end of TM6 (Phe293)

His296 is known to locate in 3<sup>rd</sup> binding site in hydrophobic pocket formed by Asp300, Asn301 and Ile303 of  $\beta$ 2Ar and known for its function in stabilizing  $Zn^{2+}$ . [120] Val295 in this manner seems to function in communication with TM2 in the activation of G-protein signalling and  $Zn^{2+}$  stabilization. [121]

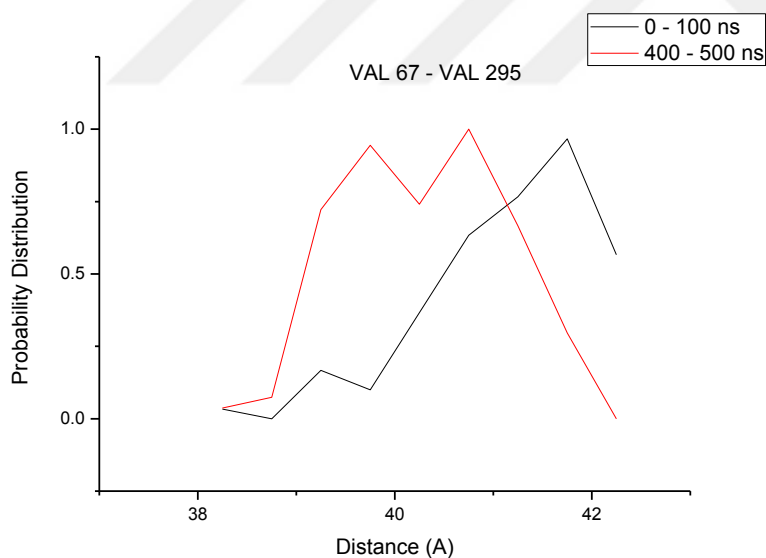


Figure 4.1.7.2.11 Probability distribution of distance between intracellular end of TM2 (Val67) and extracellular end of TM6 (Val295)

In first 100 ns domain-wise correlations, it is seen that every segment of protein is positively correlated with TM1 and ECL1 is correlated with cytoplasmic tail. ICL1 is correlated with ECL1, ECL2, ECL3, extracellular segment of TM4, intracellular segment of TM5, extracellular segment of TM6 and TM7. Intracellular segment of TM2 (position 67-73) is correlated with ECL2, ECL3, extracellular segment of TM2 (position 74-95) is correlated with ICL2 and whole TM2 is correlated with TM5. In last 100 ns domain-wise correlations TM1 correlations stays

intact, ICL1 loses its correlation with ECL1 and ECL3 and keeps its correlations with ECL2, intracellular segment of TM3, extracellular segment of TM4, intracellular segment of TM5, extracellular segment of TM6 and whole TM7. TM2 becomes correlated with ECL2 and ICL2. TM3 gets correlated with TM5 and two segment of TM3 as intracellular and extracellular becomes correlated.

In comparison of 1000 ns apo trajectory and 500 ns ligand mimicking trajectory it is seen that key players of receptor activation through Gas signaling pathway are ECL1 and TM3. In 1000 ns apo trajectory during 600 – 1000 ns period ECL1 expands the distance from ECL2, TM5 and TM6 where in 500 ns ligand mimicking trajectory it becomes closer to ICL1, TM3 and cytoplasmic tail. Transmembrane region interactions also shows a symmetrical change in lesser extent to distances between. Overall it is seen that interactions between TM2 – TM3 and ECL1 – TM3 are main indicators of conformational change.

When we have changed our scope from globular motions of protein to residue based interactions, it is seen that mostly interacting residues are conserved residues among some GPCRs. In residue based interactions, residues important for  $\beta$ 2Ar stabilization in lipid membrane, glycosylation sites, possible candidate residues which may involve in G-protein independent signaling pathways are coming to attention. Along these lines, GPCR proteins are dynamic entities which governs different signaling pathways.

## 5.1 Epinephrine Binding to $\beta$ -adrenoceptors

Epinephrine is an endogenous ligand of  $\beta$ -adrenoceptors. Epinephrine binding affinity to  $\beta_1$ -adrenoceptor is measured as 610 pM and 615 pM to  $\beta_2$ -adrenoceptor. In this chapter binding pockets of  $\beta_1, \beta_2$  and  $\beta_3$  adrenoceptors is investigated and via short steered MD simulations selectivity of epinephrine of  $\beta$ -adrenoceptors is tried to be understood both quantitatively and qualitatively. For selectivity of epinephrine binding free energies of epinephrine is calculated with steered MD method using Jarzinsky Inequality.

## 5.2 METHOD

Human  $\beta_1$ -adrenoceptor and  $\beta_3$ -adrenoceptor is modelled via homology modeling using Phyre<sup>2</sup> server. Modelled proteins are equilibrated in NVT ensemble for 4 ns and afterwards epinephrine is docked to proteins using Molegro. Best poses is chosen after 100 distinct runs. Acquired structures is equilibrated again for another 400 ps in order to maintain quasiequilibrium state.

Epinephrine is pulled from frozen backbones of  $\beta_1, \beta_2$  adrenoceptors with 0.02 Å/ps speed. Timestep is chosen 2 fs in each simulation and SMD data is collected in every 100<sup>th</sup> time step. Pulling is done till 4 Å in distance since it has been observed that rupture is occurring around 4 Å.

## 5.3 Binding Pockets

$\beta_1$ -adrenoceptor

Epinephrine is contacted with residues Pro45, Ala46, Asp356, Phe359, Lys347 and Asn344. It makes  $\pi$ -alkyl interactions with Pro45 and Ala46 and  $\pi$ - $\pi$  interactions with Phe359.

$\beta_2$ -adrenoceptor

In  $\beta_2$ -adrenoceptor epinephrine makes contact with Tyr199, Phe290, Asn293 and Ile294. In  $\beta_2$ -adrenoceptor  $\pi$ - $\pi$  interaction is got stronger via Tyr199 and Phe290.

$\beta_3$ -adrenoceptor

In  $\beta_3$ -adrenoceptor epinephrine makes contact with Asn192, Cys110, Met174 and Pro193. Hydrogen bond network in  $\beta_3$ -adrenoceptor is seen more compact than  $\beta_1$  and  $\beta_2$  and forms  $\pi$ -alkyl interactions with Pro193.

## 5.4 SMD Simulations

$\beta_1$ -adrenoceptor

Force for required pulling is observed around 100 pN and work values ranges between 1.5 kcal/mol to -1.5 kcal/mol.

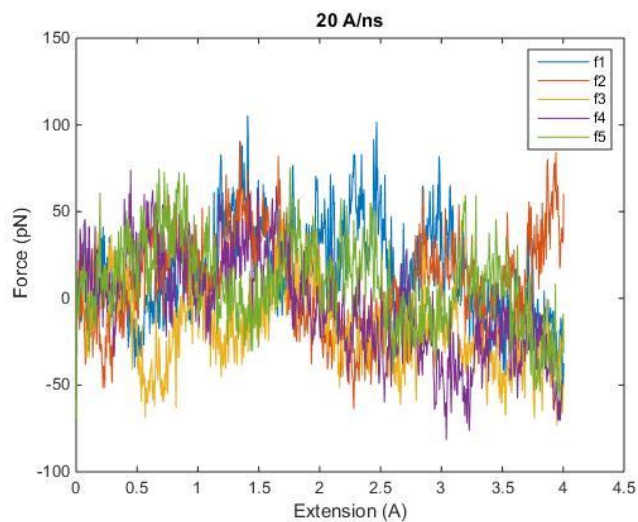


Figure 5.4.1 Force profiles of  $\beta_1$ -adrenoceptor-epinephrine pulling

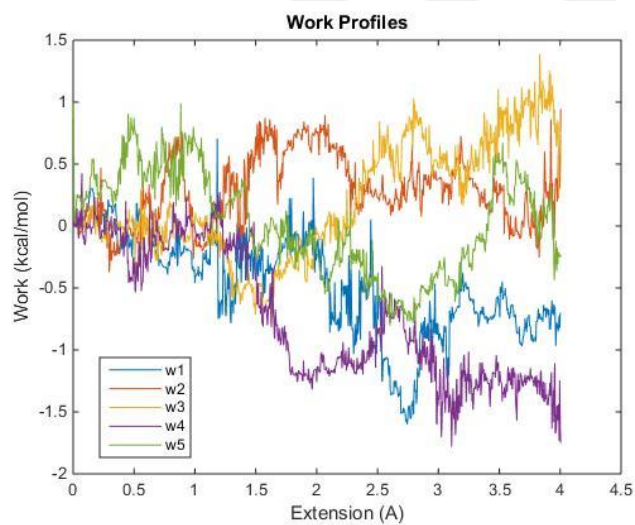


Figure 5.4.2 Work profiles of  $\beta_1$ -adrenoceptor-epinephrine pulling

## $\beta_2$ -adrenoceptor

Force profile of  $\beta_2$ -adrenoceptor increased to 300 pN. Work values are ranged from 2 kcal/mol to 12 kcal/mol.

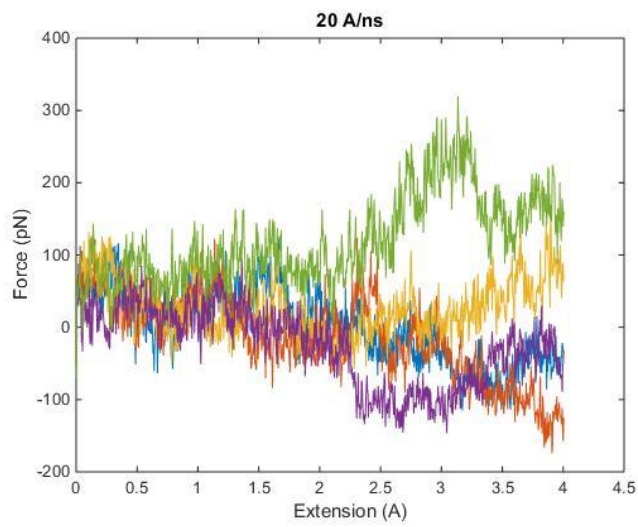


Figure 5.4.3 Force profiles of  $\beta_2$ -adrenoceptor-epinephrine pulling

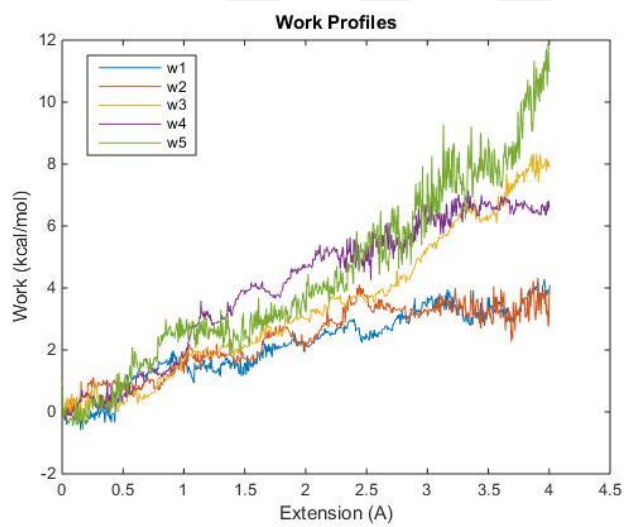


Figure 5.4.4 Work profiles of  $\beta_2$ -adrenoceptor-epinephrine pulling



### $\beta_3$ -adrenoceptor

In force profiles of  $\beta_3$  similar trend to  $\beta_1$  is observed. Work values acquired are ranged from 0.5 kcal/mol to 3.5 kcal/mol.

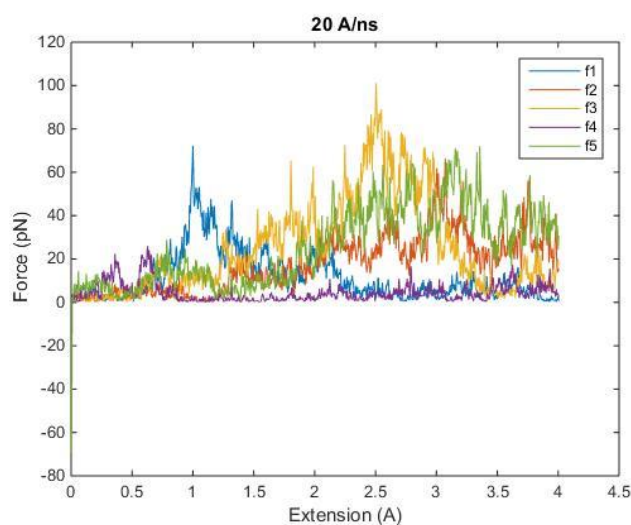


Figure 5.4.5 Force Profiles

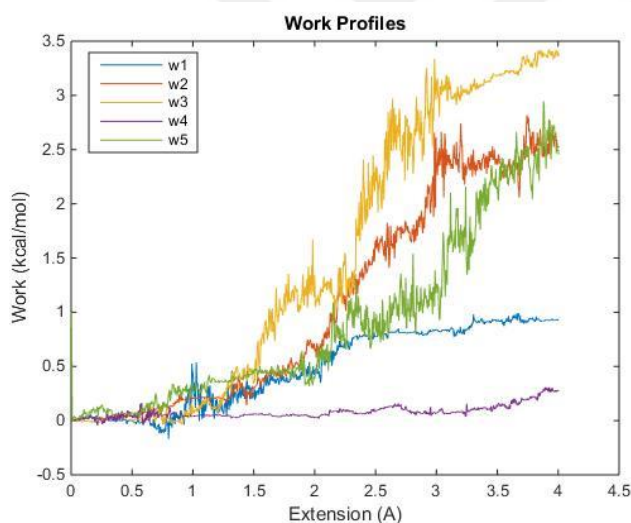


Figure 5.4.6 Force Profiles

In representative PMF calculations it seen that epinephrine selectivity is as  $\beta_2 > \beta_3 > \beta_1$ . (Figure 5.4.7). These values should not be accounted for exact binding free energy due to high statistical variance in acquired work values.

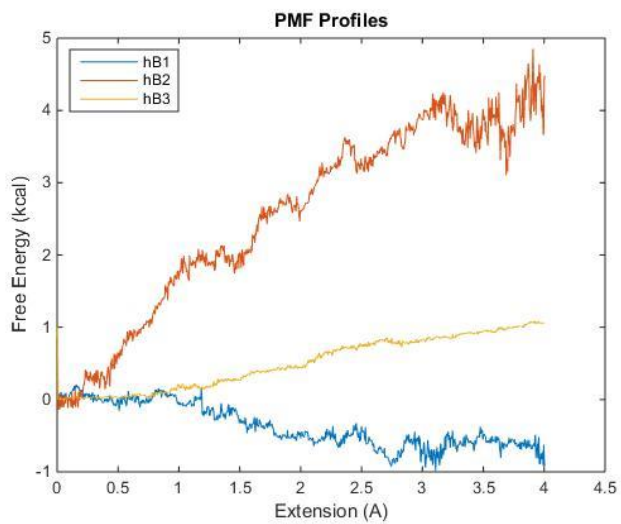


Figure 5.4.7 Force Profiles



## Chapter 6

### Conclusion

In this study, we aimed to understand  $\beta_2$ Ar in a comprehensive manner and tried to approach from different perspectives. GPCR proteins are highly dynamic and flexible proteins. In the investigation of inactive-conformation of ICL3, we found that the required energy for ICL3 to adopt an active state conformation is more or less equal to the energy emanates from G-protein binding to  $\beta_2$ Ar. This indicates that ICL3 is mostly involved in G-protein dependent signaling pathways. This proposition can also be supported by the diversity of ICL3 regions observed among GPCRs.

In investigation of allosteric mechanisms underlying  $\beta_2$ Ar, we observed that intracellular and extracellular regions are in cross communication whereas in membrane region we are seeing correlations from nearby regions for stabilization/localization of protein into lipid membrane. In 1000 ns apo trajectory the distance between ICL1 and intracellular end of TM6 shows a 7 Å change in distances between Leu64, Gln65, Thr66 and Cys265, Lys267 occurs during last 400 ns whereas the distance between intracellular end of TM2 (Ala59) and (TM6) Lys267 distance gets 8 Å apart which brings intracellular ends of TM1 and TM6 together. This formation is controlled by the ligands Phe101 and Trp99 which governs ECL2, TM5 and TM6 formation which allows a wider space in the cavity between 7 TM helices. In 500 ns ligand mimicking trajectory, main actors changes from ICL1, TM5 and TM6 to ECL1 (Met98, Gly102) and TM3 (Ile 135) which probably involves in stabilization of membrane region in lipid and the distance between TM1 and TM2 mostly.

In residue based inspections it is seen that these residues also include post-translational modification sites such as Asn69 and G-protein independent signaling pathways such as Phe104 and Leu342. The latter may be due to an absence of G-protein during simulation. For instance Tyr141 is an important player in G-protein docking and its distance with Phe293 in 500 ns ligand mimicking trajectory is not stabilizing. This may show that  $\beta_2$ Ar also sends signals for G-protein recognition. Another important factor is Ile291 is proposed as a rotamer switch in  $\beta_2$ Ar. Our findings in Cys184-Ile291 distance in 1000 ns apo trajectory support the existence of a such mechanism. Also, not observing any ligand binding pocket residues in this analysis implies that binding pocket residues have limited but strategic contacts. Due to flexible nature of GPCRs, it may be speculated that GPCR proteins have trigger alarms on several levels for activation of the specific pathway.

Binding studies of endogenous ligand epinephrine show that epinephrine is selective of  $\beta_2 > \beta_3 > \beta_1$ . This result seems logical when expression sites of  $\beta_1$ ,  $\beta_2$ ,  $\beta_3$  are taken into account.

## Appendix I

### Constraints and Steered MD Section inserted in Configuration File

```
# Fixed Atoms Constraint (set PDB beta-column to 1)
```

```
if {1}{
```

```
fixedAtoms on
```

```
fixedAtomsFile 8A_equil.ref
```

```
fixedAtomsCol B
```

```
fixedAtomsForces on }
```

```
# Extra constraints
```

```
extraBonds on
```

```
extraBondsFile constraint.txt
```

```
# Steered MD set up
```

```
SMD on
```

```
SMDFile 8A_eq_smd.ref
```

```
SMDk 6
```

```
SMDVel 0.0000024
```

```
SMDDir 0.0125 -0.9896 0.1434
```

```
SMDOutputFreq 100
```

## Appendix II

### Force and Work Calculations

---

```
filename=input('Enter force file name: ','s');
delimiterIn=' ';
ft=importdata(filename,delimiterIn);
ftemp=ft(:,2)*0.0144;
f1 = size(54000);

filename2=input('Enter extension file name: ','s');
e=importdata(filename2,delimiterIn);
d=e(:,2);
d=d(1:end)-d(1);
t=e(:,1)/1000;
k=6;
fsum=0;
v=input('Velocity (A/ps): ');
dt=input('Time Interval(ps): ');

c1=size(54001);
j=1;
initial=0;
temp=0;
while j < 54002
    c1(j)= temp+v*dt;
    temp=initial+c1(j);
    j=j+1;
end
c=transpose(c1);
for i=2:54001
    if i ==1;
        fsum=fsum+k*(v*dt-(d(i)));
        f1(i)=fsum;
    else
        fsum=fsum+k*(v*dt-(d(i)-d(i-1)));
        f1(i)= fsum;
    end
end
```

```

end
f1uc=transpose(f1);
force=f1uc-f1;
w1 = size(54000);
j=1;
%worksum=0;
for i=2:54001
    if i==1;
        w1(j)= force(j)*(d(i)); % initial worksum = 0; if i==1 worksum = worksum+
force(i)*(d(i)), w1(i)=worksum
        w1(j)=worksum;
        j=j+1;
    else
        w1(j)= force(j)*(d(i)-d(i-1)); % else worksum = worksum + force(i)*(d(i)-
d(i-1)), w1(i)=worksum
        w1(j)=worksum;
        j=j+1;
    end
end
end
work = transpose(w1);

```

## Appendix III

### Production of Covariance Matrices

#### For frames (stations) 0 - 400 ns

---

##### *Extracting coordinates*

```
for i=0:399
    filename = num2str(i);
    s = pdbread ([filename '.pdb']);
    coordinates=zeros(311,3);
    j=1;
    for k=1:5055
        if strcmp(s.Model.Atom(k).AtomName,'CA')==1;
            coordinates(j,:)=[s.Model.Atom(k).X s.Model.Atom(k).Y
s.Model.Atom(k).Z];
            j=j+1;
        end
    end
    my_field = strcat('frame_',num2str(i));
    variable.(my_field) = coordinates;
end
```

##### Calculating average positions

```
r_bar = zeros(size(variable.frame_0));
for i = 0:398
    index=num2str(i);
    frame_index = ['frame_' index];
    temp = variable.(frame_index);
    r_bar = r_bar + temp;
end
```

##### *Calculating the difference in each frame relatively to the average positions*

```
r_bar = r_bar/398;
for i=0:399
    index=num2str(i);
    a = ['frame_' index];
    delta = zeros(size(variable.frame_0));
    delta = (variable.(a) - r_bar);
    my_field = strcat('deltaR_',num2str(i));
    deltaR.(my_field) = delta;
```

end

*Calculating  $\Delta R_i \cdot \Delta R_j$  for each time frame*

```
product = zeros(311,311);  
for k= 0 : 398  
    c = ['deltaR_' num2str(k)];  
    mat = deltaR.(c);  
    for i=1:311  
        for j = 1:311  
            product(i,j)= dot (mat(i,:),mat(j,:));  
        end  
    end  
    my_field = strcat('pre_',num2str(k));  
    pre.(my_field) = product;  
end
```

end

*Calculating  $\langle \Delta R_i \cdot \Delta R_j \rangle$  for the first 400 ns*

*Sum of each dot product and division of the total time frame*

```
connect = zeros(311,311);  
for l=0:398  
    d = ['pre_' num2str(l)];  
    connect = connect + pre.(d);  
end
```

end

```
connect = connect/400;
```

*Calculating  $\Delta R_i / |\Delta R_i| \cdot \Delta R_j / |\Delta R_j|$*

```
second_product = zeros(311,311);  
for k= 0 : 398  
    c = ['deltaR_' num2str(k)];  
    mat = deltaR.(c);  
    for i=1:311  
        for j = 1:311  
            second_product(i,j)= dot  
            ((mat(i,:)/norm(mat(i,:))),(mat(j,:)/norm(mat(j,:))));  
        end  
    end
```

```
end
```

```
end
```

```
my_field = strcat('pre_norm_',num2str(k));
```



```
pre_norm.(my_field) = second_product;  
end
```

*Normalized correlaton matrice for total 0-400 ns period*

```
second_connect = zeros(311,311);  
for z=0:398  
    e = ['pre_norm_' num2str(z)];  
    second_connect = second_connect + pre_norm.(e);  
end  
second_connect = second_connect/399;
```



# Bibliography

## GPCR Structure and Function

- [1] G protein-coupled receptors. Accessed on 18/04/2017. IUPHAR/BPS Guide to PHARMACOLOGY, <http://www.guidetopharmacology.org/GRAC/FamilyDisplayForward?familyId=694>.
- [2] Stevens et.al (2013).The GPCR Network: a large-scale collaboration to determine human GPCR structure and function. *Nature Reviews*
- [3] Umemori, H., Inoue, T., Kume, S., Sekiyama, N., Nagao, M., Itoh, H., ... Yamamoto, T. (1997). Activation of the G protein Gq/11 through tyrosine phosphorylation of the alpha subunit. *Science*, 276(5320), 1878–81.
- [4] Wang C., Jiang Y., Ma J., Wu H., Wacker D., Katritch V., Han G.W., Liu W., Huang X.P., Vardy E., McCorvy J.D., Gao X., Zhou X.E., Melcher K., Zhang C., Bai F., Yang H., Yang L. (2013) .Structural basis for molecular recognition at serotonin receptors. *Science*,340(6132), 610-614
- [5] Wacker D., Wang C., Katritch V., Han G.W., Huang X.P., Vardy E., McCorvy J.D., Jiang Y., Chu M., Siu F.Y., Liu W., Xu H.E., Cherezov V., Roth B.L., Stevens R.C. (2013).Structural features for functional selectivity at serotonin receptors. *Science* 340:615-619
- [6] Ullmer C., Boddeke H.G., Schmuck K., Lubbert H. (1996).5-HT2B receptor-mediated calcium release from ryanodine-sensitive intracellular stores in human pulmonary artery endothelial cells.*Br. J. Pharmacol.* 117:1081-1088
- [7] Kruse A.C., Ring A.M., Manglik A., Hu J., Hu K., Eitel K., Hubner H., Pardon E., Valant C., Sexton P.M., Christopoulos A., Felder C.C., Gmeiner P., Steyaert J., Weis W.I., Garcia K.C., Wess J., Kobilka B.K. (2013) Activation and allosteric modulation of a muscarinic acetylcholine receptor.*Nature* 504:101-106
- [8] Umemori, H., Inoue, T., Kume, S., Sekiyama, N., Nagao, M., Itoh, H., ... Yamamoto, T. (1997). Activation of the G protein Gq/11 through tyrosine phosphorylation of the alpha subunit. *Science*, 276(5320), 1878–81.
- [9] Ryberg E., Vu H.K., Larsson N., Groblewski T., Hjorth S., Elebring T., Sjoegren S., Greasley P.J. (2005).Identification and characterisation of a novel splice variant of the human CB1 receptor.*FEBS Lett.* 579:259-264
- [10] Doranz B.J., Orsini M.J., Turner J.D., Hoffman T.L., Berson J.F., Hoxie J.A., Peiper S.C., Brass L.F., Doms R.W. (1999).Identification of CXCR4 domains that support coreceptor and chemokine receptor functions.*J. Virol.* 73:2752-2761
- [11] Villar V.A.M., Jones J.E., Armando I., Palmes-Saloma C., Yu P., Pascua A.M., Keever L., Arnaldo F.B., Wang Z., Luo Y., Felder R.A., Jose P.A. (2009).G protein-coupled receptor kinase 4 (GRK4) regulates the phosphorylation and function of the dopamine D3 receptor.*J. Biol. Chem.* 284:21425-21434
- [12] Webb M.L., Chao C.-C., Rizzo M., Shapiro R.A., Neubauer M., Liu E.C.K., Aversa C.R., Brittain R.J., Treiger B. (1995). Cloning and expression of an endothelin receptor subtype B from human prostate that mediates contraction. *Mol. Pharmacol.* 47:730-737
- [13] Briscoe C.P., Tadayon M., Andrews J.L., Benson W.G., Chambers J.K., Eilert M.M., Ellis C., Elshourbagy N.A., Goetz A.S., Minnick D.T., Murdock P.R., Sauls H.R. Jr., Shabon U., Spinage L.D., Strum J.C., Szekeres P.G., Tan K.B., Way J.M. Muir A.I. (2003).The orphan G protein-coupled receptor GPR40 is activated by medium and long-chain fatty acids. *J. Biol. Chem.* 278:11303-11311
- [14] Sum C.S., Tikhonova I.G., Neumann S., Engel S., Raaka B.M., Costanzi S., Gershengorn M.C. (2007). Identification of residues important for agonist recognition and activation in GPR40.*J. Biol. Chem.* 282:29248-29255
- [15] Tanaka K., Masu M., Nakanishi S.(1990).Structure and functional expression of the cloned rat neurotensin receptor.*Neuron*.4(6): 847-854
- [16] Feng, Yuan.,He, Xiaozhou., Yang, Yilin., Chao, Dongman., H. Lazarus, Lawrence., Xia, Ying. Current Drug Targets, Volume 13, Number 2, February 2012, pp. 230-246(17)
- [17] Rios C., Gomes I., Devi L.A. (2006).mu opioid and CB1 cannabinoid receptor interactions: reciprocal inhibition of receptor signaling and neuritogenesis.*Br. J. Pharmacol.* 148:387-395

- [18] Serhan C.N., Fierro I.M., Chiang N., Pouliot M. (2001). Nociceptin stimulates neutrophil chemotaxis and recruitment: Inhibition by aspirin-triggered-15-Epi-lipoxin A4. *J. Immunol.* 166:3650-3654
- [19] Sakurai T., Amemiya A., Ishii M., Matsuzaki I., Chemelli R.M., Tanaka H., Williams S.C., Richardson J.A., Kozlowski G.P., Wilson S., Arch J.R.S., Buckingham R.E., Haynes A.C., Carr S.A., Annan R.S., McNulty D.E., Liu W.-S., Terrett J.A., Yanagisawa M. (1998). Orexins and orexin receptors: a family of hypothalamic neuropeptides and G protein-coupled receptors that regulate feeding behavior. *Cell* 92:573-585
- [20] Jin J., Daniel J.L., Kunapuli S.P. (1998). Molecular basis for ADP-induced platelet activation. II. The P2Y1 receptor mediates ADP-induced intracellular calcium mobilization and shape change in platelets. *J. Biol. Chem.* 273:2030-2034
- [21] Hoare, S. R. J. (2005). Mechanisms of peptide and nonpeptide ligand binding to Class B G-protein-coupled receptors. *Drug Discovery Today*, 10(6), 417–427. [https://doi.org/10.1016/S1359-6446\(05\)03370-2](https://doi.org/10.1016/S1359-6446(05)03370-2)
- [22] Koth C.M., Murray J.M., Mukund S., Madjidi A., Minn A., Clarke H.J., Wong T., Chiang V., Luis E., Estevez A., Rondon J., Zhang Y., Hotzel I., Allan B.B. (2012). Molecular basis for negative regulation of the glucagon receptor. *Proc. Natl. Acad. Sci. U.S.A.* 109:14393-14398
- [23] Wu H., Wang C., Gregory K.J., Han G.W., Cho H.P., Xia Y., Niswender C.M., Katritch V., Meiler J., Cherezov V., Conn P.J., Stevens R.C. (2014). Structure of a class C GPCR metabotropic glutamate receptor 1 bound to an allosteric modulator. *Science* 344:58-64
- [24] Martin S.C., Russek S.J., Farb D.H. (1999). Molecular identification of the human GABABR2: cell surface expression and coupling to adenylyl cyclase in the absence of GABABR1. *Mol. Cell. Neurosci.* 13:180-191
- [25] White J.H., Wise A., Main M.J., Green A., Fraser N.J., Disney G.H., Barnes A.A., Emson P., Foord S.M., Marshall F.H. (1998). Heterodimerization is required for the formation of a functional GABA(B) receptor. *Nature* 396:679-682
- [26] Endoh-Yamagami S., Evangelista M., Wilson D., Wen X., Theunissen J.W., Phamluong K., Davis M., Scales S.J., Solloway M.J., de Sauvage F.J., Peterson A.S. (2009). The mammalian Cos2 homolog Kif7 plays an essential role in modulating Hh signal transduction during development. *Curr. Biol.* 19:1320-1326
- [27] Yoo B1, Lemaire A, Mangmool S, Wolf MJ, Curcio A, Mao L, Rockman HA., (2009) Beta1-adrenergic receptors stimulate cardiac contractility and CaMKII activation in vivo and enhance cardiac dysfunction following myocardial infarction. *Am J Physiol Heart Circ Physiol.* 297(4):H1377-86.
- [28] Warne et.al. (2008) Structure of a b1-adrenergic G-protein coupled receptor V.454:486-492
- [29] Doughty and Sharpe. (1997) BETA-ADRENERGIC BLOCKING AGENTS IN THE TREATMENT OF CONGESTIVE HEART FAILURE: Mechanisms and Clinical Results. *Annu. Rev. Med.* 48:103–14
- [30] Smith and Teitler. (1999). Beta-Blocker Selectivity at Cloned Human Beta1- and Beta2-Adrenergic Receptors. *Cardiovascular Drugs and Therapy.* 13:123–126
- [31] Balaraman, Bhattacharya and Vaidehi. Structural Insights into Conformational Stability of Wild-Type and Mutant b1-Adrenergic Receptor. *Biophysical Journal.* 99:568-577.
- [32] Sugimoto, Y., Fujisawa, R., Tanimura, R., Lattion, A. L., & Cotecchia, S. (2002). with Key Amino Acids Responsible for  $\beta_1$ -Selective Binding in Resting and Active States. *Pharmacology*, 301(1), 51–58.
- [33] Vanni, S., Neri, M., Tavernelli, I., & Rothlisberger, U. (2009). Observation of “ionic lock” formation in molecular dynamics simulations of wild-type  $\beta_1$  and  $\beta_2$  adrenergic receptors. *Biochemistry*, 48(22), 4789–4797. <https://doi.org/10.1021/bi900299f>
- [34] Kobilka, B. K., & Deupi, X. (2007). Conformational complexity of G-protein-coupled receptors. *Trends in Pharmacological Sciences*, 28(8), 397–406. <https://doi.org/10.1016/j.tips.2007.06.003>
- [35] Schwartz, T.W. et al. (2006) Molecular mechanism of 7TM receptor activation – a global toggle switch model. *Annu. Rev. Pharmacol. Toxicol.* 46, 481–519
- [36] Katritch, V., & Abagyan, R. (2011). GPCR agonist binding revealed by modeling and crystallography. *Trends in Pharmacological Sciences*, 32(11), 637–643. <https://doi.org/10.1016/j.tips.2011.08.001>

- [37] Liu, J. J., Horst, R., Katritch, V., Stevens, R. C., & Wüthrich, K. (2012). Biased Signaling Pathways in. *Science (New York, N.Y.)*, 1106(March), 1106–1111. <https://doi.org/10.1126/science.1215802>
- [38] Shukla, A. K., Singh, G., & Ghosh, E. (2014). Emerging structural insights into biased GPCR signaling. *Trends in Biochemical Sciences*, 39(12), 594–602. <https://doi.org/10.1016/j.tibs.2014.10.001>
- [39] Daaka, Y., Luttrell, L.M. and Lefkowitz, R.J. (1997) Switching of the coupling of the  $\beta_2$ -adrenergic receptor to different G proteins by protein kinase A. *Nature*, 390, 88–91
- [40] Pin, J. P. (2000). EMBO MEMBERS ' REVIEW Molecular tinkering of G protein-coupled receptors : an evolutionary success. *EMBO Journal*, 18(7), 1723–1729. Retrieved from <http://www.nature.com/emboj/journal/v18/n7/abs/7591600a.html>
- [41] Luttrell et al. (1999)  $\beta$ -Arrestin-Dependent Formation of  $\beta_2$  Adrenergic Receptor-Src Protein Kinase Complexes. *Science*. 655-661
- [42] A. W. Kahsai et al., (2011). *Nat. Chem. Biol.* 7, 692
- [43] Kenakin, T. (2003) Ligand-selective receptor conformations revisited: the promise and the problem. *Trends Pharmacol. Sci.* 24, 346–354
- [44] Ghanouni, P. et al. (2000) The Effect of pH on  $\beta_2$  adrenoceptor function. Evidence for protonation-dependent activation. *J. Biol. Chem.* 275, 3121–3127
- [45] Samama, P. et al. (1993) A mutation-induced activated state of the  $\beta_2$ - adrenergic receptor. Extending the ternary complex model. *J. Biol. Chem.* 268, 4625–4636
- [46] Ozgur, C., Doruker, P., & Akten, E. D. (2016). Investigation of allosteric coupling in human  $\beta_2$ -adrenergic receptor in the presence of intracellular loop 3. *BMC Structural Biology*, 16(1), 9. <https://doi.org/10.1186/s12900-016-0061-9>
- [47] Strader, C.D. et al. (1989) Identification of two serine residues involved in agonist activation of the beta-adrenergic receptor. *J. Biol. Chem.* 264, 13572–13578
- [48] Liapakis, G. et al. (2000) The forgotten serine. A critical role for Ser- 2035.42 in ligand binding to and activation of the beta2-adrenergic receptor. *J. Biol. Chem.* 275, 37779–37788
- [49] Wieland, K. et al. (1996) Involvement of Asn-293 in stereospecific agonist recognition and in activation of the beta 2-adrenergic receptor. *Proc. Natl. Acad. Sci. U.S.A.* 93, 9276–9281
- [50] Altenbach, C. et al. (2008) High-resolution distance mapping in rhodopsin reveals the pattern of helix movement due to activation. *Proc. Natl. Acad. Sci. U.S.A.* 105, 7439–7444
- [51] Kufareva, I. et al. (2011) Status of GPCR modeling and docking as reflected by community wide GPCR Dock 2010 assessment. *Structure* 19, 1108–1126
- [52] Rasmussen, S.G. et al. (2011) Crystal structure of the beta adrenergic receptor-Gs protein complex. *Nature* DOI: 10.1038/nature10361
- [53] Changeux, J.P. and Edelstein, S.J. (2005) Allosteric mechanisms of signal transduction. *Science* 308, 1424–1428
- [54] Katritch, V. et al. (2009) Analysis of full and partial agonists binding to beta2-adrenergic receptor suggests a role of transmembrane helix V in agonist-specific conformational changes. *J. Mol. Recognit.* 22, 307–318
- [55] Strader, C. D., Candelore, M. R., Hill, W. S., Sigal, I. S., and Dixon, R. A. F. (1989) Identification of 2 serine residues involved in agonist activation of the beta-adrenergic-receptor. *J. Biol. Chem.* 264:13572–13578
- [56] Wieland, K., Zuurmond, H. M., Krasel, C., Ijzerman, A. P., and Lohse, M. J. (1996) Involvement of Asn-293 in stereospecific agonist recognition and in activation of the beta2-adrenergic receptor. *Proc. Natl. Acad. Sci. U.S.A.* 93, 9276–9281

- [57] Isogaya, M., Yamagiwa, Y. M., Fujita, S., Sugimoto, Y., Nagao, T., and Kurose, H. (1998) Identification of a key amino acid of the beta (2)-adrenergic receptor for high affinity binding of salmeterol *Mol. Pharmacol.* 54, 616–622
- [58] Palm D, Munch G, Dees C, Hekman M. Mapping of  $\beta$ -adrenoceptor coupling domains to Gs-protein by site-specific synthetic peptides. *FEBS Lett.* 1989;254:89–93
- [59] Okamoto T, Murayama Y, Hayashi Y, Inagaki M, Ogata E, Nishimoto I. Identification of a Gs activator region of the  $\beta$ 2-adrenergic receptor that is autoregulated via protein kinase A-dependent phosphorylation. *Cell.* 1991; 67(4):723–30
- [60] Ferrer-Lorente R, Cabot C, Fernández-López JA, Alemany M (September 2005). “Combined effects of oleoylestrone and a  $\beta$ 3-adrenergic agonist (CL316,243) on lipid stores of diet-induced overweight male Wistar rats”. *Life Sciences* 77 (16): 2051–8. doi:10.1016/j.lfs.2005.04.008.
- [61] Rang, H. P. (2003). *Pharmacology*. Edinburgh: Churchill Livingstone. ISBN 0-443-07145-4. Page 163
- [62] Nahmias et.al. (1991) “Molecular characterization of the mouse  $\beta$ 3-adrenergic receptor: relationship with the atypical receptor of adipocytes” *The EMBO Journal* 10 (12): 3721 -3727
- [63] O'Dowd, B.F., Hnatowich, M., Caron, M.G., Lefkowitz, R.J. and Bouvier, M. (1989b) *J. Biol. Chem.* 264: 7564-7589.
- [64] Ligett et.al. (1993) “Structural basis for receptor subtype-specific regulation revealed by a chimeric  $\beta$ 3/ $\beta$ 2-adrenergic receptor.” *Proc. Natl. Acad. Sci.* 90: 3665-3669
- [65] Consoli D, Leggio GM, Mazzola C, Micale V, Drago F (November 2007). “Behavioral effects of the  $\beta$ 3 adrenoceptor agonist SR58611A: is it the putative prototype of a new class of antidepressant/anxiolytic drugs?”. *European Journal of Pharmacology* 573 (1–3): 139–47.
- [66] Overstreet DH, Stemmelin J, Griebel G (June 2008). “Confirmation of antidepressant potential of the selective  $\beta$ 3 adrenoceptor agonist amibegron in an animal model of depression”. *Pharmacology, Biochemistry, and Behavior* 89 (4): 623–6.
- [67] Fu L, Isobe K, Zeng Q, Suzukawa K, Takekoshi K, Kawakami Y (2008). “The effects of beta(3)-adrenoceptor agonist CL-316,243 on adiponectin, adiponectin receptors and tumor necrosis factor-alpha expressions in adipose tissues of obese diabetic KKAy mice”. *European Journal of Pharmacology* 584 (1):202–6.
- [68] Candelore MR, Deng L, Tota L, Guan XM, Amend A, Liu Y, Newbold R, Cascieri MA, Weber AE (Aug 1999). “Potent and selective human beta(3)-adrenergic receptor antagonists.”. *The Journal of Pharmacology and Experimental Therapeutics* 290 (2): 649–55.
- [69] Larsen TM, Toubro S, van Baak MA, Gottesdiener KM, Larson P, Saris WH, Astrup A (2002). “Effect of a 28-d treatment with L-796568, a novel  $\beta$ 3-adrenergic receptor agonist, on energy expenditure and body composition in obese men”. *The American Journal of Clinical Nutrition* 76 (4): 780–8.
- [70] Gras J (2012). “Mirabegron for the treatment of overactive bladder”. *Drugs of today (Barcelona, Spain : 1998)* 48 (1): 25–32.
- [71] Hicks A, McCafferty GP, Riedel E, Aiyar N, Pullen M, Evans C, Luce TD, Coatney RW, Rivera GC, Westfall TD, Hieble JP (2007). “GW427353 (solabegron), a novel, selective beta3-adrenergic receptor agonist, evokes bladder relaxation and increases micturition reflex threshold in the dog”. *The Journal of Pharmacology and Experimental Therapeutics* 323 (1): 202–9.
- [72] Arch, J.R.S., Ainsworth, A.T., Cawthome, M.A., Piercy, V., Sennitt, M.V., Thody, V.E., Wilson, C. and Wilson, S. (1984) *Nature*, 309:163-165.
- [73] McLaughlin, D.P. and MacDonald, A. (1990) *Br. J. Pharmacol.*, 101:569-574.
- [74] Lewis et.al. (1977) “Thyroid Hormone Regulation of Beta-adrenergic Number”. *The Journal of Biological Chemistry*. 252 (8):2787-2789.
- [75] Benovic, J. L., De Blasi, A., Stone, W. C., Caron, M. G. & Lefkowitz, R. J. (1989) *Science* 246: 235-240.
- [76] Jarzynski, C. (1996). A nonequilibrium equality for free energy differences, 78(14), 0–3.  
<https://doi.org/10.1103/PhysRevLett.78.2690>

- [77] Gullingsrud, J. R., Braun, R., & Schulten, K. (1999). Reconstructing Potentials of Mean Force through Time Series Analysis of Steered Molecular Dynamics Simulations. *Journal of Computational Physics*, 151(1), 190–211. <https://doi.org/10.1006/jcph.1999.6218>
- [78] Dixon, R. A. F., Sigal, I. S., Rands, E., Register, R. B., Candelore, M. R., et al. (1987). Ligand binding to the  $\beta$ -adrenergic receptor involves its rhodopsin-like core. *Nature* 326:73-77
- [79] Straßer, A., & Wittmann, H. J. (2013). Molecular modeling studies give hint for the existence of a symmetric h $\beta$ 2R-G $\alpha\beta\gamma$ -homodimer. *Journal of Molecular Modeling*, 19(10), 4443–4457. <https://doi.org/10.1007/s00894-013-1923-8>
- [80] Süel, G. M., Lockless, S. W., Wall, M. A., & Ranganathan, R. (2003). Evolutionarily conserved networks of residues mediate allosteric communication in proteins. *Nature Structural Biology*, 10(1), 59–69. <https://doi.org/10.1038/nsb881>
- [81] Kenakin, T., & Miller, L. J. (2010). Seven Transmembrane Receptors as Shapeshifting Proteins: The Impact of Allosteric Modulation and Functional Selectivity on New Drug Discovery. *Pharmacological Reviews*, 62(2), 265–304. <https://doi.org/10.1124/pr.108.000992>
- [82] May, L. T., Leach, K., Sexton, P. M., & Christopoulos, A. (2007). Allosteric Modulation of G Protein–Coupled Receptors. *Annual Review of Pharmacology and Toxicology*, 47(1), 1–51. <https://doi.org/10.1146/annurev.pharmtox.47.120505.105159>
- [83] Kumar S, Ma B, Tsai CJ, Sinha N, and Nussinov R (2000) Folding and binding cascades: dynamic landscapes and population shifts. *Protein Sci* 9:10–19.
- [84] Monod J, Changeux JP, Jacob F. 1963. Allosteric proteins and cellular control systems. *J. Mol. Biol.* 6:306–29
- [85] Collier, G., & Ortiz, V. (2013). Emerging computational approaches for the study of protein allostery. *Archives of Biochemistry and Biophysics*, 538(1), 6–15. <https://doi.org/10.1016/j.abb.2013.07.025>
- [86] E. P. Sablin, I. N. Krylova, R. J. Fletterick, H. A. Ingraham, *Mol. Cell* 11, 1575 (2003)
- [87] Koshland DE (1958) Application of a theory of enzyme specificity to protein synthesis. *Proc Natl Acad Sci USA* 44:98–104.
- [88] Del Carmine, R., Molinari, P., Sbraccia, M., Ambrosio, C., & Costa, T. (2004). “Induced-fit” mechanism for catecholamine binding to the beta2-adrenergic receptor. *Molecular Pharmacology*, 66(2), 356–63. <https://doi.org/10.1124/mol.66.2.356>
- [89] Canals, M., Lane, J. R., Wen, A., Scammells, P. J., Sexton, P. M., & Christopoulos, A. (2012). A Monod-Wyman-Changeux mechanism can explain G protein-coupled receptor (GPCR) allosteric modulation. *Journal of Biological Chemistry*, 287(1), 650–659. <https://doi.org/10.1074/jbc.M111.314278>
- [90] Bridges, T. M., & Lindsley, C. W. (2008). G-protein-coupled receptors: From classical modes of modulation to allosteric mechanisms. *ACS Chemical Biology*, 3(9), 530–541. <https://doi.org/10.1021/cb800116f>
- [91] Christopoulos, A. (2002) Allosteric binding sites on cell-surface receptors: novel targets for drug discovery, *Nat. Rev. Drug Discovery* 1, 198–210
- [92] Christopoulos, A., and Kenakin, T. (2002) G protein-coupled receptor allosterism and complexing, *Pharmacol. Rev.* 54, 323–374.
- [93] Ehlert FJ. 1985. The relationship between muscarinic receptor occupancy and adenylate cyclase inhibition in the rabbit myocardium. *Mol. Pharmacol.* 28:410–21
- [94] Hall DA. 2000. Modeling the functional effects of allosteric modulators at pharmacological receptors: an extension of the two-state model of receptor activation. *Mol. Pharmacol.* 58:1412–23
- [95] Conigrave, A. D., & Franks, A. H. (2003). Allosteric activation of plasma membrane receptors - Physiological implications and structural origins. *Progress in Biophysics and Molecular Biology*, 81(3), 219–240. [https://doi.org/10.1016/S0079-6107\(03\)00020-8](https://doi.org/10.1016/S0079-6107(03)00020-8)

- [96] Hardy JA, Wells JA. 2004. Searching for new allosteric sites in enzymes. *Curr. Opin. Struct. Biol.* 14:706–15
- [97] Swain JF, Gierasch LM. 2006. The changing landscape of protein allostery. *Curr. Opin. Struct. Biol.* 16:102–8
- [98] Lockless, S. W. (1999). Evolutionarily Conserved Pathways of Energetic Connectivity in Protein Families. *Science*, 286(5438), 295–299. <https://doi.org/10.1126/science.286.5438.295>
- [99] Jolliffe, I. T. (2002). Principal Component Analysis, Second Edition. *Encyclopedia of Statistics in Behavioral Science*, 30(3), 487. <https://doi.org/10.2307/1270093>
- [100] Berendsen, H. J., & Hayward, S. (2000). Collective protein dynamics in relation to function. *Current Opinion in Structural Biology*, 10(2), 165–169. [https://doi.org/10.1016/S0959-440X\(00\)00061-0](https://doi.org/10.1016/S0959-440X(00)00061-0)
- [101] Manuscript, A. (2008). NIH Public Access. *Growth (Lakeland)*, 23(1), 1–7. <https://doi.org/10.1038/jid.2014.371>
- [102] Ma, J. (2005). Usefulness and limitations of normal mode analysis in modeling dynamics of biomolecular complexes. *Structure*, 13(3), 373–380. <https://doi.org/10.1016/j.str.2005.02.002>
- [103] Cui, Q., Li, G., Ma, J., & Karplus, M. (2004). A normal mode analysis of structural plasticity in the biomolecular motor F1-ATPase. *Journal of Molecular Biology*, 340(2), 345–372. <https://doi.org/10.1016/j.jmb.2004.04.044>
- [104] Keunwan Park, Dongsup Kim, *BMC Bioinf.* 12 (Suppl. 1) (2011) 1–9.
- [105] Nigar Kantarci-Carsibasi, Turkan Haliloglu, Pemra Doruker, *Biophys. J.* 95 (2008) 5862–5873.
- [106] Kei-ichi Okazaki, Shoji Takada, *Proc. Natl. Acad. Sci.* 105 (2008) 11182–11187
- [107] Del Carmine, R., Molinari, P., Sbraccia, M., Ambrosio, C., & Costa, T. (2004). “Induced-fit” mechanism for catecholamine binding to the beta2-adrenergic receptor. *Molecular Pharmacology*, 66(2), 356–63. <https://doi.org/10.1124/mol.66.2.356>
- [108] Lakkaraju, S. K., Lemkul, J. A., Huang, J., & MacKerell, A. D. (2016). DIRECT-ID: An automated method to identify and quantify conformational variations - Application to  $\beta$ 2-adrenergic GPCR. *Journal of Computational Chemistry*, 37(4), 416–425. <https://doi.org/10.1002/jcc.24231>
- [109] G. Milligan, *Mol. Pharmacol.* 66, 1 (2004)
- [110] Krystek, S. R., Kimura, S. R., & Tebben, A. J. (2006). Modeling and active site refinement for G protein-coupled receptors: application to the beta-2 adrenergic receptor. *Journal of Computer-Aided Molecular Design*, 20(7–8), 463–70. <https://doi.org/10.1007/s10822-006-9065-z>
- [111] Hanson, M. A., Cherezov, V., Griffith, M. T., Roth, C. B., Jaakola, V. P., Chien, E. Y. T., ... Stevens, R. C. (2008). A Specific Cholesterol Binding Site Is Established by the 2.8 Å Structure of the Human  $\beta$ 2-Adrenergic Receptor. *Structure*, 16(6), 897–905. <https://doi.org/10.1016/j.str.2008.05.001>
- [112] Pogozheva, I. D., Lomize, A. L., & Mosberg, H. I. (1997). The transmembrane 7- $\alpha$ -bundle of rhodopsin: distance geometry calculations with hydrogen bonding constraints. *Biophysical Journal*, 72(5), 1963–1985. [https://doi.org/10.1016/S0006-3495\(97\)78842-8](https://doi.org/10.1016/S0006-3495(97)78842-8)
- [113] Noda, K., Saad, Y., Graham, R. M., & Karnik, S. S. (1994). The high affinity state of the  $\beta$ 2-adrenergic receptor requires unique interaction between conserved and non-conserved extracellular loop cysteines. *Journal of Biological Chemistry*, 269(9), 6743–6752.
- [114] Wu, H., Wacker, D., Katritch, V., Mileni, M., Han, G. W., Liu, W., ... Stevens, R. C. (2012). Structure of the human kappa opioid receptor in complex with JD1c. *Nature*, 485(7398), 327–332. <https://doi.org/10.1038/nature10939>
- [115] Takada, K., Clark, D. J., Davies, M. F., Tonner, P. H., Krause, T. K. W., Bertaccini, E., & Maze, M. (2002). Meperidine exerts agonist activity at the  $\alpha$ 2B-adrenoceptor subtype. *Anesthesiology*, 96(6), 1420–1426. <https://doi.org/10.1097/0000542-200206000-00022>

- [116] Cordini, A., Ismail, S., Matsoukas, M. T., Escriuet, C., Gherardi, M. J., Pardo, L., & Fourmy, D. (2015). Functional elements of the gastric inhibitory polypeptide receptor: Comparison between secretin- and rhodopsin-like G protein-coupled receptors. *Biochemical Pharmacology*, 96(3), 237–246. <https://doi.org/10.1016/j.bcp.2015.05.015>
- [117] Frielle, T., Daniel, K. W., Caron, M. G., & Lefkowitz, R. J. (1988). Structural basis of beta-adrenergic receptor subtype specificity studied with chimeric beta 1/beta 2-adrenergic receptors. *Proceedings of the National Academy of Sciences of the United States of America*, 85(December), 9494–9498.
- [118] Warne, T., Serrano-vega, M. J., Baker, J. G., Moukhametianov, R., Edwards, C., Henderson, R., ... Schertler, F. X. (2010). Structure of a  $\beta 1$  -adrenergic G protein-coupled receptor. *Nature*, 454(7203), 486–491. <https://doi.org/10.1038/nature07101.Structure>
- [119] Daniel McNaught-Flores, M. a S.-U. (2013). Histamine Modulates Isoproterenol Efficacy at the  $\beta 2$  Adrenoceptor: Inferences Regarding Allosteric Modulation by Imidazole-Containing Compounds. *Biochemistry & Physiology: Open Access*, 2(3). <https://doi.org/10.4172/2168-9652.1000114>
- [120] Swaminath G, Lee TW, Kobilka B (2003) Identification of an allosteric binding site for Zn<sup>2+</sup> on the beta2 adrenergic receptor. *J Biol Chem* 278: 352-356
- [121] Liu, Y., Chen, B., & Wei, J. (2012). Computational studies of the binding modes of CCR1 antagonists, 7022(September), 37–41. <https://doi.org/10.1080/08927022.2012.679617>
- [122] Couvineau, A., Rouyer-Fessard, C., & Laburthe, M. (2004). Presence of a N-terminal signal peptide in class II G protein-coupled receptors: Crucial role for expression of the human VPAC1 receptor. *Regulatory Peptides*, 123(1–3 SPEC. ISS.), 181–185. <https://doi.org/10.1016/j.regpep.2004.06.025>
- [123] Cang, X., Yang, L., Yang, J., Luo, C., Zheng, M., Yu, K., ... Jiang, H. (2014). Cholesterol- $\beta 1$ AR interaction versus cholesterol- $\beta 2$ AR interaction. *Proteins: Structure, Function and Bioinformatics*, 82(5), 760–770. <https://doi.org/10.1002/prot.24456>
- [124] Deflorian, F., & Jacobson, K. A. (2011). Comparison of three GPCR structural templates for modeling of the P2Y<sub>12</sub> nucleotide receptor. *Journal of Computer-Aided Molecular Design*, 25(4), 329–338. <https://doi.org/10.1007/s10822-011-9423-3>
- [125] Abaffy, T., Malhotra, A., & Luetje, C. W. (2007). The molecular basis for ligand specificity in a mouse olfactory receptor: A network of functionally important residues. *Journal of Biological Chemistry*, 282(2), 1216–1224. <https://doi.org/10.1074/jbc.M609355200>
- [126] Kobilka, B. K. and Y. Zou (2016). GPCR fusion protein containing an N-terminal autonomously folding stable domain, and crystals of the same, Google Patents.
- [127] Hogan K., Peluso S., Gould S., Parsons I., Ryan D., Wu L., Visiers I. (2006) Mapping the binding site of melanocortin 4 receptor agonists: a hydrophobic pocket formed by I3.28(125), I3.32(129), and I7.42(291) is critical for receptor activation *J. Med. Chem.* 49, 911-922