

Molecular Dynamics Simulations and Computational Analysis of K-Ras4B and c-Raf Interaction

by

Ayşe Seda Yazgılı

A Dissertation Submitted to the
Graduate School of Sciences and Engineering
in Partial Fulfillment of the Requirements for
the Degree of
Master of Science
in
Molecular Biology and Genetics



October, 2018

**Molecular Dynamics Simulations and Computational Analysis of
K-Ras4B and c-Raf Interaction**

Koç University

Graduate School of Sciences and Engineering

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

Ayşe Seda Yazgılı

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

Committee Members:

Prof. Özlem Keskin (Advisor)

Prof. Attila Gürsoy (Co-Advisor)

Prof. Halil Kavaklı

Assist. Prof. Sefer Baday

Assist. Prof Elif Nur Fırat Karalar

Date : _____



“Wenn das Schicksal uns etwas nimmt,

Vertrau der Zeit!

Denn immer, immer wieder geht die Sonne auf,

Und wieder bringt ein Tag für uns sein Licht.

Ja, immer, immer wieder geht die Sonne auf,

Denn Dunkelheit für immer gibt es nicht.”

Udo Jürgens.

ABSTRACT

Ras proteins are activated through their upstream regulators. Upon activation, Ras proteins recruit their downstream effectors to the plasma membrane and activate them. Activated downstream effectors, then activate other proteins and this signaling cascades lead several responses within the cell. Although Ras proteins are activated through guanine exchange factor protein, specific mutations result with constitutively activate Ras. Those mutations are found in many cancer types and several developmental diseases. Therefore, targeting the mutations on Ras has been the main focus in the treatment of Ras-related diseases. However, Ras proteins are still classified as “undruggable protein”. Effects of the mutations on Ras proteins and the interaction dynamics between their effectors are unclear. In this thesis, K-Ras4B and c-Raf interaction studied within two subsystems through molecular dynamics simulations. Computationally, K-Ras4B and c-Raf interaction were predicted by using PRISM (PRotein Interactions by Structural Matching). Residues on Ras and Raf interaction complex were identified using HotRegion. The results gave an idea about the specific amino acids involved in K-Ras4B and c-Raf interaction. To confirm the predictions, R41E/K42D double charge reversal mutations, called B3 mutation, were introduced to the control (G12D). MD simulations were performed on control and B3 mutated subsystems. In the system, which contains K-Ras4B₁₋₁₆₆, the mutation slightly changed the binding free energy levels. According to the result, B3 mutation decreased the binding free energy levels. However, the difference in the energy levels was not significant enough. The unexpected energy levels can be studied further by introducing a membrane and a cysteine rich domain (CRD) to the system.

ÖZETÇE

Ras proteinleri üst yolaklarında bulunan efektör proteinleri tarafından aktive edilirler. Aktivasyonları üzerine Ras proteinleri hücre zarında toplanır ve alt yolaklarında bulunan efektör proteinleri aktive ederler. Aktifleştirilmiş alt yolaklarda bulunan Ras efektörleri, daha sonra diğer proteinleri indükler ve bu sinyal kaskadları, hücre içindeki çeşitli tepkilere yol açar. Ras proteinleri, guanin değişim faktörü proteini ile aktive edilmesine rağmen, Ras üzerindeki spesifik mutasyonlar, Ras'ın aktif edilmiş formu ile sonuçlanır. Bu mutasyonlar birçok kanser türünde ve çeşitli gelişimsel hastalıklarda bulunur. Bu nedenle, Ras mutasyonlarını hedeflemek Ras ile ilgili hastalıkların tedavisinde ana odak noktası olmuştur. Yine de Ras proteinleri hala “ilaçlandırılmayan” proteinler olarak sınıflandırılmaktadır. Mutasyonların Ras proteinleri üzerindeki ve efektörleri ile etkileşim dinamikleri üzerindeki etkileri açık değildir. Bu tezde K-Ras4B ve c-Raf etkileşimi, iki alt sistemde moleküler dinamik simülasyonları ile incelenmiştir. Öncelikle hesaplamalı olarak, K-Ras4B ve c-Raf etkileşimini PRISM kullanarak tahmin edildi. Ras ve Raf etkileşimindeki kalıntılar HotRegion kullanılarak tanımlandı. Sonuçlar, K-Ras4B ve c-Raf etkileşiminde yer alan spesifik amino asitler hakkında bir fikir verdi. Tahminleri doğrulamak için, kontrol (G12D) sistemine iki adet yükleri tersine dönüştürülmüş B3 mutasyonu olarak adlandırdığımız (R41E / K42D) mutasyonlar eklendi. Kontrol ve B3 mutasyonuna uğramış alt sistemleri üzerinde MD simülasyonları yapıldı. K-Ras4B₁₋₁₆₆ içeren sistemde, B3 mutasyonu bağlanma enerji seviyelerini kısmen etkilemiştir. Bu sonuçlara göre, B3 mutasyonu bağlanma enerji değerini düşürmüştür. Fakat, enerji seviyelerindeki fark yeterince büyük değildir. Beklenmeyen enerji seviyeleri, sisteme bir membran ve bir sistein açısından zengin alan (CRD) ekleyerek daha fazla çalışılabilir.

ACKNOWLEDGEMENT

I would like to thank my thesis advisors Prof. Özlem Keskin and Prof. Attila Gürsoy, I appreciate their guidance on my thesis project and graduate studies. I am also grateful to our collaborators Prof. Ruth Nussinov, Dr. Hyunbum Jang and Nussinov group. I would also like to thank my thesis jury members: Prof. Halil Kavaklı, Asst. Prof. Elif Nur Fırat Karalar, and Asst. Prof. Sefer Baday for their valuable time.

I would like to mention my dear friends and colleagues Serena Muratcioglu, Sıla Özdemir, Bilgesu Erdoğan, Efe Elbeyli, Ali Tuğrul Balcı, Derya Soner Cavga, Farideh Halakou, Kayra Köseoğlu, Jackson Weako, Meltem Eda Ömür, Asma Hakouz, Damla Övek, Zerina Kurtovic, Harun Kulak, Aykut İşleyen and Mustafa Oğuzhan Karakaya who have always been by my side during these two years both in social and academic life.

I owe my deepest gratitude to my family, especially my parents, sister and brother: Vesile Yazgılı, İsmail Yazgılı, Neslihan Gökçe Yazgılı Kahveci, and Ahmet Kahveci. Without their love, emotional support, understanding and endless encouragement I would not be able to pursue this research work. Thank you for always being there and for making this pursuit worthwhile.

I also acknowledge TUBITAK, the foundation financially supporting this project (Research Grant No: 114M196).

TABLE OF CONTENTS

LIST OF FIGURES	IX
LIST OF TABLES	X
NOMENCLATURE	XI
CHAPTER 1	1
INTRODUCTION	1
CHAPTER 2	3
LITERATURE REVIEW	3
2.1. A Brief History of Ras Oncogenes	3
2.2. Ras Family at a Glance	3
2.2.1. <i>The Structure of Ras</i>	4
2.2.2. <i>Ras trafficking and localization in cell</i>	7
2.3. Ras Effectors and Signaling pathways.....	9
2.3.1. <i>Ras-Raf Interaction</i>	10
2.3.2 <i>Ras – PI3K Interaction</i>	12
2.3.3. <i>Ras – RalGDS Interaction</i>	13
2.4. Ras Proteins in Diseases	14
2.4.1. <i>Pancreatic Cancer</i>	14
2.4.2. <i>Colorectal Cancer</i>	15
2.4.3. <i>Lung Cancer</i>	16
2.4.4. <i>Skin Melanoma</i>	17
2.4.5. <i>Ras in Developmental Diseases</i>	18
CHAPTER 3	19
MATERIALS AND METHODS	19
3.1. Computational prediction of K-Ras4B and c-Raf RBD interaction, hot region and hot spot.....	19
3.2. Molecular Dynamics Simulation	20
3.2.1. <i>Generating initial configurations of the K-Ras4B and RBD domain of c-Raf</i>	21
3.2.2. <i>Atomistic MD simulations</i>	23
3.2.3. <i>MM-GBSA calculations</i>	23

3.2. 4. <i>Stability analyses</i>	25
3.2.5. <i>Interaction analysis</i>	26
CHAPTER 4	27
RESULTS AND DISCUSSION	27
4.1. K-Ras4B interaction to c-Raf RBD	27
4.2. K-Ras4B ₁₋₁₆₆ and c-Raf RBD complex in MD simulations system	29
4.2.1. <i>Bonding interaction analysis of the system</i>	32
4.2.2. <i>Calculating the changes in binding free energy upon charge reversal mutations for the system</i>	34
4.3. Discussion	37
CHAPTER 5	40
CONCLUSION	40
APPENDIX A	41
APPENDIX B	47
BIBLIOGRAPHY	86

LIST OF FIGURES

Figure 2.1: GTPases act as molecular switches..	4
Figure 2.2: K-Ras4B protein structure.	5
Figure 2.3: Multiple sequence alignment of K-Ras4B, K-Ras4A, H-Ras and N-Ras. ..	6
Figure 2.4: Catalytic domain of Ras protein.	7
Figure 2.5: Upstream signaling of Ras.....	9
Figure 2.6: Downstream signaling of Ras.....	10
Figure 2.7: Ras-Raf interaction in ERK regulation.....	11
Figure 3.1. Crystal structures of K-Ras4B and c-Raf RBD.	19
Figure 3.2. Flowchart of the Molecular Dynamics simulation.	21
Figure 3.3. Initial configurations of the simulated K-Ras4B ₁₋₁₆₆ and c-Raf complex system.	22
Figure 4.1. Predicted complex structure of K-Ras4B and RBD of c- Raf.	28
Figure 4.2. Crystal structure of the system.....	30
Figure 4.3. The dynamics of complex changes upon mutation.....	32
Figure 4.4. Hydrogen bond and salt bridge analysis for the system.	33
Figure 4.5. Binding free energies of the system.....	35
Figure 4.6: Energy component plots as a function of time for the system.	36
Figure 4.6: Energy component plots as a function of time for the system..	37

LIST OF TABLES

Table 1. Gibbs binding free energy of the system.....	34
--	----



NOMENCLATURE

CRD	Cysteine-Rich Domain
ERK	Extracellular signal-regulated kinase
GAP	GTPase activating protein
GDP	Guanosine diphosphate
GEF	Guanine-nucleotide exchange factor
GTP	Guanosine-5'-triphosphate
HVR	Hypervariable region
MEK	MAPK/Erk kinase
PDB	Protein Data Bank
PI3K	Phosphoinositide 3-kinase
PM	Plasma Membrane
PPI	Protein-Protein Interaction
RalGDS	Ral Guanine Nucleotide Dissociation Stimulator
Ras	Rat sarcoma
RBD	Ras Binding Domain
SOS	Son of Sevenless

Chapter 1

INTRODUCTION

Ras-family proteins are small membrane-associated GTPases and are involved in various cellular functions. They function on the plasma membrane and relay the signals through their effectors. Ras signaling pathways regulate the functioning of proliferation, differentiation, cell morphology, motility, intracellular trafficking, and gene expression. Ras mutations are detected in various types of human cancer. The oncogenic mutations on Ras proteins are result with the constitutively active form of Ras. Although there is tremendous effort to target Ras proteins, these proteins are still “undruggable”. After three decades of effort to understand the structure and mechanism of Ras proteins, researchers try different approaches to target oncogenic Ras. Two major strategies to block oncogenic Ras signaling are targeting its downstream effectors and plasma membrane localization.

Active Ras proteins transmit the information through a physical interaction with their downstream effector proteins which are involved in crucial pathways such as MAPK and PI3K. Mutations on 12nd, 13rd, and 61st residues of Ras proteins result in constitutive activation of the proteins, thus activating their downstream effectors intensely [1]. Understanding why and how the mutations affect the dynamic state of Ras proteins is therefore vital. In this thesis, important residues of K-Ras4B and Ras-Binding-Domain (RBD) of c-Raf interaction complex were identified by using PRISM and HotRegion. Later, molecular dynamics simulations method was used to simulate and validate the interaction of RBD of c-Raf. As the control, K-Ras4B₁₋₁₆₆ with the G12D mutation was used in the system. R41E/K42D double mutation was introduced to the control to construct B3 mutant. These mutated residues are on the effector binding region of Ras.

The result of the system shows that the mutations on effector binding regions alter the binding energy of K-Ras4B to c-Raf.

Chapter 2 gives a comprehensive review of Ras proteins, their shuttling as well as their interaction with the plasma membrane. Ras regulation and signaling, and Ras-related diseases are also further reviewed. Several important studies were given to emphasize how a mutation on Ras proteins affect the function of these proteins.

Chapter 3 provides a detailed explanation of the materials and the methods used in this study. Molecular dynamics (MD) simulations and the parameters that were used in the simulations defined. MMGBSA and the interaction analysis methods also described in detail.

Chapter 4 focuses on the K-Ras4B₁₋₁₆₆ and RBD of c-Raf interaction as the MD simulations system. The effect of charge reversal mutation on the K-Ras4B and RBD interaction analyzed and further discussed. This chapter presents computational results for two different subsystems, control and B3 mutant, and provides a detailed discussion of the results.

Chapter 5 presents an overview of the thesis study, gives a conclusion and future improvements, suggestions, and perspectives.

Chapter 2

LITERATURE REVIEW

2.1. A Brief History of Ras Oncogenes

Rat sarcoma virus (Ras) genes were first identified and characterized as transduced oncogenes in acutely transforming retroviruses [2, 3]. Later, Harvey and Kirsten Ras genes (HRAS and KRAS, respectively) were showed to encode 21 kDa proteins which can bind GDP/GTP [4-6]. The majority of the Ras proteins localize to the cytoplasmic surface of the plasma membrane (PM) where they are activated upon incoming signals [7]. In 1983, a third human Ras gene found in neuroblastoma and named as Neuroblastoma Ras (NRAS) [8].

2.2. Ras Family at a Glance

Ras-family proteins are small GTPases that are located on the plasma membrane. Multiple signaling pathways turn on upon several extracellular stimuli. Those pathways regulate cell proliferation, migration, survival, and differentiation [9, 10]. All Ras proteins share the activity of GTP (guanosine triphosphate) binding and GTP hydrolysis to GDP (guanosine diphosphate) [11]. The change between GTP to GDP causes conformational changes in proteins. When GTP is bound, proteins are activated so they can transmit signals. Conversely, when GTP is hydrolyzed to GDP, proteins switch to an off state that unable to relay a signal [12]. Large, trimeric GTP-binding proteins (G proteins) and small monomeric GTPases are the two major types of GTP-binding proteins. Specific regulatory proteins control both types of the GTP-binding proteins. The rate of the hydrolysis of GTP to GDP is increased by the GTPase-activating proteins (GAPs). This effect is conversely regulated by guanine nucleotide exchange factors (GEFs). GEFs allows binding of a new GTP molecule to the protein by promoting the release of GDP (Figure 2.1) [12]. Ras protein activation depends on GEFs, which catalyze the switch of GDP to GTP.

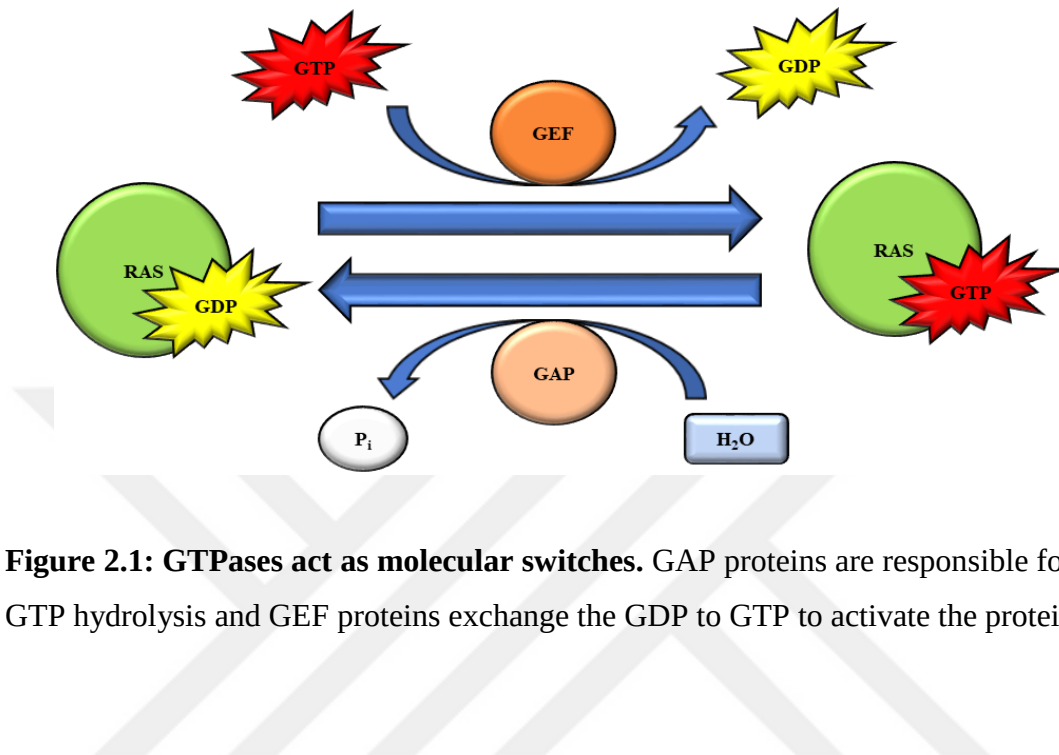


Figure 2.1: GTPases act as molecular switches. GAP proteins are responsible for the GTP hydrolysis and GEF proteins exchange the GDP to GTP to activate the protein.

2.2.1. The Structure of Ras

Three human RAS genes encode four isoforms which are H-Ras, N-Ras, K-Ras4A and K-Ras4B. K-Ras isoforms are formed by a mechanism called alternative splicing [13, 14]. All isoforms have G-domain and hypervariable region (HVR) (Figure 2.2). However, there are several key differences in gene level and protein level between the Ras isoforms.

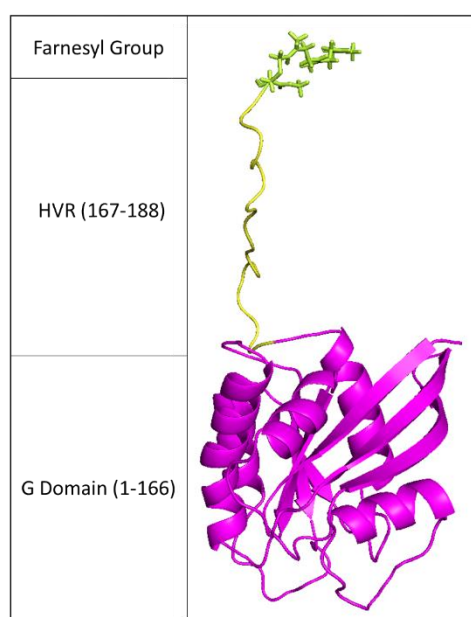


Figure 2.2: K-Ras4B protein structure. G domain of K-Ras4B protein shown in magenta and HVR shown in yellow, whereas farnesyl group shown in lime.

Three isoforms (N-Ras, H-Ras, and K-Ras4A) have 189 amino acids whereas K-Ras4B has 188 amino acids (Figure 3). All isoforms are similar between the 1st and 166th residues (89%). However, the hypervariable region (HVR) has low sequence identity between variants (167-189 for H-Ras, N-Ras, and K-Ras4B; 167-188 for K-Ras4B) [15-17].

K-Ras4A	MTEYKLWVGAGGVGKSALTIQLIQNHFVDEYDPTIEDSYRKQVVIDGETCLLDILDAG	60
K-Ras4B	MTEYKLWVGAGGVGKSALTIQLIQNHFVDEYDPTIEDSYRKQVVIDGETCLLDILDAG	60
H-Ras	MTEYKLWVGAGGVGKSALTIQLIQNHFVDEYDPTIEDSYRKQVVIDGETCLLDILDAG	60
N-Ras	MTEYKLWVGAGGVGKSALTIQLIQNHFVDEYDPTIEDSYRKQVVIDGETCLLDILDAG	60

K-Ras4A	QEEYSAMRDQYMRTGEGFLCVFAINNTKSFEDIHHYREQIKRVKDSDDVPMVLVGNKCDL	120
K-Ras4B	QEEYSAMRDQYMRTGEGFLCVFAINNTKSFEDIHHYREQIKRVKDSDDVPMVLVGNKCDL	120
H-Ras	QEEYSAMRDQYMRTGEGFLCVFAINNTKSFEDIHQYREQIKRVKDSDDVPMVLVGNKCDL	120
N-Ras	QEEYSAMRDQYMRTGEGFLCVFAINNTKSFADINLYREQIKRVKDSDDVPMVLVGNKCDL	120
	*****:*** **: *****:*****	
K-Ras4A	PSRTVDTKQAQDLARSYGIPFIETSAKTRQVEDAFYTLVREIRQYRLKKISKE-EKTPG	179
K-Ras4B	PSRTVDTKQAQDLARSYGIPFIETSAKTRQGVDDAFYTLVREIRKHKEKM-SKD-GKKKK	178
H-Ras	AARTVESRQAQDLARSYGIPYIETSAKTRQGVDDAFYTLVREIRQHKLRLNPPDESGPG	180
N-Ras	PTRTVDTKQAHELAKSYGIPFIETSAKTRQGVDDAFYTLVREIRQYRMKKLNSSDDGTQG	180
	:***::*:**::*****:***** *:*****::: : .	
K-Ras4A	CVKIKKCIIM	189
K-Ras4B	KKSKTKCVM	188
H-Ras	CMSCK-CVLS	189
N-Ras	CMGLP-CVM	189
	*::	

Figure 2.3: Multiple sequence alignment of K-Ras4B, K-Ras4A, H-Ras and N-Ras.

Acidic residues are shown as blue, Small (small+ hydrophobic (incl.aromatic -Y)) are shown as red, Basic – H are shown as magenta while Hydroxyl + sulfhydryl + amine + G are shown in green. Asterisks indicate positions which have a single, fully conserved residue. Colons indicate conservation between groups of strongly similar properties. Dot indicates conservation between groups of weakly similar properties. [18, 19]

The overall topology of the catalytic domain of Ras protein is globular fold. Catalytic domain contains six β -Sheets (β 1, β 2, β 3, β 4, β 5, and β 6) and five α helices (α 1, α 2, α 3, α 4, α 5) (Figure 2.4). The effectors of Ras proteins interact with Ras through Switch I and Switch II regions which are on α 1 and β 3, respectively. Ras isoforms differ in the protein level possibly due to the differences in HVR. The differences among amino acids relate to the different conformational state and interactions, thus provide isoform-specific information in Ras signaling [20]. On the other hand, some basic functions need similarity on HVR such as Ras recruitment to the plasma membrane [13, 21, 22]. CAAX

motif (C, cysteine; A, aliphatic amino acid; X, any amino acid) in HVR is conserved in C-terminal of Ras proteins [23-25]. CAAX motif importance has shown with a study in which mutations in this motif resulted with the PM localization and Ras signaling demolishment [26].

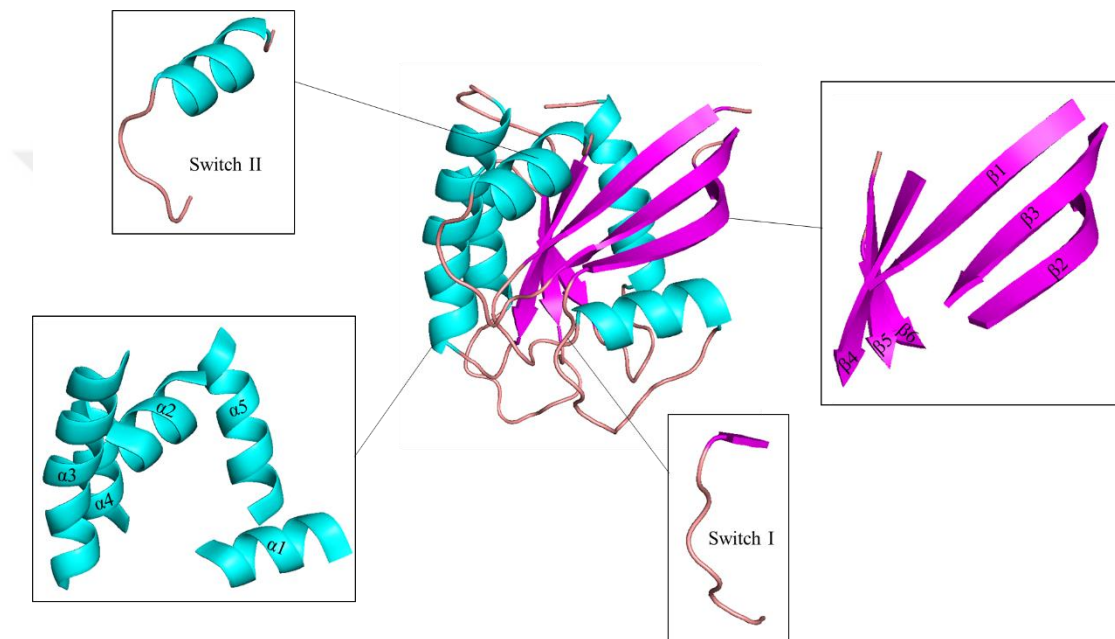


Figure 2.4: Catalytic domain of Ras protein. Catalytic domain of Ras protein is conserved through different Ras isoforms. The similarity of catalytic domain between isoforms are 89% [15-17].

2.2.2. Ras trafficking and localization in cell

Ras proteins are synthesized in the cytosol and they are functional when associated with the plasma membrane. Post-translational modifications (PTMs) are needed for a stable association between the Ras proteins and the membrane. As mentioned before, Ras isoforms have CAAX motif on HVR. The motif on HVR is the most common place where the post-translational mechanisms occur. Firstly, the cysteine is prenylated by farnesyl protein transferase. This process facilitates binding to the endoplasmic

reticulum (ER). Later, the AAX motif is cleaved off by the Ras-converting enzyme (Rce1). Finally, the farnesylated cysteine is methylated by isoprenylcysteine carboxymethyl transferase (Icmt) [27]. This farnesyl group provides weak membrane binding affinity and it is stabilized by another signal motifs. Second signal motifs are essential to correct trafficking and microlocalisation of Ras isoforms [28]. One of the second signal mechanism includes lipid modifications with palmitoyl groups. Another alternative method enables electrostatic interactions with negatively charged phospholipid headgroups of the positively charged lysine residues [28].

Farnesyl group together with the basic hexalysine patch is necessary and sufficient for K-Ras4B membrane localization. K-Ras4A is monopalmitoylated, but contains an adjacent basic patch [29]. For the placement of H-Ras to the plasma membrane via Golgi and transport vesicles, the farnesyl group and two palmitoyl groups on cysteine 181 and 184 are the two key residues for post-translational modifications [30, 31]. On the other hand, additional hydrophobic/basic residues on the HVR is required for monopalmitoylated N-Ras [29].

Ras proteins lay in a mixture of immobile clusters and they are freely diffusing monomers. Each cluster is estimated to be in the range of 12–24 nm in diameter. The clusters comprise approximately 7 Ras proteins which have a lifetime of 0.1–1 s [32, 33]. Ras clusters transient and there are constant formation and disassembly of the clusters. The clusters lately referred to nanoclusters because of the size. Ras nanocluster formations are caused by the interplay which is between the C-terminal lipid anchor, adjacent HVR and G-domain of Ras with specific phospholipids and cholesterol in lipid bilayer [34]. Each Ras isoform nanoclusters have their unique structural properties. H-Ras and N-Ras nanoclusters are shown to cholesterol-dependent, but K-Ras4A and K-Ras4B nanoclusters show different behavior [33, 35–38]. The Ras nanoclusters act as sites for Ras effector recruitment and activation [39].

2.3. Ras Effectors and Signaling pathways

On active state, Ras interacts with a variety of downstream proteins, called “Ras effectors”. Ras effectors share little sequence homology. The association of Ras to those effectors (Raf kinase, PI3K and RalGDS) requires a conserved region called the Ras-binding domain (RBD) which shows ubiquitin fold [40, 41]. Activated Ras strongly interacts with the RBD of its downstream effectors. However, when inactive, attraction to these domain decreases approximately 1000-fold [42].

Purified Ras shows intrinsic GTPase activity. Although the activity is very low, bound GTP slowly switches to GDP. Intrinsic conversion of GDP to GTP is also regulated slowly. However, these processes are catalyzed within the cell by GTPase enzymes [9]. The reactions accelerated with GAPs and GEFs, respectively. Those proteins are the upstream regulator proteins of Ras [43].

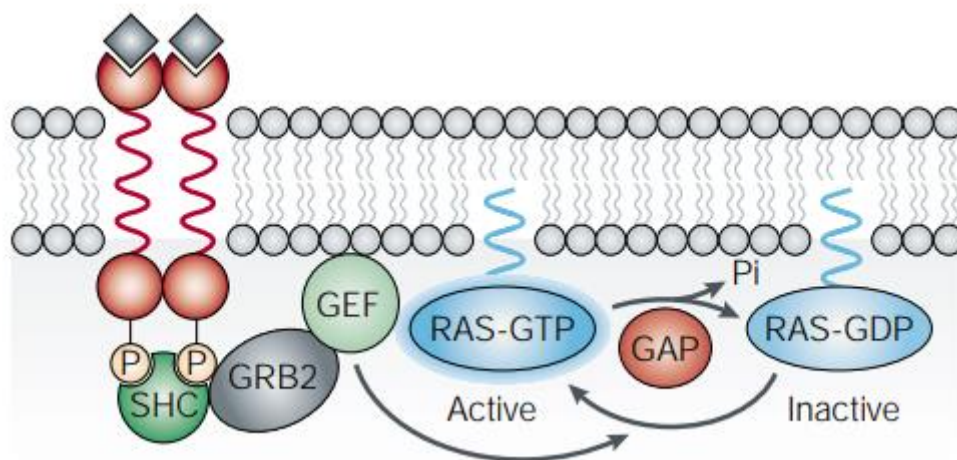


Figure 2.5: Upstream signaling of Ras. Ras protein is activated through GEFs and inactivated through GAPs actions [44].

In human, around 100 genes contain RAS-association (RA) or RBDs. However, not each of these proteins shown to have a direct interaction with Ras proteins in physiological conditions. Only 20 effectors of Ras proteins have been confirmed using

direct affinity measurements [45, 46]. One of the most studied effectors of Ras proteins is Raf protein. First identified in 1993, Raf has three closely related proteins: A-Raf, B-Raf, and c-Raf, also known as Raf1 [47-49]. The second most studied Ras effector is phosphatidylinositol 3-kinases (PI3Ks) [50, 51]. PI3K activity is necessary for the activation of the Akt family Ser/Thr kinases which have roles in the inhibition of apoptosis and the promotion of cell survival [52, 53]. Another well-studied effector of Ras proteins is the exchange factor for Ras-related Ral protein named as Ral guanine nucleotide dissociation stimulator (RALGDS) [54, 55]. RAL proteins are involved in vesicular trafficking in the cell and activation of NF-kB (Figure 2.6).

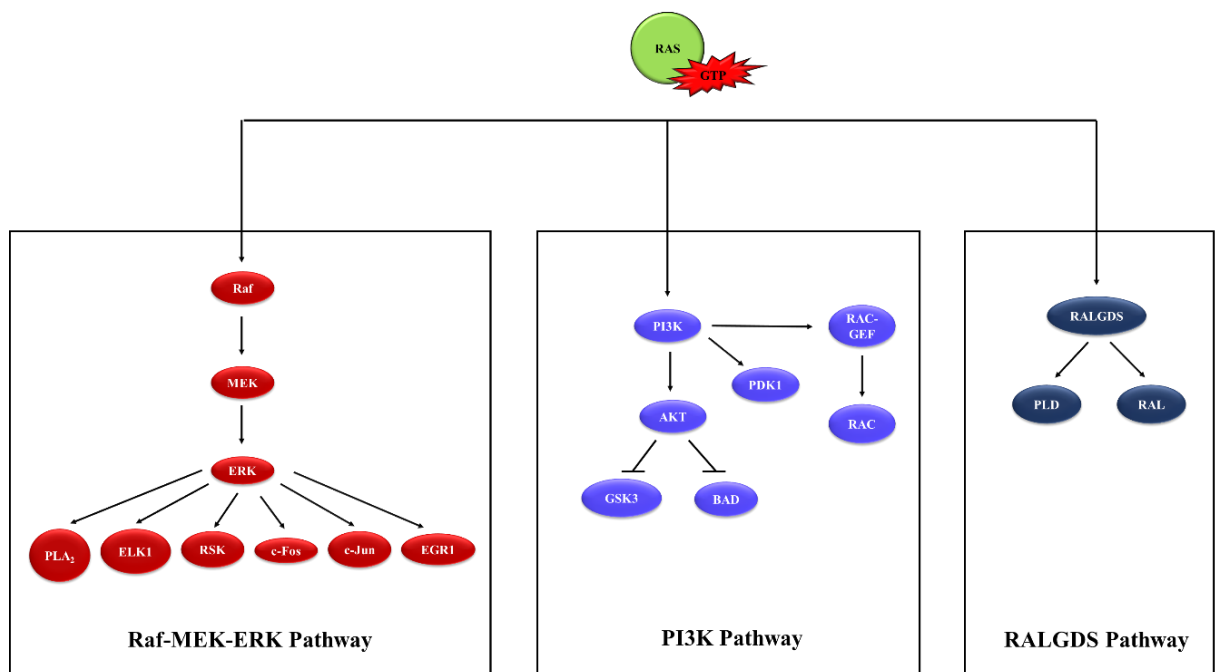


Figure 2.6: Downstream signaling of Ras. Ras proteins are involved in several different pathways through their interactions with their effectors.

2.3.1. Ras-Raf Interaction

The most studied Ras downstream effector is the protein serine/threonine kinase Raf. Active Ras interacts with three closely related Raf proteins: c-Raf, B-Raf and A-Raf [56]. Upon activation, Ras contributes to the activation of Raf proteins as it gathers Raf proteins to the plasma membrane [57]. The activation of Raf protein causes the phosphorylation and activation of MEK proteins. Subsequently, activated MEK proteins

phosphorylate and activate ERK proteins. This pathway includes cytosolic and nuclear proteins; therefore, they can be transported into the nucleus where they can activate a couple of transcription factors. Those transcription factors can regulate and promote cell-cycle progression (Figure 2.7) [58].

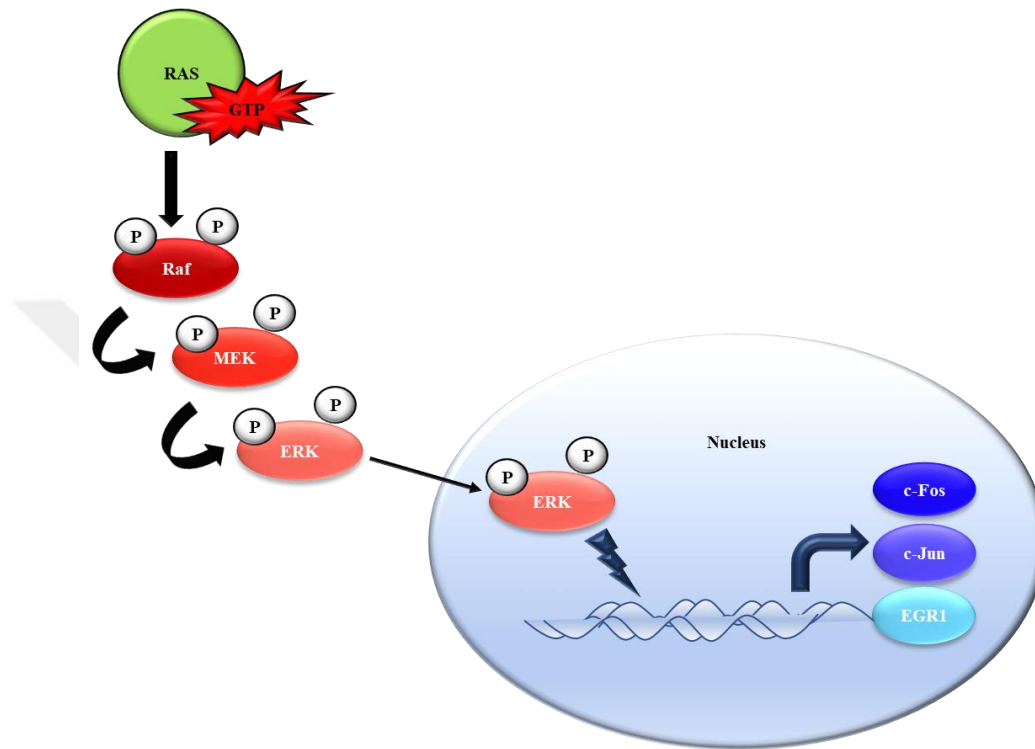


Figure 2.7: Ras-Raf interaction in ERK regulation.

Previous experiments showed the Ras Binding Domain (RBD) of c-Raf (between 55th and 131st residues) has direct an interaction with the H-Ras-GTP [59-61]. Moreover, mutations on the 31st, 35th, and 37th residues of H-Ras found to significantly impair the interaction of RBD [62]. 31st, 35th, and 37th residues are on the Switch I region, thus this study suggests that the RBD domain of c-Raf interacts directly with the switch I region of H-Ras protein [62]. The switch I region is highly conserved between all isoforms of Ras proteins (Figure 2.4). Contacts between the switch I region of Ras proteins and RBD is crucial for the stabilization of the interaction of Ras to its effectors [63]. Thus, switch I region is important for other Ras isoforms with the interaction of c-Raf protein as well as the interaction for other Ras effectors.

2.3.2 Ras – PI3K Interaction

Phosphatidylinositol 3-kinases (PI3Ks) are members of a conserved family of lipid kinases and another effector which binds directly to the Ras [64, 65]. PI3Ks are split into three classes according to sequence homology and their substrate preference. Class I PI3Ks include p110 α , p110 β , p110 γ and p110 δ encoded by the genes PIK3CA, PIK3CB, PIK3CG, and PIK3CD, respectively. Class I group divides into two sub-group Class IA (p110 α , p110 β and p110 δ) and Class IB (p110 γ). Class IA group binds p85 type of regulatory subunit whereas Class IB binds either p101 or p87. Class II comprises PI3K-C2 α , PI3K-C2 β and PI3K-C2 γ . Class III PI3K has only signals feeding into vacuolar protein sorting 34 (VPS34) [66]. Recent studies showed that VSP34 involves in the regulation of insulin sensitivity and autophagy [67-69].

There are several upstream signals that activate PI3Ks. These activator signals both dependent and independent of Ras proteins. Thus, studying the role of Ras proteins as the activator of PI3Ks is challenging [63]. G-protein-coupled receptors (GPCRs) and receptor tyrosine kinases (RTKs) are the two families that can activate PI3Ks. Although GPCRs and RTKs regulate PI3Ks independently from Ras proteins by binding p85, binding of RTKs to Ras can result in binding of RTKs to p110 α which activates PI3K signaling [70]. p110 α , p110 γ , and p110 δ bind to Ras proteins through their RBD, however, p110 β RBD interacts with Rab, Rac and Rho GTPase family members [66, 71-74]. Class II PI3Ks have RBD, yet the function is not clear. On the other hand, Class III PI3K does not have RBD.

Ras-dependent PI3K signaling requires a cascade of events. First, RTKs dimerize and are autophosphorylated at tyrosine residues upon growth factor stimulation. Autophosphorylation activates Src homology 2 (SH2) domain containing molecules like Grb2. Grb2 then interacts with a Ras-GAP, Son of Sevenless (SOS). Activation of Ras leads to the direct interaction of the PI3K-RBD in the p110 subunit [75, 76]. The interaction is stabilized through the switch I region of Ras protein and RBD of PI3K.

This interaction causes a change in the orientation of PI3K. Thus PI3K can access its lipid substrate which is inserted in the membrane more easily [41].

Ras-PI3K interaction leads to the activation of a lipid kinase. PI3K phosphorylates phosphatidylinositol 4,5-bisphosphate (PIP₂) to produce phosphatidylinositol-3,4,5-trisphosphate (PIP₃). PIP₃ serves as a second messenger that carries signals from receptor to a target and leads to the control of many downstream enzyme activations [50]. The binding affinity of Raf to Ras is significantly stronger than the binding affinity of PI3K to Ras. [73, 77]

2.3.3. Ras – RalGDS Interaction

RAL guanine nucleotide dissociation stimulator (RALGDS), RALGDS-like gene (RGL/RSB2) and RGL2/RLF exchange factors are another effector family for Ras proteins. RalGDS gene was first identified in 1993 by using sequences from yeast CD25 gene [78]. All members of RalGDS share common protein structure: a central CDC25 homology domain, with an upstream Ras Exchange Motif (REM) and RBD [79].

RalGDS proteins bind to Ras family proteins through their C-terminal RBD. This RBD region first identified by N- and C-terminal deletion experiments [80-84]. RalGDS RBD contains ~90 amino acids. The 3D structure of the RBD domain has been determined by X-ray crystallography, NMR spectroscopy and molecular modeling studies [80, 85-87]. Raf RBD and RalGDS RBD similarities, in terms of hydrophobicity and folding, have been shown, yet the sequence identity is very low [88].

RalGDS proteins are involved in the process of dissociation of the guanine nucleotides from Ras family member Ral. All the members of RalGDS family are effectors for Ras proteins. These proteins enable Ras to stimulate RAL which then involves in the regulation of different transcription factors. The interaction between Ras and RalGDS is formed through the antiparallel intermolecular β -sheet between RBD β 2 and Ras β 2. The interaction is stabilized by charged amino acid side chain interactions similar to the

interaction of Ras and Raf RBD [46, 89-91]. RALGDS pathway is associated with cell-cycle arrest and apoptosis processes [92]. As previously explained, effector binding site (switch I, between 30th and 40th residues) is a critical region for all known Ras downstream effectors. Several different studies showed the importance of this region. For example, T35 and D38 substitution affect the binding of RalGEFs and PI3K, while E37 and Y40 are important for the interaction of Ras to c-Raf & PI3K and c-Raf & RalGEFs, respectively [93].

2.4. Ras Proteins in Diseases

Ras protein mutations are involved in several different cancer types. Among all isoforms, mutations in KRAS is the most commonly found (85%) followed by NRAS (12%) and HRAS (3%). Several studies showed that specific Ras isoforms are associated with specific types of tumors. KRAS mutations are found mostly in pancreas cancer (91%), colon cancer (42%) and lung cancer (33%) whereas NRAS mutations mainly found in skin cancer (27%) and HRAS in adrenal glands (10%) [94, 95]. Those mutations are found mostly (98%) at the codons G12, G13 or Q61. G12 mutation accounts for 83% of KRAS mutations whereas it is ~10% in NRAS. Most mutations of NRAS occurs at Q61 (63%). Those mutations typically alter RAS proteins function by making them constitutively active [96]. Various mutations have found with recent studies; however, the biological importance of those mutations is largely unknown. A recent study showed that prognosis of the disease can be predicted by mutations on exon 4 [97].

Ras signaling is also shown to have an impact on several developmental disorders such as neurofibromatosis-type1, Costello syndrome, and Noonan syndrome [98, 99].

2.4.1. Pancreatic Cancer

Pancreatic adenocarcinomas count as one of the aggressive and the worst prognosis even though it is not the cancer type that most frequently seen. Pancreatic cancer is expected to be the 2nd leading cause of cancer-related deaths by 2030 [100]. Among all other types of human cancers, incidences of Ras mutations are the highest in pancreatic cancer including 90% or more of pancreatic ductal adenocarcinoma [101, 102]. Early efforts for targeting pancreatic cancer was through targeting H-Ras and N-Ras [103]. Lately,

therapy studies focused on K-Ras protein. Blocking the plasma membrane localization and downstream signaling of Ras proteins as well as the direct inhibition are the recent hot topics for the treatment of pancreatic cancer [104].

The discovery of lipid modification of the cysteine at the CAAX sequence for the membrane localization lead researches to design specific inhibitors [105, 106]. However, the clinical applications gave different results for different isoforms of Ras proteins [107]. Small molecule inhibitors targeting the functioning of farnesyltransferase enzyme blocked the H-Ras PM localization whereas it did not affect K-Ras PM localization [108]. Studies later showed that for K-Ras isoform there is another related enzyme called geranylgeranyltransferase which can modify K-Ras protein for PM localization [109]. Therefore, for an effective targeting PM localization of K-Ras, both enzymes should be targeted [110].

Substitution of glycine 12 (G12) to either aspartic acid, arginine, or valine has commonly seen specifically in K-Ras variant. This change results in the constitutive activation of the K-Ras protein and accelerates its functioning with the downstream effector proteins such as c-Raf and PI3K [44, 103, 107, 111]. Targeting Raf and PI3K pathway is, therefore, another methodology. However, a single agent that targets a single pathway is not very effective. This outcome is probably due to compensatory activation of another pathway.

2.4.2. Colorectal Cancer

Colorectal cancer (CRC) is the second most common cancer in woman whereas it is the third in men in the world [112]. 1.4 million people are diagnosed with colorectal cancer and around 700 000 of them die in a year [113]. The 5-year survival of CRC is estimated to be 60% [112, 114]. There are several forms of colorectal cancer, however, they share certain common genetic alterations [115, 116]. K-Ras mutations are present in 50-95% of dysplastic ACF, in 15-78% of colorectal adenomas and in 36-42% of CRCs [117-121]. In the timeline of colorectal cancer progress, K-Ras mutations rather early, yet it can appear any stage [115, 122]. Around 42% of colorectal cancer caused by Ras

mutations which are generally found in exon 2 (codon 12 and 13) [123, 124]. Mutations in KRAS cause constitutive activation of the EGFR pathway. However, studies on this pathway with K-Ras mutations revealed that anti-EGFR monoclonal antibodies in metastatic colorectal cancers are not effective clinically [125]. Thus, K-Ras screening is important for colorectal cancer since it can help to choose treatment type. In this case, if K-Ras mutation detected, patients with mCRC better have another therapy rather than anti-EGFR due to the predicted failure [126, 127].

2.4.3. Lung Cancer

Lung cancer is the leading cause of death among all cancer types with the most common incidence and mortality rates [114, 128]. Lung cancer is related with the lifestyle as tobacco smoking has an important impact on both tumor formation and progression. Two types of lung cancer have been identified, one is non-small cell lung cancer (NSCLC) and the other is small-cell lung cancer [129]. NSCLC is the predominant subtype of lung cancer, being adenocarcinoma the most common histology whereas small-cell lung cancer is the aggressive one [130]. Non-small lung cancer accounts for 80% of all lung cancer cases [131]. Early diagnosis in NSCLC followed by surgical resection increases the survival rates [132]. However, the majority of patients have already at late stages when the disease is diagnosed, and the survival is generally ~18 months after diagnosis [133]. K-Ras mutations in lung cancer mainly found in non-small cell lung cancer (NSCLC). These mutations are also found in squamous cell carcinoma but not in small-cell lung cancer [134, 135]. KRAS mutations are frequently found in tumor cells of smokers compared to those in non-smokers [136]. As in prostate and colorectal cancer, KRAS mutations in codon 12 and 13 are predominant for lung cancers. Most of KRAS-mutant cases in NSCLC present single point mutations at codon 12, while mutations in other positions are relatively rare (in codons 13 and 61) [137]. Within codon 12, the most frequent point mutations are G12C (42%), G12V (21%), G12D (17%) and G12A (7%) [138]. G12V, G12R, and G12C mutations have a worse prognosis compared to the G12D mutations [139]. Those mutations also seem to relate with shortened survival and time to relapse [140].

There are several studies to find a drug targeting the inhibition of MEK or the inhibition of the PI3K pathway [141]. However, finding a drug that is directly acting on K-Ras would be more efficient since it is the one which involves both MEK and PI3K pathway. K-Ras inhibitors have been challenged by the complexity of K-Ras biochemistry. The high affinity of GTP for K-Ras as well as the limited number of accessible active binding sites explain the difficulty in developing specific KRAS inhibitors [42, 142].

2.4.4. Skin Melanoma

Skin cancers are due to the abnormal growth cells that arise from the skin and can invade or spread to the other parts of the body [143]. Three main types of skin cancer are basal-cell skin cancer (BCC), squamous-cell skin cancer (SCC) and melanoma [144]. Melanoma is a malignant transformation of melanocytes and the deadliest skin cancer type [145]. Ultraviolet (UV) light through excess sun exposure as well as the use of tanning beds is a predominant risk factor for melanoma. However, light exposure with family history, genetic predisposition, and multiple nevi (moles) are also risk factors for developing cutaneous melanoma [146].

Two predominant mutations observed in skin melanoma are BRAF and NRAS with 40-50% and 15-30% incidence, respectively [147, 148]. Interest in direct targeting of N-Ras mutant melanoma has grown in recent years as BRAF inhibitors resulted in ERK activation and stimulation of tumor growth in NRAS mutant melanoma [149-153]. In normal melanocytes, N-Ras signals through B-Raf dimers. However, NRAS mutation leads to Raf class switching whereby c-Raf becomes the predominant effector rather than B-Raf [154].

In one of every three melanoma cases, Ras mutations have been detected [147, 154, 155]. Unlike K-Ras and its highly mutated residues (12th and 13th), 61st residue mutation is the most dominant form in N-Ras protein related melanomas [94, 156, 157]. Although NRAS mutations currently do not take in consideration of the treatment strategy in melanoma, several studies showed that BRAF mutations together with NRAS or KRAS mutations showed resistance to the treatment [158, 159].

2.4.5. Ras in Developmental Diseases

Since Ras proteins first identified, large experimental data accumulated. Although, Ras proteins mainly related to various cancer types, Ras mutations are also associated with hereditary familial developmental syndromes. Deeper studies showed that such developmental diseases are not only caused by germline Ras mutations, but also germline mutations in upstream and downstream components of the Ras pathway [160-166].

The hyperactivation of the RAS/MAPK pathway is associated with RASopathies, a group of clinical syndromes. NF1 is an autosomal disease caused by inactivating genetic alterations in the related NF1 gene. The incidence of CNS tumors (gliomas and astrocytomas) are higher in NF1 patients [167, 168]. A possible reason for the high incidence rates due to the hyperactivation of Ras signaling caused by the absence of the downregulatory GAP activity of NF1 [169-171].

Costello Syndrome (CS) is another Ras related disease. CS is autosomal dominant disease and its main driving force is mutations in the H-Ras gene [172, 173]. Glycine 12 substitutions found 80% of total CS mutations. However, mutations in CS are different than the mutations occur in various cancer types. While G12C, G12D and G12V found in different cancer types, CS patients generally have G12S, G12A or G13D [173]. The main symptoms of CS are delayed development, mental retardation, cardiomyopathy, coarse face, or loose skin [99]. CS patients also susceptible to developing rhabdomyosarcomas, neuroblastomas or bladder cancer [162].

Noonan Syndrome (NS) is the most common rasopathy and affects 1 in 2,000-2,500 [95, 99]. NS is the most genetically diverse form. NS is mainly caused by mutations in the PTPN11 locus (~50%) but mutations in other genes also found such as Sos1, Sos2, K-Ras, N-Ras, B- and C-Raf, MEK1, MEK2, Sprouty 1, CBL, SHOC2, PP1, MYST4, LZTR1 and A2ML1 [163, 174-180].

Chapter 3

MATERIALS and METHODS

3.1. Computational prediction of K-Ras4B and c-Raf RBD interaction, hot region and hot spot

To predict K-Ras4B interaction with c-Raf RBD, 3GFT and 1C1Y PDB files were used [181, 182]. Chain A from 3GFT for K-Ras4B (166 residues) and chain B for c-Raf RBD (77 residues) were selected to study this interaction. PRISM web server was used to predict the interaction between these two proteins. [183-185]. To see the contacting residues in detail, HotRegion web server was used [186].

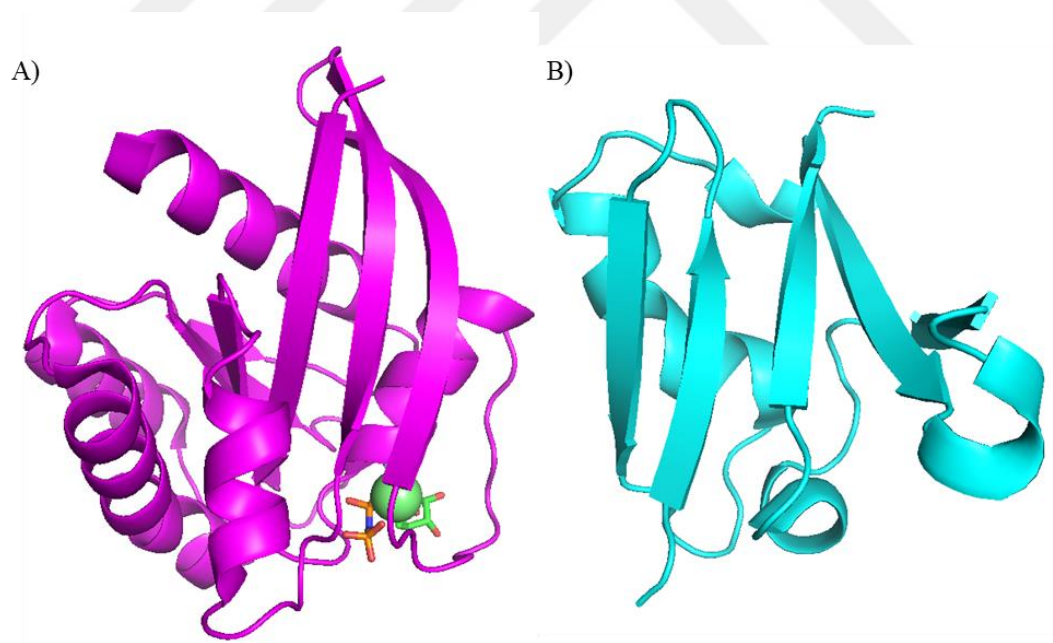


Figure 3.1. Crystal structures of K-Ras4B and c-Raf RBD. A) K-Ras4B (3GFT) crystal structure colored in magenta, Mg²⁺ ion colored in lime, GNP colored by element type, and **B)** c-Raf RBD in cyan by using PyMOL.

3.2. Molecular Dynamics Simulation

Molecular dynamics (MD) simulation developed in the late 70s [187, 188]. Since then, MD advanced from simulating simple and small size chemicals to complex and large biological systems such as proteins in solution or membrane, nucleosomes, and ribosomes by the improvement of computers [189-192].

MD simulations consist of numerical, step-by-step calculation of the classical Newton's equations of motion. The steps involved in MD simulation in this thesis simplified with the flowchart in Figure 3.2. To start with, the molecule of interest obtained either by experimental structures (X-Ray, NMR, etc.) or by modeling the system by comparative modeling data. During the solvation step, molecules are either implicitly or explicitly solvated [193-195]. In the implicit solvation method, water effect is modeled. However, in the explicit solvation water molecules are included in the system. Implicit solvation is computationally cost less, however explicit solvation is better at covering the real solvation effects including the hydrophobic effect [196]. After the solvation, the system assigns the initial velocities and does heating dynamics. If the temperature and the structure are good to go, the next step is production dynamics. In this thesis, NPT ensemble (constant number of atoms, pressure, and temperature) was used as the production dynamics. After the MD simulations are done, the next step is analysis. For this thesis, Root Mean Square Deviation (RMSD) and Fluctuation (RMSF) values were analyzed. Bond formation and losses were studied, and lastly, binding free energy values for each system were calculated.

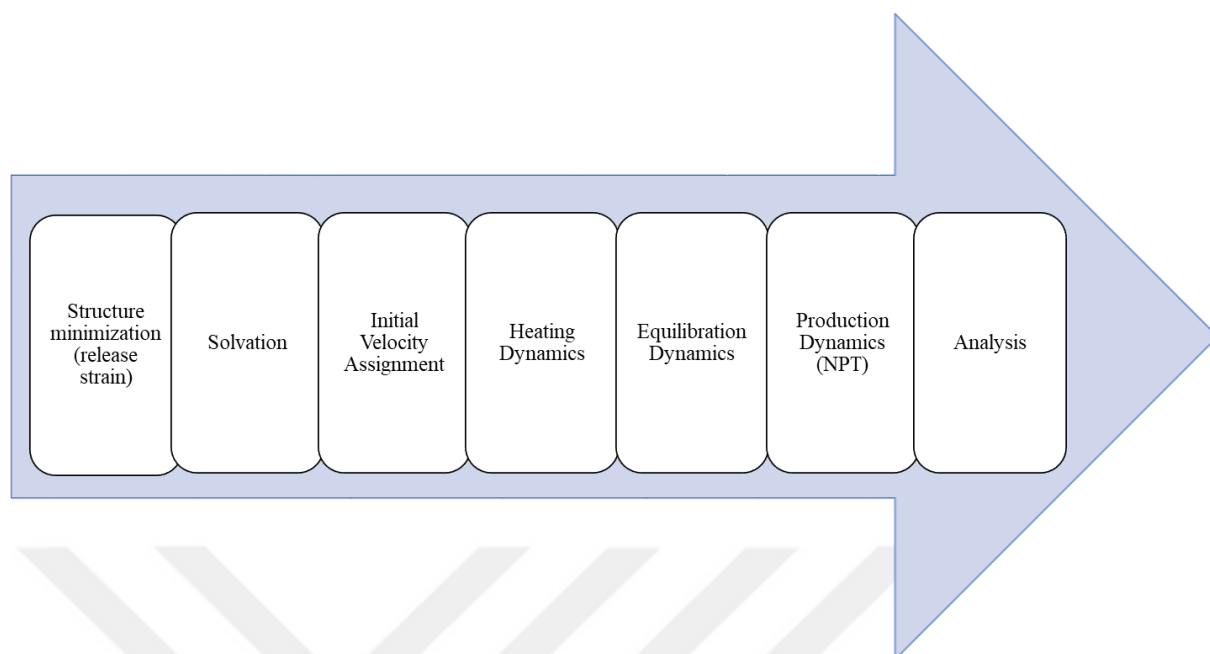


Figure 3.2. Flowchart of the Molecular Dynamics simulation. Each box shows a step in the MD algorithm, and the arrow head shows the direction of next step.

3.2.1. Generating initial configurations of the K-Ras4B and RBD domain of c-Raf

The simulation inputs were built using CHARMM c39b2 with the CHARMM 36 force field for the K-Ras4B₁₋₁₆₆ and RBD domain of c-Raf. Bonded parameters for the GTP was adopted from previous works [197-200] (See Appendix B). For all two subsystems, temperature set to 310 Kelvin and same parameter and topology files were used. The unit cell boxes were set as $100 \text{ \AA} \times 100 \text{ \AA} \times 100 \text{ \AA}$. Local interaction distances to van der Waals (VDW) truncated at $12 \text{ \AA}$ and a smooth switching the function used to truncate VDW potential energy smoothly starting at the $10 \text{ \AA}$. Exclude parameter is set as scaled1-4. This parameter specifies which pairs of bonded atoms should be excluded from non-bonded interactions. Initial velocities of the subsystems were adjusted to remove the center of mass motion of the system. This parameter does not preclude later center-of-mass motion due to external forces such as random noise in Langevin dynamics, boundary potentials, and harmonic restraints. Atom patches assigned as hydrogen, thus hydrogen atoms are kept on the same patch as their parents, allowing faster distance checking and rigid bonds. Long- and short-range forces split according

to the C^1 continuous shifting function [201]. Details of the set-up are given in Appendix B.

K-Ras4B-GTP and c-Raf RBD proteins were constructed using the crystal structures (PDB ID: 3GFT and 4DSO (for K-Ras4B), 4G0N and 1C1Y (for c-Raf RBD)) as described previously [198, 200, 202]. The four PDB files were used to align the Ras and Raf proteins. The sequences were modified into human sequence and GTP analogs were converted to GTP for these crystal structures. The complex structures were modeled using PRISM web server [184, 185]. The structures were used as a construct in the system. The control subsystem was set with the G12D mutation. Additional to G12D, 41st, and 42nd residues were mutated into glutamic acid and aspartic acid in the B3 mutated subsystem, respectively. Discovery Studio 2017 R2 was used to insert these mutations [203]. PyMOL was used for alignment of K-Ras4B proteins and taking snapshots [204]. In the system, K-Ras4B₁₋₁₆₆ and RBD of the c-Raf was simulated (Figure 3.3). The subsystems are control and B3 mutant having the G12D mutation and G12D/R41E/K42D mutations, respectively.

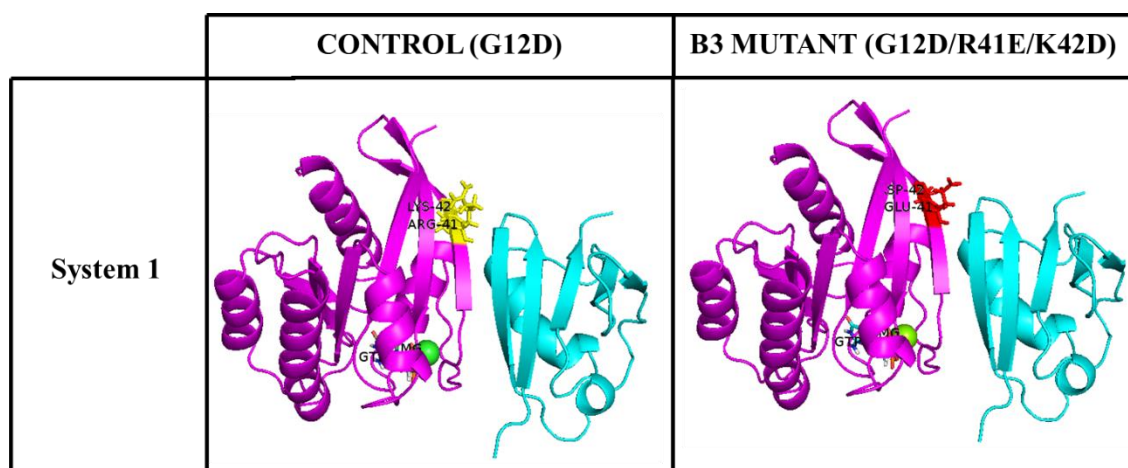


Figure 3.3. Initial configurations of the simulated K-Ras4B₁₋₁₆₆ and c-Raf complex system. The system has K-Ras4B (1-166) and c-Raf RBD.

3.2.2. Atomistic MD simulations

To align the two proteins well with each other, the K-Ras4B-GTP/c-Raf-RBD complex system was minimized with a rigid body motion for the K-Ras4B catalytic domain and RBD in order to enhance the interaction between them. Conjugate gradient energy minimization steps were performed for the subsystems. K-Ras4B and c-Raf-RBD complexes were solvated by the TIP3P water model and the isometric unit cell boxes containing the protein complexes were created. The unit cell box of the two subsystems contain around 102,500 atoms with ions, 1 Mg^{2+} , and 30 Cl^- for both subsystems, and 36 Na^+ and 40 Na^+ for control and B3 mutant, respectively. Minimization was performed for 10,000 steps. In the dynamics, long-range electrostatic interactions were calculated by using the particle mesh Ewald (PME) method. The initial Langevin temperature set as 10 K and it increased to 310 K. The Langevin temperature control maintained the constant temperature at 310 K, and the Nosé-Hoover Langevin piston pressure was kept at 1 atm in the production runs. Waters were made rigid with *rigidBonds*, parameter and the bond between each hydrogen and the atom to which it is bonded is similarly constrained.

The simulations for were performed in the NPT (constant number of atoms, pressure, and temperature) ensemble. A total of 1.0 μs of simulations for two subsystems, each with 500 ns. The NAMD 2.10 parallel-computing code on the Biowulf cluster at the National Institutes of Health (Bethesda, MD) was used for production runs [205]. The first 150 ns trajectories were removed to exempt initial transients. Thus, averages were taken afterward. The system simulated with CHARMM c39 version. The simulated trajectories were analyzed using the CHARMM all-atom additive force field (version C42) programming package [206-208].

3.2.3. MM-GBSA calculations

To compare the interaction strengths of the subsystems, binding energy calculations have been performed. The sum of gas-phase energies and solvation free energies were calculated using the molecular mechanics energies combined with the GB and surface

area continuum solvation (MM/GBSA) method within the CHARMM program [209, 210]. The average binding free energy is derived from the Gibbs free energy,

$$\langle \Delta G_b \rangle = \langle \Delta G_{gas} \rangle + \langle \Delta G_{sol} \rangle - T\Delta S \quad (1)$$

$$\Delta G_{gas} = \Delta H_{intra} + \Delta H_{vdW} + \Delta H_{elec} \quad (2)$$

$$\Delta G_{sol} = \Delta G_{sol}^{elec} + \Delta G_{sol}^{nonpolar} \quad (3)$$

$$T\Delta S = T\Delta S_{trans} + T\Delta S_{rot} + T\Delta S_{vib} \quad (4)$$

where $\langle \Delta G_b \rangle$ is the average binding free energy and $\langle \Delta G_{gas} \rangle$ is the average gas phase contribution. The gas phase contribution to the binding free energy comes from the partition of internal energy, the van der Waals interaction, and the electrostatic energy. [211, 212] (Equation 2) $\langle \Delta G_{sol} \rangle$ is the solvation energy contribution and it can be divided into the electrostatic contribution and the nonpolar contribution [211] (Equation 3). $G_{sol}^{nonpolar} = \gamma SASA + \beta$, where SASA denotes the solvent accessible surface area with a set of the surface tension parameter. $T\Delta S$ is the entropic contribution and divided into the translational, rotational, and vibrational terms (Equation 4). The translational and rotational entropies are evaluated from the calculation of principal moment of inertia, and the vibrational entropy is obtained from the quasiharmonic mode calculation in the VIBRAN module of the CHARMM program [210]. Overall, the average binding free energy was calculated as sum up all these contributions;

$$\Delta G = \Delta H_{intra} + \Delta H_{vdW} + \Delta H_{elec} + \Delta G_{sol}^{elec} + \Delta G_{sol}^{nonpolar} - (T\Delta S_{trans} + T\Delta S_{rot} + T\Delta S_{vib}) \quad (5)$$

The change in binding free energy due to the complex formation is calculated as follows for K-Ras4B and c-Raf-RBD complexes:

$$\Delta G_b = G_b^{complex} - (G_b^{K-Ras4B} + G_b^{RBD}) \quad (6)$$

3.2. 4. Stability analyses

The equilibrium state of each subsystem after the simulations and residue fluctuations were measured by RMSD and RMSF analyses, respectively. Root-mean-square Deviation (RMSD) is a measurement indicating how two structures are different than each other by calculating root mean square distance between each corresponding atoms [213, 214].

$$RMSD = \sqrt{\frac{1}{N} \sum_{i=1}^N \delta_i^2} \quad (7)$$

RMSD is calculated according to the above equation, where N is the number of atoms, δ_i is the distance between i^{th} atoms from two different structures (Equation 7). Root-mean-square fluctuation (RMSF) is the measurement of the average displacement of each residue through the whole simulations. RMSF is calculated to indicate the fluctuating residues and regions. It gives an idea of the overall behavior of specific residues and regions in the whole structures.

$$RMSF = \sqrt{\frac{1}{T} \sum_{t_j}^T (x_i(t_j) - \bar{x})^2} \quad (8)$$

RMSF is calculated according to the above equation, where T is the duration of the simulation in terms of time steps, $x_i(t_j)$ is the coordinates of i^{th} residue at time t_j , $\bar{x}$ is the mean coordinates for i^{th} residue (Equation 8).

3.2.5. Interaction analysis

Inter- and intramolecular salt bridges and hydrogen bonds (H-bonds) established through the simulations were detected with VMD plug-in (version 1.9.2) [215]. For salt bridge formation, atoms closer than 3.2 Å were considered forming the salt bridges [216]. In this thesis, salt bridges and H-bonds, which were detected over 10% of total time frames of the simulations were considered as stable interactions.



Chapter 4

Results and Discussion

4.1. K-Ras4B interaction to c-Raf RBD

Raf proteins are one of the most studied effectors of Ras protein. Although there is H-Ras and c-Raf RBD crystal structure in Protein Data Bank (PDB), there is no K-Ras4B and c-Raf RBD complex crystal structure up to date. To understand the interaction between the K-Ras4B and c-Raf RBD interaction, PRISM 2.0 was used [184, 185, 217]. PRISM predicts the binding residues of two proteins by structural similarity and evolutionary conservation of the putative binding residues named as hot spots [185]. The PDB ID used for target 1 was as the 3GFT (chain A), and 1C1Y (chain B) for target 2. No specific template was chosen. Two different interface templates gave three different energy values. For the interface 1c1yAB, binding energy score (BES) found as -51.19 and -44.63, and for the interface, 3o44AI BES found as -5.04. Two of the predicted structures were used with the lowest BES for further hot spot predictions. Later, HotRegion was used to predict the hot spot clusters (Figure 4.1) [186]. According to the analysis results for BES: -51.19, K-Ras4B has five hot spots: ILE21, GLU37, ASP38, TYR40, and LEU56. These residues reside on $\alpha 1$, Switch I, $\beta 2$ and $\beta 3$. HotRegion predicted four hot spots on RBD: GLN66, THR68, VAL69, and ARG89 (Figure 4.1b). The other 1c1yAB interface with BES: -44.63 also analyzed with HotRegion. According to the analysis results for BES: -44.63, K-Ras4B has five hot spots: ILE21, GLU37, TYR40, ASP54 and LEU56. HotRegion predicted four hot spots on RBD: GLN66, THR68, VAL88, and ARG89 (Figure 4.1d). The other 1c1yAB interface with BES: -44.63 also analyzed with HotRegion.

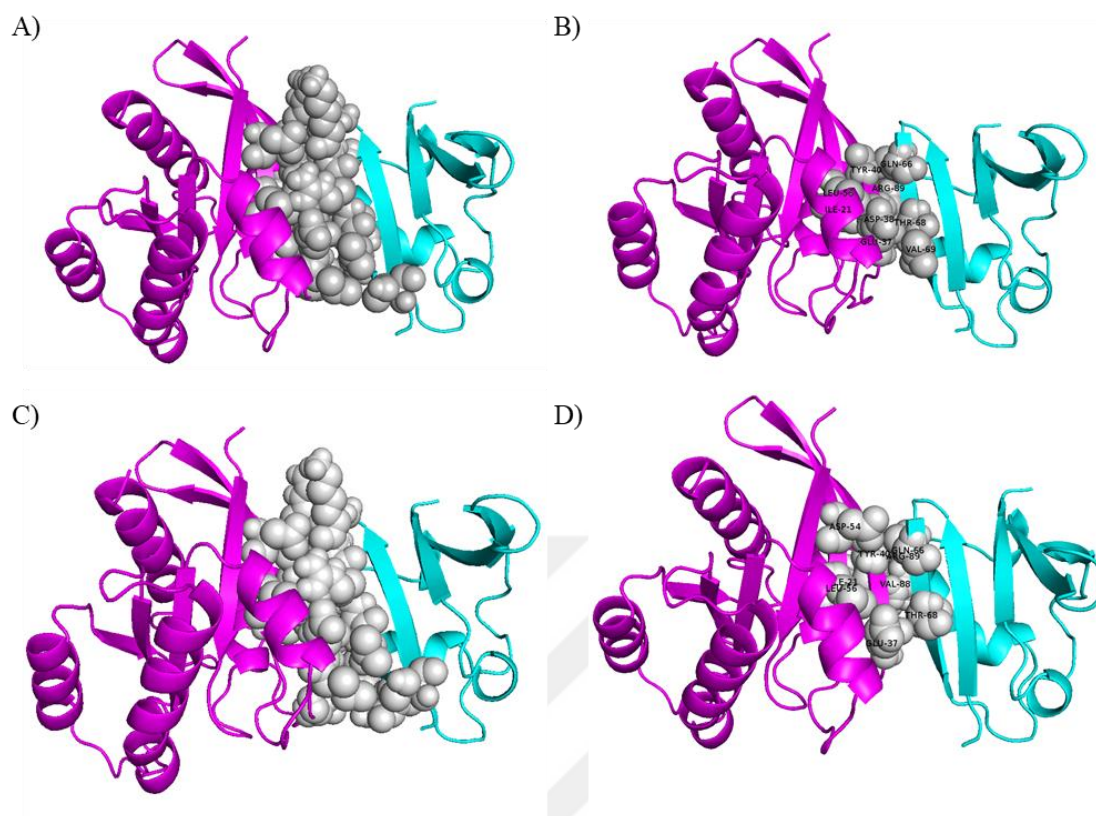


Figure 4.1. Predicted complex structure of K-Ras4B and RBD of c- Raf. **A)** Residues participating in the interface with BES: -51.19 are shown in ball-and-stick representation. Interface residues of K-Ras4B are: ILE21, GLN25, VAL29, ASP33, ILE36, GLU37, ASP38, SER39, TYR40, ARG41, ASP54, and LEU56. Interface residues of c-Raf RBD are: THR57, ASN64, LYS65, GLN66, ARG67, THR68, VAL69, LYS84, VAL88, ARG89, and GLY90. **B)** Hotspot residues on K-Ras4B and RBD with BES: -51.19 labeled. (K-Ras4B: magenta, RBD: cyan, hot regions and hot spots: gray) **C)** Residues participating in the interface with BES: -44.63 are shown in ball-and-stick representation. Interface residues of K-Ras4B are: ILE21, GLN25, ASP33, ILE36, GLU37, ASP38, SER39, TYR40, ARG41, ASP54, and LEU56. Interface residues of c-Raf RBD are: LEU56, THR57, ASN64, LYS65, GLN66, ARG67, THR68, VAL69, ASN71, LYS84, VAL88, ARG89, and GLY90. **D)** Hotspot residues on K-Ras4B and RBD with BES: -44.63 labeled. (K-Ras4B: magenta, RBD: cyan, hot regions and hot spots: gray)

4.2. K-Ras4B₁₋₁₆₆ and c-Raf RBD complex in MD simulations system

There are one GTP, one Mg^{2+} ion, one c-Raf RBD and one K-Ras4B G-domain (1-166) in the crystal structure of the K-Ras4B₁₋₁₆₆ – c-Raf RBD complex system. The complex formation occurs between the switch I region (30th -38th residues), $\beta 2$, $\beta 3$, and $\alpha 1$ of K-Ras4B (Figure 2.4) and Ras Binding Domain (RBD) of c-Raf.

Atomistic MD simulations were performed on the system for 500 ns. K-Ras4B (G12D) was used as the control for the system. Two new mutations inserted into the control before the MD simulations. The double mutation is charge reversal (hereinafter called as B3 mutant), meaning two positively charged residue mutated into two negatively charged residue. Arginine is mutated into glutamic acid (R41E) and lysine is mutated into aspartic acid (K42D) (Figure 4.2). Arginine and lysine are positively charged amino acids whereas glutamic acid and aspartic acid are negatively charged amino acids in physiological pH 7.4. The equilibrium conditions were tested checking RMSD and energy plots as a function of time. RMSD analysis showed that both two subsystems, the control and the B3 mutant, equilibrated around 150 ns (Figure 4.2c). Binding free energy analysis data also support that the system equilibration balanced around 150 ns (Figure A.1).

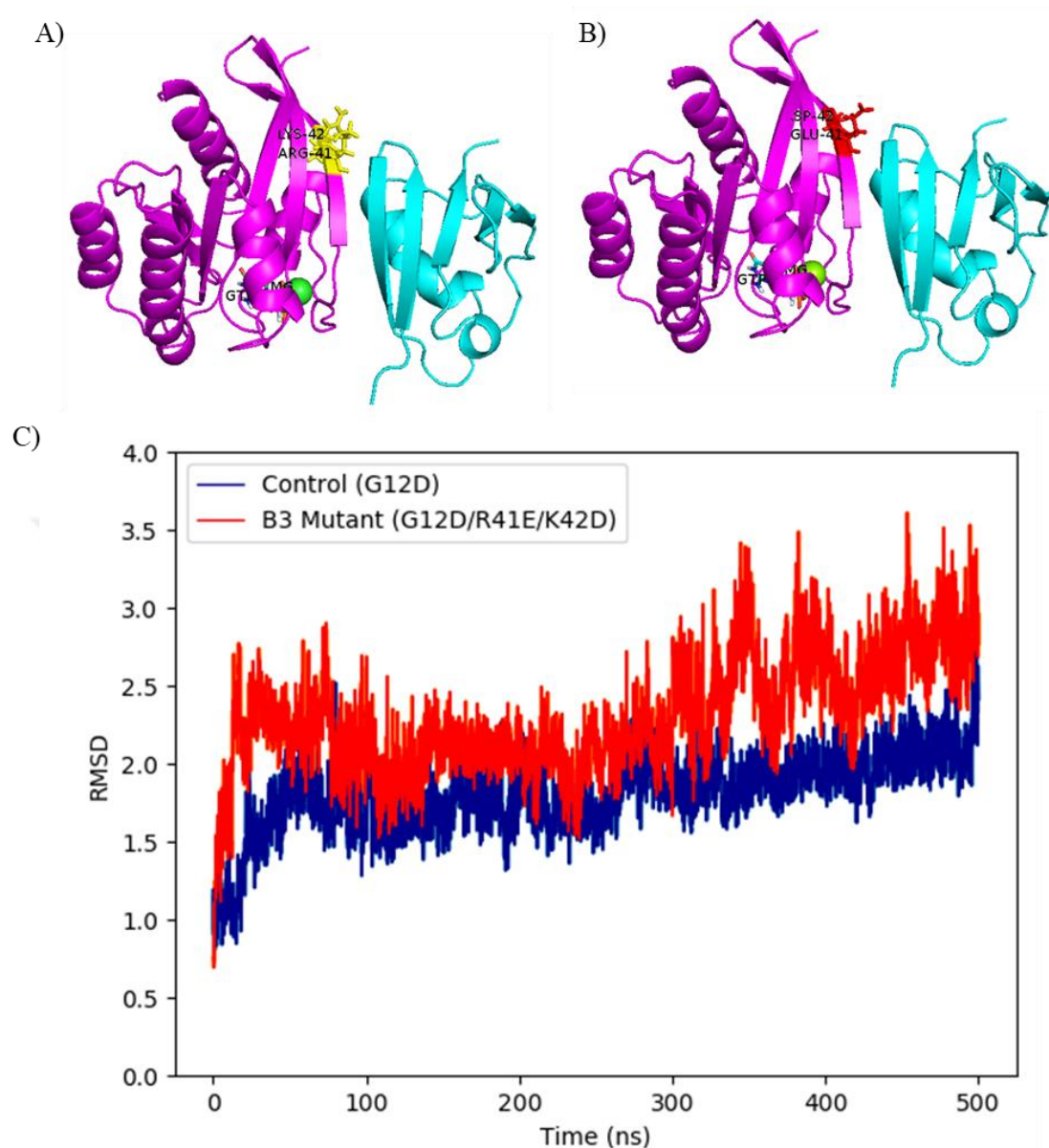


Figure 4.2. Crystal structure of the system. **A)** The control for the system, contains G12D mutation, **B)** B3 mutant contains G12D/R41E/K42D mutations. Both control and B3 mutant have GTP and Mg²⁺ (shown in colors according to the atom and lime, respectively). Target residues to be mutated are shown in yellow and mutated residues are shown in red. K-Ras4B is shown in magenta and c-Raf RBD is shown in cyan. **C)** RMSD values for both two subsystems between 0-500 ns.

Throughout the simulations, RBD fluctuates more compared with the G-domain of K-Ras4B upon mutation (Figure 4.3). The fluctuation of RBD first observed by snapshots as control and B3 mutant (Figure 4.3a and Figure 4.3b, respectively). To have a deeper understanding of the fluctuations, Root Mean Square Fluctuation (RMSF) analysis was done. The RMSF analysis showed the fluctuation differences of G-domain and RBD upon mutation (Figure 4.3c and Figure 4.3d, respectively). According to the RMSF calculations, both G-domain and RBD showed a similar pattern in control and B3 mutant for each. The pattern differs interestingly at the 34th residue, proline, of K-Ras4B. This residue resides in the Switch I region, where many of the Ras effectors bind. The RMSF values for this residue are 1.17 and 0.06 for control and B3 mutant, respectively. Decrease in the RMSF value indicates that the residue loses its flexibility. Proline has exceptional rigidity, and this affects secondary structures [218]. B3 mutations on the 41st and 42nd residues somehow further affected the proline, making it less flexible. Another pattern difference observed on the 62nd residue of K-Ras4B, which is glutamic acid and located on the Switch II. Switch II is a nucleotide binding region and one of the highly studied regions. Several covalent binders specific to this region is identified. In a recent study, a new ligand, 2C07, found to engage the active GTP state as well as the GDP state of K-Ras, unlike many other ligands which target only GDP state. The interaction of the covalent ligand stabilized the GDP state, prevented SOS binding and catalyze exchange [219]. 2C07 derived electrophiles found to inhibit PI3K, but not Raf-1-RBD binding. Thus, a change in the stability of the Switch II may affect binding of the Ras to its effector. No other certain G-domain fluctuation difference between control and B3 mutant is not observed. RBD fluctuation during simulation is decreased in general upon charge reversal mutations.

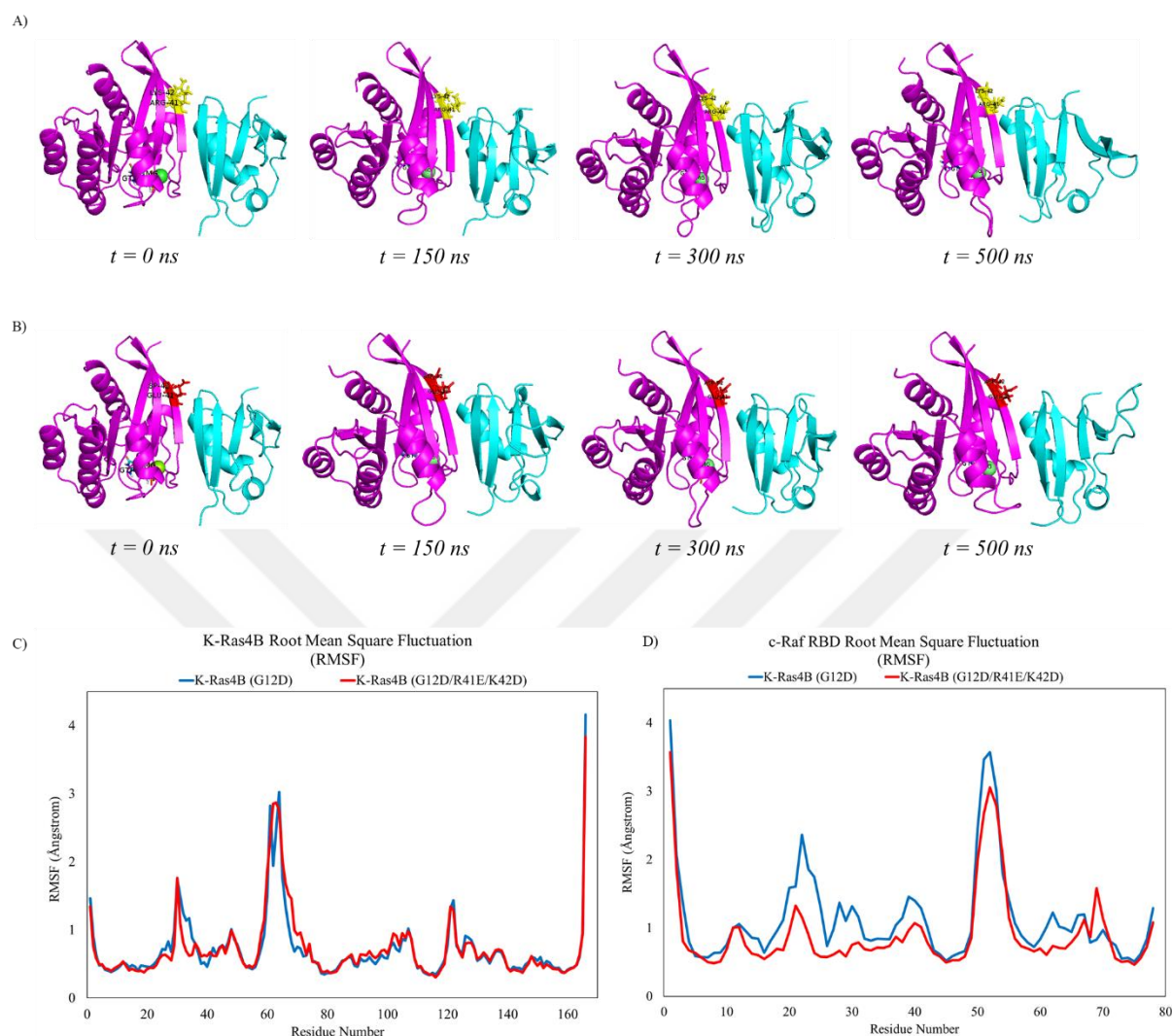


Figure 4.3. The dynamics of complex changes upon mutation. A) Initial, 150 ns, 300 ns and 500 ns position snapshots of the control, and **B)** Initial, 150 ns, 300 ns and 500 ns position snapshots of the B3 mutant visualized with PyMOL. Target residues to be mutated are shown in yellow and mutated residues are shown in red. RMSF calculations of **C)** G-Domain and **D)** RBD are done by using CHARMM c42b2.

4.2.1. Bonding interaction analysis of the system

To understand the dynamics, bond formations and losses were studied (Table A.1 and A.2). Only the hydrogen bonds having more than 20% occupancy throughout simulations with 10% difference between control and B3 mutant is taken into consideration for further discussions.

Hydrogen bond percentages diminished near the effector binding region of K-Ras4B (Figure 4.4a). Mutation on the 41st residue lead to a new hydrogen bonding with the 12th residue on the RBD. The charge reversal mutations mainly affected residues on β 2. However several different regions also affected. Switch I, an important region for K-Ras4B binding to many of its effectors also affected by this mutation. The hydrogen bond analysis might show a tendency for a loss between the interaction of K-Ras4B and RBD. However, a decrease in the number of H-bond does not give an exact clue whether the interaction between the two proteins are weakened or not. Increased percentages of hydrogen bonding between the several residues might be more effective than the decrease of the others.

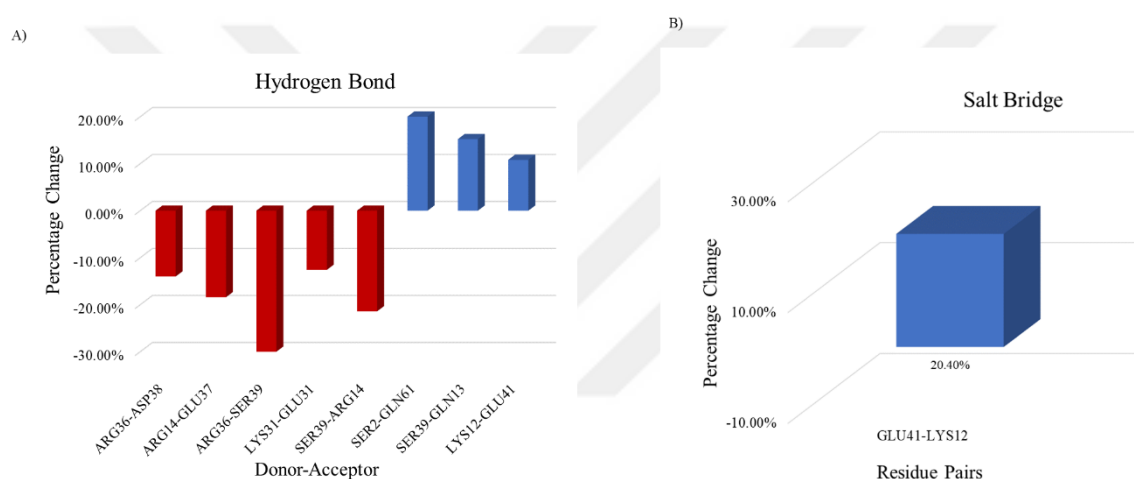


Figure 4.4. Hydrogen bond and salt bridge analysis for the system. **A)** Hydrogen bond decrease and increase upon mutation are shown in red and blue, respectively. **B)** In the system, several salt bridges forms both in control and B3 mutant. However, only 10% change between control and B3 mutant is considered.

In both control and B3 mutant, salt bridges on several residues detected (Table A.2). However, no significant change observed except that the mutated residue GLU41. Mutation of arginine to glutamic acid on 41st residue lead to a new and significant salt bridge (Figure 4.4b). Like the newly formed hydrogen bond between LYS12 and GLU41, the new salt bridge also formed between these residues. This might involve increasing the interaction of the K-Ras4B and c-Raf.

4.2.2. Calculating the changes in binding free energy upon charge reversal mutations for the system

For further analysis of interaction strengths of the system and how bonding dynamics affect the dynamics of the interaction, the binding free energy of control and B3 mutant calculated. Molecular mechanics energies combined with the generalized Born surface are continuum solvation (MM-GBSA) method used for the binding free energy analysis (Table A.3 and Table A.4). Energy calculations were done between 150-500 ns. The first 150 ns omitted for letting the system to be equilibrated. There are differences both in ΔG_{sol} , solvation energy, and ΔG_{gas} , gas phase energy of control and B3 mutant. However, together with the entropy contribution and other components of the binding free energy, the difference decreases (Figure 4.5 and Figure 4.6). The calculated binding free energies for the control and the B3 mutant K-Ras4B-RBD interactions are -14.9 ± 8.1 and -15.3 ± 8.6 kcal/mol, respectively (Table 1). The binding free energy difference is not significant but shows a tendency for the B3 mutant to be energetically favorable.

Table 2. Gibbs binding free energy of the system.

SYSTEM	$\Delta G_b = G_b^{\text{complex}} - (G_b^{K-Ras4B} + G_b^{RBD})$
CONTROL (G12D)	-14.9080 ± 8.1251 (kcal/mol)
B3 MUTANT (G12D/R41E/K42D)	-15.3391 ± 8.5921 (kcal/mol)

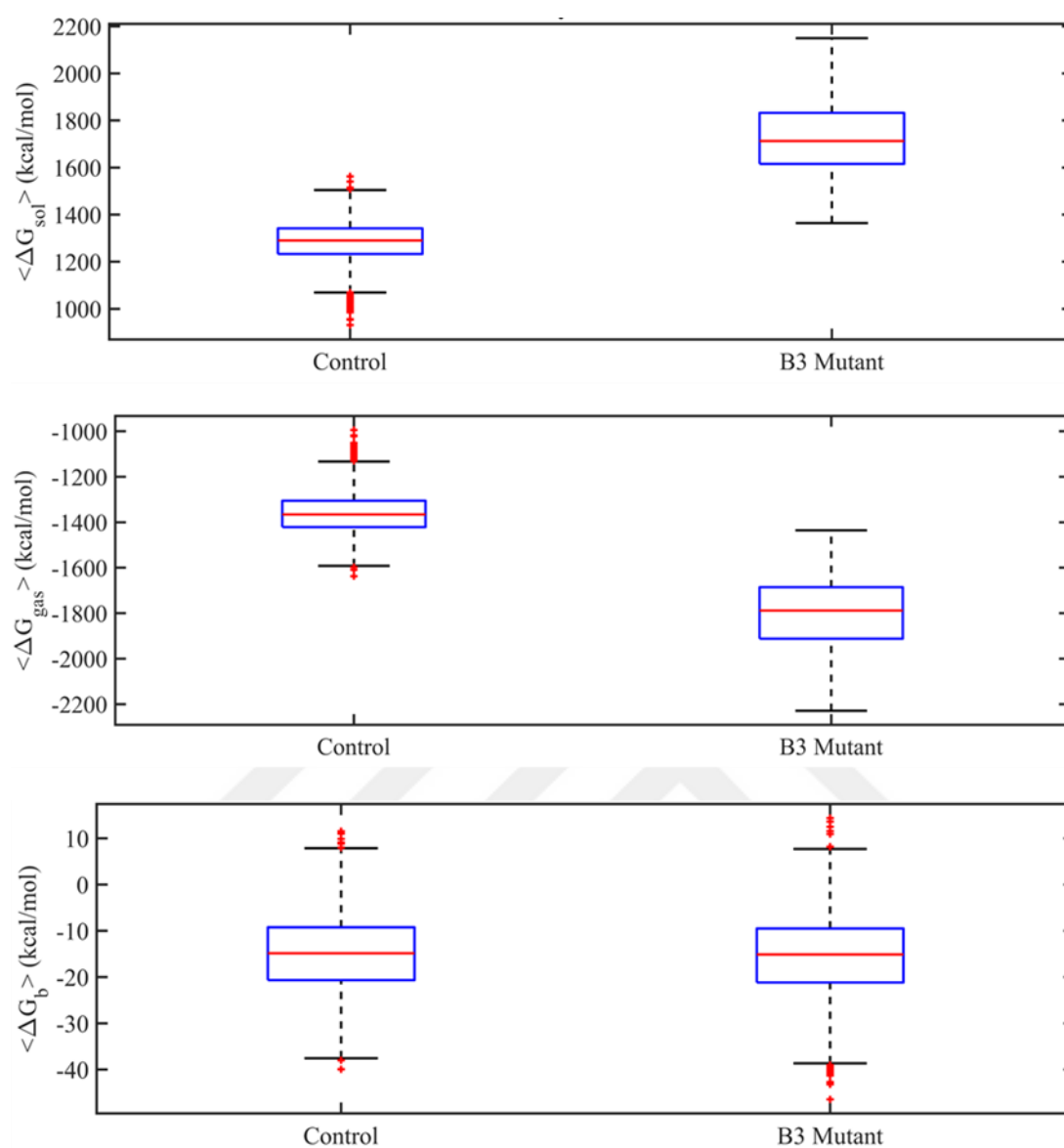


Figure 4.5. Binding free energies of the system. The solvation energy contribution, $\langle \Delta G_{sol} \rangle$, gas phase contribution, $\langle \Delta G_{gas} \rangle$ and the binding free energy $\langle \Delta G_b \rangle$ in kcal/mol calculated by the MM-GBSA method. On each box, the central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually using the '+' symbol [220-222].

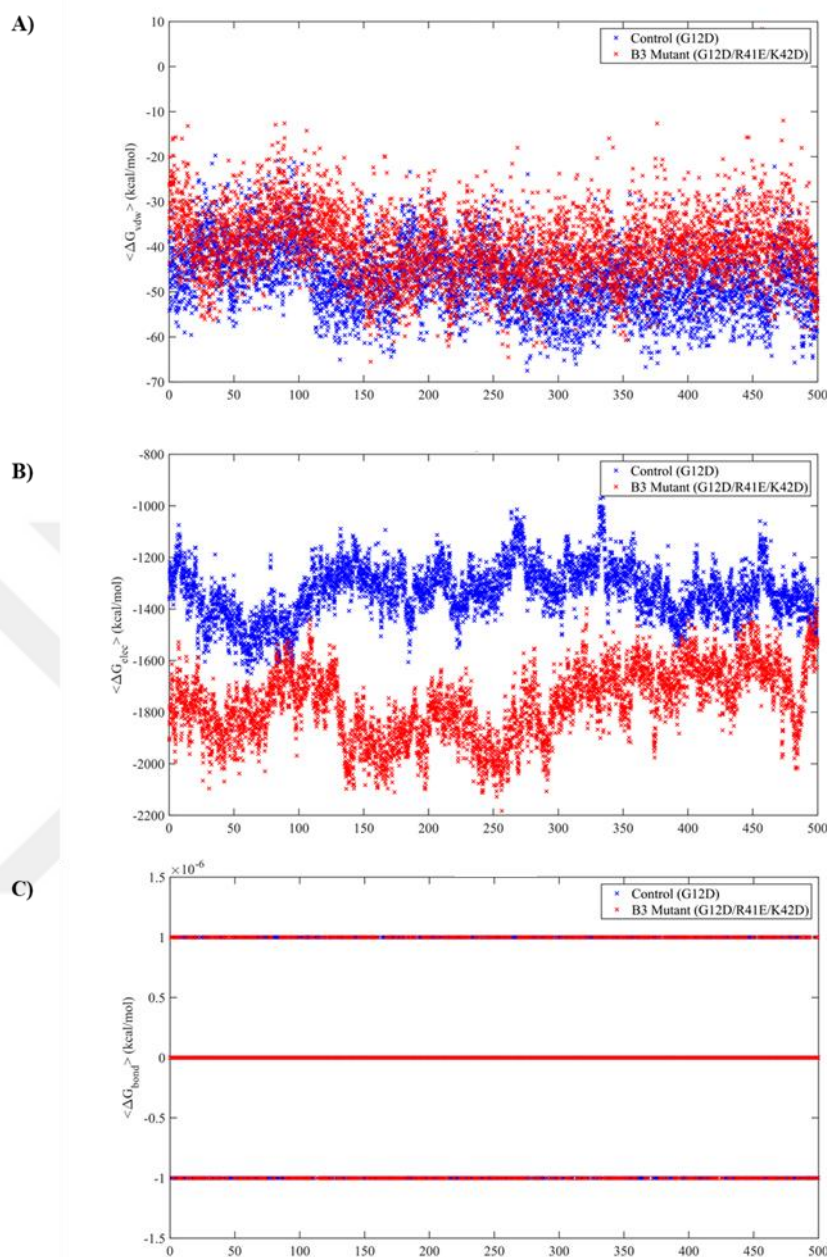


Figure 4.6: Energy component plots as a function of time for the system. A) ΔG_{vdw} , the van der Waal energy, **B)** ΔG_{elec} , the electrostatic energy, **C)** ΔG_{bond} , bond energy, **D)** ΔG_{angl} , angle energy, **E)** ΔG_{dih} , dihedral energy between 0-500 ns.

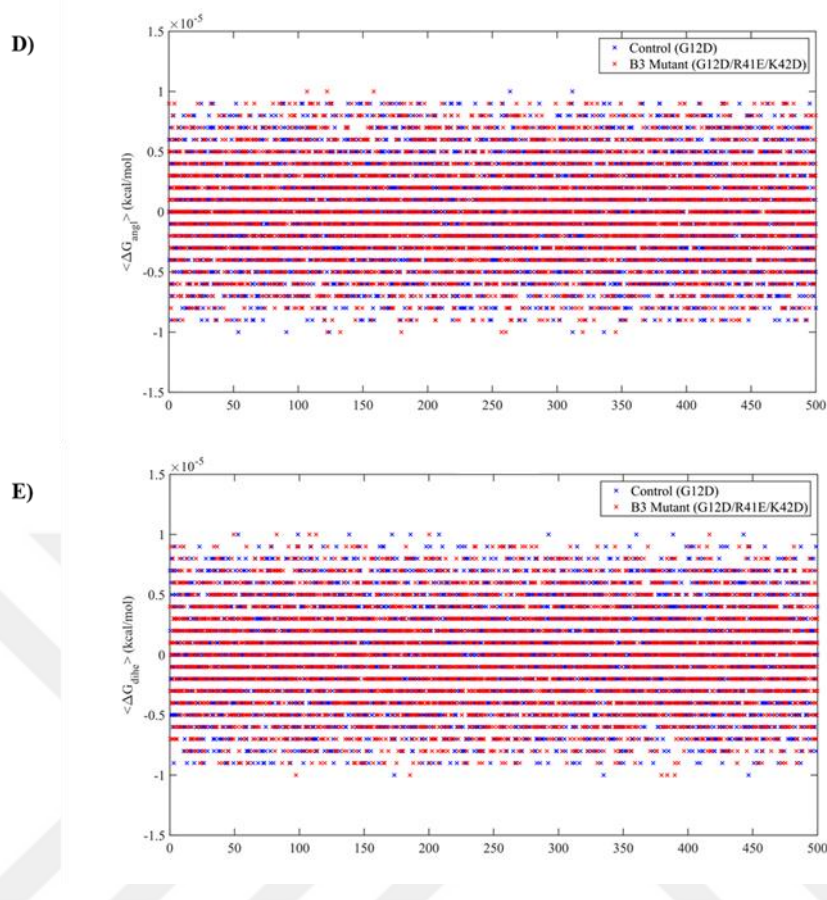


Figure 4.6: Energy component plots as a function of time for the system. A) ΔG_{vdw} , the van der Waal energy, **B)** ΔG_{elec} , the electrostatic energy, **C)** ΔG_{bond} , bond energy, **D)** ΔG_{angl} , angle energy, **E)** ΔG_{dih} , dihedral energy between 0-500 ns.

4.3. Discussion

Since first identified as an oncogene, Ras proteins are one of the most interested proteins [223]. Mutations in Ras proteins discovered in different cancer types and developmental diseases [95, 99]. Despite tremendous efforts to target Ras proteins in diseases, they are still considered as “undruggable”. There are four different Ras isoforms: H-Ras, N-Ras, K-Ras4A, and K-Ras4B. The sequence identity between the isoforms quite similar for G-domain (89%), but the sequence identity for the hypervariable region (HVR) is quite low [15, 29, 224]. To develop a ligand that targets Ras proteins, it is essential to understand the molecular requirements of Ras to function. First, researchers tried to target Ras proteins directly. However, these efforts have not succeeded. Latest studies focused on targeting the interactions of Ras with the plasma

membrane and its effectors. Another focus is to find synthetic lethal interactions with the Ras oncogene. A potential ‘Achilles’ heel’ interaction, therefore would be lethal only in the cell with Ras mutation [103]. Hence, understanding the molecular dynamics of the Ras proteins and their interaction dynamics with the effector proteins are vital. Here, K-Ras4B and c-Raf RBD complex were simulated. The system composed of K-Ras4B G-Domain between the 1st and 166th residues and RBD of c-Raf. In the control subsystem, K-Ras4B had the G12D mutation whereas B3 mutant subsystem had two additional mutations: R41E and K42D.

In the system, the complex simulated for 500 ns. The system observed to equilibrate around 150 ns by RMSD analysis and energy plot functions. The flexibility of the two residues on K-Ras4B is changed drastically upon mutation. PRO34 and GLU62 flexibility are decreased upon B3 mutation. These residues are located on two important regions: Switch I and Switch II, respectively. Switch I region is important in effector binding while Switch II is important for ligand binding. The B3 mutant decreased the flexibility of RBD in general, except for the 69th residue on RBD. VAL69 on RBD is a hot spot within the interaction of K-Ras4B and c-Raf RBD (Figure 4.1b). In RMSF values for this residue are 0.83 and 1.58 in control and B3 mutant, respectively. The mutations on K-Ras4B thus affected the flexibility on c-Raf RBD. It was previously shown that mutations on Ras can attenuate the HVR-catalytic domain interactions either by direct or longer range interactions at Switch I region [202]. Therefore, the change in flexibility of the 69th is arguably caused by the B3 mutations. R41E, arginine to glutamic acid mutation, introduce two new bonds. Between GLU41 and LY12 hydrogen bond and salt bridge formations are observed. Fluctuations and bond formations gave the idea that the mutation affects the complex interaction. To see whether this mutation is favorable by energetically or not, MMGBSA calculations were done. The binding energies for the control and B3 were calculated as -14.9080 ± 8.1251 (kcal/mol) and -15.3391 ± 8.5921 (kcal/mol), respectively. The results Show the B3 mutant subsystem is more favorable than control. However, the difference is not significant enough.

In this thesis, only the interaction between K-Ras4B₁₋₁₆₆ and RBD was simulated. However, HVR accommodates the significant sequence identity difference between the four Ras isoforms.

Ras proteins anchors to the plasma membrane through the interactions between the HVR and plasma membrane [28]. HVR also have a role as it functions to target Ras proteins to the ER and Golgi apparatus [31]. All these roles for HVR show its importance in Ras protein dynamics. K-Ras4B functions with its effector when it is attached to the membrane. Therefore, in the future addition of HVR to the system can be done. The system contains only the Ras Binding Domain (RBD) but no other regions, such as Cysteine-Rich Domain (CRD). Recently, the role of CRD in K-Ras and Raf interaction is studied. According to the new data, CRD found to have a role in decreasing the K-Ras-RBD fluctuations at the membrane which leads to increasing affinity in the complex interaction [225]. Another recent study supports the importance of CRD by showing, primary interaction of CRD with either Ras or Raf affects the topology and membrane orientations [226]. If the CRD primarily interacts with Ras, orientation favors RBD away from the membrane, whereas the interaction between CRD-Raf induces an orientation that favors RBD making membrane interactions.

Taken all together, we showed that B3 mutation affects the Ras interaction to Raf. Experimental analysis showed the mutations are unfavored. However, the computational result showed that B3 mutation has slightly less binding free energy compared to control. Since in this thesis only the molecular dynamics of the Ras-Raf interaction complex was studied, the interaction dynamics of Ras to other effectors, such as PI3K is unclear. Notably, RBD is not the only contributor to Ras- Raf dynamics [225, 226]. Thereby, the work in this thesis suggests better simulation conditions to study Ras-Raf interaction which contains full-length K-Ras4B anchored to the membrane and c-Raf RBD with CRD.

Chapter 5

CONCLUSION

The main focuses of the thesis are to predict K-Ras4B and c-Raf interaction, understand the interaction dynamics, and validate the effect of mutations near the switch I region. It also discusses improvement strategies for the simulated system.

Ras proteins regulate a variety of cell functions through their interactions to effector proteins [9, 10]. Ras binds to its effectors through Ras Binding Domain found on the effectors, such as Raf [40, 41]. Raf is recruited to the membrane upon Ras activation [57]. Previous studies showed that Raf RBD interacts to Ras through Switch I region found on Ras [59-61]. By using a powerful structural prediction algorithm, PRISM, the interacting residues reside on the Ras-Raf protein complex were confirmed. To emphasize the importance of the switch I region, molecular dynamics simulations method was used and the effect of two charge reversal mutations (R41E/K42D) was studied. K-Ras4B₁₋₁₆₆ and c-Raf RBD were used in the system. The results for RMSF and interaction analysis showed a slight oscillation between the control and B3 mutant. According to the free binding energy results, B3 mutant is energetically more favorable than control. However, differences were not significant in this system. Considering the latest studies, this thesis discusses a better simulation environment by adding HVR, a membrane for HVR attachment and CRD region for c-Raf [225, 226].

To summarize, the thesis' study supports the importance of the switch I region in K-Ras4B and c-Raf interaction, discusses the interaction dynamics and suggests improvements.

APPENDIX A

Table A. 1. Hydrogen bond analysis between K-Ras4B (1-166) and c-Raf RBD in the system

Donor	Acceptor	Control (G12D)	B3 Mutant (G12D + R41E + K42D)
ARG14-Main	SER39-Side	6	11.6
ARG14-Main	SER39-Main	0	0.2
ARG14-Side	GLU37-Side	54.4	36
ARG14-Side	GLU41-Side	0	0.4
ARG14-Side	ASP54-Side	0	7.2
ARG14-Side	SER39-Side	0	0.2
ARG20-Side	ASP33-Side	41.6	48.6
ARG20-Side	GLU31-Side	1.4	0.2
ARG20-Side	TYR32-Side	0.2	0
ARG36-Side	ASP38-Side	65.6	51.6
ARG36-Side	SER39-Main	44.4	14.4
ARG41-Main	GLN13-Side	9.8	0
ARG41-Side	ASN11-Main	0.8	0
ARG41-Side	ASN11-Side	0.2	0
ARG41-Side	LYS12-Main	1.4	0
ARG6-Side	GLU37-Side	68.4	0
ARG6-Side	TYR71-Side	0.6	0
ARG6-Side	GLU62-Side	0	3.2
ASN11-Side	GLU41-Side	0	0.2
ASN18-Side	ASP33-Side	1	0.2

ASN18-Side	PRO34-Main	0.2	0
GLN13-Side	GLU41-Main	0	1.6
GLN13-Side	GLN25-Side	0	1.8
GLN13-Side	GLU41-Side	0	1.6
GLN25-Side	VAL35-Main	2.8	3.6
GLN61-Main	THR1-Side	0	0.2
GLU41-Main	GLN13-Side	0	6
GLU63-Main	SER2-Side	0	0.6
HSD27-Side	VAL35-Main	0	0.4
LYS12-Side	GLU3-Side	2.6	0.2
LYS12-Side	GLU41-Side	0	10.8
LYS12-Side	ASP54-Side	0	0.8
LYS12-Side	GLN43-Side	0	0.2
LYS31-Side	GLU31-Side	35.6	23
LYS31-Side	ASP33-Side	24.8	20.4
LYS31-Side	ASP30-Side	6	0
LYS31-Side	GLU31-Main	0.2	1.2
LYS31-Side	ASP30-Main	0	1
LYS31-Side	TYR32-Main	0	0.6
LYS34-Side	GLU31-Side	0.2	7.8
LYS34-Side	HSD27-Side	0.2	0
LYS34-Side	ASP30-Side	0	2.2
SER2-Main	GLU63-Side	0	3
SER2-Main	GLN61-Side	0	0.2
SER2-Side	GLU62-Side	0.2	5.4
SER2-Side	GLU62-Side	0	5.6

SER2-Side	GLU63-Side	0	3.4
SER2-Side	GLN61-Side	0	0.2
SER39-Main	ARG14-Main	22.2	8.8
SER39-Side	ARG14-Main	0.4	0.8
SER39-Side	GLN13-Side	2.2	17.4
THR15-Side	ASP38-Side	68.4	64.2
THR15-Side	GLU37-Main	0.2	0
THR1-Main	GLU62-Side	12.6	5
THR1-Main	GLN61-Side	0.4	0.4
THR1-Main	GLN61-Main	0.8	0.2
THR1-Main	GLU63-Side	0.6	6
THR1-Side	GLU62-Side	1	0.6
THR1-Side	GLN61-Side	0.4	1
THR1-Side	GLU63-Side	0	1
THR4-Side	GLU62-Side	0	0.4
TYR64-Main	THR1-Side	0	0.2
TYR64-Side	SER2-Side	0	0.2
VAL16-Main	GLU37-Main	7.6	5.2

Table A.2. Salt bridge analysis of protein complex in the system.

Residue Pairs	Control (G12D)	B3 Mutant (G12D + R41E + K42D)
ASP30_chainR --- LYS31_chainC	12	4.800
ASP30_chainR --- LYS34_chainC	0	2.400
ASP33_chainR --- ARG20_chainC	0.8	1.200
ASP33_chainR --- LYS31_chainC	17.6	18
ASP38_chainR --- ARG36_chainC	0	1.000
ASP54_chainR --- ARG14_chainC	0	1.000
ASP54_chainR --- LYS12_chainC	0.2	1
GLU3_chainR --- LYS12_chainC	4.600	0.4
GLU31_chainR --- ARG20_chainC	0	0
GLU31_chainR --- LYS31_chainC	17.6	13.8
GLU31_chainR --- LYS34_chainC	0.2	9.000
GLU37_chainR --- ARG14_chainC	0	0
GLU37_chainR --- ARG6_chainC	0.8	0.2
GLU41_chainR --- ARG14_chainC	0	0

GLU41_chainR --- LYS12_chainC	0	20.4
GLU62_chainR --- ARG6_chainC	0	0

Table A.3. Enthalpy and Entropy contributions of complex, K-Ras4B (1-166) and c-Raf RBD in the system.

BIOMOLECULE NAME	CONTROL (G12D)		B3 MUTANT (G12D + R41E + K42D)	
	ΔH (kcal/mol)	$-T\Delta S$ (kcal/mol)	ΔH (kcal/mol)	$-T\Delta S$ (kcal/mol)
Complex	$-8.0842 \times 10^3 \pm 47.9075$	-768.89	$-7.9510 \times 10^3 \pm 49.4238$	-763.657
K-Ras4B	$-5.9079 \times 10^3 \pm 40.4231$	-520.099	$-5.7823 \times 10^3 \pm 39.2600$	-513.741
c-Raf	$-2.1002 \times 10^3 \pm 27.2189$	-309.897	$-2.0936 \times 10^3 \pm 27.6111$	-309.66
Complex - (K-Ras4B + RBD)	-76.0140 ± 8.1251	61.1060	-75.0831 ± 8.5921	59.7440

Table A.4. Gibbs binding free energy of the system.

SYSTEM	$\Delta G_b = G_b^{complex} - (G_b^{K-Ras4B} + G_b^{RBD})$
CONTROL (G12D)	-14.9080 ± 8.1251 (kcal/mol)
B3 MUTANT (G12D/R41E/K42D)	-15.3391 ± 8.5921 (kcal/mol)

APPENDIX B

Bonded Parameters for GTP

RESI GTP -4.00 ! guanosine triphosphate, by Hyunbum Jang

GROUP

ATOM C4' CN7 0.16 !

ATOM H4' HN7 0.09 !

ATOM O4' ON6B -0.50 ! O6

ATOM C1' CN7B 0.16 ! ||

ATOM H1' HN7 0.09 ! C6

GROUP ! / \

ATOM N9 NN2B -0.02 ! H1-N1 C5--N7\\

ATOM C4 CN5 0.26 ! | || C8-H8

ATOM N2 NN1 -0.68 ! C2 C4--N9/

ATOM H21 HN1 0.32 ! /\ / \

ATOM H22 HN1 0.35 ! H21-N2 N3 \

ATOM N3 NN3G -0.74 ! | \

ATOM C2 CN2 0.75 ! H22 \

ATOM N1 NN2G -0.34 ! \

ATOM H1 HN2 0.26 ! (-)O3G O2B O1A H5' H4' O4' \

ATOM C6 CN1 0.54 ! | | | | \ / \ \

ATOM O6 ON1 -0.51 ! O1G=PG-O3B-PB-O3A-PA-O5'-C5'---C4' C1'

ATOM C5 CN5G 0.00 ! | | | | \ /\

ATOM N7 NN4 -0.60 ! (-)O2G (-)O1B (-)O2A H5" C3'--C2' H1'

ATOM C8 CN4 0.25 ! /\ /\

ATOM H8 HN3 0.16 ! O3' H3' O2' H2"

GROUP ! | |

ATOM C2' CN7B 0.14 ! H3T H2'

ATOM H2" HN7 0.09

ATOM O2' ON5 -0.66

ATOM H2' HN5 0.43

GROUP

ATOM C3' CN7 0.14

ATOM H3' HN7 0.09

ATOM O3' ON5 -0.66

ATOM H3T HN5 0.43

GROUP

ATOM C5' CN8B -0.08

ATOM H5' HN8 0.09

ATOM H5" HN8 0.09

ATOM O5' ON2 -0.62

ATOM PA P 1.50

ATOM O1A ON3 -0.82

ATOM O2A ON3 -0.82

ATOM O3A ON2 -0.74

ATOM PB P3 1.50

ATOM O1B ON3 -0.82

ATOM O2B ON3 -0.82

ATOM O3B ON2 -0.86 ! charge adjusted to yield total triP of -4.0

ATOM PG P3 1.10

ATOM O1G ON3 -0.90

ATOM O2G ON3 -0.90

ATOM O3G ON3 -0.90

BOND O5' C5' O5' PA PA O1A PA O2A PA O3A

BOND O3A PB PB O1B PB O2B PB O3B O3B PG

BOND PG O1G PG O2G PG O3G

BOND C5' C4' C4' O4' C4' C3' O4' C1'

BOND C1' N9 C1' C2' N9 C4 N9 C8 C4 N3

BOND C2 N2 C2 N1 N2 H21

BOND N2 H22 N1 H1 N1 C6 C6 C5 C5 N7

BOND C2' C3' C2' O2' O2' H2' C3' O3' O3' H3T

BOND C1' H1' C2' H2'' C3' H3' C4' H4' C5' H5'

BOND C5' H5'' C8 H8

DOUBLE C2 N3 C4 C5 N7 C8 C6 O6

IMPR C2 N3 N1 N2 C6 N1 C5 O6 N2 H21 C2 H22

DONO H21 N2

DONO H22 N2

DONO H1 N1

DONO H2' O2'

ACCE O6 C6

ACCE N3

ACCE N7

ACCE O1A PA

ACCE O2A PA

ACCE O2'

ACCE O3'

ACCE O4'

ACCE O5'

ACCE O3A

ACCE O2B

ACCE O1B

ACCE O3B

ACCE O3G

ACCE O2G

ACCE O1G

IC PA O3A PB O3B 0.0000 000.00 180.0 000.00 0.0000

IC O3A PB O3B PG 0.0000 000.00 180.0 000.00 0.0000

IC PB O3B PG O1G 0.0000 000.00 180.0 000.00 0.0000

IC C5' O5' PA O1A 0.0000 000.00 60.0 000.00 0.0000

IC C5' O5' PA O2A 0.0000 000.00 -60.0 000.00 0.0000

IC PA O3A PB O1B 0.0000 000.00 60.0 000.00 0.0000

IC PA O3A PB O2B 0.0000 000.00 -60.0 000.00 0.0000

IC PB	O3B	PG	O2G	0.0000	000.00	60.0	000.00	0.0000
IC PB	O3B	PG	O3G	0.0000	000.00	-60.0	000.00	0.0000
IC PA	O5'	C5'	C4'	1.5996	119.00	-151.39	110.04	1.5160
IC O5'	C5'	C4'	C3'	1.4401	108.83	-179.85	116.10	1.5284
IC C5'	C4'	C3'	O3'	1.5160	116.10	76.70	115.12	1.4212
IC H3T	O3'	C3'	C4'	0.9650	105.47	38.18	111.98	1.5386
IC O4'	C3'	*C4'	C5'	1.4572	104.06	-120.04	116.10	1.5160
IC C2'	C4'	*C3'	O3'	1.5284	100.16	-124.08	115.12	1.4212
IC C4'	C3'	C2'	C1'	1.5284	100.16	39.58	102.04	1.5251
IC C3'	C2'	C1'	N9	1.5284	101.97	147.80	113.71	1.4896
IC O4'	C1'	N9	C4	1.5251	113.71	-97.2	125.59	1.3783
IC C1'	C4	*N9	C8	1.4896	125.59	-179.99	106.0	1.374
IC C4	N9	C8	N7	1.377	106.0	0.0	113.5	1.304
IC C8	N9	C4	C5	1.374	106.0	0.0	105.6	1.377
IC N9	C5	*C4	N3	1.377	105.6	180.0	128.4	1.355
IC C5	C4	N3	C2	1.377	128.4	0.0	111.8	1.327
IC C4	N3	C2	N1	1.355	111.8	0.0	124.0	1.375
IC N1	N3	*C2	N2	1.375	124.0	180.0	119.7	1.341
IC N3	C2	N2	H21	1.327	119.7	180.0	127.0	1.01
IC H21	C2	*N2	H22	1.01	127.0	-180.0	116.5	1.01
IC N3	C2	N1	C6	1.327	124.0	0.0	124.9	1.393
IC C6	C2	*N1	H1	1.393	124.9	180.0	117.4	1.03
IC C5	N1	*C6	O6	1.415	111.7	180.0	120.0	1.239

```

IC N9 N7 *C8 H8 0.0 0.0 180.0 0.0 0.0
IC C1' C3' *C2' O2' 1.5284 102.04 -114.67 110.81 1.4212
IC H2' O2' C2' C3' 0.9600 114.97 148.63 111.92 1.5284
IC O4' C2' *C1' H1' 0.0 0.0 -115.0 0.0 0.0
IC C1' C3' *C2' H2" 0.0 0.0 115.0 0.0 0.0
IC C2' C4' *C3' H3' 0.0 0.0 115.0 0.0 0.0
IC C3' O4' *C4' H4' 0.0 0.0 -115.0 0.0 0.0
IC C4' O5' *C5' H5' 0.0 0.0 -115.0 0.0 0.0
IC C4' O5' *C5' H5" 0.0 0.0 115.0 0.0 0.0

```

PATC FIRS NONE LAST NONE

Water box generation input file sample

- * File: k12dgtwhvr_water_water_overlay_1st.inp
- * Generates the bulk overlay of water for the system
- * Use simplified method for water addition

!!

! set directory

set type k12dgtwhvr_water

set dir data/muratcioglus/@type

!!

! set initial numbers of system

set bs 100.0 ! cell dimension

! set number of water; read from water directory

set nwt1 8682

set nwt2 8681

set nwt3 8578

set nwt4 8578

!!

bomlev -2

! read topology

open read card unit 10 name /data/muratcioglus/charmm_input/top_all36_ras_c39b2_hbj.inp

read rtf card unit 10

close unit 10

! read parameters

open read card unit 10 name /data/muratcioglus/charmm_input/par_all36_ras_c39b2_hbj.inp

read param flex card unit 10

close unit 10

!!

open read card unit 1 name /@dir/coord/@type_rot_initial_conf.crd

read sequence coord unit 1

generate JANG setup noangle nodihedrals first none last none

close unit 1

open read card unit 1 name /@dir/coord/@type_rot_initial_conf.crd

read coord card unit 1

close unit 1

delete atom sele .not. INIT end

scalar charge set 0.0

set N ?nres

read sequence TIP3 @nwt1

generate OV11 noangle nodihedral

read sequence TIP3 @nwt2

generate OV12 noangle nodihedral

open read card unit 1 name /@dir/water/water_large_box_@bs_1.crd

read coord card unit 1 offset @N

close unit 1

```
set N ?nres
read sequence TIP3 @nwt3
generage OV21 noangle nodihedral
read sequence TIP3 @nwt4
generage OV22 noangle nodihedral
open read card unit 1 name /@dir/water/water_large_box_@bs_2.crd
read coor card unit 1 offset @N
close unit 1
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
! Initial overlay of each water box
! Lower layer
coor stat sele segid OV1* end
define COUNT1 sele ((type OH2 .and. segid OV1*) .and. (segid JANG .around. 2.6)) end
delete atoms sele .byres. COUNT1 end
! Upper layer
coor stat sele segid OV2* end
define COUNT2 sele ((type OH2 .and. segid OV2*) .and. (segid JANG .around. 2.6)) end
delete atoms sele .byres. COUNT2 end
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
delete atoms sele segid JANG end
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
! Write water coordinates
open write card unit 1 name /@dir/coor/temp_water_ov11.crd
write coor sele segid OV11 end card unit 1
* overlay of waters
delete atoms sele segid OV11 end
open write card unit 1 name /@dir/coor/temp_water_ov12.crd
write coor sele segid OV12 end card unit 1
```

```
* overlay of waters
delete atoms sele segid OV12 end
! Write water coordinates
open write card unit 1 name /@dir/coor/temp_water_ov21.crd
write coor sele segid OV21 end card unit 1
* overlay of waters
delete atoms sele segid OV21 end
! Write water coordinates
open write card unit 1 name /@dir/coor/temp_water_ov22.crd
write coor sele segid OV22 end card unit 1
* overlay of waters
stop
```

Salt addition input file sample

```
* File: k12dgtwhvr_water_add_salt_mk_coor.inp
* Add salts and make coordinates
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
! set directory
set type k12dgtwhvr_water
set dir data/muratcioglus/@type
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
! set initial numbers of system
set nna 36      ! number of Na+
set ncl 30      ! number of Cl-
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
bomlev -2
! read topology
open read card unit 10 name /data/muratcioglus/charmm_input/top_all36_ras_c39b2_hbj.inp
```

```
read rtf card unit 10
close unit 10
! read parameters
open read card unit 10 name /data/muratcioglus/charmm_input/par_all36_ras_c39b2_hbj.inp
read param flex card unit 10
close unit 10
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
! Get k-ras (1-166) coordinates
read sequence card
* The sequence is RasK 1 to 166
166
MET THR GLU TYR LYS LEU VAL VAL VAL GLY ALA ASP GLY
VAL GLY LYS SER ALA LEU THR ILE GLN LEU ILE GLN ASN
HSD PHE VAL ASP GLU TYR ASP PRO THR ILE GLU ASP SER
TYR ARG LYS GLN VAL VAL ILE ASP GLY GLU THR CYS LEU
LEU ASP ILE LEU ASP THR ALA GLY GLN GLU GLU TYR SER
ALA MET ARG ASP GLN TYR MET ARG THR GLY GLU GLY PHE
LEU CYS VAL PHE ALA ILE ASN ASN THR LYS SER PHE GLU
ASP ILE HSD HSD TYR ARG GLU GLN ILE LYS ARG VAL LYS
ASP SER GLU ASP VAL PRO MET VAL LEU VAL GLY ASN LYS
CYS ASP LEU PRO SER ARG THR VAL ASP THR LYS GLN ALA
GLN ASP LEU ALA ARG SER TYR GLY ILE PRO PHE ILE GLU
THR SER ALA LYS THR ARG GLN GLY VAL ASP ASP ALA PHE
TYR THR LEU VAL ARG GLU ILE ARG LYS HSD

generate RASK first NTER last CT1 setup
! Get GTP in crystal structure
read sequence GTP 1
```

generate GTP setup first none last none

! Get Mg²⁺ in crystal structure

read sequence MG 1

generate MAGN noangle nodihedral

open read card unit 1 name /@dir/coord/@type_water_overlay_system_ras.crd

read coord card unit 1

close unit 1

set N ?nres

! Get Raf-1 RBD (54-131) coordinates

read sequence card

* The sequence is Raf-1 RBD 54 to 131

78

THR SER ASN THR ILE ARG VAL PHE LEU PRO ASN LYS GLN
ARG THR VAL VAL ASN VAL ARG ASN GLY MET SER LEU HSD
ASP CYS LEU MET LYS ALA LEU LYS VAL ARG GLY LEU GLN
PRO GLU CYS CYS ALA VAL PHE ARG LEU LEU HSD GLU HSD
LYS GLY LYS LYS ALA ARG LEU ASP TRP ASN THR ASP ALA
ALA SER LEU ILE GLY GLU GLU LEU GLN VAL ASP PHE LEU

generate CRBD first NTER last CTER setup

open read card unit 1 name /@dir/coord/@type_water_overlay_system_rbd.crd

read coord card unit 1 offset @N

close unit 1

set N ?nres

open read card unit 1 name /@dir/coord/@type_water_overlay_system_bulk_1.crd

read sequence coord unit 1

generate BLK1 nodihedrals noangles first none last none

close unit 1

open read card unit 1 name /@dir/coord/@type_water_overlay_system_bulk_1.crd

```
read coor card unit 1 offset @N
close unit 1
set N ?nres
open read card unit 1 name /@dir/coor/@type_water_overlay_system_bulk_2.crd
read sequence coor unit 1
generate BLK2 nodihedrals noangles first none last none
close unit 1
open read card unit 1 name /@dir/coor/@type_water_overlay_system_bulk_2.crd
read coor card unit 1 offset @N
close unit 1
set N ?nres
open read card unit 1 name /@dir/coor/@type_water_overlay_system_bulk_3.crd
read sequence coor unit 1
generate BLK3 nodihedrals noangles first none last none
close unit 1
open read card unit 1 name /@dir/coor/@type_water_overlay_system_bulk_3.crd
read coor card unit 1 offset @N
close unit 1
set N ?nres
open read card unit 1 name /@dir/coor/@type_water_overlay_system_bulk_4.crd
read sequence coor unit 1
generate BLK4 nodihedrals noangles first none last none
close unit 1
open read card unit 1 name /@dir/coor/@type_water_overlay_system_bulk_4.crd
read coor card unit 1 offset @N
close unit 1
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
! Write protein coordinates for water only
```

```
define JANG sele segid RASK .or. segid GTP .or. segid MAGN end
dele atom sele JANG end
dele atom sele segid CRBD end
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!

! Stat waters
coor stat sele resn TIP3 .and. type OH2 end
coor stat sele segid BLK1 .and. type OH2 end
set nt1 ?nsl
coor stat sele segid BLK2 .and. type OH2 end
set nt2 ?nsl
coor stat sele segid BLK3 .and. type OH2 end
set nt3 ?nsl
coor stat sele segid BLK4 .and. type OH2 end
set nt4 ?nsl
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!

! Randomly choose waters for replacing salts
! random number generation
RAND UNIF ISEED 318098
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!

! Na+ selection, nsod = total # of Na
set nsod @nna
label loop_sod1
    set r ?RAND
    if r gt 0.0 set NAME BLK1
    if r gt 0.25 set NAME BLK2
    if r gt 0.50 set NAME BLK3
    if r gt 0.75 set NAME BLK4
    set n ?RAND
```

```

if @NAME eq BLK1 mult n by @nt1
if @NAME eq BLK2 mult n by @nt2
if @NAME eq BLK3 mult n by @nt3
if @NAME eq BLK4 mult n by @nt4
if n lt 10000 trim n to 4
if n lt 1000 trim n to 3
if n lt 100 trim n to 2
if n lt 10 trim n to 1
if n eq 0 set n 1
! salt selection
print coor sele segid @NAME .and. resid @n .and. type OH2 end
if ?nsel eq 0 goto loop_sod1
! no salts around 4A each other
define temp sele (type CLA .or. type SOD .or. type MG) .and. (segid @NAME .and. resid
@n .and. type OH2) .around. 4.0 end
if ?nsel gt 0 goto loop_sod1
! replace water to salt
delete atom sele segid @NAME .and. type H* .and. resi @n end
renew atom SOD sele segid @NAME .and. type OH2 .and. resi @n end
coor stat sele type SOD end
if ?nsel lt @nsod goto loop_sod1
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
! Cl- selection, ncla = total # of Cl
set ncla @ncl
label loop_cla1
set r ?RAND
if r gt 0.0 set NAME BLK1
if r gt 0.25 set NAME BLK2
if r gt 0.50 set NAME BLK3

```

```

if r gt 0.75 set NAME BLK4
set n ?RAND
if @NAME eq BLK1 mult n by @nt1
if @NAME eq BLK2 mult n by @nt2
if @NAME eq BLK3 mult n by @nt3
if @NAME eq BLK4 mult n by @nt4
if n lt 10000 trim n to 4
if n lt 1000 trim n to 3
if n lt 100 trim n to 2
if n lt 10 trim n to 1
if n eq 0 set n 1
! salt selection
print coor sele segid @NAME .and. resid @n .and. type OH2 end
if ?nsl eq 0 goto loop_cla1
! no salts around 4A each other
define temp sele (type CLA .or. type SOD .or. type MG) .and. (segid @NAME .and. resid
@n .and. type OH2) .around. 4.0 end
if ?nsl gt 0 goto loop_cla1
! replace water to salt
delete atom sele segid @NAME .and. type H* .and. resi @n end
renew atom CLA sele segid @NAME .and. type OH2 .and. resi @n end
coor stat sele type CLA end
if ?nsl lt @ncla goto loop_cla1
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
! Write salt coordinates
open write card unit 1 name /@dir/coor/temp_sod.crd
write coor card unit 1 sele type SOD end
* temp salt coordinates
delete atoms sele type SOD end

```

Psf generation input file sample

[illegible]

```
set type k12dgtwhvr_water
set dir data/muratcioglus/@type
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
bomlev -2
! read topology
open read card unit 10 name /data/muratcioglus/charmm_input/top_all36_ras_c39b2_hbj.inp
read rtf card unit 10
close unit 10
! read parameters
open read card unit 10 name /data/muratcioglus/charmm_input/par_all36_ras_c39b2_hbj.inp
read param flex card unit 10
close unit 10
!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!!
! Get k-ras (1-166) coordinates
read sequence card
* The sequence is RasK 1 to 166
166
MET THR GLU TYR LYS LEU VAL VAL VAL GLY ALA ASP GLY
VAL GLY LYS SER ALA LEU THR ILE GLN LEU ILE GLN ASN
HSD PHE VAL ASP GLU TYR ASP PRO THR ILE GLU ASP SER
TYR ARG LYS GLN VAL VAL ILE ASP GLY GLU THR CYS LEU
LEU ASP ILE LEU ASP THR ALA GLY GLN GLU GLU TYR SER
ALA MET ARG ASP GLN TYR MET ARG THR GLY GLU GLY PHE
LEU CYS VAL PHE ALA ILE ASN ASN THR LYS SER PHE GLU
ASP ILE HSD HSD TYR ARG GLU GLN ILE LYS ARG VAL LYS
ASP SER GLU ASP VAL PRO MET VAL LEU VAL GLY ASN LYS
CYS ASP LEU PRO SER ARG THR VAL ASP THR LYS GLN ALA
GLN ASP LEU ALA ARG SER TYR GLY ILE PRO PHE ILE GLU
```

THR SER ALA LYS THR ARG GLN GLY VAL ASP ASP ALA PHE
TYR THR LEU VAL ARG GLU ILE ARG LYS HSD

generate RASK first NTER last CT1 setup

! Get GTP in crystal structure

read sequence GTP 1

generate GTP setup first none last none

! Get Mg²⁺ in crystal structure

read sequence MG 1

generate MAGN noangle nodihedral

open read card unit 1 name /@dir/coor/@type_water_overlay_system_ras.crd

read coor card unit 1

close unit 1

set N ?nres

! Get Raf-1 RBD (54-131) coordinates

read sequence card

* The sequence is Raf-1 RBD 54 to 131

78

THR SER ASN THR ILE ARG VAL PHE LEU PRO ASN LYS GLN
ARG THR VAL VAL ASN VAL ARG ASN GLY MET SER LEU HSD
ASP CYS LEU MET LYS ALA LEU LYS VAL ARG GLY LEU GLN
PRO GLU CYS CYS ALA VAL PHE ARG LEU LEU HSD GLU HSD
LYS GLY LYS LYS ALA ARG LEU ASP TRP ASN THR ASP ALA
ALA SER LEU ILE GLY GLU GLU LEU GLN VAL ASP PHE LEU

generate CRBD first NTER last CTER setup

open read card unit 1 name /@dir/coor/@type_water_overlay_system_rbd.crd

read coor card unit 1 offset @N

close unit 1

```
set N ?nres
open read card unit 1 name /@dir/coord/temp_blk1.crd
read sequence coord unit 1
generate BLK1 nodihedrals noangles first none last none
close unit 1
open read card unit 1 name /@dir/coord/temp_blk1.crd
read coord card unit 1 offset @N
close unit 1
set N ?nres
open read card unit 1 name /@dir/coord/temp_blk2.crd
read sequence coord unit 1
generate BLK2 nodihedrals noangles first none last none
close unit 1
open read card unit 1 name /@dir/coord/temp_blk2.crd
read coord card unit 1 offset @N
close unit 1

set N ?nres
open read card unit 1 name /@dir/coord/temp_blk3.crd
read sequence coord unit 1
generate BLK3 nodihedrals noangles first none last none
close unit 1
open read card unit 1 name /@dir/coord/temp_blk3.crd
read coord card unit 1 offset @N
close unit 1
set N ?nres
open read card unit 1 name /@dir/coord/temp_blk4.crd
read sequence coord unit 1
```

```
generate BLK4 nodihedrals noangles first none last none
close unit 1
open read card unit 1 name /@dir/coord/temp_blk4.crd
read coord card unit 1 offset @N
close unit 1
set N ?nres
open read card unit 1 name /@dir/coord/temp_sod.crd_mod
read sequence coord unit 1
generate SODI noangle nodihedral
close unit 1
open read card unit 1 name /@dir/coord/temp_sod.crd_mod
read coord card unit 1 offset @N
close unit 1
set N ?nres
open read card unit 1 name /@dir/coord/temp_cla.crd_mod
read sequence coord unit 1
generate CHLO noangle nodihedral
close unit 1
open read card unit 1 name /@dir/coord/temp_cla.crd_mod
read coord card unit 1 offset @N
close unit 1
!!!!!!!!!!!!!!!!!!!!!!
! Stat for protein
coord stat sele segid RASK end
coord stat sele segid RASK .and. type CA end
coord stat sele segid CRBD end
coord stat sele segid CRBD .and. type CA end
coord stat sele segid GTP end
```

```
! Stat for salt
coor stat sele segid MAGN end
coor stat sele resn SOD end
coor stat sele resn CLA end

! Stat for water
coor stat sele segid BLK1 .and. type OH2 end
coor stat sele segid BLK2 .and. type OH2 end
coor stat sele segid BLK3 .and. type OH2 end
coor stat sele segid BLK4 .and. type OH2 end
!!!!!!!!!!!!!!!!!!!!
!IOFormat NOEXtended
!!!!!!!!!!!!!!!!!!!!
! Write psf file in charmm format
open write card unit 1 name /@dir/psf/@type_namd_salt_charmm.psf
write psf card unit 1
close unit 1

! Write psf file in charmm format
open write card unit 1 name /@dir/psf/@type_namd_salt_xplor.psf
write psf card xplor unit 1
close unit 1
!!!!!!!!!!!!!!!!!!!!

! Get the final coordinates in pdb
open write form unit 1 name /@dir/coor/@type_namd_salt_charmm.pdb
write coor pdb unit 1
close unit 1
stop
```

A pre-equilibrium configuration file sample

```
#####  
# FILE: k12dgtwhvr_water_namd_preequ_mini.cnfg #  
# Minimization of solvent with protein fix #  
#####  
#####  
#####  
# Set variable values #  
#####  
set temp 310  
set dir /data/muratcioglus/k12dgtwhvr_water  
set type k12dgtwhvr_water  
#####  
# Parameters #  
#####  
paraTypeCharmm on  
parameters /data/muratcioglus/charmm_input/par_all36_ras_c39b2_hbj.inp  
#####  
# Input files #  
#####  
structure ${dir}/psf/${type}_namd_salt_xplor.psf ;# Name of structure file  
coordinates ${dir}/coord/${type}_namd_salt_charmm.pdb ;# Name of input pdb file  
#####  
# Restart files #
```

```

#####
#####

# NOTE: Do not set the initial restart velocity if you have also specified a temperature!

if {0} {
  velocities      ${dir}/vel/                ;# Initial velocities in pdb file
  binVelocities   ${dir}/vel/                ;# Initial velocities in binary format
                                     ;# Remove the "temperature" entry if you use this!

  binCoordinates  ${dir}/coor/               ;# Initial coor in binary format
  extendedSystem  ${dir}/xsc/               ;# Input periodic cell para read from
}

#####
#####

# Simulation temperature                #

#####
#####

# NOTE: Do not set the initial velocity temperature if you have also specified a .vel restart
file!

temperature      $temp

#####
#####

# Output files                        #

#####
#####

outputName       ${dir}/coor/${type}_namd_preequ_mini_pdb ;# Name of final output
coor file

binaryoutput     no                        ;# Save final coor in binary format

restartname      ${dir}/coor/${type}_namd_preequ_mini    ;# Name of output(restart)
coor & vel file

binaryrestart     yes                      ;# Save restart file in binary format

restartfreq      10000                    ;# Freq saving coor & vel (1000 steps = every
1ps)

```

```

restartsave      yes                      ;# Append time step at end of file name
dcdfile          ${dir}/trj/${type}_namd_preequ_mini.dcd    ;# Name of trajectory file
dcdfreq          100                      ;# Freq saving trajectory
dcdUnitCell      yes                      ;# Write unit cell data to dcd file
#velcdcdfile     ${dir}/vel/${type}_namd_preequ_mini_vel.dcd ;# Name of vel trajectory
file
#velcdcfreq      500                      ;# Freq saving vel trajectory
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%#
# Output configurations                                     #
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%#
outputEnergies   100                      ;# Freq for printing Energies
outputTiming     100                      ;# Freq for printing Timing info
outputPressure   100                      ;# Freq for printing Pressure
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%#
# Periodic Boundary conditions                             #
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%#
# NOTE: Do not set the periodic cell basis if you have specified an .xsc restart file!
if {1} {
cellBasisVector1 100.0 0. 0.
cellBasisVector2 0. 100.0 0.
cellBasisVector3 0. 0. 100.0
cellOrigin       0. 0. 0.
}
#xstFile          ${dir}/xsc/${type}_namd_preequ_mini.xst    ;# Name of cell param
trajectory file
#xstfreq          100                      ;# Freq for writing cell param trajectory

```

```

wrapWater      on                      ;# Wrap water around periodic boundary
wrapAll        on                      ;# Wrap all atoms around periodic boundary
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
# Force-field parameters                #
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
cutoff         12.                     ;# Local interaction dist to elec & VdW
calculation

switching      on                      ;# Smoothing func for cutoff, otherwise
truncated

switchdist     10.                     ;# Should be less than or equal to cutoff value
pairlistdist   14.                     ;# Cutoff for generating pair list, larger than
cutoff

exclude        scaled1-4               ;# Exclude pairs in nonbonding interactions
1-4scaling     1.0                     ;# Scaling factor for 1-4 interaction
COMmotion      no                      ;# No center of mass motion for entire
system

splitPatch     hydrogen                ;# Set hydrogen to allow faster distance
check

hgroupCutoff   2.5                     ;# Use for group-based distance testing
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
# Time step parameters                  #
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
timestep       2.                      ;# 2 fs/step
firsttimestep   0                      ;# Starting time step value
stepspcycle    20                      ;# Freq updating nonbond list
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
# Multiple time step parameters for long range elec  #

```

```

#####
#####

fullElectFrequency  1                ;# Calc long range elec stat every time step
nonbondedFreq       1                ;# Freq for calc short range nonbond interac
MTSAlgorithm        impulse           ;# Integr method for long range elec
longsplitting       c1                ;# Method for split elec forces between long
and short

#####
#####

# PME parameters                                #
#####
#####

if {0} {
  PME          no                ;# Use particle mesh Ewald for electrostatics
  PMETolerance  1.0e-6           ;# Affect value of Ewald coefficient
  PMEInterpOrder 4                ;# For cubic 4
  PMEGridSizeX   92              ;# Number of grid ponts in x
  PMEGridSizeY   92              ;# Number of grid ponts in y
  PMEGridSizeZ   92              ;# Number of grid ponts in z
}

#####
#####

# Constant temperature control                    #
#####
#####

if {0} {
  langevin      on                ;# Use langeving dynam for temp control
  langevinTemp   $temp
  langevinDamping 2.5              ;# Friction coeff (gamma) (1/ps), use a
small value
  langevinHydrogen off            ;# Don't couple langevin bath to hydrogens

```

```

}

#####

# Temperature control for heating or cooling          #

#####

if {0} {
    reassignFreq      50000                ;# Frequency to reassign temperature
    reassignTemp       10.0                ;# Initial temperature
    reassignIncr       30.0                ;# Temperature increment or decrement
    reassignHold       $temp               ;# Target temperature
}

#####

# Constant pressure control                          #

#####

useGroupPressure      yes                  ;# Should use w/SHAKE, needed for 2fs
steps

useFlexibleCell       no                  ;# Allows dimensions of cell to fluctuate
                                       ;# no for water box, yes for membrane

useConstantArea       no                  ;# x-y fixed and z dim can vary
                                       ;# no for water box, yes for membrane

if {1} {
    LangevinPiston     on                  ;# Use Nose-Hoover Langevin Piston
    Algorithm

    LangevinPistonTarget 1.01325          ;# Target pressure to run simu at (bars)

    LangevinPistonPeriod 200              ;# Oscillation period

    LangevinPistonDecay 100               ;# Damping time scale

    LangevinPistonTemp  $temp             ;# Should be set equal to the target
    temperature

```

```

SurfaceTensionTarget 0.0                                ;# Apply surface tension in xy plane
(dyn/cm)

}

#####

# Fix hydrogens                                         #

#####

if {1} {
  rigidBonds      all                                ;# Fix h-bonds to none, water, or all h-atoms
  rigidTolerance   1.0e-8                            ;# Allowable bond-length error for ShakeH
  rigidIterations  100                                ;# Maximum ShakeH iterations
  rigidDieOnError   on                                ;# Exit and report an error
  useSettle        off                                ;# Keep waters rigid
}

#####

# Harmonic constraints : constraint protein backbone with k=5 #

#####

if {1} {
  constraints      on                                ;# Turn on the use of constraints
  consexp          2                                ;# Exponent to be use in harmonic constraint
  energy

  consref          ${dir}/cons/${type}_preequ_constr_atoms_ref_5.pdb ;# A pdb file to use for
  reference positions

  conskfile        ${dir}/cons/${type}_preequ_constr_atoms_kfile_5.pdb ;# A pdb file to use for
  k values

  conskcol         B                                ;# Column of pdb file to use for harmonic
  constraint

```

```
}

#####

# Fixed atoms constraint : fix protein                                #

#####

if {1} {

  fixedAtoms      on                                ;# Turn on the use of constraints

  fixedAtomsForces on                                ;# Calculate forces between fixed atoms

  fixedAtomsFile   ${dir}/cons/${type}_preequ_fix_protein.pdb    ;# A pdb file to use for
  reference positions

  fixedAtomsCol    O                                ;# Set O to read occupancy column in pdb
  file

}

#####

# Minimization parameters                                           #

#####

if {1} {

  minimization      on                                ;# Perform conjugate gradient energy
  minimization

  minTinyStep       1.0e-8                            ;# First initial step for line minimizer

  minBabyStep       1.0e-2                            ;# Max initial step for line minimizer

  minLineGoal       1.0e-4                            ;# Gradient reduction factor for line
  minimizer

}

#####
```

```
# Extra parameters                                     #
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

# Put here any custom parameters that are specific to this job

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

# Execution script                                     #
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

#-----#
minimize          10000                               ;# Minimization step
#-----#
#-----#
#run              10000                               ;# Dynamics step
#-----#

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

# End of script
```

Product run configuration file sample

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

# FILE: k12dgtwhvr_water_namd_prod_10_ns.cnfg          #

# 10 ns dynamics and production                          #
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

# Set variable values                                  #
```

```
#####  
#####  
  
set temp          310  
  
set dir           /data/muratcioglus/k12dgtwhvr_water  
  
set type          k12dgtwhvr_water  
  
#####  
#####  
  
# Parameters                                     #  
  
#####  
#####  
  
paraTypeCharmm    on  
  
parameters        /data/muratcioglus/charmm_input/par_all36_ras_c39b2_hbj.inp  
  
#####  
#####  
  
# Input files                                     #  
  
#####  
#####  
  
structure          ${dir}/psf/${type}_namd_salt_xplor.psf    ;# Name of structure file  
  
coordinates        ${dir}/coord/${type}_namd_preequ_ewald_last_pdb.coor ;# Name of input pdb  
file  
  
#####  
#####  
  
# Restart files                                   #  
  
#####  
#####  
  
# NOTE: Do not set the initial restart velocity if you have also specified a temperature!  
  
if {1} {  
  
#velocities        ${dir}/vel/                               ;# Initial velocities in pdb file  
  
binVelocities      ${dir}/vel/${type}_namd_preequ_ewald_last.50000.vel ;# Initial velocities in  
binary format  
  
                                ;# Remove the "temperature" entry if you use this!
```

```

binCoordinates ${dir}/coor/${type}_namd_preequ_ewald_last.50000.coor ;# Initial coor in
binary format

extendedSystem ${dir}/xsc/${type}_namd_preequ_ewald_last.50000.xsc ;# Input periodic
cell para read from

}

#####

# Simulation temperature                                #

#####

# NOTE: Do not set the initial velocity temperature if you have also specified a .vel restart
file!

#temperature      $temp

#####

# Output files                                           #

#####

outputName        ${dir}/coor/${type}_namd_prod_10_ns_pdb ;# Name of final output
coor file

binaryoutput      no ;# Save final coor in binary format

restartname       ${dir}/coor/${type}_namd_prod_10_ns ;# Name of output(restart)
coor & vel file

binaryrestart      yes ;# Save restart file in binary format

restartfreq       500000 ;# Freq saving coor & vel (1000 steps = every
1ps)

restartsave       yes ;# Append time step at end of file name

dcdfile           ${dir}/trj/${type}_namd_prod_merged_10_ns.dcd ;# Name of trajectory
file

dcdfreq           50000 ;# Freq saving trajectory

dcdUnitCell       yes ;# Write unit cell data to dcd file

```

```

#velcdfile      ${dir}/vel/${type}_namd_prod_10_ns_vel.dcd ;# Name of vel trajectory
file

#velcdfreq      1000                                     ;# Freq saving vel trajectory

#####

# Output configurations                                     #

#####

outputEnergies  50000                                     ;# Freq for printing Energies
outputTiming    50000                                     ;# Freq for printing Timing info
outputPressure  50000                                     ;# Freq for printing Pressure

#####

# Periodic Boundary conditions                             #

#####

# NOTE: Do not set the periodic cell basis if you have specified an .xsc restart file!
if {0} {
cellBasisVector1 100.0 0. 0.
cellBasisVector2 0. 100.0 0.
cellBasisVector3 0. 0. 100.0
cellOrigin       0. 0. 0.
}

xstFile          ${dir}/xsc/${type}_namd_prod_10_ns.xst ;# Name of cell param
trajectory file

xstfreq          50000                                     ;# Freq for writing cell param trajectory
wrapWater        on                                       ;# Wrap water around periodic boundary
wrapAll          on                                       ;# Wrap all atoms around periodic boundary

```

```

#####
#####
# Force-field parameters                                     #
#####
#####

cutoff          12.                                         ;# Local interaction dist to elec & VdW
calculation

switching       on                                         ;# Smoothing func for cutoff, otherwise
truncated

switchdist      10.                                         ;# Should be less than or equal to cutoff value
pairlistdist    14.                                         ;# Cutoff for generating pair list, larger than
cutoff

exclude         scaled1-4                                   ;# Exclude pairs in nonbonding interactions
1-4scaling      1.0                                         ;# Scaling factor for 1-4 interaction
COMmotion       no                                         ;# No center of mass motion for entire
system

splitPatch      hydrogen                                   ;# Set hydrogen to allow faster distance
check

hgroupCutoff    2.5                                         ;# Use for group-based distance testing
#####
#####

# Time step parameters                                     #
#####
#####

timestep        2.                                         ;# 2 fs/step
firsttimestep    0                                         ;# Starting time step value
stepspcycle     20                                         ;# Freq updating nonbond list
#####
#####

# Multiple time step parameters for long range elec       #
#####

```

```

#####
#####

fullElectFrequency  1                      ;# Calc long range elec stat every time step
nonbondedFreq       1                      ;# Freq for calc short range nonbond interac
MTSAlgorithm        impulse                ;# Integr method for long range elec
longsplitting       c1                    ;# Method for split elec forces between long
and short

#####
#####

# PME parameters                                #
#####
#####

if {1} {
  PME          yes                      ;# Use particle mesh Ewald for electrostatics
  PMEGridSpacing 1.0                  ;# NAMD assigns automatically the grid
space
  #PME Tolerance 1.0e-6                ;# Affect value of Ewald coefficient
  #PME InterpOrder 4                    ;# For cubic 4
  #PME GridSizeX 92                     ;# Number of grid ponts in x
  #PME GridSizeY 92                     ;# Number of grid ponts in y
  #PME GridSizeZ 92                     ;# Number of grid ponts in z
}

#####
#####

# Constant temperature control                #
#####
#####

if {1} {
  langevin      on                      ;# Use langeving dynam for temp control
  langevinTemp  $temp

```

```

langevinDamping    2.5                ;# Friction coeff (gamma) (1/ps), use a
small value

langevinHydrogen    off                ;# Don't couple langevin bath to hydrogens
}

#####
# Temperature control for heating or cooling                #
#####
#####
if {0} {
reassignFreq        50000              ;# Frequency to reassign temperature
reassignTemp        10.0               ;# Initial temperature
reassignIncr        30.0               ;# Temperature increment or decrement
reassignHold        $temp              ;# Target temperature
}
#####
# Constant pressure control                                #
#####
#####
useGroupPressure    yes                ;# Should use w/SHAKE, needed for 2fs
steps

useFlexibleCell      no                ;# Allows dimensions of cell to fluctuate
                                   ;# no for water box, yes for membrane

useConstantArea      no                ;# x-y fixed and z dim can vary
                                   ;# no for water box, yes for membrane

if {1} {
LangevinPiston      on                ;# Use Nose-Hoover Langeving Piston
Algorithm

LangevinPistonTarget 1.01325          ;# Target pressure to run simu at (bars)

```

```

LangevinPistonPeriod 200                ;# Oscillation period
LangevinPistonDecay 100                  ;# Damping time scale
LangevinPistonTemp $temp                 ;# Should be set equal to the target
temperature
SurfaceTensionTarget 0.0                 ;# Apply surface tension in xy plane
(dyn/cm)
}

#####
#####

# Fix hydrogens                          #

#####
#####

if {1} {
  rigidBonds      all                    ;# Fix h-bonds to none, water, or all h-atoms
  rigidTolerance  1.0e-8                 ;# Allowable bond-length error for ShakeH
  rigidIterations 100                    ;# Maximum ShakeH iterations
  rigidDieOnError  on                    ;# Exit and report an error
  useSettle       off                    ;# Keep waters rigid
}

#####
#####

# Harmonic constraints                    #

#####
#####

if {0} {
  constraints      on                    ;# Turn on the use of constraints
  consexp          2                    ;# Exponent to be use in harmonic constraint
  energy
  consref          ${dir}/cons/         ;# A pdb file to use for reference positions
  conskfile        ${dir}/cons/         ;# A pdb file to use for k values
}

```

```

conskcol      B                      ;# Column of pdb file to use for harmonic
constraint

}

#####
#####

# Fixed atoms constraint                      #

#####

if {0} {
    fixedAtoms      on                      ;# Turn on the use of constraints
    fixedAtomsForces off                    ;# Calculate forces between fixed atoms
    fixedAtomsFile   ${dir}/coord/          ;# A pdb file to use for reference
    positions
    fixedAtomsCol    O                      ;# Set O to read occupancy column in pdb
    file
}

#####

# Minimization parameters                      #

#####

if {0} {
    minimization      on                      ;# Perform conjugate gradient energy
    minimization
    minTinyStep        1.0e-8                ;# First initial step for line minimizer
    minBabyStep         1.0e-2                ;# Max initial step for line minimizer
    minLineGoal         1.0e-4                ;# Gradient reduction factor for line
    minimizer
}

#####
#####

```


BIBLIOGRAPHY

1. Lu, S., et al., *The structural basis of oncogenic mutations G12, G13 and Q61 in small GTPase K-Ras4B*. Scientific reports, 2016. **6**: p. 21949.
2. Kirsten, W. and L.A. Mayer, *Morphologic responses to a murine erythroblastosis virus*. Journal of the National Cancer Institute, 1967. **39**(2): p. 311-335.
3. Harvey, J., *An unidentified virus which causes the rapid production of tumours in mice*. Nature, 1964. **204**(4963): p. 1104.
4. Shih, T.Y., et al., *Identification of a sarcoma virus-coded phosphoprotein in nonproducer cells transformed by Kirsten or Harvey murine sarcoma virus*. Virology, 1979. **96**(1): p. 64-79.
5. Scolnick, E.M., A.G. Papageorge, and T.Y. Shih, *Guanine nucleotide-binding activity as an assay for src protein of rat-derived murine sarcoma viruses*. Proceedings of the National Academy of Sciences, 1979. **76**(10): p. 5355-5359.
6. Scolnick, E.M. and W.P. Parks, *Harvey sarcoma virus: a second murine type C sarcoma virus with rat genetic information*. Journal of virology, 1974. **13**(6): p. 1211-1219.
7. Willingham, M.C., et al., *Localization of the src gene product of the Harvey strain of MSV to plasma membrane of transformed cells by electron microscopic immunocytochemistry*. Cell, 1980. **19**(4): p. 1005-1014.
8. Hall, A., et al., *Identification of transforming gene in two human sarcoma cell lines as a new member of the ras gene family located on chromosome 1*. Nature, 1983. **303**(5916): p. 396.
9. Reuther, G.W. and C.J. Der, *The Ras branch of small GTPases: Ras family members don't fall far from the tree*. Current opinion in cell biology, 2000. **12**(2): p. 157-165.
10. Campbell, S.L., et al., *Increasing complexity of Ras signaling*. Oncogene, 1998. **17**(11): p. 1395.
11. Colicelli, J., *Human RAS superfamily proteins and related GTPases*. Sci. STKE, 2004. **2004**(250): p. re13-re13.
12. Alberts, B., *Molecular biology of the cell*. 2017: Garland science.
13. Downward, J., *Ras signalling and apoptosis*. Current opinion in genetics & development, 1998. **8**(1): p. 49-54.
14. Khosravi-Far, R., et al., *Increasing complexity of Ras signal transduction: involvement of Rho family proteins*, in *Advances in cancer research*. 1997, Elsevier. p. 57-107.
15. Castellano, E. and E. Santos, *Functional specificity of ras isoforms: so similar but so different*. Genes & cancer, 2011. **2**(3): p. 216-231.
16. Ellis, C.A. and G. Clark, *The importance of being K-Ras*. Cellular signalling, 2000. **12**(7): p. 425-434.
17. Quinlan, M.P. and J. Settleman, *Isoform-specific ras functions in development and cancer*. 2009.
18. Sievers, F., et al., *Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega*. Molecular systems biology, 2011. **7**(1): p. 539.
19. Goujon, M., et al., *A new bioinformatics analysis tools framework at EMBL-EBI*. Nucleic acids research, 2010. **38**(suppl_2): p. W695-W699.
20. Nussinov, R., H. Jang, and C.J. Tsai, *Oligomerization and nanocluster organization render specificity*. Biological Reviews, 2015. **90**(2): p. 587-598.
21. Buday, L. and J. Downward, *Many faces of Ras activation*. Biochimica et Biophysica Acta (BBA)-Reviews on Cancer, 2008. **1786**(2): p. 178-187.

22. Shields, J.M., et al., *Understanding Ras: 'it ain't over'til it's over'*. Trends in cell biology, 2000. **10**(4): p. 147-154.
23. Long, S.B., P.J. Casey, and L.S. Beese, *The basis for K-Ras4B binding specificity to protein farnesyl-transferase revealed by 2 Å resolution ternary complex structures*. Structure, 2000. **8**(2): p. 209-222.
24. Newman, C.M. and A.I. Magee, *Posttranslational processing of the ras superfamily of small GTP-binding proteins*. Biochimica et Biophysica Acta (BBA)-Reviews on Cancer, 1993. **1155**(1): p. 79-96.
25. Reid, T.S., et al., *Crystallographic analysis of CaaX prenyltransferases complexed with substrates defines rules of protein substrate selectivity*. Journal of molecular biology, 2004. **343**(2): p. 417-433.
26. Willumsen, B.M., et al., *The p21 ras C-terminus is required for transformation and membrane association*. Nature, 1984. **310**(5978): p. 583.
27. Wright, L.P. and M.R. Philips, *Thematic review series: lipid posttranslational modifications CAAX modification and membrane targeting of Ras*. Journal of lipid research, 2006. **47**(5): p. 883-891.
28. Hancock, J.F., H. Paterson, and C.J. Marshall, *A polybasic domain or palmitoylation is required in addition to the CAAX motif to localize p21ras to the plasma membrane*. Cell, 1990. **63**(1): p. 133-139.
29. Laude, A.J. and I.A. Prior, *Palmitoylation and localisation of RAS isoforms are modulated by the hypervariable linker domain*. J Cell Sci, 2008. **121**(4): p. 421-427.
30. Apolloni, A., et al., *H-ras but not K-ras traffics to the plasma membrane through the exocytic pathway*. Molecular and cellular biology, 2000. **20**(7): p. 2475-2487.
31. Choy, E., et al., *Endomembrane trafficking of ras: the CAAX motif targets proteins to the ER and Golgi*. Cell, 1999. **98**(1): p. 69-80.
32. Hancock, J.F. and G. Robert, *Ras plasma membrane signalling platforms*. Biochemical Journal, 2005. **389**(1): p. 1-11.
33. Plowman, S.J., et al., *H-ras, K-ras, and inner plasma membrane raft proteins operate in nanoclusters with differential dependence on the actin cytoskeleton*. Proceedings of the National Academy of Sciences, 2005. **102**(43): p. 15500-15505.
34. Rotblat, B., et al., *Three separable domains regulate GTP-dependent association of H-ras with the plasma membrane*. Molecular and cellular biology, 2004. **24**(15): p. 6799-6810.
35. Eisenberg, S., et al., *Clustering of raft-associated proteins in the external membrane leaflet modulates internal leaflet H-ras diffusion and signaling*. Molecular and cellular biology, 2006. **26**(19): p. 7190-7200.
36. Niv, H., et al., *Activated K-Ras and H-Ras display different interactions with saturable nonraft sites at the surface of live cells*. The Journal of cell biology, 2002. **157**(5): p. 865-872.
37. Prior, I.A., et al., *Direct visualization of Ras proteins in spatially distinct cell surface microdomains*. The Journal of cell biology, 2003. **160**(2): p. 165-170.
38. Roy, S., et al., *Individual palmitoyl residues serve distinct roles in H-ras trafficking, microlocalization, and signaling*. Molecular and cellular biology, 2005. **25**(15): p. 6722-6733.
39. Zhou, Y. and J.F. Hancock, *Ras nanoclusters: Versatile lipid-based signaling platforms*. Biochimica et Biophysica Acta (BBA)-Molecular Cell Research, 2015. **1853**(4): p. 841-849.
40. Koide, H., et al., *GTP-dependent association of Raf-1 with Ha-Ras: identification of Raf as a target downstream of Ras in mammalian cells*. Proceedings of the National Academy of Sciences, 1993. **90**(18): p. 8683-8686.
41. Herrmann, C., *Ras-effector interactions: after one decade*. Current opinion in structural biology, 2003. **13**(1): p. 122-129.

42. Spoerner, M., et al., *Dynamic properties of the Ras switch I region and its importance for binding to effectors*. Proceedings of the National Academy of Sciences, 2001. **98**(9): p. 4944-4949.
43. Takai, Y., T. Sasaki, and T. Matozaki, *Small GTP-binding proteins*. Physiological reviews, 2001. **81**(1): p. 153-208.
44. Downward, J., *Targeting RAS signalling pathways in cancer therapy*. Nature Reviews Cancer, 2003. **3**(1): p. 11.
45. Kiel, C., et al., *Recognizing and defining true Ras binding domains II: in silico prediction based on homology modelling and energy calculations*. Journal of molecular biology, 2005. **348**(3): p. 759-775.
46. Wohlgemuth, S., et al., *Recognizing and defining true Ras binding domains I: biochemical analysis*. Journal of molecular biology, 2005. **348**(3): p. 741-758.
47. Moodie, S.A., et al., *Complexes of Ras. GTP with Raf-1 and mitogen-activated protein kinase kinase*. Science, 1993. **260**(5114): p. 1658-1661.
48. Roskoski Jr, R., *RAF protein-serine/threonine kinases: structure and regulation*. Biochemical and biophysical research communications, 2010. **399**(3): p. 313-317.
49. Warne, P.H., P.R. Vician, and J. Downward, *Direct interaction of Ras and the amino-terminal region of Raf-1 in vitro*. Nature, 1993. **364**(6435): p. 352.
50. Pacold, M.E., et al., *Crystal structure and functional analysis of Ras binding to its effector phosphoinositide 3-kinase γ* . Cell, 2000. **103**(6): p. 931-944.
51. Rodriguez-Viciana, P. and F. McCormick, *RaIGDS comes of age*. Cancer Cell, 2005. **7**(3): p. 205-6.
52. Datta, S.R., A. Brunet, and M.E. Greenberg, *Cellular survival: a play in three Akts*. Genes & development, 1999. **13**(22): p. 2905-2927.
53. Khwaja, A., et al., *Matrix adhesion and Ras transformation both activate a phosphoinositide 3-OH kinase and protein kinase B/Akt cellular survival pathway*. The EMBO journal, 1997. **16**(10): p. 2783-2793.
54. Feig, L.A., T. Urano, and S. Cantor, *Evidence for a Ras/Ral signaling cascade*. Trends in biochemical sciences, 1996. **21**(11): p. 438-441.
55. Hofer, F., et al., *Activated Ras interacts with the Ral guanine nucleotide dissociation stimulator*. Proceedings of the National Academy of Sciences, 1994. **91**(23): p. 11089-11093.
56. Leever, S.J., H.F. Paterson, and C.J. Marshall, *Requirement for Ras in Raf activation is overcome by targeting Raf to the plasma membrane*. Nature, 1994. **369**(6479): p. 411.
57. Marais, R., et al., *Ras recruits Raf-1 to the plasma membrane for activation by tyrosine phosphorylation*. The EMBO journal, 1995. **14**(13): p. 3136-3145.
58. Yordy, J.S. and R.C. Muise-Helmericks, *Signal transduction and the Ets family of transcription factors*. Oncogene, 2000. **19**(55): p. 6503.
59. Vojtek, A.B., S.M. Hollenberg, and J.A. Cooper, *Mammalian Ras interacts directly with the serine/threonine kinase Raf*. Cell, 1993. **74**(1): p. 205-214.
60. Chuang, E., et al., *Critical binding and regulatory interactions between Ras and Raf occur through a small, stable N-terminal domain of Raf and specific Ras effector residues*. Molecular and cellular biology, 1994. **14**(8): p. 5318-5325.
61. Ghosh, S. and R.M. Bell, *Identification of discrete segments of human Raf-1 kinase critical for high affinity binding to Ha-Ras*. Journal of Biological Chemistry, 1994. **269**(49): p. 30785-30788.
62. Gohlke, H., C. Kiel, and D.A. Case, *Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RaIGDS complexes*. Journal of molecular biology, 2003. **330**(4): p. 891-913.

63. Krygowska, A.A. and E. Castellano, *PI3K: a crucial piece in the RAS signaling puzzle*. Cold Spring Harbor perspectives in medicine, 2018. **8**(6): p. a031450.
64. Rodriguez-Viciana, P., et al., *Phosphatidylinositol-3-OH kinase direct target of Ras*. Nature, 1994. **370**(6490): p. 527.
65. Vanhaesebroeck, B., P.K. Vogt, and C. Rommel, *PI3K: from the bench to the clinic and back*, in *Phosphoinositide 3-kinase in Health and Disease*. 2010, Springer. p. 1-19.
66. Vanhaesebroeck, B., et al., *The emerging mechanisms of isoform-specific PI3K signalling*. Nature reviews Molecular cell biology, 2010. **11**(5): p. 329.
67. Su, H., et al., *VPS34 acetylation controls its lipid kinase activity and the initiation of canonical and non-canonical autophagy*. Molecular cell, 2017. **67**(6): p. 907-921. e7.
68. Stjepanovic, G., et al., *Vps34 kinase domain dynamics regulate the autophagic PI 3-kinase complex*. Molecular cell, 2017. **67**(3): p. 528-534. e3.
69. Bilanges, B., et al., *Vps34 PI 3-kinase inactivation enhances insulin sensitivity through reprogramming of mitochondrial metabolism*. Nature communications, 2017. **8**(1): p. 1804.
70. Hemmings, B.A. and D.F. Restuccia, *The PI3K-PKB/Akt Pathway*. Cold Spring Harbor perspectives in biology, 2015. **7**(4).
71. Christoforidis, S., et al., *Phosphatidylinositol-3-OH kinases are Rab5 effectors*. Nature cell biology, 1999. **1**(4): p. 249.
72. Dbouk, H.A., et al., *G protein-coupled receptor-mediated activation of p110 δ by G6 γ is required for cellular transformation and invasiveness*. Sci. Signal., 2012. **5**(253): p. ra89-ra89.
73. Fritsch, R., et al., *RAS and RHO families of GTPases directly regulate distinct phosphoinositide 3-kinase isoforms*. Cell, 2013. **153**(5): p. 1050-1063.
74. Guillermet-Guibert, J., et al., *The p110 δ isoform of phosphoinositide 3-kinase signals downstream of G protein-coupled receptors and is functionally redundant with p110 γ* . Proceedings of the National Academy of Sciences, 2008. **105**(24): p. 8292-8297.
75. Castellano, E. and J. Downward, *Role of RAS in the regulation of PI 3-kinase*, in *Phosphoinositide 3-kinase in Health and Disease*. 2010, Springer. p. 143-169.
76. Ong, S., et al., *Stimulation of phosphatidylinositol 3-kinase by fibroblast growth factor receptors is mediated by coordinated recruitment of multiple docking proteins*. Proceedings of the National Academy of Sciences, 2001. **98**(11): p. 6074-6079.
77. Rubio, I., et al., *Interaction of Ras with phosphoinositide 3-kinase gamma*. Biochemical Journal, 1997. **326**(Pt 3): p. 891.
78. Albright, C.F., et al., *Characterization of a guanine nucleotide dissociation stimulator for a ras-related GTPase*. The EMBO journal, 1993. **12**(1): p. 339-347.
79. Quilliam, L.A., J.F. Rebhun, and A.F. Castro, *A growing family of guanine nucleotide exchange factors is responsible for activation of Ras-family GTPases*. 2002.
80. Ferro, E. and L. Trabalzini, *RalGDS family members couple Ras to Ral signalling and that's not all*. Cellular signalling, 2010. **22**(12): p. 1804-1810.
81. Herrmann, C., et al., *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, 1996. **271**(12): p. 6794-6800.
82. Koyama, S., et al., *Ras-interacting domain of RGL blocks Ras-dependent signal transduction in Xenopus oocytes*. FEBS letters, 1996. **380**(1-2): p. 113-117.
83. O'gara, M.J., et al., *Characterization of the Ras binding domain of the RalGDS-related protein, Rlf*. Biochemical and biophysical research communications, 1997. **238**(2): p. 425-429.
84. Wolthuis, R., et al., *RalGDS-like factor (Rlf) is a novel Ras and Rap 1A-associating protein*. Oncogene, 1996. **13**(2): p. 353-362.

85. Geyer, M., et al., *Structure of the Ras-binding domain of RalGEF and implications for Ras binding and signalling*. Nature Structural and Molecular Biology, 1997. **4**(9): p. 694.
86. Kigawa, T., et al., *Solution structure of the Ras-binding domain of RGL*. FEBS letters, 1998. **441**(3): p. 413-418.
87. Pistolesi, S., et al., *Molecular motion of spin labeled side chains in the C-terminal domain of RGL2 protein: A SDSL-EPR and MD study*. Biophysical chemistry, 2006. **123**(1): p. 49-57.
88. Ponting, C.P. and D.R. Benjamin, *A novel family of ras-binding domains*. Trends in biochemical sciences, 1996. **21**(11): p. 422-425.
89. Huang, L., et al., *Structural basis for the interaction of Ras with RalGDS*. Nature Structural and Molecular Biology, 1998. **5**(6): p. 422.
90. Huang, L., et al., *Three-dimensional structure of the Ras-interacting domain of RalGDS*. Nature structural biology, 1997. **4**(8): p. 609-615.
91. Vetter, I.R., et al., *Structural and biochemical analysis of Ras-effector signaling via RalGDS*. FEBS letters, 1999. **451**(2): p. 175-180.
92. De Ruiter, N.D., B.M. Burgering, and J.L. Bos, *Regulation of the Forkhead transcription factor AFX by Ral-dependent phosphorylation of threonines 447 and 451*. Molecular and cellular biology, 2001. **21**(23): p. 8225-8235.
93. Bos, J.L., *All in the family? New insights and questions regarding interconnectivity of Ras, Rap1 and Ral*. The EMBO journal, 1998. **17**(23): p. 6776-6782.
94. Pylayeva-Gupta, Y., E. Grabocka, and D. Bar-Sagi, *RAS oncogenes: weaving a tumorigenic web*. Nature Reviews Cancer, 2011. **11**(11): p. 761.
95. Simanshu, D.K., D.V. Nissley, and F. McCormick, *RAS proteins and their regulators in human disease*. Cell, 2017. **170**(1): p. 17-33.
96. Zhou, B., C.J. Der, and A.D. Cox. *The role of wild type RAS isoforms in cancer*. in *Seminars in cell & developmental biology*. 2016. Elsevier.
97. Janakiraman, M., et al., *Genomic and biological characterization of exon 4 KRAS mutations in human cancer*. Cancer research, 2010. **70**(14): p. 5901-5911.
98. Cox, A.D. and C.J. Der, *Ras history: The saga continues*. Small GTPases, 2010. **1**(1): p. 2-27.
99. Fernández-Medarde, A. and E. Santos, *Ras in cancer and developmental diseases*. Genes & cancer, 2011. **2**(3): p. 344-358.
100. Rahib, L., et al., *Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States*. Cancer research, 2014.
101. Raphael, B.J., et al., *Integrated genomic characterization of pancreatic ductal adenocarcinoma*. Cancer cell, 2017. **32**(2): p. 185-203. e13.
102. Almoguera, C., et al., *Most human carcinomas of the exocrine pancreas contain mutant cK-ras genes*. Cell, 1988. **53**(4): p. 549-554.
103. Cox, A.D., et al., *Drugging the undruggable RAS: Mission possible?* Nat Rev Drug Discov, 2014. **13**(11): p. 828-51.
104. Choi, M., et al. *Challenges in Ras therapeutics in pancreatic cancer*. in *Seminars in cancer biology*. 2017. Elsevier.
105. He, B., et al., *RAM2, an essential gene of yeast, and RAM1 encode the two polypeptide components of the farnesyltransferase that prenylates a-factor and Ras proteins*. Proceedings of the National Academy of Sciences, 1991. **88**(24): p. 11373-11377.
106. Reiss, Y., et al., *Inhibition of purified p21ras farnesyl: protein transferase by Cys-AAX tetrapeptides*. Cell, 1990. **62**(1): p. 81-88.
107. Berndt, N., A.D. Hamilton, and S.M. Sebti, *Targeting protein prenylation for cancer therapy*. Nature Reviews Cancer, 2011. **11**(11): p. 775.

108. James, G., J.L. Goldstein, and M.S. Brown, *Resistance of K-RasBV12 proteins to farnesyltransferase inhibitors in Rat1 cells*. Proceedings of the National Academy of Sciences, 1996. **93**(9): p. 4454-4458.
109. Whyte, D.B., et al., *K-and N-Ras are geranylgeranylated in cells treated with farnesyl protein transferase inhibitors*. Journal of Biological Chemistry, 1997. **272**(22): p. 14459-14464.
110. Sun, J., et al., *Both farnesyltransferase and geranylgeranyltransferase I inhibitors are required for inhibition of oncogenic K-Ras prenylation but each alone is sufficient to suppress human tumor growth in nude mouse xenografts*. Oncogene, 1998. **16**(11): p. 1467.
111. Katz, M.E. and F. McCormick, *Signal transduction from multiple Ras effectors*. Current opinion in genetics & development, 1997. **7**(1): p. 75-79.
112. Arvelo, F., F. Sojo, and C. Cotte, *Biology of colorectal cancer*. Ecancermedicallscience, 2015. **9**.
113. De Angelis, R., et al., *Cancer survival in Europe 1999–2007 by country and age: results of EUROCARE-5—a population-based study*. The lancet oncology, 2014. **15**(1): p. 23-34.
114. Siegel, R., C. DeSantis, and A. Jemal, *Colorectal cancer statistics, 2014*. CA: a cancer journal for clinicians, 2014. **64**(2): p. 104-117.
115. Bogaert, J. and H. Prenen, *Molecular genetics of colorectal cancer*. Annals of gastroenterology, 2014. **27**(1): p. 9.
116. Jass, J., *Classification of colorectal cancer based on correlation of clinical, morphological and molecular features*. Histopathology, 2007. **50**(1): p. 113-130.
117. Einspahr, J.G., et al., *Associations of Ki-ras proto-oncogene mutation and p53 gene overexpression in sporadic colorectal adenomas with demographic and clinicopathologic characteristics*. Cancer Epidemiology and Prevention Biomarkers, 2006. **15**(8): p. 1443-1450.
118. Morkel, M., et al., *Similar but different: distinct roles for KRAS and BRAF oncogenes in colorectal cancer development and therapy resistance*. Oncotarget, 2015. **6**(25): p. 20785.
119. Raskov, H., et al., *Colorectal carcinogenesis-update and perspectives*. World Journal of Gastroenterology: WJG, 2014. **20**(48): p. 18151.
120. Takayama, T., et al., *Analysis of K-ras, APC, and β -catenin in aberrant crypt foci in sporadic adenoma, cancer, and familial adenomatous polyposis*. Gastroenterology, 2001. **121**(3): p. 599-611.
121. Yadamsuren, E.-A., et al., *Characteristics of advanced-and non advanced sporadic polypoid colorectal adenomas: correlation to KRAS mutations*. Pathology & Oncology Research, 2012. **18**(4): p. 1077-1084.
122. Ishii, M., et al., *Analysis of Ki-ras gene mutations within the same tumor using a single tumor crypt in colorectal carcinomas*. Journal of gastroenterology, 2004. **39**(6): p. 544-549.
123. Piton, N., et al., *Real-life distribution of KRAS and NRAS mutations in metastatic colorectal carcinoma from French routine genotyping*. Cancer Epidemiology and Prevention Biomarkers, 2015. **24**(9): p. 1416-1418.
124. Colussi, D., et al., *Molecular pathways involved in colorectal cancer: implications for disease behavior and prevention*. International journal of molecular sciences, 2013. **14**(8): p. 16365-16385.
125. Lievre, A., et al., *KRAS mutation status is predictive of response to cetuximab therapy in colorectal cancer*. Cancer research, 2006. **66**(8): p. 3992-3995.
126. Allegra, C.J., et al., *American Society of Clinical Oncology provisional clinical opinion: testing for KRAS gene mutations in patients with metastatic colorectal carcinoma to predict response to anti-epidermal growth factor receptor monoclonal antibody therapy*. J Clin Oncol, 2009. **27**(12): p. 2091-2096.
127. Engstrom, P.F., et al., *Colon cancer*. Journal of the National Comprehensive Cancer Network, 2009. **7**(8): p. 778-831.

128. Ferlay, J., et al., *Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012*. International journal of cancer, 2015. **136**(5): p. E359-E386.
129. Stewart, B.W. and P. Kleihues, *World cancer report*. Vol. 57. 2003: IARC press Lyon.
130. Lynch, T.J., et al., *Activating mutations in the epidermal growth factor receptor underlying responsiveness of non-small-cell lung cancer to gefitinib*. New England Journal of Medicine, 2004. **350**(21): p. 2129-2139.
131. Soda, M., et al., *Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer*. Nature, 2007. **448**(7153): p. 561.
132. Knight, S.B., et al., *Progress and prospects of early detection in lung cancer*. Open biology, 2017. **7**(9): p. 170070.
133. Wang, T., et al., *Five-year lung cancer survival: which advanced stage nonsmall cell lung cancer patients attain long-term survival?* Cancer: Interdisciplinary International Journal of the American Cancer Society, 2010. **116**(6): p. 1518-1525.
134. Ding, L., et al., *Somatic mutations affect key pathways in lung adenocarcinoma*. Nature, 2008. **455**(7216): p. 1069.
135. Mitsudomi, T., et al., *Mutations of ras genes distinguish a subset of non-small-cell lung cancer cell lines from small-cell lung cancer cell lines*. Oncogene, 1991. **6**(8): p. 1353-1362.
136. Ahrendt, S.A., et al., *Cigarette smoking is strongly associated with mutation of the K-ras gene in patients with primary adenocarcinoma of the lung*. Cancer, 2001. **92**(6): p. 1525-1530.
137. Brose, M.S., et al., *BRAF and RAS mutations in human lung cancer and melanoma*. Cancer research, 2002. **62**(23): p. 6997-7000.
138. Karachaliou, N., et al., *KRAS mutations in lung cancer*. Clinical lung cancer, 2013. **14**(3): p. 205-214.
139. Gnanasampanthan, G., et al., *Ki-ras mutation type and the survival benefit from adjuvant chemotherapy in Dukes' C colorectal cancer*. The Journal of pathology, 2001. **195**(5): p. 543-548.
140. Masciaux, C., et al., *The role of RAS oncogene in survival of patients with lung cancer: a systematic review of the literature with meta-analysis*. British journal of cancer, 2005. **92**(1): p. 131.
141. Matikas, A., et al., *Targeting KRAS mutated non-small cell lung cancer: a history of failures and a future of hope for a diverse entity*. Critical reviews in oncology/hematology, 2017. **110**: p. 1-12.
142. Gysin, S., et al., *Therapeutic strategies for targeting ras proteins*. Genes & cancer, 2011. **2**(3): p. 359-372.
143. Diepgen, T. and V. Mahler, *The epidemiology of skin cancer*. British Journal of Dermatology, 2002. **146**: p. 1-6.
144. Alimirah, F., et al., *Cellular senescence drives skin carcinogenesis*. 2018, AACR.
145. Lo, J.A. and D.E. Fisher, *The melanoma revolution: from UV carcinogenesis to a new era in therapeutics*. Science, 2014. **346**(6212): p. 945-949.
146. Jenkins, R.W. and R.J. Sullivan, *NRAS mutant melanoma: an overview for the clinician for melanoma management*. Melanoma Management, 2016. **3**(1): p. 47-59.
147. Akbani, R., et al., *Genomic classification of cutaneous melanoma*. Cell, 2015. **161**(7): p. 1681-1696.
148. Davies, H., et al., *Mutations of the BRAF gene in human cancer*. Nature, 2002. **417**(6892): p. 949.
149. Fedorenko, I.V., G.T. Gibney, and K.S. Smalley, *NRAS mutant melanoma: biological behavior and future strategies for therapeutic management*. Oncogene, 2013. **32**(25): p. 3009.

150. Hatzivassiliou, G., et al., *RAF inhibitors prime wild-type RAF to activate the MAPK pathway and enhance growth*. Nature, 2010. **464**(7287): p. 431.
151. Johnson, D.B. and I. Puzanov, *Treatment of NRAS-mutant melanoma*. Current treatment options in oncology, 2015. **16**(4): p. 15.
152. Kelleher, F.C. and G.A. McArthur, *Targeting NRAS in melanoma*. The Cancer Journal, 2012. **18**(2): p. 132-136.
153. Poulikakos, P.I., et al., *RAF inhibitors transactivate RAF dimers and ERK signalling in cells with wild-type BRAF*. Nature, 2010. **464**(7287): p. 427.
154. Dumaz, N., et al., *In melanoma, RAS mutations are accompanied by switching signaling from BRAF to CRAF and disrupted cyclic AMP signaling*. Cancer research, 2006. **66**(19): p. 9483-9491.
155. Ball, N.J., et al., *Ras mutations in human melanoma: a marker of malignant progression*. Journal of Investigative Dermatology, 1994. **102**(3): p. 285-290.
156. Burd, C.E., et al., *Mutation-specific RAS oncogenicity explains NRAS codon 61 selection in melanoma*. Cancer discovery, 2014.
157. Forbes, S.A., et al., *COSMIC: mining complete cancer genomes in the Catalogue of Somatic Mutations in Cancer*. Nucleic acids research, 2010. **39**(suppl_1): p. D945-D950.
158. Shi, H., et al., *Acquired resistance and clonal evolution in melanoma during BRAF inhibitor therapy*. Cancer discovery, 2013.
159. Van Allen, E.M., et al., *The genetic landscape of clinical resistance to RAF inhibition in metastatic melanoma*. Cancer discovery, 2013.
160. Denayer, E., T. de Ravel, and E. Legius, *Clinical and Molecular Aspects of RAS-related disorders*. Journal of medical genetics, 2008.
161. Schubbert, S., G. Bollag, and K. Shannon, *Deregulated Ras signaling in developmental disorders: new tricks for an old dog*. Current opinion in genetics & development, 2007. **17**(1): p. 15-22.
162. Schubbert, S., K. Shannon, and G. Bollag, *Hyperactive Ras in developmental disorders and cancer*. Nature Reviews Cancer, 2007. **7**(4): p. 295.
163. Schubbert, S., et al., *Germline KRAS mutations cause Noonan syndrome*. Nature genetics, 2006. **38**(3): p. 331.
164. Tartaglia, M., G. Zampino, and B. Gelb, *Noonan syndrome: clinical aspects and molecular pathogenesis*. Molecular syndromology, 2010. **1**(1): p. 2-26.
165. Tidyman, W.E. and K.A. Rauen, *Noonan, Costello and cardio-facio-cutaneous syndromes: dysregulation of the Ras-MAPK pathway*. Expert reviews in molecular medicine, 2008. **10**.
166. Tidyman, W.E. and K.A. Rauen, *The RASopathies: developmental syndromes of Ras/MAPK pathway dysregulation*. Current opinion in genetics & development, 2009. **19**(3): p. 230-236.
167. Ferner, R.E., *Neurofibromatosis 1 and neurofibromatosis 2: a twenty first century perspective*. The Lancet Neurology, 2007. **6**(4): p. 340-351.
168. Korf, B.R., *Malignancy in neurofibromatosis type 1*. The oncologist, 2000. **5**(6): p. 477-485.
169. Bollag, G., et al., *Loss of NF1 results in activation of the Ras signaling pathway and leads to aberrant growth in haematopoietic cells*. Nature genetics, 1996. **12**(2): p. 144.
170. Jett, K. and J.M. Friedman, *Clinical and genetic aspects of neurofibromatosis 1*. Genetics in Medicine, 2010. **12**(1): p. 1.
171. Rasmussen, S.A. and J. Friedman, *NF1 gene and neurofibromatosis 1*. American journal of epidemiology, 2000. **151**(1): p. 33-40.
172. Aoki, Y., et al., *Germline mutations in HRAS proto-oncogene cause Costello syndrome*. Nature genetics, 2005. **37**(10): p. 1038.
173. Aoki, Y., et al., *The RAS/MAPK syndromes: novel roles of the RAS pathway in human genetic disorders*. Human mutation, 2008. **29**(8): p. 992-1006.

174. Cirstea, I.C., et al., *A restricted spectrum of NRAS mutations causes Noonan syndrome*. Nature genetics, 2010. **42**(1): p. 27.
175. Cordeddu, V., et al., *Mutation of SHOC2 promotes aberrant protein N-myristoylation and causes Noonan-like syndrome with loose anagen hair*. Nature genetics, 2009. **41**(9): p. 1022.
176. Nava, C., et al., *CFC and Noonan syndromes due to mutations in RAS/MAPK signaling pathway: genotype/phenotype relationships and overlap with Costello syndrome*. Journal of medical genetics, 2007.
177. Pandit, B., et al., *Gain-of-function RAF1 mutations cause Noonan and LEOPARD syndromes with hypertrophic cardiomyopathy*. Nature genetics, 2007. **39**(8): p. 1007.
178. Razzaque, M.A., et al., *Germline gain-of-function mutations in RAF1 cause Noonan syndrome*. Nature genetics, 2007. **39**(8): p. 1013.
179. Roberts, A.E., et al., *Germline gain-of-function mutations in SOS1 cause Noonan syndrome*. Nature genetics, 2007. **39**(1): p. 70.
180. Tidyman, W.E. and K.A. Rauen, *Pathogenetics of the RASopathies*. Human molecular genetics, 2016. **25**(R2): p. R123-R132.
181. Nassar, N., et al., *The 2.2 Å crystal structure of the Ras-binding domain of the serine/threonine kinase c-Raf1 in complex with Rap1A and a GTP analogue*. Nature, 1995. **375**(6532): p. 554.
182. Tong, Y., Tempel, W., Shen, L., Arrowsmith, C.H., Edwards, A.M., Sundstrom, M., Weigelt, J., Bochkarev, A., Park, H., Structural Genomics Consortium (SGC). *Human K-Ras in complex with a GTP analogue*. 2009 [cited 2018; Available from: <https://www.rcsb.org/structure/3gft>].
183. Aytuna, A.S., A. Gursoy, and O. Keskin, *Prediction of protein–protein interactions by combining structure and sequence conservation in protein interfaces*. Bioinformatics, 2005. **21**(12): p. 2850-2855.
184. Baspinar, A., et al., *PRISM: a web server and repository for prediction of protein–protein interactions and modeling their 3D complexes*. Nucleic acids research, 2014. **42**(W1): p. W285-W289.
185. Tuncbag, N., et al., *Predicting protein-protein interactions on a proteome scale by matching evolutionary and structural similarities at interfaces using PRISM*. Nature protocols, 2011. **6**(9): p. 1341.
186. Cukuroglu, E., A. Gursoy, and O. Keskin, *HotRegion: a database of predicted hot spot clusters*. Nucleic acids research, 2011. **40**(D1): p. D829-D833.
187. Warshel, A. and M. Levitt, *Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme*. Journal of molecular biology, 1976. **103**(2): p. 227-249.
188. McCammon, J.A., B.R. Gelin, and M. Karplus, *Dynamics of folded proteins*. Nature, 1977. **267**(5612): p. 585.
189. Levitt, M. and R. Sharon, *Accurate simulation of protein dynamics in solution*. Proceedings of the National Academy of Sciences, 1988. **85**(20): p. 7557-7561.
190. Kandt, C., W.L. Ash, and D.P. Tieleman, *Setting up and running molecular dynamics simulations of membrane proteins*. Methods, 2007. **41**(4): p. 475-488.
191. Tinoco Jr, I. and J.-D. Wen, *Simulation and analysis of single-ribosome translation*. Physical biology, 2009. **6**(2): p. 025006.
192. Brandman, R., Y. Brandman, and V.S. Pande, *A-site residues move independently from P-site residues in all-atom molecular dynamics simulations of the 70S bacterial ribosome*. PloS one, 2012. **7**(1): p. e29377.
193. Lazaridis, T. and M. Karplus, *Effective energy function for proteins in solution*. Proteins: Structure, Function, and Bioinformatics, 1999. **35**(2): p. 133-152.

194. Roux, B. and T. Simonson, *Implicit solvent models*. Biophysical chemistry, 1999. **78**(1-2): p. 1-20.
195. Kleinjung, J. and F. Fraternali, *Design and application of implicit solvent models in biomolecular simulations*. Current opinion in structural biology, 2014. **25**: p. 126-134.
196. Anandakrishnan, R., et al., *Speed of conformational change: comparing explicit and implicit solvent molecular dynamics simulations*. Biophysical journal, 2015. **108**(5): p. 1153-1164.
197. Jang, H., et al., *Membrane-associated Ras dimers are isoform-specific: K-Ras dimers differ from H-Ras dimers*. Biochemical Journal, 2016: p. BCJ20160031.
198. Nussinov, R., et al., *A new view of Ras isoforms in cancers*. Cancer research, 2015.
199. Jang, H., et al., *Mechanisms of membrane binding of small GTPase K-Ras4B farnesylated hypervariable region*. Journal of Biological Chemistry, 2015: p. jbc. M114. 620724.
200. Chavan, T.S., et al., *High-affinity interaction of the K-Ras4B hypervariable region with the Ras active site*. Biophysical journal, 2015. **109**(12): p. 2602-2613.
201. Humphrey, W. and A. Dalke, *VMD user guide (Version 0.94)*. 1994, Beckman Institute Technical Report TB-94-07, University of Illinois.
202. Lu, S., et al., *GTP binding and oncogenic mutations may attenuate hypervariable region (HVR)-catalytic domain interactions in small GTPase KRAS4B, exposing the effector binding site*. Journal of Biological Chemistry, 2015: p. jbc. M115. 664755.
203. BIOVIA, D.S., *Discovery studio modeling environment, release 2017*. 2016, Dassault Systèmes San Diego, CA.
204. DeLano, W.L., *The PyMOL user's manual*. DeLano Scientific, San Carlos, CA, 2002. **452**.
205. Phillips, J.C., et al., *Scalable molecular dynamics with NAMD*. Journal of computational chemistry, 2005. **26**(16): p. 1781-1802.
206. Hynninen, A.P. and M.F. Crowley, *New faster CHARMM molecular dynamics engine*. Journal of computational chemistry, 2014. **35**(5): p. 406-413.
207. Brooks, B.R., et al., *CHARMM: the biomolecular simulation program*. Journal of computational chemistry, 2009. **30**(10): p. 1545-1614.
208. Venable, R. *CHARMM on Biowulf*. 2018, June 07; Available from: <https://hpc.nih.gov/apps/charmm/>.
209. Wang, W. and P.A. Kollman, *Free energy calculations on dimer stability of the HIV protease using molecular dynamics and a continuum solvent model*. Journal of molecular biology, 2000. **303**(4): p. 567-582.
210. Brooks, B.R., et al., *CHARMM: a program for macromolecular energy, minimization, and dynamics calculations*. Journal of computational chemistry, 1983. **4**(2): p. 187-217.
211. Genheden, S. and U. Ryde, *The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities*. Expert opinion on drug discovery, 2015. **10**(5): p. 449-461.
212. Rastelli, G., et al., *Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA*. Journal of computational chemistry, 2010. **31**(4): p. 797-810.
213. Maiorov, V.N. and G.M. Crippen, *Significance of root-mean-square deviation in comparing three-dimensional structures of globular proteins*. 1994.
214. Cohen, F.E. and M.J. Sternberg, *On the prediction of protein structure: the significance of the root-mean-square deviation*. Journal of molecular biology, 1980. **138**(2): p. 321-333.
215. James Gumbart, D.L. *HBonds Plugin Version 1.2*. 2011, March 09; Available from: <https://www.ks.uiuc.edu/Research/vmd/plugins/hbonds/>.
216. Leonardo Trabuco, E.V. *Salt Bridges Plugin Version 1.1*. 2011, March 09; Available from: <https://www.ks.uiuc.edu/Research/vmd/plugins/saltbr/>.
217. Ogmen, U., et al., *PRISM: protein interactions by structural matching*. Nucleic acids research, 2005. **33**(suppl_2): p. W331-W336.

-
218. Li, S.-C., et al., *Alpha-helical, but not beta-sheet, propensity of proline is determined by peptide environment*. Proceedings of the National Academy of Sciences, 1996. **93**(13): p. 6676-6681.
 219. Gentile, D.R., et al., *Ras binder induces a modified switch-II pocket in GTP and GDP states*. Cell chemical biology, 2017. **24**(12): p. 1455-1466. e14.
 220. Nelson, L.S., *Evaluating overlapping confidence intervals*. Journal of quality technology, 1989. **21**(2): p. 140-141.
 221. Velleman, P.F. and D.C. Hoaglin, *Applications, basics, and computing of exploratory data analysis*. 1981: Duxbury Press.
 222. McGill, R., J.W. Tukey, and W.A. Larsen, *Variations of box plots*. The American Statistician, 1978. **32**(1): p. 12-16.
 223. Barbacid, M., *Ras genes*. Annual review of biochemistry, 1987. **56**(1): p. 779-827.
 224. Hobbs, G.A., C.J. Der, and K.L. Rossman, *RAS isoforms and mutations in cancer at a glance*. J Cell Sci, 2016. **129**(7): p. 1287-92.
 225. Li, S., et al., *Raf-1 Cysteine-Rich Domain Increases the Affinity of K-Ras/Raf at the Membrane, Promoting MAPK Signaling*. Structure, 2018. **26**(3): p. 513-525. e2.
 226. Travers, T., et al., *Molecular recognition of RAS/RAF complex at the membrane: Role of RAF cysteine-rich domain*. Scientific reports, 2018. **8**(1): p. 8461.