Allosteric effects in protein dynamics and their interactions with membranes

By Devanand T LIFE10201404002

The Institute of Mathematical Sciences, Chennai

A thesis submitted to the Board of Studies in Life Sciences In partial fulfillment of requirements For the Degree of DOCTOR OF PHILOSOPHY

of

HOMI BHABHA NATIONAL INSTITUTE



January, 2020

Homi Bhabha National Institute

Recommendations of the Viva Voce Board

As members of the Viva Voce Committee, we certify that we have read the dissertation prepared by Devanand T entitled "Allosteric effects in protein dynamics and their interactions with membranes" and recommend that it may be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Margen	Date: 24 June 2020
Chairman - Prof. Gautam I. Menon	
atymani. N	Date: 24 June 2020
Guide/Convener - Prof. Satyavani Vemparala	Date: 24 June 2020
Examiner - Prof. Prabal Kumar Maiti	Date: 24 June 2020
Member 1 - Prof. Areejit Samal	Date: 24 June 2020
Member 2 - Prof. Rahu Siddharthan Member 3 - Prof. Prasanna Venkatraman	Date: 24 June 2020

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to HBNI.

I hereby certify that I have read this thesis prepared under my direction and recommend that it may be accepted as fulfilling the thesis requirement.

Date: 24 JUNE 2020

ani No Guide

Place: IMSc, Chennai

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at Homi Bhabha National Institute (HBNI) and is deposited in the Library to be made available to borrowers under rules of the HBNI.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgement of source is made. Requests for permission for extended quotation from, or reproduction of this manuscript in whole or in part, may be granted by the Competent Authority of HBNI when in his or her judgement the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

Devanand T

DECLARATION

I hereby declare that the investigation presented in the thesis has been carried out by me. The work is original and has not been submitted earlier as a whole or in part for a degree/diploma at this or any other Institution/University.

Dit

Devanand T

List of Publications arising from the thesis

Journal

- Phosphorylation promotes binding affinity of Rap-Raf complex by allosteric modulation of switch loop dynamics
 Devanand T, Prasanna Venkatraman, and Satyavani Vemparala
 Scientific reports 8.1 (2018): 12976.
- Interdigitation of lipids induced by membrane active proteins Devanand T, Sankaran Krishnaswamy, and Satyavani Vemparala Journal of Membrane Biology, 2019

Under preparation

 Phosphorylation of interfacial phosphosite: Effect on complex stability of Rap-Raf
Devanand T, Prasanna Venkatraman, and Satyavani Vemparala

Other publications (not included in the thesis)

1. Aggregation dynamics of charged peptides in water: Effect of salt concentration

Susmita Ghosh, **Devanand T**, Upayan Baul, and Satyavani Vemparala J. Chem. Phys. 151, 074901 (2019);

Dit

Devanand T

List of presentation and participation in conferences

- "Aspects of Gene Regulation 2014" A one day seminar IMSc Chennai, December 2014
- "3rd Workshop and Conference on Modeling Infectious diseases" IMSc Chennai, November 2015
- "ICTP-ICTS Winter School on Quantitative Systems Biology 2015" ICTS Bangalore, December 2015
- "Summer School Computational Approaches to Memory and Plasticity (CAMP)"

NCBS Bangalore, July 2016

• "Network theory: conceptual advances and practical applications" A one day seminar

 IMSc Chennai, April 2016

- "Aspects of Gene Regulation 2016" A two day seminar IMSc Chennai, August 2016
- "GIAN course on Nonequilibrium Statistical Mechanics and Molecular dynamics"

IIT Kharagpur, December 2016

• "The Interface of Biology and Theoretical Computer Science" A three day discussion meeting

NCBS Bangalore, December 2016

• "EMBO lecture course: Experimental and Theoretical Approaches to Cell Mechanics" NCDS & DBI Bargeland, April/Mar 2017

NCBS & RRI Bangalore, April/May 2017

- A Discussion Meeting "Recent advances in Molecular Simulations" IISc Bangalore, February 2018 | POSTER: Phosphorylation induces a unique conformation in Rap-Raf complex by allosteric modulation of switch loop dynamics
- "Indo-US conference on Sculpting the future of medicine Gateway to the postproteogenome era"

ACTREC Mumbai, December 2018 | POSTER: Phosphorylation promotes binding affinity of Rap-Raf complex by allosteric modulation of switch loop dynamics

• "Workshop on Free Energy Calculations For Chemical And Biological Systems"

IIT Kanpur, March 2019

 "CCP5/CECAM Summer School 2019"
Durham University, July 2019 | TALK: Membrane–Active protein Nogo66 induces interdigitation in phospholipid membranes

This Thesis is dedicated to my dear mom

San tha

Acknowledgements

I am indebted to multiple people in the journey towards the successful completion of this thesis. First of all, I would like to thank my Supervisor Prof. Satyavani Vemparala for her time and effort in mentoring me both academically and otherwise. I will be always thankful to her for bringing clarity to my thoughts during difficult times. Her extreme patience and support towards me during my mood swings helped me to overcome those tough periods. Besides her great mentoring, one of the things which I admire in her is the elegant teaching and science communication skills.

I would like to thank my collaborators Prof. Prasanna Venkatraman and Prof. Sankaran Krishnaswamy. I would also like to thank my doctoral committee members Prof. Gautam Menon, Prof. Areejit Samal and Prof. Rahul Siddharthan for their valuable suggestions and encouragement. I would like to thank Prof. Sitabhra Sinha for his exceptional teaching in our course work days.

I would like to thank IMSc for the High-Performance computing facilities, especially Annapurna and Nandadevi, and for library access. Thanks to members of IMSc administration and scientific staff for a very research-friendly environment at IMSc. Special thanks to Vishnu Prasad, Indra, Prema, Baskaran, Geetha, Gopi, Vasan, Mangal, and Jahir.

I would like to thank #STEMsocial science blogging community for supporting me and encouraging me to write about my research and technical works. Specifically, I would like to thank @lemouth, @jtm, @mobbs, @suesa, @mathowl, @scienceblocks, @katerinaramm, @gentleshaid, @samminator, @kingabesh and @scienceangel for their support and love towards me. I learned a lot about science communication from this community and it is always great to interact with them.

I thank friends here at IMSc and those whom I met via various academic workshops and conferences. Among IMSc friends, first of all, I would like to thank Prashanth Raman for giving timely and critical advice in the beginning of my PhD days which helped very much in choosing the right option. I always enjoyed the time with my dear friend Abdul Majith, who always will have his weird and different, but innocent thoughts about any topic. Time with comrade Javakrishnan was mostly spent on discussing politics and LGBT+ related stuff. My time with former office mates Raja, Joydeep and latter mates Kamal Tripathi and Karthik were enjoyable. Raja and Kamal were patient enough to suffer my frequent queries about a variety of topics. It was always fun to spend time with Deepika, Tuhin, Srivatsa, Arghya Mondal, Tanmay Mitra, Trilochan, Ankit Agrawal, Pritam, Vigneshwar, Prashanth, Ria Ghosh, Ria Sain, Shivani, Gaurav Sood, Rakesh, Sreevidya, and Vinay. I cherish all those cricket moments which happened in IMSc, specifically the moments with Sohan Lal Saini, Rakesh, Raghavendra and Sahil. When it comes to career advice, friends Rishu, Sudhir, Upayan Baul, Raja, Bijoy Daga, and Tanmay Mitra were very helpful. Discussions regarding statistical mechanics with Rishu, Rajesh, Kamal, Tanmay, Vigneshwar and Vinay were very helpful. One nostalgia thing which I will cherish always is my entry to an almost completely Bengali gang of TT players in room 123 and learning that game. I am thankful to Tuhin, Nirmalya, late Senti Imsong, Sunando, Anirban, Abhrajit and others for those good days.

Lastly but not least, I would like to thank my mother Santha and my sister Lakshmi for not just tolerating, but supporting my aim to do research at a very inappropriate timing and phase in my life. I should really thank my sister for supporting our family in my absence. I don't think I can pay back all the loneliness and health issues my mom suffered due to my absence. I also thank my Doctor Ganesan Venkatasubramanian of NIMHANS, Bangalore who routinely checked my status of mental well-being and gave timely advice. I would like to also mention happiness, my 1.5year-old nephew, Kunjunni brought to our home. Seeing his activities and innocent smile gives me hope for a better future.

Contents

Sy	/nops	sis		v
1	Intr	oducti	on	1
	1.1	Single	residue phosphorylation and its effects on Rap:Raf complex	2
		1.1.1	The Ras protein	3
		1.1.2	The Rap protein	4
		1.1.3	The regulator proteins GEF and GAP	7
		1.1.4	Motivation for the study	8
	1.2	Memb	rane–active protein:phospholipid membrane interactions	10
		1.2.1	Nogo66 : The membrane–active protein	11
		1.2.2	Lipid membranes and phases	12
		1.2.3	Motivation for the study	14
	1.3	Thesis	Organization	16
2	Met	thods		19
	2.1	All At	om Classical Molecular Dynamics	19
		2.1.1	Forcefield	21
		2.1.2	Basic steps of an MD simulation for biomolecules	25
	2.2	Analys	sis Techniques	30
		2.2.1	Covariance Analysis	30
		2.2.2	Principal Component Analysis	31
		2.2.3	Community Network Analysis	32
		2.2.4	MM-GBSA Free Energy Calculations	33

3	Pho allo	osphor steric	ylation promotes binding affinity of Rap-Raf complex by modulation of switch loop dynamics	37
	3.1	Introd	uction	37
	3.2	MD S	imulations	40
	3.3	Result	S	42
		3.3.1	Effects of phosphorylation on structure and dynamics of Rap .	42
		3.3.2	Conformational sampling of Rap: Effects of phosphorylation .	50
		3.3.3	Effects of phosphorylation on complex and interactions at the interface	55
		3.3.4	Effects of phosphorylation on binding energy of the complex .	59
	3.4	Discus	ssion	61
4	Pho bilit	osphor ty of R	ylation of interfacial phosphosite: Effect on complex sta- Cap-Raf	67
	4.1	Introd	uction	67
	4.2	MD S	imulations	69
	4.3	Result	S	71
		4.3.1	Effects of SER39 phosphorylation on structure and dynamics of Rap:Raf complex	71
		4.3.2	Effects of phosphorylation on the interactions at the Rap:Raf complex interface	73
		4.3.3	Effects of SER39 phosphorylation on binding free energy of the complex	74
	4.4	Discus	ssion and Concluding Remarks	75
5	Inte	erdigit	ation of lipids induced by membrane active proteins	77
	5.1	Introd	uction	77
	5.2	Metho	ds	81
		5.2.1	MD Simulations	81
		5.2.2	Lipid defect analysis	83
	5.3	Result	S	85
		5.3.1	Emergence of interdigitation in DMPC membrane patch with Nogo-66 interaction	85

	5.3.2	Interactions of Nogo-66 with DMPC membrane patch	. 93
5.4	Discus	ssion and Concluding Remarks	. 98
Biblio	graphy		103

List of Figures

Rap-Raf protein complex visualization (Synopsis)	viii
The communities detected in (a) GTP- and (b) GTP-PSER11 cases. (Synopsis)	viii
The communities detected in GTP-PSER39 case. (Synopsis)	ix
The time evolution of averaged interdigitation distance between ter- minal carbon atoms of lipid chains (Synopsis)	xiii
The time evolution of electrostatic and van der Waals interaction energy between Nogo-66 and DMPC membrane (Synopsis)	xiv
The MAPK/ERK pathway. See the Ras and Raf proteins in the cascade.	5
Depiction of Ras-like protein bound to GTP (active-state) and GDP (inactive-state).	6
Sequence similarity in Ras isoforms. Also note the unique differences in HVR region.	7
Ras isoforms binds to unique regions in lipid bilayers via HVR	7
A single DMPC lipid	13
Different lipid phases	13
A typical phase transition profile of a model membrane from DSC experiments looks like this plot.	15
The form of potential energy function used in all-atom classical molecular dynamics simulation codes(Forcefield)	21
The bonded and nonbonded interactions	23
The nature of the 6-12 Lennard-Jones(LJ) interaction.	24
Illustration of Ewald Summation.	26
	Rap-Raf protein complex visualization (Synopsis)

2.5	Illustration of Thermodynamic cycle utilized in MM-GBSA binding free enrgy calculations.	34
3.1	Rap-Raf protein complex visualization showing the location of important functional loops like P-loop, Switch I, Switch II, RBD loop regions and SER11 residue.	
3.2	Time evolution of root mean squared deviation(RMSD) of Rap with GTP ligand with and without SER11 phosphorylated	43
3.3	The RMSF of residues of Rap C_{α} atoms averaged over 350 to 400 ns of simulation for both unphosphorylated (green) and phosphorylated (blue).	44
3.4	RMSF of Rap chain carbon-alphas corresponding to the last 50ns of MD trajectories	45
3.5	Time evolution of SASA value of unphosphorylated SER11 residue in Rap protein (with GTP ligand).	45
3.6	Overlapped snapshots of GTP bound Rap protein showing the con- formation of functional Switch I and Switch II loops	47
3.7	The time evolution of (a) distance between GLY60(O) and GLY12(N) atoms, (b) distance between GLY60(center of mass) and GTP(O3G) atoms for both unphosphorylated (green) and phosphorylated (blue) cases.	48
3.8	Snapshots of nucleotide pocket of Rap-GTP with and without phosphorylation of SER11.	49
3.9	The time evolution of distance between the catalytic residue THR61 and GTP ligand without (green) and with (blue) phosphorylation	49
3.10	Cross-correlation plots of the Rap complex for 300 to 400ns trajectory data of GTP- and GTP-PSER11 cases.	52
3.11	Conformer plots of Rap domain (375 to 400ns simulation data)	53
3.12	Largest cavity location of Rap domain in original crystal structure and last 5ns averaged structures of different cases	53
3.13	The time evolution of (a) distance between Rap:TYR32(CA) and Raf:LYS84(CA) atoms, (b) distance between Rap:ASP38(OD1) and Raf:THR68(OG1) atoms.	57
3.14	Time evolution of distances between critical residues at Rap:Raf com- plex interface (Rap bound with GTP) with and without SER11 phos- phorylated.	58
3.15	The communities detected in a)GTP- and b)GTP-PSER11 cases	59

4.1	Rap-Raf protein complex visualization showing the location of impor- tant functional loops like P-loop, Switch I, Switch II and RBD loop regions and SER39 residue.	69
4.2	The RMSF of residues of (a)Rap and (b) Raf C_{α} atoms	71
4.3	The communities detected in a)GTP-SER11 and b)GTP-PSER39(sys2) and c)GTP-PSER39(sys3) cases.	72
4.4	Conformer plot of Rap:Raf complex considering the last 100ns of the trajectory.	72
4.5	The saltbridge distances between Rap and Raf domains for the last 100ns	74
5.1	A single DMPC lipid(hydrogen atoms not shown)	84
5.2	(a)Top view showing DMPC membrane and (b)Side view of mem- brane and protein of last frame in SIM1-300K simulation case	86
5.3	(a)Top view showing DMPC membrane and (b)Side view of mem- brane and protein(hydrogen atoms not shown) of last frame in SIM1- 310K simulation case	87
5.4	The time evolution of averaged interdigitation distance between ter- minal carbon atoms of lipid chains,C314, belonging to two leaflets for different simulation systems, described in Table 5.1 is shown	88
5.5	DMPC-Nogo-66 systems:Spatial distribution of the interdigitated gap parameter averaged over 2 X 2Å grids, perpendicular to the mem- brane normal, and further averaged over 25 ns(175 to 200 ns) of MD trajectory.	89
5.6	Lipid interdigitation map consider C314 tail atom done with a resolution of 2\AA averaged over 175 to 200ns MD trajectory	90
5.7	The distribution of (a) membrane thickness and (b) lipid chain tilt angles for various systems, averaged over 25ns (175-200 ns) of MD trajectory are shown	91
5.8	The order parameter plots corresponding to 175 to 200ns MD trajectory.	91
5.9	DMPC-Nogo-66 systems: cross correlation covariance maps spanning all the lipid molecules in the system.	92
5.10	Averaged cross-correlation map of P atoms from lipid heads over the duration 175 to 200ns of respective molecular dynamics trajectories	93
5.11	Density profiles corresponding to (a)SIM1(300K) and (b)SIM1(310K) with respect to the z-direction.	94

5.12	The time evolution of (a)electrostatic and (b)van der Waals inter- action energy between Nogo-66 and DMPC membrane for three in-	
	stances of simulations at 300K.	95
5.13	The interaction energy between the protein and the membrane systems.	96
5.14	The histogram of total "deep defects area" computed throughout the DMPC membrane trajectory.	97
5.15	(a)The average interdigitation distance plot and (b)VDW interac- tion energy plot between lipid tail atoms C314 and C214 with water molecules in 4 Å of those atoms.	99

List of Tables

1	Deep defect constants (Synopsis)	xiv
3.1	Details of systems simulated and analysed	42
3.2	The systems for which the pocket analysis is done	55
3.3	Free energy contribution of GTP liganded simulations	60
3.4	Quasiharmonic entropy contribution of GTP liganded simulations	61
4.1	Details of systems simulated and analysed	70
4.2	Free energy contribution (last 100ns) of GTP liganded simulations	74
5.1	Details of systems simulated and analysed	83
5.2	Deep defect constants.	98

Summary

In this thesis titled "Allosteric effects in protein dynamics and their interactions with membranes", we report classical all-atom molecular dynamics(MD) studies about conformational dynamics and allosteric signal propagation in Rap:Raf protein complex as a result of single residue phosphorylation and induction of interdigitation in lipid bilayer due to interaction with membrane–active proteins.

The effects of phosphorylation of a serine residue on the structural and dynamic properties of Ras-like protein, Rap, and its interactions with effector protein Ras binding domain (RBD) of Raf kinase, in the presence of GTP, are investigated via molecular dynamics simulations. The simulations show that phosphorylation significantly effects the dynamics of functional loops of Rap which participate in the stability of the complex with effector proteins. The effects of phosphorylation on Rap are significant and detailed conformational analysis suggest that the Rap protein, when phosphorylated and with GTP ligand, samples different conformational space as compared to non-phosphorylated protein. In addition, phosphorylation of SER11 opens up a new cavity in the Rap protein which can be further explored for possible drug interactions. Residue network analysis shows that the phosphorylation of Rap results in a community spanning both Rap and RBD and strongly suggests transmission of allosteric effects of local alterations in Rap to distal regions of RBD, potentially affecting the downstream signalling. Binding free energy calculations suggest that phosphorylation of SER11 residue increases the binding between Rap and Raf corroborating the network analysis results. The increased binding of the Rap-Raf complex can have cascading effects along the signalling pathways where availability of Raf can influence the oncogenic effects of Ras proteins. These simulations underscore the importance of post translational modifications like phosphorylation on the functional dynamics in proteins and can be an alternative to drug-targeting, especially in notoriously undruggable oncoproteins belonging to Ras-like GTPase family.

As an extension to the aforementioned study, we used atomistic molecular dynamics simulations to look into the effect of phosphorylation of SER39 residue in the Rap protein of Rap:Raf complex. Since SER39 being a residue in the Rap:Raf interface, we explored two cases. Blind phosphorylation of Rap was the first case for simulation. And in the next case, Rap alone with SER39 phosphorylation was simulated first and later joined it with Raf for further simulations. The latter resulted in formation of stable salt bridges between Rap and Raf proteins. Also the free energy calculations indicate better stability between these proteins. Network analysis studies clearly shows that the SER39 phosphorylation reorganizes the community network to include the entire region of Raf chain comprising of Raf L4 loop which can take the signal downstream in the pathway.

The next study was about the membrane-active protein Nogo-66 and its interactions with dimyristoylphosphocholine membranes. Nogo-66 is found to induce interdigitation in dimyristoylphosphocholine membranes. Extensive molecular dynamics simulations have been employed to probe the interactions of Nogo-66 with these model membranes. This phase change happens when the temperature is close to the main transition temperature of the membrane (T_m) and only in the presence of the protein. No similar interdigitation of the membrane lipids was observed temperatures well above T_m in the presence of the protein. In addition, in protein-free simulations, no interdigitation of the membrane lipids was found both at temperatures near or well above T_m indicating that the observed effect is caused by the interactions of Nogo-66 with the membrane. Analysis of the simulations suggest protein-membrane interactions, even if transient, alter the lifetimes of lipid head defects and can potentially alter the effective T_m and cause interdigitation. This study emphasize the importance of membrane-active proteins and their interactions with membranes leading to phase transitions which would affect other membrane-related processes such as domain formation.

Synopsis

1 Introduction

Proteins are the workhorse molecules that enable the cells to function. It is important to study the conformational dynamics of proteins, both globular and membraneassociated ones. Post-translational modifications (PTMs) and mutations even at single residue resolution in proteins are capable of bringing global changes in the conformational dynamics via long-range signal cascades like allostery mechanisms. Binding affinity between proteins in protein complexes can also be modulated under PTMs or mutations[1]. On the other hand, interactions of proteins with cell membranes (however transient) are known to induce structural and dynamical changes in cell membranes. Examples include membrane-active peptides such as antimicrobial peptides or intrinsically disordered proteins[2, 3]. Biopolymers such as proteins and RNA have been shown to induce phase changes when they interact with membranes[4, 5]. For deeper understanding of the above mentioned biologically relevant processes, we need to study them at an atomistic resolution.

In this synopsis, we discuss two problems which are addressed using the all-atom classical molecular dynamics techniques:

• Conformational dynamics and allosteric signal propagation in Rap:Raf protein complex as a result of single residue phosphorylation in Rap protein

• Induction of interdigitation in lipid bilayer due to interaction with membraneactive proteins

2 Single residue phosphorylation induces conformational dynamics and allosteric signal propagation in Rap:Raf protein complex

Ras superfamily of small GTPase proteins is a major player in many cellular functions like cell proliferation, cytoskeletal morphology, and cellular adhesion, etc[6, 7]. The GTPases act as molecular switches in the sense that they get activated or inactivated depending on nucleotide bound to them. They gets activated and deactivated when bound to guanine nucleotides GTP (guanosine triphosphate) and to GDP (guanosine diphosphate) respectively [8]. Few prominent proteins belonging Ras superfamily include Ras protein, Rap protein, Ral protein and Rho protein etc. Ras genes were first identified as transduced oncogenes by Harvey and Kirsten [9, 10]. These small proteins are of the order of 20-25kD and have two main domains namely G-Domain and HVR (Hyper Variable region) containing a linker and a C-terminal region which interacts with plasma membrane^[11]. Ras proteins exist in 3 major isoforms called H-Ras, K-Ras, and N-Ras and are crucial players in the MAPK/ERK pathway. MAPK/ERK pathway is very significant in cell division, apoptosis and processes like differentiation and the signal starts from a EGF (Extracellular Growth factor) at cell membrane propagates towards nucleus from cell membrane^[12]. Although the major effecter lobe of the Ras isoforms share almost same sequence, they have unique, but overlapping functions and this uniqueness can be assigned to the HVR (hypervariable region) which is different for each isoform and associates these proteins to interact towards different regions in cellular membrane via C-terminal tails^[13]. These proteins have key regions like Switch-I (25-40 positions in amino

acid sequence), Switch-II (57-74) and P-loop (10-17) regions (see Fig.1). The ligand (GTP/GDP) site along with a Mg^{2+} ion is situated between the Switch-I. Switch-II and P-loop. The Rap proteins, which belong to the Ras family are structurally similar to the Ras proteins with aminoacids ALA11 and GLN61 in Ras replaced with SER11 and THR61 respectively for Rap proteins. Few isoforms of Rap protein include Rap1A, Rap1B, Rap2A, and Rap2B[14]. Rap proteins are antagonists to Ras proteins, in the sense that they bind competitively with the effector MAP-KKK (MAP kinase kinase kinase) protein Raf (Rapidly Accelerated Fibrosarcoma). Other important functions of Rap proteins include modulating Hedgehog signaling which is crucial in morphogenesis, coordinates cell migration and cell-cell adhesion etc[15, 16, 17]. Raf kinase has three conserved regions namely RBD (Ras binding domain which is very similar to ubiquitin fold), CRD (Cysteine Rich Domain) and Kinase domain. The CRD region has a tendency to anchor itself to the plasma membrane thus stabilizing the interaction with Ras-like proteins. Raf has three isoforms called A-Raf, B-Raf, and C-Raf and these proteins are involved in processes like cell cycle regulation, cell proliferation, apoptosis, and development etc[18].

The first problem which we investigate in this thesis is about the conformational dynamics and allostery signal propagation as a result of a single residue phosphorylation in Rap:Raf complex. Ras proteins are considered to be notoriously "undruggable" i.e., lacking specific binding pockets which can host drug molecules[19]. Malignant Ras proteins can cause cancer and so exploring alternate techniques which can inhibit such oncoproteins is essential. In this problem, we study the effects of phosphorylation of a single residue (SER11 in P-loop and SER39 in Switch-I loop at the interface of Rap:Raf complex) of GTP liganded Rap protein on conformational dynamics and allosteric signal propagation in Rap:Raf complex using all-atom classical molecular dynamics simulations. In addition to MD simulations, techniques such as free energy calculations, residue interaction community network analysis, pocket analysis and PCA are used.



Figure 1: Rap-Raf protein complex (crystal structure, PDB ID 1C1Y) showing the location of important functional loops like P-loop, Switch I, Switch II and RBD loop regions.



Figure 2: The communities detected in (a) GTP- and (b) GTP-PSER11 cases. The yellow community represents the L4 loop.

Most often phosphosites are located in well accessible or highly flexible regions and loops in proteins[20, 21]. However, while analyzing large scale data sets, proteins for which many of these phosphosites are not readily exposed to the solvent are also observed. SER11 is found to be a buried site from the solvent accessible surface



Figure 3: The communities detected in GTP-PSER39 case.

area (SASA) calculations. These disallowed phosphosites might require further investigations using techniques like Normal mode analysis or MD simulation to see if protein dynamics expose buried sites[22]. The systems studied using molecular dynamics simulations were GTP liganded Rap protein bound with Raf-RBD with and without SER11 phosphorylation. The simulations show that phosphorylation significantly effects the dynamics of functional loops (Switch-I, Switch-II and Ploop) of Rap protein which participate in the stability of the complex with the effector Ras binding domain (RBD) of Raf protein. It also shows that in simulations of non-phosphorylated Rap, SER11 was getting exposed to the solvent, which means that the inherent protein dynamics can potentially expose this site to relevant kinases. We performed different analyses to understand the effect of SER11 and SER39 phosphorylations individually in Rap:Raf complex. Some of the major results are summarized below:

• Community detection [23, 24] analysis revealed that there are 4 common communities spanning both Rap and Raf domains. (See Fig.2 b) This implies that allosteric signal can propagate from SER11 of Rap protein till the distal L4 loop in Raf domain which is around 40Å far. The L4 loop in Raf is important since it communicates with the downstream CRD (Cysteine rich domain). Hence this finding suggests that even a single residue phosphorylation can affect downstream signalling.

- Principal component analysis of Rap domain in phosphorylated and nonphosphorylated cases revealed that the conformational space sampled by the SER11 phosphorylated Rap is smaller and very different from the non-phosphorylated case. This suggests that Rap protein is less dynamic when SER11 is phosphorylated.
- Analyzing different functional pockets in phosphorylated case revealed that there is a unique pocket in Rap, which can be a potential target, unlike the non-phosphorylated case.
- Free energy calculations reveal that SER11 phosphorylation promotes binding affinity between Rap and Raf-RBD. This increased binding affinity of Rap protein towards Raf protein compared to Ras protein suggests that it can have important consequences in the downstream MAPK signaling pathways. Increased binding of Rap protein to Raf-RBD, due to SER11 phosphorylation, can potentially make Raf protein even less available to Ras protein. Free energy calculations reveal that SER39 phosphorylation also prmotoes binding affinity between Rap and Raf-RBD.
- For SER39 phosphorylation simulations, community detection analysis revealed that there are 3 common communities spanning both Rap and Raf domains. (See Fig.3) One of the communities (colored in red) span SER39 node in Rap and almost all of Raf residues. This imply that the allostery signal can propagate from SER39 of Rap protein till L4 loop in Raf domain via different paths suggesting alternate paths of propagation.

These results strongly suggest that phosphorylation of potential and identified phosphosites of Rap protein can be used as an alternate technique to inhibit antagonists like malignant and "undruggable" Ras proteins, since phosphorylated Rap proteins can bind to Raf protein more competitively making Raf less available to Ras proteins, thus affecting the downstream MAPK pathway. We believe that these case studies are examples of how integrating tools that can probe dynamics can yield a wealth of biological information hidden in crystal structures using high-throughput studies.

3 Structural changes in membranes on interaction with Membrane–active agents

In biology, lipid membranes play a very important role in compartmentalizing functional units whether they be cells or organelles inside the cell. Lipid membranes are built mainly from amphipathic phospholipid molecules, and also other molecules like cholesterol and membrane proteins. These lipid molecules being amphipathic, in general, have a polar and hydrophilic head structure and non-polar and hydrophobic hydrocarbon tail chains. One important aspect of the cell membranes is the nature of the phase in which they exist, which can affect their function [25, 26]. The lipid bilayers are known to exhibit varying degrees of lipid molecule packing resulting in different phases. Lamellar lipid bilayers exist predominantly in two phases: gel (L_{β}) and liquid-crystalline or fluid (L_{α}) phases. Gel phase is observed at low temperatures and is characterised by ordered lipid molecules with lipid tails tilted or untilted with respect to bilayer normal. Gel phase can also exist in an interdigitated ($L_{\beta}I$ phase) state where the upper and lower membrane leaflets overlap with each other [27]. The fluid phase is the most common phase observed and is required for many normal biological functions, in which the lipid tails are disordered and along the membrane normal and it is known that certain proteins have the ability to sense the fluidity of the membranes for interaction [28]. Hence the lipid bilayers exhibit thermotropic behavior and change phases as the temperature is increased and the critical temperature at which the lipid bilayer changes from gel to fluid phase is called the main transition temperature (T_m) and varies with many factors including lipid chain length, degree of unsaturation in the lipid tails, water content and possibile mixing of more than one lipid species etc[29]. Apart from the gel and fluid phase, a third phase called ripple phase (P_β) has been discovered in 1973[30] and are corrugated in nature with coexisting gel and fluid phases[31]. Certain microdomains in lipid membranes called "lipid-rafts" enriched with cholesterol and sphingolipids exhibit very ordered structures among the lipid tails (known as Liquid ordered state). These lipid domains play a key role in signal transduction and modulate a variety of biochemical pathways[32].

There are broadly two classes of membrane proteins: peripheral and integral proteins. Peripheral proteins interact with membrane surface by interacting mostly with the lipid head groups. Some of the classic examples include antimicrobial peptides (AMPs) like melittin [33], heat shock protein Hsp12[5], etc. They are capable of rupturing the membrane (as in the case of AMPs), inducing a phase transition in membranes (as in the case of Hsp12), remodeling the membrane in form of domain formation and membrane thinning, etc[34, 3, 5]. These peripheral proteins which are capable of inducing membrane remodeling or folds at the interface of lipid head groups are also referred to as membrane-active proteins in literature. Integral proteins, on the other hand, are transmembrane proteins like ion channels and immune receptors *etc* which are capable of modulating lipid order based on the hydrophobic mismatch between the transmembrane protein and the membrane [35] and inducing microdomains in lipid membranes^[36]. In this work, we explore the phase changes in membranes when membrane-active proteins interact with them and the role of the temperature using Nogo66 as an example. Nogo is a protein which is identified to inhibit the axonal regeneration in the central nervous system (CNS) and is expressed by oligodendrocytes (a type of glial cell) which is a component in CNS myelin sheath. Nogo66 is a 66 residue extracellular domain of Nogo protein, which inhibits axonal extension [37, 38]. Nogo66 is an intrinsically disordered protein (IDP) when in solution, but requires interaction with phosphocholine (like DMPC, 1,2-dimyristoyl-sn-glycero-3-phosphocholine) lipid membranes to attain a folded structure. Nogo protein is also found in Endoplasmic Reticulum (ER) and stabilizes the curvature of ER membrane[2, 39].



Figure 4: The time evolution of averaged interdigitation distance between terminal carbon atoms (considering both C314 and C214 atoms) of lipid chains, belonging to two leafets for different simulation systems.

Here, we describe our work which shows that a membrane–active protein can induce interdigitation among DMPC (dimyristoylphosphocholine) lipid bilayer even beyond the main-transition temperature. Extensive MD simulations were done with the Membrane–active protein Nogo66 on DMPC membrane patch (in presence of water and counter-ions). Some of the important results are summarized below:

• We observe interdigitation among lipids of both upper and lower leaflets, membrane thinning and formation of asymmetric ripple even beyond the maintransition temperature of DMPC lipids ($T_m = 297K$). Nogo66 plays a key role in membrane remodeling via inducing regions of interdigitated lipids. Proteinfree simulations beyond T_m do not show any interdigitation, meanwhile, three



Figure 5: The time evolution of a electrostatic and b van der Waals interaction energy between Nogo-66 and DMPC membrane for three instances of simulations at 300 K. The time period over which Nogo66 experiences most favourable electrostatic and van der Waals energies, simultaneously, with the DMPC membrane are highlighted

simulations of Nogo66 with DMPC membrane simulations with different initial conditions reproduced asymmetric ripples. At a higher temperature of 310K, no interdigitation was observed.

No.	System	Deep Defect Constant (Å ²)
1	CTRL(300K)	7.79 ± 0.05
2	SIM1(300K)	13.22 ± 4.58
3	SIM1(310K)	9.48 ± 0.81

Table 1: Deep defect constants.

• Lipid defects were observed while investigating the Nogo66 with DMPC membrane (at 300K) trajectory which revealed interdigitation. The transient interaction of lipids with protein was inducing unfavourable exposure of the lipid acyl chain groups to water and such spatial regions of hydrophobic exposure are main constituents of interfacial lipid packing defects. Deep defects are those membrane voids where aliphatic atoms are deeper than a threshold distance (usually 1 Å), determined by the position of glycerol carbon atom. The probability, p(A), of finding a defect with area A is given by

$$p(A) = b \times e^{-A/\pi} \tag{1}$$

where b is the pre-exponential factor and π is the defect constant (in Å²), with larger values of π indicating higher probability of finding larger defects. The defect constant π in Eq. 2.1 is a parameter that indicates probability of finding larger defects and indirectly estimates the survival of the defects measured. This value is computed for the three systems shown in Table 1. The larger value of π for DMPC+Nogo-66 system at 300K compared to both control system at 300K and membrane-protein system at 310K strongly indicates higher survival of the defects when Nogo-66 is interacting with the DMPC bilayer at 300K, close to transition temperature T_m .

• The interaction of Nogo-66 with the membrane is not very strong and during the MD trajectories, in all three instances at 300K, there are time periods, lasting 20-30 ns when both electrostatic and van der Waals interactions between Nogo-66 and DMPC membrane are very favourable. (See Fig.5) In all the three simulations, the onset of interdigitation coincided with these very favourable interaction time periods. (See Fig.4)

This study is important in the sense that protein/peptide induced membrane remodeling via interdigitation or phase transition to ripple or liquid ordered phase *etc* can initiate formation of microdomains or "lipid-rafts" *etc* which plays important role in various signal transduction processes in biology.

Chapter 1

Introduction

Proteins are the workhorses in the cells and these molecules enable us to function, whether it be seeing this world through our eyes or protecting our body from harmful viruses and bacteria or accelerating biochemical reactions in the form of enzymes etc. Learning about how proteins function can give key insights into the process of fighting against diseases. Targeting a malfunctioning protein via a specific small molecule drug can be a strategy to tackle a particular disease. But there are cases when proteins are "undruggable" [19] and such cases need alternate strategies.

Another important component in biology is lipid membranes. Lipid membranes not only function as the outer covering of cells and organelles but also play an active role in biological processes like cell adhesion and cell signaling etc. The lipid bliayer membranes exhibit different phases under different temperature conditions. As temperature increases, the lipid bilayer membrane transitions from a very ordered 'gel' phase to very disordered 'fluid' phase at a characteristic temperature called main-transition temperature T_m . The fluidity of the membrane is modulated by different parameters like the content of cholesterol in the membrane, interaction of membrane with other agents like alcohol, RNA, nanoparticles, and proteins etc[40, 4, 5, 41, 42, 43]. In this thesis, we focus on two problems. The first problem is about the allostery propagation due to single residue phosphorylation on the Rap:Raf protein complex. The second problem is about a membrane-active protein Nogo66 inducing interdigitation in the phospholipid membrane. In the next two sections of this chapter, we will explain about Ras family of proteins (including Rap protein), GEF and GAP regulator proteins, Raf protein, Nogo66 protein, and DMPC (1,2-Dimyristoyl-sn-glycero-3-phosphocholine) lipid membranes. The final section of the chapter focuses on the details of the thesis organization.

1.1 Single residue phosphorylation and its effects on Rap:Raf complex

The proteins belonging to Ras superfamily are small GTPases which are major players in many cellular functions like cell proliferation, cytoskeletal morphology, and cellular adhesion etc. These proteins take part in the crucial MAPK/ERK (mitogen-activated protein kinases or extracellular signal-regulated kinases) pathway (See Figure 1.1) which is very important in cellular processes like cell division. Missense mutations in Ras proteins have been implicated in variety of cancers and hence are also oncoproteins. Over 30% of human cancers exhibit oncogenic Ras mutations[44]. These GTPases which function as molecular switches get activated and inactivated based on whether they are bound to guanine nucleotides GTP (guanosine triphosphate) or GDP (guanosine diphosphate) respectively. A GTP bound active Ras protein can get inactivated after the event of bound GTP getting hydrolyzed to GDP. The activity of these proteins is also regulated by other proteins such as GEFs (Guanine nucleotide exchange factors) and GAPs (GTPase-activating proteins), which can accelerate the process of activating and deactivating the GTPase molecular switches respectively. Few prominent proteins belonging Ras superfamily are Ras protein (including the isoforms H-,K- and N-Ras), Rap protein, Ral protein and Rho protein etc. See Figure 1.2 (a) for a cartoon depiction of a Ras-like protein under active and inactive states. Figure 1.2 (b) shows a H-Ras (an isoform of Ras protein) system where the Switch-I and Switch-II are closely bound to P-loop region via the THR35 and GLY60 residues with the aid of GTP analogue and Magnesium ion. In the case of GDP bound H-Ras it is clearly visible that the configuration of THR35 and GLY60 is in the opposite direction of the GDP ligand. The GTP-bound active Ras-like protein show higher affinity towards its effector protein (such as Raf protein) as compared to corresponding GDP-bound inactive form.

Proteins belonging to common Ras family are involved in a diverse functions like cell growth, differentiation, and survival processes. The functions of Rap proteins include cell adhesion, cell-cell junction formation and regulation of the actin cytoskeleton. Ral proteins are involved in biological processes like cell motility, apoptosis, cell division, etc. Rho proteins play a key role in the processes like cytoskeletal remodeling and deciding the cell polarity etc.[45, 46, 47] In the coming subsections, we will explain about Ras and Rap proteins.

1.1.1 The Ras protein

Ras (rat sarcoma) genes were first identified as transduced oncogenes by Harvey and Kirsten[9, 10]. These are small proteins in the order of 20-25kD and having two main domains namely G-Domain (which is the main catalytic unit) and HVR (Hyper Variable region) containing a linker and a C-terminal region which interacts with plasma membrane[11]. Ras proteins exists in 3 major isoforms called H-Ras, K-Ras, and N-Ras and are crucial players in the MAPK/ERK pathway (See Figure 1.1). Although the major effecter lobe of the isoforms share 100% sequence identity, they have unique, but overlapping outputs to decrease functional redundancy. See the Figure 1.3. And this uniqueness can be assigned to the C-terminal tail region
called HVR (hypervariable region) which is unique for each isoform and assists these proteins to interact with different regions in cellular membrane (See the Figure 1.4). These proteins have key regions like Switch-I (25-40 positions in amino acid sequence), Switch-II (57-74) and P-loop (10-17) regions. The ligand (GTP/GDP) site along with a Mg²⁺ ion is in between the regions of Switch-I, Switch-II and P-loop. The dynamic nature of Switch I and Switch II loops and the switching between active and inactive states has been long established. Switch I and Switch II are interaction sites to the GEF, GAP, and downstream effector proteins[48, 49]. Although the classic picture of Ras protein is switiching between two states, it was shown that GTP-bound active state is distributed between two states in which one has weak coordination of THR35 with γ - phosphate and Magnesium ion resulting in a displacement of Switch I away from the binding pocket and the other dominant one which has active form of the Ras protein with higher affinity for downstream effectors and a higher GTPase activity[50, 51].

1.1.2 The Rap protein

The Rap proteins, which belongs to the Ras family of proteins are very similar structurally with the Ras proteins. Amino acids ALA11 and GLN61 in Ras are replaced with SER11 and THR61 respectively for Rap proteins. Rap proteins are antagonists to Ras proteins, in the sense that they competitively bind with the effector MAP-KKK (MAP kinase kinase kinase) protein Raf. Other important functions of Rap proteins include modulating Hedgehog signaling which is crucial in morphogenesis, coordinates cell migration and cell-cell adhesion etc[15, 16, 17]. Few isoforms of Rap protein include Rap1A, Rap1B, Rap2A, and Rap2B.

Similar to Ras, Rap consists of domains which are involved in the binding of the guanine nucleotide, GTP hydrolysis, binding of regulatory proteins GAPs, GEFs and downstream effector proteins etc. The main domains involved in nucleotide



Figure 1.1: The MAPK/ERK pathway. See the Ras and Raf proteins in the cascade. GTP liganded Rap protein is shown to represent that it competitively tries to bind with Raf protein. This figure is adapted from wikimedia commons page about MAPK/ERK pathway, which is under Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0) license.

binding, GTP hydrolysis and GAP, GEF and effector binding are the P-loop motif (amino acids 10-17)[53] and the Switch I (amino acids 25-40) and Switch II (amino acids 57-75) domains.[54, 55] These switch loops are highly mobile regions that change conformation depending on the bound nucleotide.

It has been found to be a repressor to Ras proteins and thus inhibits Raf/MEK/ERK pathway. Rap and Ras proteins have been found to have a very similar effector binding domain (Ras binding domain or RBD) which indicates that they might have common downstream effector proteins. Rap1 protein is capable of interacting with c-Raf protein and form a inactive complex, thereby inactivating Raf/MEK/ERK



Figure 1.2: (a)A cartoon depiction of Ras-like protein bound to GTP (active-state) and GDP (inactive-state).(b) An example system visualized using VMD. X-RAY crystal structures H-Ras (a Ras isoform) in inactive GDP bound (PDB ID: 4Q21) and active GTP analogue bound (PDB ID: 5P21) states. The transitions between active and inactive bound are regulated by GEF and GAP proteins. The yellow colored molecule is GDP and GTP analogue respectively, blue colored region is Switch-I, red colored region is Switch-II and orange colored region is P-loop. THR35 and GLY60 residues are black and magenta respectively.

pathway[56]. On one hand Rap1 interaction with c-Raf inactivates Raf/MEK/ERK pathway, but activates it under the interaction with c-Raf's homologue, B-Raf[57]. Initially Rap was thought to have only inhibitory role to Ras proteins. But now it



Figure 1.3: Sequence similarity in Ras isoforms. Also note the unique differences in HVR region. This figure was taken from Ian A. Prior et al paper[52] with permissions from author and journal of Cancer Research.



Figure 1.4: Ras isoforms binds to unique regions in lipid bilayers via HVR (shown as the yellow tail of protein). You can note that even the bound ligand (GTP/GDP) modulates the region of interaction. This figure was taken from Jillian A. Parker et al paper[13] with permission from journal of Molecular Cancer Research.

is established that Rap takes part in cell-adhesion via integrins and cadherins etc.

1.1.3 The regulator proteins GEF and GAP

As mentioned before, GEFs (Guanine nucleotide exchange factors) and GAPs (GTPaseactivating proteins) regulate the activity of small GTPases like Ras family proteins

to control their cellular functions. GEFs switches ON signalling by catalyzing the exchange of bound GDP for GTP in G-protein and GAPs switches OFF the signalling by accelerating GTP hydrolysis. There are different GEFs and GAPs to modulate the regulation of multiple G-proteins. GEFs accelerate the dissociation of the GDP from the G-protein by altering the nucleotide-binding site such that the nucleotide affinity is reduced and, thus, the GDP is released and subsequently replaced by GTP. The affinity of the GTP are for GTP and GDP are similar, and so GEF does not favour GTP or GDP. The increase in GTP-bound GTPase over GDP-bound GTP is due to the higher concentration of GTP compared to GDP in cell. On the other hand different GAPs employ different techniques for GTP hydrolysis. Residues from GTPase and GAP take part in hydrolysis process. In the case of Ras:RasGAP complex, RasGAP stabilizes the GLN61 of Ras, which in turn coordinates the water molecule for an in-line nucleophilic attack to the γ - phosphate. An arginine residue (from GAP), called the arginine finger, is positioned into the phosphate-binding site and stabilizes the transition state by neutralizing negative charge at the γ – phosphate. The proper catalysis of phosphoryl transfer by GAPs involves the right orientation of the attacking water molecule, avoiding water from the active site, and the stabilization of the transition state.[8]

1.1.4 Motivation for the study

As previously mentioned Rap is an antagonist to Ras, when it comes to binding to the effector protein Raf (Rapidly Accelerated Fibrosarcoma). Raf kinase has three conserved regions namely RBD (Ras binding domain which is very similar to ubiquitin fold), CRD (Cysteine Rich Domain) and Kinase domain. The CRD region has a tendency to anchor itself to the plasma membrane thus stabilizing the interaction with Ras-like proteins. Raf protein has three isoforms called A-Raf,B-Raf, and C-Raf. These proteins are involved in processes like cell cycle regulation, cell proliferation, apoptosis, and development, etc[18].

Ras proteins are known to be notoriously undruggable and no inhibitors have been discovered yet [19]. The term 'undruggable' is used to describe proteins that could not be targeted pharmacologically and in the context of cancer, proteins like Ras and Myc can be classified as undruggable[58]. So the research works for finding inhibitor molecule and corresponding pocket for docking is still going on. Although recently there are studies which reported novel pockets and small molecule inhibitors, they are yet to pass the clinical trials. One such example is an inhibitor molecule called BI-2852 (1) which is capable of targeting KRas[59]. It is necessary to investigate other strategies which can be used to inhibit mutated Ras proteins. One of the mechanisms in particular and to target similar undruggable proteins in general is post-translational modifications (PTMs). Proteins are synthesized as polypeptides from the ribosome machinery. Many of these proteins undergo PTMs and that too at different points in their life cycle. PTMs occur at sidechains or terminals of the proteins. Protein phosphorylation is the most well-studied reversible PTM, especially with respect to its functional characterization. Kinases catalyze the protein by phosphorylating it and phosphatases reverse the process by dephosphorylation. This reversible feature of protein phosphorylation makes it possible for a protein to act as an ON/OFF molecular switch. A classical example is the plasma membrane H⁺-ATPase AHA2 for which a total of eight different phosphorylation sites have been identified using mass spectrometry experiments. [60, 61] Two of these sites increase proton pump activity, whereas two other sites inhibit activity. PTMs like phosphorylation can induce an allosteric signal in the protein structure thereby modulating the local or global dynamics of protein [62, 63]. Earlier studies show that phosphorylation induces conformational changes of proteins via disrupting hydrogen bond network^[64]. Such conformational changes affect distal regions in a protein or a protein-protein complex and this is called allostery.

Most often, phosphorylation is found at a site which belongs to well accessible and disordered regions like loops in the proteins. However, while analyzing large scale data it has been documented that many of these phosphosites are not readily exposed to the solvent [22, 65, 66]. The presence of significant number of possible phosphosites in the disallowed region of phosphoconformation suggests that factors including dynamic conformational changes of proteins, binding to other proteins or regulatory factors can potentially expose these buried phosphosites to the solvent and to a kinase. Therefore understanding the role of protein dynamics in exposing such sites to solvents, interactions involved in transmitting the effects of phosphorylation to other functional sites is essential. A Possible way to modulate Ras functioning would be to disrupt Ras:Raf interaction. We want to explore the possibility of increasing Rap:Raf interaction, via PTMs in Rap, and effectively make Raf less available to Ras. Potential phosphosites including SER11 and SER39 were identified earlier^[67]. The effects of PTMs like phosphorylation on Rap:Raf complex (using structures available in protein data bank) can be studied at atomistic resolution using Molecular Dynamics as the tool.

1.2 Membrane–active protein:phospholipid membrane interactions

There are broadly two classes of membrane proteins: peripheral and integral proteins. Peripheral proteins interact with membrane surface via interacting mostly with the lipid head groups. Some of the classic examples are antimicrobial peptides (AMPs) like melittin, heat shock protein Hsp12, etc. They are capable of rupturing the membrane (as in the case of AMPs), inducing a phase transition in membranes (as in the case of Hsp12), remodeling the membrane in form of domain formation and membrane thinning, etc[34, 3, 5]. These peripheral proteins which are capable of inducing membrane remodeling or folds at the interface of lipid head groups are also referred to as membrane–active proteins in literature[68, 69]. Integral proteins are transmembrane proteins like ion channels and immune receptors etc. These proteins are capable of modulating lipid order based on the hydrophobic mismatch between the transmembrane protein and the membrane[35] and inducing microdomains in lipid membranes[36].

Nogo66, a segment of Nogo protein is a membrane–active peptide which transitions from disordered to folded structure upon interacting with phosphocholine membrane surface[2]. Nogo protein is also found in Endoplasmic Reticulum (ER) and is known for stabilizing the curvature of ER membrane[2, 39]. In the next sections, we will introduce the membrane–active protein Nogo66, lipid membranes, and interaction between them, etc.

1.2.1 Nogo66 : The membrane–active protein

Nogo-A, also known as reticulon 4-A is identified to inhibit the axonal regeneration in the central nervous system. Two domains of Nogo-A have been identified that inhibit neurite growth and cell migration: Nogo-A- $\Delta 20$ and Nogo66[70]. Nogo-A- $\Delta 20$ (residues 544-725 of the Nogo-A sequence) is a 182 residue long domain located in the middle of the 803 residues long Nogo-A-specific segment. Nogo66 (residues 1026-1091 of the Nogo-A sequence) is a 66 residue extracellular domain situated between two long hydrophobic transmembrane regions at the C-terminus of Nogo-A, which inhibits axonal extension[37, 38]. Neurons in the central nervous system (CNS) are known for not regenerating upon neurotrauma like spinal cord or brain injury. This is contrary to the fact that neurons in peripheral nervous system (PNS) regrow nerve fibres to a much higher extent[71, 72]. The inhibitory molecules in myelin associated with CNS are the reason for this property[73]. The three proteins present in CNS myelin, Nogo, MAG and OMgp are responsible for growth cone collapse and inhibiting neurite outgrowth in vitro[74]. Unlike other Nogo inhibitory domains, the structure of Nogo66 is known via NMR technique. Nogo66 is intrinsically disordered when free in solution, but requires interaction with phosphocholine (for example, 1,2-Dimyristoyl-sn-glycero-3-phosphocholine lipid) lipid membranes for getting folded into a five-helix bundle[2]. This along with the aforementioned property of modulating ER membrane curvature makes Nogo66 an ideal model system to computationally study about membrane–active protein:phospholipid interactions. Recently, a study showed that GLU26 residue plays a key role in stabilizing the helical conformation of the lipid-bound state via interactions with choline. Also it was seen that GLU26ALA mutation improved the helicity of the protein[75].

1.2.2 Lipid membranes and phases

In biology, lipid membranes play a very important role in compartmentalizing functional units whether it be cells or organelles inside the cell. These lipids being amphipathic, in general, have a polar and hydrophilic head structure and two non-polar and hydrophobic hydrocarbon tail chains (See Figure 1.5). Cell membranes consist of lipids like phosphocholines, cholesterol and sphingolipids and other molecules. Cholesterol is a sterol, a type of lipid which modulates the fluidity of cell membranes and sphingolipids, on the other hand are a class of lipids containing a backbone of sphingoid bases, a set of aliphatic amino alcohols that includes sphingosine. Certain microdomains in lipid membranes called "lipid-rafts" enriched with cholesterol and sphingolipids exhibit very ordered structures among the lipid tails when compared to the surrounding lipids (known as Liquid ordered state). The aforementioned order of the tail (acyl chain) is quantified by a parameter called deuterium order parameter which is defined as $S_{CD} = \langle \frac{3cos^2(\theta)-1}{2} \rangle$, where θ is the angle which C-H vector makes with bilayer normal and angle brackets imply the time average of S_{CD} values. These lipid raft domains play a key role in signal transduction and modulate a variety of biochemical pathways [32].



Figure 1.5: A single DMPC lipid (hydrogen atoms not shown). Polar head and hydrophobic tails highlighted.



Figure 1.6: Different lipid phases:(A) spherical micelles; (B) cylindrical micelles (tubules); (C) disks; (D) inverted micelles; (E) a fragment of a rhombohedral phase; (F) lamellae (G) inverted hexagonal phase; (H) inverted micellar cubic phase. This image was adapted from Rumiana Koynova's paper[76], which is under Creative Commons Attribution 3.0 Unported (CC BY 3.0) license.

Researchers have studied model membranes both experimentally and theoretically [77, 78, 79]. Based on the physical environment, hydration and lipid shapes, lipids can self-organize into a variety of morphologies like lamellar structures, micelles, inverted micelles, etc[80] (See Figure 1.6). Since the hydrocarbon tails of lipids are hydrophobic, it is energetically unfavorable for them to get exposed to water. So these lipids form lamellar structures or micelles, etc where tails are prevented from any kind of exposure to water. From DSC (Differential scanning calorimetry) studies, it was revealed that all model lipid membranes have two characteristic transition temperatures. See Figure 1.7 (a) and (b) to see the DSC profile of a typical model membrane and a cartoon depiction of different phases of the lamellar lipid bilayer. When the temperature is lower than so called pre-transition temperature T_p , the lipid membrane organizes itself into a "gel phase" (L_{β}) , which is very ordered in nature and the tails are in "all-trans" configuration. Beyond main-transition temperature T_m , the lipid membrane attains a disordered "liquid" (fluid, (L_α)) phase where the tails are in "gauche" configuration. In between pre-transition temperature T_p and main-transition temperature T_m , the lipid membrane attains a "Ripple" (P_{β}) structure which has both regions of ordered gel phase and disordered fluid phase. One main feature of gel phased lipid membrane is that the lateral diffusion of lipids is nearly absent. At the same time in the fluid phase, the lipids diffuse laterally and even exhibits flip-flops (trans-layer lipid diffusion), but at a higher time scale (hours)[81, 82].

1.2.3 Motivation for the study

Various studies in the past have demonstrated the capability of molecules like local anesthetics, alcohols, si-RNA, membrane–active proteins and charged nanoparticles in modulating the phase transition profile by inducing interdigitated/rippled domains in them[40, 4, 5, 41, 84, 42, 43]. Vasudevan et al reported[2] that the Nogo66



Figure 1.7: (a) A typical phase transition profile of a model membrane from DSC experiments looks like this plot. Below the pre-transition temperature T_p , the lipid bilayer membrane exist as gel phase which is very ordered in nature. Beyond the main-transition temperature T_m , the membrane exists in fluid state (disordered) which is the biologically relavant state for cell membranes etc. In between the pre-transition temperature T_p and the main-transition temperature T_m , the membrane exist in ripple form which has both ordered gel phase regions and disordered fluid state. (b) A cartoon depiction of different phases of lipid bilayer sheets. This image was adapted from Justyna B. Startek et al paper[83], which is under Creative Commons Attribution-ShareAlike 4.0 International (CC BY-SA 4.0) license.

(the extracellular domain of the neurite outgrowth inhibitor protein Nogo) is an intrinsically disordered in solution, but folds at the phosphocholine membrane interface. Another study on a membrane–active protein called Hsp12 (a heat shock protein) in yeast system showed that the protein folds at DMPG (1,2-Dimyristoyl-snglycero-3-(phospho-s-(1-glycerol))) membrane interface and induces ripple structure in the membrane even beyond transition temperature due to the interaction with protein[5]. Given that Nogo-66 is a membrane-active protein, we investigate the effect of Nogo-66 on the membrane structure, in particular the effect of interaction on the phase transition of membrane. All-atom classical molecular dynamics simulations have been used as the tool to study the interaction of Nogo66 with the DMPC membrane.

1.3 Thesis Organization

The rest of the thesis is organized as described below:

In Chapter 2, we describe fundamentals about classical all-atom molecular dynamics simulations employed in our work and description of various analyses performed on the simulation data.

In Chapter 3, the effects of phosphorylating SER11 residue of Rap protein in a Rap:Raf complex is described. We use classical all-atom molecular dynamics as the tool to examine this problem. We elucidate the mechanism of allostery propagated in the complex and binding free energy enhancement between Rap and Raf proteins in the complex as a result of phosphorylation. We conclude this chapter by highlighting:

- The importance of studies like this which can reveal dynamics from structures available in PDB database.
- The capability of these studies which can reveal new potential pockets in the proteins.
- A possible alternate strategy of using single-site post-translational modifica-

tion/mutation of proteins which can compete with "undruggable" malfunctioning antagonists and thereby preventing the wrong downstream signals.

In Chapter 4, the effects of phosphorylating interfacial residue SER39 residue in the Rap:Raf complex is studied using all-atom classical molecular dynamics simulations. The main conclusion from this chapter is interfacial residue perturbation via phosphorylation resulted in better interaction between the two proteins Rap and Raf, which is revealed by free energy calculations. The favourable interaction between proteins is driven primarily by enhanced electrostatic component due to the interfacial phosphorylation. The Network Analysis study revealed allostery propagation till L4 loop.

Chapter 5 describes the phenomenon of a membrane-active protein Nogo66 inducing interdigitation in the DMPC membrane patch even beyond its main transition temperature. We explain about the energetically unfavorable lipid defects induced in the membrane patch due to the interaction with Nogo66. We also monitored the interaction energies between Nogo66 and the membrane and identified that the onset of interdigitation is correlated with the instances where there is favorable electrostatic and van der Waals interaction. We conclude this chapter by explaining how membrane-active proteins can play an important role in membrane remodeling like domain formation.

Chapter 2

Methods

In this chapter, we will describe the simulation and analysis techniques which have been employed in our work. There are multiple kinds of simulation techniques that can be employed to study biomolecules, including Monte Carlo(MC) simulations, Molecular dynamics(MD), lattice simulations, etc. The simulation technique which we have used in our work is all-atom classical molecular dynamics. In the first section, we will give an introduction to all-atom classical MD simulations and then explain the concept of "forcefield" which defines the interactions between atoms(or particles) in the system. And the next section is about concepts like thermodynamic ensembles and thermostats. And finally, we explain basic analysis techniques and free-energy calculation method called MM-GBSA(Molecular Mechanics-Generalized Born surface area) which have been used in this work.

2.1 All Atom Classical Molecular Dynamics

Molecular dynamics(MD) simulations are techniques used to simulate a system of N interacting particles(many-particle systems). It is called classical MD simulations if the law which governs the motion of these particles is based on Newton's equation

of motions. And if the simulation is at the resolution of atoms, it is called all-atom classical molecular dynamics simulations. This technique can be used to simulate different systems including a variety of biological systems like a single molecule of protein in water, a protein complex, protein-membrane interactions, etc.

As mentioned before, solving Netwon's laws of equations is at the heart of molecular dynamics simulations. We use Newtonian mechanics to calculate the net force and acceleration experienced by each atom i(among N total atoms), considered as a particle of mass m_i and a fixed charge q_i . And the acceleration (a_i) experienced by each atom i, is determined by:

$$\vec{a_i} = \frac{\vec{F_i}}{m_i} = -\frac{1}{m_i} \frac{\partial U(r_1, r_2, \dots r_N)}{\partial r_i} = -\frac{\nabla U(r_1, r_2, \dots r_N)}{m_i}$$
(2.1)

where r_i and a_i is the coordinate and the acceleration of the i^{th} atom respectively, $U(r_1, r_2, ..., r_N)$ is the manybody potential energy function and defines the interaction between particles here. Positions, velocities and accelerations at time t are used to calculate new positions, and velocities at time $t + \delta T$. This process is continued for the desired number of steps.

Since it is impractical to simulate particles in the order of thermodynamic limit($\sim 10^{23}$), periodic boundary conditions are often used in the simulations. This allows a better estimation of bulk properties from simulations of finite nanoscale systems and they also address issues like finite-size effects. Another important aspect of any MD simulation is that it assumes ergodicity if the simulations are run for for a long time. Ergodicity is the ability to sample the whole of phase space volume for the given ensemble conditions. One of the consequences is the ensemble averages versus time averages of observables tend to be the same value. Since MD simulations are supposed to run huge systems($\sim 10^6$ particles), the programs should be written keeping in mind for making use of parallel computing facilities. Some of the popular packages for MD simulations are NAMD[85], GROMACS[86], Amber[87] and LAMMPS[88].

2.1.1 Forcefield

Interatomic forces are derived from analytically approximate interatomic potential function (U in equation 2.1). So the main task is to formulate this potential function which defines how particles interact with each other. The form of potential energy function used in molecular dynamics simulations is called forcefield. A general form of such function is shown in Figure 2.1. The interaction potential U consists of both bonded and nonbonded interaction terms. Bonded interactions model covalently bonded atoms and nonbonded interactions include VDW interactions and electrostatic bonding etc. Some of the popular forcefields used in biological contexts are CHARMM(Chemistry at Harvard Macromolecular Mechanics)[89, 90], AMBER(Assisted Model Building and Energy Refinement)[91] and OPLS(Optimized Potentials for Liquid Simulations)[92, 93], etc.



Figure 2.1: The form of potential energy function U used in all-atom classical molecular dynamics simulation codes(Forcefield). This figure was adapted from Durrant and McGammon paper[94], which is under Creative Commons Attribution 2.0 Generic (CC BY 2.0) license.

Bonded Interactions

To model covalently bonded interactions in a system, there are three kinds of interactions that are considered. One is bonded interactions in bond distance(r) between 2 adjacently bonded atoms realized in form of a harmonic well potential, second is a bonded interaction in form of angle(θ) between three consecutively bonded atoms which again is realized as another harmonic well and finally a dihedral interaction which occurs when you consider 4 consecutive atoms where the oscillations happen at the angle(ϕ) between the planes spanned by first three atoms and the last three atoms. The multiple minimas arise from the fact that these dihedral interactions are optimized to reproduce the experimentally observed energy differences between trans and gauche conformations etc.

The first term in Figure 2.1 represents the bond energy between two covalently bonded particles *i* and *j*. r_{ij} is the instantaneous distance and r_{eq} is the equilibrium bond length between these atoms. $K_{r_{ij}}$ is the force constant associated with that particular bond. Similarly, for the angles and dihedrals, there are other force constants $K_{\theta_{ij}}$ and $V_{n_{ij}}$. Besides these classic interactions, there are also terms like improper dihedral interactions between four atoms whose bond connectivity as in the vertices of a tetrahedron. Such interactions are used to maintain planarity in the molecular structure. The potential can be represented as a quadratic harmonic function.

Nonbonded Interactions

The nonbonded interactions refer to VDW and electrostatic interactions between atoms. See the Figure 2.2 for an example case. Lennard-Jones interactions are included in the forcefield in order to capture the "non-polar" interactions and electrostatic interactions capture the interactions between particles due to their charges.



Figure 2.2: An illustration of the bonded and nonbonded interactions in the case of a protein Rap visualized in VMD. Lysine $31(N\xi \text{ atom in blue})$ and Aspartic acid $38(O\delta 1 \text{ and } O\delta 2 \text{ atoms in red})$ residues are interacting via a saltbridge interaction.

In Figure 2.1, the Lennard-Jones interactions (the 6-12 LJ potential) are those which includes the $(1/R)^{12}$ and $(1/R)^6$ terms. The $(1/R)^{12}$ is the repulsion term and the $(1/R)^6$ term is the attraction term. The nature of LJ interaction is shown in Figure 2.3. The ϵ is the depth of the potential well, σ is the finite distance at which the inter-particle potential is zero, r is the distance between the particles, and r_m is the distance at which the potential reaches its minimum. At r_m , the potential function has the value $-\epsilon$. In practice, the LJ potential is a short-range interaction as the distance between particles increase. For efficient computation, the LJ interactions are truncated beyond a certain cut-off distance. To avoid singularities (discontinuities) in the function due to cut-off, the truncation is implemented through an associated "switch-function", which takes the potential to zero continuously between the switch and cut-off distances. The forcefields usually assign LJ interactions between particles of same type. When different types of particles interact, the parameters are assigned values based on Lorentz-Berthelot mixing rule[95]. According to this mixing rule,

$$\epsilon_{ij} = \sqrt{\epsilon_{ii} \times \epsilon_{jj}}$$

$$R_{minij} = (R_{minii} + R_{minjj})/2$$
(2.2)



Figure 2.3: The nature of the 6-12 Lennard-Jones(LJ) interaction. V is the Lennard-Jones(LJ) potential and ϵ is the depth of the potential well. r is the distance between the particles and σ is the finite distance at which the inter-particle potential is zero. This figure was taken from wikimedia commons, which is under Creative Commons Attribution-ShareAlike 3.0 Unported (CC BY-SA 3.0) license.

The second kind of non-bonded interaction is the electrostatic interactions (Coulombic potential) where the interaction between two charged atoms (with charges q_i and q_j) are described as $\frac{1}{4\pi\epsilon_0\epsilon_r}\frac{q_iq_j}{r}$ where r is the separation between the particles, ϵ_0 and ϵ_r is the vacuum permittivity and relative permittivity respectively. Electrostatic interactions are long-range interactions and so it is the most computationally difficult

to compute in MD simulations. Also due to the use of periodic boundary conditions in simulations, the sites interact with all the periodic images and this demands special methods to deal with electrostatic interactions. Ewald Summation is one of the popular techniques to deal with this issue. To understand this, consider the Poisson equation in which Coulombic potential $\rho(x, y, z)$ written in differential form:

$$\nabla^2 \phi(x, y, z) = -\rho(x, y, z)/\epsilon \tag{2.3}$$

where $\phi(x, y, z)$ is the potential, $\rho(x, y, z)$ is the charge density at positions x, y, z, and ϵ is the permittivity of the medium. Since we have point charges in the system, the system cannot be solved by discretizing this equation. The Ewald summation replaces the point charges(delta functions) with some smooth functions like Gaussian functions, but inverted ones. Now the sum of these point charges and respective inverted Gaussians comprise the short-range interactions which can be computed in real space(r-space). The complementary part which constitutes the long-range interactions is computed in the Fourier domain(reciprocal space or k-space) for computational efficiency. (See Figure 2.4.) In MD simulations, one of the popular ways to compute the electrostatic interactions is via the Particle Mesh Ewald(PME) technique[96]. Other methods include multilevel summation method(MSM)[97] and fast multipole method(FMM)[98] etc. In PME, the calculation in reciprocal space is done using Fast Fourier transforms(FFT) using charges assigned to discrete grid points. And the computational complexity of this algorithm is $\mathcal{O}(N \log N)$, whereas the direct evaluation will take $\mathcal{O}(N^2)$.

2.1.2 Basic steps of an MD simulation for biomolecules

The initial step in the MD simulation is to obtain coordinates of the system. For biological systems, the coordinate information are usually obtained from experimental



Figure 2.4: Illustration of Ewald Summation by decomposing it into Real and Reciprocal spaces.

sources such as crystal structure from protein data bank(PDB)[99]. If the system includes lipid membranes etc, CHARMM-GUI[100, 101] software can be used to construct the system. An important consideration while creating the initial structure is to avoid creating systems with energetically very unfavorable structures like atoms with merging coordinates or steric hindrances etc. Protonation states should be carefully assigned to proteins depending on the pH conditions which are usually specified in experimental data(like PDB). The second step is minimization and its purpose is to take the initial structure to a state of local energy minima in the potential energy landscape. Steepest descent method and Conjugate gradient method etc are common minimization algorithms [102]. Minimization only assigns or "corrects" the positions of atoms. The next step is assigning velocities to the particles which is usually sampled from Maxwell-Boltzmann distribution. The reason for using Maxwell-Boltzmann distribution in assigning velocities stems from the fact that the average kinetic energy $\langle K \rangle$ is related to the temperature of the system. The correct temperature T can only be reproduced if all the N atoms follow the relation $\langle K \rangle = 3Nk_BT/2$, where k_B is the boltzmann constant. Using random velocities instead of Maxwell-Boltzmann distribution might take more time to settle to final distribution. The simulations are run under the thermodynamic ensembles like NVE, NPT or NVT ensemble, where 'N' stands for number of particles, 'V' for volume of the system, 'T' for temperature, 'E' for the total energy of the system and 'P' for the Pressure of the system. In order to relax the system away from any artificially induced metastable starting states, we perform equilibration simulations. Once equilibration is done, the production simulation runs are executed.

Integrators

The newtonian equations of motion(equation 2.1) are integrated to generate time evolution of a system and for this various integration schemes are proposed in literature. Certain conditions like reversibility of dynamics and the symplectic property have to be maintained by the integrators. One of the popular integrator is the Verlet algorithm. To derive it, the position is expanded according to Taylor series:

$$\mathbf{r}(t+\delta t) = \mathbf{r}(t) + \delta t \mathbf{v}(t) + \frac{1}{2} \delta t^2 \mathbf{a}(t) + \dots$$

$$\mathbf{r}(t-\delta t) = \mathbf{r}(t) - \delta t \mathbf{v}(t) + \frac{1}{2} \delta t^2 \mathbf{a}(t) + \dots$$
 (2.4)

where \mathbf{r} , \mathbf{v} and \mathbf{a} are the position, velocity and acceleration vectors of particles respectively.

The above 2 equations gives :

$$\mathbf{r}(t+\delta t) = 2\mathbf{r}(t) - \mathbf{r}(t-\delta t) + \delta t^2 \mathbf{a}(t)$$
(2.5)

and the velocity:

$$\mathbf{v}(t) = \frac{\mathbf{r}(t+\delta t) - \mathbf{r}(t-\delta t)}{2\delta t}$$
(2.6)

To evolve the system to future positions, this Verlet algorithm requires the current position vectors, position vectors of the previous time step, and the current acceleration vectors. This is a simple algorithm, but has the disadvantage of having no information about velocity in the equation (and so has to be calculated separately) and the truncation error associated with the positions and velocity in the algorithm is of the order $\mathcal{O}(\delta t^4)$ and $\mathcal{O}(\delta t^2)$ respectively, which adds to the lack of precision. And the biggest problem is that the algorithm has to keep previous positions in the memory. And for these reasons, the modified algorithm called Velocity-Verlet algorithm is the integrator used by most MD engines. Here the algorithm contains the following equations:

$$\mathbf{r}(t+\delta t) = \mathbf{r}(t) + \delta t \mathbf{v}(t) + \frac{1}{2} \delta t^2 \mathbf{a}(t)$$

$$\mathbf{v}(t+\delta t) = \mathbf{v}(t) + \frac{1}{2} \delta t (\mathbf{a}(t) + \mathbf{a}(t+\delta t))$$
(2.7)

The above formulae are executed in three steps. As can be seen in second equation above, to calculate the new velocities requires the accelerations at both t and $t + \delta t$. From the second equation, the positions at $t + \delta t$ are calculated. Next, $\mathbf{v}(t + \delta t/2) =$ $\mathbf{v}(t) + \frac{1}{2}\delta t\mathbf{a}(t)$ is calculated. Forces are computed from current positions $\mathbf{r}(t + \delta t)$, giving $\mathbf{a}(t + \delta t)$.

The final step is to calculate the velocities at $t + \delta t$ with the formula:

$$\mathbf{v}(t+\delta t) = \mathbf{v}(t+\delta t/2) + \frac{1}{2}\delta t\mathbf{a}(t+\delta t)$$
(2.8)

The Velocity-Verlet algorithm is better in precision comparing to Verlet algorithm, since its error is of the order $\mathcal{O}(\delta t^2)$.

Thermostats and Barostats

Thermostats and barostats are used to maintain the conditions of constant temperature and pressure in the system to fixed values. The temperature of a system is measured from kinetic energies as in equipartition theorem. Thermostat algorithms work by modifying the Newtonian equations of motion which are inherently in NVE ensemble. Thermostats can be implemented either in a deterministic or stochastic form and be either global or local, depending on whether they are coupled to the entire system or separately to parts of the system. Some popular thermostats include:

- Velocity Rescaling thermostat
- Berendsen thermostat
- Andersen thermostat
- Langevin thermostat
- Nosé-Hoover thermostat

Velocity rescaling method is very crude(non-physical), but simple to implement. It relies on rescaling the velocity of the particles such that the instantaneous temperature exactly matches the target temperature. In Berendsen Thermostat the system is coupled with a heat bath with a target temperature(T_0). The temperature is adjusted according to the equation below, where the deviation exponentially decays with a time constant τ :

$$dT/dt = (T_0 - T/\tau)$$
 (2.9)

Berendsen thermostat is known for not sampling from the NVT ensemble and so simulation artifacts are common with it. The principle of Andersen thermostat is to select particles randomly from the system and making them collide with a heat bath by assigning the particles new velocities sampled from the Maxwell-Boltzmann distribution. Although this thermostat samples from NVT ensemble, the system dynamics can be affected by the *collisions* with the heat bath. Langevin thermostat is a stochastic thermostat which supplements NVE ensemble with Brownian dynamics, which has both components of viscosity and random collision effects in the form of noise. Nosé-Hoover thermostat removes the concept of an external heat bath from other cases and adds an extra degree of freedom in the form of a fictitious mass that interacts with the particles in the system. The dynamics are well preserved, but ergodicity is somewhat sacrificed for smaller systems. This can be improved with the use of Nosé-Hoover chains[103]. If the simulations are done in the NPT ensemble, barostats are necessary to assign constant pressure. The popular barostats are:

- Volume rescaling
- Berendsen barostat
- Andersen barostat
- Parrinello-Rahman barostat
- Nosé-Hoover Langevin piston barostat

Volume rescaling works by adjusting the system volume to attain target pressure. It doesn't sample the ensemble properly and cannot be used for production runs. Berendsen barostat works using the similar principle as in the Berendsen thermostat and uses an improved version of volume rescaling and couples the system with a weakly interacting pressure bath. Andersen barostat is an isotropic barostat which uses the principle similar to Nosé-Hoover thermostat by adding an additional degree of freedom to the equations of motion. This barostat is good in sampling the ensemble. Parrinello-Rahman barostat is an extension of Andersen barostat which can also be used for realizing pressure anisotropy. Nosé-Hoover Langevin piston barostat uses Langevin dynamics to control fluctuations in the barostat.

2.2 Analysis Techniques

2.2.1 Covariance Analysis

Covariance Analysis can be done on the MD trajectory of protein C_{α} atoms in order to analyze the coupling between residues. The covariance matrix constructed from the displacements with respect to the average structure can give information regarding correlated/anti-correlated movements between various regions of the protein. The covariance matrix is defined as following:

$$C_{ij} = \frac{\langle \Delta r_i \cdot \Delta r_j \rangle}{\sqrt{\langle \Delta r_i \cdot \Delta r_i \rangle \langle \Delta r_j \cdot \Delta r_j \rangle}}$$
(2.10)

where $\langle \cdot \rangle$ stands for the averaged values(across the snapshots in molecular dynamics trajectory), Δr_i and Δr_j are the ith and jth atom's displacements with respect to the corresponding averaged structure atoms.

2.2.2 Principal Component Analysis

Principal Component Analysis (PCA) is a very popular dimensionality reduction technique used in multidimensional datasets. It performs an orthogonal transformation on the possibly correlated data to produce linearly uncorrelated variables called principal components (PCs). The first PC will be in the direction of maximum variance from data and so on. The first step is to fit the molecule with the structure of the initial snapshot of MD trajectory, and remove the translations and rotations etc. And the second step is to generate the covariance matrix from the data. In our case, the data is the molecule coordinates from the molecular dynamics trajectory. The data is in the form of $M \times 3N$ format, where M is the number of snapshots in molecular dynamics trajectory and 3N is the x,y,z coordinates of the number of atoms. The covariance matrix is defined as in equation 2.10. The reference atoms in this equation can be from the crystal structure instead of averaged structure coordinates. The next step is to do Singular Value Decomposition on this Covariance matrix and project that data on to the coordinates. What this essentially gives is a scatter plot between different dominant orthogonal components from decomposing the RMSD(Root Mean Square Distance) data.

PCA in MD context can be used:

- To study about the global motions in molecules.
- To filter out the high frequency motions in the MD data.
- To see the conformational space sampled by molecules.
- To find significant reaction coordinates of the system and use it for free-energy calculations etc.

Although PCA is very popular, two caveats to keep in mind:

- If the weight of eigen values captured by the first few vectors are less, it means that PCA will not be a meaningful tool in such cases!
- It can also give meaningless PC directions say in the case of a double well potential. In such cases other techniques like t-ICA(time lagged Independent Component Analysis) would be useful.

2.2.3 Community Network Analysis

Community network analysis is a technique that can be used to investigate the propagation of an allosteric signal in a protein of protein-protein complex etc. This technique converts the protein or protein-protein complex into an equivalent network(graph) representation. If two atoms in the trajectory are interacting within a threshold distance(say 4.5Å) for the 75% of the time duration of the considered MD trajectory, then those two atoms are considered as nodes connected via an edge in the network. *NetworkView*[104, 24] plugin in VMD is used to perform community network analysis on the MD trajectory data. The community detection analysis was done using software "gncommunities" and the algorithm behind this is Girvan-Newman algorithm[24, 23]. The edges are weighted using the correlation matrix(C_{ij})

data between the C_{α} atoms using the relation:

$$w_{ij} = -\log(abs(C_{ij})) \tag{2.11}$$

Once the communities are detected in the network representation of the protein or a protein-protein complex, it becomes easy to infer the signal propagation(possible allostery) as a result of specific events in the network.

2.2.4 MM-GBSA Free Energy Calculations

MM-GBSA(Molecular Mechanics-Generalized Born surface area) method[105, 106, 107, 108] is the most efficient method to calculate the binding free energy of large systems like protein-ligand or protein-protein complex, etc. This method is computationally cheaper than other free energy methods such as steered molecular dynamics[109], free energy perturbation, and metadynamics. It calculates binding free energy by making use of the thermodynamic cycle depicted as in Figure 2.5.

This method is parameterized within the additivity approximation where the net free energy change is treated as the summation of individual energy components[110]. The free energy was calculated using the equation below:

$$G_{TOT} = H_{MM} + G_{solv} - T\Delta S_{conf} \tag{2.12}$$

where H_{MM} is the sum of the bonded, electrostatic and Lennard-Jones energy terms, G_{solv} is the sum of polar and non-polar solvation energies, T is the temperature and S_{conf} is the configurational entropy. The polar component of G_{solv} is estimated using the Generalized Born model and the non-polar component is calculated from the linear relation between SASA(solvent accessible surface area) and surface tension of the solvent(water). G_{solv} is included in the electrostatic measurements done via



Figure 2.5: Illustration of Thermodynamic cycle utilized in MM-GBSA binding free enrgy calculations. The ligand and receptor can be a small molecule and protein, two different proteins etc.

NAMD script for GBSA calculations. In the single trajectory method, employed here, the bonded energy contribution to the change in H_{MM} will be zero. Regarding the inclusion of entropic term in the free energy of binding, we would like to note that the usual method of computing the entropy via normal modes[111] is computationally very expensive, especially for protein-protein complexes and other methods may have convergence issues and has been omitted in many earlier works, which we follow in this work as well. [106, 107, 112, 105, 113, 114, 115, 116, 117, 118] The binding free energy for complex formation is then calculated as:

$$\Delta G_{bind} = G_{TOT}^{Protein-Ligand} - \left(G_{TOT}^{Protein} + G_{TOT}^{Ligand}\right)$$
(2.13)

where $G_{TOT}^{Protein-Ligand}$, $G_{TOT}^{Protein}$ and G_{TOT}^{Ligand} are the free energies corresponding to Protein-Ligand complex, only Protein, and only Ligand trajectories. In addition to the binding free energy measurements, we have to calculate the change in entropy of the individual as well as the complex using the quasi-harmonic approach.

Chapter 3

Phosphorylation promotes binding affinity of Rap-Raf complex by allosteric modulation of switch loop dynamics

3.1 Introduction

Rap belongs to the family of small Ras-like GTPases, which have many roles in cellular activities like cell proliferation, apoptosis and differentiation etc[119, 6, 120, 7, 121]. These GTPases act like molecular switches, active when GTP is bound and inactive when GDP is bound. These conformations are interconvertible by the action of the guanine nucleotide exchange factors (GEFs) and the GTPase activating proteins (GAPs). GEFs exchange GDP for GTP and GAPs catalyse the hydrolysis of GTP converting the active form into the inactive protein[8]. The Ras GTPases participate in many signalling pathways, including the MAPK/ERK , PI3K[122, 44, 123]. Many factors including their cellular location, bound ligand molecule and

phosphorylation can affect how these molecules interact with downstream signalling proteins [124, 125], which is crucial in transmitting signal from Ras to the mitogenactivated protein kinase. Because of their role in key signalling events which are often deregulated in cancer and due to their prominent role as oncogenes, Ras family members (specifically H-, K- and N-Ras proteins) have garnered considerable attention over the years[126, 123, 19, 52]. Considering all cancers where at least 20 tumours were counted and weighted equally, pan ras mutations were found at an incident rate of 16% [52] and activating Ras mutations are associated with approximately 30% of all human cancers[44]. Many of these mutations render the tumour aggressive and are responsible for the death of patients. Yet there are no targeted therapies for these class of proteins as they are considered notoriously undruggable lacking specific binding pockets[19].

The Rap proteins such as the Rap1A and Rap1B rose to prominence because of their high degree of identity to the Ras proteins[127, 128, 7, 129, 130, 119, 131]. Rap1A was identified as a suppressor of Ras activity in screening assays, a function attributed to its ability to competitively bind (in the presence of GTP) to downstream Raf without activating it and hence disrupting the signal transmission along the MAPK pathway[132]. Binding of Rap1 to RafB on the other hand results in activation as seen with the Ras family of proteins[133]. Many mutations, domain swapping experiments have indicated regions other than the RBD domain of Raf are responsible for these differences[134, 135]. All known Ras effectors share a common Ras-binding Domain (RBD). Besides competing with Ras, Rap proteins are involved in many other crucial cellular functions such as cell adhesion, cell-cell junction formation and regulation of the actin cytoskeleton[136, 137, 138, 139].

Although not as well studied as the role of GEF proteins and GAP catalyzed changes in the nucleotide bound conformations, phosphorylation is known to regulate the functions of Ras and Rap proteins[140, 141, 142, 143, 144]. Phosphorylation is the most common reversible post translational modification(PTM) of proteins with a role in regulation of essentially all cellular functions [145, 146]. The mechanism of how phosphorylation acts as a molecular switch that allows cells to respond instantaneously to various stimuli without the need for new protein synthesis, how phosphorylation at a remote site often influences the activity at a completely different site continue to be active areas of research. Most often phosphorylation is observed in disordered, well accessible or highly flexible regions and loops in proteins [20, 21]. However, multiple investigators, including our collaborator, while analyzing large scale data sets have documented that many of these phosphosites are not readily exposed to the solvent [22, 65, 66, 21]. The presence of significant number of possible phosphosites in the disallowed region of phosphoconformation suggests that factors including dynamic conformational changes of proteins, binding to other proteins or regulatory factors can potentially expose these buried phosphosites to the solvent and to a kinase. Therefore understanding the role of protein dynamics in exposing such sites to solvents, interactions involved in transmitting the effects of phosphorylation to other functional sites is essential. Inherent loop dynamics of proteins play a critical role in functioning of the protein [147, 148, 149, 150] and it is very important to understand the effects of local mutations or PTMs on the global structure of proteins. However it is not trivial to obtain phosphoproteins in amounts large enough for experimental investigations even when the kinase involved is known. MD simulations are invaluable tools successfully employed on multiple occasions to understand the effect of phosphorylation on the structure, dynamics, allosteric effect, conformational stabilization and map the electrostatic interactions in the proteins and thereby deduce the effect on functions [151, 152, 153, 154, 155].

In this work, we study the effects of phosphorylation of a single residue, SER11, identified as a possible phosphosite in Rap1A [67] (see Figure 3.1), on the conformational dynamics of the Rap1A and its interactions with the effector protein kinase c-Raf1. It is to be noted that this particular residue, serine, occurs only in Rap GTPases
and is replaced by alanine in Ras GTPases[156]. Phosphorylation at the same site has been observed in tumor samples of lung cancer patients as well [157] and SER11 phosphosite in RAP1 carries motifs for many kinases, some with high and others with moderate scores: for example putative sites CK1, Aurora and ATM kinases, are predicted by KinasePhos2 a webserver for phosphosite predictions[158]. In addition to SER11 residue, SER39, SER179 and SER180 are other possible phosphosites in Rap1 that are either predicted or experimentally determined. Proximity of SER11 to the nucleotide ligand, which alters the activity of GTPases such as Rap1, renders the investigation of effects of phosphorylation particularly interesting. We also explore the effects of such phosphorylation on the dynamics of functional loops such as Switch I and Switch II loops to characterize the allosteric pathways within Rap-Raf complex and subsequently gain some insight into possible mechanisms through which Rap may affect the downstream MAPK signalling pathway.

3.2 MD Simulations

The effects of phosphorylation on the dynamics of complex Rap-Raf are studied using the available complex structure of Ras-related protein Rap1A (referred to as Rap) liganded with GMPPNP which is a GTP-analogue, Mg²⁺ ion and Raf-RBD(Ras binding Domain) from c-Raf1 (referred to as Raf in this work) downloaded from protein data bank with PDB id 1C1Y with a resolution of 2.2Å[159]. 1C1Y comprises of GTP analogue molecule(GppNHp) liganded protein Rap1A complexed with Ras binding domain(RBD) of cRaf-1. For simulations performed in this work, we replaced the GTP analogue with GTP molecule. All-atom classical MD simulations were done for the systems listed in Table 3.1, using CHARMM36[160, 161] forcefield with the aid of NAMD[85] software. The visualization was done using software VMD[162] and analysis of data using Tcl scripting language which is embedded with VMD, Matlab and Grcarma[163] software. The protein complex(as in



Figure 3.1: Rap-Raf protein complex (crystal structure, PDB ID 1C1Y) visualization showing the location of important functional loops like P-loop, Switch I, Switch II and RBD loop regions. The phosphosite SER11 is shown in green.

PDB file 1C1Y) was crystallized at pH 7.6. The protonation states for all ionizable residues were found using PDB2PQR server [164] before solvating them. We found out that patch for neutral Lysine has to used at Rap residue number 16. And for the phosphorylated SER11 case simulations, dianionic phosphoserine SP2 patch was applied. PSFGEN tool from VMD software was employed for these patches. Each system was solvated in a TIP3P[165] water box and overall charge neutrality was realized by adding appropriate counter ions usingVMD. The systems were then subjected to energy minimization runs using the conjugate gradient method for 5000 steps, followed by molecular dynamics simulation runs in isothermal-isobaric(NPT) ensemble. The GTP analogue in the crystal structure,GppNHp was replaced by GTP and GDP molecules for the present simulations. The simulations of Rap-Raf complex with GTP and GDP ligands with and without SER 11 phosphorylated were done for 400 ns with a time step of 2 fs. The Nosé-Hoover-Langevin piston with a decay period of 100fs and a damping time of 50fs was used to maintain a constant pressure of 1atm[166, 167]. Berendsen thermostat[168] was used to control temperature at 298 K. A cut-off distance of 12Å was used to compute all short-range van der Waals (VDW) interactions and the long-range electrostatics interactions was treated with the Particle Mesh Ewald(PME) method[96, 169].

No.	Description	Length of Simulation
1	Rap-Raf complex with GTP and Mg^{2+}	400ns
2	Rap-Raf complex with GTP and Mg^{2+} with phospho-Serine11	400ns
3	Rap alone with GTP	200ns
4	Rap with Phospho-serine11 with GTP	191ns

Table 3.1: Details of systems simulated and analysed.

We used CASTp server[170] for pocket identification in the Rap domain. The largest cavities of the last 5ns time averaged structure of Rap domain for different cases were computed. These calculations can give information about the flexibility of the protein cavities and provides the necessary insights into comparative analysis of different forms of the protein as well as how each simulation is different from the original starting structures.

3.3 Results

3.3.1 Effects of phosphorylation on structure and dynamics of Rap

The stability of the structures monitored via the root mean squared deviation of the GTP-bound Rap protein with and without the SER11 phosphorylated is shown in Figure 3.2. The mobility of various parts of the protein and the effect of phosphorylation on the mobility of the Rap protein is measured through the root mean squared



Figure 3.2: Time evolution of root mean squared deviation(RMSD) of Rap with GTP ligand with and without SER11 phosphorylated. Only C_{α} atoms of Rap chain are considered for computing RMSD.

fluctuations (RMSF) of each residue, averaged over last 50 ns, 350-400 ns, (Figure 3.3). In the GTP bound form, the effects of SER11 phosphorylation are centered around the Switch I and Switch II loop regions. Proteins belonging to Ras superfamily are known to inhabit multiple conformational states and the two important loops that determine such conformational flexibility are Switch I (residues 25-40) and Switch II (residues 57-74) loops. The mobility of the loop region containing residues 80-85, which is spatially proximal to the phosphosite, is increased in the case of phosphorylated SER11. This is due to the acquired favourable interactions between the phosphorylated SER11 and the Switch II loop (discussed later) and consequently disrupt the interaction between the phosphosite and the loop containing residues 80-85. To understand the effects of complex formation on the structure of Rap, we also simulated the Rap protein alone (not in complex with Raf), with GTP ligand. The RMSF plots for all the four cases of Rap protein with or without complex formation with RBD of Raf protein and with and without phosphorylation of SER11 is shown

in Figure 3.4. The results suggest that the loop with residues 138-141 is mobile in all cases except the case of Rap-Raf complex with SER11 phosphorylated. This is due to favourable electrostatic interactions between ASN140 and ASP108. Results (See Figure 3.4) suggest that the mobility of both switch loops is very high when Rap is not in complex with RBD of Raf but undergoes reduction of mobility when SER11 is phosphorylated. The Solvent accessible surface area (SASA) of SER11 residue calculated for the molecular dynamics trajectory of Raf-Rap complex with GTP (measured using VMD keeping the probe radius to 1.4Å) indicates that the dynamical nature of the P-loop allows the buried phosphosite SER11 to be exposed to water for a significant time on the present simulation timescale, suggesting a strong possibility of a kinase phosphorylating the SER11 residue and lends validity to our simulation studies involving phosphorylated Rap protein (See Figure 3.5). In the following sections we describe how phosphorylation at SER11 influences the dynamics that affects a) Switich I loop interaction with nucleotide and RBD domain b) Switch II loop interaction with the nucleotide.



Figure 3.3: The RMSF of residues of Rap C_{α} atoms averaged over 350 to 400 ns of simulation for both unphosphorylated (green) and phosphorylated (blue). The relevant Switch I and Switch II loops are indicated in the figure.



Figure 3.4: RMSF of Rap chain carbon-alphas corresponding to the last 50ns of MD trajectories.



Figure 3.5: Time evolution of SASA value of unphosphorylated SER11 residue in Rap protein (with GTP ligand). For comparison, the SASA value of SER11 in the crystal structure is shown as blue line.

Switch II loop interacts primarily with the GDP/GTP exchange factors (known as GEFs), which accelerate the release of the previously bound GDP to the proteins and to be replaced by GTP. The conformation of Switch II loop undergoes profound

changes when GDP is exchanged with GTP ligand. [171, 8] The GEF proteins are known to make extensive contacts with residues in the Switch II loop inducing local conformational changes near the nucleotide binding site, which results in the release of the bound nucleotide. Thus the mobility of the Switch II loop plays an important role in binding the GEF proteins leading to the GTP-bound conformations of G-proteins and consequently affect their ability to bind to downstream effectors. From the Figure 3.3, it can be seen that the Switch II loop is most mobile in the GDP bound form of Rap, whose mobility reduces in the GTP bound form. This can be understood as the conformational stability that the loop acquires upon its interaction with the GEF proteins. Phosphorylation of the SER11 residue further reduces the mobility of this important functional loop in the GTP bound form of the protein. To understand this difference in the dynamics of the Switch II loop, its interactions with residues in the nearby P-loop (to which SER11 residue belongs) were investigated. In the Rap protein, as with the Ras proteins, the nucleotide pocket is flanked by primarily three loops: Switch I, Switch II and P-loop. As can be seen in Figure 3.6, the conformation of Switch II loop is drastically altered when SER11 is phosphorylated in Rap-GTP protein. There are two strong electrostatic interactions which underlie such a significant conformational change. The Switch II loop contains several polar and charge residues including ARG68, which is positively charged. On phosphorylation of the SER11, the pocket region close to this residue acquires more negative charge compared to the unphosphorylated Rap. The simulations strongly suggest there is a phosphorylation induced change in conformation of Arg 68, which forms a stable salt bridge with the phosphate group of the SER11. This strong interaction results in pulling of the Switch II region into the nucleotide binding pocket which results in formation of another stable electrostatic interaction between the main chain carbonyl oxygen and the amide nitrogen atoms belonging to GLY 60 (Switch II) and GLY 12(P-loop) respectively. The distance between the two residues in the Rap-GTP protein with and without phosphorylation throughout

the simulation timescale are shown in Figure 3.7 (a) and the difference is almost of the order of 7 Å. The GLY60 residue also forms stable favourable interactions with the oxygen atoms of the GTP ligand when SER11 is phosphorylated (Figure 3.7) (b). The resulting favourable electrostatic interaction locks the Switch II loop into a conformation that reduces drastically its mobility.



Figure 3.6: Overlapped snapshots of GTP bound Rap protein showing the conformation of functional Switch I and Switch II loops with (red) and without (yellow) phosphorylation towards the end of the simulation. The positions of residues GLY 12 and GLY60 with (red) and without(yellow) phosphorylation is also marked.

Switch I loop of Rap interacts directly with the effector protein Raf kinase, and the strength of the interaction strongly depends on the bound nucleotide. Compared to the inactive state of Ras, when GDP is the ligand, RBD binds to active Ras-GTP is the ligand almost 1000 times more strongly[172]. This mode of interaction is conserved in Ras superfamily of proteins[124, 173, 174]. Hence, the dynamic mobility of the Switch I loop is of crucial importance in the interaction between proteins in Ras superfamily and their effectors. The RMSF plot in Figure 3.3 shows that the Switch I loop has slightly increased mobility when SER11 is phosphorylated, compared to the unphosphorylated Rap-GTP protein (in the GTP bound forms). The increased mobility is largely due to the movement of Glu30 of Switch I loop region (Figure 3.8 (a)). In the unphosphorylated Rap-GTP case, the GTP ligand



Figure 3.7: The time evolution of (a) distance between GLY60(O) and GLY12(N) atoms, (b) distance between GLY60(center of mass) and GTP(O3G) atoms for both unphosphorylated (green) and phosphorylated (blue) cases.

forms several favourable interactions with residues lining the nucleotide pocket. This includes a strong and persistent hydrogen bond between the the oxygen molecules attached to the γ -phosphate atom and the hydroxyl group of TYR32, which has been observed in many crystal structures of Ras super family including Rap. [159, 175] The other favourable electrostatic interaction of GTP ligand with Switch I loop is between GLU 30 and hydroxyl groups attached to ribose moieties of GTP ligand (see Figure 3.8). These interactions reduce the mobility of the Switch I loop and participate in the stability of the complex formation with RBD loop of Raf kinase. With the inward movement of the Switch II loop into the nucleotide pocket region, as mentioned above, the position of GTP ligand changes and the interaction between GLU30 and GTP ligand is broken (see Figure 3.8 (b)), resulting in increased mobility of the residue GLU30 (see Figure 3.3). It is to be noted that the position of TYR32 in the crystal structure used in the present simulations (1C1Y) is in the conformation in which the residue is located within the active site [176, 177]. This conformation of TYR32 is expected to play a crucial role for catalysis of Rap proteins [178], independent of GAP proteins, and that this conformation of TYR32

is preserved even when nucleotide pocket is significantly perturbed when SER11 is phosphorylated. This alteration of dynamics of Switch I loop by phosphorylation of SER11, which is located spatially and sequentially away from Switch I, clearly shows that allosteric mechanism is involved in communication.



Figure 3.8: Snapshots of nucleotide pocket of Rap-GTP with and without phosphorylation of SER11. The functional loops Switch I and Switch II loops are shown in red and yellow colour respectively. The relevant residues in the two loops are also shown and the hydrogen bonds between GTP and residues in Switch I loop, when relevant, are shown in white dashed lines.



Figure 3.9: The time evolution of distance between the catalytic residue THR61 and GTP ligand without (green) and with (blue) phosphorylation.

The most significant effect of phosphorylation of SER11 is in the position of THR61 residue on the Switch II loop, with respect the bound ligand GTP as shown in Figure 3.9. In Ras proteins, the residue 61 (which is GLN) plays a very crucial role,

along with the GAP proteins, in the GTP hydrolysis. The residue 61 is one of the most mutated site found in human tumours which impairs or abolishes the hydrolysis of GTP can lend the Ras protein to be in perpetually ON state^[52]. In Rap proteins, this important residue is replaced by a threenine and experimental studies have shown that THR61, unlike GLN61 in Ras, plays a predominant role in binding of GAP proteins and does not participate in GTP hydrolysis [176, 179]. The solved structure of Rap1 in complex with Rap1GAP has shown that the conformation of THR61 is away from the active site [176]. In the Rap-RBD structure used in the present simulation also, the Switch II is in a disordered state and THR61 is pointed away from the nucleotide. The simulations with the unphosphorylated Rap-RBD complex shows that the Switch II loop remains mobile and the THR61 moves away from the nucleotide during the course of the simulation. However, the phosphorylation of SER11 residue brings the THR61 into the nucleotide active site by forming a stable bond with GTP (as seen in Figure 3.9). This conformation and location of THR61 inside the active site can have profound effect on the ability of the GAP proteins to hydrolyse the GTP ligand and can potentially affect Rap's interaction with Raf.

3.3.2 Conformational sampling of Rap: Effects of phosphorylation

Having found that there is a reciprocal relationship in the dynamics of the switch loops that has a strong influence on nucleotide binding and Raf interaction the role of phophorylation on the conformation of different forms of the protein and effect on Rap activity is further explored in this section. Covariance analysis, using the cross correlation matrix as defined in the Methods chapter, is a very useful tool in getting insights into the relative correlated motions of different parts of the protein. The cross correlation matrix is computed by measuring the positional deviations of individual residues from an averaged structure and it is further averaged over equilibrium trajectory time scale. For all the simulations considered in this study, the cross correlation matrix is constructed over last 100 ns of simulation time in a run of 400 ns. We would like to emphasise that in our simulations, all rotations and translations were removed before performing the cross correlation analysis, as is the norm. The presence of hinges and possible large scale movement about the hinges can potentially complicate the positional cross correlation measurements, but no such global changes have been observed in all our simulations. It is very evident from the results (shown in Figure 3.10) that the phosphorylation significantly alters the correlation between various functional loops. In the unphosphorylated Rap, the Switch II loop is anticorrelated with both Switch I and P-loop (shaded in blue 1 and 2 respectively in Figure 3.10), which disappears in the case of phosphorylated Rap protein. In addition, a strongly positively correlated motion (region 4 in Figure 3.10) appears between Switch II and P-loop region. These results are consistent with the observations made in the previous section: the attractive interactions between the GLY60 in Switch II loop and GLY12 in P-loop triggers the observed positively correlated motion between the two loops. The conformational change in Switch II loop also removes strong anti-correlation between Switch II and helix 4 (residues 75 to 100) in Rap (region 3 in Figure 3.10), further suggesting an overall increase in the correlated motion between different parts of the Rap protein when SER11 is phosphorylated.

To understand the change in configurational space explored by the Rap protein when SER11 is phosphorylated, PCA analysis is employed. As has been described earlier, the two major regions which experience considerable changes after phosphorylation are localized regions in Switch I and Switch II loops. It has been long proposed that the proteins involved in the complex formation undergo changes in conformational entropy to compensate for the loss of translational entropy, due to complex formation, and it would be interesting to see if the phosphorylation can affect such



Figure 3.10: Cross-correlation plots of the Rap complex for 300 to 400ns trajectory data of GTP- and GTP-PSER11 cases.

conformational sampling. The mobility of the Switch II loop is considerably reduced when SER11 is phosphorylated and the loop's configuration also changes such that the residues on the loop are pulled significantly towards the nucleotide binding site. The Switch I experiences an increase in its mobility on phosphorylation, but the change is much less compared to the reduction of mobility of Switch II loop. From these results, it can be expected that the Rap molecule acquires an overall tighter configuration on phosphorylation which can be verified through monitoring the subspace defined by the two largest principal components (PC1 and PC2) of the projected MD trajectory. Towards this we have analyzed the trajectory data generated over the last 25 ns of MD simulations (375-400 ns) by fitting the coordinates of all the frames of C_{α} atoms. The results clearly show that the Rap protein occupies a different conformational space compared to the unphosphorylated form and that the over all conformational density is smaller, strongly indicating a tighter conformation (see Figure 3.11). The first two principal component vectors PC1 and PC2 captured nearly 83% of the information content from the last 25ns data of MD trajectories of Rap domain C_{α} atoms for the GTP liganded forms with and without SER11 phosphorylation.



Figure 3.11: Conformer plots of Rap domain (375 to 400ns simulation data). The plot shows conformational space sampled by Rap protein in terms of PC1(80.13%) and PC2(2.36%).



Figure 3.12: Largest cavity location of Rap domain in original crystal structure and last 5ns averaged structures of GTP (red; volume: 605.96 Å³), Crystal structure(blue; volume: 521.01 Å³) and GTP-PSER11(black; 278.01 Å³) cases.

The conformational changes in Rap, observed both in terms of interactions between various loop regions via covariance analysis and overall conformational flexibility of the protein, prompted us to look for phosphorylation induced changes in the distribution of pockets within the protein. The results are shown in Figure 3.12. The largest pocket in the Rap protein, by volume, is identified (averaged over last 5ns of simulation) and the results are shown for Rap-GTP with and without phosphorylation and the original crystal structure are also shown for comparison. The pocket location remains more or less the same for all the forms of Rap protein, except when Rap is phosphorylated at SER11 and with ligand GTP (shown in dark grey in fig Figure 3.12). As can be seen from Table 3.2, the residues lining the pocket are very similar in all the three forms of Rap-PSER11. Most of the residues that line the nucleotide pocket, not surprisingly, belong to Switch I, Switch II and P-loop (as seen from the colouring of the residues in Table 3.2). However in the case of GTP-PSER11, the residues are predominantly only from P-loop and Switch II and the Switch I loop residues are not part of the residues lining the nucleotide pocket, which can be rationalised in terms of increased mobility of Switch I loop. The pocket volume is also changed and is much smaller for the phosphorylated case. As discussed earlier, the change in the pocket location can be rationalised in terms significant perturbation to the nucleotide binding pocket of the Rap protein upon phosphorylation. The results clearly show that the location of the largest pocket is near the nucleotide in all the cases except when the protein is phosphorylated in the presence of GTP (GTP-PSER11 case). The observation that the largest pocket size in the case of phosphorylated Rap protein, in the presence of GTP, is smaller than all the other case, suggests a more compact structure when phosphorylated and in the presence of GTP which is consistent with the PCA results in Figure 3.11.

No.	System-Name	Residues lining the largest pocket	Pocket Vol.
1	Rap-crystal	GLY12 GLY13 VAL14 GLY15 LYS16 SER17 ALA18	521.01 Å ³
	structure:	PHE28 VAL29 GLU30 LYS31 TYR32 ASP33 PRO34 THR35	
	1C1Y.pdb	ASP57 THR58 ALA59 GLY60 THR61 ASN116 LYS117 ASP119 ALA148 LYS149	
2	GTP Case	SER11 GLY13 VAL14 GLY15 LYS16 SER17 ALA18 VAL21	605.96 Å^3
		PHE28 VAL29 GLU30 LYS31 TYR32 ASP33 PRO34 THR35 ASP38 TYR40	
		ASN116 LYS117 ASP119 SER147 ALA148 LYS149	
3	GTP-PSER11	VAL8 LEU9 GLY10 LYS16 ILE36 GLU37	278.01 $Å^3$
	Case	ASP57 THR58 ALA59 GLU62 GLN63 PHE64 MET67 ARG68 TYR71 LEU96	

Table 3.2: The systems for which the pocket analysis is done. Other than for crystal structure, the pocket analysis was done on averaged structure of 395 to 400ns of trajectory data. The color code used is red for P-loop, blue for Switch I and green for Switch II residues.

3.3.3 Effects of phosphorylation on complex and interactions at the interface

In this section we describe in depth the effect of phosphorylation on the interaction between Rap and Raf. Studies have shown that one of the functions of Rap proteins is to bind to the Raf effector via the RBD domain effectively trapping the Raf protein in an inactive $\operatorname{complex}[136]$. The interaction between Ras and Raf is essential for activation of Raf kinase domain (which is located in the C-terminal half of the Raf protein), which plays a crucial signalling role in the MAPK pathway and Rap proteins can interefere with this mechanism by making Raf unavailable to Ras proteins. Proteins in Ras superfamily, including Rap protein, interact with effector proteins like Raf via Switch I loop[124, 174]. Experimental studies [50] have shown that single mutation of a conserved residue like THR 35 can significantly alter the dynamics of the Switch I loop and consequently affect the interaction of Ras superfamily proteins with the effector proteins like Raf. Switch I region (also called effector loop) is identical between Ras and Rap proteins. Protein-protein binding exploits the inherent flexibility of the proteins to undergo conformational changes and form a $\operatorname{complex}[180, 181]$. As seen in previous section, phosphorylation of SER11 residue leads to considerable changes in the mobility of the two main functional loops: Switch I and II of Rap protein which liganded with GTP. RMSF of Rap protein in the presence of GTP ligand shows that, the mobility of the Switch II and Switch I loop decreases and increases respectively (Figure 3.3). In this section we explore the changes to the interface of Rap with Raf when SER11 of Rap is phosphorylated.

Unlike the majority of protein-protein interfaces which have more hydrophobic contacts, the crystal structure of Rap-Raf suggests that there are many polar interactions between the Switch I loop of Rap and RBD of Raf [159, 182]. The Switch I loop is not in close contact with the nucleotide binding pocket region in Rap protein but residue TYR32 and its conformational changes when GDP is exchanged for GTP, plays a crucial role in binding of the Switch I loop with the RBD of Raf protein [128]. This conformational change in Tyr 32 presumably facilitates the formation of a polar contact between residues ASP38 of Rap and ARG89 of Raf. As shown in previous section, the phosphorylation of SER11 significantly perturbs and reorganizes the nucleotide binding region and the surrounding loop conformations (See Figure 3.6and Figure 3.8). Due to breaking of strong polar interaction between GLU 30 and hydroxyl groups attached to ribose moieties of GTP ligand on phosphorylation of SER11, the Switch I acquires additional mobility and the loop moves away significantly compared to the unphosphorylated Rap-GTP case (See Figure 3.6). This movement of Switch I loop, on phosphorylation, results in introduction of additional interactions between RBD of Raf protein and Switch I loop. Figure 3.13 (a) shows the evolution of distance between TYR32 of Rap and LYS84 of Raf and in the case of phosphorylated Rap-GTP protein, the movement of the Switch I loop decreases the distance between Switch I loop and LYS84 of Rap by more than 6 Å. The consequences of such movement can be seen in the polar interactions between the RBD of Raf and residues of Rap at their interface (Figure 3.13 (b) and in Figure 3.14). The time evolution of distance profiles suggest that few polar interactions like SER39-ARG89, GLU37-ARG59, ASP33-ARG84 remain unchanged with phosphorylation of SER11. However phosphorylation leads to changes in other polar interactions

at the interface with disruptions in ASP33-ARG73 and GLU37-ARG67 interactions and enhancement in GLU54-ARG67 interaction. The polar interactions involving ASP38 of Rap are required interactions for effector binding to the Rap protein [128] show positive enhancement upon phosphorylation. As can be seen in Figure 3.13 (b) and FIG S4(c), in phosphorylated Rap case, the ASP38 residue forms longsurviving ionic interactions with both ARG89 and THR68 residues of RBD of Raf protein suggesting increased binding between Rap and Raf with phosphorylation. Allosteric network analysis confirms this enhanced interaction (Figure 3.15).



Figure 3.13: The time evolution of (a) distance between Rap:TYR32(CA) and Raf:LYS84(CA) atoms, (b) distance between Rap:ASP38(OD1) and Raf:THR68(OG1) atoms.

From the allosteric network analysis, we see that the number of detected communities remain the same (i.e 10 communities) however, the organization especially at the interface is drastically different when phosphorylated is drastically different, especially at the interface, when Rap is phosphorylated. The most striking aspect of the network analysis is that the GTP-PSER11 case has atleast 3 communities in common to the interface between Rap and Raf proteins (cyan, purple and red communities in Figure 3.15 (b)), which is absent in the unphosphorylated case. New community (shown in purple in Figure 3.15 (b)) that connects the Switch II loop



Figure 3.14: Time evolution of distances between critical residues at Rap:Raf complex interface (Rap bound with GTP) with and without SER11 phosphorylated.

with the L4 loop of the Raf protein (which is closer to the Cysteine Rich Domain part of Raf), via the protein-protein interface emerges in the phosphorylated form . In a recent combined experimental and simulation study on a Ras-Raf complex, a similar result was obtained due to mutation of GLN61LEU[175]. The mutation resulted in altering the allosteric pathways in which a single community network was found to form between the interface of Ras-Raf complex and the distant L4 loop of



Figure 3.15: The communities detected in a)GTP- and b)GTP-PSER11 cases. The yellow community represents the L4 loop.

Raf. The present observations of global effects of phosphorylation of a single residue SER11 in Rap protein reiterates the fact that such changes can be allosterically communicated to spatially distant regions in the complex and suggest how a local mutation can have global effects.

3.3.4 Effects of phosphorylation on binding energy of the complex

In the previous sections, we described how the phosphorylation of SER11 affects the dynamics of functional Switch I and Switch II loops and consequently how interacting communities spanning the complex interface emerges. In this section we look at the effect of phosphorylation on the binding free energy of the Rap-Raf complex, in particular to gain insight into the emergence of community across the complex interface. To do this, we used the standard MM-GBSA technique, as described in the Methods chapter. The results of the MM-GBSA calculations are shown in Table 3.3.

Contribution	GTP-SER11 system (Kcal/mol)	GTP-PSER11 system (Kcal/mol)
ΔG_{bind}	-62.88 ± 8.59	-68.67 ± 8.35
ΔE_{elec}	-498.66 ± 109.20	-520.18 ± 61.72
ΔE_{vdW}	-45.88 ± 5.69	-48.48 ± 5.92
ΔG_{solv}	481.66 ± 104.23	500.00 ± 57.16
$\Delta G_{solv-np}$	-8.95 ± 0.54	-10.01 ± 0.51
$\Delta G_{solv-polar}$	490.62 ± 104.17	510.02 ± 57.19

Table 3.3: Free energy contribution of GTP liganded simulations.

The binding energy values show that ΔG_{bind} of the complex is lower when the SER11 is phosphorylated suggesting an increased binding of Rap-Raf complex. It can be seen that the contribution to the increased binding energy of the complex has main contribution from the and ΔE_{elec} and $\Delta G_{solv-polar}$ terms. The binding free energy results and its contributions are consistent with increased interactions at the interface as seen in formation of long-surviving ionic interactions between ASP38 residue of Rap with both ARG89 and THR68 residues of RBD of Raf protein in Figure 3.13 and also the emergent community at the complex interface shown in Figure 3.15 (b). The predominant contribution of electrostatic interactions to the binding free energy, seen here, is consistent with earlier work on thermodynamic analysis of Ras/Effector Complex Interfaces [183]. The entropy calculations were performed to understand any possible effects of phosphorylation of SER11, using the quasi-harmonic approach, and are shown in Table 3.4. The calculations clearly suggest that effects of phosphorylation on the overall complex is minimal but, the Rap protein is less dynamic when SER11 is phosphorylated. This result is consistent with our PCA results, which suggests that the conformational sampling of the Rap protein is smaller when SER11 is phosphorylated.

Soloction	GTP-SER11 system		GTP-PSER11 system	
Delection	$\Delta S (kcal/K/mol)$	T ΔS (kcal/mol)	$\Delta S (kcal/K/mol)$	T ΔS (kcal/mol)
Rap-Raf complex	1.7842	531.6916	1.7845	531.781
Rap chain	1.1707	348.8686	1.1624	346.3952
Raf chain	0.5978	178.1444	0.6137	182.8826
Rap-Raf Interface(3Å)	0.2005	59.749	0.2339	69.7022

Table 3.4: Quasiharmonic entropy contribution of GTP liganded simulations.

3.4 Discussion

The focus of the present work is to understand the effects of phosphorylation of a single residue SER11 in Rap on the interactions between Rap and RBD of Raf and speculate on how network of new interactions, that emerge as a consequence of phosphorylation, are used to communicate the changes at the interface to the distal region of Raf. The SER11 phosphosite is present in the P-loop (see Figure 3.1), close to the Switch I region and is relatively buried. This phosphorylation was identified earlier in an independent experimental study. [67]. The mobility of the Rap protein, and in particular those of various functional loops, as seen from the simulations here suggest that the dynamic nature of the loop regions can potentially expose this buried phosphosite. Indeed solvent accessible surface area (SASA) measurements of SER11 residue in a 400 ns long simulation (when unphosphorylated) and with GTP ligand revealed that for a significant time on the simulation time scale, the SASA value of SER11 is greater than that of its crystal structure value (See Figure 3.5). This also corroborates the observation in Figure 3.3 that the Ploop region is marginally dynamic. These two observations together lend credence to our assumption that though the phosphosite is buried in the crystal structure, the inherent protein dynamics can potentially expose the site to solvent and other relevant kinases which can phosphorylate the SER11 residue. Further simulations of the SER11 Phosphorylated GTP bound Raf Rap complex helped us to understand the effects of the phosphorylation on the possible global conformational changes in Rap protein, its interface with RBD of Raf and possible allosteric effects transmitted

to spatially far away locations.

All of the observed changes upon phosphorylation converge on one unifying theme: SER11 phosphorylation stabilizes the GTP bound Rap structure, likely to prevent GTP hydrolysis by pulling and closing the Switch II loop over the active site and establishes an allosteric network that can potentially transmit these changes to the distal L4 loop of Raf. Furthermore the PCA results indicated that the conformational space occupied by this SER11 GTP bound Rap is dramatically different from the unphosphorylated form and reflects a tighter conformation of the protein. The simulations also reveal that there is significant alteration in the pocket location and its size when the SER11 residue in Rap is phosphorylated. This new pocket is flanked predominantly by residues in Switch II loop, and is different from the unphosphorylated form in which the largest pocket is surrounded by Switch I loop residues. The allosteric network analysis suggests that there is an increase in communities across the interface with phosphorylation. More significantly a single community spanning residues in the Switch II loop all the way to distal L4 loop emerges on phosphorylation. The net effect of SER11 phosphorylation is an allosteric relay of signals from Switch II region in Rap to L4 loop in RBD of Raf kinase which could result in constitutive activation of Rap and consequently that of Raf kinase potentially affecting the downstream signalling. A recent work on Hsp90 family of proteins also suggested that such buried post translational modification sites can play an important role in allosteric conformational changes and can potentially act as mediators of global dynamics in the Hsp90 structures [184]. The binding free energy calculations concur with the other results and show that the binding of Rap, with GTP ligand, with RBD of Raf is stronger when SER11 is phosphorylated, though the difference may not be very strong given the possible errors in such calculations. This plausible increased binding has its origin in favourable electrostatic interactions between residues of the two proteins at the interface due to cascading effects of phosphorylation of SER11 and can have important consequences in the downstream MAPK

signaling pathways. One of the functions of Rap proteins is to competitively bind to Raf, without activating it, and disrupt the Ras-Raf binding [132]. Increased binding of Rap to RBD, due to phosphorylation of SER11 residue, can potentially make Raf even less available to Ras. In addition, such phosphorylation is also seen to induce possible long-range allosteric communication between Rap (via Switch II) and L4 loop of RBD of Raf which connects RBD to the cystein rich domain (CRD) of Raf. The entropy calculations suggest that the phosphorylation of SER11 does not change the overall entropy of the complex significantly but when considered individually, the Rap protein has lower entropy inits phosphorylated state. This result is quite consistent with our PCA results and together they suggest that the Rap protein has a 'tighter' conformational sampling when phosphorylated. This possibly due to increase in favourable electrostatic attractions within the Rap protein, which are are cascading effects of phosphorylation of SER11. Whether such increased binding of phosphorylated Rap with RBD of Raf can trigger conformational changes in the CRD of Raf, further affecting the binding of Raf to Ras, since it has been suggested that CRD of Raf also binds to Ras[185, 134, 186, 187, 188], is open to speculation.

There is a significant parallel between these changes observed in the present MD simulations with those of Ras protein GTP bound crystal structure when GLN61 is mutated to a LEU. This mutation is a well known oncogenic mutation that prevents GTP hydrolysis locking Ras in a constitutively active form[175]. The authors predicted that the extended long range allosteric effect transmitted across the interface to the L4 loop is responsible for the kinase activity of Raf. The similarity between the two observations strongly suggests that the phosphorylation at SER11 mimics the oncogenic mutation in Ras which when extrapolated to function suggests that Rap may be constitutively activated by such phoshorylation. This SER11 phosphorylation in Rap was observed in Hela cells when the EGFR is activated. This phosphorylation increases upon nacodazole treatment[67]. Other highthroughput studies have detected the same phosphoryation in tumors[157]. Although this phos-

phorylation does not seem to occur at high enough occupancy, it is nevertheless detected with high confidence. These observations suggests a likely scenario. It is possible that this phosphorylation of Rap happens in normal cells during ligand binding to receptors. It is also possible that aberrant signalling due to some over active kinase often seen in cancers may phosphorylate Rap and activate it. Consequently this may lead to disruption of Rap-Mapk signaling or it may independently activate other RBD domain containing effector proteins. Such a possibility is supported by the significant parallel between the observations reported here and those of Ras protein GTP.

A concern which can arise regarding our study is about the length of simulations. We believe that the length of simulations (400ns for GTP-PSER11 case) in this study is adequate enough to deduce the present results. Earlier simulations on similar systems [175] had a time scale of \sim 100ns long and we believe our longer time simulations can improve the equilibration times considerably. However, we do acknowledge that for more accurate results, more number of simulations with different initial conditions are required for improved sampling. New and advanced simulation methodologies like Markov state modeling and enhanced sampling simulations can help in this matter considerably. The limited number of simulations, coupled with aspects of sub-microsecond long simulations and possible limitations of forcefields used in the simulations in general. It would be also very useful if experimental evidence can also corroborate the quantitative results like free energy calculations obtained via simulations.

In summary we believe that this case study is an example as to how integrating tools that can probe dynamics can yield wealth of biological information hidden in crystal structures and highthroughput studies. They can provide probable mechanism by which single site PTM or point mutations affect functions of a protein. In addition these results reveal new binding pockets in proteins not evident in static crystal structures but evolve due to dynamic changes in proteins. Such dynamic pockets may be trapped by small molecules to inhibit the functions of the protein thus expanding the repertoire of druggable genome space. Arguably one may target the kinase responsible for the phosphorylation of such proteins thus providing alternative strategies to inhibit the functions of notoriously un druggale and elusive protein such as the Ras GTPases.

Chapter 4

Phosphorylation of interfacial phosphosite: Effect on complex stability of Rap-Raf

4.1 Introduction

Rap proteins belong to the small Ras-like GTPase family which are involved in important cellular functions such as cell adhesion, cell-cell junction formation and regulation of the actin cytoskeleton[136, 137, 138, 139]. GTPases acts like molecular switches which switch between "active" GTP-bound and "inactive" GDP-bound states. The rates of switching activity of these molecular switches are determined by factors like guanine exchange factors(GEFs), which catalyze Rap/Ras proteins to GTP-bound active state and GTPase activating proteins(GAPs) which enhance the intrinsic hydrolysis of bound GTP to GDP[189]. Ras-like GTPases and their mutations have been identified as some of the key drivers of oncogenesis in several human tumors and intense efforts have been made to understand the basis of modulation of their activation and signaling via different methods. Refer to the Introduction section of Chapter 3 for more details.

Reversible PTMs like phosphorylation^[190] are integral part of many signaling pathways and it is no surprise that ligand-based activity switching proteins like Ras-like GTPases have been shown to possess multiple PTM sites, which can be potentially targeted for the development of small molecule inhibitors that might alter the function of these GTPases. This alternate paradigm of limiting the activity of Ras-like GTPases becomes especially relevant given the failure to find successful direct inhibitors of Ras proteins. We have recently showed that the phosphorylation of a single serine residue (SER11), located proximally to the bound GTP ligand triggers a cascade of interactions resulting in enhanced binding affinity of the Rap-Raf complex^[191]. A possible implication of this increased binding affinity of the complex can potentially inhibit Ras-Raf interaction by making Raf less available to Ras proteins thus affecting the downstream signalling pathway. Our simulations also demonstrated that a new pocket is revealed in the phosphorylated Rap protein and a possible allosteric pathway opens up, spanning more than 40Å between the phosphorylated SER11 and the L4 loop of RBD of Raf. Experimental studies [67] revealed more possible phosphosites in Rap protein including SER39 which is located on the Switch I loop, which is implicated in many complexes of Rap and its corresponding effector proteins [192, 159]. Switch I loop plays an important role in many interactions of Rap with the Ras binding domain (RBD) regions of effector or regulator protein like Raf and the inactive (GDP-bound) to active (GTP-bound) switching of protein is accompanied largely by conformational changes in Switch I and Switch II loops.

In this work, we explore the effects of phosphorylation of SER39 residue which is located on functionally relevant Switch I loop and at the Rap-Raf complex interface(See the Rap:Raf complex visualization in Fig. 4.1). The buried location of this phosphosite at the interface of the complex presents two issues to be probed:(1) what are the effects of phosphorylation of a residue at the complex interface? (2) does the residue gets phosphorylated before or after formation of the complex and what are the structural differences at the complex interface in both these cases?



Figure 4.1: Rap-Raf protein complex (crystal structure, PDB ID 1C1Y) visualization showing the location of important functional loops like P-loop, Switch I, Switch II and RBD loop regions. The phosphosites SER11 and SER39 is shown in green.

4.2 MD Simulations

To understand the role of phosphorylation of SER39 on the dynamics of Rap-Raf complex, we performed a series of simulations details of which are given in Table 4.1. The systems simulated include a control simulation in which no residue in the complex was phosphorylated (SYS1), a system in which the residue SER39 was phosphorylated in the complex (SYS2) and a third system where only the Rap protein with phosphorylated SER39 was simulated first and the RBD-Raf was docked on to the phosphorylated Rap and simulated further (SYS3). We attempt to understand if the sequence of phosphorylation and complex formation plays an important role in complex stability as well as structural rearrangement of the Rap and Raf proteins with respect to each other, given that the proposed phosphosite SER39 is buried at the complex interface.

The initial structure for all the simulations is taken from Protein Data Bank with id 1C1Y[159], with a resolution of 2.2Å, and contains Rap1A (referred to as Rap henceforth) liganded with a GTP-analogue molecule, Mg²⁺ ion and RBD (Ras binding Domain) region of Raf protein variant c-Raf1 (referred to as Raf henceforth). The missing atoms were modelled using MODELLER software [193] and the appropriate protonation states for all ionizable residues were determined using PDB2PQR server [164]. All atom molecular dynamics (MD) simulations were performed on all the systems listed in Table 4.1 using NAMD[85] software and CHARMM36[160, 161] forcefield. VMD software [162] was used to visualize the system, do the basic system setup and data analysis(with the help of inbuilt Tcl scripting language) etc. Each system was solvated in a water box (using TIP3P[165] water model) and overall charge neutrality was achieved through the addition of appropriate counter ions and to maintain 150 mM salt concentration. The systems were then subjected to energy minimization runs using the conjugate gradient method for 5000 steps, followed by MD simulation runs in NPT ensemble with the integrator time step set to 2fs during equilibration and production runs. A constant pressure of 1 atm was maintained by using the Nosé-Hoover-Langevin piston algorithm [166, 167] and the temperature of 298K was maintained using Berendsen thermostat [168]. A cut-off distance of 12Å was used to compute all short-range van der Waals (VDW) interactions and the long-range electrostatics interactions was treated with the Particle Mesh Ewald(PME) method [96, 169].

No.	System-Name with temperature	Description (System is in TIP3P water box)
1	Rap:Raf-GTP(SYS1)	Rap:Raf protein complex with GTP ligand (400ns)
2	Rap:Raf-GTP-PSER39 (SYS2)	Rap:Raf protein complex with GTP ligand and PSER39 (550ns)
3	Only Rap-GTP-PSER39 (SYS3a)	Rap protein with GTP ligand and PSER39 (300ns)
4	Rap:Raf-GTP-PSER39 (SYS3)	Rap:Raf protein complex with GTP ligand and PSER39* (550ns)
		*(Rap chain taken from the final frame of system no. 3 'SYS3a')

Table 4.1: Details of systems simulated and analysed.



Figure 4.2: The RMSF of residues of (a)Rap and (b) Raf C_{α} atoms averaged over last 50 ns of simulation for SYS2, SYS3 and CTRL cases. The relevant Switch I, Switch II and L4 loops are indicated in the figure.

4.3.1 Effects of SER39 phosphorylation on structure and dynamics of Rap:Raf complex

RMSF calculations of Rap protein shown in Fig. 4.2 (a) indicate that the mobility of Switch I is higher in GTP-PSER39(SYS2) case when comparing to GTP-PSER39(SYS3) and unphosphorylated CTRL(GTP- case as mentioned in the previous chapter[191]) cases. This can be attributed to the fact that Rap protein interacts in a very strong and stable manner with Raf via Switch I in the GTP-PSER39(SYS3) and CTRL cases comparing to GTP-PSER39(SYS2) case. This will also be evident from the saltbridge distance calculations and binding free energy calculations which are presented in the coming parts. The Switch II mobility is lower for GTP-PSER39(SYS2) case comparing to other cases. The L4 loop of Raf protein shows higher mobility in all cases.(See Fig. 4.2 (b))

The network analysis also lead us to the same conclusion as mentioned before. See Fig. 4.3 for the network representation of GTP-PSER39 cases. There are 8 communities in the case of GTP-PSER39(sys2) case and 5 communities in the GTP-PSER39(sys3) case. The important point to note is the connections at the interface and how it is linked to Loop L4 region. There are 3 common networks in both cases. But the GTP-PSER39(sys3) network which includes SER39 residue in Rap chain spans a community which encompasses the loop L4 of Raf chain(this community is colored in red, see Fig. 4.3). This imply that the effect of phosphorylation can affect the distal Raf L4 loop possibly via multiple paths in that single community and influence the downstream signalling. L4 loop of Raf-RBD is significant because it can possibly interact with the cystein rich domain(CRD) of Raf and there by propagate downstream signal, due to the allosteric signal which originates at the interfacial SER39 residue.



Figure 4.3: The communities detected in a)GTP-SER11 and b)GTP-PSER39(sys2) and c)GTP-PSER39(sys3) cases. The yellow community represents the L4 loop.



Figure 4.4: Conformer plot of Rap:Raf complex considering the last 100ns of the trajectory. The plot shows conformational space sampled by Rap:Raf complex in terms of PC1(42.80%) and PC2(23.97%).

We see that the Rap chain in GTP-SER39(sys3) case explores a larger conformational space compared to GTP-SER39(sys2) case (See Fig. 4.4). This is because of the higher fluctuation in the Switch II loop of Rap and aminoacids surrounding 115th residue of Raf in the GTP-SER39(sys3) case(which can also be seen in RMSF plots from Fig. 4.2). Although the Rap chain in GTP-PSER39(sys3) case exhibits stable and strong interaction with Raf chain, the Rap and Raf chains sample a higher configurational space.

4.3.2 Effects of phosphorylation on the interactions at the Rap:Raf complex interface

Saltbridge distance analysis between Rap and Raf proteins along with community network analysis elucidate the effects of phosphorylation at the Rap:Raf complex interface. In the pre-simulated Rap with SER39 phosphorylation accompanied by Raf docked case(SYS3), there are 5 stable saltbridges namely GLU37(Rap)-ARG59(Raf), GLU37(Rap)-ARG67(Raf), ASP33(Rap)-ARG73(Raf), ASP33(Rap)-LYS84(Raf) and ASP38(Rap)-ARG89(Raf). The distances alongside with their histogram calculations shown in Fig. 4.5 reveals that these are very tightly bound interactions. The saltbridge bonds between Rap and Raf chains reflects the strong and stable interaction in the SYS3 case. The saltbridge distances in SYS2 and CTRL cases exhibit higher variance than SYS3 case. SYS3 exhibits tight saltbridge interactions unlike other two cases. The Raf interacting Rap residues (GLU37, ASP33, and ASP38) are predominantly negative residues and Raf residues at interface are mostly positive residues (ARG59, ARG67, ARG73, LYS84, and ARG89). This was reported in past[194] and explains why this Rap:Raf interface is usually stable. SER39 phosphorylation have made these interaction way more favourable electrostatically, which is also visible in binding free energy calculations which will be discussed in coming subsection.



Figure 4.5: The saltbridge distances between Rap and Raf domains for the last 100ns shown in time series data and their corresponding normalized histograms.

4.3.3 Effects of SER39 phosphorylation on binding free en-

ergy of the complex

Contribution	GTP-SER11	GTP-PSER11	GTP-PSER39	GTP-PSER39
(KCal/mol)	300-400ns	300-400ns	(sys2) 450-550ns	(sys3) 450-550ns
ΔG_{bind}	-62.88 ± 8.59	-68.67 ± 8.35	-47.01 ± 12.09	-67.52 ± 7.47
ΔE_{elec}	-498.66 ± 109.20	-520.18 ± 61.72	-612.50 ± 88.47	-638.40 ± 48.94
ΔE_{vdW}	-45.88 ± 5.69	-48.48 ± 5.92	-27.25 ± 12.68	-31.13 ± 5.61
ΔG_{solv}	481.66 ± 104.23	500.00 ± 57.16	592.74 ± 88.76	602.01 ± 45.01
$\Delta G_{solv-np}$	-8.95 ± 0.54	-10.01 ± 0.51	-9.87 ± 0.53	-9.24 ± 0.47
$\Delta G_{solv-polar}$	490.62 ± 104.17	510.02 ± 57.19	602.62 ± 88.71	611.25 ± 45.00

Table 4.2: Free energy contribution(last 100ns) of GTP liganded simulations.

We carried out binding free energy calculations via MM-GBSA technique as men-

tioned in the Methods chapter. The binding free energy calculations suggest that the Rap:Raf complex interaction in GTP-PSER39(sys3) is very strong and stable than in GTP-PSER39(sys2) case. See Table 4.2 for the free energy calculations. It is very evident from the average ΔG_{bind} values of GTP-PSER39(sys2) case and GTP-PSER39(sys3) case which are -47.01 Kcal/mol and -67.52 Kcal/mol respectively. The most favourable interaction between Rap and Raf complex occurs in the GTP-PSER11 case[191]. The values corresponding to GTP-PSER39(sys3) case are very comparable to that of in the case of GTP-PSER11 case. The binding free energy between Rap and Raf for GTP-PSER39(sys3) case(-67.52 ± 7.47 KCal/mol) is comparable to that of GTP-PSER11 case(-68.67 ± 8.35 KCal/mol)[191], but the electrostatic component dominates in the former case(-638.40 ± 48.94 KCal/mol) versus -520.18 ± 61.72 KCal/mol). The binding free energy calculations are also consistent with the saltbridge distance calculations(See Fig. 4.5). All the GTP-PSER39(sys3) case saltbridge distances between Rap and Raf are very stable.

4.4 Discussion and Concluding Remarks

Our molecular dynamics study about SER39 phosphorylation in Rap:Raf complex looked into both scenarios namely phosphorylating blindly at SER39 in the complex and the case where pre-simulated SER39 phosphorylated Rap docked with Raf. The results reveal that phosphorylating SER39 prior to Rap:Raf complex formation case has higher binding affinity between Rap and Raf proteins. This can be seen from the binding free energy calculations, especially the electrostatic component of binding free energy which is the most favourable in the above mentioned scenario. These results are also in agreement with the saltbrige distance calculations between Rap and Raf proteins. The SER39 phosphorylation event is more probable to occur before complex formation, since complex formation will hinder the possibility for the kinase to phosphorylate at this buried interfacial region. As a consequence of
interfacial phosphorylation at SER39 in this Rap:Raf complex, we expect that Rap interface interacts more strongly with Raf chain. And this is indeed what we see via community network analysis, where we observe that almost all Raf residues form into a single community and overlaps with Rap SER39 itself[See Fig. 4.3 (c)]. This also imply about the possibility of a strong allosteric signal propagation originating at SER39 as a result of phosphorylation till Raf terminal.

It is known that Rap proteins are antagonists to Ras proteins when it comes to binding to Raf proteins and so Rap competitively bind to Raf, without activating it, and disrupt the Ras-Raf binding[132]. So a conclusion which can be deduced from favourable Rap:Raf binding affinity upon SER39 phosphorylation is that stronger Rap interaction with Raf possibly imply that malignant Ras proteins are denied access to Raf proteins thereby halting the faulty cancer causing signal. One point to note is that the favourable binding energy in the pre-simulated SER39 phosphorylated case(SYS3) is dominated by the electrostatic component. And that is assumed to have origins in the proximity of phosphosite SER39 towards the Rap:Raf interface versus distal SER11 site.

Molecular dynamics studies like this can provide insights about tackling malignant undruggable proteins(example: Ras GTPases) via alternate techniques like post translation modifications. Recent computational studies were also significant in understanding allosteric signal propagation in protein complexes and conformational dynamics of proteins. These in-silico studies along with wet lab experiments will hopefully lead the fight against undruggable oncoproteins.

Chapter 5

Interdigitation of lipids induced by membrane active proteins

5.1 Introduction

Many small molecules and proteins (both peripheral and integral) are known to interact with the cell membranes in a non-specific way and affect the integrity of the membrane structure [195]. These collective membrane active agents can interact both with the hydrophobic tails and polar or charged head groups and can affect structural properties of membranes like membrane thickness, active potential, deuterium order parameters of the acyl chains, and partial pressure etc. Dynamical properties of the cell membranes such as diffusion, permeability, fluidity and lipid flip rates can also be affected by membrane active agents [196]. Alteration of membrane properties by such membrane-active agents is important to explore since lipid membrane environment hosts many membrane proteins and such agents may also affect the functioning of these proteins in an indirect way [197, 198]. In addition to small drug-like molecules, studies using model membranes have also looked at how peptides or biomimetic polymers interact and effect the cell membrane integrity. The class of peptides and polymers belonging to antimicrobial agents have been the most studied in this category. Various experiments and simulations have shown that these membrane active petides/polymers affect the membrane structure profoundly [199, 200]. In some cases, this leads to domain formation, which eventually results in cell lysis. Domain formation occurs in cell membranes in the absence of such membrane active agents as well [201, 202, 203]. However, usually the cell membrane has enough time to recover from such local inhomogenieties but when membrane active agents interact with the cell membranes, there is not sufficient time for the membrane to recover and integrity is severely compromised. All these studies strongly suggest that the agent-membrane interactions exploit certain inherent "defect-forming" behavior of membrane lipid head groups [204].

One important aspect of the cell membranes is the nature of the phase in which they exist, which can affect their function [25, 26]. The lipid bilayers are known to exhibit varying degrees of lipid molecule packing resulting in different phases. Lamellar lipid bilayers exist predominantly in two phases: $gel(L_{\beta})$ and liquid-crystalline or fluid(L_{α}) phases. Gel phase is observed at low temperatures and is characterised by ordered lipid molecules with lipid tails tilted or untilted with respect to bilayer normal. Gel phase can also exist in an interdigitated $(L_{\beta}I \text{ phase})$ fashion where the upper and lower membrane leaflets overlap each other [27]. The fluid phase is the most common phase observed and is required for many normal biological functions, in which the lipid tails are disordered and along the membrane normal and certain proteins have the ability to sense the fluidity of the membranes for interaction [28]. Hence the lipid bilayers exhibit thermotropic behavior and change phases as the temperature is increased and the critical temperature at which the lipid bilayer changes from gel to fluid phase is called the main transition temperature (T_m) and varies with many factors including lipid chain length, degree of unsaturation in the lipid tails, water content and possibile mixing of more than one lipid species etc. Apart from the gel and fluid phase, a third phase called ripple phase (P_{β}) has been discovered in 1973 [30] and are corrugated in nature with coexisting gel and fluid phases [31]. Simulations also reported existence of such phases [205, 202, 206, 207]. Ripple phase itself is categorized into asymmetric and symmetric ripples [208, 209]. The asymmetric ripple contains interdigitated lipids, whereas symmetric ripple has no interdigitated lipids.[210] Biological membranes are mostly composed of amphipathic phospholipid molecules which form bilayers. The most common phospholipids observed in cells are PhosphatidylCholines(PC). Recent allatom MD experiment studies have reproduced Ripple phase in PhosphatidylCholine membranes[206].

There are many studies which probe the effect of small molecules on the phase transition temperature of the membrane [211, 212, 213, 214]. These studies have shown that the molecule not only perturb the membrane structure but can also affect the transition temperature between gel and liquid phases and that the response of the lipid membrane to these small molecule interactions is sensitive to temperature especially in the vicinity of the transition temperature. Apart from small molecules, membrane active proteins also perturb the membrane structure significantly. Inherent lateral inhomogeneity in the lipid membranes is now well established in terms of presence of ordered and densely packed domains amidst disordered and loosely packed lipid molecules [215, 216]. This has been described in the literature as 'lipid rafts' and much research has been devoted to detect and understand the functional role of such domains which are both small in size and also have small lifetimes, as the membranes self-heal on a time scale of the order of milli seconds [217, 218]. Proteins are known to induce order into the membranes, especially when they partition into them, almost to the order of 50 Å or more [219, 220].

Experiments have shown that alcohols, when interacting with model membranes, affect the transition temperature and in many cases induce interdigitation and/or ripple phase. Recently experiments on local anesthetic PEtOH have also shown

that the anesthetic molecule affects the thermotropic behaviour of the lipids and causes interdigitation in the DPPC membranes in a concentration dependant manner [40, 84]. In the past, certain case studies have reported that interaction of phospholipids with other molecules like siRNA, charged nanoparticles, and membraneactive proteins can modulate their transition temperature [41, 4, 5]. Prates et al [41]study was a coarse-grained MD work which showed that charged nanoparticles on interaction with DPPC(1,2-dipalmitoyl-sn-glycero-3-phosphocholine) lipids brings disorder to the system and lower their intrinsic transition temperature from 315K to even 285K. Choubey et al^[4] performed an all atom MD study with siRNA interacting with DPPC lipids and found that the interaction induces L_{α} to P_{β} transition at 323K. Welker et al [5] in their experimental study of membrane-active protein named Hsp12 in interaction with DMPG(dimyristoylphosphatidylglycerol, $T_m=298$ K) lipids observed ripple phase structures even at 313K. Interdigitation in lipids can be induced by molecules mentioned before via voids/defects which they create in membranes. The amphiphilic molecules like alcohols with short chain length and hydroxyl group interacts with membrane head groups and increases the head group surface area and create voids. As these hydrophobic voids are energetically unfavourable they are filled with opposite monolayer molecules which induce interdigitation[84, 221, 222, 223]. Other molecules which participate in membrane remodeling processes are membrane active proteins. For example, high concentration of membrane active protein α -Synuclein induces membrane thinning, lateral expansion of lipid molecules and positive curvature in membrane [34]. On the other hand there are membrane active peptides like ALPS(Amphipathic Lipid Packing Sensor) motifs which binds to lipid defects and senses curvature of membrane^[224]. Antimicrobial peptides (AMPs) are a class of membrane active peptides which disrupts the bacterial membrane integrity via multiple variety of mechanisms[3].

Nogo-66 is part of the extracellular domain of the neurite outgrowth inhibitor (Nogo) protein and has been shown to undergo disorder-order transition in the presence

of phosphocholine membrane environment [2]. One of the functions of Nogo-66 is to act as a scaffolding protein in defining the membrane architecture and is a membrane active protein. Given the relevance of the membrane to the structure and consequently the function of Nogo-66, all-atom molecular dynamics simulations have been undertaken to assess the effect of interaction of Nogo-66 with phosphocholine model lipid bilayers as a function of temperature. It needs to be stated that Nogo-66 is used as a model protein to understand the general effects of membrane active proteins on phase behavior of lipid bilayers, a general understanding of the same is sought in the growing examples of different biological molecules inducing dramatic global structural changes in the lipid bilayers. Particular emphasis is laid on the delicate temperature range around the transition temperature, the role of packing defects in lipids in conjuction with the temperature, which results in interdigitated states in lipid bilayer systems.

5.2 Methods

5.2.1 MD Simulations

Atomistic molecular dynamics(MD) simulations under NPT ensemble were performed to study the interaction of Nogo-66 with DMPC(1,2-Dimyristoyl-sn-glycero-3-phosphocholine) bilayers, whose main transition temperature T_m is 297.15K and 300K experimentally [225] and in simulations [226] respectively. The simulations in the present study were performed at two different temperatures: 300K and 310K, near and above T_m to understand the role of temperature in conjunction with protein interaction on the membrane structure. A control simulation, consisting of only the bilayer, was performed to observe the interdigitated state of the membrane at 290K, well below the T_m of DMPC. The NMR structure of Nogo-66 protein was downloaded from protein data bank with PDB id 2KO2. 2KO2 comprises of 10 NMR conformations of Nogo-66 peptide. The NMR structure was prepared at pH 4 in water.^[2] Equilibrated membrane patches made using a CHARMM-GUI Membrane builder module. [100, 101] Each leaflet of DMPC patch has 144 lipid molecules. Simulations were done for the systems listed in Table 5.1 and in cases where the interdigitation of the membrane was observed in the presence of protein, multiple simulations were performed with different initial conditions to reproduce the same. The forcefield used for simulations is CHARMM36[160, 161] and the simulations are implemented using NAMD software. [85] The visualization was done using VMD[162] software and analysis of MD data using Tcl language (embedded in VMD), Matlab. Some analysis like cross-correlation maps are done in Grearma^[163] software and membrane analysis is done with the aid of VMD plugin called MEMBPLUGIN^[227]. The NMR structure determination of Nogo-66 (as in PDB file 2KO2) was done at pH 4. The protonation states for all ionizable residues were determined using PDB2PQR server [164] before solvating them. PSFGEN tool from VMD software was employed to assign correct protonation states. The Nogo-66 system was solvated in a water box (using TIP3P[165] water model) and overall charge neutrality was achieved through the addition of appropriate Na⁺ and Cl⁻ counter-ions. The Nogo-66 is then subjected to energy minimization runs using the conjugate gradient method for 5000 steps, followed by a 50ns MD simulation run in NPT ensemble at temperature 300K and pressure 1 atm with time step 2fs. The final Nogo-66 structure of this simulation was used for the DMPC membrane-associated NPT simulations. As an initial step, a 2ns equilibration was done using a slowly releasing harmonic potential on the protein with a step size 1fs. Another 2ns equilibration was done with no restraints. In these simulations the Na^+ and Cl^- counter-ions had a concentration of 150 mM/L. The Nosé-Hoover-Langevin piston barostat was used to maintain a constant pressure of 1atm[166, 167]. A cut-off distance of 12Å was used to compute all short-range van der Waals (VDW) interactions and the long-range electrostatic interactions were treated with the Particle Mesh Ewald (PME) method. [96, 169] The total simulation

No.	System-Name with temperature	Description (System is in TIP3P water box)
1	CTRL(300K)	DMPC bilayer lipid patch (230ns)
2	SIM1(300K)	Nogo-66 on DMPC patch $(300ns)$
3	SIM2(300K)	Nogo-66 on DMPC patch $(300ns)$
4	SIM3(300K)	Nogo-66 on DMPC patch $(320ns)$
5	CTRL(310K)	DMPC bilayer lipid patch (200ns)
6	SIM1(310K)	Nogo-66 on DMPC patch $(300ns)$
7	CTRL(290K)	DMPC bilayer lipid patch $(320ns)$

Table 5.1: Details of systems simulated and analysed.

time was nearly 2 microseconds.

5.2.2 Lipid defect analysis

PackMem tool[228] was used for the analysis of lipid defects in the membrane trajectories. The PackMem tool calculates upper and lower leaflet defects and categories those defects into "Deep", "Shallow" and "All" defects. Deep defects are those membrane voids where aliphatic atoms are deeper than a threshold distance, determined by the position of glycerol carbon atom. Shallow defects are those membrane voids where the aliphatic tail atoms are found beyond the glycerol central atom. "All" defects are union of both deep and shallow defects. Figure 5.1 shows a single DMPC lipid showing the glycerol central carbon atom. The distribution of large lipid-packing defect areas(only areas considered greater than 15 Å²) in membrane trajectories has been observed to follow a mono-exponential decay[229, 230]. For this analysis, membrane trajectory frames separated by 100 picoseconds were used in this study. The complete trajectory in each case is used for the analysis(See Table 5.1 for the trajectory lengths). The probability, p(A), of finding a defect with area A is given by

$$p(A) = b \times e^{-A/\pi} \tag{5.1}$$

where b is the pre-exponential factor and π is the defect constant(in Å²), with larger values of π indicating higher probability of finding larger defects.



Figure 5.1: A single DMPC lipid(hydrogen atoms not shown). The central carbon atom of glycerol group is highlighted. This is the reference point from where the defects are defined. Consider a threshold distance of 1 Angstrom below to the glycerol central atom. If the exposed tails are below that threshold, it is defined as deep defect.

5.3 Results

5.3.1 Emergence of interdigitation in DMPC membrane patch with Nogo-66 interaction

We observe interdigitation in the DMPC membrane system with Nogo-66 protein(SIM1(300K), SIM2(300K) and SIM3(300K) simulation systems in Table 5.1)at 300K. The coexistence of tilted interdigitated gel phase and fluid-like lipid phase regions is very pronounced in all the three simulations and snapshots of the membrane in one of the simulations, at the end of 300ns, is shown in Figure 5.2. The images are coloured according to whether the lipids are interdigitated (blue) or in fluid (red) phase respectively. The snapshot of the system at 310K is shown in Figure 5.3, which shows the fluid phase of the membrane system. Apart from the visual identification of the interdigitated lipids, a selection is also made in terms of the lipid head group defects and the corresponding residence of water molecules near the exposed lipid tails in such defect sites. A more detailed description of identifying such defect sites and the subsequent characterization of the same can be found in later sections. The figure also shows that the bilayer experiences varying thickness, with the interdigitated part of the membrane experiencing lower thickness than the fluid phase. From Figure 5.2 (b), it can also be seen that the lipid molecules in the interdigitated phase are oriented perpendicular to the membrane surface, and not tilted as seen in many classic ripple phase structures. This could likely be because the headgroups of the lipid molecules in simulations do not experience enough head group strain to orient the lipid tails in tilted fashion as suggested in [208].

To quantify the observed interdigitation of the lipid molecules, an average interdigitation gap parameter (measured as distance between terminal carbon atoms of lipid chains of both leaflets) is monitored as a function of simulation time and the results are shown in Figure 5.4. It can be seen that this gap decreases significantly in all the



Figure 5.2: (a)Top view showing DMPC membrane and (b)Side view of membrane and protein(hydrogen atoms not shown, side view angle changed to properly show the interdigitated phase) of last frame in SIM1-300K simulation case. The blue coloured lipids are the interdigitated ones and red coloured lipids are the non-interdigitated ones.

three simulations drops significantly strongly suggesting the emergence of repeatable interdigitated phase in the DMPC bilayer at 300K in the presence of Nogo-66. It should be noted that the onset of such interdigitated state differs in the three simulations, but this is to be expected given different initial conditions of the membrane structures. The emergence of ripple structure in one of the control simulations well below the T_m of DMPC (at 290 K) is also shown in the Figure 5.4. However, in



Figure 5.3: (a)Top view showing DMPC membrane and (b)Side view of membrane and protein(hydrogen atoms not shown) of last frame in SIM1-310K simulation case

other simulations (control simulations at 300K and 310K and DMPC+Nogo-66 at 310K) no emergence of interdigitated phase is observed, strongly indicating that, atleast in the simulations considered here, interdigitated phase emerges at temperatures well below the T_m of the membrane or when a membrane active protein is present in the system close to T_m . The spatial distribution of the interdigitation



Figure 5.4: The time evolution of averaged interdigitation distance between terminal carbon atoms of lipid chains, C314, belonging to two leaflets for different simulation systems, described in Table 5.1 is shown.

gap parameter across the membrane surface, over 25 ns (175 to 200 ns), is shown in Figure 5.5. The membrane surface, perpendicular to the membrane normal, is divided into regions of dimensions 2X2Å and the interdigitation gap of lipids in these regions is measured and averaged over 25 ns of simulation time for DMPC systems interacting with Nogo-66 protein at 300K and 310K. The spatial distribution of the interdigitation gap is quite uniform for membranes at 310K (Figure 5.5 (right)) and the non-zero average value of this parameter suggests that the most of the lipids are in the fluid phase. This is in striking contrast to the spatial distribution of the interdigitation gap at 300K (Figure 5.5 (left)) which shows the lateral separation of lipids into interdigitated (blue) and fluid phases, in agreement with the visual demonstration of the same in Figure 5.2. Similar spatial maps of for control systems (without Nogo-66), CTRL(300K) and CTRL(310K), are shown in Figure 5.6 and show no variation in the interdigitation gap across the membrane surface, confirming results in Figure 5.4, when no membrane active protein is present.

The spatial distribution of the interdigitation gap parameter in Figure 5.5 suggested that the thickness of the membrane also may not be uniform in the presence of the



Figure 5.5: DMPC-Nogo-66 systems:Spatial distribution of the interdigitated gap parameter averaged over 2 X 2Å grids, perpendicular to the membrane normal, and further averaged over 25 ns(175 to 200 ns) of MD trajectory. The blue patch in SIM1(300K) system shows much reduced values of interdigitated gap parameter values, a signature of strongly interdigitated lipids. This quantitative result matches with the visual image in Figure 5.2(The scale bar units in Å)

membrane active protein and this was monitored and shown in Figure 5.7 (a). The membrane thickness is measured as distance between the P atoms of upper and lower leaflets and averaging over all the lipid molecules and over 25 ns (175-200 ns) of simulations. Both the control system at 310K and DMPC-Nogo-66 system at 310K show a single peak distribution of thickness (average value 34\AA), consistent with uniform distribution of the interdigitated gap parameter as seen in Figure 5.5 (right) and Figure 5.3. The control simulation at 300K also shows almost a single peak distribution of thickness values (average value 37Å), albeit at higher values of average thickness than the previously mentioned systems. However, for systems with DMPC and Nogo-66, the membrane thickness distribution shows a bimodal distribution with two peaks around 28\AA and 35\AA strongly suggesting the simultaneous presence of interdigitated and fluid phases consistent with other analysis. The tilt angle of the lipid acyl chains, measured with respect to membrane normal, also gives an indication of disorder in the lipid tails and is calculated for all the systems simulated and shown in Figure 5.7 (b). The systems where Nogo-66 interacts with DMPC membrane (SIM1(300K), SIM2(300K)), the tilt angle distribution is skewed



Figure 5.6: Lipid interdigitation map consider C314 tail atom done with a resolution of 2Å averaged over 175 to 200ns MD trajectory. The blue patch in SIM1(300K) case shows strong interdigitation which is a signature of ripple phase.(The scale bar units in Å.)

towards much smaller tilt angles, suggesting almost parallel orientations of the lipid chains with respect membrane normal. For all the other systems, the tilt angle distribution is more uniform with an average value of $\sim 30^{\circ}$.

In addition to the tilt angle distribution, the acyl chain order can also be characterized in terms of deuterium order parameter, $S_{CD} = \langle \frac{3cos^2(\theta)-1}{2} \rangle$, where θ is the angle which C-H vector makes with bilayer normal and angle brackets imply the time average of S_{CD} values. The results of S_{CD} calculation are shown in Figure 5.8. The results clearly show that the lipid chains are more ordered for control simulations well below the T_m (290K) and simulations of DMPC membrane in the presence of Nogo-66 at 300K, compared to the membrane-protein system at 310K and control membrane-only systems at 300K and 310K.



Figure 5.7: The distribution of (a) membrane thickness and (b) lipid chain tilt angles for various systems, averaged over 25ns (175-200 ns) of MD trajectory are shown



Figure 5.8: The order parameter plots corresponding to 175 to 200ns MD trajectory. The lipids in SIM1(300K) and SIM2(300K) exhibit higher values on average indicating higher order as a result of ripple formation in these cases.

To probe the correlated motion of lipid molecules in the interdigitated phase, corsscorrelation covariance analysis is performed on the phosphorous atoms in DMPC lipid chains. The covariance matrix was constructed from the displacements with respect to the average structures over 25ns (175-200 ns) of simulation time. The covariance matrix is defined as following:

$$C_{ij} = \frac{\langle \Delta r_i \cdot \Delta r_j \rangle}{\sqrt{\langle \Delta r_i \cdot \Delta r_i \rangle \langle \Delta r_j \cdot \Delta r_j \rangle}}$$
(5.2)

where $\langle \cdot \rangle$ stands for the averaged values, Δr_i and Δr_j are the ith and jth atom's displacements with respect to the corresponding averaged structure atoms. The results of such covariance analysis is shown in Figure 5.9. The membrane system in the presence of nogo-66 (Figure 5.9 (a)) exhibits striking correlated motion among many lipid molecules that most likely belong to interdigitated phase and have C_{ij} values close to 1. On the other hand, the membrane system in the presence of nogo-66 at 310K shows predominantly no correlated motions of lipid head groups, corresponding to fluid phase. Similar analysis for other systems is shown in Figure 5.10. This analysis underscores the dynamical and correlated nature of lipid movements in interdigitated phase.



Figure 5.9: DMPC-Nogo-66 systems: cross correlation covariance maps spanning all the lipid molecules in the system. The higher correlations among the lipid molecules is indicated by values close to 1 and if the lipid molecules are uncorrelated, the values are closer to 0.



Figure 5.10: Averaged cross-correlation map of P atoms from lipid heads over the duration 175 to 200ns of respective molecular dynamics trajectories. Strong red patches which imply strong correlated motions among the lipids is only visible in SIM1 (300K) which exhibits ripple phase.

5.3.2 Interactions of Nogo-66 with DMPC membrane patch

In this section, we probe the interactions of the Nogo-66 with the DMPC membrane to understand the origin of the induced phase change in the membrane system, only in the presence of the protein, at temperatures close to T_m . We compute the z-density profiles (membrane normal along z-axis) of various components of the membrane+protein system to understand their relative positioning. The computed density profiles over 25 ns of simulation time (175-200ns) is shown in Figure 5.11 for both the systems at 300K and 310K. The interdigitation of lipid molecules gives



Figure 5.11: Density profiles corresponding to (a)SIM1(300K) and (b)SIM1(310K) with respect to the z-direction. The DMPC C210 atom density profile(grey curve) is more or less unimodal in (a) and bimodal in (b) which indicates in SIM1(300K) case they overlap in z-dimension, which is because of the interdigitation of upper and lower lipids in the rippled membrane.

rise to the single peak of the membrane acyl chains around z = 0Å, center of the membrane system at 300K, which is in contrast to the exprected double peak seen in the fluid phase at 310K. Regarding the distribution of various functionalvgroups of Nogo-66 along the membrane normal, it can be seen that the overall interaction of Nogo-66 with the membrane is not very strong in terms of partitioning of the protein into the membrane. However, within this framework of membrane-protein interactions, there is a greater tendency of the charged and hydrophobic residues to interact with the headgroup atoms of the DMPC membrane. In the following part of the section, we explore to significance of the such interactions, though transient and weak, on the onset of the interdigitated phase in the membrane.

From the density profiles in Figure 5.11 and from visual inspection of the trajectories, the protein Nogo-66 does not partition into the membrane or interact very strongly like many reported membrane active proteins in literature [231, 232, 233, 234, 235, 236]. In all the systems considered here, the Nogo-66 interacts transiently



Figure 5.12: The time evolution of (a)electrostatic and (b)van der Waals interaction energy between Nogo-66 and DMPC membrane for three instances of simulations at 300K. The time period over which Nogo-66 experiences most favourable electrostatic and van der Waals energies, simultaneously, with the DMPC membrane are highlighted.

with the membrane and in the following sections we attempt to demonstrate that even these transient interactions can induce phase changes into the DMPC membrane, when the temperature is close to (but not lower than) the main transition temperature T_m . The interaction of Nogo-66 with the membrane is monitored via the electrostatic and van der Waals energy between protein and membrane as a function of time and attempts are made to correlate the interaction energy behavior with the onset of interdigitation in the DMPC membrane. For this analysis, we only considered DMPC+Nogo-66 systems at 300K and the time evolution of the interaction energy is shown in Figure 5.12. As mentioned earlier, the interaction of Nogo-66 with the membrane is not very strong and during the MD trajectories, in all three instances, there are time periods, lasting 20-30 ns (as highlighted in the Figure 5.12) when both electrostatic and van der Waals interactions between Nogo-66 and DMPC membrane are very favourable. In all the three simulations, the onset of interdigitation coincided with these very favourable interaction time periods, as indicated in Figure 5.4. We also probed whether the protein has preferntial interactions with one of the phases when interdigitation is present in the membrane systems and the interaction energy of Nogo-66 with interdigitated and fluid phase is computed for one of the systems for the last 25 ns and the results are shown in Figure 5.13. The results strongly suggest that the protein Nogo-66 preferentially interacts with the interdigitated phase over the fluid phase and in this sense it can be considered as a phase-sensing membrane active agent as well. To further probe the effect of Nogo-66 interactions with the DMPC membrane that initiates the interdigitated phases in three repeat simulations, we looked at the aspect of membrane headgroup defects and their potential life times.



Figure 5.13: The interaction energy between the protein and the membrane systems. For systems at 300K, the interaction energy of the protein is partitioned between the fluid and the interdigitated phases. At 310K, the interaction energy between the protein and all the lipids in the fluid phase are computed.

Lipid packing defects have been a topic of much interest lately due to their implication in recognition, anchoring and subsequent partitioning of many amphiphilic



Figure 5.14: The histogram of total "deep defects area" computed throughout the DMPC membrane trajectory.

membrane active agents [237, 224, 238, 204, 239]. Such non-ideal interfacial packing defects of lipid headgroups have been shown to play crucial role in many peptides/proteins which are unstructured in solution but have an ability to adopt stable secondary structures at the water-membrane interface [240, 241, 242, 34]. The dynamical water-membrane interface, subject to many fluctuations including protein interaction, often results in unfavourable exposure of the lipid acyl chain groups to water and such spatial regions of hydrophobic exposure are main constituents of interfacial lipid packing defects. The lipid defects in various membrane systems considered in this study were characterized by using PackMem tool^[228] and the distribution of areas of deep defects (characterizing the exposure of hydrophobic tails located deeper than the position of nearest glycerol atoms) for three systems are shown in Figure 5.14. The results show that for a control system of only DMPC at 300K, the average and spread of total area of deep defects is much smaller than that when Nogo-66 is interacting with DMPC at the same temperature. The distribution of the deep defect area of DMPC+Nogo-66 for both 300K and 310 K is similar suggesting that the nature of the defects is not temperature dependent. The defect constant π in Eq.1. is a parameter that indicates probability of finding larger defects and indirectly estimates the survival of the defects measured. This value is computed for all the three systems shown in Figure 5.14 and the results are in

No.	System	Deep Defect Constant (Å ²)
1	CTRL(300K)	7.79 ± 0.05
2	SIM1(300K)	13.22 ± 4.58
3	SIM1(310K)	9.48 ± 0.81

Table 5.2: Deep defect constants.

Table 5.2. The larger value of π for DMPC+Nogo-66 system at 300K compared to both control system at 300K and membrane-protein system at 310K strongly indicates higher survival of the defects when Nogo-66 is interacting with the DMPC bilayer at 300K, close to transition temperature T_m . We now calculate the interaction between water molecules and exposed acyl chain groups to connect the defect production with the onset of interdigitated phase and the result is shown in Figure 5.15. The results are computed after a selection is made regarding the number of water molecules which are within 4Å of terminal acyl chain groups and if identified, computing the interaction energy between them. Figure 5.15 clearly shows a remarkable correlation between onset of the interdigitated phase and the interaction between water molecules and the exposed terminal group atoms.

5.4 Discussion and Concluding Remarks

Lipid packing defects in the headgroups have been implicated in membrane interface processes including recognition, binding of membrane active agents. Presence of such interfacial defects, which can lead to increased lateral spacing between headgroups and unfavourable exposure of lipid acyl chains to water, can also play crucial role in membrane remodeling including formation of interdigitated phases [243, 244, 245]. Nogo-66 is a protein that exists in an unstructured conformation in solution and has been shown to fold at the membrane interface [2] and the primary focus of this work is to understand the interaction of such membrane active proteins on the possible membrane remodeling including inducing interdigitated phases into the



Figure 5.15: (a)The average interdigitation distance plot and (b)VDW interaction energy plot between lipid tail atoms C314 and C214 with water molecules in 4 Å of those atoms. There is a very strong positive correlation between these interdigitation and interaction plots.

membrane as a function of temperature. We find that when the simulations are performed at temperatures slightly higher than the main transition temperature T_m of the lipids, protein interaction with membranes (even in a non-specific way) has higher probability of inducing interdigitated phases than when the temperatures are much higher than T_m . The simulations of membrane-protein systems at 300K were repeated three times to confirm the results and the interdigitated phases appear in all three simulations (with different initial conditions). We characterize the membrane remodeling and interdigitated phases via various analysis: The binding of Nogo-66 protein on DMPC membrane induces lipid interdigitation at temperatures close to transition temperature of the membrane. Multiple analyses including membrane thickness, tilt angle and deep defect calculations and interaction energies support this observation. The emergence of the interdigitated phase results in a bimodal distribution of membrane thickness, smaller tilt angles and more ordered lipid chains in systems that result in partial interdigitated phases.

In order to understand the specific way in which the interaction of Nogo-66 with DMPC causes the interdigitated phases, we monitored the interaction energies and identified that when Nogo-66 interacts with membrane with most favourable electrostatic and van der Waals interaction for considerable time scale (usually 20-30 ns), the onset of interdigitation follows such events. It is to be noted that Nogo-66 does not bind permanently to the membrane surface nor partitions into the membrane but interacts transiently and somewhat weakly. Such non-specific weak interactions leading to membrane remodeling processes such as membrane thinning and/or interdigitation of membrane lipids has been seen in other disordered protein such as α -Synuclein [34, 245, 246, 247, 248]. α -Synuclein is presynaptic protein whose unfavourable aggregation has been implicated in neurodegenerative diseases is also disordered in solution and adapts an amphiphilic alpha-helical structure near the membrane [249]. In all these systems, a proposed mechanism for membrane remodeling upon membrane-protein interaction includes lipid packing defects. It has been suggested that the membrane active agents can detect the lipid packing defects and interact with the membrane via these defects and membrane in turn remodels itself to accommodate the protein interaction and retain structural integrity [34, 250, 251]. In the present simulations, we have similar strong indications regarding the role of lipid packing defects, in conjunction with temperature, in inducing interdigitated phases in the DMPC membrane upon interaction with Nogo-66.

Various studies in past have demonstrated the capability of molecules like local anes-

thetics, alcohols, si-RNA, membrane-active proteins and charged nanoparticles in modulating the phase transition profile by inducing interdigitated/rippled domains in them [40, 4, 5, 41, 84, 42, 43]. Membrane remodeling processes as a result of lipids interacting with membrane-active peptides/proteins such as domain formation, interdigitation and membrane disruption etc have been reported in numerous studies. Membrane-active peptides like α -Synuclein were reported to induce membrane thinning via interdigitation [34]. The antimicrobial peptide human cathelicidin LL-37 is a membrane-active peptide which was reported to induce interdigitation in negatively charged phosphatidylglycerols (PG which mimics bacterial membrane) and zwitterionic phosphatidylcholines (PC which mimics mammalian membrane) lipids. But LL-37 interacts strongly with PG heads and disrupts this particular membrane^[252]. Another example is the experimental study about intrinsically disordered protein Hsp12 modulating the phase transition profile of negatively charged DMPG lipid membrane^[5], which is in the same direction as our results in this work. The high shift in transition temperature from 298K to 313K for DMPG lipid membrane upon interaction with Hsp12 is perhaps reflective of the fact that PG heads are smaller in size(comparing to PC heads) and is negatively charged.

In many experimental studies involving cell-penetrating peptides and their interaction with the cell membranes, it has been shown that a possible mechanism adopted by the membranes to reduce the peptide effects on its structure is to favor formation of liquid ordered raft-like phases [253, 254, 235]. The peptides in these studies interact very strongly with the membranes. There are also studies, especially in the literature regarding the class of antimicrobial peptides, that strong interactions of peptides can induce significant phase changes in the membrane system at temperatures higher than main transition temperature of the membrane. [255]. In the same study the interaction of a weakly interacting peptide produced minimal changes to the membrane phase. From our simulations, we posit that the weakly interacting membrane active agents can also induce phase changes in the membrane systems, when the temperatures are close to (but higher than) the T_m . The membraneprotein system at 310K does not show any interdigitated phase, likely due to lower survival probabilities of forming large deep defects and also due to self-healing nature of the membrane due to higher thermal energies, which is less likely to occur at lower temperatures. It is noteworthy that the 'lower temperature' simulations in this study are only 10 degrees lower, but that makes the temperature (300K) close to the main transition temperature of DMPC ($T_m = 297.15K$) and the tendency of membrane remodeling via interdigitation may just be enhanced by this fact. Using Nogo-66 as a model membrane active protein, that interacts in a non-specific way with the membranes supporting similar experimental observations of other disordered proteins. Whether presence of other lipid molecules like cholesterol, which can modulate the fluidity of the membranes in which they are present, alters such membrane remodeling upon membrane active protein interaction is a part of future studies.

Bibliography

- Athi N Naganathan. Modulation of allosteric coupling by mutations: from protein dynamics and packing to altered native ensembles and function. *Current opinion in structural biology*, 54:1–9, 2019.
- [2] Sheeja V Vasudevan, Jessica Schulz, Chunyi Zhou, and Melanie J Cocco. Protein folding at the membrane interface, the structure of nogo-66 requires interactions with a phosphocholine surface. *Proceedings of the National Academy* of Sciences, 107(15):6847–6851, 2010.
- [3] Upayan Baul and Satyavani Vemparala. Membrane-bound conformations of antimicrobial agents and their modes of action. In Advances in Planar Lipid Bilayers and Liposomes, volume 22, pages 97–128. Elsevier, 2015.
- [4] Amit Choubey, Ken-ichi Nomura, Rajiv K Kalia, Aiichiro Nakano, and Priya Vashishta. Small interfering ribonucleic acid induces liquid-to-ripple phase transformation in a phospholipid membrane. *Applied Physics Letters*, 105(11):113702, 2014.
- [5] Sylvia Welker, Birgit Rudolph, Elke Frenzel, Franz Hagn, Gerhard Liebisch, Gerd Schmitz, Johannes Scheuring, Andreas Kerth, Alfred Blume, Sevil Weinkauf, et al. Hsp12 is an intrinsically unstructured stress protein that folds upon membrane association and modulates membrane function. *Molecular cell*, 39(4):507–520, 2010.

- [6] Henry R Bourne, David A Sanders, and Frank McCormick. The gtpase superfamily: a conserved switch for diverse cell functions. *Nature*, 348(6297):125– 132, 1990.
- Judith H Raaijmakers and Johannes L Bos. Specificity in ras and rap signaling. Journal of biological chemistry, 284(17):10995–10999, 2009.
- [8] Johannes L Bos, Holger Rehmann, and Alfred Wittinghofer. Gefs and gaps: critical elements in the control of small g proteins. *Cell*, 129(5):865–877, 2007.
- [9] WH Kirsten and LA Mayer. Malignant lymphomas of extrathymic origin induced in rats by murine erythroblastosis virus. *Journal of the National Cancer Institute*, 43(3):735–746, 1969.
- [10] JJ Harvey. An unidentified virus which causes the rapid production of tumours in mice. *Nature*, 204(4963):1104, 1964.
- [11] Ingrid R Vetter and Alfred Wittinghofer. The guanine nucleotide-binding switch in three dimensions. *Science*, 294(5545):1299–1304, 2001.
- [12] Jack McCain. The mapk (erk) pathway: investigational combinations for the treatment of braf-mutated metastatic melanoma. *Pharmacy and Therapeutics*, 38(2):96, 2013.
- [13] Jillian A Parker and Carla Mattos. The ras-membrane interface: isoformspecific differences in the catalytic domain. *Molecular cancer research*, 13(4):595–603, 2015.
- [14] Xiao-Xi Guo, Su An, Yang Yang, Ying Liu, Qian Hao, and Tian-Rui Xu. Rapinteracting proteins are key players in the rap symphony orchestra. *Cellular Physiology and Biochemistry*, 39(1):137–156, 2016.

- [15] Suresh Marada, Ashley Truong, and Stacey K Ogden. The small gtpase rap1 is a modulator of hedgehog signaling. *Developmental biology*, 409(1):84–94, 2016.
- [16] Ketki Sawant, Yujun Chen, Nirupama Kotian, Kevin M Preuss, and Jocelyn A McDonald. Rap1 gtpase promotes coordinated collective cell migration in vivo. *Molecular biology of the cell*, 29(22):2656–2673, 2018.
- [17] Leo S Price, Amra Hajdo-Milasinovic, Jun Zhao, Fried JT Zwartkruis, John G Collard, and Johannes L Bos. Rap1 regulates e-cadherin-mediated cell-cell adhesion. *Journal of Biological Chemistry*, 279(34):35127–35132, 2004.
- [18] Deborah T Leicht, Vitaly Balan, Alexander Kaplun, Vinita Singh-Gupta, Ludmila Kaplun, Melissa Dobson, and Guri Tzivion. Raf kinases: function, regulation and role in human cancer. *Biochimica et Biophysica Acta (BBA)*-*Molecular Cell Research*, 1773(8):1196–1212, 2007.
- [19] Adrienne D Cox, Stephen W Fesik, Alec C Kimmelman, Ji Luo, and Channing J Der. Drugging the undruggable ras: mission possible? *Nature reviews*. *Drug discovery*, 13(11):828, 2014.
- [20] Lilia M Iakoucheva, Predrag Radivojac, Celeste J Brown, Timothy R OConnor, Jason G Sikes, Zoran Obradovic, and A Keith Dunker. The importance of intrinsic disorder for protein phosphorylation. *Nucleic acids research*, 32(3):1037–1049, 2004.
- [21] José L Jiménez, Björn Hegemann, James RA Hutchins, Jan-Michael Peters, and Richard Durbin. A systematic comparative and structural analysis of protein phosphorylation sites based on the mtcptm database. *Genome biology*, 8(5):R90, 2007.
- [22] Arun Kumar Somavarapu, Satish Balakrishnan, Amit Kumar Singh Gautam, David S Palmer, and Prasanna Venkatraman. Structural interrogation of phos-

phoproteome identified by mass spectrometry reveals allowed and disallowed regions of phosphoconformation. *BMC structural biology*, 14(1):9, 2014.

- [23] Michelle Girvan and Mark EJ Newman. Community structure in social and biological networks. Proceedings of the national academy of sciences, 99(12):7821–7826, 2002.
- [24] Anurag Sethi, John Eargle, Alexis A Black, and Zaida Luthey-Schulten. Dynamical networks in trna: protein complexes. Proceedings of the National Academy of Sciences, 106(16):6620–6625, 2009.
- [25] Marieke Kranenburg and Berend Smit. Phase behavior of model lipid bilayers. The Journal of Physical Chemistry B, 109(14):6553–6563, 2005.
- [26] Hermann-Josef Kaiser, Daniel Lingwood, Ilya Levental, Julio L Sampaio, Lucie Kalvodova, Lawrence Rajendran, and Kai Simons. Order of lipid phases in model and plasma membranes. *Proceedings of the National Academy of Sciences*, 106(39):16645–16650, 2009.
- [27] Eric A Smith and Phoebe K Dea. Differential scanning calorimetry studies of phospholipid membranes: the interdigitated gel phase. In Applications of Calorimetry in a Wide Context-Differential Scanning Calorimetry, Isothermal Titration Calorimetry and Microcalorimetry. InTech, 2013.
- [28] Emma I O'Leary, Zhiping Jiang, Marie-Paule Strub, and Jennifer C Lee. Effects of phosphatidylcholine membrane fluidity on the conformation and aggregation of n-terminally acetylated α-synuclein. Journal of Biological Chemistry, 293(28):11195–11205, 2018.
- [29] Avanti Polar Lipids Inc. What is the transition temperature of the lipid? https://avantilipids.com/tech-support/faqs/ transition-temperature. Accessed: 2019-10-16.

- [30] A Tardieu, Vittorio Luzzati, and FC Reman. Structure and polymorphism of the hydrocarbon chains of lipids: a study of lecithin-water phases. *Journal of molecular biology*, 75(4):711–733, 1973.
- [31] TC Lubensky and FC MacKintosh. Theory of ripple phases of lipid bilayers. *Physical review letters*, 71(10):1565, 1993.
- [32] Kai Simons and Derek Toomre. Lipid rafts and signal transduction. Nature reviews Molecular cell biology, 1(1):31, 2000.
- [33] Geert Van Den Bogaart, Jeanette Velásquez Guzmán, Jacek T Mika, and Bert Poolman. On the mechanism of pore formation by melittin. *Journal of Biological Chemistry*, 283(49):33854–33857, 2008.
- [34] Myriam M Ouberai, Juan Wang, Marcus J Swann, Celine Galvagnion, Tim Guilliams, Christopher M Dobson, and Mark E Welland. α-synuclein senses lipid packing defects and induces lateral expansion of lipids leading to membrane remodeling. Journal of Biological Chemistry, pages jbc–M113, 2013.
- [35] Bhagyashree D Rao, Sandeep Shrivastava, and Amitabha Chattopadhyay. Hydrophobic mismatch in membranes: when the tail matters. In *Membrane Organization and Dynamics*, pages 375–387. Springer, 2017.
- [36] Caroline Dart. Symposium review: Lipid microdomains and the regulation of ion channel function. *The Journal of physiology*, 588(17):3169–3178, 2010.
- [37] Tadzia GrandPré, Fumio Nakamura, Timothy Vartanian, and Stephen M Strittmatter. Identification of the nogo inhibitor of axon regeneration as a reticulon protein. *Nature*, 403(6768):439, 2000.
- [38] Alyson E Fournier, Tadzia GrandPré, Graham Gould, Xingxing Wang, and Stephen M Strittmatter. Nogo and the nogo-66 receptor. In *Progress in brain research*, volume 137, pages 361–369. Elsevier, 2002.

- [39] Gia K Voeltz, William A Prinz, Yoko Shibata, Julia M Rist, and Tom A Rapoport. A class of membrane proteins shaping the tubular endoplasmic reticulum. *Cell*, 124(3):573–586, 2006.
- [40] Elizabeth S Rowe and Janet M Campion. Alcohol induction of interdigitation in distearoylphosphatidylcholine: fluorescence studies of alcohol chain length requirements. *Biophysical journal*, 67(5):1888–1895, 1994.
- [41] João Paulo Prates Ramalho, Paraschiva Gkeka, and Lev Sarkisov. Structure and phase transformations of dppc lipid bilayers in the presence of nanoparticles: insights from coarse-grained molecular dynamics simulations. *Langmuir*, 27(7):3723–3730, 2011.
- [42] Da-Cheng Wang, Theodore F Taraschi, Emanuel Rubin, and Nathan Janes. Configurational entropy is the driving force of ethanol action on membrane architecture. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1145(1):141–148, 1993.
- [43] Thomas J McIntosh, Hainan Lin, Shusen Li, and Ching-hsien Huang. The effect of ethanol on the phase transition temperature and the phase structure of monounsaturated phosphatidylcholines. *Biochimica et Biophysica Acta* (BBA)-Biomembranes, 1510(1):219–230, 2001.
- [44] Shaoyong Lu, Hyunbum Jang, Shuo Gu, Jian Zhang, and Ruth Nussinov. Drugging ras gtpase: a comprehensive mechanistic and signaling structural view. *Chemical Society Reviews*, 45(18):4929–4952, 2016.
- [45] Antoine E Karnoub and Robert A Weinberg. Ras oncogenes: split personalities. Nature reviews Molecular cell biology, 9(7):517, 2008.
- [46] Won Do Heo and Tobias Meyer. Switch-of-function mutants based on morphology classification of ras superfamily small gtpases. *Cell*, 113(3):315–328, 2003.

- [47] Åsa Fransson, Aino Ruusala, and Pontus Aspenström. Atypical rho gtpases have roles in mitochondrial homeostasis and apoptosis. *Journal of Biological Chemistry*, 278(8):6495–6502, 2003.
- [48] Michael V Milburn, Liang Tong, Abraham M DeVos, Adel Brunger, Ziro Yamaizumi, Susumu Nishimura, and Sung-Hou Kim. Molecular switch for signal transduction: structural differences between active and inactive forms of protooncogenic ras proteins. *Science*, 247(4945):939–945, 1990.
- [49] Barry J Grant, Alemayehu A Gorfe, and J Andrew McCammon. Ras conformational switching: simulating nucleotide-dependent conformational transitions with accelerated molecular dynamics. *PLoS Comput Biol*, 5(3):e1000325, 2009.
- [50] Michael Spoerner, Christian Herrmann, Ingrid R Vetter, Hans Robert Kalbitzer, and Alfred Wittinghofer. Dynamic properties of the ras switch i region and its importance for binding to effectors. *Proceedings of the National Academy of Sciences*, 98(9):4944–4949, 2001.
- [51] Chigusa Kobayashi and Shinji Saito. Relation between the conformational heterogeneity and reaction cycle of ras: molecular simulation of ras. *Biophysical journal*, 99(11):3726–3734, 2010.
- [52] Ian A Prior, Paul D Lewis, and Carla Mattos. A comprehensive survey of ras mutations in cancer. *Cancer research*, 72(10):2457–2467, 2012.
- [53] Alfred Wittinghofer and Emil F Pal. The structure of ras protein: a model for a universal molecular switch. *Trends in biochemical sciences*, 16:382–387, 1991.
- [54] Brian E Hall, Shao Song Yang, P Ann Boriack-Sjodin, John Kuriyan, and Dafna Bar-Sagi. Structure-based mutagenesis reveals distinct functions for ras

switch 1 and switch 2 in sos-catalyzed guanine nucleotide exchange. *Journal* of *Biological Chemistry*, 276(29):27629–27637, 2001.

- [55] Nina van den Berghe, Robbert H Cool, and Alfred Wittinghofer. Discriminatory residues in ras and rap for guanine nucleotide exchange factor recognition. *Journal of Biological Chemistry*, 274(16):11078–11085, 1999.
- [56] Kendall D Carey, Robert T Watson, Jeffrey E Pessin, and Philip JS Stork. The requirement of specific membrane domains for raf-1 phosphorylation and activation. *Journal of Biological Chemistry*, 278(5):3185–3196, 2003.
- [57] Takashi Fujita, Toru Meguro, Ryo Fukuyama, Hiromichi Nakamuta, and Masao Koida. New signaling pathway for parathyroid hormone and cyclic amp action on extracellular-regulated kinase and cell proliferation in bone cells checkpoint of modulation by cyclic amp. *Journal of Biological Chemistry*, 277(25):22191–22200, 2002.
- [58] Chi V Dang, E Premkumar Reddy, Kevan M Shokat, and Laura Soucek. Drugging the'undruggable'cancer targets. *Nature Reviews Cancer*, 17(8):502, 2017.
- [59] Dirk Kessler, Michael Gmachl, Andreas Mantoulidis, Laetitia J Martin, Andreas Zoephel, Moriz Mayer, Andreas Gollner, David Covini, Silke Fischer, Thomas Gerstberger, et al. Drugging an undruggable pocket on kras. Proceedings of the National Academy of Sciences, 116(32):15823–15829, 2019.
- [60] A Harvey Millar, Joshua L Heazlewood, Carmela Giglione, Michael J Holdsworth, Andreas Bachmair, and Waltraud X Schulze. The scope, functions, and dynamics of posttranslational protein modifications. *Annual review* of plant biology, 70:119–151, 2019.
- [61] Charles S Rubin and Ora M Rosen. Protein phosphorylation. Annual review of biochemistry, 44(1):831–887, 1975.

- [62] Fuxiao Xin and Predrag Radivojac. Post-translational modifications induce significant yet not extreme changes to protein structure. *Bioinformatics*, 28(22):2905–2913, 2012.
- [63] Yassmine Chebaro, Ismail Amal, Natacha Rochel, Cécile Rochette-Egly, Roland H Stote, and Annick Dejaegere. Phosphorylation of the retinoic acid receptor alpha induces a mechanical allosteric regulation and changes in internal dynamics. *PLoS computational biology*, 9(4):e1003012, 2013.
- [64] Weiwei Han, Jingxuan Zhu, Song Wang, and Dong Xu. Understanding the phosphorylation mechanism by using quantum chemical calculations and molecular dynamics simulations. *The Journal of Physical Chemistry B*, 121(15):3565–3573, 2016.
- [65] Dmitry A Karasev, Darya A Veselova, Alexander V Veselovsky, Boris N Sobolev, Victor G Zgoda, and Alexander I Archakov. Spatial features of proteins related to their phosphorylation and associated structural changes. *Proteins: Structure, Function, and Bioinformatics*, 86(1):13–20, 2018.
- [66] Jeffrey R Johnson, Silvia D Santos, Tasha Johnson, Ursula Pieper, Marta Strumillo, Omar Wagih, Andrej Sali, Nevan J Krogan, and Pedro Beltrao. Prediction of functionally important phospho-regulatory events in xenopus laevis oocytes. *PLoS computational biology*, 11(8):e1004362, 2015.
- [67] Kirti Sharma, Rochelle CJ DSouza, Stefka Tyanova, Christoph Schaab, Jacek R Wiśniewski, Jürgen Cox, and Matthias Mann. Ultradeep human phosphoproteome reveals a distinct regulatory nature of tyr and ser/thr-based signaling. *Cell reports*, 8(5):1583–1594, 2014.
- [68] Federica Sgolastra, Brittany M Deronde, Joel M Sarapas, Abhigyan Som, and Gregory N Tew. Designing mimics of membrane active proteins. Accounts of chemical research, 46(12):2977–2987, 2013.
- [69] Fatma Avci, Berna Sariyar Akbulut, and Elif Ozkirimli. Membrane active peptides and their biophysical characterization. *Biomolecules*, 8(3):77, 2018.
- [70] Thomas Oertle, Marjan E van der Haar, Christine E Bandtlow, Anna Robeva, Patricia Burfeind, Armin Buss, Andrea B Huber, Marjo Simonen, Lisa Schnell, Christian Brösamle, et al. Nogo-a inhibits neurite outgrowth and cell spreading with three discrete regions. *Journal of Neuroscience*, 23(13):5393–5406, 2003.
- [71] Samuel David and Albert J Aguayo. Axonal elongation into peripheral nervous system" bridges" after central nervous system injury in adult rats. *Science*, 214(4523):931–933, 1981.
- [72] ME Schwab and H Thoenen. Dissociated neurons regenerate into sciatic but not optic nerve explants in culture irrespective of neurotrophic factors. *Journal* of Neuroscience, 5(9):2415–2423, 1985.
- [73] Martin E Schwab and Pica Caroni. Oligodendrocytes and cns myelin are nonpermissive substrates for neurite growth and fibroblast spreading in vitro. *Journal of Neuroscience*, 8(7):2381–2393, 1988.
- [74] D Hunt, RS Coffin, and PN Anderson. The nogo receptor, its ligands and axonal regeneration in the spinal cord; a review. *Journal of neurocytology*, 31(2):93–120, 2002.
- [75] Ali Alhoshani, Rosemarie Vithayathil, Jonathan Bandong, Katherine M Chrunyk, Gabriel O Moreno, Gregory A Weiss, and Melanie J Cocco. Glutamate provides a key structural contact between reticulon-4 (nogo-66) and phosphocholine. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1838(9):2350–2356, 2014.
- [76] Rumiana Koynova. Lipid phases eye view to lipofection. cationic phosphatidylcholine derivatives as efficient dna carriers for gene delivery. *Lipid Insights*, 2:LPI–S864, 2008.

- [77] Sandrine Morandat, Slim Azouzi, Estelle Beauvais, Amira Mastouri, and Karim El Kirat. Atomic force microscopy of model lipid membranes. Analytical and bioanalytical chemistry, 405(5):1445–1461, 2013.
- [78] Leonor Saiz and Michael L Klein. Computer simulation studies of model biological membranes. Accounts of chemical research, 35(6):482–489, 2002.
- [79] Siewert J Marrink, Jelger Risselada, and Alan E Mark. Simulation of gel phase formation and melting in lipid bilayers using a coarse grained model. *Chemistry and physics of lipids*, 135(2):223–244, 2005.
- [80] Thomas Heimburg. Thermal biophysics of membranes. John Wiley & Sons, 2008.
- [81] Harden M McConnell and Roger D Kornberg. Inside-outside transitions of phospholipids in vesicle membranes. *Biochemistry*, 10(7):1111–1120, 1971.
- [82] Giulia Parisio, Maria Maddalena Sperotto, and Alberta Ferrarini. Flip-flop of steroids in phospholipid bilayers: effects of the chemical structure on transbilayer diffusion. Journal of the American Chemical Society, 134(29):12198– 12208, 2012.
- [83] Justyna B Startek, Brett Boonen, Karel Talavera, and Victor Meseguer. Trp channels as sensors of chemically-induced changes in cell membrane mechanical properties. *International journal of molecular sciences*, 20(2):371, 2019.
- [84] S Thirupathi Reddy, Sandeep Shrivastava, and Amitabha Chattopadhyay. Local anesthetics induce interdigitation and thermotropic changes in dipalmitoylphosphatidylcholine bilayers. *Chemistry and physics of lipids*, 210:22–27, 2018.
- [85] James C. Phillips, Rosemary Braun, Wei Wang, James Gumbart, Emad Tajkhorshid, Elizabeth Villa, Christophe Chipot, Robert D. Skeel, Laxmikant

Kal, and Klaus Schulten. Scalable molecular dynamics with namd. *Journal of Computational Chemistry*, 26(16):1781–1802, 2005.

- [86] Herman JC Berendsen, David van der Spoel, and Rudi van Drunen. Gromacs: a message-passing parallel molecular dynamics implementation. Computer physics communications, 91(1-3):43–56, 1995.
- [87] David A Case, Thomas E Cheatham III, Tom Darden, Holger Gohlke, Ray Luo, Kenneth M Merz Jr, Alexey Onufriev, Carlos Simmerling, Bing Wang, and Robert J Woods. The amber biomolecular simulation programs. *Journal* of computational chemistry, 26(16):1668–1688, 2005.
- [88] Steve Plimpton. Fast parallel algorithms for short-range molecular dynamics. Journal of computational physics, 117(1):1–19, 1995.
- [89] Bernard R Brooks, Robert E Bruccoleri, Barry D Olafson, David J States, S a Swaminathan, and Martin Karplus. Charmm: a program for macromolecular energy, minimization, and dynamics calculations. *Journal of computational chemistry*, 4(2):187–217, 1983.
- [90] Bernard R Brooks, Charles L Brooks III, Alexander D Mackerell Jr, Lennart Nilsson, Robert J Petrella, Benoît Roux, Youngdo Won, Georgios Archontis, Christian Bartels, Stefan Boresch, et al. Charmm: the biomolecular simulation program. Journal of computational chemistry, 30(10):1545–1614, 2009.
- [91] Paul K Weiner and Peter A Kollman. Amber: Assisted model building with energy refinement. a general program for modeling molecules and their interactions. Journal of Computational Chemistry, 2(3):287–303, 1981.
- [92] William L Jorgensen and Julian Tirado-Rives. The opls force field for proteins. energy minimizations for crystals of cyclic peptides and crambin. J. Am. Chem. Soc, 110(6):1657–1723, 1988.

- [93] William L Jorgensen, David S Maxwell, and Julian Tirado-Rives. Development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*, 118(45):11225–11236, 1996.
- [94] Jacob D Durrant and J Andrew McCammon. Molecular dynamics simulations and drug discovery. BMC biology, 9(1):71, 2011.
- [95] Dezső Boda and Douglas Henderson. The effects of deviations from lorentz– berthelot rules on the properties of a simple mixture. *Molecular Physics*, 106(20):2367–2370, 2008.
- [96] Tom Darden, Darrin York, and Lee Pedersen. Particle mesh ewald: An n·log(n)methodforewaldsumsinlargesystems.TheJournalofchemicalphysics, 98(12): 10089 - -10092, 1993.
- [97] David J Hardy, Zhe Wu, James C Phillips, John E Stone, Robert D Skeel, and Klaus Schulten. Multilevel summation method for electrostatic force evaluation. Journal of chemical theory and computation, 11(2):766–779, 2015.
- [98] Leslie Greengard and Vladimir Rokhlin. A new version of the fast multipole method for the laplace equation in three dimensions. Acta numerica, 6:229– 269, 1997.
- [99] Helen M Berman, John Westbrook, Zukang Feng, Gary Gilliland, Talapady N Bhat, Helge Weissig, Ilya N Shindyalov, and Philip E Bourne. The protein data bank. *Nucleic acids research*, 28(1):235–242, 2000.
- [100] Sunhwan Jo, Taehoon Kim, Vidyashankara G Iyer, and Wonpil Im. Charmmgui: a web-based graphical user interface for charmm. *Journal of computational chemistry*, 29(11):1859–1865, 2008.

- [101] Emilia L Wu, Xi Cheng, Sunhwan Jo, Huan Rui, Kevin C Song, Eder M Dávila-Contreras, Yifei Qi, Jumin Lee, Viviana Monje-Galvan, Richard M Venable, et al. Charmm-gui membrane builder toward realistic biological membrane simulations. *Journal of computational chemistry*, 35(27):1997–2004, 2014.
- [102] R Leach Andrew. Molecular modeling principles and applications. 2nd, editor.: Pearson Education Limited, 2001.
- [103] Glenn J Martyna, Michael L Klein, and Mark Tuckerman. Nosé-hoover chains: The canonical ensemble via continuous dynamics. The Journal of chemical physics, 97(4):2635–2643, 1992.
- [104] John Eargle and Zaida Luthey-Schulten. Networkview: 3d display and analysis of protein. rna interaction networks. *Bioinformatics*, 28(22):3000–3001, 2012.
- [105] S. Genheden and U. Ryde. The mm/pbsa and mm/gbsa methods to estimate ligand-binding affinities. *Expert opinion on drug discovery*, 10, 2015.
- [106] Ariela Vergara-Jaque, Jeffrey Comer, Luis Monsalve, Fernando D Gonzalez-Nilo, and Claudia Sandoval. Computationally efficient methodology for atomic-level characterization of dendrimer-drug complexes: a comparison of amine-and acetyl-terminated pamam. *The Journal of Physical Chemistry B*, 117(22):6801–6813, 2013.
- [107] Tingjun Hou, Junmei Wang, Youyong Li, and Wei Wang. Assessing the performance of the mm/pbsa and mm/gbsa methods. 1. the accuracy of binding free energy calculations based on molecular dynamics simulations. *Journal of chemical information and modeling*, 51(1):69–82, 2010.
- [108] Peter A Kollman, Irina Massova, Carolina Reyes, Bernd Kuhn, Shuanghong Huo, Lillian Chong, Matthew Lee, Taisung Lee, Yong Duan, Wei Wang, et al.

Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Accounts of chemical research*, 33(12):889–897, 2000.

- [109] Michel A Cuendet and Olivier Michielin. Protein-protein interaction investigated by steered molecular dynamics: the tcr-pmhc complex. *Biophysical journal*, 95(8):3575–3590, 2008.
- [110] Ken A Dill. Additivity principles in biochemistry. Journal of Biological Chemistry, 272(2):701–704, 1997.
- [111] S. Genheden, O. Kuhn, P. Mikulskis, D. Hoffmann, and U. Ryde. The normalmode entropy in the mm/gbsa method: effect of system truncation, buffer region, and dielectric constant. *Journal of chemical information and modeling*, 52, 2012.
- [112] N. Foloppe and R. Hubbard. Towards predictive ligand design with free-energy based computational methods? *Current medicinal chemistry*, 13, 2006.
- [113] J. Wang, T. Hou, and X. Xu. Recent advances in free energy calculations with a combination of molecular mechanics and continuum models. *Current Computer-Aided Drug Design*, 2, 2006.
- [114] Nadine Homeyer and Holger Gohlke. Free energy calculations by the molecular mechanics poisson- boltzmann surface area method. *Molecular Informatics*, 31(2):114–122, 2012.
- [115] T. Yang. Virtual screening using molecular simulations. Proteins: Structure, Function, and Bioinformatics, 79, 2011.
- [116] P. Das, A. R. Chacko, and G. Belfort. Alzheimer's protective cross-interaction between wild-type and a2t variants alters ab42 dimer structure. ACS chemical neuroscience, 8, 2016.

- [117] Ronald D Gorham Jr, Wilson Rodriguez, and Dimitrios Morikis. Molecular analysis of the interaction between staphylococcal virulence factor sbi-iv and complement c3d. *Biophysical journal*, 106(5):1164–1173, 2014.
- [118] L. Zhang. Probing immobilization mechanism of alpha-chymotrypsin onto carbon nanotube in organic media by molecular dynamics simulation. *Scientific reports*, 5, 2015.
- [119] EW Frische and FJT Zwartkruis. Rap1, a mercenary among the ras-like gtpases. *Developmental biology*, 340(1):1–9, 2010.
- [120] Adrienne D Cox and Channing J Der. Ras history: The saga continues. Small GTPases, 1(1):2–27, 2010.
- [121] Daniel L Altschuler, Scott N Peterson, Michael C Ostrowski, and Eduardo G Lapetina. Cyclic amp-dependent activation of rap1b. Journal of Biological Chemistry, 270(18):10373–10376, 1995.
- [122] Fahad GB Alanazi, Badi Q Alenazi, Abdulaziz D Al-Faim, Omar Bagadir, Ali Al-Shangiti, Omar Kujan, Bassel Tarakji, Saeed M Daboor, Mohamed W Al-Rabea, Waleed Tamimi, et al. Ras signaling pathway, historical view. *Journal* of Cell Science & Therapy, 5(4):1, 2014.
- [123] Julian Downward. Targeting ras signalling pathways in cancer therapy. Nature Reviews Cancer, 3(1):11–22, 2003.
- [124] Shaoyong Lu, Hyunbum Jang, Serena Muratcioglu, Attila Gursoy, Ozlem Keskin, Ruth Nussinov, and Jian Zhang. Ras conformational ensembles, allostery, and signaling. *Chemical reviews*, 116(11):6607–6665, 2016.
- [125] Julian R Molina and Alex A Adjei. The ras/raf/mapk pathway. Journal of Thoracic Oncology, 1(1):7–9, 2006.

- [126] Esther Castellano and Eugenio Santos. Functional specificity of ras isoforms so similar but so different. *Genes & cancer*, 2(3):216–231, 2011.
- [127] Simon J Cook, Bonnee Rubinfeld, Iris Albert, and Frank McCormick. Rapv12 antagonizes ras-dependent activation of erk1 and erk2 by lpa and egf in rat-1 fibroblasts. *The EMBO journal*, 12(9):3475, 1993.
- [128] Stephen R Sprang. How ras works: structure of a rap-raf complex. Structure, 3(7):641-643, 1995.
- [129] Johannes L Bos. Ras-like gtpases. Biochimica et Biophysica Acta (BBA)-Reviews on Cancer, 1333(2):M19–M31, 1997.
- [130] Pablo Rodriguez-Viciana, Celine Sabatier, and Frank McCormick. Signaling specificity by ras family gtpases is determined by the full spectrum of effectors they regulate. *Molecular and cellular biology*, 24(11):4943–4954, 2004.
- [131] Vassiliki A Boussiotis, Gordon J Freeman, Alla Berezovskaya, Dwayne L Barber, and Lee M Nadler. Maintenance of human t cell anergy: blocking of il-2 gene transcription by activated rap1. *Science*, 278(5335):124–128, 1997.
- [132] Hitoshi Kitayama, Yoshikazu Sugimoto, Tomoko Matsuzaki, Yoji Ikawa, and Makoto Noda. A ras-related gene with transformation suppressor activity. *Cell*, 56(1):77–84, 1989.
- [133] Toshihisa Ohtsuka, Kazuya Shimizu, Bunpei Yamamori, Shinya Kuroda, and Yoshimi Takai. Activation of brain b-raf protein kinase by rap1b small gtpbinding protein. *Journal of Biological Chemistry*, 271(3):1258–1261, 1996.
- [134] Tomoyo Okada, Chang-Deng Hu, Tai-Guang Jin, Ken-ichi Kariya, Yuriko Yamawaki-Kataoka, and Tohru Kataoka. The strength of interaction at the raf cysteine-rich domain is a critical determinant of response of raf to ras family small gtpases. *Molecular and cellular biology*, 19(9):6057–6064, 1999.

- [135] Chang-Deng Hu, Ken-ichi Kariya, George Kotani, Mikako Shirouzu, Shigeyuki Yokoyama, and Tohru Kataoka. Coassociation of rap1a and ha-ras with raf-1 n-terminal region interferes with ras-dependent activation of raf-1. *Journal of Biological Chemistry*, 272(18):11702–11705, 1997.
- [136] Johannes L Bos, Johan de Rooij, and Kris A Reedquist. Rap1 signalling: adhering to new models. Nature reviews Molecular cell biology, 2(5):369–377, 2001.
- [137] Matthijs RH Kooistra, Nadia Dubé, and Johannes L Bos. Rap1: a key regulator in cell-cell junction formation. *Journal of cell science*, 120(1):17–22, 2007.
- [138] Johannes L Bos. Linking rap to cell adhesion. Current opinion in cell biology, 17(2):123–128, 2005.
- [139] Martijn Gloerich and Johannes L Bos. Regulating rap small g-proteins in time and space. *Trends in cell biology*, 21(10):615–623, 2011.
- [140] Pamela Y Ting, Christian W Johnson, Cong Fang, Xiaoqing Cao, Thomas G Graeber, Carla Mattos, and John Colicelli. Tyrosine phosphorylation of ras by abl allosterically enhances effector binding. *The FASEB Journal*, 29(9):3750– 3761, 2015.
- [141] C Frémin, JP Guégan, C Plutoni, J Mahaffey, MR Philips, G Emery, and S Meloche. Erk1/2-induced phosphorylation of r-ras gtpases stimulates their oncogenic potential. *Oncogene*, 35(43):5692–5698, 2016.
- [142] Naji Sahyoun, O Bradley McDonald, Francis Farrell, and Eduardo G Lapetina. Phosphorylation of a ras-related gtp-binding protein, rap-1b, by a neuronal ca2+/calmodulin-dependent protein kinase, cam kinase gr. Proceedings of the National Academy of Sciences, 88(7):2643-2647, 1991.

- [143] Severa Bunda, Pardeep Heir, Tharan Srikumar, Jonathan D Cook, Kelly Burrell, Yoshihito Kano, Jeffrey E Lee, Gelareh Zadeh, Brian Raught, and Michael Ohh. Src promotes gtpase activity of ras via tyrosine 32 phosphorylation. Proceedings of the National Academy of Sciences, 111(36):E3785–E3794, 2014.
- [144] X. Zhang. Cyclase-associated protein 1 (cap1) is a prenyl-binding partner of rap1 gtpase. Journal of Biological Chemistry, 293, 2018.
- [145] George A Khoury, Richard C Baliban, and Christodoulos A Floudas. Proteome-wide post-translational modification statistics: frequency analysis and curation of the swiss-prot database. *Scientific reports*, 1, 2011.
- [146] Jonathan D Graves and Edwin G Krebs. Protein phosphorylation and signal transduction. *Pharmacology & therapeutics*, 82(2):111–121, 1999.
- [147] S. Matsunaga. Structural transition of solvated h-ras/gtp revealed by molecular dynamics simulation and local network entropy. J Mol Graph Model, 77, 2017.
- [148] S. Lukman, D. P. Lane, and C. S. Verma. Mapping the structural and dynamical features of multiple p53 dna binding domains: Insights into loop 1 intrinsic dynamics. *PloS One*, 8, 2013.
- [149] A. Karshikoff, L. Nilsson, and R. Ladenstein. Rigidity versus flexibility: the dilemma of understanding protein thermal stability. *FEBS J*, 282, 2015.
- [150] S. Vemparala, S. Mehrotra, and H. Balaram. Role of loop dynamics in thermal stability of mesophilic and thermophilic adenylosuccinate synthetase: a molecular dynamics and normal mode analysis study. BBA, 1814, 2011.
- [151] L Michel Espinoza-Fonseca, David Kast, and David D Thomas. Molecular dynamics simulations reveal a disorder-to-order transition on phosphorylation of smooth muscle myosin. *Biophysical journal*, 93(6):2083–2090, 2007.

- [152] Albert J Lyons, Neha S Gandhi, and Ricardo L Mancera. Molecular dynamics simulation of the phosphorylation-induced conformational changes of a tau peptide fragment. *Proteins: Structure, Function, and Bioinformatics*, 82(9):1907–1923, 2014.
- [153] Yasushige Yonezawa. Molecular dynamics study of the phosphorylation effect on the conformational states of the c-terminal domain of rna polymerase ii. *The Journal of Physical Chemistry B*, 118(17):4471–4478, 2014.
- [154] Juan A Bueren-Calabuig and Julien Michel. Impact of ser17 phosphorylation on the conformational dynamics of the oncoprotein mdm2. *Biochemistry*, 55(17):2500–2509, 2016.
- [155] Jason L Smart and J Andrew McCammon. Phosphorylation stabilizes the n-termini of α-helices. *Biopolymers*, 49(3):225–233, 1999.
- [156] Hiroki Noguchi, Takahisa Ikegami, Aritaka Nagadoi, Yuji O Kamatari, Sam-Yong Park, Jeremy RH Tame, and Satoru Unzai. The structure and conformational switching of rap1b. *Biochemical and biophysical research communications*, 462(1):46–51, 2015.
- [157] Devin K Schweppe, James R Rigas, and Scott A Gerber. Quantitative phosphoproteomic profiling of human non-small cell lung cancer tumors. *Journal* of proteomics, 91:286–296, 2013.
- [158] Y.-.. H. Wong. Kinasephos 2.0: a web server for identifying protein kinasespecific phosphorylation sites based on sequences and coupling patterns. Nucleic acids research, 35, 2007.
- [159] Nicolas Nassar, Gudrun Horn, Christian Herrmann, Anna Scherer, et al. The 2.2 angstrom crystal structure of the ras-binding domain of the serine/threonine kinase c-raf1 in complex with rap1a and a gtp analogue. Nature, 375(6532):554, 1995.

- [160] Alexander D MacKerell Jr, Donald Bashford, MLDR Bellott, Roland Leslie Dunbrack Jr, Jeffrey D Evanseck, Martin J Field, Stefan Fischer, Jiali Gao, H Guo, Sookhee Ha, et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *The journal of physical chemistry B*, 102(18):3586–3616, 1998.
- [161] Robert B Best, Xiao Zhu, Jihyun Shim, Pedro EM Lopes, Jeetain Mittal, Michael Feig, and Alexander D MacKerell Jr. Optimization of the additive charmm all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain χ1 and χ2 dihedral angles. Journal of chemical theory and computation, 8(9):3257–3273, 2012.
- [162] William Humphrey, Andrew Dalke, and Klaus Schulten. VMD Visual Molecular Dynamics. Journal of Molecular Graphics, 14:33–38, 1996.
- [163] Panagiotis I Koukos and Nicholas M Glykos. Grcarma: A fully automated task-oriented interface for the analysis of molecular dynamics trajectories. *Journal of computational chemistry*, 34(26):2310–2312, 2013.
- [164] Todd J Dolinsky, Jens E Nielsen, J Andrew McCammon, and Nathan A Baker. Pdb2pqr: an automated pipeline for the setup of poisson–boltzmann electrostatics calculations. *Nucleic acids research*, 32(suppl 2):W665–W667, 2004.
- [165] William L Jorgensen, Jayaraman Chandrasekhar, Jeffry D Madura, Roger W Impey, and Michael L Klein. Comparison of simple potential functions for simulating liquid water. *The Journal of chemical physics*, 79(2):926–935, 1983.
- [166] Glenn J Martyna, Douglas J Tobias, and Michael L Klein. Constant pressure molecular dynamics algorithms. The Journal of Chemical Physics, 101(5):4177–4189, 1994.

- [167] Scott E Feller, Yuhong Zhang, Richard W Pastor, and Bernard R Brooks. Constant pressure molecular dynamics simulation: the langevin piston method. *The Journal of chemical physics*, 103(11):4613–4621, 1995.
- [168] Herman JC Berendsen, JPM van Postma, Wilfred F van Gunsteren, ARHJ DiNola, and JR Haak. Molecular dynamics with coupling to an external bath. *The Journal of chemical physics*, 81(8):3684–3690, 1984.
- [169] Ulrich Essmann, Lalith Perera, Max L Berkowitz, Tom Darden, Hsing Lee, and Lee G Pedersen. A smooth particle mesh ewald method. *The Journal of chemical physics*, 103(19):8577–8593, 1995.
- [170] Joe Dundas, Zheng Ouyang, Jeffery Tseng, Andrew Binkowski, Yaron Turpaz, and Jie Liang. Castp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic acids research*, 34(suppl_2):W116–W118, 2006.
- [171] Jean-Bernard Créchet, Alberto Bernardi, and Andrea Parmeggiani. Distal switch ii region of ras2p is required for interaction with guanine nucleotide exchange factor. *Journal of Biological Chemistry*, 271(29):17234–17240, 1996.
- [172] Christian Herrmann, Gudrun Horn, Marcel Spaargaren, and Alfred Wittinghofer. Differential interaction of the ras family gtp-binding proteins h-ras, rap1a, and r-ras with the putative effector molecules raf kinase and ral-guanine nucleotide exchange factor. Journal of Biological Chemistry, 271(12):6794– 6800, 1996.
- [173] Helen R Mott and Darerca Owen. Structures of ras superfamily effector complexes: What have we learnt in two decades? *Critical reviews in biochemistry* and molecular biology, 50(2):85–133, 2015.
- [174] Dhirendra K Simanshu, Dwight V Nissley, and Frank McCormick. Ras proteins and their regulators in human disease. *Cell*, 170(1):17–33, 2017.

- [175] Susan K Fetics, Hugo Guterres, Bradley M Kearney, Greg Buhrman, Buyong Ma, Ruth Nussinov, and Carla Mattos. Allosteric effects of the oncogenic rasq611 mutant on raf-rbd. *Structure*, 23(3):505–516, 2015.
- [176] Andrea Scrima, Christoph Thomas, Delia Deaconescu, and Alfred Wittinghofer. The rap-rapgap complex: Gtp hydrolysis without catalytic glutamine and arginine residues. *The EMBO journal*, 27(7):1145–1153, 2008.
- [177] Greg Buhrman, Genevieve Holzapfel, Susan Fetics, and Carla Mattos. Allosteric modulation of ras positions q61 for a direct role in catalysis. Proceedings of the National Academy of Sciences, 107(11):4931–4936, 2010.
- [178] Jacqueline Cherfils, Julie Ménétrey, Gérard Le Bras, Gisèle Le Bras, Isabelle Janoueix-Lerosey, Jean de Gunzburg, Jean-Renaud Garel, and Isabelle Auzat. Crystal structures of the small g protein rap2a in complex with its substrate gtp, with gdp and with gtpγs. The EMBO Journal, 16(18):5582–5591, 1997.
- [179] Partha P Chakrabarti, Oliver Daumke, Yan Suveyzdis, Carsten Kötting, Klaus Gerwert, and Alfred Wittinghofer. Insight into catalysis of a unique gtpase reaction by a combined biochemical and ftir approach. *Journal of molecular biology*, 367(4):983–995, 2007.
- [180] Shane Atwell, Mark Ultsch, Abraham M De Vos, and James A Wells. Structural plasticity in a remodeled protein-protein interface. *Science*, 278(5340):1125–1128, 1997.
- [181] Yina Gu, Da-Wei Li, and Rafael Brüschweiler. Statistical database analysis of the role of loop dynamics for protein–protein complex formation and allostery. *Bioinformatics*, 33(12):1814–1819, 2017.
- [182] Changhui Yan, Feihong Wu, Robert L Jernigan, Drena Dobbs, and Vasant Honavar. Characterization of protein–protein interfaces. *The protein journal*, 27(1):59–70, 2008.

- [183] C. Kiel, L. Serrano, and C. Herrmann. A detailed thermodynamic analysis of ras/effector complex interfaces. *Journal of molecular biology*, 340, 2004.
- [184] Gabrielle Stetz, Amanda Tse, and Gennady M Verkhivker. Dissecting structure-encoded determinants of allosteric cross-talk between posttranslational modification sites in the hsp90 chaperones. *Scientific reports*, 8(1):6899, 2018.
- [185] Zhijun Luo, Bruce Diaz, Mark S Marshall, and Joseph Avruch. An intact raf zinc finger is required for optimal binding to processed ras and for rasdependent raf activation in situ. *Molecular and cellular biology*, 17(1):46–53, 1997.
- [186] Helen R Mott, John W Carpenter, Sheng Zhong, Sujoy Ghosh, Robert M Bell, and Sharon L Campbell. The solution structure of the raf-1 cysteine-rich domain: a novel ras and phospholipid binding site. *Proceedings of the National Academy of Sciences*, 93(16):8312–8317, 1996.
- [187] Shuai Li, Hyunbum Jang, Jian Zhang, and Ruth Nussinov. Raf-1 cysteine-rich domain increases the affinity of k-ras/raf at the membrane, promoting mapk signaling. *Structure*, 26(3):513–525, 2018.
- [188] Zhen-Lu Li, Priyanka Prakash, and Matthias Buck. A tug of war maintains a dynamic protein-membrane complex: molecular dynamics simulations of craf rbd-crd bound to k-ras4b at an anionic membrane. ACS central science, 4(2):298–305, 2018.
- [189] IG Macara, KM Lounsbury, SA Richards, C a McKiernan, and D Bar-Sagi. The ras superfamily of gtpases. *The FASEB Journal*, 10(5):625–630, 1996.
- [190] Alexandr P Kornev and Susan S Taylor. Dynamics-driven allostery in protein kinases. Trends in biochemical sciences, 40(11):628–647, 2015.

- [191] T Devanand, Prasanna Venkatraman, and Satyavani Vemparala. Phosphorylation promotes binding affinity of rap-raf complex by allosteric modulation of switch loop dynamics. *Scientific reports*, 8(1):12976, 2018.
- [192] Hao Zhang, Yu-Chung Chang, Mark L Brennan, and Jinhua Wu. The structure of rap1 in complex with riam reveals specificity determinants and recruitment mechanism. *Journal of molecular cell biology*, 6(2):128–139, 2013.
- [193] Benjamin Webb and Andrej Sali. Protein structure modeling with modeller. Protein Structure Prediction, pages 1–15, 2014.
- [194] Frank McCormick. Ras biology in atomic detail. Nature structural biology, 3(8):653, 1996.
- [195] Mahendra Kumar Jain and Nora Min Wu. Effect of small molecules on the dipalmitoyl lecithin liposomal bilayer: Iii. phase transition in lipid bilayer. The Journal of Membrane Biology, 34(1):157–201, 1977.
- [196] AG Lee. Lipid-protein interactions in biological membranes: a structural perspective. Biochimica et Biophysica Acta (BBA)-Biomembranes, 1612(1):1– 40, 2003.
- [197] Thomas R Weikl. Membrane-mediated cooperativity of proteins. Annual review of physical chemistry, 69:521–539, 2018.
- [198] Z Vivian Feng, Steve Granick, and Andrew A Gewirth. Modification of a supported lipid bilayer by polyelectrolyte adsorption. *Langmuir*, 20(20):8796– 8804, 2004.
- [199] Veerendra Kumar Sharma and Shuo Qian. Effect of an antimicrobial peptide on lateral segregation of lipids, a structure and dynamics study by neutron scattering. *Langmuir*, 2019.

- [200] Richard M Epand, Chelsea Walker, Raquel F Epand, and Nathan A Magarvey. Molecular mechanisms of membrane targeting antibiotics. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1858(5):980–987, 2016.
- [201] Nicolas Destainville, Thomas H Schmidt, and Thorsten Lang. Where biology meets physics converging view on membrane microdomain dynamics. In *Current topics in membranes*, volume 77, pages 27–65. Elsevier, 2016.
- [202] Siewert J Marrink, Alex H De Vries, and D Peter Tieleman. Lipids on the move: simulations of membrane pores, domains, stalks and curves. *Biochimica* et Biophysica Acta (BBA)-Biomembranes, 1788(1):149–168, 2009.
- [203] Friederike Schmid. Physical mechanisms of micro-and nanodomain formation in multicomponent lipid membranes. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1859(4):509–528, 2017.
- [204] Guillaume Drin and Bruno Antonny. Amphipathic helices and membrane curvature. FEBS letters, 584(9):1840–1847, 2010.
- [205] Alex H de Vries, Serge Yefimov, Alan E Mark, and Siewert J Marrink. Molecular structure of the lecithin ripple phase. *Proceedings of the National Academy* of Sciences, 102(15):5392–5396, 2005.
- [206] Pouyan Khakbaz and Jeffery B Klauda. Investigation of phase transitions of saturated phosphocholine lipid bilayers via molecular dynamics simulations. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1860(8):1489– 1501, 2018.
- [207] Svetlana Baoukina and D Peter Tieleman. Computer simulations of phase separation in lipid bilayers and monolayers. In *Methods in Membrane Lipids*, pages 307–322. Springer, 2015.

- [208] Kiyotaka Akabori and John F Nagle. Structure of the dmpc lipid bilayer ripple phase. Soft Matter, 11(5):918–926, 2015.
- [209] WJ Sun, Stephanie Tristram-Nagle, Robert M Suter, and John F Nagle. Structure of the ripple phase in lecithin bilayers. *Proceedings of the National Academy of Sciences*, 93(14):7008–7012, 1996.
- [210] Olaf Lenz and Friederike Schmid. Structure of symmetric and asymmetric ripple phases in lipid bilayers. *Physical review letters*, 98(5):058104, 2007.
- [211] AR Verde, MB Sierra, LM Alarcón, VI Pedroni, GA Appignanesi, and MA Morini. Experimental and computational studies of the effects of free dha on a model phosphatidylcholine membrane. *Chemistry and Physics of Lipids*, 217:12–18, 2018.
- [212] Ellyn Gray, Joshua Karslake, Benjamin B Machta, and Sarah L Veatch. Liquid general anesthetics lower critical temperatures in plasma membrane vesicles. *Biophysical journal*, 105(12):2751–2759, 2013.
- [213] Monika Laner and Philippe H Hünenberger. Effect of methanol on the phasetransition properties of glycerol-monopalmitate lipid bilayers investigated using molecular dynamics simulations: in quest of the biphasic effect. Journal of Molecular Graphics and Modelling, 55:85–104, 2015.
- [214] Thomas Mavromoustakos, Petros Chatzigeorgiou, Catherine Koukoulitsa, and Serdar Durdagi. Partial interdigitation of lipid bilayers. *International Journal* of Quantum Chemistry, 111(6):1172–1183, 2011.
- [215] Matthew B Stone, Sarah A Shelby, Marcos F Núñez, Kathleen Wisser, and Sarah L Veatch. Protein sorting by lipid phase-like domains supports emergent signaling function in b lymphocyte plasma membranes. *Elife*, 6:e19891, 2017.

- [216] Edward Lyman, Chia-Lung Hsieh, and Christian Eggeling. From dynamics to membrane organization: Experimental breakthroughs occasion a" modeling manifesto". *Biophysical journal*, 2018.
- [217] Daniel Lingwood and Kai Simons. Lipid rafts as a membrane-organizing principle. science, 327(5961):46–50, 2010.
- [218] Philip R Nicovich, Joanna M Kwiatek, Yuanqing Ma, Aleš Benda, and Katharina Gaus. Fscs reveals the complexity of lipid domain dynamics in the plasma membrane of live cells. *Biophysical Journal*, 114(12):2855–2864, 2018.
- [219] MAHENDRA K JAIN and HAROLD B WHITE III. Long-range order in biomembranes. In Advances in lipid research, volume 15, pages 1–60. Elsevier, 1977.
- [220] Michael Edidin. The state of lipid rafts: from model membranes to cells. Annual review of biophysics and biomolecular structure, 32(1):257–283, 2003.
- [221] Marieke Kranenburg, Maddalena Venturoli, and Berend Smit. Phase behavior and induced interdigitation in bilayers studied with dissipative particle dynamics. *The Journal of Physical Chemistry B*, 107(41):11491–11501, 2003.
- [222] TJ McIntosh, RV McDaniel, and SA Simon. Induction of an interdigitated gel phase in fully hydrated phosphatidylcholine bilayers. *Biochimica et Biophysica* Acta (BBA)-Biomembranes, 731(1):109–114, 1983.
- [223] Kathryn L Griffin, Chih-Yin Cheng, Eric A Smith, and Phoebe K Dea. Effects of pentanol isomers on the phase behavior of phospholipid bilayer membranes. *Biophysical chemistry*, 152(1-3):178–183, 2010.
- [224] Stefano Vanni, Lydie Vamparys, Romain Gautier, Guillaume Drin, Catherine Etchebest, Patrick FJ Fuchs, and Bruno Antonny. Amphipathic lipid packing

sensor motifs: probing bilayer defects with hydrophobic residues. *Biophysical journal*, 104(3):575–584, 2013.

- [225] Susan Mabrey and Julian M Sturtevant. Investigation of phase transitions of lipids and lipid mixtures by sensitivity differential scanning calorimetry. *Proceedings of the National Academy of Sciences*, 73(11):3862–3866, 1976.
- [226] Bachir Aoun, Eric Pellegrini, Marcus Trapp, Francesca Natali, Laura Cantù, Paola Brocca, Yuri Gerelli, Bruno Demé, Michael Marek Koza, Mark Johnson, et al. Direct comparison of elastic incoherent neutron scattering experiments with molecular dynamics simulations of dmpc phase transitions. *The European Physical Journal E*, 39(4):48, 2016.
- [227] Ramon Guixa-González, Ismael Rodriguez-Espigares, Juan Manuel Ramírez-Anguita, Pau Carrió-Gaspar, Hector Martinez-Seara, Toni Giorgino, and Jana Selent. Membplugin: studying membrane complexity in vmd. *Bioinformatics*, 30(10):1478–1480, 2014.
- [228] Romain Gautier, Amélie Bacle, Marion L Tiberti, Patrick F Fuchs, Stefano Vanni, and Bruno Antonny. Packmem: A versatile tool to compute and visualize interfacial packing defects in lipid bilayers. *Biophysical journal*, 115(3):436– 444, 2018.
- [229] Lydie Vamparys, Romain Gautier, Stefano Vanni, WF Drew Bennett, D Peter Tieleman, Bruno Antonny, Catherine Etchebest, and Patrick FJ Fuchs. Conical lipids in flat bilayers induce packing defects similar to that induced by positive curvature. *Biophysical journal*, 104(3):585–593, 2013.
- [230] Haosheng Cui, Edward Lyman, and Gregory A Voth. Mechanism of membrane curvature sensing by amphipathic helix containing proteins. *Biophysical jour*nal, 100(5):1271–1279, 2011.

- [231] Ming-Tao Lee, Fang-Yu Chen, and Huey W Huang. Energetics of pore formation induced by membrane active peptides. *Biochemistry*, 43(12):3590–3599, 2004.
- [232] Ming-Tao Lee, Tzu-Lin Sun, Wei-Chin Hung, and Huey W Huang. Process of inducing pores in membranes by melittin. *Proceedings of the National Academy* of Sciences, 110(35):14243–14248, 2013.
- [233] Shuo Qian and William T Heller. Peptide-induced asymmetric distribution of charged lipids in a vesicle bilayer revealed by small-angle neutron scattering. *The Journal of Physical Chemistry B*, 115(32):9831–9837, 2011.
- [234] Chun-Jen Su, Shiuan-Shiaou Wu, U-Ser Jeng, Ming-Tao Lee, An-Chung Su, Kuei-Fen Liao, Wei-Yu Lin, Yu-Shan Huang, and Chun-Yu Chen. Peptideinduced bilayer thinning structure of unilamellar vesicles and the related binding behavior as revealed by x-ray scattering. *Biochimica et Biophysica Acta* (BBA)-Biomembranes, 1828(2):528–534, 2013.
- [235] Claudia Almeida, Antonin Lamazière, Angélique Filleau, Yohann Corvis, Philippe Espeau, and Jesus Ayala-Sanmartin. Membrane re-arrangements and rippled phase stabilisation by the cell penetrating peptide penetratin. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1858(11):2584–2591, 2016.
- [236] Margherita Di Pisa, Gerard Chassaing, and Jean-Marie Swiecicki. Translocation mechanism (s) of cell-penetrating peptides: biophysical studies using artificial membrane bilayers. *Biochemistry*, 54(2):194–207, 2014.
- [237] Stefano Vanni, Hisaaki Hirose, Hélène Barelli, Bruno Antonny, and Romain Gautier. A sub-nanometre view of how membrane curvature and composition modulate lipid packing and protein recruitment. *Nature communications*, 5:4916, 2014.

- [238] Hang Yin and Aaron D Flynn. Drugging membrane protein interactions. Annual review of biomedical engineering, 18:51–76, 2016.
- [239] Upayan Baul and Satyavani Vemparala. Influence of lipid composition of model membranes on methacrylate antimicrobial polymer-membrane interactions. Soft matter, 13(41):7665–7676, 2017.
- [240] Patrick L Harrison, George R Heath, Benjamin RG Johnson, Mohamed A Abdel-Rahman, Peter N Strong, Stephen D Evans, and Keith Miller. Phospholipid dependent mechanism of smp24, an α-helical antimicrobial peptide from scorpion venom. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1858(11):2737–2744, 2016.
- [241] Joëlle Bigay, Jean-François Casella, Guillaume Drin, Bruno Mesmin, and Bruno Antonny. Arfgap1 responds to membrane curvature through the folding of a lipid packing sensor motif. *The EMBO journal*, 24(13):2244–2253, 2005.
- [242] Harvey T McMahon and Jennifer L Gallop. Membrane curvature and mechanisms of dynamic cell membrane remodelling. *Nature*, 438(7068):590, 2005.
- [243] Marieke Kranenburg, Martin Vlaar, and Berend Smit. Simulating induced interdigitation in membranes. *Biophysical journal*, 87(3):1596–1605, 2004.
- [244] Taiki Shigematsu, Kenichiro Koshiyama, and Shigeo Wada. Stretch-induced interdigitation of a phospholipid/cholesterol bilayer. The Journal of Physical Chemistry B, 122(9):2556–2563, 2018.
- [245] James L Slater and Ching-Hsien Huang. Interdigitated bilayer membranes. Progress in lipid research, 27(4):325–359, 1988.
- [246] William T Heller, Alan J Waring, Robert I Lehrer, Thad A Harroun, Thomas M Weiss, Lin Yang, and Huey W Huang. Membrane thinning effect of the β-sheet antimicrobial protegrin. *Biochemistry*, 39(1):139–145, 2000.

- [247] Almut Mecke, Dong-Kuk Lee, Ayyalusamy Ramamoorthy, Bradford G Orr, and Mark M Banaszak Holl. Membrane thinning due to antimicrobial peptide binding: an atomic force microscopy study of msi-78 in lipid bilayers. *Biophysical journal*, 89(6):4043–4050, 2005.
- [248] Nicholas P Reynolds, Alice Soragni, Michael Rabe, Dorinel Verdes, Ennio Liverani, Stephan Handschin, Roland Riek, and Stefan Seeger. Mechanism of membrane interaction and disruption by α-synuclein. Journal of the American Chemical Society, 133(48):19366–19375, 2011.
- [249] Giuliana Fusco, Maximo Sanz-Hernandez, and Alfonso De Simone. Order and disorder in the physiological membrane binding of α-synuclein. Current opinion in structural biology, 48:49–57, 2018.
- [250] Jobin Varkey, Jose Mario Isas, Naoko Mizuno, Martin Borch Jensen, Vikram Kjøller Bhatia, Christine C Jao, Jitka Petrlova, John C Voss, Dimitrios G Stamou, Alasdair C Steven, et al. Membrane curvature induction and tubulation are common features of synucleins and apolipoproteins. *Journal of Biological Chemistry*, 285(42):32486–32493, 2010.
- [251] Naoko Mizuno, Jobin Varkey, Natalie C Kegulian, Balachandra G Hegde, Naiqian Cheng, Ralf Langen, and Alasdair C Steven. Remodeling of lipid vesicles into cylindrical micelles by α-synuclein in an extended α-helical conformation. Journal of Biological Chemistry, 287(35):29301–29311, 2012.
- [252] E Sevcsik, G Pabst, A Jilek, and K Lohner. How lipids influence the mode of action of membrane-active peptides. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 1768(10):2586–2595, 2007.
- [253] Antonin Lamazière, Gérard Chassaing, Germain Trugnan, and Jesus Ayala-Sanmartin. Tubular structures in heterogeneous membranes induced by the

cell penetrating peptide penetratin. Communicative & integrative biology, 2(3):223-224, 2009.

- [254] Antonin Lamazière, Claude Wolf, Olivier Lambert, Gérard Chassaing, Germain Trugnan, and Jesus Ayala-Sanmartin. The homeodomain derived peptide penetratin induces curvature of fluid membrane domains. *PLoS One*, 3(4):e1938, 2008.
- [255] VK Sharma, E Mamontov, M Tyagi, S Qian, DK Rai, and VS Urban. Dynamical and phase behavior of a phospholipid membrane altered by an antimicrobial peptide at low concentration. *The journal of physical chemistry letters*, 7(13):2394–2401, 2016.

Thesis Highlight

Name of the Student: Devanand TName of the CI/OCC: IMSc, ChennaiEnrolment No.: LIFE10201404002Thesis Title: Allosteric effects in protein dynamics and their interactions with membranesDiscipline: Life SciencesSub-Area of Discipline: Computational BiophysicsDate of viva voce: 24/06/2020

Ras superfamily of small GTPase proteins is a major player in many cellular functions like cell proliferation, cytoskeletal morphology, and cellular adhesion, etc. The effects of phosphorylation of a serine residue on the structural and dynamic properties of Ras-like protein, Rap, and its interactions with effector protein Ras binding domain (RBD) of Raf kinase, in the presence of GTP, are investigated via molecular dynamics simulations. The simulations show that phosphorylation significantly effects the dynamics of functional loops of Rap which participate in the stability of the complex with effector proteins. The effects of phosphorylation on Rap are significant and detailed conformational analysis suggest that the Rap protein, when phosphorylated and with GTP ligand, samples different conformational space as compared to non-phosphorylated protein. In addition, phosphorylation of SER11 opens up a new cavity in the Rap protein which can be further explored for possible drug interactions. Residue network analysis shows that the phosphorylation of Rap results in a community spanning both Rap and RBD and strongly suggests transmission of allosteric effects of local alterations in Rap to distal regions of RBD, potentially affecting the downstream signalling. In the figure below, we illustrate the long range interaction at Raf L4 loop as a result of SER11 phosphorylation.



Figure 1 : The communities detected in (a) GTP- (wild type) and (b) GTP-PSER11 (phosphorylated) cases. The yellow community represents the L4 loop.

Binding free energy calculations suggest that phosphorylation of SER11 residue increases the binding between Rap and Raf corroborating the network analysis results. The increased binding of the Rap-Raf complex can have cascading effects along the signalling pathways where availability of Raf can influence the oncogenic effects of Ras proteins. These simulations underscore the importance of post-translational modifications like phosphorylation on the functional dynamics in proteins and can be an alternative to drug-targeting, especially in notoriously undruggable oncoproteins belonging to Ras-like GTPase family.