University of South Bohemia in České Budějovice

Faculty of Science

Action at a Distance in Arginine Repressor and the Store-operated Calcium Channel Orai: Molecular Modeling and Simulations

PhD thesis

Saurabh Kumar Pandey

Supervisor: **Prof. Rudiger H. Ettrich, PhD**
University of South Bohemia České Budějovice
Center for Nanobiology and Structural Biology,
Institute of Microbiology,
Academy of Sciences of the Czech Republic

České Budějovice 2018

This thesis should be cited as:

Pandey, S.K. 2018: Action at a Distance in Arginine Repressor and the Store-operated Calcium Channel Orai: Molecular Modeling and Simulations. PhD Thesis Series, No 6, University of South Bohemia, Faculty of Science, České Budějovice, Czech Republic, 230 pp.

Annotation

In this thesis molecular modeling tools have been applied to investigate the phenomenon of allostery in two systems: Arginine repressor and human Orai channel. Arginine repressor protein binds to DNA in response to the intracellular concentration of L-arginine and controls the arginine metabolism in bacteria. Using molecular dynamics simulations sampling of the conformational space of arginine repressor and the allosteric effects of L-arginine binding on structure and dynamics of systems were studied. Orai channels are responsible for entry of calcium ion in the cell in response to Ca^{2+} depletion in the endoplasmic reticulum. Homology modeling was applied to prepare a structural model of human Orai3. Molecular dynamics simulations of the earlier published Orai1 model and the Orai3 model described here were performed. Results pointed to differences in the structures and dynamics of the two Orai isoforms, Orai1 and Orai3 and consequent effects on the channel which are probably responsible for the different behavior of Orai isoforms. A putative cholesterol binding site was identified using *in silico* docking approach and possible effects of cholesterol binding on the Orai1 channel structure and function were reported. Allosteric effects of mutations on a distant cholesterol-binding pocket was investigated.

Declaration [in Czech]

Prohlašuji, že svoji disertační práci jsem vypracoval samostatně pouze s použitím pramenů a literatury uvedených v seznamu citované literatury.

Prohlašuji, že v souladu s § 47b zákona č. 111/1998 Sb. v platném znění souhlasím se zveřejněním své disertační práce, a to v úpravě vzniklé vypuštěním vyznačených částí archivovaných Přírodovědeckou fakultou elektronickou cestou ve veřejně přístupné části databáze STAG provozované Jihočeskou univerzitou v Českých Budějovicích na jejích internetových stránkách, a to se zachováním mého autorského práva k odevzdanému textu této kvalifikační práce. Souhlasím dále s tím, aby toutéž elektronickou cestou byly v souladu s uvedeným ustanovením zákona č. 111/1998 Sb. zveřejněny posudky školitele a oponentů práce i záznam o průběhu a výsledku obhajoby kvalifikační práce. Rovněž souhlasím s porovnáním textu mé kvalifikační práce s databází kvalifikačních prací Theses.cz provozovanou Národním registrem vysokoškolských kvalifikačních prací a systémem na odhalování plagiátů.

České Budějovice .....

Saurabh Kumar Pandey

This thesis originated from a partnership of the Faculty of Science, University of South Bohemia and the Center for Nanobiology and Structural Biology, Institute of Microbiology of the Czech Academy of Sciences, supporting doctoral studies in the Biophysics study program.

Financial support

Financial support was provided by Czech Science Foundation (13-21053S).

Acknowledgement

I would like to thank my supervisor Prof. Rudiger H. Ettrich for providing me the opportunity to work on several interesting projects. It was a great pleasure to work in your supervision and to learn by seeing how every time you came up with new ideas and suggestions to tackle an issue. Thank you very much for giving me freedom along with guidance and always believing in me.

I would like to acknowledge Prof. Jannette Carey who always shared her ideas and brought her experience to the project . It was a great pleasure to work with you. Your advices and teachings always helped me.

I would like to thank our group members David, Babak, Daniel, Vasilina and Natalie with whom I worked on several projects and got to learn so many new things. Thank you for helping me during this time, for interesting discussions and for creating a friendly atmosphere at work. The memories of time we shared at work and on our trips for conferences will always cheer me up.

I would like to thank Milan who was always very generous to share his ideas and to help me during all these years. Working with you was always fun and very fruitful.

I would like to thank Dr. Jost Ludwig Jost with whom I worked and also shared the office. Thank you for all the interesting discussions about science as well as about politics and culture. It was always fun.

I would like to thank our collaborators from Linz for all the interesting discussions and for sharing their data and exciting results.

I would like to thank my friends Pavel, Pradeep, Deepika, Anja, Katja, Vitali, Julia, Avishek for all the memories and fun we had. You guys made it easier to be here and made it possible to enjoy the place along with work. Special thanks for all the late night discussions over “coffee”.

I would like to thank my friends Lalit, Manoj, Alok, Navneet, KK, Dharmendra, Abhigyan, Tapan, Rajeev and Sudhanshu who were always there to encourage me during ups and downs in my life. Without you guys it would not have been possible. I would like to thank my childhood friend Pradeep for all his support throughout my good and bad days. Thanks for being there with all the encouraging words, care and support.

I would like to thank my friends Valja, Gurkan and Zoltan for wonderful time and several memorable trips. You guys were always there to cheer and support me. I will always carry your friendship with me.

It is difficult to express my feelings in words when it comes to acknowledge my family. The love and support I got from you all were unmatched. Thank you for believing in me and encouraging me to learn and to follow my own path in life.

List of Papers and Author's Contribution

Papers included in thesis:

I. Pandey, S.K., Reha, D., Zayats, V., Melichercik, M., Carey, J., Ettrich, R. (2014) Binding-competent states for L-arginine in E. coli arginine repressor apoprotein. J Mol Model. 20: 2330.

SP performed molecular dynamics simulations, participated in analysis of results and writing of the manuscript.

II. Pandey, S.K., Melichercik, M., Reha, D., Carey, J., Ettrich, R. Allosteric activation of Arginine repressor protein by L-arginine. (manuscript)

SP participated in performing molecular dynamics simulations, analyzing the results and in writing of the manuscript.

III. Fahrner, M., **Pandey, S.K**., Muik, M., Traxler, L., Butorac, C., Stadlbauer, M., Zayats, V., Krizova, A., Plenk, P., Frischauf, I., Schindl, R., Gruber, H.J., Hinterdorfer, P., Ettrich, R., Romanin, C., and Derler, I. (2018) Communication between N terminus and loop2 tunes Orai activation. **293:** 1271–1285.

SP participated in performing homology modeling, molecular dynamics simulations, analyzing the results and writing of the manuscript.

IV. Derler I, Jardin I, Stathopulos PB, Muik M, Fahrner M, Zayats V, **Pandey SK**, Poteser M, Lackner B, Absolonova M, Schindl R, Groschner K, Ettrich R, Ikura M, Romanin C (2016) Cholesterol modulates Orai1 channel function. Sci Signal. 9 (412):ra10.

SP participated in analysis of results.

Pandey, S., Bonhenry, B., Ettrich, R.H. (2018) Biomolecular Simulations in Structure-Based Drug Discovery: Ion Channel Simulations. Willey Publications, Gervasio, Francesco L., Spiwok, Vojtech (Editor), First Edition, pp: 247-280.

ISBN:978-3-527-34265-5

SP participated in writing the manuscript.

Papers not included in thesis:

Zayats V., Stockner T., Pandey S.K., Worz K., Etrrich R., Ludwig J (2015). A refined atomic scale model of the Saccharomyces cerevisiae K^+-translocation protein Trk1p combined with experimental evidence confirms the role of

selectivity filter glycines and other key residues. Biochimica et Biophysica Acta (BBA)–Biomembranes, 1183-1195.
SP performed a subset of MD simulations, participated in analysis of results and writing of manuscript.

Preface

Since the time Monod and Jacob gave the term “allostery” in 1961 concept of allostery has evolved and found its importance in every aspect of biological systems. Recent applications of allostery in drug designing and discovery of allosteric drugs shows that it was not an exaggeration when Monod and Jacob called allostery as “second secret of life”. Allostery is essentially the regulation of activity at a distant site or phenomenon of action at a distance. Since the early phenomenological models, “concerted model” and “sequential model” a great progress in the thermodynamics, biophysical and computational techniques have occurred which has given us more detailed understanding of allosteric phenomenon and also broadened our view. As allostery in proteins is highly dynamic and complex process computational approaches such as molecular dynamics simulations have been proved very powerful tool to understand the mechanistic and dynamic details of allosteric processes. The minute details like changes in side chain orientation of protein residues as well as large conformational changes such as gating of ion channels to domain motions have been successfully investigated using molecular dynamics simulations. In this thesis two systems have been studied using bioinformatics and molecular modeling tools: Orai channels and Arginine repressor protein.

Arginine repressor (ArgR) protein is the master regulator of arginine-regulon and found in a variety of bacteria. It senses the intracellular concentration of L-arginine and sends feedback signal to DNA operator regions in order to regulate the biosynthesis of L-arginine. Despite of rich amount of structural, biochemical and biophysical data and information available about arginine biosynthesis, the mechanism by which binding of L-arginine to the core oligomerization-domain of ArgR allosterically affects the affinity of peripheral

DNA-binding domain to DNA remains unclear. Sequence alignment, phylogenetic analysis and molecular dynamics simulation methods were used here in order to understand the effects of L-arginine binding on structure and dynamics of ArgR and consequent allosteric activation.

We aimed to investigate following points:

- Describing the transition of system from apo to holo states, starting with the identification of L-arginine binding competent states.
- Allosteric activation of ArgR protein by L-arginine binding; is the trimer-trimer rotation a universal mode of action despite different drivers (residues) responsible for it in different bacteria?

Ca^{2+} are ubiquitously present in all types of cells and play an essential role as second messenger in a wide variety of functions ranging from cell proliferation to apoptosis. Here we studied human Orai proteins found in the plasma membrane. Orai proteins together with STIM protein of endoplasmic reticulum (ER) constitute Ca^{2+} release activated Ca^{2+} (CRAC) channels which are mainly responsible for store-operated Ca^{2+} entry (SOCE) into the cell. The mutations in Orai1 protein has been associated with various diseases such as severe combined immunodeficiency, muscular dystrophy, autoimmunity etc. The availability of Orai crystal structure in Drosophila and its high sequence similarity with human Orai proteins gave us a good platform to investigate the structural-functional properties of Orai channels using molecular modeling tools. All our studies of Orai channels reported in this thesis were done in close collaboration with Christoph Romanin group at Linz which gave us opportunity to investigate the system from theoretical as well experimental point of view. Our work was focused upon human Orai1 and Orai3 isoforms. We aimed to investigate:

- Structural and functional differences in Orai1 and Orai3: the role of ETON region in Orai1 and Orai3; why is a different length of ETON needed for channel function in these two isoforms.
- Effect of cholesterol on Orai1 structure and function: identification of cholesterol binding sites; allosteric effect of mutations in ETON region upon binding of cholesterol to full-length Orai1.

This thesis consists of three parts: Chapter 1 contains an introduction of the systems being studied and their importance, Chapter 2 contains a brief background of molecular modeling methods used in the thesis, Chapter 3 contains the summary of the work which is described in attached papers and also the additional results and discussion of unpublished data complementary to the papers and Chapter 4 contains conclusions of the thesis work.

LIST OF ABBREVIATIONS

AMBER	Assisted Model Building with Energy Refinement
ArgR	Arginine Repressor
BLAST	Basic Local Alignment Search Tool
CHARMM	Chemistry at HARvard Macromolecular Mechanics
CHL	Cholesterol
CRAC	Ca^{2+} Release-Activated Ca^{2+}
DAG	DiAcylGlycerol
EM	Energy Minimization
GROMACS	GROningen MAchine for Chemical Simulations
IP3	Inositol trisphosphate
MD	Molecular Dynamics
NMR	Nuclear Magnetic Resonance
PDB	Protein Data Bank
PIP2	Phosphatidylinositol 4,5-bisphosphate
PLC	Phospholipase C
POPC	1-Palmitoyl-2-oleoylphosphatidylcholine
SOCE	Store-Operated Ca^{2+} Entry
YASARA	Yet Another Scientific Artificial Reality Application

LIST OF FIGURES

TABLE OF CONTENTS

INTRODUCTION

1.1 Allostery: Second secret of Life

Allostery is the phenomenon of action at a distance i. e. binding of a ligand at a distant and non-overlapping site affects the binding of same or other type of ligands to other binding sites. The term “allosteric inhibition” was first coined by Monod and Jacob in 1961 (Monod and Jacob 1961) when they named the Novick-Szilard-Umbarger effect “the inhibitor is not a steric analogue of the substrate” and also explained feedback control mechanism by L-isoleucine in *Escherichia coli* threonine kinase (Changeux 1961).

1.1.1 Concerted and Sequential model

In 1965 Monod, Wyman and Changeux (Monod et al. 1965) defined a model known as Concerted model (also known as MWC model) where they emphasized on the preservation of symmetry during the conformational transition even in the partially liganded structure. This model assumes that protein samples two conformations tensed (T) and relaxed (R) even in the absence of effector; where the R subunits readily bind substrate than those in T subunits. The allosteric effector has different affinities towards these states so they pull the equilibrium to the state for which they have more binding affinity.

In 1935 Pauling proposed a model to explain positive cooperativity of oxygen binding to hemoglobin protein (Pauling 1935). In 1966 Koshland proposed a model (Koshland et al. 1966) known as Sequential or KNF model which was related to Pauling model. KNF model proposes the binding as a sequential process where binding of a ligand to protein monomer changes its tertiary structure. Allosteric signal from one binding site to others is transmitted via structure and it does not require necessarily an overall change in the quaternary structure of protein or expect it to be symmetric. Thus binding of a ligand to a monomer changes it from T to R states and changes the affinity of binding sites

in the other monomers, but without necessarily converting them to R states. KNF model explains both positive and negative cooperativity found in the protein-ligand binding.

1.1.2 Why symmetricity and cooperativity?

In their article Monod, Wyman and Changeux (Monod et al. 1965) emphasized on maintenance of symmetry and cooperativity between subunits of a quaternary structures. To understand the importance of this we should ask another question: why enzymes mostly prefer polymeric structures?. There are few points which may be considered as possible reasons for this, e.g. proper geometry of binding sites and stability of protein in terms of surface to volume ratio and protein specificity. The position of residues constituting binding sites should be fixed in space in a very precise and distinctive manner so as to provide the specificity to binding site. For this reason peptide chain should be long enough to provide specific position to residues and to provide them degrees of freedom to reorient themselves; as well as there should be enough length of peptide chain around the binding site to restrain those residues in the binding site.

These factors perhaps favor the formation of polymeric globular structures rather than long monomeric structures. A monomer may have many binding site residues which have affinity to interact with each other. If these monomers arrange themselves antiparallely in a protein, these residues can interact with each other. Here many combinations of these arrangements are possible, the other pairs get themselves distributed symmetrically around the diad axis formed due to the first two pairs . Now any mutation even in one residue would have an enhanced effect upon the geometry and/or functional properties of the protein in a cooperative and symmetric manner (Monod et al. 1965).

1.1.3 Assumptions and drawbacks of concerted and sequential models:

First assumption of concerted model was that allosteric proteins must be polymeric but now allosteric effect is known in many monomeric proteins. Ma and Cui studied a signaling protein, chemotaxis Y protein (CheY) in Escherichia coli, 129 residue long monomer (Ma and Cui 2007). Here they used transition path sampling (TPS) in combination with free energy (potential of mean force) simulations to understand coupling between different transition events at atomic resolution. Allostery definitely plays a role here and it is well explained on the basis of both classical models.

Other assumption about symmetricity and quaternary structure change is also not always true and it has been found that neither the change in quaternary structure is always necessary to manifest the allosteric effect nor the symmetricity is essentially conserved. Bacterial chaperone GroEL protein has been extensively studied using experimental and simulation methods which show that allosteric effect is due to fine tuned motion of individual domains rather than in whole heptamer (Ma et al. 2000; Vaart et al. 2004).

Negative cooperativity was explained in KNF model but not in MWC model. In a number of cases negative cooperativity has been found. Therefore in many systems such as Arginine repressor (ArgR) in *E. coli* allostery has been explained using both models. In *E. coli* ArgR even partially liganded states are symmetric in nature but experiments have shown negative cooperativity among subsequent binding events (Jin et al. 2005; Strawn et al. 2010).

Also, since both above models being phenomenological in nature if taken in strict sense obviously are not able to answer the question how the ligand binding brings the change in quaternary structure of proteins.

1.1.4 Modern views of allostery

Further development of new techniques/methods such as molecular dynamics simulations, NMR, FRET etc have opened the way to calculate thermodynamic, kinetic, structural and functional properties of systems.

In 1984, Cooper and Dryden proposed a “dynamics allostery” model where they proposed that even without a conformational change the allosteric communication between binding and allosteric sites is possible (Cooper & Dryden 1984). This model incorporates the entropic contribution due to thermal fluctuations in the system.

A “new view” or “population shift” model proposed by Nussinov group which states that a population of the activated conformers already exists even in the absence of ligand and binding of ligand shifts the equilibrium between activated and non-activated population towards the former (Kumar et al. 1999; Ma et al. 1999; Gunasekaran 2004; Swain and Gierasch 2006; Sol et al. 2009). However this “new view” is not so different than classical MWC model as population shift or shift of equilibrium was also proposed in MWC model (Cui and Karplus 2008).

1.1.5 Allostery in membrane proteins: signal transduction and diseases

The allostery phenomena has been successfully applied to study the signal transduction in numerous membrane proteins e. g. G-protein-coupled-receptor, voltage gated ion channels, ligand gated ion channels and nuclear hormone receptors (Changeux 2013) . The existence of two (or more) conformer populations as proposed in MWC model (or extension of MWC model) and shift in the equilibrium between these populations by ligand binding fits with the results in various studies (Changeux and Edelstein 2005).

Many mutations lead to opening of ion channels even in the absence of ligands

and these mutations has been assigned responsible for various diseases (Changeux 2012). MWC model have been applied in these ion channels where a mutation lead to the open state (similar to R state) and shifts the equilibrium between the open (R) and close (T) states towards the former. The phenomenon of “action at a distance” can be applied to analyze the signal transduction here.

Classically a drug designing approach was to look for a lead molecule which can either compete or mimic the natural ligand and bind to the active (orthosteric) site at their target protein. With the discovery of several molecules which bind at the allosteric site away from the orthosteric sites and bring the changes in physiological activity (Taly et al. 2009; Changeux 2013). This opened a new approach to the drug design where the lead molecules or drugs could be designed to modulate the protein function such as closing/opening of ion channel via binding at allosteric sites.

1.2 Arginine repressor

Arginine repressor (ArgR) protein is responsible for regulation of genes of arginine regulon (Maas 1994). It senses the intracellular arginine concentration and sends feedback signals to arginine regulon and thus controls the biosynthesis and metabolism of L-arg. It is a very simple, elegant and ancient example of allostery. ArgR has been extensively studied by various experimental and theoretical methods thus genetic, biochemical and structural data and information is available for ArgR from many bacteria (Cunin et al. 1986; Maas 1994).

1.2.1 Biosynthesis of arginine in bacteria

Biosynthesis of arginine in bacteria is a topic of great research interest due to its more complicated metabolic pathway than others (Cunin et al. 1986).

Biosynthesis of arginine starts from glutamate. Glutamate is first converted into ornithine via N-acetylated compounds in five steps. Conversion of ornithine to arginine takes place in three steps and carbamoyl phosphate is utilized during this process. In the first step of this pathway glutamate is acetylated into N-acetyl glutamate (Vyas and Maas 1963) which is then phosphorylated and reduced into N-acetyl semialdehyde (Vogel & McLellan 1970). Transamination of N-acetyl semialdehyde results into N^2-acetyl ornithine. After this the removal of acetyl group follows two different pathways in different bacteria:

1) Linear pathway where acetyl group is removed from N^2-acetyl ornithine by the enzyme acetylornithinase. This pathway is used in Enterobacteriaceae, Baceillaceae etc. In *E. coli* this linear pathway is followed and enzyme N-acetyl glutamatesynthatase is the target enzyme for feedback inhibition (Vyas and Maas 1963). (Figure 1.1)

2) Cyclic pathway where transacetylation of N^2-acetyl ornithine and glutamate occurs by acetyltransferase enzyme. In cyanobacteria, pseudomonas etc. this cyclic pathway is followed. Since in these organisms N-acetyl glutamate is produced during transacetylation, the second enzyme of the ornithine biosynthesis pathway N-acetyl glutamate 5-phosphotransferase is the target enzyme for feedback inhibition (Udaka 1966).

Thus we see that the arginine feedback regulation has adapted the target enzyme according to the arginine biosynthesis pathway (Cunin et al. 1986, Udaka 1966). Formation of arginine from ornithine takes place in three steps. In the first step carbamoyl moiety of carbamoyl phosphate is transferred to ornithine, forming citrulline by OTCase enzyme. In the next step citrulline converts to argininosuccinate accepting amino group of aspartate, and finally arginine is synthesized (Lu 2006; Lehninger et al. 2008).

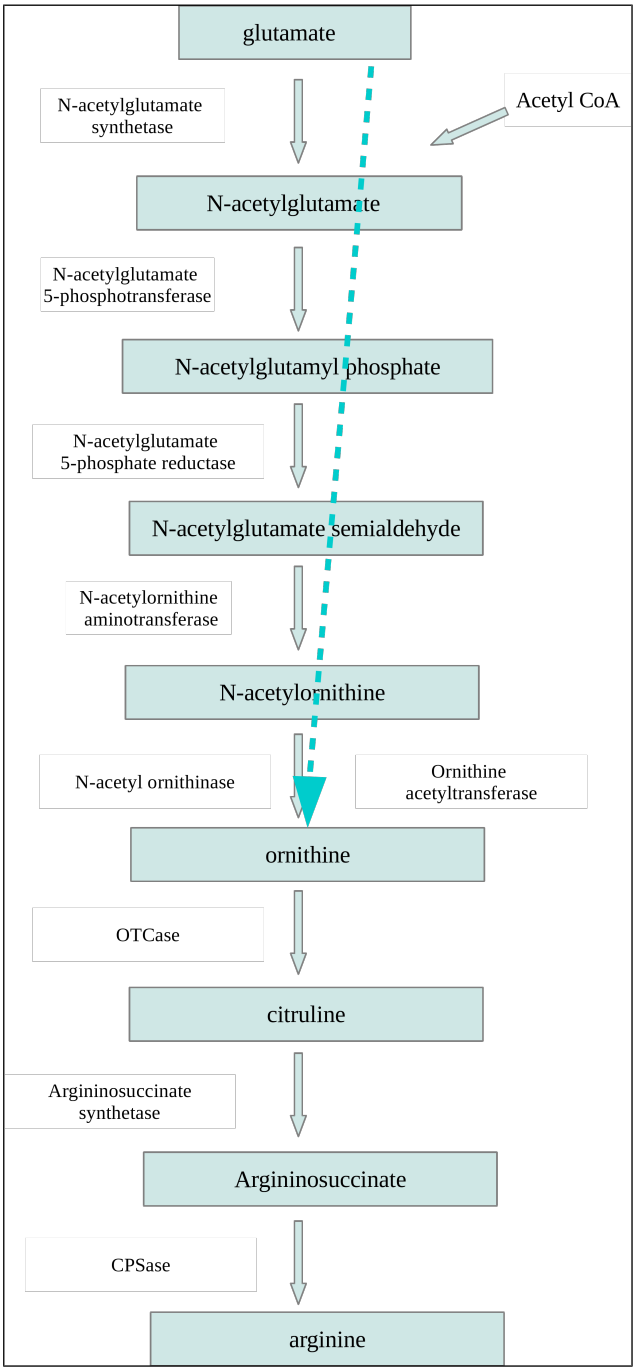

Figure 1.1: Arginine biosynthesis pathway in *E. coli* (Cunin et al. 1986).

1.2.2 Control of gene expression in *E. coli*

History of arginine regulation and expression system goes back to 1950s and the term 'regulon' was coined by Werner K. Maas and John Clark to describe the group of multiple genes controlled by a single regulatory gene (Maas 1994). The model of arginine regulon came at the same time when Jacob and Monod proposed their famous “operon model” which changed the way of thinking of scientists about gene regulatory system of biochemical pathways. In independent studies Maas and Vogel both found that concentration of arginine affects the synthesis of enzymes of ornithine metabolic pathway. Vogel gave the name ‘repressor’ to the product of the gene which can control a group of genes (Vogel 1957, Maas 1994).

In *E. coli* arg regulon there are 12 genes, organized into nine transcriptional units. The product of gene ArgR controls the synthesis of more than 10 enzymes and also its own synthesis (Tian et al. 1992). The complete sequences as well as control regions of these genes are known. Each transcriptional unit contains an operator region overlapping the promoter sequence. The operator contains two palindromic sequences 18 base pairs long, known as “arg box” separated by three nucleotides in structural genes and by two nucleotides in regulatory gene ArgR (Cunin et al.1983). These arg boxes share a common consensus sequence and at least one arg box in an operator overlaps the promoter region (Cunin et al.1986; Lim et al. 1987). The binding of ArgR to the arg boxes results in repression response due to affecting RNA polymerase binding. The level of repression response depends on number of Arg boxes (Cunin et al. 1986). Studies suggested when only one Arg box (argH) is present the repression is lower in comparison to genes with two boxes (argECBH, argF and argI). Also the strongest response was observed for genes where the dyad symmetry was extended which make the overlap with promoter stronger (Cunin et al.1983).

Mutations in the right hand side of argECBH arg box which are neither conserved nor symmetrical also affect repression-derepression response. Thus it was concluded that number, symmetry as well as sequence of arg box affects the binding specificity and affinity of ArgR (Cunin et al.1983). In the mutational experiments in argECBH when 1 base pair was deleted between two Arg boxes the repression response was lowered (Cunin et al. 1986). It was later confirmed by a number of experiments such as gel retardation and in vitro DNA footprinting that the repressor molecule binds simultaneously to both arg boxes of the operator region and the binding affinity is lowered in case of single arg box (Tian et al. 1992). The binding of ArgR in E. coli and other repressors is similar, it binds in one face of DNA double helix and contacts with both major and minor grooves of control region of arg genes (Tian et al. 1992; Sunnerhagen et al. 1997, Wang et al 1998). These facts suggest that probably binding of ArgR to adjacent boxes at DNA might be cooperative (Cunin et al. 1986).

1.2.3 Arginine repressor (ArgR): Structure

ArgR is a hexameric protein where each monomer consists of N-terminal and C-terminal domains which are of basic and acidic nature respectively (Maas 1994). The C-terminal domains form the core of the hexamer and are responsible for oligomerization and L-arginine binding while the N-terminal domains form the periphery of ArgR hexamer and are responsible for DNA binding (Figure 1.2) (Dennis et al. 2002, Sunnerhagen et al. 1997; Burke et al. 1994; Tian and Mass 1994). The ArgR polypeptide chains of B. subtilis and E. coli share very low sequence identity between them, only 19% and 35% in N-terminal and C-terminal part respectively with a 27% overall sequence identity (Czaplewski et al., 1992). Despite of this very low sequence identity structure

and functional division between N- and C-terminal domains remains conserved (Smith et al. 1989).

Crystal structure of ArgR from different bacteria are available: ArgR in apo state (PDB: 1F9N) and ArgR C-terminal domain (ArgRC) in holo state (PDB: 2P5M) from *Bacillus subtilis* (Dennis et al. 2002, Garnett et al. 2007); ArgR in apo state (PDB: 1B4A) and ArgR-C terminal domain in holo state (PDB: 1B4B) from *Bacillus stearothermophillus* (Ni et al. 1999); ArgRC in apo form (PDB: 3BUE) and holo form (PDB: 2ZFZ) from *Mycobacterium tuberculosis* (Cherney et al 2008); ArgRC in apo (PDB: 1XXC) and holo form (PDB: 1XXA) (Duyne et al. 1996) and ArgR N-terminal domain in *E. coli* (PDB: 1AOY) (Sunnerhagen et al. 1997).

N-terminal domain belongs to winged helix-turn-helix (wHTH) family. It comprises a three α-helix bundle followed by a wing composed of two β-strands and loops (Sunnerhagen et al. 1997). The structural scaffold of N-domain is very similar in E. coli (PDB: 1AOY) and B. subtilis (PDB: 1F9N).

C-terminal domain consists of four stranded β-sheet and two α-helices thus belonging to α/β fold family. In holoEcArgRC (PDB:1XXA) six L-arginine molecules bind at the trimer-trimer interface. Each L-arginine forms eight hydrogen bonds with two of the three subunits of the same trimer and two hydrogen-bonds with a subunit of opposing trimer (Duyne et al. 1996).

The crystal structures of ArgRC from *E. coli* in apo and holo state are very similar showing the RMS of 0.76 Å based upon superposition of α-carbon atoms (Figure 1.3) (Duyne et al. 1996). The crystal structures of the protein in *B. stearothermophilus* in apo form (apoBstArgR) and C-terminal domain in holo form (holoBstArgRC) show the rotation of 15 degrees about the trimer-trimer interface (Ni et al. 1999). This rotation was inferred as an effect of L-arginine binding to the core of protein and this rotational motion is supposed to

be further transmitted to DNA binding N-terminal domain (Ni et al. 1999). Structure of arginine repressor in *Bacillus subtilis* and *Mycobacterium tuberculosis* (Figure 1.4) also show such rotation and suggest that this rotation might be present across all the species and necessary for DNA binding (Cherney et al 2008; Cherney et al 2009; Cherney et al. 2010; Dennis et al. 2002). Crystal structures of intact ArgR in both apo and holo forms in the same organism is unavailable till date therefore direct comparison of the effect of L-arginine binding on N-terminal domain is not possible.

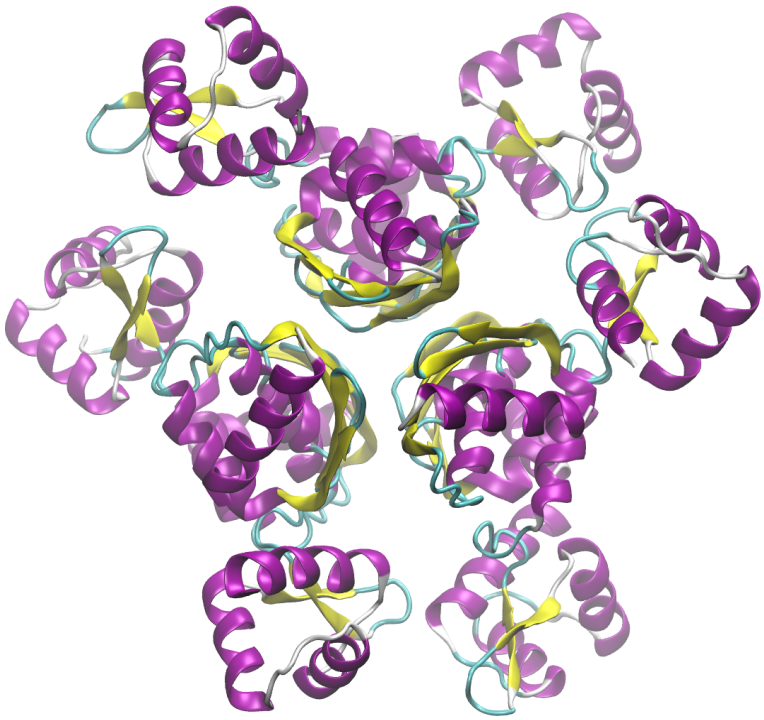

Figure 1.2: Cartoon representation of apo-arginine repressor from *Bacillus subtilis*: top view (PDB: 1f9n). β-sheet and α-helices and β-strands are shown in vilet and yellow colors respectively.

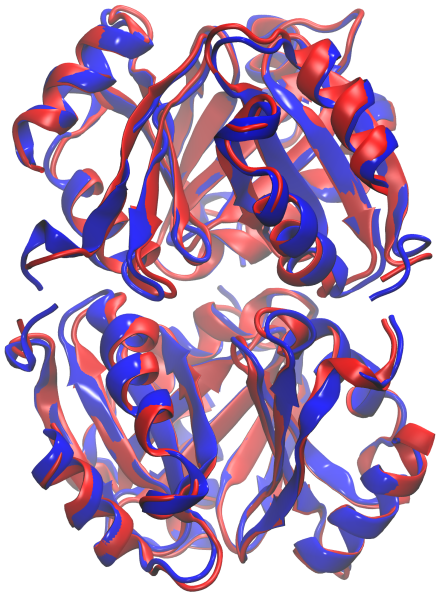

Figure 1.3: Comparison of crystal structures for C-terminal domain of ArgR in E. coli in apo (blue) and holo (red) states.

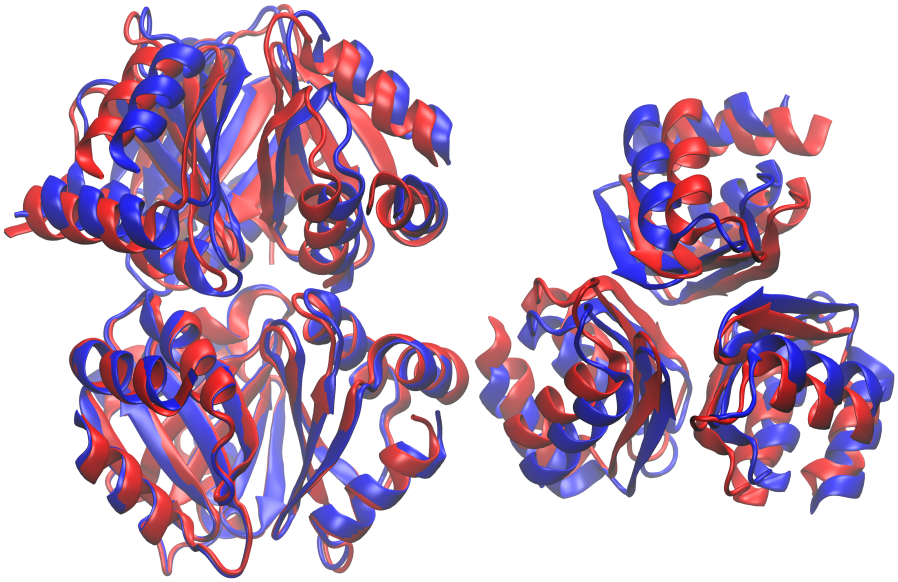

Figure 1.4: Rotation in crystal structures for ArgR C-terminal domain in Mycobacterium tuberculosis in apo (blue) and holo (red) states: side view (left panel) and top view (right panel).

The results from isothermal titration calorimetry (ITC) experiments for quantitative analysis of L-arginine binding to EcArgR and EcArgRC showed that the binding of first L-arginine is ~100 folds stronger than subsequent binding events. Based on multiphasic isotherms 1+5 model for L-arg binding was proposed where first event of binding causes a global conformational change in the hexamer (Jin et al. 2005). The molecular dynamics simulations performed on EcArgRC apo and holo structures showed rotation of trimers about the interface in the absence of L-arginine and a salt bridge between R110 and D128 was shown responsible for rotation (Strawn et al. 2010). Further presence of L-arginine molecule was shown to be sufficient in order to stop this rotation.

Despite of the above knowledge and computational, experimental and structural data available, very little is known about the mechanism of ArgR function, relation between L-arginine binding to its core region and response of this to the DNA binding at periphery. The study of ArgR system using computational approach here will certainly help us to gain an insight of structural details of ArgR system at atomic level, conformational landscape of ArgR and to understand the mechanism of ArgR activation by L-arginine binding. Hopefully the information gained here would also help to understand allosteric processes in other proteins.

1.3 Ion channels

Cells are functional and structural unit of an organism. Cells separate themselves from outer environment by a plasma membrane (PM). In order to communicate and to exchange the necessary organic or inorganic compounds with outer environment cells possess ion channels in their plasma membrane. Ion channels play an important role in a wide range of functions such as signal transmission by regulating membrane potential, exchange of ions across the membrane, muscle contraction, homeostasis etc (Lipscombe and Toro 2013). Ion channels allow the passage to ions across the membrane passively down their concentration gradient and hence differ from ion transporters which allow active transport using energy.

The movement of ions through the pore of the channels occurs in multiple steps, first the ions reach to the channel by diffusion then very often it is attracted by oppositely charged residues of the channel near the pore opening, further movement of ions usually is checked by selectivity filter; after passing this point movement of ion is regulated by gating due to the pore lining residues of the channel (Catterall 1995).

Ion channels usually allow the certain type of ions based on size or charge etc. thus are selective for ions (Maffeo et al. 2012). Many highly K^+-selective ion channels are primarily responsible for maintaining the membrane potential at resting stage of the cell (Ketchum et al. 1995).

The opening or closing (called gating) of ion channels is governed by various stimuli based on which ion channels can be classified such as voltage-gated, light gated, mechanosensitive, light gated etc.

Voltage-gated ion channels respond to change in membrane potential. Large number of these channels are found in neuronal cells where these participate in signal transmission (Catterall 1995). These channels are highly selective and

might differ from each other based upon their mode of activation, gating properties and inactivation time period (Purves et al. 2004). Some channels such as Na^+-selective channels in squid axon open and close very quickly whereas some in mammalian axons do not get inactivated quickly thus maintaining the action potential for a longer time period (Hille 2001).

Ligand gated ion channels respond to the extracellular or extracellular chemical signals such as neurotransmitters or second messengers such as cAMP, cGMP, Ca^{2+}, PIP2 etc (Tovar and Westbrook 2012). These channels receive the chemical signal and convert it into electrical information leading to certain event. Ligand gated ion channels participate in sense of smell, light, pain etc.

1.3.1 Store operated calcium entry (SOCE) and Ca^{2+} release-activated Ca^{2+} (CRAC) channels

Ca^{2+} ions are of vital importance in organisms, taking part in functions ranging from sub-cellular to inter-cellular processes (Berridge et al. 2000). Ca^{2+} are involved in protein synthesis, gene expression, cell proliferation, muscle contraction, neurotransmitter release, synaptic signaling etc (Berridge 2001; Clapham 2007). The intracellular concentration of Ca^{2+} ranges in nanomolar and extracellular concentration stays in millimolar range. Subcellular organelles like endoplasmic reticulum (ER) serve as reservoirs of Ca^{2+} in the cell with a Ca^{2+} concentration in micromolar (μM) range (Demaurex and Frieden 2003; Raffaello et al. 2016). In order to maintain this difference which is important to carry out vital functions, a wide variety of calcium channels are present in plasma membrane and intracellular organelles (La Rovere et al. 2016).

There are two ways by which Ca^{2+} enter the cell: store operated Ca^{2+} entry (SOCE) (Parekh & Putney 2005, Smyth 2010) and receptor operated Ca^{2+} entry (ROCE) (Clapham 2007; Putney 1986). These two mechanisms work

independently. In ROCE the calcium channels in plasma membranes are activated by second messengers. In store operated channels the entry of calcium ions into cell is mediated by ER membrane proteins and channels in plasma membrane.

Cell surface receptors once bound to agonists activate the phospholipase C (PLC) which in turn cleaves PIP2 into 1,4,5-trisphosphate (IP3) and DAG. The DAG then induces Ca^{2+} entry into the cell and IP3 induces Ca^{2+} release from ER via IP3 receptor (IP3R). Store operated channels (SOCs) are activated by this depletion of Ca^{2+} in the ER lumen, hence get the name (Parekh & Putney 2005, Smyth 2010).

First, Putney in 1986 proposed that influx of Ca^{2+} was due to Ca^{2+} depletion in ER rather than due to IP3 (Putney 1986). This process of refilling the cytoplasm via ER was named capacitative Ca^{2+} entry (CCE) which was later renamed as SOCE since it was later shown that Ca^{2+} enters the cytoplasm from extracellular space (Putney 1990). The idea proposed by Putney was later supported when thapsigargin (TG) was shown to block the Ca^{2+} pumps and depleted the stores of Ca^{2+} and this depletion was followed by influx of Ca^{2+} (Takemura 1989). Using patch clamp experiments Hoth & Penner (Hoth and Penner 1992) showed a Ca^{2+} selective current activated by store depletion of Ca^{2+}. The channels involved in this were named as Ca2+ release-activated Ca^{2+} (CRAC) channels (Hoth & Penner 1992). The main characteristic feature of CRAC channels is its activation by stored Ca^{2+} depletion. CRAC channels are extremely selective for Ca^{2+} over Na^{+} with a permeability ratio of 1000:1 (Hoth 1995; Prakriya and Lewis 2015). The single channel conductance of CRAC channels is 100 to 1000 folds smaller than other channels (Prakriya 2006).

1.3.2 STIM-ORAI complex

STIM (stromal interaction molecule) proteins are single pass membrane proteins localized mostly in ER (Roos et al 2005). STIMs are highly conserved and are expressed in all cell types (Roos et al. 2005; Soboloff et al. 2012). STIMs contain many protein interaction motifs: Signal peptide (SP), EF-hand, sterile alpha motif (SAM), CRAC activation domain (CAD), inactivation domain (ID), proline serine rich domain (P/S), EB1 binding domain (EB) and polybasic domain (PBD) (Stathopulos 2008; Stathopulos 2013; Yang et al. 2012). N-terminal part which includes SP, EF-hand and SAM domains is situated on luminal side. C-terminal part containing coiled-coiled domain, CAD, P/S, EB and PBD domain lies on the cytoplasmic side. A single pass transmembrane domain connects the luminal and cytoplasmic domains.

The luminal EF-SAM domain is responsible for detecting the change in concentration of Ca^{2+} in ER (Liou 2005; Roos et al 2005). It consists of cEF, nEF and SAM domain. At rest (normal concentration of Ca^{2+} in ER) cEF hand binds Ca^{2+} while nEF-hand stabilizes cEF. Together these two EF hands stay in open conformation and create a hydrophobic interaction network with SAM domain resulting into a compact monomeric structure of EF-Sam domain (Stathopulos 2008; Zheng 2011). The release of Ca^{2+} from cEF-hand results into unfolding of EF-SAM domain and further resulting into conformational changes in luminal and cytoplasmic domains. As a consequence of this unfolding and conformational changes hydrophobic part of EF-SAM domain is exposed and favors the dimerization or higher order oligomerization of STIMs (Stathopulos et al. 2008; Prakriya & Lewis 2015).

Cytosolic part of STIM contains CAD which is also known as STIM-Orai activation region (SOAR) (Part et al. 2009). The CAD comprises coiled coils domains CC2 and CC3 which are responsible for the binding and activating

Orai channel.

During the resting stage of cell Orai and STIM are distributed throughout the plasma membrane and ER respectively. The depletion of $[Ca^{2+}]_{ER}$ leads to redistribution and accumulation at ER-PM junctions where Orai is activated by STIM binding (Soboloff et al 2012). This process of SOCE activation is very unique as it requires the contacts between proteins from two different membranes.

In vertebrates two homologous forms of STIMs are found: STIM1 and STIM2 with a sequence similarity of (Collins and Meyer 2011). STIM1 has higher affinity for Ca^{2+} and it responds to large scale depletion of $[Ca^{2+}]_{ER}$. On the other hand STIM2 with its lower affinity for Ca^{2+} can respond to small changes in $[Ca^{2+}]_{ER}$ even when cell is at resting stage with normal $[Ca^{2+}]_{ER}$.

The defect in CRAC channel activity was assigned responsible for SCID (severe combined immunodeficiency) syndrome in human (Feske et al. 2005; Partiseti et al. 1994). Using modified linkage analysis and RNAi screening identified mutated gene and named the protein as Orai1 (Feske et al. 2006). In human three homologs of Orai: Orai1-3 are found. Three homologs (Orai1-3) share a sequence identity of ~62% over whole polypeptide chain and ~92% in transmembrane region (Prakriya and Lewis 2015). Orai proteins are unique in the sense that though conserved among themselves they share no significant sequence similarity with other ion channels known.

In Drosophila only one form of Orai is found. Crystal structure of Orai in Drosophila (dOrai) is available (Hou et al. 2012). The crystal structure shows six subunits arranged around a central axis to form a six fold symmetric hexamer in the transmembrane region (Figure 1.5).

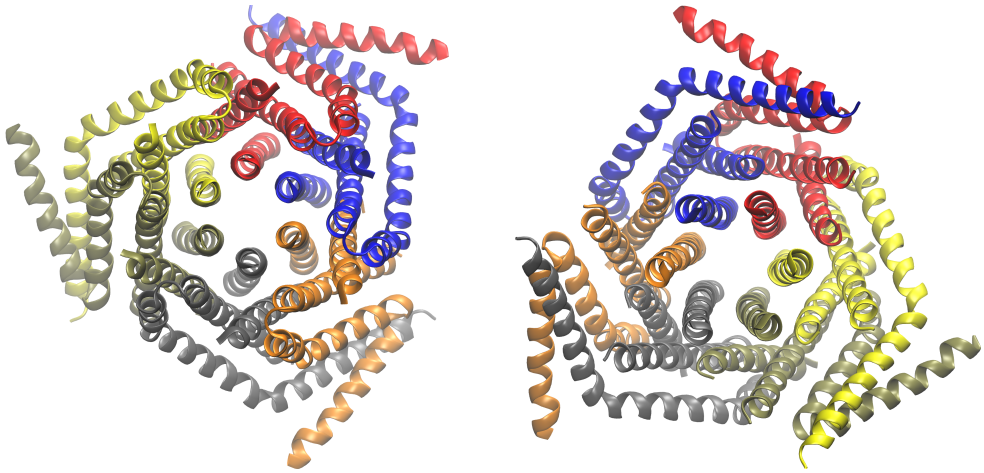

Figure 1.5: Extracellular view (left) and intracellular view (right) of dOrai channel showing six subunits in different colors.

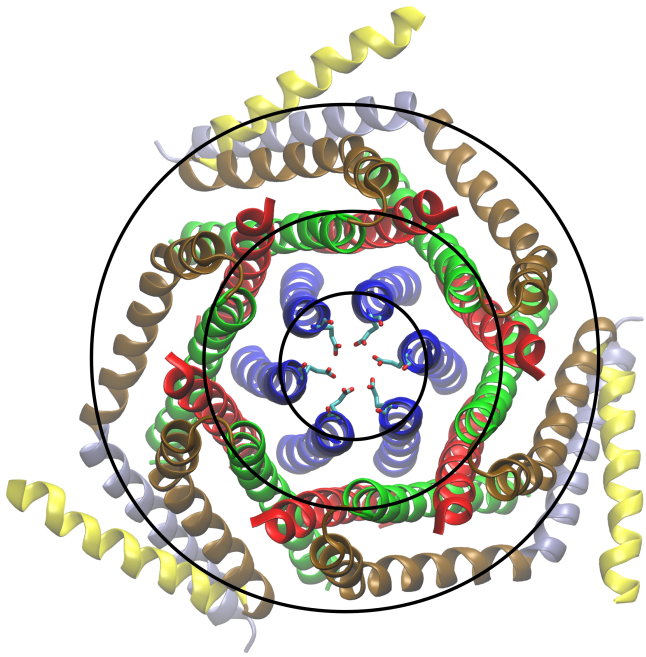

Figure 1.6: transmembrane helices 1-4 arranged in three concentric ring around the central axis going through the pore in dOrai channel. Glu178 residues on the extracellular side of the channel are shown as sticks.

Each subunit comprises four transmembrane helices (TM1-TM4) and an TM4 extension helix which is cytosolic. TM4 extension helices form a pair where one of the two helices lies just below and almost parallel to the plasma membrane and other helix of the pair goes deeper in the cytosol. Thus taking TM4 extension helices into account the hexamer is trimer of dimers (two subunits) and has a three fold symmetry. The TM helices arranged in three concentric rings: TM1 helices forming the pore of the channel, TM2 and TM2 helices forming the middle ring and TM4 helices forming the outer ring. N- and C- terminal of each subunit lies in the cytoplasmic side (Figure1.6).

Based upon the nature of pore lining TM1 residues the pore of Orai channel can be divided into four parts: selectivity filter region containing glutamate ring on the extracellular side, hydrophobic region, basic region and cytosolic region. The pore radius differs in different region of the channels (Figure 1.7).

STIM1 have been proposed to bind Orai at C-terminal TM4 extension helices and this C-terminus is necessary for STIM-Orai coupling (Li et al. 2007; Muik et al. 2008; Park et al. 2009). The binding at C-terminus causes a conformational change in Orai and allows Orai-TM1(N-terminus) - STIM interactions and finally leading to opening of the Orai channel (Amcheslavsky et al 2015; Hou et al. 2012).

The N-terminus of human Orai has also been identified as binding partner of STIM and necessary for CRAC activity. The cytosolic extension of TM1 helices known as extended transmembrane Orai NH2-terminal (ETON) region contains several residues mutations of which resulted into loss of function (Derler et al. 2016). ETON region is conserved in Orai proteins, however different length of ETON regions are necessary in Orai1 and Orai3 (Bergsmann et al. 2011; Derler et al. 2013). The reason for different behavior and importance of ETON for coupling/gating in these two Orai-isoforms is still not very clear.

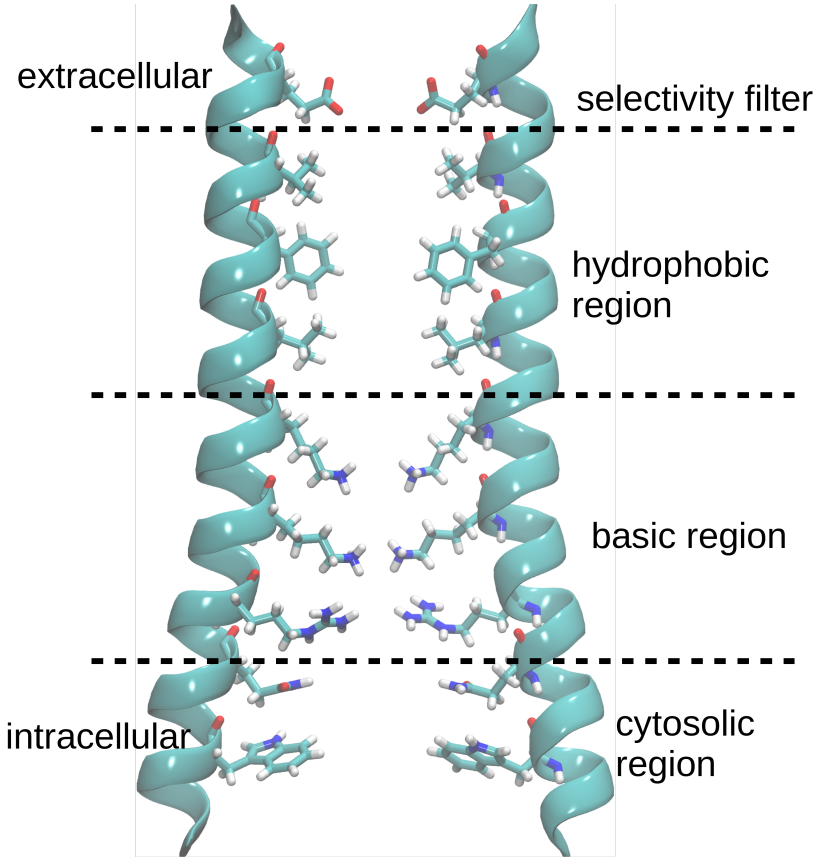

Figure 1.7: Pore diagram of dOrai showing pore lining residues (as sticks) and four regions in the pore.

dOrai sequence is 73% identical to human Orai1 (hOrai1) sequence. Thus dOrai crystal structure gives a promising platform to study the hOrai1 using molecular modeling methods. Recently hOrai1 model was prepared by homology modeling method using dOrai structure as a template (Frischauf et al. 2015). Here, molecular dynamics simulations of human Orai1 model helped to identify

a calcium-accumulating region (CAR) at the extracellular side of the pore opening. The CAR increases the Ca^{2+} permeation by enhancing the accumulation of Ca^{2+} near the pore opening which is necessary in the tissues where extracellular Ca^{2+} concentration is low (Frischauf et al. 2015).

Since Orai and STIM are distributed throughout in all type cells in human and are involved in various cellular and subcelluar functions through Ca^{2+}, absence of activity or structural defects in these proteins lead to various diseases (Prakriya 2015). Loss of STIM/Orai function has been associated with lower I_{CRAC}, impaired cytokine production, B-cells proliferation etc. These functional abnormalities result into various immunodeficiencies, dental enamel calcification defects, pulmonary diseases and congenital nonprogressive myopathy etc (Berna-Erro 2012).

MATERIALS AND METHODS

An overview of applications of computational methods to study different properties of ion channels have been given in **Paper V**. Various MD simulation methods are available to study ion-conduction, selectivity and permeation of channel, free-energy profile of channels along the pore, channel gating, etc. A comparison of such popular methods, their applications as well as limitations have been illustrated with examples of different ion-channels and membrane proteins. These methods are also applicable to study any bimolecular systems in general.

An introduction of the computational methods used in thesis and the principles behind these methods have been given in the following part.

2.1 Homology modeling

In order to study the structure, function and mechanism of action of a protein, it is necessary to have a reliable three dimensional structure of protein. Currently X-ray crystallography, nuclear magnetic resonance (NMR) and electron microscopy (Rawson et al. 2016) are the main techniques to determine the structures of macromolecules. X-ray crystallography is the major techniques through which most of the biological macromolecules structures have been solved in high resolution and majority of protein structures deposited in Protein Data Bank (PDB) have been solved by this method (Shi 2014). A bottle neck of this technique is obtaining the proteins in highly pure form and to get the crystals which might be time consuming and difficult in some cases. For example determining the structure of membrane protein through X-ray crystallography is challenging and requires additional work because of the water insolubility of these proteins which makes it difficult to get the pure sample and then to get the crystals (Shi 2014). Despite of recent development of

techniques such as electron crystallography or neutron diffraction (Blakeley 2008) The determination techniques are alone not sufficient and fast enough which is evident if we compare the size of proteome database and that of solved structures. Here computational methods to predict the 3D-structures of proteins come into play and have proved very helpful in studying structure-function relationship.

Homology modeling or comparative modeling is one of these computational approaches. It is based upon the assumption that similar amino acid sequences tend to form similar three-dimensional structures (Richardson 1981). This indeed becomes obvious when we see that over 100,000 known protein structures belong to merely ~500 folds (Chothia 1992, Zhang 1998).

The very first and important step in structure prediction using homology modeling approach is to find a suitable template with a known 3D-structure. Usually this could be obtained by using protein-BLAST (Altschul 1990; Altschul 1997) tool against PDB database. A template structure with a reasonable sequence identity to target sequence is necessary to obtain a reliable structural model as it minimizes the risk arising in automated sequence alignment and also gives more confidence that two structures (target and template) belong to the same protein fold. Sequence identity of >30% has been widely accepted as a confidence point in various studies (Chothia 1986; Chung 1996). If there are multiple templates available having a common 3D-structure pattern, multiple sequence alignment can be better as it will resolve the small errors in local secondary structure alignments and it will also give reliability to the chosen templates if they have overlap in the conserved regions (Lushington 2015). The resolution and R-factor of template structure should also be taken into account as the obtained model is dependent on quality of template structure used.

After selecting the template next step is the sequence alignment of target to that of template. One needs to carefully check few points such as gaps within the region corresponding to secondary structure elements, position of cysteine and proline residues in alignment, border of transmembrane segments, alignment of conserved segments, etc (Lushington 2015). Alignments can be manually shifted according to the requirement and gaps in the loop region or solvent exposed parts and usually should not affect the results. Once the sequence alignment is done, next step is to build the initial model. There are many approaches to do this which mostly differ in modeling the backbone structures (Lushington 2008). Fragment based modeling approaches identify the non-gapped regions in the alignment and models backbone structure for those fragments; or in some cases first prepare a consensus backbone structure for all templates and the torsional angles are obtained from backbone fragment libraries after which the target sequence is modeled based on the consensus structure. This is followed by loop modeling and side chain modeling (Sternberg 1996; Sutcliffe et al. 1987).

Yasara (Krieger et al. 2002) software used in this thesis adopts fragment based approach. It is based on the fact that the side chain rotamers depend upon the backbone conformations and therefore prefer certain orientations relative to φ and ψ dihedral angles of backbone (Chinea 1995). For loop modeling, Yasara retrieves the loop conformations from non-redundant sequences from PDB (Wang & Dunbrack 2003; Michalsky 2003). Segments are extracted based upon the two anchoring points of loop sequence and then scoring of these conformations obtained is done based upon sequence identity, structural quality, etc. After these loops are anchored using cyclic coordinate descent algorithm (Canutescu & Dunbrack 2003), final best conformations are selected using graph theory algorithm for predicting side chain conformations of loop as well

as the rest of the protein scaffold (Canutescu et al. 2003).

In restraint based approach the model is built by satisfying special restraints such as distances, angles, dihedral angles, pairs of dihedral angles, chirality and other structural features. The restraint based modeling has been implemented in a very popular tool: Modeller (Sali & Blundell 1993).

Once the structural model of template has been obtained, it is necessary to evaluate the quality of the model. There are several tools available which essentially check the quality based on deviation of bond length, angle, or dihedral in target structure to that of known structures. Few tools such as Procheck (Laskowski et al. 1993), ProSa (Wiederstein & Sippl 2007), WHATIF (Vriend 1990) and Anolea (Melo 1997) are very popular.

2.2 Molecular dynamics

Since the first application of molecular dynamics (MD) method to study the liquid system by Alder and Wainwright in 1957 (Alder & Wainwright 1957) and later by Rahman and Stillinger (Rahman & Stillinger 1971), MD has emerged as a very powerful tool to study the structure, dynamics and function of complex biological systems such as of proteins, DNA, lipids, small organic and inorganic molecules at atomic level resolution (Adcock & McCammon 2006).

The very first molecular dynamics simulation of protein (bovine pancreatic trypsin inhibitor) performed by McCammon showed that proteins are not rigid structures and rather had fluid like interiors where atoms are always in motion (McCammon et al. 1977). Structures derived from X-ray crystallography or nuclear magnetic resonance (NMR) show a static picture or a few snapshots of a molecule, and are unable to provide mechanistic and dynamic details of a system (Shi 2014). Molecular dynamics covers and provides details about the smallest events in the structure of biomolecules like interactions and

conformational changes of amino acid side chains to changes like domain motions and protein folding and unfolding. It has been successfully used to study the protein-ligand interactions, protein folding problems, ion-channels and metabolic pathways, free-energy landscapes, allostery and many more biological phenomena occurring in a cell (Zhuravlev and Papoian, 2010).

In general there are two main families of simulation methods to study of molecular systems: Monte Carlo (MC) simulation and molecular dynamics (MD) simulation.

Monte Carlo simulations are based on stochastic approaches for generating new set of configurations of a system (Earl & Deem 2008). Here an initial set of configurations of particles in a system is taken and a change in the configurations is brought by a MC move. An acceptance criterion is applied to either accept or reject these changes in the configurations. Here the acceptance criterion makes sure that the new generated configurations are in a particular statistical mechanics ensemble and have been assigned proper weight. Thus a large number of configurations are generated and at the end of simulation average of the calculated property of interest is obtained. MC simulations are straightforward and since there is no need to solve Newton's equation of motion the MC moves can be chosen or modified accordingly in order to generate the new trial configurations within the particular ensemble desired. (Earl & Deem 2008; Schlick 2011). In classical molecular dynamics simulation, Newton's equation of motion are applied to a set of interacting particles to evaluate the time evolution of system (Lindahl 2015). The atoms are considered as balls and covalent bonds are considered as harmonic springs connecting those balls. Here an initial set of coordinates (positions) of interacting particles are taken and Newton's equation of motion F=m.a is resolved in order to calculate the new positions and momenta of particles and this procedure is iterated for a definite

period of time based upon the thermodynamic property of interest. Thus trajectories generated here by iterating the calculations represent time evolution of system which is an advantage of MD over MC simulations (Adcock and McCammon 2006). The forces between particles and potential energies are derived using interaction functions known as molecular mechanics force fields (Guvench and Alexander 2008).

2.2.1 Force fields

A force field is essentially a set of equations and atomic parameters which helps to describe all necessary interactions between particles in the system (Lindahl 2015). The equations here evaluate the potential energies and then their derivatives, force on particles i at a given time as:

$$F_i = \nabla V_i \qquad \text{(eq. 1)}$$

where V is potential energy and F is force.

The particles in a system can interact in two ways, either via bonded interactions or non-bonded interactions. Thus potential energy of a system can be described as follows:

$$V_{system} = V_{bonded} + V_{nonbonded} \qquad \text{(eq. 2)}$$

Bonded interactions

There are four types of bonded interactions: bond stretching, bond angle, dihedral (or torsion) angle and improper dihedral (Guvench & Alexander 2008).

The bonds here are treated as springs with a force constant K_b, bond length equilibrium and the energy required in order to stretch or compress the bond is calculated using Hooke's law as:

$$\boxed{V_{bond} = \sum^{bonds} K_b (b - b_0)^2} \qquad \text{(eq. 3)}$$

where K_b is the force constant representing stiffness of the bond, b is bond length at a given time and b_0 is the equilibrium bond length.

The potential energy (V_{angle}) required to bend the angle between a triplet of atoms from its equilibrium value (θ_0) is also treated harmonically and can be written as :

$$V_{angle} = \sum^{angles} K_\theta (\theta - \theta_0)^2 \qquad \text{(eq. 4)}$$

where K_θ is stiffness of the bond angle and θ and θ_0 are angle at given time and equilibrium (angle) respectively.

Energy required to rotate a bond is modeled via torsion or dihedral interaction potential ($V_{torsion}$) of the following form

$$V_{torsion} = \sum^{torsions} K_\chi (1 + \cos(n\chi - \delta)) \qquad \text{(eq. 5)}$$

where K_χ is barrier height parameter, χ is dihedral, n is multiplicity and δ is phase. Cosine of the value is taken because rotation is periodic.

The improper dihedral term is used to maintain planarity, chirality or to define out of plane geometry. This is also a 4-body interaction. The proper dihedral (or torsional) angle potential is calculated on four consecutive atoms where atom 1 and 4 are separated from each other by three bonds. The improper dihedral interaction term depends upon three atoms around a fourth central atom. The improper dihedral energy term can be also modeled using harmonic potential:

$$V_{impropers} = \sum^{impropers} K_\varphi (\varphi - \varphi_0)^2 \qquad \text{(eq. 6)}$$

where K_φ is force constant for improper dihedral and φ_0 is the equilibrium angle.

Non-bonded interactions

Non-bonded interactions are calculated between atoms which are not connected via covalent bonds. Non-bonded interactions include Coulomb interactions and

van der Waals interactions.

Van der Waals interaction includes an attraction and a repulsion term and can be represented as Lennard-Jones interaction combining these terms as:

$$V_{vdW} = \sum^{nonbonded} 4\varepsilon_{ij}\left(\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right) \qquad \text{(eq. 7)}$$

where ε_{ij} is potential minimum for interaction between i and j atoms, σ_{ij} is the separation distance corresponding to zero energy and r_{ij} is distance between atoms. Here the first and second interaction terms are for repulsion and attraction respectively. As the atoms start to come closer the 6-term which describes model dispersion, comes into play and since it is negative term the energy becomes more attractive. Once the atoms come to a distance corresponding to energy minimum, further decrease in distance leads to the dominance of 12-term and increase in the energy which results into repulsion. The values of parameters σ^{12} and σ^{6} depend upon the atom types (Guvench & Alexander 2008).

Electrostatic interaction energy (V_E) between charged atoms is described by Coulomb law. This term describes effect of atomic charge and distance between atoms. The opposite and same charges attract and repel each other respectively. The electrostatic interaction energy are calculated as

$$V_E = \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \qquad \text{(eq. 8)}$$

where q_i and q_j are charges on atoms i and j, ε_0 is the permittivity of vacuum and r_{ij} is the distance between atoms.

Most popular force fields for molecular dynamics simulation of biomolecules are CHARMM (Mackerell et al. 1998; Mackerell et al. 2004), AMBER (Cornell

at al 1995; Duan et al. 2003; Lindorff-Larsen et al. 2010) , GROMOS (Oostenbrink et al 2004) and OPLS (Jorgensen et al. 1996). These force fields have above functional form for bonded and non-bonded terms in common but differ in many ways such as use of improper dihedral, combination rule for Lennard-Jones parameters or different scaling of 1-4 non-bonded interactions. All the force fields use the parameters which are optimized to reproduce the data produced by crystallographic, spectroscopic, thermodynamics and quantum mechanics calculations. However different force fields might use different subset of target data for the parameter optimization and thus have slightly different values of adjusted parameters as well (Guvench & Alexander 2008).

2.2.2 Numerical integration algorithms

Several integration algorithms such as Verlet (Verlet 1967), velocity Verlet (Swope 1982) or leap-frog (Hockney 1970) have been developed to solve Newton's equations of motion. In all these algorithms the position, velocity and acceleration can be written as Taylor series expansion. Verlet algorithm is one of the most popular algorithm and velocity Verlet and leap-frog are based on it. These algorithms have been implemented in the Gromacs tools used in this thesis work.

Verlet algorithm calculates new positions of an atom by using *a)* the position and acceleration at time point t and *b)* position at a previous time point t-δt as:

$$r(t+\delta t) = 2r(t) - r(t-\delta t) + (\delta t)^2 a(t) \qquad \text{(eq. 9)}$$

where t is time, δt time step, r is position and a is acceleration.

Since it requires the position from a previous step r(t − δt), it is not self-initiating and requires other methods to calculate the position at previous time.

Another disadvantage is that Verlet algorithm does not calculate velocities simultaneously with positions and needs next position in order to calculate the velocity later on.

In velocity Verlet algorithm the position, velocity and acceleration at next positions (t+δt) are calculated simultaneously as follows:

$$r(t+\delta t) = r(t) + v(t)\delta t + \frac{1}{2}a(t)(\delta t)^2 \qquad \text{(eq. 10)}$$

$$v(t+\delta t) = v(t) + \frac{1}{2}(a(t) + a(t+\delta t)).(\delta t) \qquad \text{(eq. 11)}$$

Thus we see that position at next step depends on position, velocity and acceleration at current step. The velocity at next step depends on velocity and acceleration at current step and acceleration at next step.

In leapfrog algorithm first the velocity at next half time is calculated from previous half-step velocity and acceleration at current time. Using this new half velocity and current position at time t, position at next step (t+δt) is calculated.

$$v\left(t+\frac{1}{2}\delta t\right) = v\left(t-\frac{1}{2}\delta t\right) + \delta t.a(t) \qquad \text{(eq. 12)}$$

$$r(t+\delta t) = r(t) + \delta t.v\left(t+\frac{1}{2}\delta t\right) \qquad \text{(eq. 13)}$$

Thus, here velocity and position are not calculated at same time rather alternate each other by time step.

The advantage of leap frog is that only v (t – 1/2(δt)) and r(t) needs to be stored and rest can be calculated based on these two. A disadvantage is that since the positions and velocities are not calculated at same time, kinetic and potential energies are not available at the same time. Therefore total energy for a given time point can not be calculated directly (Adcock & McCammon 2006).

2.3 Energy minimization

The structures obtained through experimental or theoretical methods sometimes contain overlapping atoms, distorted bond angle or bond length or simply occupy an unfavorable position. These structures represent a higher energy state and thus are not stable. The molecular dynamics simulations starting with these structure might crash due to a higher force originating due to some unfavorable or abnormal interactions. Energy minimization is used in such cases to remove these potential clashes from the system (Adcock & McCammon 2006; Leach 2001). Several algorithms are available which refine the structure with a goal of finding a set of coordinates corresponding to minimum potential energy. As the potential energy is a function of set of atomic coordinates, a system with N atoms will have a potential energy function of 3N variables. Thus finding a global minimum is not straightforward and it is a nonlinear optimization problem. Several popular algorithms such as steepest descent, conjugate gradient or Newton raphson (Allwright 1976; Leach 2001) are used in molecular modeling which can be classified based on the derivatives order used. Steepest descent and conjugate gradient are first order derivative methods and use gradient of potential energy to find out the next point and finally a path on the potential energy surface leading to a minimum energy point. Steepest descent method searches for reduced potential energy by directly calculating force on an atom and adjusts the coordinates accordingly, thus this approach is considered physically meaningful. Here, initially a change in coordinates is brought by step size of arbitrarily defined value and potential energy of new structure is calculated via new forces due to the changed coordinates. The step sizes are increased or decreased accordingly and this procedure is iterated. Steepest descent method is very robust and capable to find nearest local minima

but unable to find the global minima where potential energy surface is irregular and have several local minima. Thus it is useful in removing the clashes or small distortions which requires small changes in the structure.

Newton-Raphson is second order derivative method and it relies on the assumption that potential energy is quadratic to variables in the system in the region of energy minima. This method is more suitable for small molecules and has difficulties with biological macromolecules where the potential energy surface has many local minima and system is non quadratic. Sometimes it is useful to run steepest descent before applying Newton-Raphson method (McCammon & Harvey 1987; Adcock & McCammon 2006).

2.4 Protein-ligand docking

In silico protein-ligand docking have been successfully applied to study the protein-ligand interactions and found its applications in search for lead molecules in drug design (Grinter & Zou 2014; Wishart 2008). In order to know structural and functional effects of binding of small molecules (ligand) to a protein of interest it is important to know the structure of protein-ligand complex (Xie & Hwang 2015). Since the progress in experimental structure determination techniques like X-ray crystallography, NMR, etc. a large number of protein structures have become available, however sometimes determination of the protein-ligand complex structure via these experimental techniques could prove to be a very difficult and time demanding task (Shi 2014); also due to recent advancement in computational methods for protein structure prediction, as well as computational power, more and more protein structure models become available every year. In these situations, *in silico* docking methods come into play by predicting the structure of protein-ligand complexes which was a difficult task otherwise (Grinter & Zou 2014).

Docking procedure consists of two parts: predicting the correct binding poses using search algorithms and evaluation of conformations using scoring functions (Pagadala 2017).

In order to search for the most probable conformation of protein-ligand complex all the degrees of freedom should be searched. Usually ligand is kept flexible and protein is kept rigid in the first step, it is called rigid docking (Wishart 2008). There are two methods to search the binding site: local search and global search. In case where no information is available about the possible binding pocket the global algorithm is used which takes whole protein structure into account in order to search. Once the binding pocket has been predicted local search can be applied by keeping ligand as well as side chains of few residues of protein as flexible (Pagadala 2017).

Once the binding poses (probable conformations and orientations) of ligand have been predicted the evaluation of these binding poses is done using force-field based, empirical or statistical potentials are used (Grinter & Zou 2014).

Yasara software used for docking in this thesis implements AutoDock-Vina which uses iterated local search global optimizer algorithm for prediction binding poses and semiempirical scoring function for evaluation and ranking of binding poses (Trott & Olson 2010).

3. RESULTS AND DISCUSSION

3.1 Arginine repressor

Arginine repressor works as master regulator of genes of arginine regulon in various bacteria. It senses the concentration of L-arginine (L-arg) in the cell and sends feedback signals to DNA operators in order to control the L-arg metabolism. Here we studied the ArgR system in *E. coli* and *B. subtilis* using molecular modeling methods. The results have been divided into two parts: first section (**3.1.1**) describes the study of ArgRC in *E. coli* and L-arginine binding competent states; second section (**3.1.2**) describes the study on ArgR full-length structure of *B. subtilis* and its activation by DNA binding

3.1.1 Arginine repressor of E. coli: Study of C-terminal domain using molecular modeling

Crystal structures of C-terminal domain of arginine repressor in *E. coli* (EcArgR) are available in apo (without L-arginine; apoEcArgRC) and holo (with six molecules of L-arginine; holoEcArgRC) forms. The crystal structures for apoEcArgRC and holoEcArgRC are almost identical and thus provide no information about the structural effect of L-arginine binding on the C-domain and how the binding of L-arginine results into the activation of protein. The previous molecular dynamics simulation studies using GROMOS87 force field have shown a rotational movement of trimers over each other and presence of one L-arginine molecule was shown to stop this rotation.

This prompted us to study this system using molecular dynamics using more updated force field. The aim here was to simulate the conformational transition of conformer populations from apo to holo states and to investigate the effect of L-arginine binding upon this transition. MD simulations of EcArgRC in apo and holo forms were performed using amber99sb force field. Visual analysis of trajectories revealed a rotation of trimers over the trimer-trimer interface. The

principal component analysis of apoArgRC trajectories showed a global rotational motion of the trimers. Structural alignment of two extreme conformations along first eigenvector showed 18° of rotation. The same analysis for holoArgRC showed rotation of only 8° between two extremes. In order to quantify and analyze these rotational motions with respect to simulation time, trimer rotations were calculated based upon center of mass of monomers. Thus rotations reported here represent the global motions instead of noise resulting from small fluctuations in distances or angels. The analysis of equilibrated part of trajectories revealed mean value of rotation of 4.5° and 8° for holoArgRC and apoArgRC trajectories respectively (rotation in crystal structures was taken as 0° and values were normalized). In the equilibrated part of trajectories both populations show rotational oscillations about these mean values but apoArgRC populations are more broadly distributed than holoArgRC. Interactions between R110 and Asp128 from opposite trimers were observed in apoArgRC and a clear correlation between number of these interactions and rotation was found. Since in holoArgRC Asp128 is involved in binding with L-arg therefore no such interaction was possible. During the apoArgRC simulations hydrogen bond interactions between Arg110 and Asp128 are broken time to time and thus binding pocket is open. This gave us an opportunity to dock a molecule of L-arginine in the open binding pocket and to study the dynamic behavior of the system using MD simulations. Successful docking and stable MD simulations of singly ligated systems showed that apoArgRC conformers observed and used for docking represented binding competent states. The analysis of singly ligated systems showed that the after equilibration of ~85 to ~95 ns the population of conformers oscillates around mean value of 4° rotation similar to holoArgRC simulations. Thus system provides the open binding pockets time to time and binding of one ligand brings

the system from rotated (apo-like) conformations to unrotated (holo-like) conformations.

Thus this study helped to identify two different rotational states for apoEcArgRC and holoArgRC which differ from each other by ~4° to ~5° degrees. In apoArgRC formation and release of Arg100-Asp128 salt bridges over the trimer-trimer interface governs the rotational oscillations. Occasional release of these salt bridges opens the L-arginine binding pocket. The slow temporal evolution of apoEcArgRC rotational behavior helped to correlate the rotation with the Arg100-Asp128 interactions. The sampling of binding competent states, effect of L-arginine docking on the rotational dynamics and conformational transitions from apo-like to holo-like states were observed here for the first time.

The details of this work are described in **Paper I.**

3.1.2 Arginine repressor of B. subtilis : phylogenetic analysis and molecular dynamics simulation study

In the section **3.1.1** results from study of C-terminal domains of ArgR in *E. coli* (EcArgRC) using molecular were discussed. The results showed the differences in behavior of apo and holo structures where conformers from these two forms occupy different rotational states thus show the effect of L-arginine binding. However the effect of L-arginine binding and signal transduction from C-domains to N-domains could not be studied due to unavailability of full length EcArgR structure. Phylogenetic analysis of all the available amino acid sequences revealed three major branches of ArgR proteins. ArgR from Mycobacterium tuberculosis (MtArgR) represents the oldest branch whereas EcArgR and BsArgR belong to separate younger branches. MD simulations of

EcArgR showed that R110 residue was responsible for rotation by forming salt bridge across the trimer-trimer interface. This LIAR--D motif containing the charged R110 residue is conserved throughout in the branch containing EcArgR. The MtArgR does not have this motif and has a serine residue at the position corresponding to R110(in EcArgR) however this region is highly conserved in M. tuberculosis branch and has a R133 just one helical turn away which plays a similar role as in E. coli R110 (Strawn et al. 2010). BsArgR do not possess LIAR--D motif and no charged residue near the binding pocket which can interact via saltbridges across the trimer-trimer interface, unlike in EcArgRC or MtArgRC. MD simulation of apo and holo BsArgR were performed using AMBER99SB-ILDN for 2 μs. The resulted structures for apo and holo forms acquire different rotational states which differ by ~9° to ~10° degrees from each other similar to that observed in E. coli (~5°).

MD simulations showed inter-trimer interactions between K75/R78 and D82 residues from linker helix and these interactions are crucial for rotation. A motif K/R--R---D/E containing these charged residues was identified which is present and conserved in MtArgR branch and also in the BsArgR branch. The hydrogen bond interaction analysis showed that overall number of interactions between N and C-terminal domains are less in holoBsArgR than in apoBsArgR. The entropy calculations for holoBsArgR and apoBsArgR surprisingly showed same values of entropy. However the N-domains in holoBsArgR had higher entropy which was compensated by lower entropy of C-domains. Indeed C-domains in holoBsArgR have restricted movement due to presence of L-arg in the binding site as well as due to the interactions between linker helices.; this momentum is then transferred to the peripheral N-domain.

These results together represent that different residues are responsible for rotation in B. subtilis than in E. coli however the rotation is similar and

probably inherent in both ArgR. The rotational motion due to binding of L-arginine and other interactions in the C-domain gets transferred to N-terminal domain promoting the DNA-binding competent states. Thus the L-arginine binding allosterically activates the ArgR protein via rotational motion of trimers which is inherent in ArgR despite of different amino acid composition.

The details of this work are described in **Paper II.**

3.2 Orai channel

Orai and STIM proteins are main components of CRAC channel and responsible for major Ca^{2+} entry into the cell. STIMs are mainly found in endoplasmic reticulum (ER) membrane whereas Orai are anchored in plasma membrane (PM). Depletion of Ca^{2+} in ER leads to conformational changes, redistribution and multimerization of STIMs at ER-PM junctions and STIM1-Orai binding and opening of Orai channel. The results have been divided into two parts: first section (**3.2.1**) describes the study of Orai1 and Orai3 and importance of loop2; second sections (**3.2.2**) describes the study of Orai1-cholesterol complex.

3.2.1 Communication between N terminus and loop2 tunes Orai activation

The CRAC channel gating involves the binding of STIM1 C-terminus to cytosolic N- and C-termini of Orai (Frischauf et al. 2008; Muik et al. 2008; Navarro-Borelly et al. 2008). The C-termini of Orai is the main binding partner however N-termini also have been found necessary for Orai channel function (Zheng et al. 2013; Mullins et al. 2009; McNally et al. 2013; Derler et al. 2013; Zhou et al. 2013). The approximately 3-4 helical turn long extension of TM1 into cytosol called extended transmembrane Orai N-terminus (ETON) region is fully conserved in Orai1 (73-90) and Orai3 (48-85). Despite of this highly conserved ETON region several mutational studies suggest that Orai1 and Orai3 require different length of ETON region in order to be store operated active (Derler et al. 2013; Bergsmann et al. 2011). Orai3 tolerates approximately one and half turn (5 residues) longer truncation compared to Orai1 in order to retain store-operated activation. Our aim here was to study the reasons for different behavior of these two isoforms and to find the differences in structural requirement for gating and function. Orai1 N-terminal truncation

mutant (Orai1 $\Delta N_{1–76/78}$) results into loss of function whereas analogues Orai3 truncation mutant (Orai3 $\Delta N_{1–51/53}$) remains STIM-dependent active. Electrophysiology experiments showed that swapping Orai1-loop2 with Orai3-loop2 restores function in Orai1 N-truncation mutants up to the deletion of first half of ETON region. The loop2 in Orai1 and Orai3 are ~75% identical. Mutations of 5 residues in Orai1-loop2 to those in Orai3-loop2 showed STIM1-dependent activation as observed for the above chimeras where loop2 were swapped. In order to check the effect of loop2 in both isoforms in the absence of STIM1, several mutants which are constitutively active were selected. Truncation of N-terminus in these mutants also resulted in loss of function and the swapping of loop2 restored the function, therefore loop2 effects were similar to wild types truncated mutants and chimeras.

Orai-loop2 with Orai3-loop2. Orai1 and Orai3 hexamer structures were modeled by homology modeling using dOrai structure as a template and molecular dynamics (MD) simulations were performed using OPLS force field. MD simulations of Orai1 and Orai1 $\Delta N_{1–76/78}$ truncation mutant showed that truncation of N-terminus creates space and loop2 moves toward it thus occupying different position than compared to that in Orai1 full-length structure. This also results into interactions between N-terminal residue Y80 residues and N156 residues of loop2 in Orai1 truncation mutant. Simulations of Orai3 showed that TM2 helices are ~2 helical turns longer in Orai3 therefore flexible portion of Orai3-loop2 is shorter (aa 133-138) than in Orai1-loop2 (aa 151-162). Due to the shorter flexible portion of loop2, no interaction between N-terminus and loop2 was observed in Orai3 truncation mutants during MD simulations. Molecular dynamics simulations together with experimental data suggested that loop2 in Orai1 N-terminus truncation mutants works in an

inhibitory manner and leads to non-permissive conformations of Orai channel but not in Orai3-truncation mutants. We predict that interactions between loop2 and Orai1 N-terminus results into masking the STIM1 coupling sites which are located either at N-terminus or loop2 or both; thus loop2 either directly or allosterically affects the STIM1-Orai coupling in an isoform specific manner. Thus along with the requirement of conserved ETON region a fine tuning between N-termini and loop2 conformations has an essential role in maintaining permissive conformation of channel and store-operated function of Orai channel.

The details of this work are described in **Paper III.**

3.2.2 Cholesterol modulates Orai1 channel function

By using cholesterol depleting agents cholesterol oxidase and filipin to disintegrate cholesterol-containing membranes, an increase in store-operated Ca^{2+} entry (SOCE) and CRAC currents was observed. The cholesterol depletion in RBL cells also showed an increase in CRAC currents and increased degranulation. These effects indicated a modulatory role of cholesterol in Orai1 channel function and prompted us to search for a cholesterol binding motif in primary sequence of Orai1.

A cholesterol recognition/interaction motif [Cholesterol binding (CB) motif] was identified in the ETON region of Orai1-TM1 helix. The amino acid sequence from residues 74 to 83 (LSWRKLYLSR) fulfills the criterion of being a cholesterol binding consensus motif-[L/V]-(X) 1–5 -Y-(X) 1–5 -[K/R] (Epand 2008). Mutations of residues L to I and Y to S has been known to destroy the CRAC motif function ((Epand 2008; Li & Papadopoulos 1998). Two mutants Orai1-L74I and Orai1-Y80S were prepared in order to check the necessity of this region for cholesterol binding. The experimental results showed that in both

mutants an increase in Ca^{2+} entry and currents were observed without any significant change in STIM-Orai1 coupling and amount of Orai1 protein in plasma membrane.

Furthermore, the application of cholesterol depletion agents on these mutants showed no effects on Ca^{2+} entry or currents which indicated that cholesterol affects the Orai1 channel function which is disturbed by mutation of L74 and Y80 residues.

In order to check the binding of cholesterol to the ETON region, fluorescence-binding studies were performed on a synthetic peptide of this region (aa 72-90) as well as mutant peptides L74I and Y80S. Results suggested that cholesterol binds to this peptide in vitro and cholesterol binding is 11 to 18 folds weaker for both peptide mutants in comparison to wild type peptide.

(The details about experiments have been given in **Paper IV**.)

3.2.2.1 Molecular dynamics simulation of peptide-cholesterol complex

In order to investigate the cholesterol binding to N-terminal peptide (complementary to the in vitro binding studies described in **Paper IV**) we performed MD simulations of the Orai1 N-terminal peptide in the presence of cholesterol, simulating buffer conditions as used in the in vitro measurements. The structure of this peptide (aa 72 to 90) was excised from the published human Orai1 model (Frischauf et al. 2015). The peptide was solvated in SPC water containing 2 Vol% ethanol and charges were neutralized by adding Na^{+} ions to the system. MD simulation was performed for 100 ns using OPLS-AA (Jorgensen et al. 1996) force field.

Noteworthy, the helical structure of the peptide started to unfold partially after a few nanoseconds, leading to a drastically reduced α-helical content after 20 ns.

Evaluation of the secondary structure profile using DSSP tool (Touw et. Al

2015; Kabsch & Sander 1983) revealed ~75-80% loss of α-helicity during the simulation (Fig 3.1).

Moreover, in these simulations, the peptide repeatedly sampled a conformation, where L74 and Y80 were oriented to the same side, distinct to their orientation in the ETON region of the full-length dOrai1 crystal structure. To examine for interaction of this Orai1 NT peptide with cholesterol, a cholesterol molecule was initially placed in close proximity (7 Å) to the L74 and Y80 residues and the simulation was performed for 200 ns. Cholesterol parameters (Plesnar et al. 2012; Plesnar et al. 2013) for the all-atom Optimized-Potentials-for-Liquid-Simulations (OPLS-AA) force field were used.

The analysis of peptide-cholesterol simulation revealed that after an initial increase in distance with peptide and even leaving the pocket, cholesterol molecule approached the peptide again and within 30 ns reached a stable association to a binding pocket formed by this N-terminal peptide involving L74 and Y80 (Fig. 3.2). In order to investigate the role of L74 and Y80 residues in cholesterol binding, mutations of these residues were introduced to the peptide-cholesterol complex at 100 ns and the mutant peptide-cholesterol complex were further simulated for 50 ns. Although MD simulations could not reveal altered hydrophobic interactions by the L74I mutation with cholesterol, the Y80S mutation destabilized the cholesterol binding pocket, leading to an instability of the cholesterol-mutant peptide complex and suggesting Y80 as the main binding partner for cholesterol-binding (Figure 3.3).

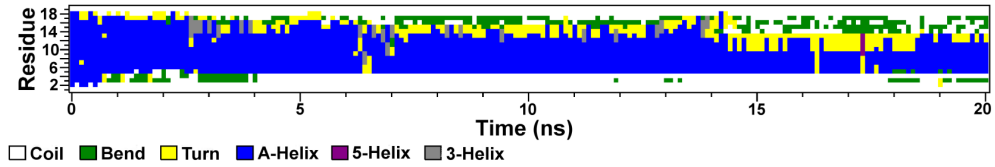

Figure 3.1 Secondary structure profile of wild type N-terminal peptide for first 20 ns, calculated using DSSP program. Only 20 ns data have been shown for making the profile picture simpler. The peptide unfolds at the beginning reaching a more or less stable secondary structure distribution after 15ns that exhibits only a small helical content and mostly random coil structure.

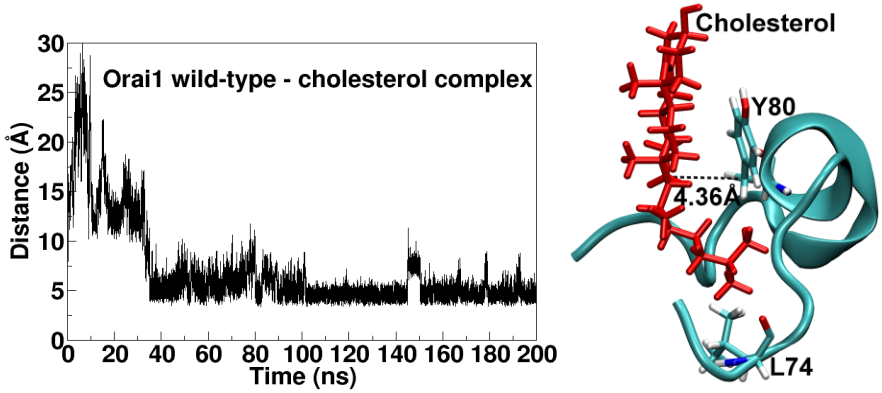

Figure 3.2 MD simulations on the Orai1 N-terminal peptide-cholesterol complex: distance between cholesterol (atom C16) and Y80 (beta-carbon) in the course of the 200ns trajectory is shown for wild type N-terminal peptide-cholesterol complex (left panel), structure of wild type peptide-cholesterol complex after 200 ns MD simulation (right panel) Cholesterol shown as red colored stick, L74 and Y80 residues are shown as sticks colored as atom name. Peptide is shown in cartoon representation (cyan color).

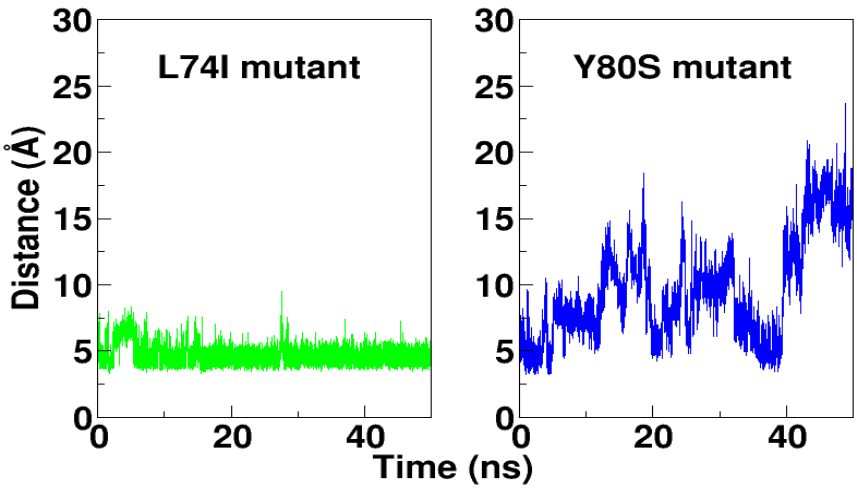

Figure 3.3 Distance between cholesterol (atom C16) and Y/S-80 (beta-carbon) calculated for 50 ns after introduction of the point mutation for both mutants of the N-terminal peptide.

Circular dichroism (CD) spectrum of the peptide also suggested a large portion of peptide secondary structure as random coil together with helical portion (**Paper IV**); which was in line with results from MD simulations of peptide. The reduced α-helicity of peptide observed in CD spectrum and MD simulations suggests that peptide adapts a less-helical secondary structure, different from the structure of this ETON region in dOrai crystal structure which is α-helical. Therefore whether the cholesterol binds to this CB motif in full length Orai1 structure or not, still remains unclear.

In order to study the cholesterol association to full-length Orai1, FRET-TIRF microscopy was used. Results from FRET measurements suggested a close association of Orai1 to cholesterol and mutations L74I and Y80S was resulted into reduced Orai1-cholesterol associations (**Paper IV**).

3.2.2.2 Docking of cholesterol to full-length Orai1

In order to search for cholesterol binding pockets in the full-length Orai1 structure we performed global docking in Yasara (Morris et al. 1998; Krieger et al. 2002). Here whole hexamer structure including the interfaces between monomers was taken into account in order to find the potential binding sites. AutoDock VINA (Trott & Olson 2010) implemented in YASARA (Krieger et al. 2002) was used for docking of cholesterol to Orai1. First, global docking in which the protein was kept rigid and the ligand was flexible was performed to identify potential cholesterol binding sites across the whole hexameric structure. After identification of potential binding sites by several global docking runs, local docking was performed with a fully flexible ligand and partially flexible binding site residues to improve the orientation of the ligand and the residues in the predicted binding pockets.

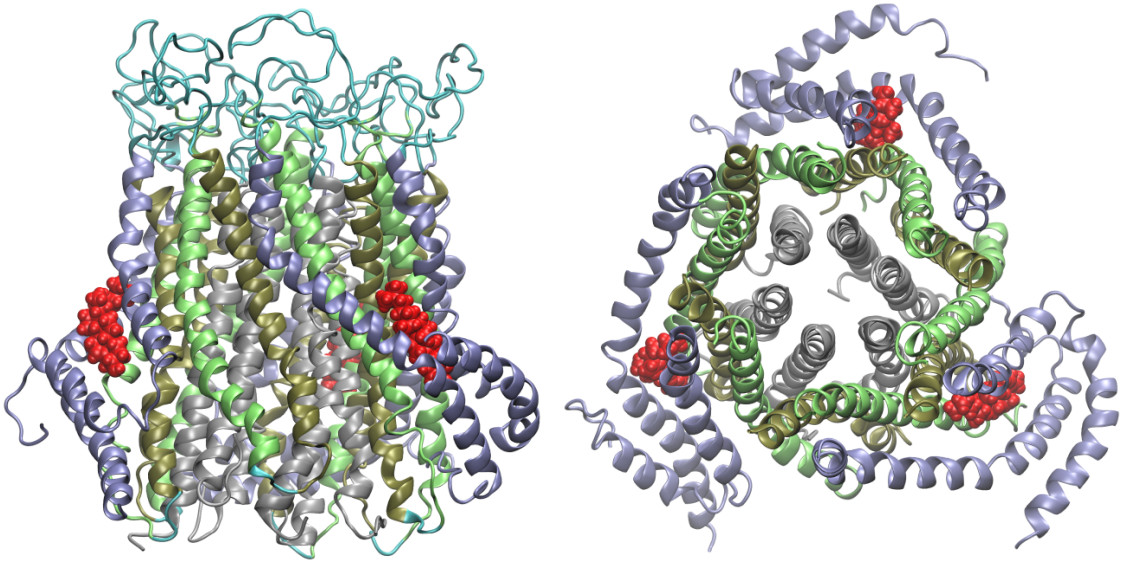

Figure 3.4 Cholesterol binding to a TM2/3/4 binding pocket in Orai1: side view (left panel) and top view (right panel). TM1: silver, TM2: tan, TM3: lime and TM4: ice blue. Cholesterol molecules are shown as vdW spheres and colored as red.

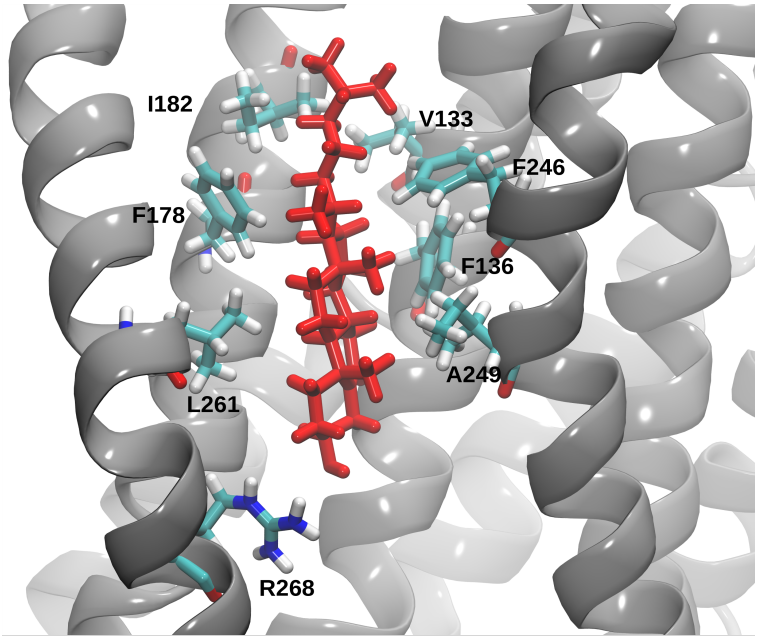

Figure 3.5 Interaction of cholesterol with Orai1 residues in the predicted binding pocket. Cholesterol has been shown as red colored sticks and protein residues have been shown as sticks colored as elements.

Global docking of cholesterol to full-length human Orai1 did indeed reveal binding. However, the cholesterol binding was not found at the helical ETON region in the Orai1 N-terminus, instead it occured at a cholesterol binding pocket located between TM2 (V133, F136), TM3 (F178, I182) and TM4 (F246, A249, L261, R268) from an Orai1 dimer (Figure 3.5)

3.2.2.3 MD simulation of Orai-CHL: wild-type and mutants

In order to investigate the effect of cholesterol on Orai1 and also the effect of mutations in ETON region we studied four different systems: Orai1 wild-type, Orai1-cholesterol complex, Orai1-L74I-cholesterol complex and Orai1-Y80-cholesterol complex.

The Orai1-cholesterol complex was embedded into a pre-equilibrated palmitoyloleoylphosphatidylcholine (POPC) bilayer. Single point mutations, L74I and Y80S were introduced in each monomer of the Orai1-cholesterol complex using YASARA. All these systems were embedded in a lipid bilayer and molecular dynamics simulations were performed for 100 ns and last 10 ns were used for analysis.

Hydrophobic contacts analysis: We performed the analysis of hydrophobic contacts of Orai1 with cholesterol, indicative for cholesterol binding, over the whole length of the Orai1 model (L66 to G297) for wild-type, L74I and Y80S mutants. Hereby, all protein heavy atoms (non-hydrogen) within 4 Å from cholesterol molecules were considered as number of contacts which was then averaged for all three cholesterol binding sites for comparison between wild-type and mutant simulations. Due to the nature of cholesterol these contacts represent nearly exclusively hydrophobic contacts.

The number of hydrophobic contacts was significantly reduced for the Orai1 L74I/Y80S N-terminal mutants compared to the wild-type form following 100 ns of simulations (Figure 3.6). The reduction in number of hydrophobic contacts was more pronounced with the Y80S than with the L74I mutation. The findings about reduction in hydrophobic contacts for mutants through MD simulations are in agreement with FRET measurements (**Paper IV**) which suggested diminished coupling of cholesterol to these Orai1 mutants compared to wild-type Orai1.

The predicted cholesterol binding site is neither in a direct contact with the ETON region (containing L74/Y80 residues) nor it shares any residues with ETON region. This indicated that cholesterol binding to the Orai1 TM2/3/4 domains is probably allosterically affected by these residues in the Orai1 N-terminus.

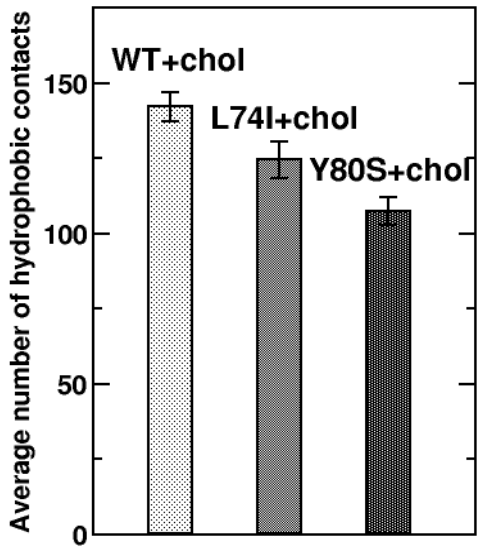

Figure 3.6 Block diagram displaying the number of close contacts of CHL molecule with Orai1 WT compared to mutants (L74I, Y80S).

Pore analysis: In order to analyse the effect of cholesterol and mutations on the pore of the channel pore analysis was performed using HOLE (Smart et al. 1996) program. Here solvent and ions were excluded and the van der Waals radii were assigned according to the AMBER united atom force field.

Pore analysis helped with visualization of the cavity running through the pore and evaluation of the pore radius over the whole TM1 helices in an Orai1 channel. Pore radius is different for different part of the channel as shown for initial homology model of Orai1 (Figure 3.7). In the absence of cholesterol as well as upon L74I or Y80S mutation, the pore radius showed significant widening at the basic region compared to wild-type Orai1 (without cholesterol) (Fig 3.8). As some gating function has been suggested for the basic region of the TM1 helix it is tempting to speculate that this widening might contribute to the increase in currents.

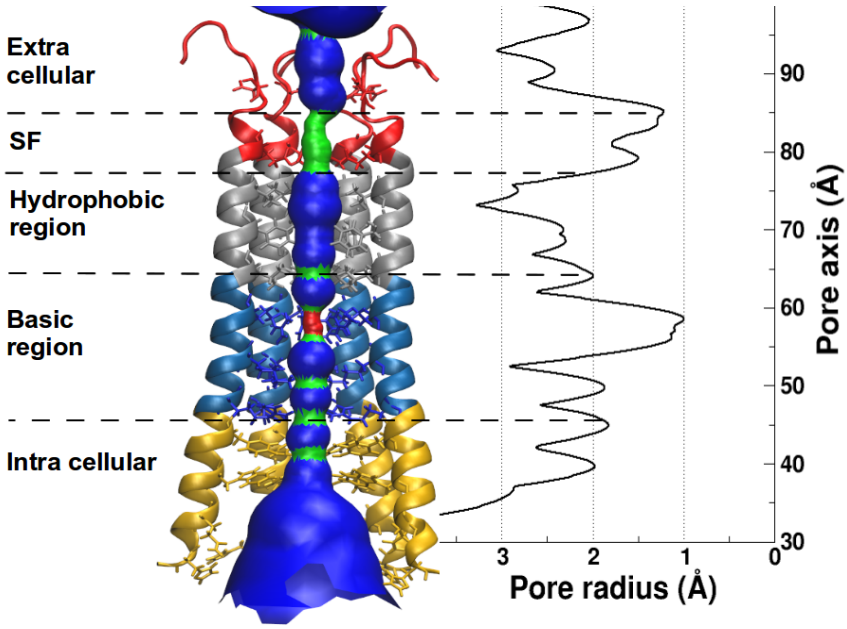

Figure 3.7 Pore radius profile: representation of the internal surface of the pore in the Orai1 WT initial homology model and the pore radius profile.

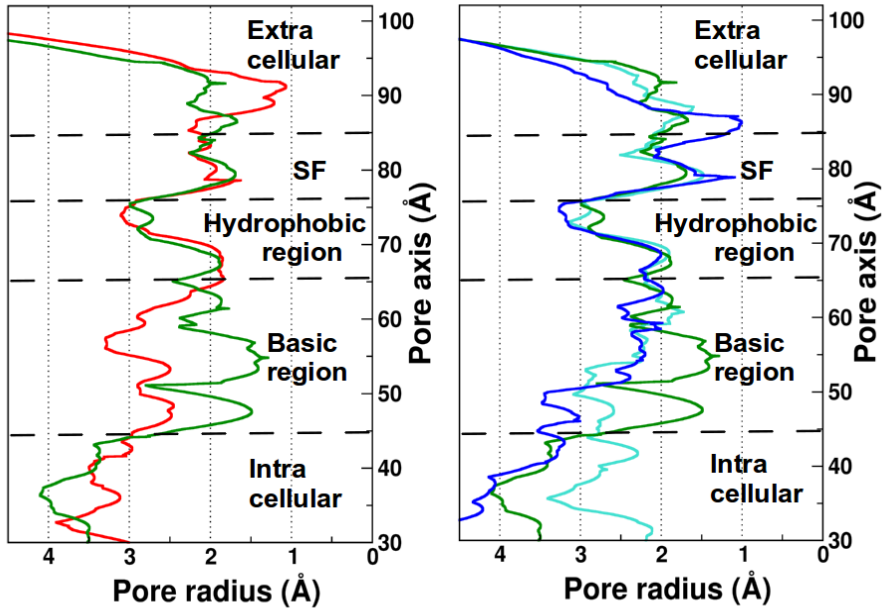

Figure 3.8 Comparison of pore size: Left Panel shows the effect of cholesterol on the pore radius: WT simulation without CHL (red) and WT+cholesterol complex (green).

Right Panel shows the effect of mutations on the pore radius in the Orai1+cholesterol complex: WT+cholesterol complex (green) without, L74I+cholesterol (blue) and Y80S+cholesterol (light blue).

Finally we were able to distill out a cascade of events that are leading to unbinding of cholesterol, and in all cases where cholesterol start to leave this seems to be the initiation. Although the changes are very subtle and not easy to observe, careful analayis using average structures over certain time windows rather than the full trajectory, allowed to get a clear picture:
While in WT K78 from TM1 is found to form a salt bridge with E166 on TM3 from the same monomer and only rarely contacts E149 on TM2 from the adjacent monomer, the mutation initiates a cascade of events. The observed broadening of the channel in the basic region (See Figure 3.8, right panel) pushes K78 closer to E149, and in the case when both start to interact the salt bridge of K78 with E166 is lost as a consequence. E166 then attracts the guanidino group of R170 that is otherwise oriented towards the lipid head groups and forms a persistent salt bridge, hereby not only switching the side chain from one to the other side but also moving backbone atoms by approximately 1 A. As a consequence the neighboring H171 that is part of the CHL binding pocket follows the movement of R170 and gets deeper into the CHL binding pocket, pushing the CHL out of the pocket on this edge of the binding site. In simulations of WT Orai1 without cholesterol the same interactions are observed, only that the cascade of events goes in the opposite direction, as initially the formation of the salt-bridge between R170 and E166 causes a breaking of the salt-bridge between K78 and E166, and as a consequence K78 shifts towards E149 hereby widening the channel in this region (Figure 3.8, left panel).

In summary, experiments showed that cholesterol depletion by application of cholesterol oxidase results into an increase in Ca^{2+} entry. A cholesterol-binding motif was identified in ETON region. Mutations L74I and Y80S in this ETON region showed a similar effect on Ca^{2+} entry and current as cholesterol depleting agents (**Paper IV**). Although fluorescence-binding studies together with MD simulations showed binding of cholesterol to the peptide (aa 72-90), CD spectrum as well as secondary structure data from MD simulations showed a large portion of peptide as random coil, different from what was observed in dOrai crystal structure or Orai1 model. In the dOrai crystal structure and in Orai1 full-length model residues I74 and Y80 are not on the same face of helix and in order to bind cholesterol the helix will have to partially lose its secondary structure. Also the structural analysis of Orai1 model shows that the predicted cholesterol binding site in ETON region is surrounded by other TM-helices and has too little space for cholesterol, therefore without a conformational rearrangement/reorientation of TM-helices cholesterol binding seems unlikely to occur.

Global docking search for cholesterol binding predicted a different binding pocket which was formed by residues from TM2, TM3 and TM4 of one Orai1 dimer subunit of hexameric Orai model. Thus, three cholesterol binding pockets were located within one hexameric channel complex. MD simulations data showed the effect of L74I and Y80I mutations on pore radius as well as on hydrophobic contacts between cholesterol and Orai1. L74I or Y80S mutation within the ETON region reduced the number of hydrophobic contacts within this binding pocket to cholesterol as revealed by MD simulations. In line with MD simulation, FRET measurements also suggested diminished coupling of cholesterol to these Orai1 mutants compared to wild-type Orai1. Also pore size profile obtained via MD simulations showed that the presence of cholesterol in

Orai1 wild type results into narrowing down of pore radius in the basic region (Figure 3.8, left panel). Mutations L74I and Y80I in ETON region (also with cholesterol) show an increase of pore size in this basic region (Figure 3.8, right panel). Together the data suggests that cholesterol modulates the pore size of Orai1 channel, mutations in the ETON region of TM1 helix affect the binding of cholesterol in an allosteric manner and consequently affect the Orai1 channel function.

4. Conclusions and future aspects

Molecular dynamics simulations of ArgRC of E. coli in apo and holo form showed different behavior of the system. A global motion was observed where both forms, apo and holo show a rotation of trimers over each other and attain conformations distinct from the one in crystal structure. Quantification of rotation showed that apo and holo states differ from each other by 4° to 5° of rotation about the trimer-trimer interface. This motion was correlated with formation and release of R100 - D128 salt bridges between two trimers. ApoArgRC system sampled a broad range of conformational space including the states where L-arginine binding pockets were open due to broken R100 - D128 interactions over a time period. One L-arginine molecule was docked in the open binding pocket. Three MD simulations of singly liganded states were performed. The singly liganded ArgRC systems remain stable during MD simulations, thus these conformers with open pocket indeed represented binding competent states. Also MD simulations successfully showed that binding of one L-arginine is sufficient to bring the system from apo-like to holo-like conformations. Further MD simulations of these singly-liganded states presented some open pockets which gives the possibility to prepare systems with two L-arginine molecules and study the negative cooperativity of L-arginine binding.

Phylogenetic analysis of available ArgR sequences showed three branches, represented by MtArgR, EcArgR and BsArgR. Unlike ArgR of *E. coli*, ArgR of *B. subtilis* does not have charged residue like R110 of E. coli at analogous position in the core domain. MD simulations of B. subtilis full-length structures in apo and holo form showed rotational states which differ by ~9 degrees from each other. The charged residues K75, R78 and D82 of the linker helix were found responsible for the rotation. Hydrogen bond and entropy analysis

revealed that while C-domains show lower entropy in holoBsArgR than compared to apoBsArgR, whereas N-domain in holoBsArgR possess higher entropy and are more free and adopts conformations which are more likely to bind DNA, these conformations were not observed for apoBsArgR. Thus the presence of rotation in ArgR in *E. coli* as well as in *B. subtilis* despite of different drivers responsible for rotations suggest that rotational motion is inherent for all ArgR and L-arginine allosterically activates the DNA binding domain by using this rotation.

Orai3 structure was modeled using homology modeling. Using molecular dynamics simulations we were able to show different behavior of Orai1 and Orai3 isoforms. Structural comparison of these two isoforms after MD simulations revealed that TM2 in Orai3 is longer than in Orai1 hence Orai1-loop2 is longer than Orai3-loop2. Truncation of TM1 N-terminus results into some space and Orai1-loop2 moves to this space and forms interaction with Y80 residue at N-terminus. This phenomena was not observed in Orai3 where loop2 is shorter. Thus MD simulations together with experimental data show that not only N-terminus but also loop2 is required for store-operated function of Orai channel.

A cholesterol binding motif was identified in ETON region of Orai1 using bioinformatics analysis. MD simulations of Orai1-peptide (aa 72-90) and cholesterol showed that cholesterol molecule was able to bind peptide and mutation Y80S had negative effect on cholesterol binding. However secondary structure analysis showed that peptide loses up to 75% of helical structure thus adapts very different secondary structure than in dOrai crystal structure or in Orai1 model. The different conformation of N-terminal peptide observed in MD simulations and also the fact that in full-length Orai1 structure, this cholesterol-binding motif is surrounded by other TM-helices shows that a conformational

rearrangement of N-terminus of Orai1 would be required in order to bind the cholesterol. This suggested a possibility of other cholesterol binding site in full-length Orai1 structure. A cholesterol binding site in full-length Orai1 was identified and MD simulations Orai1-cholesterol complex were performed. Cholesterol remained bound to all three binding sites in a hexamer. The MD simulation of mutants L74I and Y80S showed allosteric effects of these mutations in N-terminus region on cholesterol binding. In both mutants a decrease in hydrophobic contacts between protein residues and cholesterol was observed in comparison to wild type-cholesterol complex. Pore analysis of the trajectories showed that presence of cholesterol lowers the pore size in basic region and mutations L74I and Y80S (with cholesterol) showed wider pore compared to Orai1-cholesterol complex. Thus cholesterol has modulatory role in regulating Orai1 channel function via controlling the pore size and mutations in ETON region allosterically affect the cholesterol binding and consequently affect the pore diameter and store-operated function of Orai channel.

5. References

Adcock, S. A., McCammon, J. A. (2006) Molecular Dynamics: Survey of Methods for Simulating the Activity of Proteins. Chemical reviews 106(5): 1589-1615.

Allwright, J. C. (1976) Conjugate gradient versus steepest descent. J. Optimiz. Theory App. 20(1): 129–134.

Altschul, S. F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) Basic local alignment search tool. J. Mol. Biol. 215: 403-410.

Altschul, S. F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database. Nucleic Acids Res. 25(17): 3389-3402.

Amcheslavsky, A., Wood, M. L., Yeromin, A. V., Parker, I., Freites, J. A., Tobias, D. J., Cahalan, M. D. (2015). Molecular Biophysics of Orai Store-Operated Ca2+ Channels. Biophysical Journal, 108(2):237–246.

Bergsmann, J., Derler, I., Muik, M., Frischauf, I., Fahrner, M., Pollheimer, P., Schwarzinger, C., Gruber, H. J., Groschner, K., Romanin, C. (2011)Molecular determinants within N terminus of Orai3 protein that control channel activation and gating. J Biol Chem. 286(36):31565-75.

Berna-Erro, A., Woodard, G. E., & Rosado, J. A. (2012) Orais and STIMs: physiological mechanisms and disease. Journal of Cellular and Molecular Medicine, 16(3): 407–424.

Berridge, M. J., Lipp, P., Bootman, M. D. (2000) The versatility and universality of calcium signalling. Nat. Rev., Mol. Cell Biol. 1: 11-21.

Berridge, M. J. (2001) The versatility and complexity of calcium signalling. Novartis Found. Symp. 239: 52–64.

Blakeley, M. P., Langan, P., Niimura, N., Podjarny, A. (2008) Neutron crystallography: opportunities, challenges, and limitations. Current opinion in structural biology. 18(5): 593-600.

Burke, M., Merican, A. F., Sherratt D. J. (1994) Mutant Escherichia coli arginine repressor proteins that fail to bind L-arginine, yet retain the ability to bind their normal DNA-binding sites. Mol. Microbiol. 13:609-618.

Canutescu, A. A. and Dunbrack, R. L. Jr. (2003) Cyclic coordinate descent: A robotics algorithm for protein loop closure. Protein Sci. 12: 963-972.

Canutescu, A. A., Shelenkov, A. A., Dunbrack, R. L. Jr. (2003) A graph-theory algorithm for rapid protein side-chain prediction. Protein Sci. 12: 2001-2014.

Catterall, W. A. (1995) Structure and Function of Voltage-Gated Ion Channels. Annual Review of Biochemistry. 64:493-531.

Changeux, J. P. (1961) The feedback control mechanisms of biosynthetic L-threonine deaminase by Lisoleucine. Cold Spring Harb. Symp. Quant. Biol. 26: 313–18.

Changeux, J. P. and Edelstein, S. J. (2005) Allosteric mechanisms of signal transduction. Science. 308: 1424–1428.

Changeux, J. P. (2012) Allostery and the Monod–Wyman–Changeux model after 50 years. Annu. Rev. Biophys. 41:103–133.

Changeux, J. P. (2013) 50 years of allosteric interactions: the twists and turns of the models. Nature Reviews Molecular Cell Biology 14, 819-829.
Cooper, A., Dryden, D. T. (1984) Allostery without conformational change. A plausible model. Eur. Biophys. J. 11: 103–109.
Cherney, L. T, Cherney, M. M, Garen, C. R., Lu, G. J, James, M. N. (2008) Structure of the C-domain of the arginine repressor protein from Mycobacterium tuberculosis. Acta Cryst D 64: 950–956.
Cherney, L. T, Cherney, M. M, Garen, C. R., James, M. N. (2009) The structure of the arginine repressor from Mycobacterium tuberculosis bound with its DNA operator and Co-repressor, L-arginine. J Mol Biol. 388(1):85–97. pmid:19265706.
Cherney, L. T., Cherney, M. M., Garen, C. R., James, M. N. (2010) Crystal structure of the intermediate complex of the arginine repressor from Mycobacterium tuberculosis bound with its DNA operator reveals detailed mechanism of arginine repression. J Mol Biol. 399(2):240–54. pmid:20382162.
Chinea, G., Padron, G., Hooft, R. W., Sander, C., Vriend, G. (1995) The use of position-specific rotamers in model building by homology. Proteins 23: 415-421.
Chothia, C. and Lesk, A.M. (1986) The relation between the divergence of sequence and structure in proteins. EMBO J. 5: 823–826.
Chothia, C. (1992) Proteins. One thousand families for the molecular biologist. Nature 357(6379):543-4.
Chung, S.Y. and Subbiah, S. (1996) How similar must a template protein be for homology modeling by side-chain packing methods? Pac. Symp. Biocomput. 126–141.
Clapham, D. E. (2007) Calcium signaling. Cell, 14:1047-1058.
Collins, S. R., & Meyer, T. (2011) Evolutionary Origins of STIM1 and STIM2 within Ancient Ca2+ Signaling Systems. Trends in Cell Biology, 21(4), 202–211.
Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz Jr., K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. J. Am. Chem. Soc. 117: 5179–5197.
Cui, Q., Karplus, M. (2008) Allostery and cooperativity revisited, Protein Sci. 17: 1295–1307.
Cunin, R., Glansdorff, N., Piérard, A., Stalon, V. (1986). Biosynthesis and metabolism of arginine in bacteria. Microbiological Reviews 50(3): 314–352.
Cunin, R., T. Eckhardt, J. Piette, Boyen, A., Pierard, A., Glansdorff, N. (1983). Molecular basis for modulated regulation of gene expression in the regulon of Escherichia coli K12. Nucleic Acids Res. 11:5007-5019.
Czaplewski LG, North AK, Smith MC, Baumberg S, Stockley PG (1992) Purification and initial characterization of AhrC: the regulator of arginine metabolism genes in Bacillus subtilis. Mol Microbiol 6:267–275 11.
Demaurex, N., Frieden, M. (2003) Measurements of the free luminal ER Ca(2+) concentration with targeted “cameleon” fluorescent proteins. Cell Calcium 34:109–119

Dennis, C. A., Glykos, N. M., Parsons, M. R., Phillips, S. E. V. (2002) The structure of AhrC, the arginine repressor/activator protein from Bacillus subtilis. Acta Cryst D58: 421–430.
Derler, I., Plenk, P., Fahrner, M., Muik, M., Jardin, I., Schindl, R., Gruber, H. J., Groschner, K., Romanin, C. (2013) The extended transmembrane Orai1 N-terminal (ETON) region combines binding interface and gate for Orai1 activation by STIM1. J Biol Chem. 288(40): 29025-34.
Derler I., Jardin I., Stathopulos P. B., Muik M., Fahrner M., Zayats V., Pandey S. K., Poteser M., Lackner B., Absolonova M., Schindl R., Groschner K., Ettrich R., Ikura M., and Romanin C. (2016) Cholesterol modulates Orai1 channel function. Sci. Signal. 9, ra10 10.1126/scisignal.aad7808
Duan, Y., Wu, C., Chowdhury, S., Lee, M. C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J., Kollman, P. (2003). A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J. Comput. Chem., 24(16): 1999-2012.
Earl, D. J. and Deem, M. W. (2008). Monte Carlo simulations. Methods in Molecular Biology 443, 25–36.
Epand, R. M. (2008) Proteins and cholesterol-rich domains. Biochim Biophys Acta. 1778: 1576–1582.
Feske, S., Prakriya, M., Rao, A., Lewis, R. S. (2005) A severe defect in CRAC Ca2+ channel activation and altered K^+ channel gating in T cells from immunodeficient patients. J Exp Med. 202(5): 651-62.
Feske, S., Gwack, Y., Prakriya, M., Srikanth, S., Puppel, S. H., Tanasa, B., Hogan, P. G., Lewis, R.S., Daly, M., Rao, A. (2006) A mutation in Orai1 causes immune deficiency by abrogating CRAC channel function. Nature. 441(7090): 179-85.
Frischauf, I., Schindl, R., Derler, I., Bergsmann, J., Fahrner, M., and Romanin, C. (2008) The STIM/Orai coupling machinery. Channels 2: 261–268.
Frischauf, I., Zayats, V., Deix, M. et al. (2015) A calcium-accumulating region, CAR, in the channel Orai1 enhances Ca2+ permeation and SOCE-induced gene transcription. Science Signaling, 8(408), ra131.
Garnett, J.A., Baumberg, S., Stockley, P.G., Phillips, S.E. (2007) Structure of the C-terminal effector-binding domain of AhrC bound to its corepressor L-arginine. Acta Crystallogr.,Sect.F 63: 918-921.
Gunasekaran, K., Ma, B., Nussinov R. (2004) Is allostery an intrinsic property of all dynamic proteins? Proteins 57: 433–443.
Grinter, S. Z., Zou, X. (2014) Challenges, Applications, and Recent Advances of Protein-Ligand Docking in Structure-Based Drug Design. Molecules 19: 10150-10176.
Guvench, O., MacKerell, Alexander, D. Jr. (2008) Comparison of Protein Force Fields for Molecular Dynamics Simulations. In Methods Molecular Biology; Kukol, A., Ed.; Humana Press: New York, NY, USA. Volume 443: 63–88.
HILLE, B. (2001) Ion Channels of Excitable Membranes, 3rd Ed. Sunderland, MA: Sinauer Associates.

Hockney, R. W. (1970) The potential calculation and some applications. Methods Comput. Phys 9: 136–211.
Hoth, M., Penner, R. (1992) Depletion of intracellular calcium stores activates a calcium current in mast cells. Nature. 355(6358): 353-6.
Hoth, M. (1995) Calcium and barium permeation through calcium release-activated calcium (CRAC) channels. Pflügers Arch 430: 315–322.
Hou, X., Pedi, L., Diver, M. M., Long, S. B. (2012) Crystal structure of the calcium release-activated calcium channel Orai. Science. 338(6112): 1308-13.
Humphrey, W., Dalke, A., Schulten, K. (1996) VMD: Visual molecular dynamics. J Mol Graph Model 14, 33-38.
Jin, L., Xue, W. F., Fukayama, J. W., Yetter, J., Pickering, M. et al. (2005) Asymmetric allosteric activation of the symmetric ArgR hexamer. J Mol Biol 346: 43–56.
Jorgensen, W. L., Maxwell, D. S., Tirado-Rives, J. (1996). Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. J. Am. Chem. Soc. 118(45): 11225-11236.
Kabsch, W., Sander, C. (1983) Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22, 2577-2637.
Ketchum, K. A., Joiner, W.J., Sellers, A. J., Kaczmarek, L. K. Goldstein S. A. (1995) A new family of outwardly rectifying potassium channel proteins with two pore domains in tandem. Nature 376: 690–695.
Koshland, D. E. Jr, Nemethy, G, Filmer, D. (1966) Comparison of experimental binding data and theoretical models in proteins containing subunits. Biochem. 5: 365–385.
Krieger, E., Koraimann, G., Vriend, G. (2002). Increasing the precision of comparative models with YASARA NOVA-a self-parameterizing force field. Proteins Struct. Funct. Bioinf. 47(3): 393-402.
Kumar, S., Ma, B., Tsai, C. J., Wolfson, H., Nussinov, R. (1999) Folding funnels and conformational transitions via hinge-bending motions. Cell Biochem Biophys 31: 141–164.
Laskowski, R. A., Macarthur, M. W., Moss, D. S., Thornton, J. M. (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr, 26: 283–291.
La Rovere, R. M., Roest, G., Bultynck, G., Parys, J. B. (2016) Intracellular Ca(2+) signaling and Ca(2+) microdomains in the control of cell survival, apoptosis and autophagy. Cell Calcium 60:74–87.
Leach, A. R. (2001) Molecular Modelling: Principles and Applications. Pearson Education; London.
Lehninger, A. L., Nelson, D. L. 1., Cox, M. M. (2008). Lehninger principles of biochemistry (5th ed.).
Li, H., Papadopoulos, V. (1998) Peripheral-type benzodiazepine receptor function in cholesterol transport. Identification of a putative cholesterol recognition/interaction amino acid sequence and consensus pattern. Endocrinology. 139: 4991–4997.

Li, Z., Lu, J., Xu, P., Xie, X., Chen, L., Xu, T. (2007) Mapping the interacting domains of STIM1 and Orai1 in Ca2+ release-activated Ca2+ channel activation. J Biol Chem. 282(40):29448-56.

Lim, D. B., Oppenheim, J. D., Eckhardt, T., Maas, W. K. (1987) Nucleotide sequence of the argR gene of Escherichia coli K-12 and isolation of its product, the arginine repressor. Proc Natl Acad Sci USA 84: 6697–6701.

Lindahl, E. (2015) Molecular Dynamics Simulations. In: Kukol A. (eds) Molecular Modeling of Proteins. Methods in Molecular Biology (Methods and Protocols), vol 1215. Humana Press, New York, NY.

Lindorff-Larsen, K., Piana, S., Palmo, K. et al. (2010) Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins 78(8): 1950-1958.

Liou, J., Kim, M. L., Heo, W. D., Jones, J. T., Myers, J. W., Ferrell, J. E. Jr, Meyer, T. (2005) STIM is a Ca2+ sensor essential for Ca2+-store-depletion-triggered Ca^{2+} influx. Curr Biol. 2005 Jul 12; 15(13): 1235-41.

Lipscombe, D., Toro, C.P. (2013) Ion Channels. In: eLS. John Wiley & Sons, Ltd: Chichester.

Purves, D., Augustine, G. J., Fitzpatrick, D., Hall, W. C., LaMantia, A.-S., McNamara, J. O., & Williams, S. M. (Eds.). (2004). Neuroscience (3rd ed.). Sunderland, MA, US: Sinauer Associates.

Lu, C.D. (2006) Pathways and regulation of bacterial arginine metabolism and perspectives for obtaining arginine overproducing strains. Appl. Microbiol. Biotechnol. 70: 26-272.

Lushington, G. H. (2015) Comparative Modeling of Proteins. In: Kukol A. (eds) Molecular Modeling of Proteins. Methods in Molecular Biology (Methods and Protocols), vol 1215. Humana Press, New York, NY

Lushington, G. H. (2008). Comparative modeling of proteins. Methods in Molecular Biology 443: 199–212.

Ma, B., Kumar, S., Tsai, C. J., Nussinov, R. (1999) Folding funnels and binding mechanisms. Protein Eng 12: 713–720. pmid:10506280

Maas, W. K. (1994) The Arginine Repressor of Escherichia coli. Microbiological Reviews, Dec., p. 631-640.

Maas, W. K., and Clark, A. J., 1964. Studies on the mechanism of repression of arginine biosynthesis in Escherichia coli. II. Dominance of repressibility in diploids. J. Mol. Biol. 8: 365-370.

Ma, L. and Cui, Q. (2007) The activation mechanism of a signaling protein at atomic resolution from advanced computations. J. Am. Chem. Soc. 129: 10261–10268.

Ma, J., Sigler, P. B., Xu, Z., and Karplus, M. (2000) A dynamic model for the allosteric mechanism of GroEL. J. Mol. Biol. 302: 303–313.

Maffeo, C., Bhattacharya, S.,Yoo, J., Wells, D., Aksimentiev A. (2012) Modeling and Simulation of Ion Channels. Chemical Reviews 112(12): 6250-6284.

MacKerell, A. D., Bashford, D., Bellott, M, Dunbrack R.L Evanseck, J. D Field, M. J. Fischer, S Gao, J., Guo, H., Ha, S., Joseph-McCarthy, D., Kuchnir, L., Kuczera, K., Lau, F. T. K.,

Mattos, C., Michnick, S., Ngo, T., Nguyen, D. T., Prodhom, B., Reiher, W. E., Roux, B., Schlenkrich, M., Smith, J. C., Stote, R., Straub, J., Watanabe, M., WiorkiewiczKuczera, J., Yin, D., Karplus, M. (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. J. Phys. Chem. B. 102(18): 3586–3616.

Mackerell, A. D., Feig, M., Brooks, C. L. (2004) Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. J. Comput. Chem., 25(11), 1400-1415.

McCammon, J. A., Gelin, B. R., and Karplus, M.Dynamics of folded proteins. Nature (Lond.) 267, 585 (1977).

McCammon, J. A., Harvey, S. C. (1987) Dynamics of Proteins and Nucleic Acids. Cambridge University Press; Cambridge.

McNally, B. A., Somasundaram, A., Jairaman, A., Yamashita, M., and Prakriya, M. (2013) The C- and N-terminal STIM1 binding sites on Orai1 are required for both trapping and gating CRAC channels. J. Physiol. 591: 2833–2850.

Melo, F., Devos, D., Depiereux, E., Feytmans, E. (1997)ANOLEA: a www server to assess protein structures. Proc Int Conf Intell Syst Mol Biol. 5: 187–190.

Michalsky, E., Goede, A., Preissner, R. (2003) Loops In Proteins (LIP) - a comprehensive loop database for homology modelling. Protein Engineering 16: 979-985.

Monod, J. and Jacob, F. (1961) Teleonomic mechanisms in cellular metabolism, growth, and differentiation. Cold Spring Harb. Symp. Quant. Biol. 26: 389–401.

Monod, J., Wyman, J., Changeux, J. P. (1965) On the nature of allosteric transitions: a plausible model. J Mol Biol. 12: 88–118.

Monod, J. (1977) Chance and Necessity: Essay on the Natural Philosophy of Modern Biology: Penguin Books Ltd.

Morris, G.M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K., Olson, A. J. (1998) Automated Docking Using a Lamarckian Genetic Algorithm and and Empirical Binding Free Energy Function. J. Comput. Chem. 19: 1639-1662.

Muik, M., Frischauf, I., Derler, I., Fahrner, M., Bergsmann, J., Eder, P., Schindl, R., Hesch, C., Polzinger, B., Fritsch, R., Kahr, H., Madl, J., Gruber, H., Groschner, K., Romanin, C. J. (2008) Dynamic coupling of the putative coiled-coil domain of ORAI1 with STIM1 mediates ORAI1 channel activation.Biol Chem. 283(12): 8014-22.

Mullins, F. M., Park, C. Y., Dolmetsch, R. E., and Lewis, R. S. (2009) STIM1 and calmodulin interact with Orai1 to induce Ca^{2+} -dependent inactivation of CRAC channels. Proc. Natl. Acad. Sci. U.S.A. 106: 15495-15500.

Navarro-Borelly, L., Somasundaram, A., Yamashita, M., Ren, D., Miller, R. J., and Prakriya, M. (2008) STIM1-Orai1 interactions and Orai1 conformational changes revealed by live-cell FRET microscopy. J. Physiol. 586: 5383–5401.

Oostenbrink, C., Villa, A., Mark, A. E., & Gunsteren, W. F. (2004). A biomolecular force field based on the free enthalpy of hydration and solvation: the GROMOS force-field parameter sets

53A5 and 53A6. J. Comput. Chem. 25(13): 1656-1676.
Pagadala, N. S., Syed. K., Tuszynski, J. (2017) Software for molecular docking: a review. Biophysical Reviews 9(2): 91-102. doi:10.1007/s12551-016-0247-1.
Parekh, A. B., Putney, J. W. (2005) Store-operated calcium channels. Physiol Rev. 85: 757–810.
Partiseti, M., Le Deist, F., Hivroz, C., Fischer, A., Korn, H., Choquet, D. (1994) The calcium current activated by T cell receptor and store depletion in human lymphocytes is absent in a primary immunodeficiency. J Biol Chem. 269(51): 32327-35.
Plesnar, E., Subczynski, W.K., Pasenkiewicz-Gierula, M. (2012) Saturation with cholesterol increases vertical order and smoothes the surface of the phosphatidylcholine bilayer: a molecular simulation study. Biochim Biophys Acta 1818, 520-529.
Plesnar, E., Subczynski, W. K., Pasenkiewicz-Gierula, M. (2013) Comparative computer simulation study of cholesterol in hydrated unary and binary lipid bilayers and in an anhydrous crystal. The journal of physical chemistry B 117: 8758-8769.
Putney, J.W. Jr. (1986) A model for receptor-regulated calcium entry. Cell Calcium 7(1):1-12.
Putney, J. W. (1990) Capacitative calcium entry revisited. Cell Calcium 11: 611–624.
Prakriya, M., and Lewis, R. S. (2006) Regulation of CRAC Channel Activity by Recruitment of Silent Channels to a High Open-probability Gating Mode. J. Gen. Physiol. 128, 373–386.
Prakriya, M., & Lewis, R. S. (2015) Store-Operated Calcium Channels. Physiological Reviews 95(4): 1383–1436.
Ni, J., Sakanyan, V., Charlier, D., Glansdorff, N., Van Duyne G. D. (1999) Structure of the arginine repressor from Bacillus stearothermophilus. Nat Struct Biol 6: 427–432.
Ni, J., Sakanyan, V., Charlier, D., Glansdorff, N., Van Duyne, G. D. (1999)Structure of the arginine repressor from Bacillus stearothermophilus. Nat Struct Biol. 6(5): 427–32.
Pauling, L. (1935) The oxygen equilibrium of hemoglobin and its structural interpretation. Proc. Natl. Acad. Sci. 21: 181–191.
Park, C. Y., Hoover, P. J., Mullins, F. M., Bachhawat, P., Covington, E. D., Raunser, S., Walz, T., Garcia, K.C, Dolmetsch, R. E., Lewis, R. S. (2009) STIM1 clusters and activates CRAC channels via direct binding of a cytosolic domain to Orai1. Cell. 136(5): 876-90.
Park, C. Y. et al. (2009) STIM1 clusters and activates CRAC channels via direct binding of a cytosolic domain to Orai1. Cell. 136: 876–890.
Raffaello, A., Mammucari, C., Gherardi, G., Rizzuto, R. (2016) Calcium at the Center of Cell Signaling: interplay between endoplasmic reticulum, mitochondria, and lysosomes. Trends Biochem. Sci. 41:1035–1049.
Rahman, A. and Stillinger, F. H. (1971) Molecular dynamics study of liquid water. J. Chem. Phys. 55: 3336–3359.
Rawson, S., Davies, S., Lippiat, J. D., Muench, S. P. (2016) The changing landscape ofmembrane protein structural biology through developments in electron microscopy. Mol. Membr. Biol., 33(1–2), 12-22.
Richardson, J. S (1981) The anatomy and taxonomy of protein structure. Adv Protein Chem. 34:167-339.

Roos, J. et al. (2005) STIM1, an essential and conserved component of store-operated Ca^{2+} channel function. J. Cell Biol. 169: 435–445.
Sadus, R. J. (1999) Molecular Simulation of Fluids. Applications to Physical Systems. Second Edition, Elsevier, Amsterdam.
Šali, A. and Blundell, T.L. (1993). Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. 234 (3): 779–815.
Schlick, T., Collepardo-Guevara, R., Halvorsen, L. A., Jung, S., Xiao, X. (2011) Biomolecular modeling and simulation: a field coming of age. Quarterly reviews of biophysics 44(2):191-228.
Shi, Y. (2014) A Glimpse of Structural Biology through X-Ray Crystallography. Cell 159(5): 995 – 1014.
Smart, O. S., Neduvelil, J. G., Wang, X., Wallace, B. A., Sansom, M. S. P. (1996) HOLE: A program for the analysis of the pore dimensions of ion channel structural models. Journal of Molecular Graphics 14(6): 354–360.
Smith, M. C. M., Czaplewski, L., North, A. K., Baumberg, S. & Stockley, P. G. (1989) Sequences required for regulation of arginine biosynthesis promoters are conserved between Bacillus subtilis and Escherichia coli. Mol. Microbiol. 3, 23-28.
Smyth, J. T., Hwang, S. Y., Tomita, T., DeHaven, W. I., Mercer, J. C. & Putney, J. W. (2010). Activation and regulation of store-operated calcium entry. J. Cell. Mol. Med., 14, 2337-2349.
Soboloff, J., Rothberg, B.S., Madesh, M., Gill, D. L. (2012) STIM proteins: dynamic calcium signal transducers. Nat Rev Mol Cell Biol 13: 549–565.
Srinivasan, N., Blundell, T. L. (1993) An evaluation of the performance of an automated procedure for comparative modelling of protein tertiary structure. Protein Eng., 6(5): 501–512.
Strawn, R., Melichercik, M., Green, M., Stockner, T., Carey, J., Ettrich, R. (2010) Symmetric allosteric mechanism of hexameric Escherichia coli arginine repressor exploits competition between L-arginine ligands and resident arginine residues. PLoS Comput Biol 6(6):e1000801.
Sol, A. del., Tsai, C. J., Ma, B., Nussinov, R. (2009) The origin of allosteric functional modulation: multiple pre-existing pathways, Structure 17: 1042–1050.
Stathopulos, P. B., Schindl, R., Fahrner, M., Zheng, L., Gasmi-Seabrook ,G. M., Muik, M., Romanin, C., Ikura, M. (2013) STIM1/Orai1 coiled-coil interplay in the regulation of store-operated calcium entry. Nat Commun. 4: 2963.
Stathopulos, P. B., Zheng, L., Li, G.Y., Plevin, M. J., Ikura, M. (2008) Structural and mechanistic insights into STIM1-mediated initiation of store-operated calcium entry. Cell. 135(1):110-22.
Sternberg, M. J. E. (1996) Protein Structure Prediction - A Practical Approach. Oxford University Press, Oxford.
Sutcliffe MJ, Hayes FR, Blundell TL (1987) Knowledge based modeling of homologous proteins, part II: Rules for the conformations of substituted side chains. Prot Eng, 1: 385–92.
Sunnerhagen, M., Nilges, M., Otting, G., Carey, J. (1997) Solution structure of the DNA-binding domain and model for the complex of multifunctional hexameric arginine repressor with DNA. Nat Struct Biol 4: 819–826.

Swain, J. F., Gierasch, L.M. (2006) The changing landscape of protein allostery, Curr. Opin. Struct. Biol. 16: 102–108.
Swope, W. C., Andersen, H. C., Berens, P. H., Wilson, K. R. (1982) A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: application to small water clusters. J. Chem. Phys., 76(1): 637-649.
Takemura, H., Hughes, A. R., Thastrup, O., Putney, J. W. Jr. (1989) Activation of calcium entry by the tumor promoter thapsigargin in parotid acinar cells. Evidence that an intracellular calcium pool and not an inositol phosphate regulates calcium fluxes at the plasma membrane. J Biol Chem. 264(21): 12266-71.
Taly, A., Corringer, P. J., Guedin, D., Lestage, P., Changeux, J. P. (2009) Nicotinic receptors: allosteric transitions and therapeutic targets in the nervous system. Nature Rev. Drug Discov. 8: 733–750.
Tian, G., D. Lim, J. Carey, and Maas, W. K. (1992) Binding of the arginine repressor of Escherichia coli K12 to its operator sites. J. Mol. Biol. 226:387-397.
Tovar, K. R., Westbrook, G. L. (2012) Ligand-Gated Ion Channels, Editor(s): Nicholas Sperelakis, Cell Physiology Source Book (Fourth Edition), Academic Press. 549-562.
Touw, W. G. et al. (2015) A series of PDB-related databanks for everyday needs. NucleicAcids Res 43, D364-368.
Trott, O., Olson, A. J. (2010)AutoDock VINA: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. J. Comput. Chem. 31: 455-461.
Udaka, S. (1966) Pathway-specific pattern of control of arginine biosynthesis in bacteria. J. Bacteriol. 91:617-621.
Van der Vaart, A., Ma, J., and Karplus, M. (2004) The unfolding action of GroEL on a protein substrate. Biophys. J. 87: 562–573.
Van Duyne, G.D., Ghosh, G., Maas, W.K., Sigler, P.B. (1996) Structure of the oligomerization and l-arginine binding domain of the arginine repressor of Escherichia coli. J. Mol. Biol., 256: 377-391.
Verlet, L. (1967) Computer experiments on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. Phys. Rev., 159, 98-103.
Vogel, H. J. (1957) Repression and induction as control mechanisms of enzyme biogenesis: the "adaptive" formation of acetylornithinase. The chemical basis of heredity, The Johns Hopkins Press, Baltimore, p. 276-289.
Vogel, H. J., and W. L. McLellan (1970). N-acetylglutamic semialdehyde dehydrogenase (Escherichia coli). Methods Enzymol. 17A: 255-260.
Vriend, G. (1990) WHAT IF: A molecular modeling and drug design program. J. Mol. Graph., 8, 52-56.
Vyas, S., and Maas,W. K. (1963). Feedback inhibition of acetylglutamate synthetase by arginine in Escherichia coli. Arch. Biochem. Biophys. 100: 542-546.
Wang, H., Glansdorff, N., Charlier, D. (1998) The arginine repressor of Escherichia coli K-12

makes direct contacts to minor and major groove determinants of the operators. J. Mol. Biol. 277, 805-824.
Wiederstein, M. and Sippl, M. J. (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res., 35, 407-410.
Wang G and Dunbrack RL Jr (2003) PISCES: a protein sequence culling server. Bioinformatics 19: 1589-1591.
Wishart, D. S. (2008) Identifying Putative Drug Targets and Potential Drug Leads. In: Kukol A. (eds) Molecular Modeling of Proteins. Methods Molecular Biology™, vol 443. Humana Pres.
Xie, Z. R., Hwang, M. J. (2015) Methods for Predicting Protein–Ligand Binding Sites. In: Kukol A. (eds) Molecular Modeling of Proteins. Methods in Molecular Biology (Methods and Protocols), vol 1215. Humana Press, New York, NY.
Yang, X., Jin, H., Cai, X., Li, S., Shen, Y. (2012) Structural and mechanistic insights into the activation of Stromal interaction molecule 1 (STIM1). Proc Natl Acad Sci. 109(15): 5657-62.
Zheng, L., Stathopulos, P. B., Schindl, R., Li, G. Y, Romanin, C., Ikura, M. (2011) Auto-inhibitory role of the EF-SAM domain of STIM proteins in store-operated calcium entry. Proc Natl Acad Sci U S A. 108(4): 1337-42.
Zhang, C., DeLisi, C. (1998) Estimating the number of protein folds. J Mol Biol. 284(5): 1301-5.
Zheng, H., Zhou, M. H., Hu, C., Kuo, E., Peng, X., Hu, J., Kuo, L., and Zhang, S. L. (2013) Differential roles of the C and N termini of Orai1 protein in interacting with stromal interaction molecule 1 (STIM1) for Ca^{2+} release-activated Ca^{2+} (CRAC) channel activation. J. Biol. Chem. 288: 11263–11272.
Zhou, Y., Cai, X., Loktionova, N. A., Wang, X., Nwokonko, R. M., Wang, X., Wang, Y., Rothberg, B. S., Trebak, M., and Gill, D. L. (2016) The STIM1-binding site nexus remotely controls Orai1 channel gating. Nat. Commun. 7: 13725.
Zhuravlev, P. I., Papoian, G. A. (2010) Protein Functional Landscapes, Dynamics, Allostery: A Tortuous Path towards a Universal Theoretical Framework. Q. Rev. Biophys. 43: 295-332.

6. PUBLICATIONS

Paper I

Binding-competent states for L-arginine in E. coli arginine repressor apoprotein

Saurabh Kumar Pandey, David Řeha, Vasilina Zayats, Milan Melichercik, Jannette Carey, Rüdiger Ettrich

J Mol Model (2014) 20:2330.

Paper II

Allosteric activation of Arginine repressor protein by L-arginine

Saurabh Kumar Pandey, Milan Melichercik, David Řeha, Jannette Carey, Rüdiger H. Ettrich

Manuscript

Paper III

Communication between N terminus and loop2 tunes Orai activation

Marc Fahrner, **Saurabh K. Pandey**, Martin Muik, Lukas Traxler, Carmen Butorac, Michael Stadlbauer, Vasilina Zayats, Adéla Krizova, Peter Plenk, Irene Frischauf, Rainer Schindl, Hermann J. Gruber, Peter Hinterdorfer, Rüdiger Ettrich, Christoph Romanin, and Isabella Derler

J. Biol. Chem. (2018) 293(4) 1271–1285

Paper IV

Cholesterol modulates Orai1 channel function

Isabella Derler, Isaac Jardin, Peter B. Stathopulos, Martin Muik, Marc Fahrner, Vasilina Zayats, **Saurabh K. Pandey**, Michael Poteser, Barbara Lackner, Marketa Absolonova, Rainer Schindl, Klaus Groschner, Rüdiger Ettrich, Mitsu Ikura, Christoph Romanin

Science signaling. (2016) 9(412):ra10.

Paper V

Biomolecular Simulations in Structure-Based Drug Discovery: Ion Channel Simulations.

Pandey, S., Bonhenry, B., Ettrich, R.H.

Willey Publications, Gervasio, Francesco L., Spiwok, Vojtech (Editor), First Edition, pp: 247-280 (2018). ISBN:978-3-527-34265-5