Impact of D816V mutation of c-Kit on the binding of SHP1 using Molecular Dynamic Simulation approach

A Major Project dissertation submitted

in partial fulfilment of the requirement for the degree of

Master of Technology

In

Bioinformatics

Submitted by

Himani Raina

(2K11/BIO/08)

Delhi Technological University, Delhi, India

Under the supervision of

Dr Navneeta Bharadvaja

Department of Biotechnology
Delhi Technological University
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road,
Delhi-110042, INDIA

-

CERTIFICATE

This is to certify that the M. Tech. dissertation entitled **"Impact of D816V mutation of c-Kit on the binding of SHP1 using Molecular Dynamic Simulation approach"**, submitted by **HIMANI RAINA (2K11/BIO/08)** in partial fulfilment of the requirement for the award of the degree of Master of Technology, Delhi Technological University (Formerly Delhi College of Engineering, University of Delhi), is an authentic record of the candidate's own work carried out by her under my guidance.

The information and data enclosed in this dissertation is original and has not been submitted elsewhere for honouring of any other degree.

Date:

Dr Navneeta Bharadvaja
(Project Mentor)
Department of Bio-Technology
Delhi Technological University
(Formerly Delhi College of Engineering, University of Delhi)

DECLARATION

I **Ms. Himani Raina (2K11/BIO/08)**, student of M. Tech. (Bioinformatics), Delhi Technological University have completed the project titled **"Impact of D816V mutation of c-Kit on the binding of SHP1 using Molecular Dynamic Simulation approach"** for the award of Degree of M. Tech. (Bioinformatics), for academic session 2011-13. The information given in this project is true to the best of my knowledge.

Himani Raina

2K11/BIO/08

ACKNOWLEDGEMENT

It gives me immense pleasure to thank all those people who have been instrumental in the completion of my project.

I express my deep sense of thankfulness to Dr Navneeta Bharadvaja whose enthusiastic zeal and encouragement boosted me for the successful completion of my work.

I sincerely thank Dr. R.P.Tripathi, Director of INMAS, for allowing me to do this research work. I would like to extend my vote of thanks to Dr. Rajiv Vij (Scientist 'F') without whom this project would not have been possible in INMAS.

I would like to express my heartfelt gratitude to my supervisor Dr. G.U. Gurudatta (Scientist 'F') for allowing me to work under his guidance and providing me with all the facilities during my project work.

I am deeply indebted to Mr. Pawan Kumar Raghav for his keen interest, constructive criticism, unceasing encouragement and guidance. I express my gratitude to him for sincerely helping me and boosting my morale to work hard.

I gratefully acknowledge Dr. Yogesh Kumar Verma (Scientist 'C'), Dr. Neeraj Satija, Mr. Siddharth Pandey and Mr Vikas for their constant support.

Finally, yet importantly, I would like to express my heartfelt thanks to my beloved parents, for their blessings and my friends, for their help and wishes for the successful completion of this project.

HIMANI RAINA
2K11/BIO/08

CONTENTS

LIST OF FIGURES

LIST OF TABLES

TABLE **PAGE NO**

LIST OF ABBREVIATIONS

ADP	Adenosine Diphosphate
APS	Adaptor protein with PH and SH2 domains
ATP	Adenosine triphosphate
Ba/F3	Murine Pro B-cell
BMMC	Bone Marrow-Derived Mast Cell
BMP	Bone Morphogenetic Protein
CXCR4	Chemokine (C-X-C motif) Receptor 4
DAG	Diacylglycerol
EGF	Epidermal Growth Factor
ERK	Extracellular-Signal-Regulated Kinase
GDP	Guanosine Diphosphate
GIST	Gastrointestinal Stromal Tumors
Grb2	Growth Factor Receptor-Bound Protein-2
GTP	Guanosine Triphosphate
HEK293	Human Embryonic Kidney 293 cells
HMC1	Human Mast Cell Line 1
IP3	Inositol-1,4,5-Trisphosphate
JAK	Janus Kinase
JMR	Juxtamembrane Region
JNK	c- Jun N-Terminal Kinase
MAPK	Mitogen-Activated Protein Kinase
MD	Molecular Dynamics

MIHC	Myb-Immortalized Haemopoietic Cells
NAMD	Not (just) Another Molecular Dynamics program
NMR	Nuclear Magnetic Rasonance
NS	Nanosecond
PDGF	Platelet Derived Growth Factor
PH	Pleckstrin Homology
PI3K	Phosphoinositide 3-kinase
PKC	Protein Kinase C
PLC- γ	Phospholipase C-γ
PTB	Phospho-Tyrosine Binding
PTK	Protein Tyrosine Kinase
RMSD	Root Mean Square Deviations
RNA	Ribonucleic Acid
RTK	Receptor Tyrosine Kinase
SCF	Stem Cell Factor
SH2	Src Homology 2
SHP1	Src Homology 2 domain containing Phosphatase 1
SHP2	Src homology 2 domain containing phosphatase 2
SOCS	Suppressors Of Cytokine Signaling
STAT	Signal Transducers and Activators of Transcription
VMD	Visual Molecular Dynamics
Wnt	Wingless-Related Integration Site

Impact of D816V mutation of c-Kit on the binding of SHP1 using Molecular Dynamic Simulation approach

Himani Raina

Delhi Technological University, Delhi, India

1. ABSTRACT

Stem cells have the ability to differentiate into many different cell types and are associated with unique properties of self renewal and division. c-Kit is type III receptor tyrosine kinase expressed in a number of cells including stem cells. The signaling cascade of c-Kit is regulated by its positive and negative regulators. Stem cell factor (SCF) mediated signaling play important role in proliferation. On the contrary, the two important negative regulators of c-Kit are SHP1 and SHP2. Interestingly it has been observed that c-Kit signaling becomes SCF and SHP1 independent on mutating D816 residue to V. To elucidate the structural impact of this mutation on c-Kit structure, 1ns molecular dynamic simulation of the wild type and mutant c-Kit has been done in unphosphorylated and phosphorylated state. Computed energy differences in the wild type and mutant c-Kit revealed the impact of this single residue mutation on the structural stability. A thorough analysis of the hydrogen bonding pattern have led to a plausible mechanism that in case of the active (phosphorylated) mutant c-Kit new hydrogen are formed in the JMR region that make the Y570 incapable to bond with SHP1. The displacement of residues of the JMR and A-loop clearly suggests the disruption of the binding pocket of SHP1. Our study suggested that D816V mutation induces conformational changes in JMR and A-loop region of c-Kit that leads to disruption of SHP1 and c-Kit interaction.

2. INTRODUCTION

Stem cells are defined as cells that have the ability to perpetuate themselves through self-renewal and to generate mature cells of a particular tissue through differentiation. (Reya *et al*., 2001). The emerging applications of stem cells in the field of medicine and health have augmented the interest of the scientific community in stem cell research. Several factors control the balance between stem cell renewal and differentiation. The stem signaling network consists of a number of pathways mediated through their subsequent receptors. c-Kit , a type III receptor tyrosine kinase expressed in hematopoietic stem cells, dendritic, erythroid, megakaryotic, and myeloid progenitor cells, and pro-B and pro-T cells (Lyman and Jacobsen 1998). The ligand for c-Kit receptor is known Stem cell factor. Stem cell factor (SCF) and Kit, its receptor, play pivotal roles in cell differentiation, proliferation, and survival (Roskoski 2005).

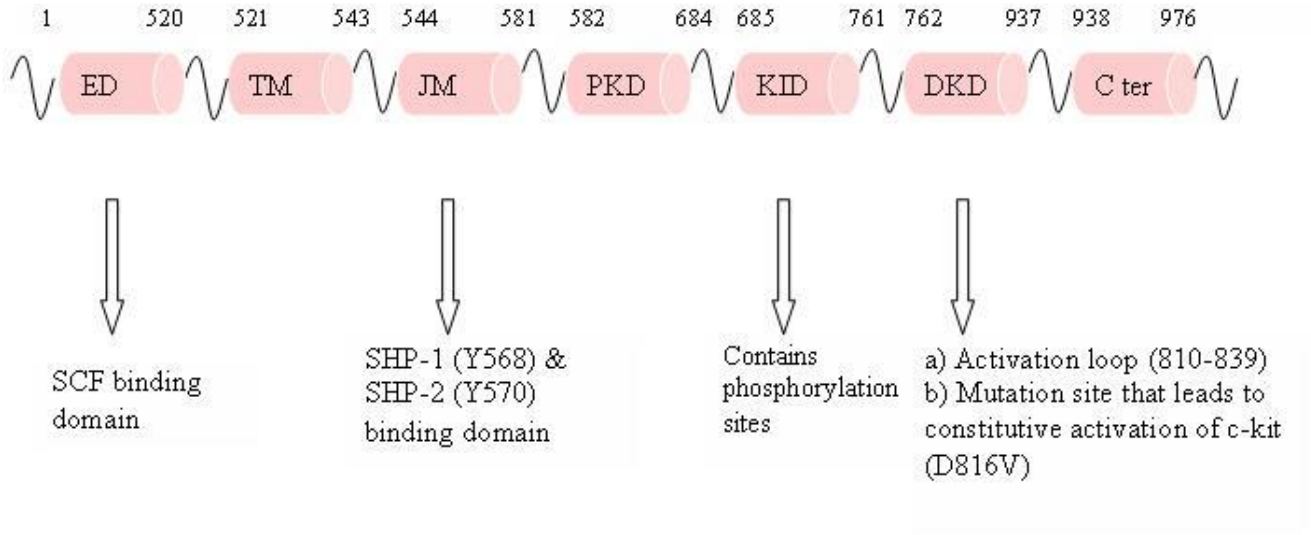


Fig 1: Domain regions of c-Kit. ED: Extracellular Domain TM: Transmembrane Domain JM: Juxtamembrane region PKD: Proximal kinase Domain KID: Kinase insert Domain DKD : Distal kinase Domain C ter: C Terminal region Important protein binding region and mutation sites shown by arrows.

Signaling cascade of c-Kit and SCF has significant contribution in a number of biological events like proliferation, differentiation and survival. c-Kit signaling is negatively regulated by phosphatase enzymes SHP1 and SHP2. Abnormal activation of c-Kit/SCF signaling is known often induced by mutations. The first evidence of activating mutation in humans c-Kit was found in mast cell leukemia cell line HMC-1 with the detection of two mutations in codons 560 (V560G) in the juxtamembrane domain and 816 (D816V) in the cytoplasmic domain of c-Kit, resulting in ligand-independent activation of the c-Kit product (Furitsu *et al*., 1993). The

biochemical studies showing the mechanism of ligand independent activation of this mutation has already been studied, but the fact why the negative regulation of D816V mutant c-Kit is impaired is unknown.

In the present work, we have studied the D816V mutation present in activation loop of c-Kit. In this study, we have carried out a detailed analysis of Kit receptor cytoplasmic region structural and dynamic changes related to D816V mutation through MD simulations. A molecular dynamics simulation of 1 ns at 310 K, of fully hydrated c-Kit wild type and mutant in the unphosphorylated and phosphorylated state has been performed to investigate the structural, dynamical and functional effects of the mutation. In order to study the impact of mutation on the binding site of SHP1 and SHP2, we have carried out MD simulation of the phosphorylated structure. A detailed analysis of the energy (total energy, electrostatic energy, vanderwaal energy), displacement per residue, root mean square deviations, H-bonding pattern and secondary structure content has been performed in order to study explicate the structural impact of a single residue mutation.

3. REVIEW OF LITERATURE

3.1 Stem cells

The cells that have the ability to perpetuate themselves through self-renewal and to generate mature cells of a particular tissue through differentiation are scientifically called as Stem cells (Reya *et al*., 2001). In addition to their role in the tissue and organ systems development and regeneration as units of biological organization but they also play important role as units in evolution by natural selection. Stem cells are generally defined as clonogenic cells capable of both self-renewal and multi-lineage differentiation (Till and McCulloch, 1961). Irrespective of their source all stem cells are associated with three general properties i.e. they are capable of dividing and renewing themselves for long periods; they are unspecialized; and they can give rise to specialized cell types. Hematopoietic stem cells are one of the best characterized stem cells. A number of cell receptors contribute to the cellular mechanisms activated by stem cells. The stem cell signaling network that contributes to its functional attributes includes a number of pathways like Wnt, Notch, CXCR4, BMP, c-Kit signaling (Chotinantakul and Leeanansaksiri 2012). The emerging potential of stem cells in the field of health and medicine has contributed to increasing interest of the scientific society in stem cell research. Among the major regulators of stem cell signaling are the receptor tyrosine kinases. Protein tyrosine kinases (PTK) are important regulators of intracellular signal transduction pathways mediating development and mutlicellular communication. Their activity is normally tightly controlled and regulated. PTK include receptor tyrosine kinase and non receptor tyrosine kinase as their major types. Further the receptor tyrosine kinases can be broadly classified into several classes according to structural variations in the receptors. Protein kinases regulate apoptosis, cell cycle progression and proliferation, cytoskeletal rearrangement, differentiation, development, the immune response, motility, nervous system function, and transcription. Owing to the myriad actions of protein kinases, it is imperative that they be stringently regulated because aberrant activity of these enzymes leads to a variety of diseases including cancer, diabetes, and autoimmune, cardiovascular, inflammatory, and nervous disorders. Considerable effort has been expended to determine the physiological and pathological functions of protein kinase signal transduction pathways. Because mutations and dysregulation of protein kinases play causal roles in human disease, these enzymes represent attractive drug targets (Cohen 2002).

3.2 Stem cell malignancies

The highly regulated and proscribed signaling cascades of stem cells leading to various biological functions are often deregulated. There are three aspects of the relationship between stem cells and tumour cells: first, the similarities in the mechanisms that regulate self-renewal of normal stem cells and cancer cells; second, the possibility that tumour cells might arise from normal stem cells; and third, the notion that tumors might contain 'cancer stem cells'- rare cells with indefinite proliferative potential that drive the formation and growth of tumors. It is also well known that cancers of the hematopoietic system (that is, leukemia's) provide the best evidence that normal stem cells are the targets of transforming mutations, and that cancer cell proliferation is driven by cancer stem cells (Reya *et al*., 2001). Stem cell signaling pathway associated with self renewal and oncogenesis are the Notch, Sonic hedgehog (Shh) and Wnt signaling pathways. Apart from this, a number of growth factors are also involved in promoting the survival and self-renewal of stem cells, and the proliferation, differentiation, and migration.

Numerous growth factors act via cell surface receptors of the receptor tyrosine kinase family (RTK). Ligand binding induces intrinsic tyrosine kinase activity of these receptors and downstream intracellular signaling pathways leading to various biological effects. It is interesting that several mutations of RTK family members have been implicated in growth factor independence and development of human malignancies. Basic cancer research has focused on identifying the genetic changes that lead to cancer. This has led to major advances in our understanding of the molecular and biochemical pathways that are involved in tumorigenesis and malignant transformation.

Signaling by stem cell factor and Kit, its receptor, play important roles in gametogenesis, hematopoiesis, mast cell development and function, and melanogenesis. Abnormal activation of c-Kit/SCF growth signal has been frequently observed in cancer. Activating mutations of c-Kit were first described in a feline model, induced by the Hardy-Zuckerman 4-feline sarcoma virus encoding the transforming oncogene v-Kit, a mutated and truncated viral homolog of c-Kit (Besmer *et al*., 1986). Gain-of-function mutations occur in a percentage of human neoplasms including mastocytomas (>90%), gastrointestinal stromal tumors (>70%), sinonasal T-cell lymphomas (17%), seminomas/dysgerminomas (9%), and acute myelogenous leukemia (1%) (Heinrich *et al*., 2002). Furthermore, autocrine or paracrine activation of Kit has been postulated in numerous other human malignancies including ovarian neoplasms and small-cell lung cancer (Krystal et al., 1996). Furthermore, a large number of human cancers express Kit. Activating Kit mutations occur in the extracellular, the juxtamembrane region (JMR), and the proximal and distal protein kinase domains.

Gastrointestinal stromal tumors (GIST) arise from the interstitial cells of Cajal; these cells play a role in intestinal motility. These tumors arise in the stomach (60%), small intestine (25%), rectum (5%), esophagus (2%), and a variety of other abdominal locations (Corless *et al*., 2004). The standard treatment for localized GIST is surgical removal. Recurrence after surgical resection is common, occurring in up to 90% of people with larger tumors, and long term survival following recurrence is unusual.

Table 1: Activating mutations in c-Kit and associated Disease

Residue	**Disease**	**Segment**
550K-558K (Del)	GIST	Juxtamembrane
550 K-586R	GIST	Juxtamembrane
559V-560V (Del)	GIST	Juxtamembrane
557 R-W	GIST	Juxtamembrane
559 V-A	GIST	Juxtamembrane
559 V-D	GIST	Juxtamembrane
560 V-G	Leukemia	Juxtamembrane
816 D-V	Mastocytosis	Activation loop
816 D-Y	AML, Mastocytosis,GIST	Activation loop
816 D-H	Tumor	Activation loop
816 D-G	Mast cell disease	Activation loop
816 D-F	Mastocytosis	Activation loop
816 D-N	Sinusoidal NK/T cell lymphoma	Activation loop
822 N-K	Germ cell tumor	Activation loop
829 A-P	Germ cell tumor	Activation loop
820 D-G	Mast disease aggressive	Activation loop

The c-Kit mutations found in gastrointestinal stromal tumors involve most commonly the JM domain (Hirota *et al*., 1998). Studies have indicated that JM domain mutations occur in about 67% of all cases of GIST (Corless *et al*., 2004). These mutations involve residues K550–R586. Although some of these mutations involve amino acid substitutions, most are deletions of 2–16 residues. The greatest frequency of mutation occurs at residues R557–V559 in the critical JM-B segment. These gain-of-function mutations emphasize the auto-inhibitory role of the JM domain that plays important in controlled c-Kit signaling. In addition to mutations in JM region, about 17% of mutations in GIST occur in the extracellular domain and involve a duplication of A501 and Y502. The extracellular mutations disrupt an inhibitory dimerization motif. About 2% of the Kit mutations involve a K642E substitution in the αC loop of the small lobe of protein tyrosine kinase domain of c-Kit. Another 2% of the c-Kit mutations involve residues D820 and N822 in

the activation loop of the large lobe. The mutations at residues K642, D820, and N822 may stabilize the active conformation of Kit.

The cases that lack c-Kit mutation in GIST, contain mutations in the PDGFα receptor (Heinrich *et al.*, 2003), also a member of the type III receptor protein-tyrosine kinases. In contrast to Kit, the majority of PDGFα mutations occur within the activation loop, and a minority involves the JM domain.

Kit mutations in nasal T-cell lymphomas involve both the JM domain (residues V559 and E561) as well as the activation loop (residues D816 and V825) (Hongyo *et al*., 2000). The most common mutation in mast cell tumors occurs at residue 816; other mutations occur at residues D820 and V560. Residues D816 and D820 occur in the activation loop while residue Val560 occurs within the JM domain (Kitamura and Hirota 2004). The most common Kit mutation in germ-cell tumors occurs at residue D816 in the activation loop. Other mutations occur at residues D820 and D822 in the activation loop and residue W557 in the JM domain.

A Kit mutation at residue D816 has also been described few cases of acute myelogenous leukemia (Heinrich *et al*., 2002). Studies examining the monomer–dimer transition in wild type, V560G, and D816V mutant mouse Kit have showed that these two mutations lead to SCF-independent Kit activation (Kitayama *et al*., 2002). The extent of phosphorylation is greater in the activation loop mutant (residue D816) than in the JM mutant (residue V560). Chemical cross-linking experiments have also been performed using bis (sulfosuccinimidyl) suberate (BS3), a bifunctional, water-soluble, membrane impermeable cross-linker that reacts with the Kit extracellular domain. In the absence of SCF, the JM mutant exists as a dimer while the activation loop mutant exists as a monomer. Receptor dimerization following ligand binding is considered to be the chief mechanism leading to receptor protein- tyrosine kinase activation (Blume-Jensen *et al*., 1991). The JM mutation in Kit leads to receptor dimerization with concomitant kinase activation. Besides auto-inhibition, as described above, the JM segment inhibits Kit dimerization and subsequent kinase activation. When cells containing the Kit D816V activation loop mutant are treated with SCF, receptor dimerization occurs.

3.3 The structure and signaling cascades of c-Kit

The viral oncogene v-Kit was identified in 1986 as the transforming gene of the Hardy-Zuckerman 4 feline sarcoma virus (Besmer *et al*., 1986), and its cellular homolog, c-Kit, was cloned and sequenced subsequently (Yarden *et al*., 1987). Studies showed that c-Kit is allelic with the dominant white spotting (W) of mice (Chabot *et al*., 1988, Geissler *et al*., 1988). As a result of alternative messenger RNA (mRNA) splicing, four isoforms of c-Kit have been identified in humans and two in mice. In both mice and humans, an alternative splicing result in isoforms characterized by the presence or absence of a tetrapeptide sequence (GNNK) in the extracellular part of the juxtamembrane region (Reith *et al*., 1991, Crosier *et al*., 1993), and occurs due to alternate use of 5 splice donor sites (Hayashi *et al*., 1991).

Stem cell factor (SCF) receptor or CD117, also known as human receptor tyrosine kinase (RTK) c-Kit (Arock and Valent 2010), belongs to the type III RTK family. Type III RTKs consist of a glycosylated extra-cellular ligand-binding domain (ectodomain) connected to a cytoplasmic region by means of a single transmembrane helix. The class III receptors are characterized by the presence of five immunoglobulin-like domains in their extracellular portion. Stem cell factor (SCF), the ligand for c-Kit binds to the second and third immunoglobulin domains while the fourth domain plays a role in receptor dimerization (Zhang *et al*., 2000). The structure of the class III receptors differs from that of other receptor tyrosyl kinases by the insertion of 70–100 amino acids near the middle of the kinase domain. The cytoplasmic region of c-Kit is composed of an auto-inhibitory juxtamembrane region (JMR) and a protein tyrosine kinase (PTK) that is subdivided into proximal and distal lobes separated by an insert sequence of variable length (70–100 amino acids). In human c-Kit, the 77-amino acid kinase insert domain (KID) possesses phosphorylation sites and provides an interface for the recognition of pivotal signal transduction proteins.

The entire Kit structure comprising 976 residues has numerous domains. Fig 2 clearly shows the domain regions of c-Kit. The extracellular region spans from residues G23 to T520, whereas the sequence M1-T22 forms the signal sequence. The transmembrane segment is L521 to L543 followed by the Juxtamembrane domain that lies within T544-K581. The protein tyrosine kinase domain is further subdivided into three domains, first is the proximal kinase domain from W582 to R684, second is Kinase insert domain from K685 to E761 and last is the distal kinase domain (L762-S937).

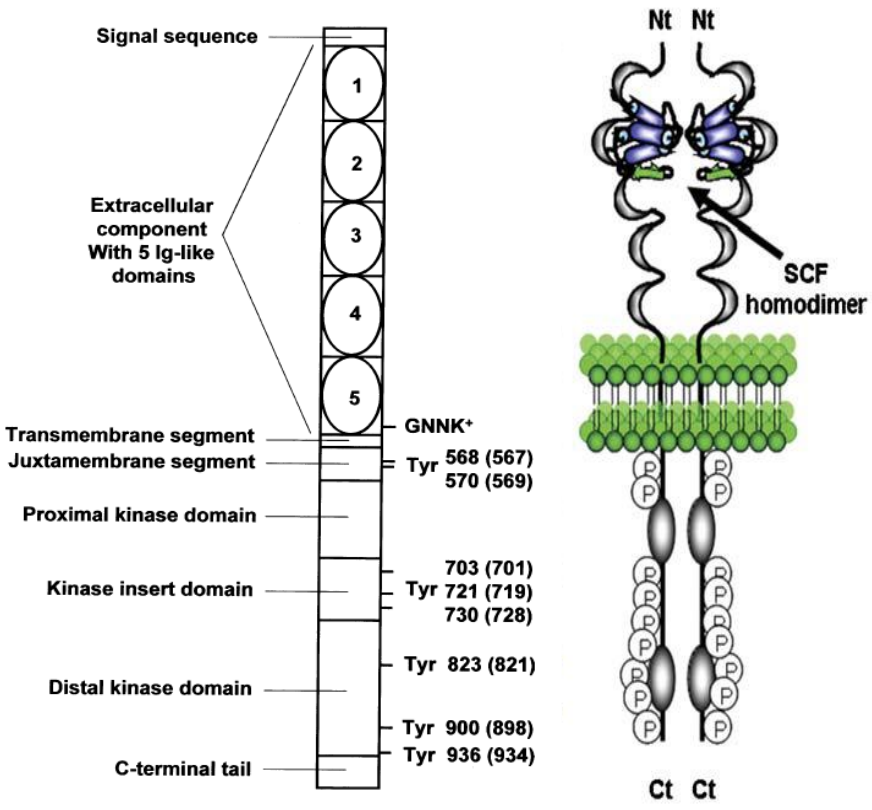


Fig 2: Structural domains in c-Kit with important phosphorylation sites and receptor dimerization state.(Reber *et al*., 2006)

Numerous three dimensional structures have been crystallized for c-Kit. The Kit protein-tyrosine kinase domain has the characteristic bilobed architecture as observed in all protein kinases. Residues W582–E671 makes up the small N-terminal lobe of the kinase, and residues L678-H953 make up the large C-terminal lobe with a hinge segment between them. The small lobe has a predominantly anti-parallel beta-sheet structure and is involved in anchoring and orienting ATP. It contains a glycine-rich (GAGAFG) ATP-phosphate- binding loop composed of residues G596–G601. The large lobe is predominantly a-helical in nature. The large lobe is responsible for binding the peptide or protein substrate. Furthermore, part of the ATP-binding site occurs in the large lobe. As described for other protein kinases, the catalytic site of Kit kinase lies in the cleft between the small and large lobes (Mol *et al*., 2003, Mol *et al*., 2004).

3.3.1 Structure of active Kit

The active form of Kit was crystallized by incubating the enzyme with Mg ATP to initiate the transphosphorylation reaction. The Y570 and Y568 in the juxtamembrane are the major phosphorylation sites. The autophosphorylation of these residues in the JM domain occurs before that of the Tyr residues in the activation loop (Mol *et al*., 2004a). The structure of the active c-kit that has crystallized is known to lacking the extracellular domain. The active structure of Kit with ADP, Mg2+, and the side chain of pTyr568 from an adjacent molecule bound at the active site is consistent with those of other active protein kinases (Mol *et al*., 2003). The αC-helix and the activation loop in this structure were in active conformations. The glutamate residue from the αC-helix (E640) forms a salt bridge with K623 (the K of K/D/D) that bridges the a- and b phosphates of ADP in the active conformation of c-Kit. Mg2+ binds to D810 (of the DFG sequence), N797 and to the phosphate of the phosphotyrosine residue. F811 (of the DFG segment) exists in its active ‘‘on’’ (Mol *et al*., 2004a) or ‘‘in’’ (Mol *et al*., 2004b) conformation that allows the binding of the adenine of ADP. The activation loop of Kit is in its extended, or active, conformation. This extended conformation of the A-loop makes the tyrosine residues of this region accessible to its dimer partner for phosphorylation. Phosphorylation of Y823 in the A-loop plays important role in stabilizing the extended conformation of the activation loop and thus maintaining the c-Kit in its active form.

3.3.2 Structure of inactive c-Kit

In the inactive (autoinhibited) c-Kit the JM segment forms a V-shaped loop that inserts directly into the interface between the small and large lobes of the protein tyrosine kinase domain (Mol *et al*., 2004a). Ribbon diagram of the enzyme as shown in Fig 3 suggests that the JM domain has the potential to inhibit Kit by displacing the a C-helix, preventing the activation loop from assuming its extended and active conformation, and preventing the movement of the small and large lobes necessary for the binding and release of substrates. The mechanism of c-Kit autoinhibition has been found to be comparable with another type III protein tyrosine kinase Flt3 (Griffith *et al*., 2004). The JM segment has been divided into three topological components: from the N- to C- direction as JM-binding motif (JM-B), the JM switch (JM-S), and the JM zipper (JM-Z). The corresponding residues to these segments are Y553–V559 (JM-B), V560–I571 (JM-S), and D572–K581 (JM-Z) followed by residues T544–M552 that forms the JM proximal segment (JM-P). The four segments of the JMR fom a V-shaped structure that lies on the surface of the Kit kinase domain. JM-B segment makes contacts with the glycine-rich nucleotide-binding loop, the activation loop, and the αC-helix that are implicated in the activation–inactivation transitions of this enzyme. The JM-B segment prevents the extension of the compact, or non-extended, form of the activation loop from assuming its extended active state in order have the active conformation of c-kit. Several residues of the JM domain of inactive Kit form hydrophobic bonds with residues in both the small and large lobes of the kinase domain of c-Kit.

Moreover, Y553 of the JM segment forms hydrogen bonds with the side chains of buried and conserved D810 of the DFG segment and E640 of the aC-helix. In the inactive conformation, F811 (of the DFG segment) has a special significance as it occurs in its inactive ‘‘off’’ (Mol *et al.,* 2004a) or ‘‘out’’ (Mol *et al.,* 2004b) conformation that prevents the binding of the adenine of ADP, and preventing activation of c-Kit. Thus, JM domain of inactive Kit sterically blocks the activation loop from assuming an active conformation. Another factor that contributes to the autoinhibited state, is the hydroxyl group of non-phosphorylated tyrosine 823 that forms a hydrogen bond with the catalytic aspartate (D792), thereby preventing the binding of protein substrates to the active site.

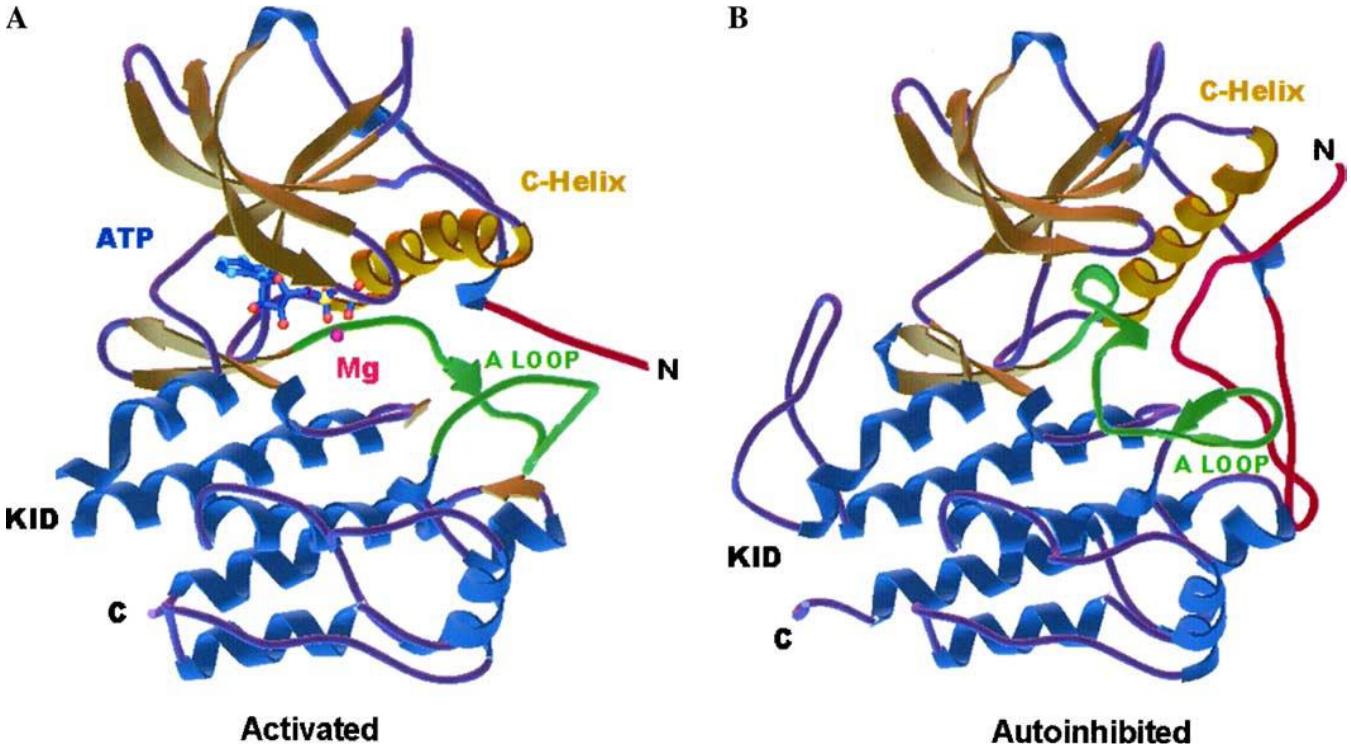


Fig 3: Ribbon diagrams of the activated and autoinhibited forms of Kit showing the N- and C-termini. (Mol *et al*., 2004b)

3.3.3 c-Kit activation by stem cell factor

The binding of SCF to c-Kit results in receptor dimerization and activation of its downstream signaling. The c-Kit ligand, stem cell factor (SCF) is expressed as a glycosylated transmembrane protein. There are two isoforms of SCF that have been obtained as result of alternative splicing and differ in the absence or presence of a particular proteolytic cleavage site (Huang *et al*., 1992). The isoform containing the cleavage site undergoes proteolysis and becomes soluble upon release from the plasma membrane, whereas the isoform lacking the cleavage site remains cell

associated. The abilities of the isoforms to transmit signals are different. Stimulation with the soluble isoform leads to rapid and transient activation and autophosphorylation of c-Kit, as well as fast degradation, whereas stimulation with the membrane-associated isoform leads to more sustained activation (Miyazawa *et al*., 1995). Differences also exist in signaling downstream of c-Kit. The membrane-bound ligand induced a more persistent activation of Erk1/2 and p38 mitogen-activated protein kinase (MAPK), as compared to the soluble ligand (Kapur *et al*., 2002). The membrane anchoring of the SCF prevent the internalization of receptor-ligand complex, thus imparting persistant activation in contrast to the soluble SCF. Stem cell factor is widely expressed during embryogenesis and can be detected in brain, endothelium, gametes, heart, kidney, lung, melanocytes, skin, and the stromal cells of the bone marrow, liver, and thymus (Galli *et al*., 1994). Stem cell factor acts synergistically with hematopoietic colony-stimulating factors such as granulocytemacrophage colony-stimulating factor, interleukin 3, and erythropoietin (Lennartsson *et al*., 2005).

3.3.4 Signal Transduction through c-Kit

Binding of SCF to c-Kit results in dimerization of the receptor followed by activation of its intrinsic tyrosine kinase activity (Blume-Jensen *et al*., 1991). It is thought that dimerization is driven by the simultaneous binding of a dimeric SCF molecule to two receptor monomers (Philo *et al*., 1996, Lemmon *et al*., 1997). The activated receptor becomes autophosphorylated on a number of tyrosine residues, mainly located outside the kinase domain, which serve as docking sites for signal transduction molecules containing Src homology 2 (SH2) or phosphotyrosine binding (PTB) domains (Pawson 2004). The initial residues that undergo autophosphorylation, which are in the juxtamembrane segment, include Y568 and Y570. Other residues that undergo phosphorylation, including three in the kinase insert domain are shown in Fig 2. Activated Kit also catalyzes the phosphorylation of substrate proteins. The phosphorylated residues interact with various signal transduction molecules and initiate signaling pathways in order to impart functional role to c-Kit. Numerous downstream signaling pathways are associated with the functional roles of c-Kit. The main function of c-kit is the proliferation of hematopoietic stem cells. The pathways that specifically regulate proliferation via c-Kit are the PI3K and MAPK pathway.

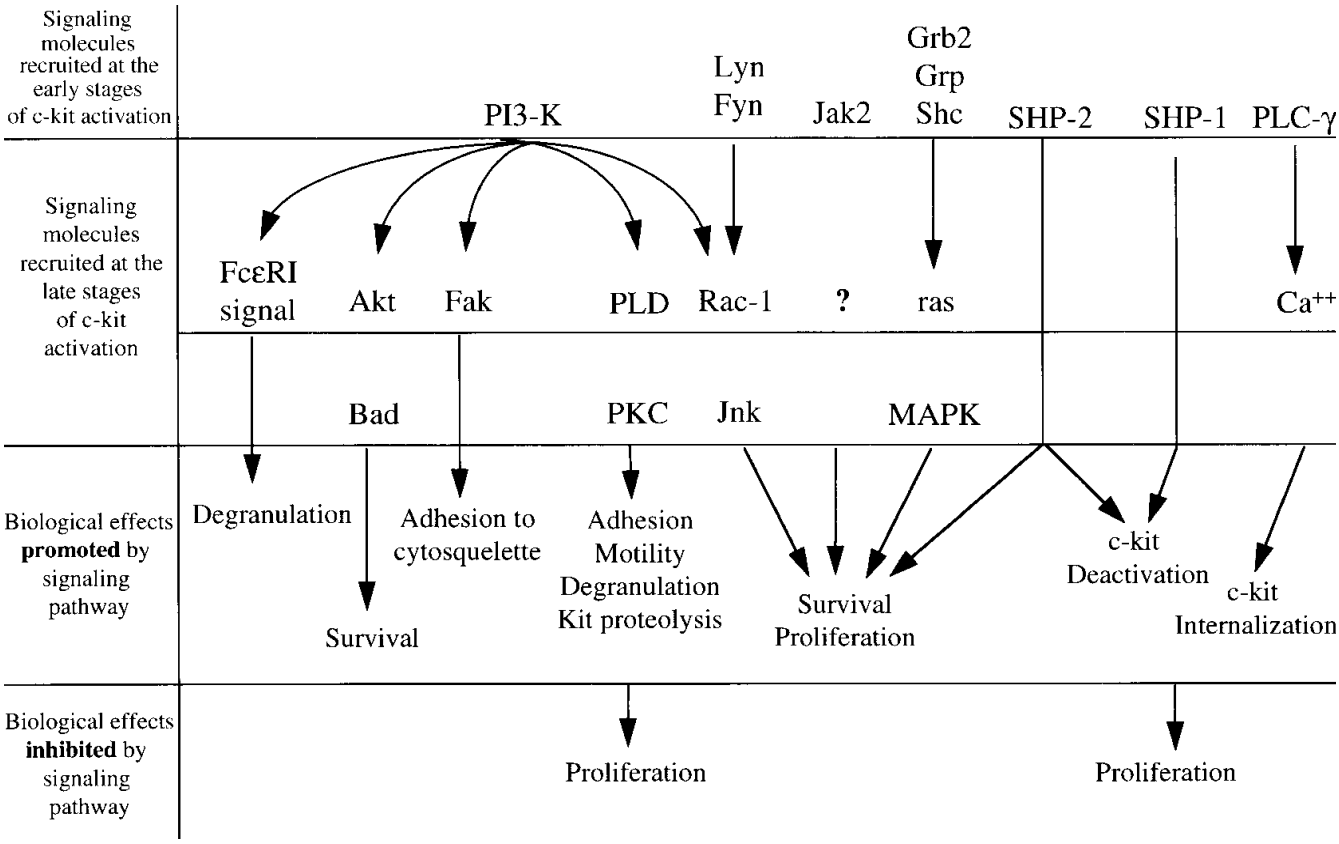


Fig 4: Biological effects of signaling molecules associated with c-Kit (Boissan et al., 2000)

Kit has the potential to participate in multiple signal transduction pathways as a result of interacting with several enzymes and adaptor proteins (Roskoski *et al.*, 2005). (Fig 4) The adaptor protein APS, Src family kinases, and SHP2 tyrosyl phosphatase bind to phosphotyrosine 568. SHP1 tyrosyl phosphatase and the adaptor protein Shc bind to phosphotyrosine 570. C-terminal Src kinase homologous kinase (Chk) and the adaptor Shc bind to both phosphotyrosines 568 and 570. These residues occur in the juxtamembrane domain of Kit. Three residues in the kinase insert domain are phosphorylated and attract: (a) the adaptor protein Grb2 (Y703), (b) phosphatidylinositol 3-kinase (Y721), and (c) phospholipase C (Y730). Phosphotyrosine 900 in the distal kinase domain binds phosphatidylinositol 3-kinase that in turn binds the adaptor protein Crk. Phosphotyrosine 936, also in the distal kinase domain, binds the adaptor proteins APS, Grb2, and Grb7 (Roskoski et al., 2005). The numerous Kit interactions cited above lead to activation of several signal transduction pathways. (Fig 5)

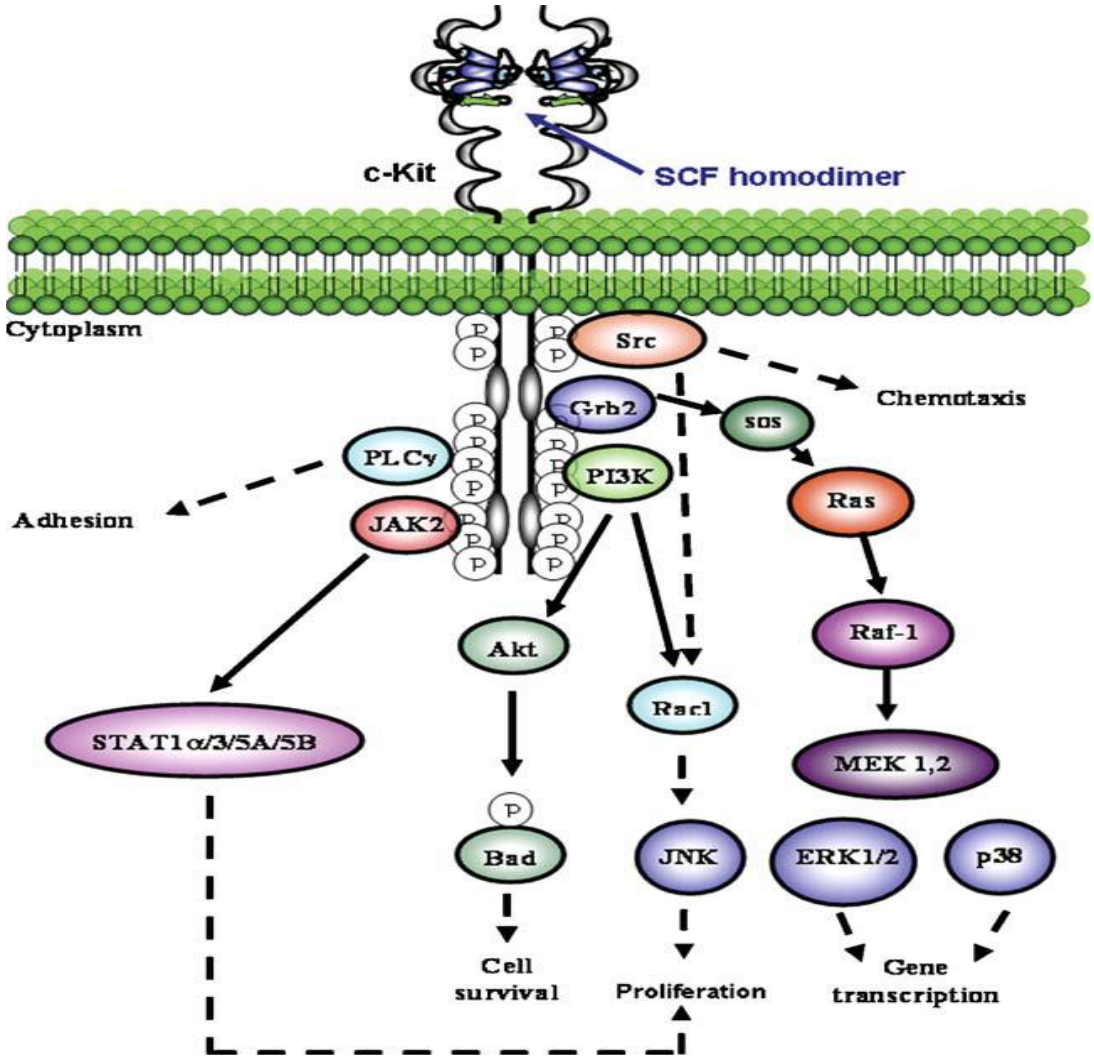


Fig 5 : c-Kit signal transduction and their biological roles. (Reber *et al*., 2006)

Role of MAP Kinase Pathway in c-Kit signaling

Growth factor receptor-bound protein-2 (Grb2) is an adaptor protein that binds to phosphorylated Y703 and Y936 of c-Kit (Thommes *et al*., 1999). The complex of Grb2 with son-of sevenless protein sos interacts with G-protein Ras , which leads to activation of Raf-1 and finally of the MAP kinases p38 , ERK1/2 and c- Jun N-terminal kinase (JNK) .These MAP kinases are known to act on transcription factors activity, and thereby on gene transcription. Numerous studies have implicated the critical importance of the Ras/Erk pathway in cell division and survival (Lewis *et al*., 1998). Ras is a small G-protein that can oscillate between an active GTP-bound form and an inactive GDP-bound form. Although Ras can activate a number of signal transduction molecules such as Rac or PI3-kinase (Rodriguez-Viciana *et al*., 1997, Qui *et al*., 1995), its role in the Ras/Erk cascade is the most well characterized.

Role of PI-3K Pathway in c-Kit signaling

Phosphatidyl-inositol-3 kinase (PI3-kinase) interacts with phosphorylated Y721 of c-Kit (Serve *et al*., 1994). SCF-induced PI3-kinase recruitment leads to Akt activation and to subsequent phosphorylation of the pro-apoptotic factor Bad. This phosphorylation inhibits Bad activity, thereby promoting cell survival (Blume-Jensen *et al*., 1998). The role of PI3K pathway in cell survival has been studied in Myb-immortalized haemopoietic cells (MIHC) expressing WT or Y721F c-Kit (GNNK+ isoform).The c-kit Y721F mutant was insufficient to prevent a high rate of cell death by SCF induced signaling and also the PI3-kinase inhibitor LY294002 inhibited SCF-induced survival of MIHC transfected with the GNNK+ isoform, in a dose-dependent manner (Young et al., 2006). PI3-kinase is also known to mediate SCF-induced proliferation of bone marrow-derived mast cells (BMMC), by activating the small guanosine-triphosphate (GTP)-binding protein Rac1 and JNK pathways (Timokhina *et al*., 1998).In PI3-kinase-/- BMMC SCF-induced JNK phosphorylation and cell proliferation was downregulated which supports the fact that PI3K plays important role in proliferation. (Fukao *et al*., 2002).

Role of PLC-γ Pathway in c-Kit signaling

The association site of PLC-γ seems to be the phosphorylated Y730 of c-Kit (Gommerman *et al*., 2000).Phospholipase C-γ (PLC-γ) exists as two isoforms, PLC-γ 1 and PLC-γ 2. They both consist of two SH2 domains, one SH3 domain, one PH domain and a catalytic domain. While PLC-γ 1 is ubiquitously expressed, PLC-γ 2 is mainly expressed in the hematopoietic system (Carpenter and Ji 1999). PLC hydrolyses the phosphoinositide PIP2, thereby generating the second messengers DAG and inositol-1,4,5-trisphosphate (IP3). DAG is an activator of the classical and novel forms of PKC, while IP3 binds to specific receptors present on the endoplasmic reticulum, triggering release of Ca2+ from internal stores and the intracellular concentration of free Ca2+ regulates a number of cellular processes (Berridge *et al*., 2003). PLC-γ associates with c-Kit in the absence of ligand in P815 transformed murine mast cells, where c-Kit is constitutively phosphorylated (Rottapel *et al*., 1991). PLC-γ associates with c-Kit after epidermal growth factor (EGF) stimulation in HEK293 cells overexpressing an EGF receptor-c-Kit chimera (Herbst *et al*., 1991). In addition, transfection of COS cells with a truncated form of c-Kit (tr-Kit) consisting of the kinase domain of Kit leads to activation of PLC-γ (Sette *et al*., 1998).

Role of Src Pathway in c-Kit signaling

SCF induces the activation of multiple Src (v-src avian sarcoma [Schmidt-Ruppin A-2] viral oncogene homolog) family members, including Src, Tec, Lyn and Fyn (Blume- Jensen *et al*., 1994; Tang *et al*., 1994; Linnekin *et al*., 1997). The Src family associates with phosphorylated Y568 and Y570 in the juxtamembrane domain of c-Kit (Price *et al*., 1997; Ueda *et al*., 2002). Src kinase and PI3-kinase signalling pathways converge to activate Rac1 and JNK after SCF stimulation in BMMC, promoting cell proliferation (Timokhina *et al*., 1998). The reduction in the expression of the Src kinase Lyn is accompanied by an inhibition of SCF induced proliferation (Linnekin *et al*., 1997). SCF induced gene transcription is also mediated by activation of members of the Src family kinases, leading to activation of the Ras/MAP kinase pathway (Lennartsson *et al*., 1999; Bondzi *et al*., 2000). The GNNK- isoform of c-Kit shows stronger activation of Src family kinases than the GNNK+ isoform (Voytyuk *et al*., 2003).

Role of JAK/STAT Pathway in c-Kit signaling

The Janus kinases (JAKs) are cytoplasmic tyrosine kinases that are activated through ligand stimulation of cytokine receptors or RTKs. Downstream of JAKs are the signal transducers and activators of transcription (STATs), which are phosphorylated by JAKs. STAT proteins are a class of transcription factors with DNA binding domains, an SH2 domain and a carboxy-terminal transactivating domain. SCF induces the activation of the Janus kinase (JAK)/Signal Transducers and Activators of Transcription (STAT) pathways. JAK2 associates with c-Kit and is phosphorylated after SCF stimulation (Brizzi *et al*., 1994b). JAK2 activation results in phosphorylation of STAT1α, 3, 5A and 5B (Brizzi *et al*., 1999). SCF-induced JAK/STAT activation is associated with fetal liver haematopoietic progenitor cells proliferation and differentiation (Radosevic *et al*., 2004). Studies have shown that JAK/STAT activation after SCF stimulation (100 ng/ml), in MO7e and HCD57 cells, but reasons for this difference remain unclear (Gotoh *et al*., 1996).

2.3.5 c-Kit activity inhibited by its negative regulators

Protein kinase C (PKC) is a family of serine/threonine kinases that are important regulators of several RTKs, including c-Kit (Blume-Jensen *et al*., 1994). Stimulation of c-Kit with soluble SCF results in phosphoinositide 3-kinase (PI3-kinase)-dependent activation of phospholipase D (Kozawa *et al*., 1997), leading to release of phosphatidic acid, which can be dephosphorylated to yield diacylglycerol (DAG), an activator of PKC. The tyrosine kinase activity of c-Kit can be modulated through phosphorylation by PKC. Down modulation of c-Kit activity by PKC occurs through dual mechanisms. Activation of PKC phosphorylates S741 and S746 in the kinase insert region of c-Kit, which leads to inhibition of kinase activity (Blume-Jensen *et al*., 1995).

The suppressors of cytokine signaling (SOCS) are a family of proteins that were originally cloned based on their ability to suppress cytokine signaling (Wormald and Hilton 2004). They have a central SH2 domain flanked by an N-terminal domain of variable length and a C-terminal domain of 40 amino acids denoted the SOCS box. SOCS-1 has been identified as an interactor with c-Kit (De Sepulveda *et al*., 1999). Its expression is induced upon stimulation of mast cells with SCF, and it associates with c-Kit via its SH2 domain. In contrast to its function in cytokine signaling, SOCS-1 selectively suppresses c-Kit-stimulated mitogenesis, while not affecting survival signals. The mechanism does not involve inactivation of the tyrosine kinase activity of c-Kit, but through binding of Grb2 via its SH3 domain to SOCS-1, which in turn binds to Vav. Targeted deletion of SOCS-1 leads to a reduced proliferative response via c-Kit upon SCF stimulation (Ilangumaran *et al*., 2003).

The protein tyrosine phosphatase SHP1interacts with Y570 of c-Kit and negatively regulates c-Kit signaling (Kozlowski *et al*., 1998; Yi and Ihle 1993). SHP1 is a cytosolic phosphotyrosyl phosphatase containing two tandem SH2 domains, a phosphatase domain and a C-terminal tail. SHP1 occurs primarily in hematopoietic and epithelial cells, and it is a negative regulator of growth factor signaling. Besides inhibiting Kit signaling, SHP1diminishes the growth-promoting properties of the colony-stimulating factor 1, erythropoietin, and interleukin 3 receptors, an effect mediated either directly by receptor dephosphorylation or indirectly by dephosphorylation of receptor-associated protein- tyrosine kinases. A hyperproliferative phenotype of their hematopoietic progenitor cells has been observed in motheaten (me) mice that express loss of function mutation in SHP1(Shultz *et al*., 1993). The role of SHP1 might to some extent be cell-type specific (Lorenz *et al*., 1996). SHP2, like SHP1, is a cytosolic phosphotyrosyl phosphatase containing two tandem SH2 domains. The ability of SHP2 to associate with activated Kit is markedly reduced by the Y568F mutation but is unaffected by the Y570F mutation. Moreover, expression of c-Kit bearing phenylalanine substitutions for either Y568 or Y570 is associated with markedly enhanced proliferation in response to SCF. pY570 mediates the association of Kit with SHP1 while pY568 mediates the association of Kit with SHP2 *in vivo*. Thus, SHP1and SHP2 can negatively modulate Kit signaling by interacting with these specific phosphotyrosine residues.

Regulation of physiological functions in the cell is mostly governed by phosphorylation-a crucial mechanism in cell signaling-catalyzed by protein kinases. Owing to the myriad of events mediated by c-Kit signaling, its controlled mechanism is necessarily required in order to have sustained cellular functions. The irregular signaling cascades are often associated diseases. The deactivation of tyrosine kinases or their oncogenic activation relates with mutations (point mutations as well as deletions and gene fusions) which affect the primary structure of the protein. The structural properties of PTKs were characterized mainly by X-ray analysis. Although crystallographic data yield valuable insights into such structural rearrangements, they represent only average conformation for a given set of crystallization conditions. Alternative experimental

techniques, such as NMR spectroscopy, and computational approaches, such as molecular dynamics (MD) provide a way to better understand the structure-dynamics-function relationships at the atomic level and further characterize the alteration of protein structure and internal dynamics induced by cancer mutations. These theoretical methods also enable to describe intermediate conformational states that can be used to guide the design of specific inhibitors acting as modulators of the enzymatic function by targeting putative allosteric sites. As already discussed, c-Kit has encountered a number of gain of function mutations that lead to its uncontrolled activation and thus cause diseases. A variety of mutations in the gene encoding the proto-oncogene KIT were found in different types of human cancer, in gastrointestinal stromal tumors (GISTs) , acute myeloid leukemia (AML) , mast cell leukemia (MCL) and human germ cell tumors , among others. The JMR mutations are considered regulatory as they disrupt the auto-inhibitory mechanism which negatively regulates the activity of the protein while the PTK mutations are considered catalytic as they directly affect the configuration of the enzymatic site probably by stabilizing the Aloop extended conformation. To this category belongs the mutation of D in position 816, most frequently substituted by V, found in most patients with mastocytosis, leukemia and germ cell tumors. D816V is also resistant to Imatinib (GleevecTM) treatment, which has motivated to study its role in KIT activation mechanisms. The molecular mechanism of the ligand-independent activation of the D816V mutant receptor has already been studied but the disruption of interaction of c-Kit and its negative regulators is not known yet.

The present work aims to account for structural changes in c-Kit in reponse to mutation D816V that leads to ligand independent activation of c-Kit and compare the structural alterations obtained after phosphorylation and also to observe the effect of mutation on the binding site of negative regulators of c-Kit, SHP1and SHP2.

4. METHODOLOGY

4.1 Sequence and structure analysis of c-Kit structures

The x-ray structures of c-Kit were obtained from Protein data bank (Berman *et al.,* 2000). The crystallographic structures of KIT cytoplasmic region (PDB entries: 1T45, 1PKG, 1QZK, 1QZJ, 1RO1, 3G0E, 3GOF) were analyzed using the bioinformatics protein visualization tool Pymol (The PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC) and multiple sequence alignment tool CLUSTALW (Larkin et al 2007). The superimposition of Kit structure was done using Pymol. The secondary structure prediction of the c-Kit structure was done with STRIDE (Heinig and Frishman 2004). Mutations were computationally created in the c-Kit structure using the SPDBV (Guex and Peitsch 1997) tool. The main purpose to analyze all the available pdb Id's of c-Kit was to select the appropriate structure for the present study.

4.2 Molecular dynamics simulations

The molecular dynamic simulation was done using NAMD (Phillips *et al.,* 2005) through the VMD (Humphrey *et al.,* 2006) interface. VMD is a molecular graphics program designed for the interactive visualization and analysis of biopolymers such as proteins, nucleic acids, lipids, and membranes. NAMD is a software used to perform molecular dynamics (MD) simulations that compute atomic trajectories by solving equations of motion numerically using empirical force fields, such as the CHARMM (Brooks *et al.,* 1983) force field, that approximate the actual atomic force in biopolymer systems. NAMD is a parallel molecular dynamics code designed for high-performance simulation of large biomolecular systems. NAMD works with AMBER (Case *et al*., 2012) and CHARMM potential functions, parameters, and file formats.

There are four basic requirements to run molecular dynamic simulation using NAMD. The first is the protein databank file of the protein/ complex under study that stores the atomic coordinates and/or the velocities for the system, second is the protein structure file (psf) that stores all molecule specific information to apply force field to the molecular system, third is the force field parameter file which contains the mathematical expression of potential which atoms in the system will experience and the last is the configuration file that which contains user specified options that NAMD should adopt in running a simulation. The CHARMM27 force field package used in this study is a combination of two files, topology file (needed to generate psf) and parameter file (supplies specific numerical values for the generic CHARMM potential function).

For an all-atom MD simulation, one assumes that every atom experiences a force specified by a model force field accounting for the interaction of that atom with the rest of the system. Today, such force fields present a good compromise between accuracy and computational efficiency. NAMD employs a common potential energy function that has the following contributions:

$$U\text{total} = U\text{bond} + U\text{angle} + U\text{dihedral} + U\text{vdW} + U\text{Coulomb} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

The first three terms in equation 1 describe the stretching, bending, and torsional bonded interactions, and The final two terms describe interactions between nonbonded atom pairs

The molecular dynamic simulation process is on a whole comprised of three steps: minimization, equilibration and finally the production run. The minimization step involves searching the landscape of the molecule for a local minimum i.e. a point at which the molecule is relaxed. On the contrary, equilibration involves molecular dynamics, whereby Newton's second law of motion is solved for each atom in the system to dictate its trajectory.

A molecular Dynamic simulation is a useful tool for investigating the structural and dynamical behaviour of a protein and also allows one to extract the temporal and structural fluctuations of both the protein and surrounding solvent at atomic resolution. The MD approach covers the time scale from femto- to nanoseconds, and it is warranted by the fact that the force fields employed are able to reproduce the spectroscopic experiments performed on the same scale.

Here we have performed MD simulation of wild type and mutant c-Kit in both phosphorylated and unphosphorylated state in order to determine the structural impact of the D816V mutation and phosphorylation on c-Kit structure.

First of all the simulation was carried out for the unphosphorylated wild type and mutant c-Kit The initial setup for the molecular dynamic simulation of c-Kit wildtype (WT) and c-Kit mutant was done with VMD 1.9.1 modules. The first step was the preparation of protein structure files (psf) for the wild type and the mutant with CHARMM27 topology file using the Autopsf plugin. Further, both the c-Kit wild type (5284 atoms) and mutant (5288) psf were subjected to solvation using the Solvate plugin in the modeling module of VMD. The water model TIP3 with box padding dimension of min (5, 5, 5) and max (10, 10, 10) was used. After solvation, the system comprised of 36155 atoms in total was ready to undergo through the first step of simulation i.e. minimization. The minimization of both the systems was done for 5000 steps using the conjugate gradient method which aims to select successive search directions (starting with the initial gradient) and eliminate repeated minimization along the same direction. The coor output file generated after minimization contains the coordinates of the minimized system and thus it is employed to save the minimized system in the pdb format. The minimized system is further subjected to equilibration in order to equilibrate the system at desired temperature prior to the final simulation run. Periodic boundary conditions with cell basis vector (x,y,z:: 73.53,68.91,77.84) was used. The system equilibration was done at 310 K for 100ps with NPT ensemble.The two NPT simulations, one for the wild type c-Kit and the other for the mutant c-Kit were run for a time period of 1ns on the equilibrated structures using NAMD through command line. The pressure was set to 1.01325 bar and temperature was set to 310 K using Langevin Dynamics . SHAKE algorithm was used for rigid bonds with integration timestep of 2.0 fs. The PME electrostatics were used to deal with the long range electrostatic interactions.

For the simulation of the phosphorylated c-Kit wild type and mutant structure specific changes were done in system setup**.** The Autopsf plugin of VMD was used to prepare the c-Kit psf with phosphorylated 568 (TYR) and 570 (TYR) residues of wildtype and mutant c-Kit. The phosphorylated residues suggest that the receptor is activated and these phosphorylated sites can now act as docking sites for SHP1 and SHP2. The phosphorylated structures were generated by using the topology file of CHARMM force field for phosphorylated residues. The phosphorylated, mutant and wild type c-Kit was subjected to solvation with TIP3 water model using the solvate plugin. After solvating the system, minimization was carried out for 5000 steps using conjugate gradient method. The minimized coordinates were saved and used for equilibration step. Equilibration was done for 100ps using the same parameters except for the force field where the structure file for the phosphorylated residues is appended in the CHARMM parameter file using the version 31 of CHARMM. Two NPT simulations, one for wild type phosphorylated c-Kit and the other for mutant phosphorylated c-Kit were run for 1ns time period on the equilibrated structures. The pressure was set to 1.01325 bar and temperature was set to 310 K using Langevin Dynamics. SHAKE algorithm was used for rigid bonds with integration timestep of 2.0 fs. The PME electrostatics were used to deal with the long range electrostatic interactions.

4.3 Analysis of MD simulations

All analyses were carried out using plugins included in VMD 1.9 and Pymol using the average structure obtained after the entire simulation run. Trajectories were subjected to energy analysis (electrostatic energy, vander waal energy and total energy) using the NAMD plot plugin as function of time. The RMSD trajectory tool of VMD was used to calculate the RMSD (root mean square deviation) of the trajectory using the initial zero frame as reference. The timeline plugin of VMD was employed for the calculation of Residue specific RMSD, Displacement per residue and number of hydrogen bonds along the trajectory. The secondary structure content prediction was done with STRIDE program. H Bond calculation (version 1.1), an online tool of Centre of Informational Biology, Ochanomizer University was used to calculate the residues involved in intra-molecular hydrogen bonding. Salt bridges between oppositely charged residues in the protein were investigated within a minimum cutoff distance of 0.5nm.

5. RESULTS

5.1 Sequence and structure analysis of c-Kit structures

We have first analyzed the available c-Kit structures which span the intracellular region in particular . The fasta sequences of the pdb id are obtained from Protein data bank are subjected to multiple sequence alignment using ClustalW as shown in Fig 6.

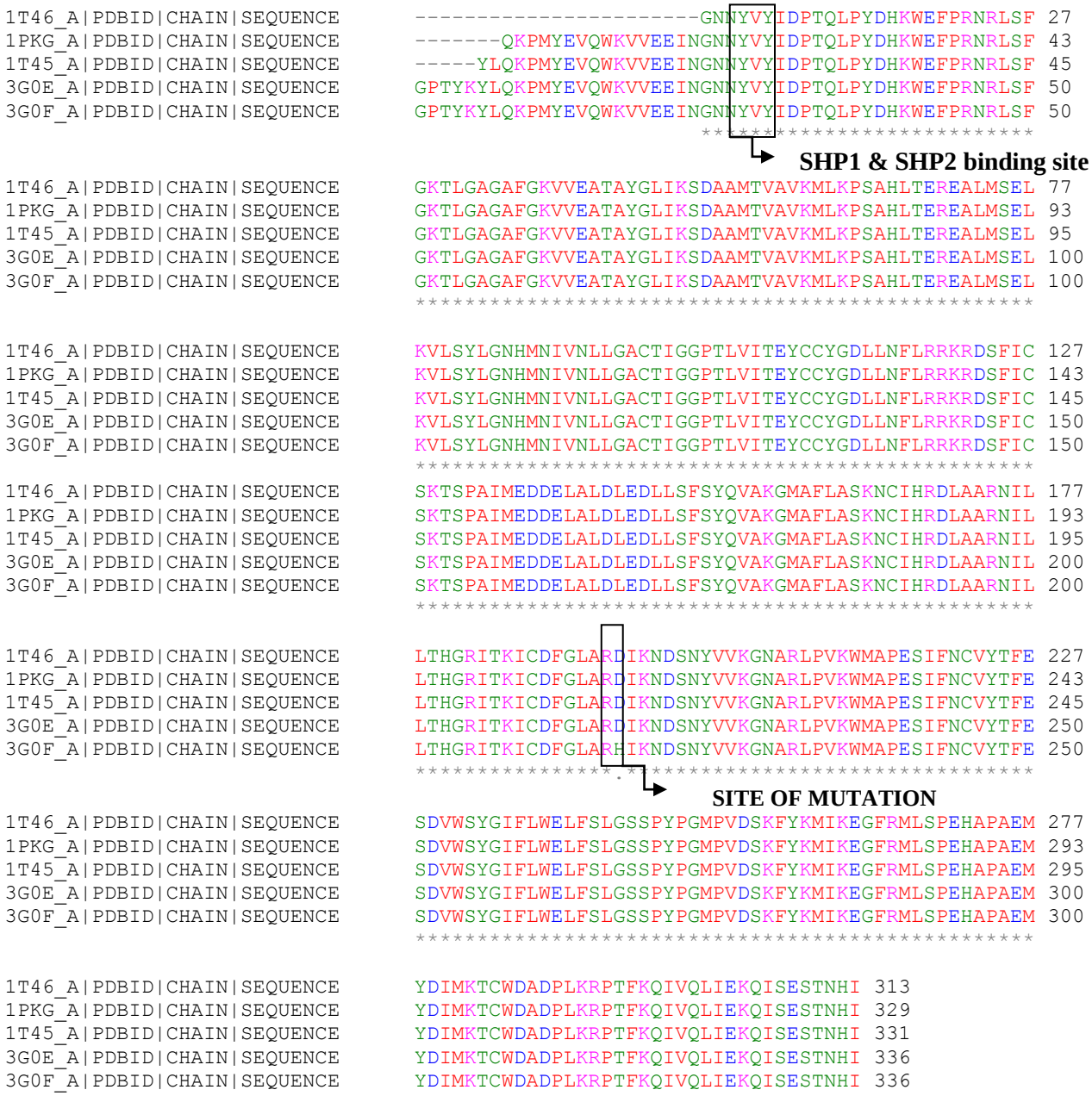


Fig 6: Muliple sequence alignment of c-Kit PDB Id's - 1PKG, 1T45, 1G0F, 1G0E and 1T46

The PDB Id's selected are such that they contain the site of SHP1 and SHP2 binding site residues i.e. 570 and 568 respectively along with the site of mutation under study i.e. 816 as shown in Fig 6.

The two very important pdb ids 1PKG (activated) and 1T45 (unactivated) of c-Kit were also subjected to secondary structure prediction using STRIDE as shown in Fig 7. Since we are here studying the structural impact of a activating mutation in c-Kit, thus structural differences in available crystals structures was used as a preliminary analysis.

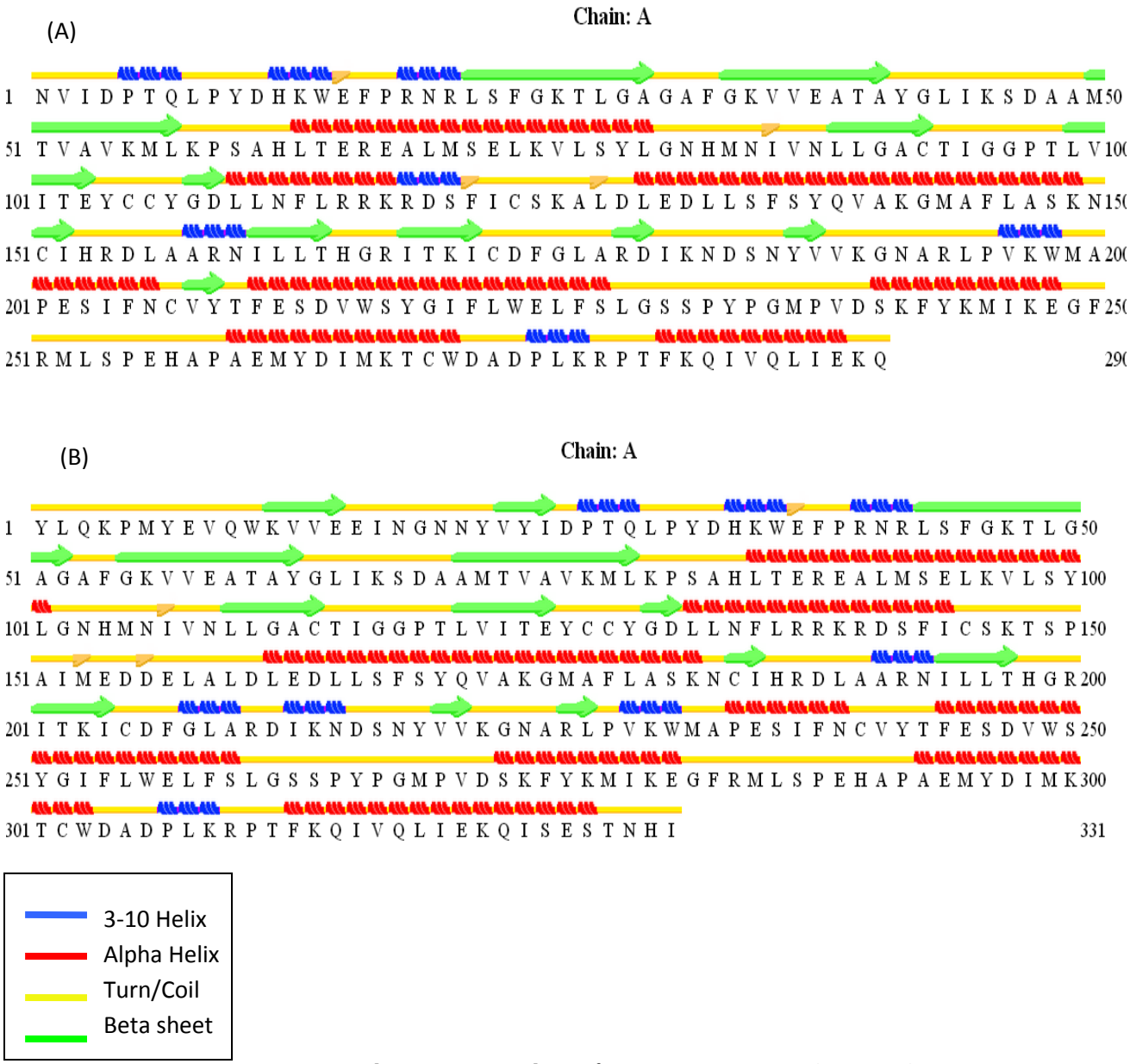


Fig 7: Secondary structure analysis of c-Kit using STRIDE A) 1PKG B) 1T45

Superimposition of PDB Id's of c-Kit was done to account for the location of D816 residue in all the structure in order to choose the appropriate structure for the present work (Fig 8).

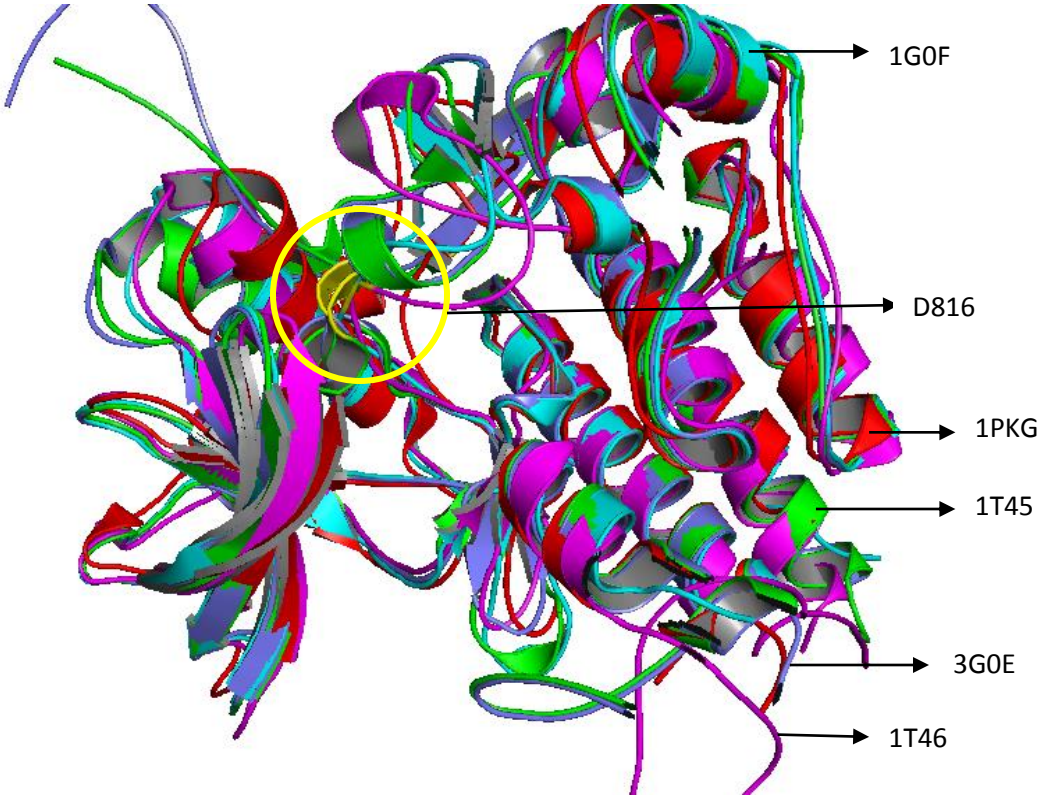


Fig 8: Superimposition of PDB ID of c-Kit. (1PKG (red), 1T45 (green), 3G0E (slate blue), 1G0F (cyan), 1T46 (pink))

The superimposition analysis of c-Kit structure was used confirm the secondary structure location of D816 residue so as to enable the selection of an appropriate structure for MD simulation analysis. Superimposition of structure shows changes in the location of D816 residue as in 1PKG it is present is turn , 1T45 it is present in helix , 1G0E it is present in helix , 1T46 it is present in loop and in 1GOF (mutant D816H) it is present in loop. Considering the wild type structures, 1T45 is chosen as the appropriate for our analysis, as it is auto-inhibited structure, so we can easily study the mechanism for ligand independent activation of c Kit on mutation , another reason is that the structure spans the residues where we wish to do the mutation as well as the residues where we wish to visualize the structural changes .

Superimposition analysis using Pymol were also done to analyze the effect mutant residue on the SHP1 and SHP2 binding residues i.e. Y 570 and Y 568 respectively. A known crystallographic mutant (D816H) structure of c-Kit 3G0F was also superimposed with 1T45 (Fig 9) to account for the structural changes in mutant site and SHP1 and SHP2 binding site on mutation.

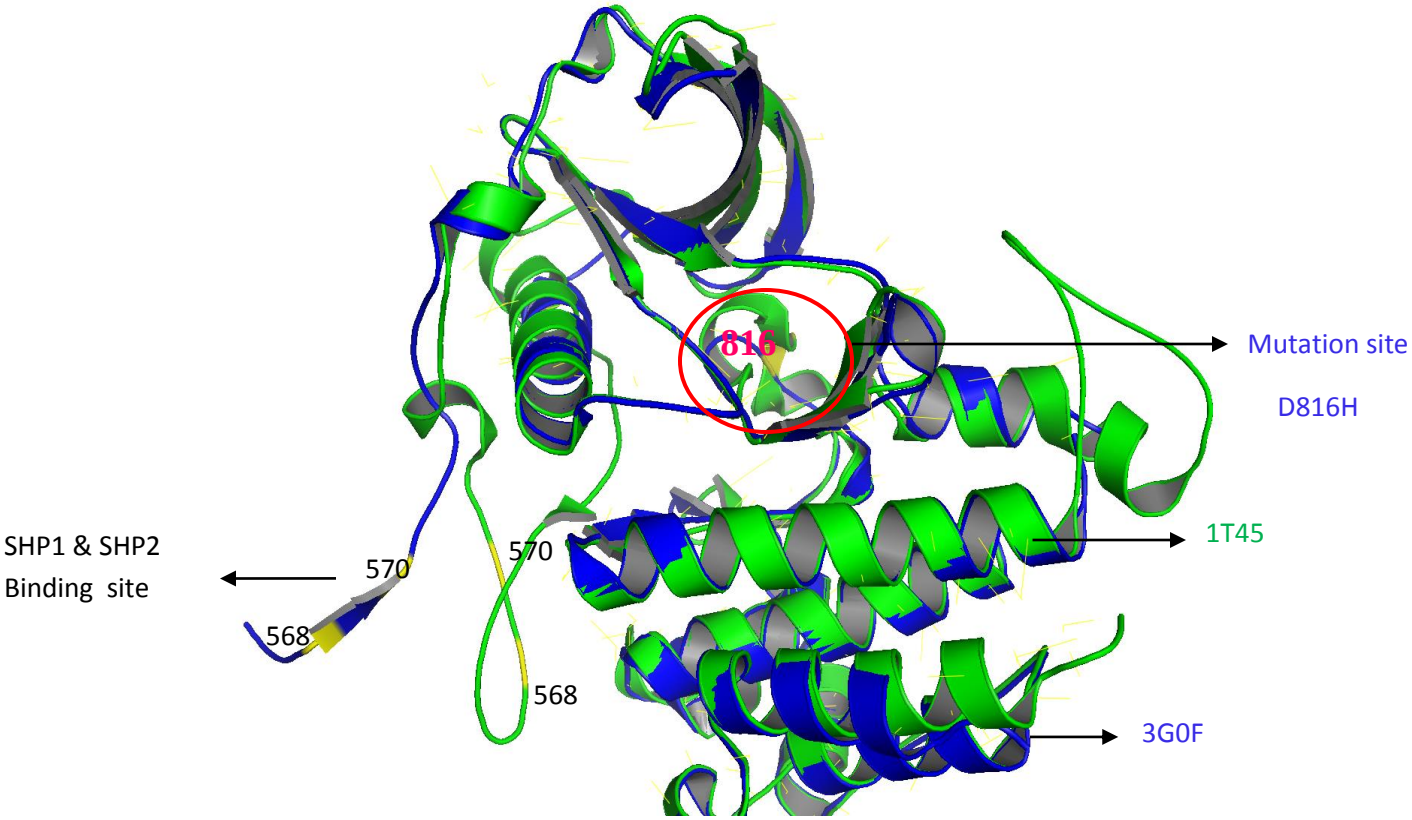


Fig 9: Superimposition of 1T45 (wild type) (green) and 3G0F (D816H mutated c Kit) (Blue). The highlighted residues in yellow marked as 570 and 568 depict the binding site for SHP-1 & SHP-2 respectively in both the structures.

As it is shown in the Fig 9, there is increase in flexibility on mutation site as in wild type the D816 residue was in helix (helix) and after mutation (D816H) the residue is seen in a loop (blue). On the other hand the residues of SHP1 (Y570) and SHP2 (Y568) binding site on mutation get converted to beta sheet from loop structure. This show how the active site of SHP binding undergoes structural modification on mutation at D816 residue. This structural difference in the known crystallographic mutant (D816H) structure was used as a basis for our present study.

5.2 Molecular Dynamic simulation (Minimization and Equilibration of the system)

Four NPT simulations, wild type c-Kit, mutant c-Kit , wild type phosphorylated c-Kit and mutant phosphorylated c-Kit were carried out using NAMD. Prior to the final production run, all the four systems were subjected to minimization for 5000 steps and equilibration for 100 ps. The achievement of minimization of the system was ensured by plotting the energy graphs after minimization. The total energy and the electrostatic energy plots obtained after minimization for the unphosphorylated (wild type and mutant) and phosphorylated (wild type and mutant) c-Kit clearly show that 5000 steps were enough to reach a local minima for the system (Fig 10 and Fig 11).

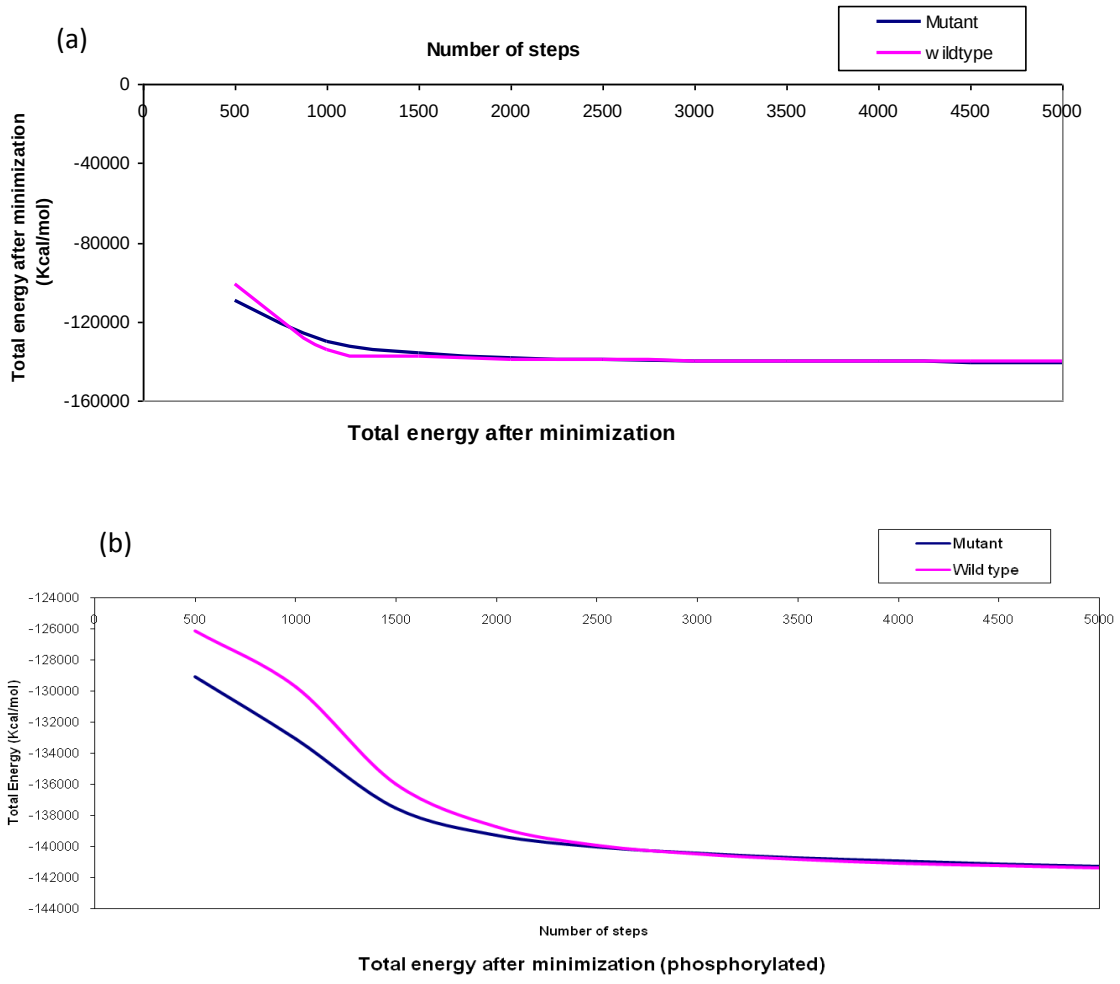


Fig 10: Total Energy after minimization for 5000 steps. a) Unphosphorylated (Wildtype and Mutant c-Kit) b) Phosphorylated (Wildtype and Mutant c-Kit)

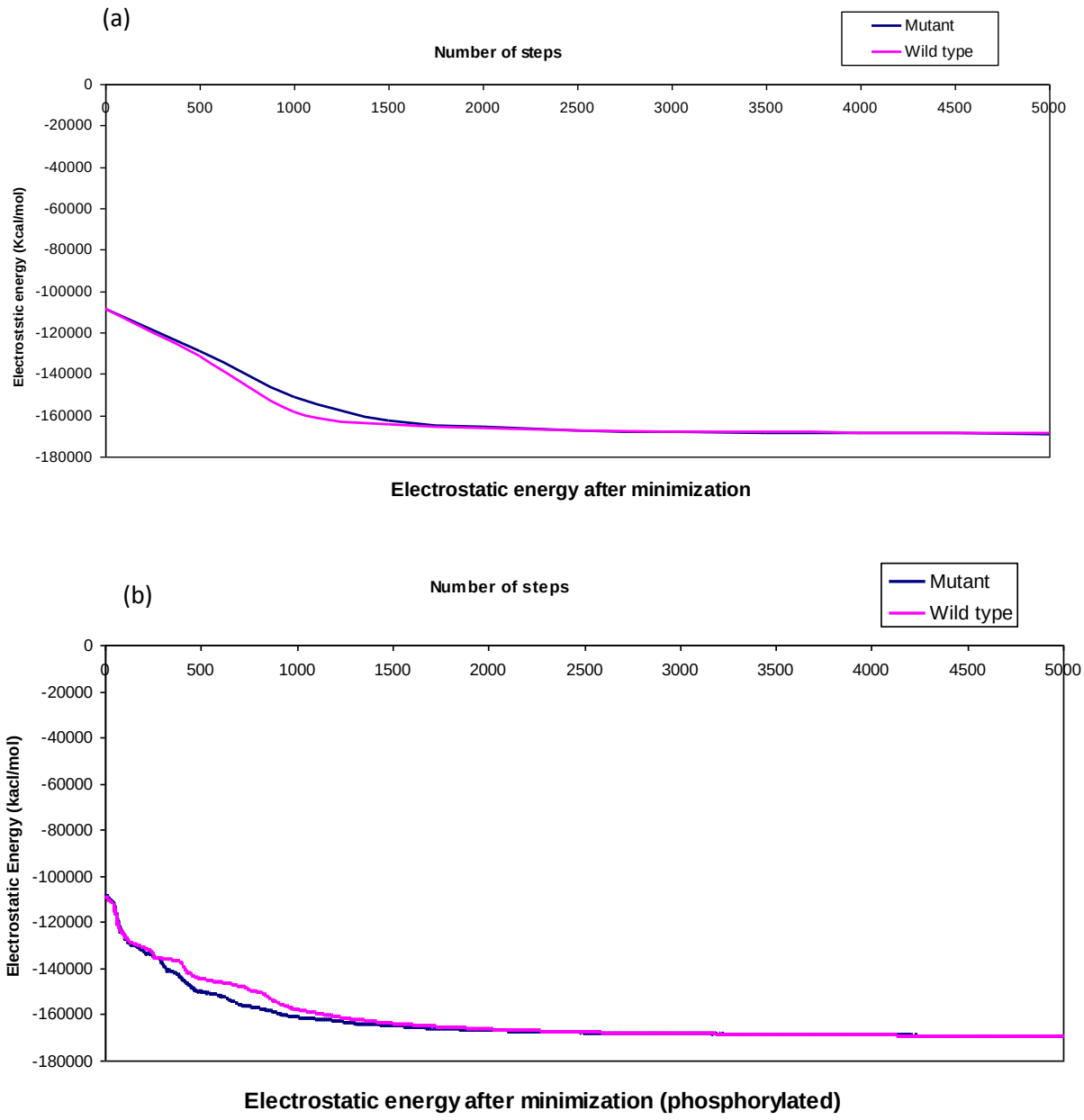


Fig 11: Electrostatic Energy after minimization for 5000 steps. a) Unphosphorylated (Wild type and Mutant c-Kit) b) Phosphorylated (Wild type and Mutant c-Kit)

After attaining the optimum minimization within 5000 steps, the minimized structures were equilibrated for 100ps in order to relax the system. Achievement of equilibration is judged by the fact how well the energy is distributed over the given amount of time. The total, electrostatic and vander waal energy plots obtained after equilibration (Fig 12, Fig 13, Fig 14) suggest that the energies are evenly distributed over100ps of time. In all the four systems, the energy profile reaches a stable landmark after 10000 steps.

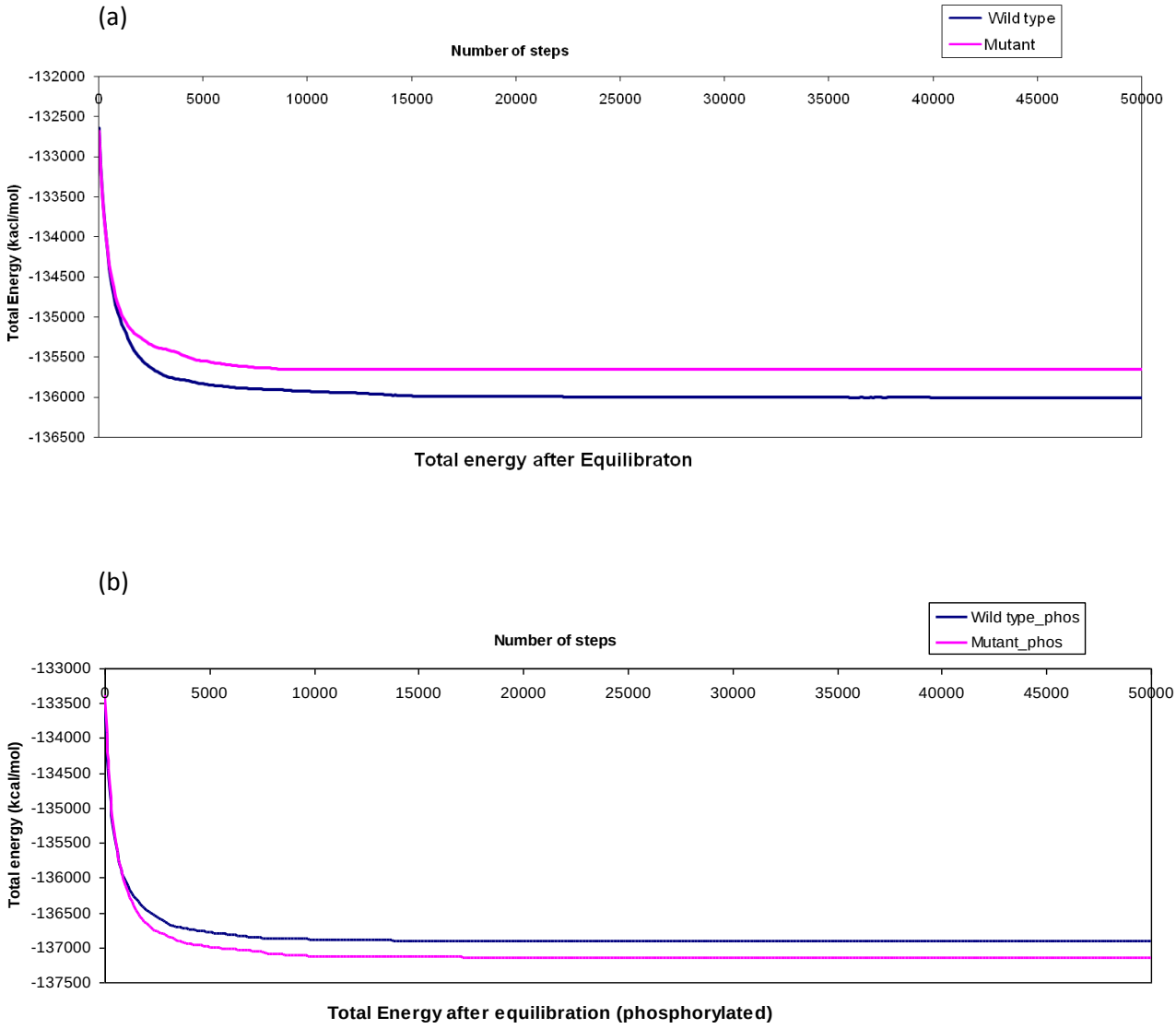


Fig 12: Total Energy after equilibration for 100ps (50000 steps). a) Unphosphorylated (Wildtype and Mutant c-Kit) b) Phosphorylated (Wildtype and Mutant c-Kit)

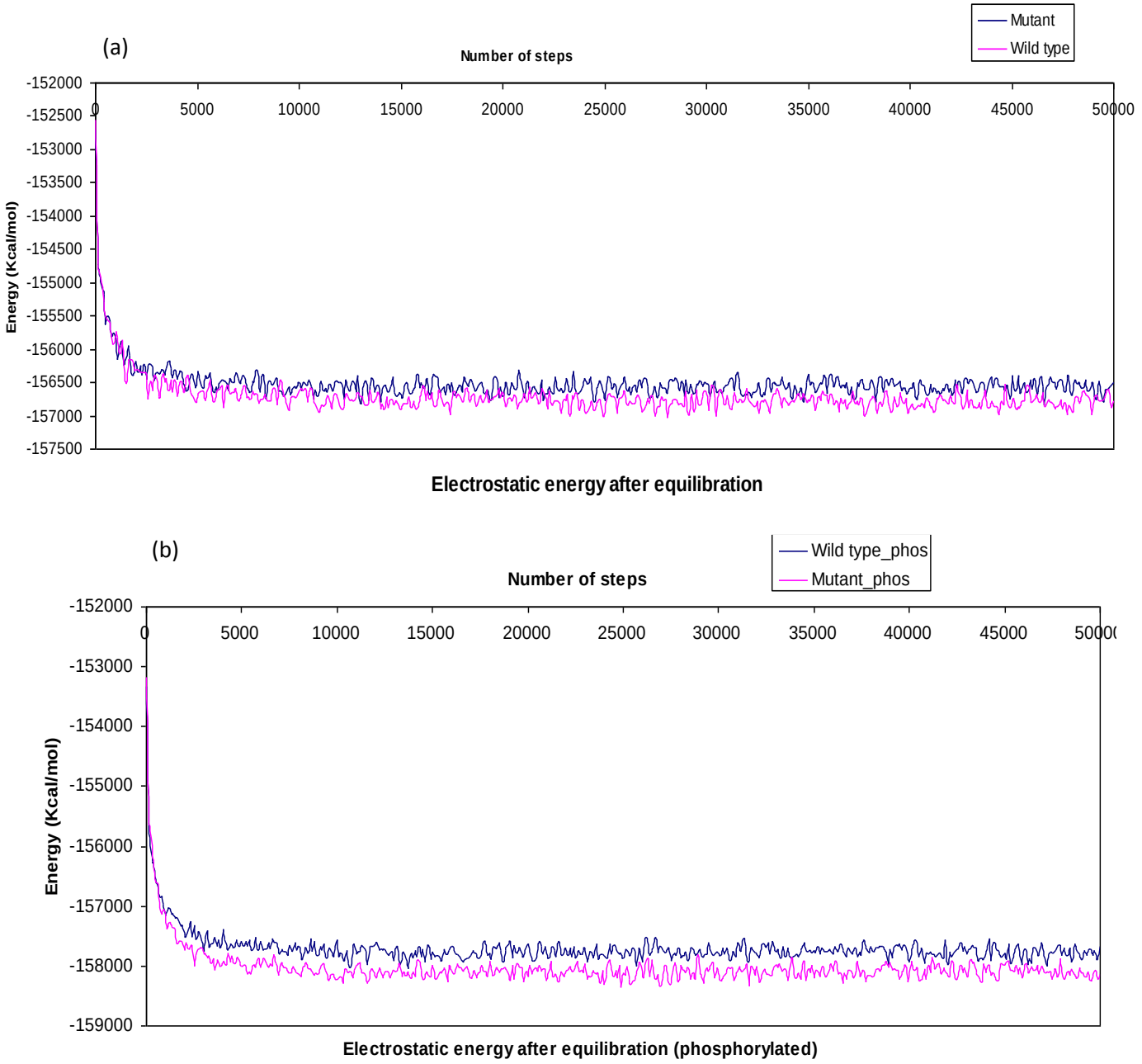


Fig 13 : Electroststic energy after equilibration for 100ps (50000 steps). a) Unphosphorylated (Wild type and mutant c-Kit) b) Phosphorylated (Wild type and mutant)

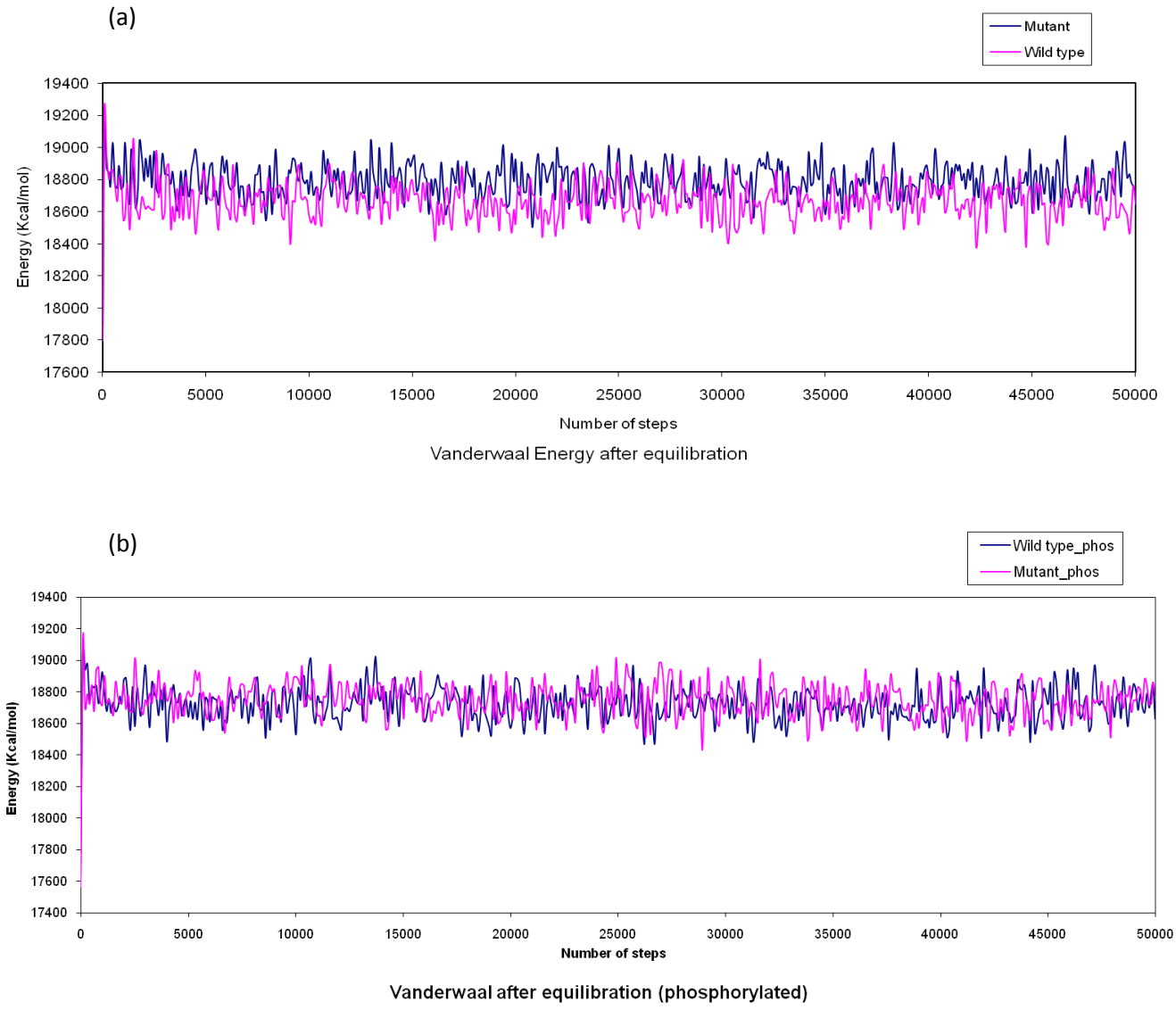


Fig 14: Vander waal Energy after equilibration for 100 ps (50000 steps) a) Unphosphorylated (Wildtype and Mutant) b) Phosphorylated (Wildtype and Mutant)

5.3 Analysis of MD Trajectories

As soon as the system setup was done after minimization and equilibration, the system was subjected to production run whose trajectories were analyzed to unleash the structural alterations in the unphosphorylated and the phosphorylated system. The main purpose of comparing the phosphorylated and the unphosphorylated system was to access the effect of mutation alone and also after activation i.e. phosphorylation. Since the binding of SHP-1 and SHP-2 occurs only upon activation, thus phosphorylation is an important player in binding of these negative regulators to c-Kit.

MD simulations were performed and analyzed to understand the protein internal motions and conformational changes within nanosecond time scale for D816V mutant c-Kit as compared to the wild type c-Kit both before phosphorylation and after phosphorylation.

Total four simulation trajectories (two for phosphorylated c-Kit and two for the unphosphorylated c-Kit) were analyzed. The wild type and mutant structures were analyzed to compare the alterations after mutation and further phosphorylation was done to access the changes in the SHP-1 and SHP-2 docking sites of c-Kit.

The main emphasis of the study was to enumerate the structural alterations caused by the D816V on the SHP-1 and SHP-2 binding residues. The study was focused on observing the conformational changes induced in the juxtamembrane region (541-581) by the mutation in the activation loop of c-Kit (810-839).

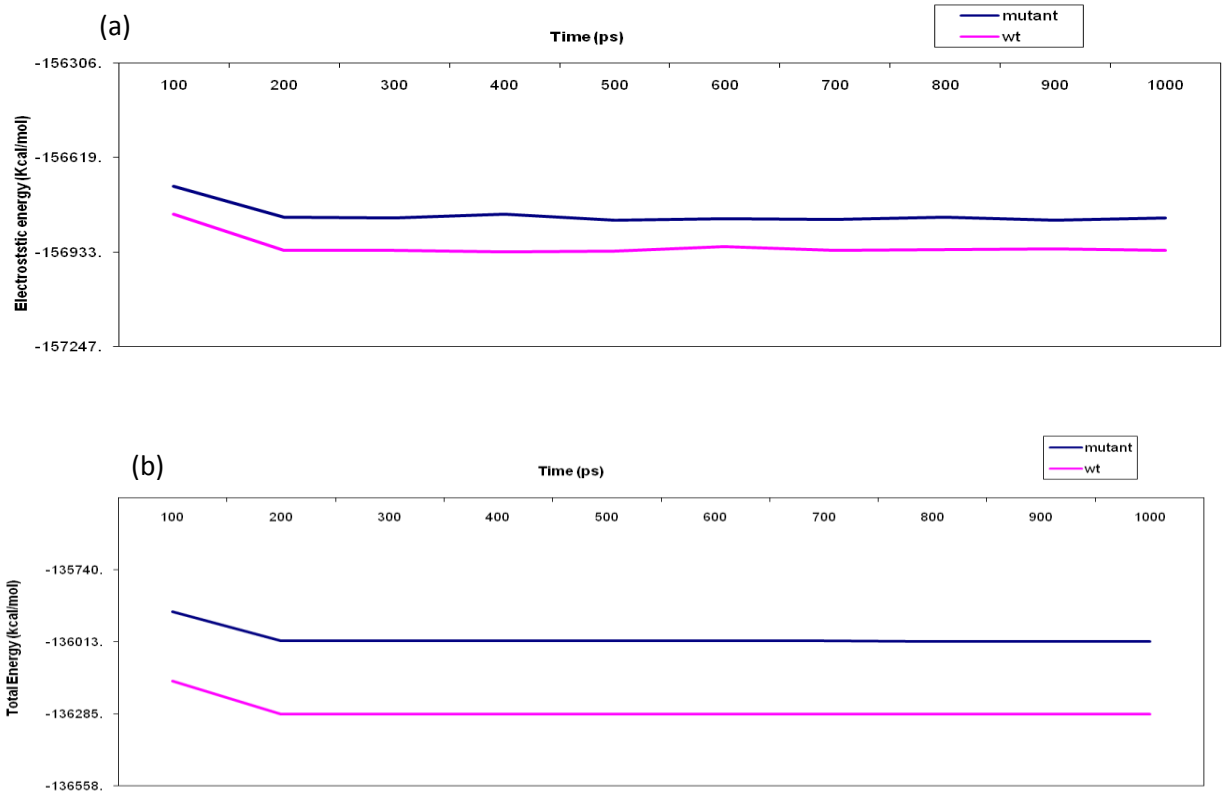
The simulation trajectories were subjected to energy analysis in order to reveal the stability of the protein along the course of time. The average energy of the four simulations is given in Table 2. The energy profiles of both the system were calculated for the entire 1000ps simulation. Fig 15 a, b and c show the trend of electrostatic energy , vander waal interations and the total enery (sum of potential and kinetic energy) respectively as a function of time for both the simulations of the unphosphorylated system. The electrostatic energy has higher for the mutant as compared to the wildtype (Fig 15a). Same can be observed in case of the vander waal interaction as shown in Fig 15c , that mutant c-Kit has a greater value than the wildtype. The greater values of non-bonded energies (electrostatic and vander waal) in case of the mutant c-Kit suggest that the trajectory in case of mutant is unstable as compared with the wildtype protein. Further stability analysis of the simulations were done by ploting the total energy along the entire MD run of 1ns. Fig 15b which plots the total energy (potential and kinetic energy) for the wild type and mutant trajectories shows that the energy tends to converge after 200ps and reaches minimum value of approx -135999 in case of mutant and -136273 in case of wild type c-Kit.

In case of the phosphorylated c-Kit simulations (wild type and mutant) , the energy plots (Fig 16 a,b,c) indicate aslo similar patterns as compared to the unphosphorylated system. The nonbonded energies i.e electrostatic and vander waal have a higher value for the phosphorylated mutant c-Kit as compared to the phosphorylation. This shows that phosphorylation along with mutation

also resulted in instablity of the structure as compared to wildtype. On the other the total energy for the phosphorylated system shows reverse pattern, the mutated has lower value as compared to wildtype.

Table 2: Average energies after 1ns MD for phosphorylated and Unphosphorylated c-Kit

Energy (Kcal/mol)	Unphosphorylated c-Kit		Phosphorylated c-Kit	
	Wildtype	Mutant (D816V)	Wildtype	Mutant (D816V)
Total Energy	-136273.2287	-135999.1517	-136597.87	-136824.47
Electrostatic Energy	-156915.2248	-156809.4667	-157996.91	-157585.81
Vanderwaal Energy	18510.91729	18693.66797	18528.53	18548.81



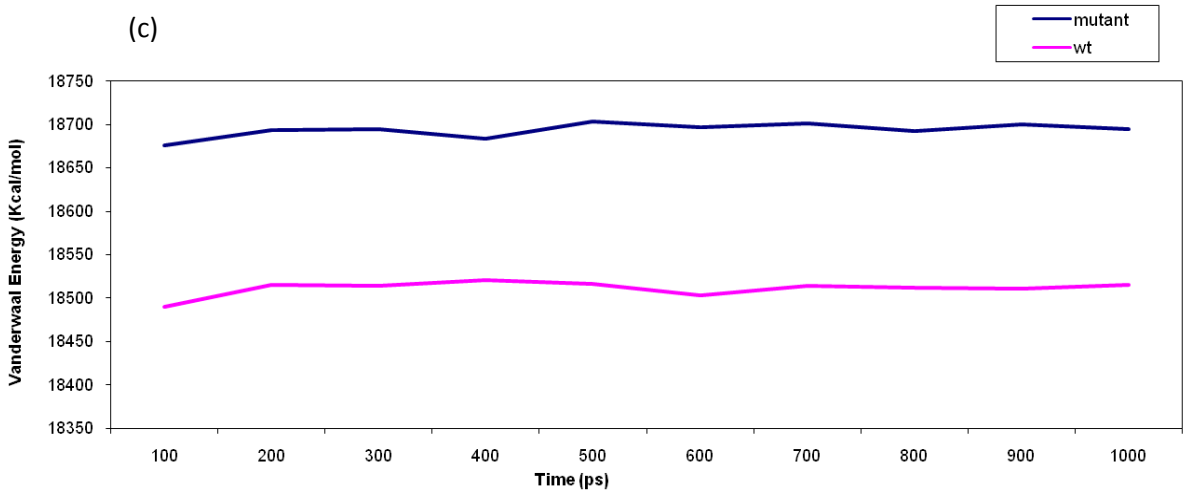
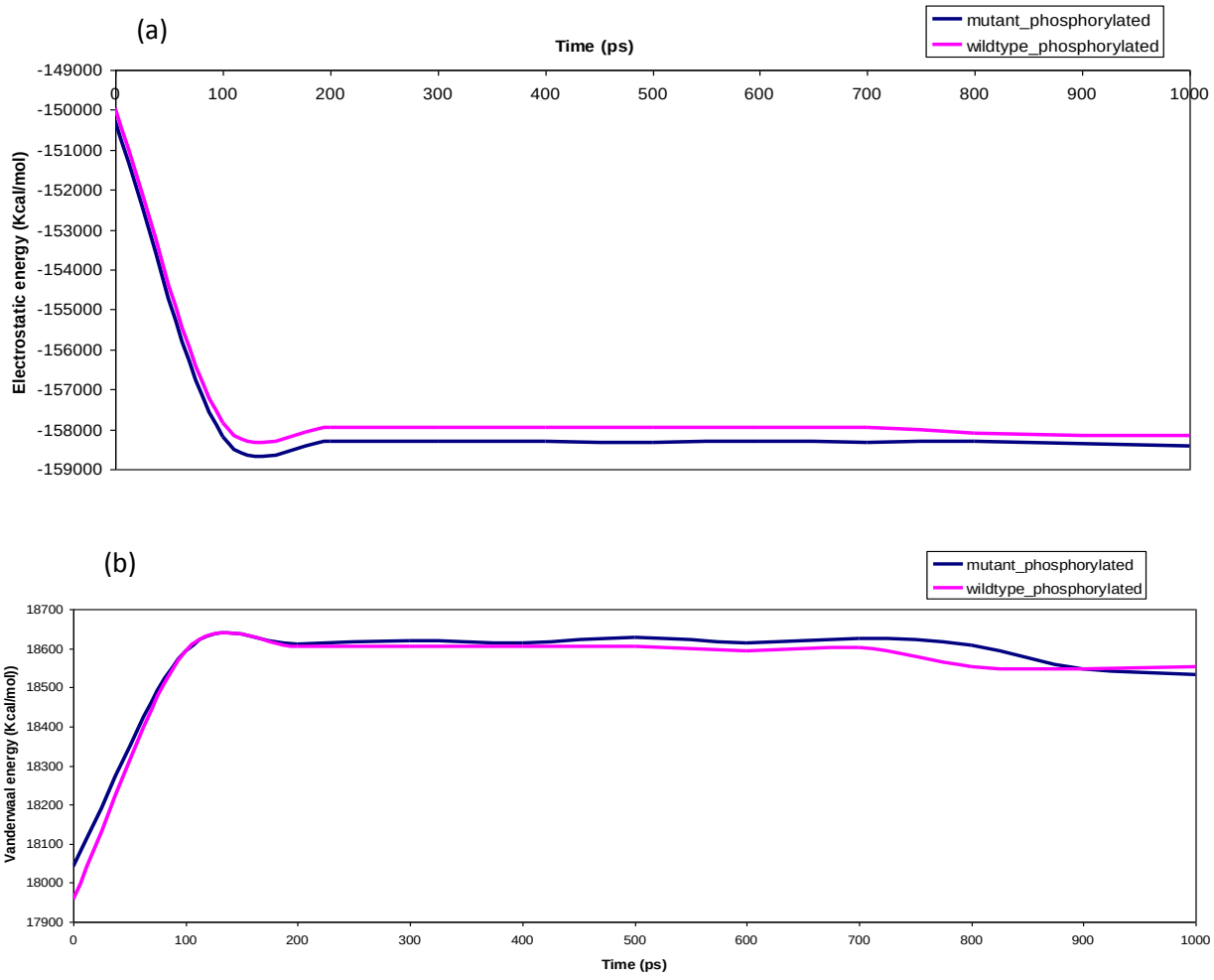


Fig 15: The energy profile along the entire trajectory for Unphophorylated c-Kit as a function of time a) Electrostatic interactions b)Vander waal interaction c) Total energy



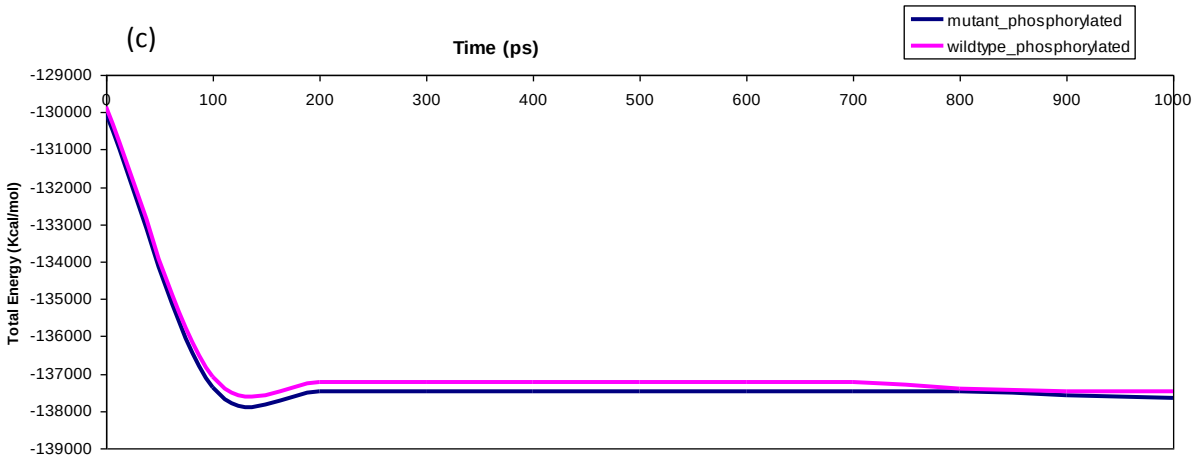


Fig 16: The energy profile along the entire trajectory for Phosphorylated c-Kit as a function of time a) Electrostatic interactions b) Vander waal interaction c) Total energy

In addition to energy calculations, the stability of the simulations was also checked by calculating the root mean square deviations (RMSDs) of the backbone atoms with respect to the starting X-ray structure was calculated and monitored for WT and D816V c-Kit for unphosphorylated and phosphorylated system, over the course of simulations and are presented in Table 3.

Table 3: RMSD over the trajectory

	Unphosphorylated		Phosphorylated	
	Wild type	Mutant	Wild type	Mutant
Average RMSD	0.274	0.297	0.282	0.275
Standard Deviation	0.007	0.007	0.008	0.008
Min	0.121	0.107	0.113	0.124
Max	0.281	0.305	0.290	0.289

The average RMSDs show that phosphorylation has induced instability in the wild type structure but the scenario is not the same for the mutant. The unphosphorylated mutant has graeter RMSD as compared to the phosphorylated.The RMSDs of the backbone atoms with respect to the refrence frame zero for the four simulations (Unphosphorylated Wt, Unphosphorylated Mutant, Phosphostylated wild type, Phosphorylated Mutant) as a function of frames are plotted in Fig 17a and 17b.

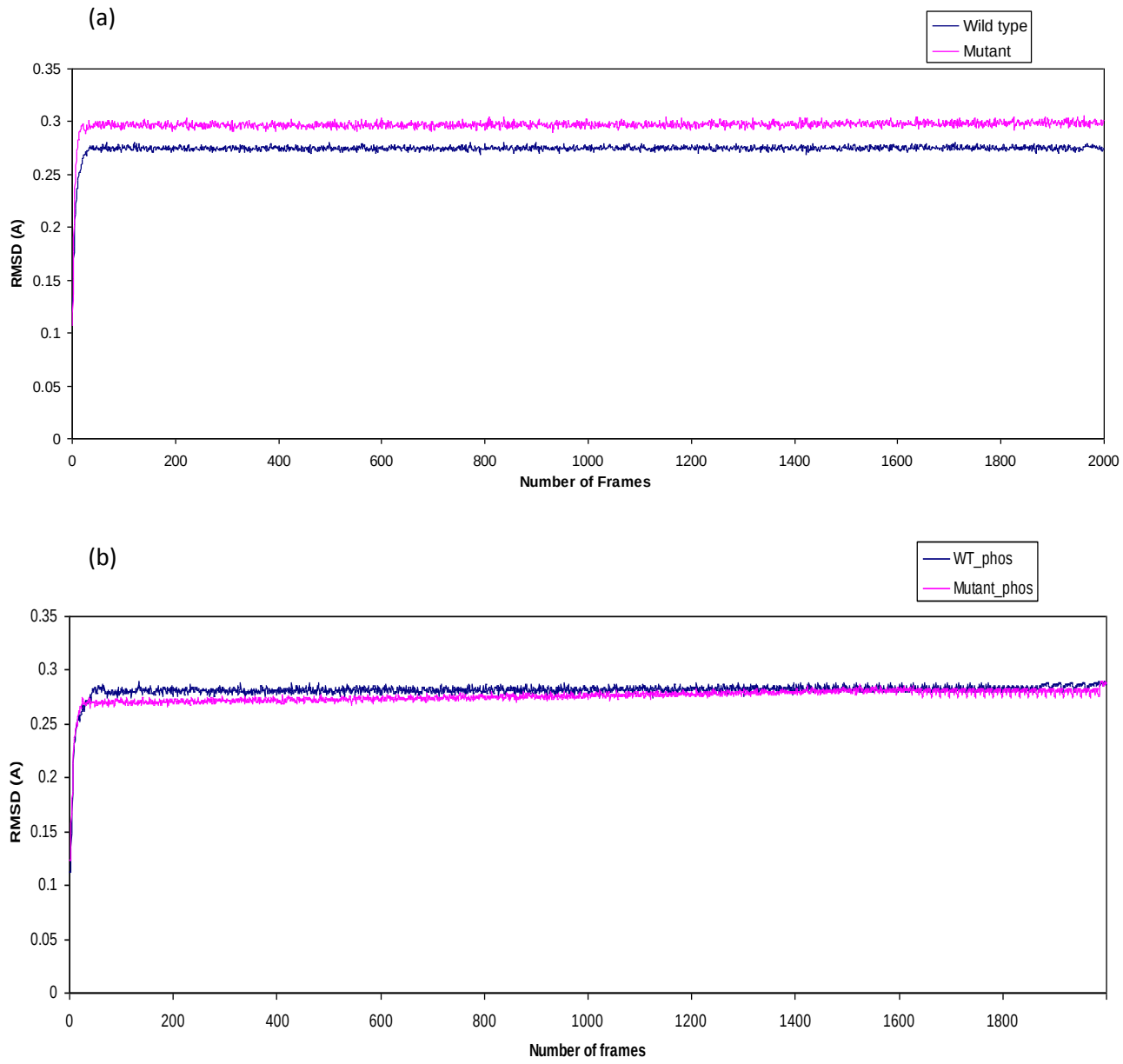


Fig 17: Backbone RMSD as a function of frame for wild type C-Kit and D816V mutant c-Kit. a) Unphosphorylated b) Phosphorylated.

From the Fig 17, it is clearly shown that for the both the systems (unphosphorylated and phosphorylated), the trajectory attains stability after a span of approximately 50 frames (25 ps). The RMSD plot for the unphosphorylated system reaches a plateau at 0.27 for wild type and 0.3 for mutant c-Kit and for the phosphorylated system the pattern is reversed i.e. wild type has a higher value of 0.282 as compared 0.275 of the mutant. The relative difference in the RMSD values and the irregular profile of the plots reflect the structural changes of highly flexible protein regions.

The average values of RMSD indicate that the mutation at 816 residue in the c-Kit causes instability in the c-Kit structure if the system is unphosphorylated. The lowering of the RMSD upon phosphorylation of the mutant suggests that phosphorylation induces stability.

As the main objective of the present study was to use MD simulation to model the effects of the D816V mutation on the Y568 and Y570 residues, thus a global analysis, such as the backbone RMSD, was not so informative in this regard. Therefore residue specific RMSD was calculated with respect to the initial frame of the trajectory. Fig 18a shows the residue specific RMSD pattern of wild type and mutant c-Kit for the unphosphorylated system. The graph shows that residues in the proximal kinase domain (582-684) undergo significant fluctuations as compared to the mutant which can be seen to have lower RMSD values in this region thus less flexibility. Another important domain probably affected by the mutation is distal kinase domain (762-935), where the residues in the mutant c-Kit have higher values of RMSD. This domain also contains the A-loop (810-839) region and plays important role in receptor activation by attaining an open conformation in the activated c-Kit. The residues of the A-loop upon mutation probably become more flexible as depicted by their high RMSD values as compared to the wild type.

Residue specific plot for the phosphorylated system show some significant peaks in case of the mutant (Fig 18b). The increase of the RMSD for the mutant can be seen in the Juxtamembrane domain (544-581) and the PTK domain (582-937) as compared to the wild type. But if we specifically observe the A-loop (810-839), then we can see that there is decrease in flexibility of this region upon phosphorylation whereas mutation alone induced flexibility.

In order to study the impact of mutation and phosphorylation on the SHP-1 and SHP-2 binding site, we can compare the RMSD values of these specific residues in all the four cases. In the unphosphorylated c-Kit, RMSD for residue 568 varies from 0.37 in wild type c-Kit to 0.46 in mutant c-Kit and for the residue 570 it varies from 0.34 in wild type c-Kit to 0.33 in mutant c-Kit. On the other hand after phosphorylation the variation for the residue 568 is 0.59 in wild type c-Kit to 0.87 in mutant c-Kit and for the residue 570, it varies from 0.29 in wild type c-Kit to 0.23. The RMSD changes at these specific residues upon mutation and phosphorylation suggest some flexibility changes at these positions.

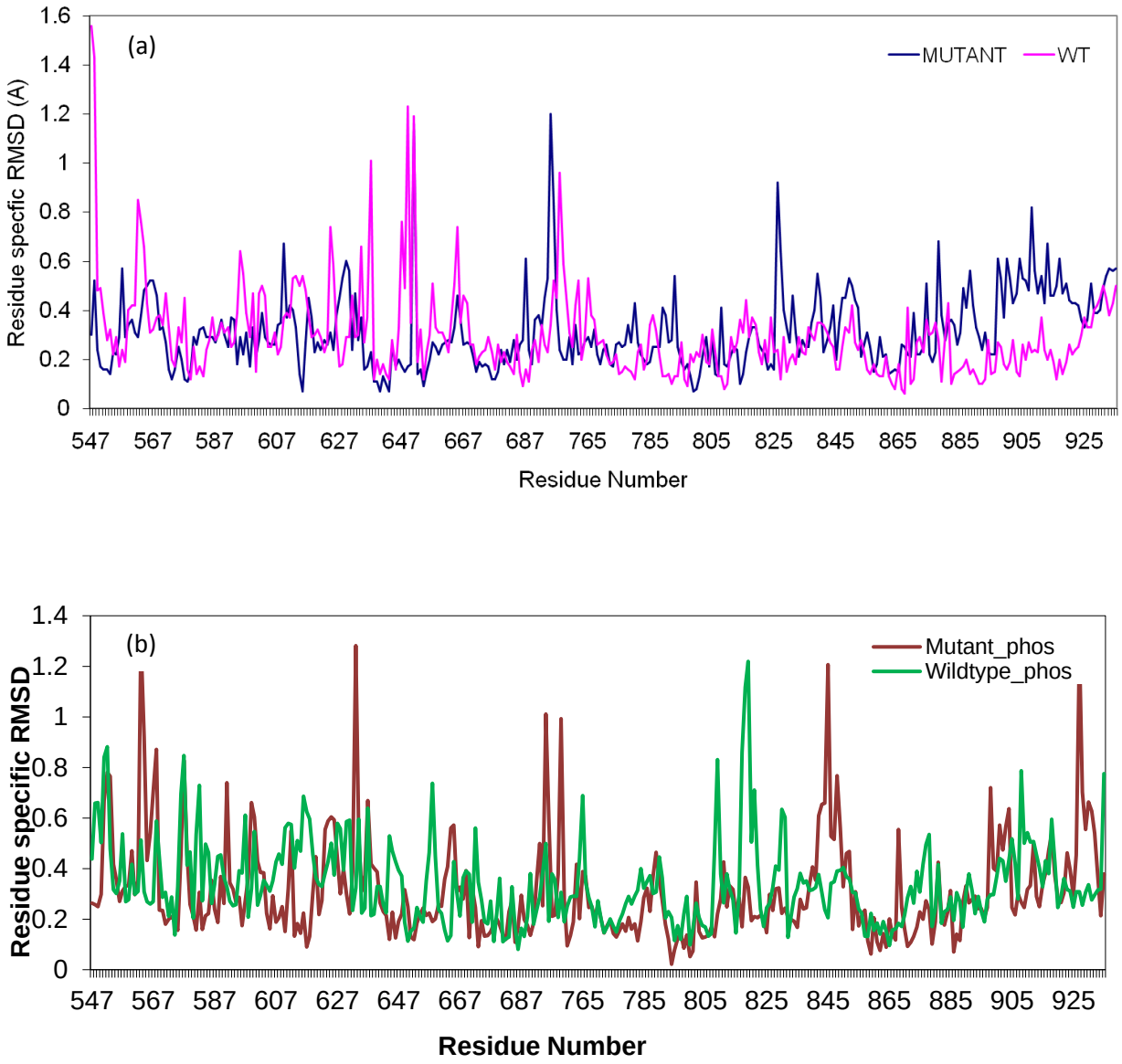


Fig 18: Residue Specific RMSD of a) Unphosphorylated b) Phosphorylated c-Kit .

The prediction of secondary structure helps to elucidate the structural modifications in the protein along the trajectory of the simulation run. The secondary structure prediction was done

using the average structure obtained after the simulations. The global analysis of the secondary structure content of the four average structures (Unphosphorylated: Wt and Mutant, Phosphorylated: Wt and Mutant) was done. In the unphosphorylated system, significant structural variations were observed in the Juxtamembrane region (544-581) along with some slight changes in the distal kianse domain (762-937) as well. The residues 795-797 for the unphosphorylated c-Kit showed a significant variation in its secondary upon mutation from 3^{10} helix to turns. Similar change was also observed in the residues 911-913. Also, in the phosphorylated system changes were in the JMR (544-581) and in the DKD (762-937). Residue 786 showed a change in structure from helix to coil upon mutation and the residues 911-913 changed from turn to 3^{10} helix.

Since the aim was to elucidate the effect of mutation and phosphorylation on SHP-1 binding, so we specifically emphasized on the secondary structure transitions in the JMR and activation loop region. Table 4 shows the secondary structure for the JMR and Activation loop for both the mutant and wild type simulated structures in the unphosphorylated and the phosphorylated states. It is clearly observed from the data that mutation in the A-loop has some effect on the secondary structure changes in the JMR region. The slight changes in the secondary structure of residues 555-557, 562-563,571-576 in JMR upon mutation suggest its role inducing structural alterations in SHP1 and SHP2 binding domains. Not only the JMR region is effected by the mutation , but also some of the residues of the domain containing the mutation (A-loop) undergo changes. The neighboring residues of the 816 i.e. 817,818,819 undergo structural alteration from 3^{10} Helix in wild type to turns in mutant. Thus these alterations suggest increase in flexibity upon mutation.

Table 4: Secondary Structure Prediction for the JMR and A-loop region for all four simulated structures (Highlighted regions indicate changes)

JMR Region Residue		Secondary structure WT	MU	WP	MP	A-loop Residue		Secondary structure WT	MU	WP	MP
547	TYR	C	C	C	C	810	ASP	C	C	C	C
548	LEU	C	C	C	C	811	PHE	C	C	C	C
549	GLN	C	C	C	C	812	GLY	G	G	G	G
550	LYS	C	C	C	C	813	LEU	G	G	G	G
551	PRO	C	C	C	C	814	ALA	G	G	G	G
552	MET	C	C	C	C	815	ARG	C	C	C	C
553	TYR	C	C	C	C	816	ASP	C	T	T	T
554	GLU	C	C	C	C	817	ILE	G	T	T	T
555	VAL	C	T	T	T	818	LYS	G	T	T	T
556	GLN	C	T	T	T	819	ASN	G	T	T	T
557	TRP	C	T	T	T	820	ASP	T	T	T	T
558	LYS	E	E	E	E	821	SER	T	T	T	T
559	VAL	E	E	E	E	822	ASN	T	T	T	T
560	VAL	E	E	E	E	823	TYR	T	T	T	T
561	GLU	E	E	E	E	824	VAL	E	E	E	E
562	GLU	T	C	C	C	825	VAL	E	E	E	E
563	ILE	T	T	T	C	826	LYS	T	T	T	T
564	ASN	T	T	T	T	827	GLY	T	T	T	T
565	GLY	T	T	T	T	828	ASN	T	T	T	T
566	ASN	T	T	T	T	829	ALA	T	T	T	T
567	ASN	T	T	T	T	830	ARG	E	E	E	E
568	TYR	C	C	C	C	831	LEU	E	E	E	E
569	VAL	E	E	E	E	832	PRO	T	C	C	C
570	TYR	E	E	E	E	833	VAL	G	G	G	G
571	ILE	E	C	C	C	834	LYS	G	G	G	G
572	ASP	C	T	T	C	835	TRP	G	G	G	G
573	PRO	G	T	T	G	836	MET	C	C	C	C
574	THR	G	T	T	G	837	ALA	C	C	C	C
575	GLN	G	T	T	G	838	PRO	H	H	H	H
576	LEU	C	T	T	C	839	GLU	H	H	H	H
577	PRO	C	C	C	C						
578	TYR	C	C	C	C						
579	ASP	C	T	C	T						
580	HSE	G	T	G	T						
581	LYS	G	T	G	T						

C:coil ,T:Turn, E: Strand ,H: Helix ,G: 310 Helix ; WT : Wild type, MU: Mutant WP: Wild type phosphorylated , MP: Mutant Phosphorylated.

It is already well known that intra-molecular hydrogen bonds play a very important role in the proper folding of the protein structure. The hydrogen distance is inversely proportional to its strength i.e more the HBond distance less strong is the bond. Hydrogen bonds also help to stabilize the 3D conformation of a protein. Fig 19 shows a plot of the number of Hbonds in the Wild type (WT) c-Kit and the mutant c-Kit along the trajectory as a function of the frames generated during the unphosphorylated and unphosphorylated simulations. The plot 19a clearly indicates that the number of HBonds has decreased in case of the unphosphorylated mutant c-Kit as compared with the wild-type c-Kit. Thus it suggests that some of the inherent Hbonds that were present in the wild type structure have been lost upon mutation at a single residue in the protein. Thus the mutation plays a major role in destabilizing the structure.

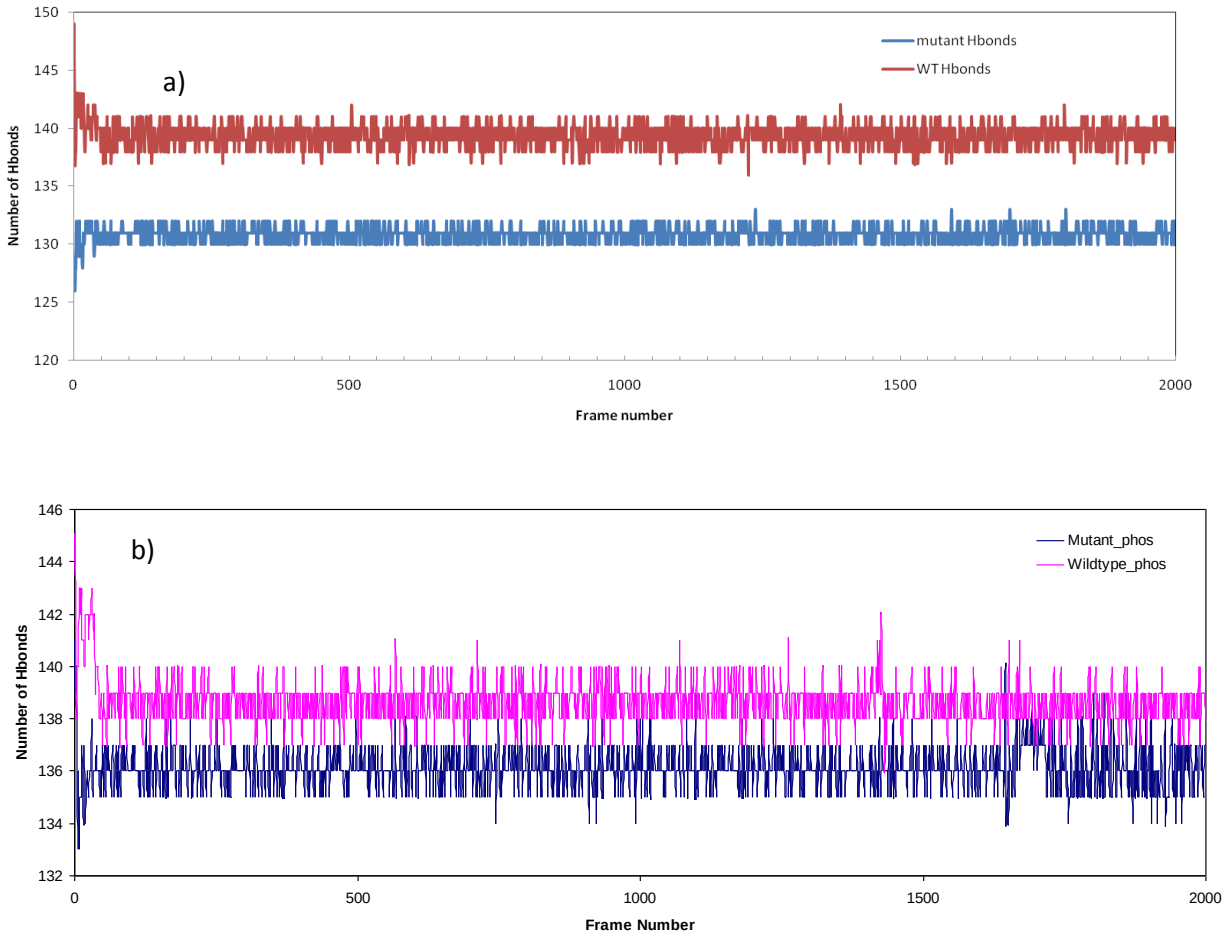


Fig 19: The number of Hydrogen bonds along the simulated trajectory as function of frame number a) Unphosphorylated b)Phosphorylated

To further account for the residues involved in the hydrogen bonding in the JMR and Activation loop of unphosphorylated and phosphorylated c-Kit, we here present the data in Table 5 and 6

below. The data shows specifically compares the hydrogen bonds formed in the JMR and A-loop region of the both the simulated structures (mutant and wild type). It is clearly visible that some of the hydrogen bonds are found only in case of mutant structure (highlighted in blue) and not in the wild type. This indicates that mutation in one domain of protein has affected the formation of new hydrogen bonds in some other domain.

Table 5 : Hydrogen bonds in JMR of unphosphorylated (wild type and mutant) and phosphorylated c-Kit (wild type and mutant) after simulation

Intramolecular Hydrogen Bonding in JMR Region of c-Kit							
M		W		MP		WP	
HBond	Dist	HBond	Dist	HBond	Dist	HBond	Dist
549 GLN-550 LYS	2.87	549 GLN- 550 LYS	2.9	549 GLN-550 LYS	2.91	549 GLN- 550 LYS	2.9
549 GLN-633 GLU	2.76	549 GLN -633 GLU	2.76	549 GLN -633 GLU	2.75	549 GLN-633 GLU	2.77
550 LYS -633 GLU	2.63	550LYS-633 GLU	2.64	550 LYS-633 GLU	2.62	550 LYS -633 GLU	2.67
553 TYR-623 LYS	3.07	553TYR -810 ASP	2.89	550 LYS-549 GLN	2.91	553 TYR -623 LYS	3.27
553 TYR-810 ASP	2.59	553 TYR -623 LYS	3.39	553 TYR -810 ASP	2.59	553 TYR-810 ASP	2.6
553 TYR - 640 GLU	3.36	554 GLU -558 LYS	2.62	554 GLU -558 LYS	2.73	554 GLU-791 ARG	3.1
554 GLU -791 ARG	2.68	555 VAL -558 LYS	2.7	554 GLU-791 ARG	3.12	554 GLU -558 LYS	2.82
554 GLU -558 LYS	2.63	556 GLN-810 ASP	1.77	555 VAL-791 ARG	3.06	556 GLN-791 ARG	2.84
555 VAL-558 LYS	2.73	556 GLN-791 ARG	2.76	556 GLN -791 ARG	1.78	557 TRP-809 CYS	2.16
555 VAL-791 ARG	3.38	558 LYS-789 ILE	2.03	558 LYS-789 ILE	1.93	558 LYS-789 ILE	1.96
556GLN -791 ARG	1.93	559 VAL-571 ILE	1.9	559 VAL-571 ILE	1.91	558 LYS -788 CYS	3.44
558 LYS-789 ILE	2.0	560 VAL-787 ASN	1.81	560 VAL-787 ASN	1.89	559 VAL-571 ILE	1.87
559 VAL-571 ILE	1.82	561 GLU-569 VAL	2.18	561 GLU-569 VAL	1.81	560 VAL- 787 ASN	1.84
560 VAL -787 ASN	1.8	561 GLU -562 GLU	2.35	565 GLY-567 ASN	2.96	561 GLU-569 VAL	1.78
561 GLU-569 VAL	1.88	563 ILE -565 GLY	1.94	570 TYR-558 LYS	3.18	564 ASN-567 ASN	1.89
564 ASN -567 ASN	1.89	564 ASN-567 ASN	2.89	571 ILE-559 VAL	1.91	565 GLY-567 ASN	2.89
565 GLY -567 ASN	3.07	565 GLY-567 ASN	3.26	572 ASP-574 THR	2.64	572 ASP-574 THR	2.58
568 TYR -849 GLU	3.06	568 TYR -849 GLU	2.99	574 THR -642 LYS	2.78	572 ASP -575 GLN	2.87
570 TYR - 846TYR	1.89	570 TYR-846 TYR	1.88	578 TYR-583GLU	2.61	574 THR-642 LYS	2.85
572 ASP-574 THR	2.56	572 ASP-574 THR	2.03	579 ASP-581 LYS	1.95	579 ASP-581 LYS	2.62
572 ASP -575 GLN	2.48	572 ASP 575 GLN	2.48	579 ASP -582 TRP	2.17	579 ASP-582 TRP	2.28
573 PRO-576 LEU	2.47	574 THR-642 LYS	2.79				
574 THR-642 LYS	2.97	578 TYR -583 GLU	2.62				
579 ASP-581 LYS	2.25	579 ASP-581LYS	2.65				
		579 ASP-582 TRP	2.3				

M: Mutant unphosphorylated, W:Wild type unphosphophorylated, MP: Mutant phosphorylated , WP: wild type phosphorylated

Table 6 : Hydrogen bonds in A-loop of unphosphorylated (wild type and mutant) and phosphorylated c-Kit after simulation

Intramolecular Hydrogen Bonding in A-Loop of c-Kit							
M		W		MP		WP	
Hbond	Dist	HBond	Dist	Hbond	Dist	Hbond	Dist
810 ASP-812 GLY	1.87	810 ASP-556 GLN	1.77	810 ASP-812 GLY	1.9	810 ASP-812 GLY	1.83
810 ASP -553 TYR	2.59	810 ASP-553 TYR	2.89	810 ASP-553 TYR	2.59	810 ASP -813 LEU	2.44
810 ASP -623 LYS	2.7	810 ASP-813 LEU	2.31	810 ASP-623 LYS	2.69	810 ASP-553 TYR	2.6
811 PHE-814 ALA	2.14	810 ASP-623 LYS	2.73	810 ASP- 809 CYS	3.35	810 ASP-623 LYS	2.67
813 LEU-598 GLY	2.03	811 PHE-814 ALA	2.25	811 PHE -814 ALA	2.29	811 PHE-797 ASN	2.34
815 ARG-796 ARG	3	812 GLY -815 ARG	2.44	812 GLY -815 ARG	2.27	812 GLY-815 ARG	2.46
815 ARG-797 ASN	3.03	813 LEU-598 GLY	2.43	813 LEU-598 GLY	2.16	815 ARG-823 TYR	3.46
815 ARG-677 ASP	2.69	815 ARG-796 ARG	3.33	815 ARG-796 ARG	3.02	815 ARG-796 ARG	3.27
816 ASP-819 ASN	1.98	815 ARG 797 ASN	2.81	815 ARG-797 ASN	2.9	815 ARG -797 ASN	2.86
817 ILE-830 ARG	3.3	815ARG-677 ASP	2.71	815 ARG -677 ASP	2.68	815 ARG-677 ASP	2.77
818 LYS-830 ARG	2.94	816 ASP-819 ASN	2.73	815 ARG-817 ILE	2.43	815 ARG-599 ALA	2.36
820 ASP-821 SER	2.04	817 ILE-830 ARG	3.2	816 VAL-819 ASN	1.85	816 ASP -819 ASN	2.79
820 ASP-822 ASN	2	818LYS-830 ARG	2.82	816 VAL-820 ASP	2.25	817 ILE-830 ARG	3.38
820 ASP-823 TYR	2.83	820 ASP-822 ASN	1.88	817 ILE-830 ARG	3.02	820 ASP-822 ASN	2.68
822 ASN-796 ARG	2.78	820 ASP-823 TYR	2.45	820 ASP -821 SER	2.09	822 ASN -796 ARG	3.04
824 VAL-831 LEU	1.81	822 ASN-796 ARG	3.05	820 ASP -822 ASN	1.91	823 TYR-792 ASP	2.61
825 VAL-880 TYR	3.11	823 TYR-792 ASP	2.61	820 ASP-823 TYR	2.77	823 TYR-830 ARG	3.48
826 LYS-829 ALA	1.92	824 VAL-831 LEU	1.83	822 ASN-796 ARG	2.83	824 VAL-831 LEU	1.91
826 LYS-841 ILE	2.67	826LYS-829 ALA	1.9	824 VAL-831 LEU	1.92	825 VAL- 880 TYR	3.24
832 PRO-835 TRP	1.98	826 LYS-841 ILE	2.73	825 VAL-880 TYR	3.29	826 LYS-829 ALA	1.83
833 VAL-836 MET	2.26	831 LEU-833 VAL	2.31	826 LYS-829 ALA	2.08	826 LYS-841 ILE	2.7
834 LYS-873 MET	2.74	832 PRO-835 TRP	1.98	826 LYS -841 ILE	2.85	829 ALA-826 LYS	1.83
834 LYS-870 TYR	3.07	833 VAL-836 MET	2.31	827 GLY-829 ALA	2.26	832 PRO-835 TRP	1.96
834 LYS-869 PRO	2.8	834 LYS-873 MET	2.77	829 ALA-826 LYS	2.08	833 VAL-836 MET	2.42
835 TRP-861 GLU	1.79	834 LYS-870 TYR	2.93	829 ALA-827 GLY	2.26	834 LYS-873 MET	2.8
835 TRP-854 SER	3.37	834 LYS-869 PRO	2.93	832 PRO-835 TRP	1.94	834 LYS-870 TYR	3
837 ALA-841 ILE	1.99	835 TRP-861 GLU	1.83	834 LYS-873 MET	2.66	834 LYS-869 PRO	2.8
838 PRO-842 PHE	2.08	837 ALA-841 ILE	2.15	834 LYS-870 TYR	3.28	835 TRP -832 PRO	1.96
839 GLU-914 ARG	2.77	837 ALA-840 SER	2.48	834 LYS-869 PRO	2.72	835 TRP-861 GLU	1.78
839 GLU-843 ASN	1.99	838 PRO-842 PHE	2.01	835 TRP -832 PRO	1.94	835 TRP-854 SER	3.49
		839 GLU-914 ARG	2.99	835 TRP-861 GLU	1.81	837 ALA-841 ILE	1.95
		839 GLU-843 ASN	2.01	835 TRP-854 SER	3.26	838 PRO-842 PHE	1.93
				837 ALA-841 ILE	2.07	839 GLU-839 GLU	1.87
				838 PRO -842 PHE	1.92	839 GLU-914 ARG	2.7
				839 GLU-839 GLU	1.94	839 GLU-843 ASN	1.89
				839 GLU-914 ARG	2.83		
				839 GLU-843 ASN	1.96		

M: Mutant unphosphorylated, W: Wild type unphosphorylated, MP: Mutant phosphorylated, WP: wild type phosphorylated

In addition to Hydrogen bonding , salt bridges are important players in stabilizing the protein structures in the folded conformations. Salt bridges were observed between positively charged

residues. The salt bridges obtained in the mutant and the wild type show no difference at all suggesting that the mutation has no effect on this protein stabilizing factor.

Table 7: Salt Bridge in unphosphorylated and phosphorylated c-Kit

	Unphosphorylated		Phosphorylated	
Sno	Wild type	Mutant	Wild type	Mutant
1	ASP579-LYS581	ASP579-LYS581	-	ASP579-LYS581
2	ASP615-LYS613	ASP615-LYS613	ASP615-LYS613	ASP615-LYS613
3	ASP677-ARG815	ASP677-ARG815	ASP677-ARG815	ASP677-ARG815
4	ASP760-ARG804	ASP760-ARG804	ASP760-ARG804	ASP760-ARG804
5	ASP760-LYS685	ASP760-LYS685	ASP760-LYS685	ASP760-LYS685
6	ASP768-ARG804	ASP768-ARG804	ASP768-ARG804	ASP768-ARG804
7	ASP792-ARG796	ASP792-ARG796	ASP792-ARG796	ASP792-ARG796
8	ASP810-LYS623	ASP810-LYS623	ASP810-LYS623	ASP810-LYS623
9	ASP901-LYS904	ASP901-LYS904	ASP901-LYS904	ASP901-LYS904
10	GLU554-ARG791	GLU554-ARG791	GLU554-ARG791	GLU554-ARG791
11	GLU554-LYS558	GLU554-LYS558	GLU554-LYS558	GLU554-LYS558
12	GLU605-LYS593	GLU605-LYS593	GLU605-LYS593	GLU605-LYS593
13	GLU633-LYS550	GLU633-LYS550	GLU633-LYS550	GLU633-LYS550
14	GLU635-LYS642	GLU635-LYS642	GLU635-LYS642	GLU635-LYS642
15	GLU640-LYS623	GLU640-LYS623	GLU640-LYS623	GLU640-LYS623
16	GLU671-LYS807	GLU671-LYS807	GLU671-LYS807	GLU671-LYS807
17	GLU758-LYS593	GLU758-LYS593	GLU758-LYS593	GLU758-LYS593
18	GLU839-ARG914	GLU839-ARG914	GLU839-ARG914	GLU839-ARG914
19	GLU861-ARG796	GLU861-ARG796	GLU861-ARG796	GLU861-ARG796
20	GLU885-LYS884	GLU885-LYS884	GLU885-LYS884	GLU885-LYS884
21	GLU930-LYS926	GLU930-LYS926	GLU930-LYS926	GLU930-LYS926

The displacement of each residue during the course of the trajectory can related be to the motions of the protein domains. The displacement of the residues can also be significant in specifying the open or closed conformation of the domains. The regions that displace the most can be considered to the most flexible and instable. Fig 20 clearly show the per residue displacement from the initial frame. For the unphosphorylated system, the juxtamembrane regions displace more in case of wild type while in the mutant the displacement is less. On the contrary towards the N terminal the residues displace more in case of mutant as compared to the wild type.On the other hand the scenario after phosphorylation is different, both the mutant and wild type show less displacement. In the JMR region, the residues in the mutant displace more than wild type and considering the A-loop region , the displacement is more in case of wild type.

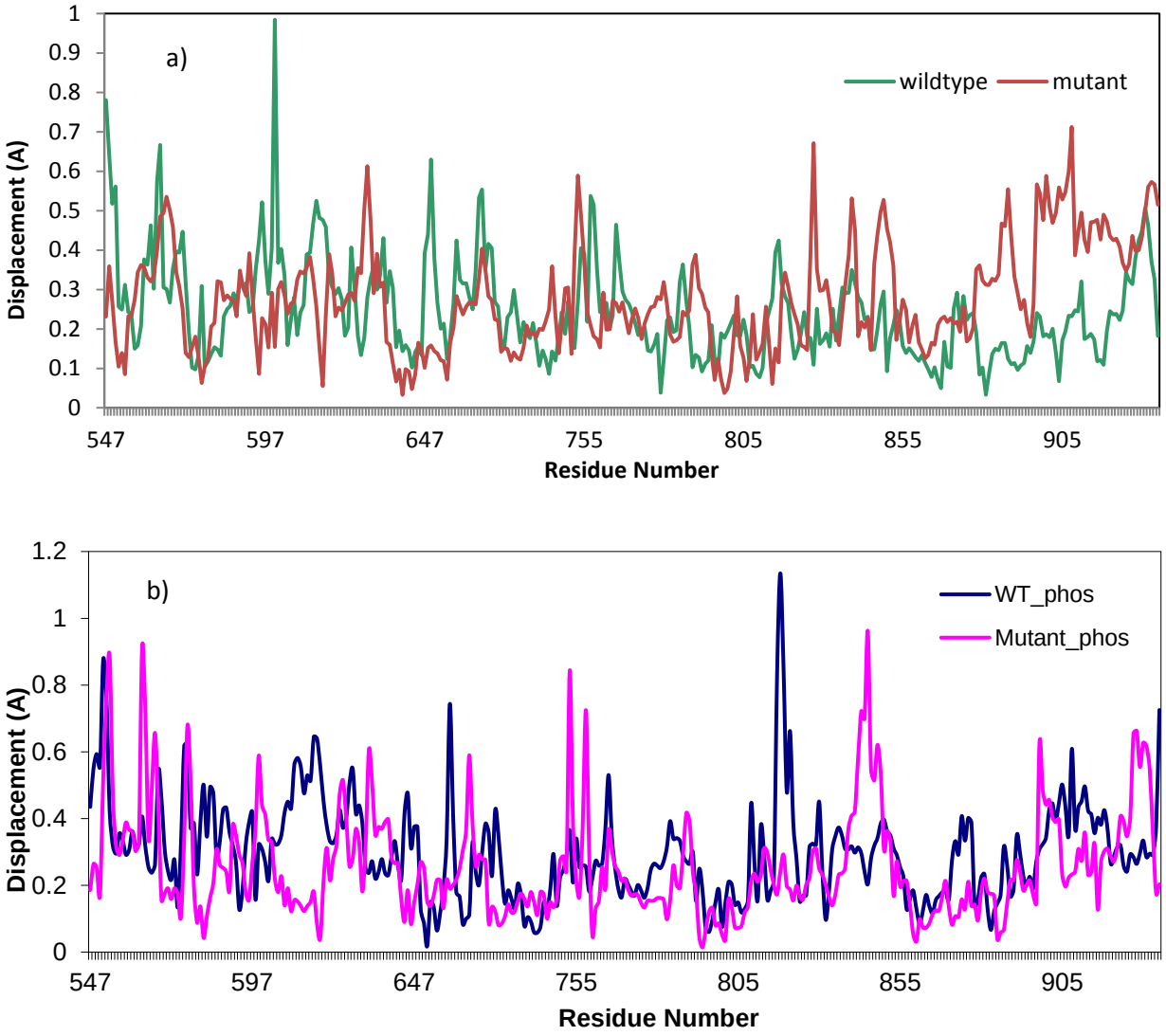


Fig 20 : Displacement per residue for c-Kit a) Unphosphrylated b) Phosphorylated

6. CONCLUSION AND FUTURE PERSPECTIVE

In this study we performed computational studies on wild type c-Kit and D816V mutant in unphosphorylated and phosphorylated state, using molecular mechanical methods. The MD trajectories of the D816V mutant, which is known to cause ligand independent activation of c-Kit and also implicated in mastocytosis showed some structural rearrangements as compared to the wild type structure. The simulation trajectories of the unphosphorylated wild type and mutant c-Kit upon analysis showed that mutation at a single residue has induced instability in the structure as indicated by the variations in RMSD and energy profiles. The structural alterations in the mutant c-Kit were specifically observed in the JMR region where significant secondary structure changes were observed from helix to turns, thus specifying the increase in flexibility of the structure upon mutation. The D816V mutation has also led to formation of new hydrogen bonds and disruption of some inherent hydrogen in the JMR and A-loop region. The fact that the mutation in the A-loop shows major significance on the JMR region is clearly indicated by the residue specific RMSD and per residue displacement plots. Thus we find that the D816V can effect the binding of SHP-1 and SHP-2 by studying the dynamic behavior of phosphorylated mutant on the nanosecond time scale and its deviation from the phosphorylated wild type.

The study proposes atomic level description level description of the regulatory impact of the D816V mutation on the c-Kit structure in general and SHP-1/SHP-2 binding site in particular.The conformational exploration of kinases, particularly the RTK KIT upon oncogenic mutation presents an obvious therapeutic interest. Understanding of the regulation/deregulation of kinase activation contributes to the design of novel generation of inhibitors targeting KIT and other related kinases.

7. DISCUSSION

The mechanisms of signal transduction by the wild type c-Kit are quite well characterized. Ligand binding to the receptor causes dimerization and activation of its intrinsic kinase activity, leading to phosphorylation of downstream signaling transduction molecules. In contrast the mechanism by which the oncogenic mutant D816V of c-Kit signals is partially known. The ligand independent activation of the D816V mutant c-Kit has been already explored (Bougherara *et al*., 2009), but the mechanism underlying the inability of SHP-1/SHP-2 to negatively regulate the mutant c-Kit is still unknown. The deactivation of tyrosine kinases or their oncogenic activation relates to mutations which affect the primary structure of the protein (Lennartsson *et al*., 2005, Kitamura and Hirotab 2004).

This study represents a detailed description at the atomic level of the impact of the D816V mutation on the KIT cytoplasmic region structure and internal dynamics in order to contribute to the basic concept that why the negative regulators of c-Kit (SHP-1 & SHP-2) do not have any effect on the mutant c-Kit. Unlike for many other kinases, the activation of c-Kit does not require the phosphorylation of the activation loop ,instead primary phosphorylation sites are present in the juxtamembrane region. A dual phosphorylated tyrosine containing conserved motif ($Y^{568}V^{569}Y^{570}$) present in the juxtamembrane region of c-Kit acts a common docking site for SH2 and SH3 domain containing protein (Lennartsson *et al*., 2005). SHP-1 and SHP-2 tyrosine phosphatases bind to Y^{570} and Y^{568} respectively. Our structural-based bioinformatics analysis of the KIT receptor auto-inhibited inactive and active states highlighted the strong polymorphous character of both A-loop and the JMR and also their crucial stabilizing role for both the conformation is known. These the conformational state of these two regions greatly influence the activation of the c-Kit (Huse and Kuriyan 2002, Nolen *et al*., 2004). The possible link of the A-loop and JMR, inspired us to study the impact of the the A-loop residue D816V mutation on the binding pockets of SHP-1 and SHP-2. Hence we chose to explore the conformational space of the Kit phosphorylated and unphosphorylated state in wild type and mutant forms using 1ns MD simulations.

In our simulations, JMR and A-loop regions demonstrated high flexibility upon mutation in the unphosphorylated state. More flexibility suggests that the conformational state of the protein is not stable in these two regions upon mutation. The flexibility of these regions is also attributed to the secondary structure transitions induced specifically in these regions. Our recording of the hydrogen bonds have also justified the flexibility of these as a number of inherent bonds have been lost upon mutation and conformational changes have induced formation of some new bonds.

From the various analysis performed on the phosphorylated system, a plausible mechanism for the disruption of SHP1 binding to c-Kit has been deduced here. One factor that possibly effects the binding is the change in the hydrogen bonding pattern observed in case of the activated mutant c-Kit. The binding site residues of SHP1 and SHP2 i.e. Y570 and Y568 respectively are specifically involved in intra-molecular hydrogen bonding with E849 and T846 residues of the distal kinase domain in the unphosphorylated wild type. But it is clearly observed that upon phosphorylation, these bonds get disrupted in the wild type c-Kit, in order to facilitate intermolecular binding with SHP1. But interestingly, the Y570 residue in case of the phosphorylated mutant c-Kit is seen to be involved in hydrogen binding with K558 residue . This intra-molecular bond is probably the reason for the inability of the phosphorylated mutant to interact with SHP1.

Another factor that has an effect on c-Kit and SHP1 interaction is the formation of new hydrogen bonds in the A-loop in case of the phosphorylated mutant. The three dimensional structure of c-Kit clearly indicates the close proximity of JMR and A-loop. Thus formation of new hydrogen bonds in A-loop (D820-S821, V816-D820, D820-Y823) reduce the capability of the A-loop to form intermolecular with SHP1 and play role in stabilizing protein-protein interaction. The formation of these new bonds can also be related to the decrease in flexibility of this domain as observed by the residue specific RMSD and per residue displacement plots.

In an attempt to elucidate the effect of mutation on SHP-1 and SHP-2 binding residues, specific residues in the A-loop are analyzed in the phosphorylated c-Kit. The intra-molecular hydrogen bonds between Asn819-Asp816 play an important role in stabilizing the A-loop and these residues are also important players in binding of SH2 domain of SHP-1 with c-Kit (Pati *et al*., 2010). The bond strength increases upon mutation, thus its disruption is difficult. This clearly suggests that in mutant c-Kit the binding pocket residues of c-Kit are unable to form hydrogen bonding with SHP-1.

8. REFRENCES

1. Arock, M; Valent, P. (2010). Pathogenesis, classification and treatment of mastocytosis: state of the art in 2010 and future perspectives. Expert Rev Hematol. ***3***, 497–516.
2. Berman, HM; Westbrook, J; Feng, Z; Gilliland, G; Bhat, TN; Weissig, H; Shindyalov, IN; Bourne, PE. (2000). The Protein Data Bank. Nucl Acids Res. ***28***: 235-242.
3. Berridge, MJ ; Bootman, MD ; Roderick, HL. (2003). Calcium signalling: dynamics, homeostasis and remodelling. Nat. Rev. Mol. Cell. Biol. ***4***,517–529.
4. Besmer, P; Murphy, JE; George, PC; Qiu, FH; Bergold, PJ; Lederman, L; Snyder, HW Jr; Brodeur, D; Zuckerman, EE; Hardy, WD. (1986) A new acute transforming feline retrovirus and relationship of its oncogene v-Kit with the protein kinase gene family. Nature. ***320***, 415–421.
5. Blume-Jensen, P ; Janknecht, R ; Hunter, T.(1998). The Kit receptor promotes cell survival via activation of PI 3-kinase and subsequent Akt-mediated phosphorylation of Bad on Ser136. Curr. Biol. ***8***, 779–782.
6. Blume-Jensen, P ; Wernstedt, C ; Heldin, CH ; Rönnstrand, L. (1995) . Identification of the major phosphorylation sites for protein kinase C in Kit/stem cell factor receptor in vitro and in intact cells. J. Biol. Chem. ***270***, 14192–14200.
7. Blume-Jensen,P ; Claesson-Welsh, L; Siegbahn, A ; Zsebo, KM ; Westermark, B ; Heldin , CH .(1991).Activation of the human c-Kit product by ligand-induced dimerization mediates circular actin reorganization and chemotaxis, EMBO J. ***10***,4121–4128.
8. Blume-Jensen,P ; Rönnstrand, L; Gout, I ; Waterfield, MD; Heldin, CH. (1994) .Modulation of Kit/stem cell factor receptor-induced signaling by protein kinase C. J. Biol. Chem. ***269***,21793–21802.
9. Blume-Jensen, P; Claesson-Welsh, L; Siegbahn, A; Zsebo, KM; Westermark,B; Heldin, CH. (1991).Activation of the human c-Kit product by ligand-induced dimerization mediates circular actin reorganization and chemotaxis. EMBO J. ***10***, 4121–4128.
10. Bondzi, C; Litz, J ; Dent, P; Krystal, GW.(2000). Src family kinase activity is required for Kit-mediated mitogen-activated protein (MAP) kinase activation, however loss of functional retinoblastoma protein makes MAP kinase activation unnecessary for growth of small cell lung cancer cells. Cell Growth Differ. ***11***, 305–314.
11. Bougherara, H; Subra, F; Crepin, R; Tauc, P; Auclair, C; et al. (2009) The aberrant localization of oncogenic Kit tyrosine kinase receptor mutants is reversed on specific inhibitory treatment. Mol Cancer Res. 7, 1525–1533.
12. Brizzi, MF ; Dentelli, P ; Rosso, A ; Yarden, Y ; Pegoraro, L.(1999). STAT protein recruitment and activation in c-Kit deletion mutants. J. Biol. Chem. ***274***, 16965–16972.
13. Brizzi, MF ; Zini, MG ; Aronica, MG; Blechmann, JM ; Yarden, Y ; Pegoraro, L.(1994).b. Convergence of signaling by interleukin-3, granulocyte- macrophage colony-stimulating factor, and mast cell growth factor on JAK2 tyrosine kinase. J. Biol. Chem. ***269***, 31680–31684.

14. Brooks, BR; Bruccoleri,RE; Olafson, BD; States,DJ; Swaminathan, S ; Karplus, M.(1983).CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations, J. Comp. Chem. ***4***, 187-217.

15. Carpenter,G ; Ji, Q.(1999).Phospholipase C-gamma as a signal-transducing element. Exp. Cell Res. ***253***,15–24.

16. Case, DA; Darden, TA; Cheatham, TE; Simmerling, CL; Wang, J; Duke,RE; Luo, R; Walker, RC; Zhang, W; Merz, KM; Roberts, B; Hayik, S; Roitberg, A; Seabra, G; Swails, J; Goetz, AW; Kolossváry, I; Wong, KF; Paesani, F; Vanicek, J; Wolf, RM; Liu, J; Wu,X; Brozell, SR; Steinbrecher, T; Gohlke, H; Cai, Q; Ye,X; Wang,J; Hsieh, MJ; Cui, G; Roe, DR; Mathews, DH; Seetin, MG; Salomon-Ferrer, R; Sagui, C; Babin, V; Luchko, T; Gusarov, S; Kovalenko, A; Kollman PA. (2012). AMBER 12. University of California, San Francisco.

17. Chabot, B; Stephenson, DA; Chapman, VM; Besmer, P; Bernstein, A. (1988). The proto-oncogene c-Kit encoding a transmembrane tyrosine kinase receptor maps to the mouse W locus. Nature. ***335***,88–89.

18. Chotinantakul, K; Leeanansaksiri, W. (2012). Hematopoietic Stem Cell Development, Niches, and Signaling Pathways. Bone Marrow Res.

19. Cohen P. (2002) Protein kinases-the major drug targets of the twenty-first century? Nat. Rev. Drug Discov. ***1*** ,309–315.

20. Corless, CL; Fletcher, JA; Heinrich, MC.(2004) .Biology of gastrointestinal stromal tumors, J. Clin. Oncol. ***22***, 3813–3825.

21. Crosier, P S ; Ricciardi ,ST; Hall ,LR ; Vitas ,MR; Clark ,SC; Crosier, KE. (1993). Expression of isoforms of the human receptor tyrosine kinase c-Kit in leukemic cell lines and acute myeloid leukemia. Blood.***82***,1151–1158.

22. De Sepulveda, P ; Okkenhaug, K ; Rose, JL ; Hawley, RG ; Dubreuil, P; Rottapel,R. (1999). Socs1 binds to multiple signalling proteins and suppresses steel factor-dependent proliferation. EMBO J. ***18***,904–915.

23. Fukao, T ; Yamada, T ; Tanabe, M ; Terauchi, Y ; Ota, T ; Takayama, T ; Asano, T ;Takeuchi, T; Kadowaki, T ; Hata Ji, J ; Koyasu, S.(2002). Selective loss of gastrointestinal mast cells and impaired immunity in PI3K-deficient mice. Nat. Immunol. ***3***, 295–304.

24. Furitsu, T; Tsujimura, T; Tono, T; Ikeda, H; Kitayama, H; Koshimizu, U; Sugahara, H; Butterfield, JH; Ashman, LK; Kanayama, Y; Matsuzawa, Y; Kitamura, Y; Kanakura, Y. (1993). Identification of mutations in the coding sequence of the proto-oncogene c-Kit in a human mast cell leukemia cell line causing ligand independent activation of c-Kit product. J Clin Invest. ***92***,1736, 1993.

25. Galli, SJ; Zsebo, KM; Geissler, EN. (1994). The Kit ligand, stem cell factor, Adv. Immun.***55***, 1–96.

26. Geissler, EN; Ryan, MA; Housman, DE. (1988). The dominant-white spotting (W) locus of the mouse encodes the c-Kit proto-oncogene. Cell ***55***,185–192

27. Gommerman, JL ; Sittaro, D ; Klebasz, NZ ; Williams, DA ; Berger, SA.(2000). Differential stimulation of c-Kit mutants by membrane-bound and soluble Steel Factor correlates with leukemic potential. Blood.***96*** , 3734–3742.

28. Gotoh, A ; Takahira, H ; Mantel, C ; Litz-Jackson, S ; Boswell, H.S ; Broxmeyer, HE.(1996). Steel factor induces serine phosphorylation of Stat3 in human growth factor-dependent myeloid cell lines. Blood.***88***, 138–145.

29. Griffith, J; Black,J; Faerman,C; Swenson, L; Wynn, M; Lu, F; Lippke,J; Saxena,K. (2004). The structural basis for autoinhibition of FLT3 by the juxtamembrane domain, Mol. Cell.***13***, 169–178.

30. Guex, N ; Peitsch, MC. (1997).SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. Electrophoresis.***18***,2714-2723.

31. Hayashi S; Kunisada,T; Ogawa, M; Yamaguchi,K; Nishikawa,S. (1991). Exon skipping by mutation of an authentic splice site of c-Kit gene in W/W mouse. Nucleic Acids Res. ***19***,1267–1271.

32. Heinig, M ; Frishman, D. (2004). STRIDE: a Web server for secondary structure assignment from known atomic coordinates of proteins. Nucl. Acids Res. ***32***, W500-2.

33. Heinrich, M; Blanke, CD; Druker, BJ; Corless,CL.(2002). Inhibition of KIT tyrosine kinase activity: a novel molecular approach to the treatment of KIT-positive malignancies, J. Clin. Oncol. ***20*** , 1692–1703.

34. Heinrich, MD; Corless, CL; Demetri,GD; Blanke,CD; Mehren, Mvon; Joensuu,H; McGreevey, LS; Chen,CJ; Van den Abbeele,AD; Druker,BJ; Kiese,B; Eisenberg,B; Roberts,PJ; Singer, S; Fletcher,CD; Silberman,S; Dimitrijevic,S; Fletcher,JA.(2003). Kinase mutations and imatinib response in patients with metastatic gastrointestinal stromal tumor, J. Clin. Oncol. ***21***, 4342–4349.

35. Herbst, R; Lammers, R; Schlessinger, J ; Ullrich, A.(1991). Substrate phosphorylation specificity of the human c-Kit receptor tyrosine kinase. J. Biol. Chem. ***266***, 19908–19916.

36. Hirota,S; Isozaki,K; Moriyama,Y; Hashimoto,K; Nishida,T; Ishiguro,S; Kawano, K; Hanada,M; Kurata,A; Takeda,M; Tunio,G; Matsuzawa,Y; Kanakura,Y; Shinomura,Y; Kitamura,Y.(1998). Gain-offunction mutations of c-Kit in human gastrointestinal stromal tumors, Science.***279*** , 577–580.

37. Hongyo,T; Li,T; Syaifudin,M; Baskar,R; Ikeda,H; Kanakura,Y; Aozasa,K; Nomura,T.(2000). Specific c-Kit mutations in sinonasal natural killer/T-cell lymphoma in China and Japan, Cancer Res. ***60***, 2345–2347.

38. Huang,EJ; Nocka,KH; Buck,J; Besmer,P. (1992). Differential expression and processing of two cell associated forms of the Kit-ligand: KL-1 and KL-2. Mol. Biol. Cell .***3***,349–362.

39. Hubbard, SR; Till, JH. (2000). Protein tyrosine kinase structure and function, Annu. Rev. Biochem.***69***, 373–398.

40. Humphrey, W ; Dalke, A ; Schulten, K.(1996). "VMD - Visual Molecular Dynamics". J. Molec. Graphics.vol. ***14***, pp. 33-38.

41. Huse M, Kuriyan J. (2002). The conformational plasticity of protein kinases. Cell. ***109***, 275–282.

42. Ilangumaran, S ; Finan,D ; Raine,J ; Rottapel,R.(2003).Suppressor of cytokine signaling 1 regulates an endogenous inhibitor of a mast cell protease. J. Biol. Chem. ***278***,41871–41880.

43. Kapur,R; Chandra,S; Cooper,R; McCarthy,J; Williams,DA. (2002).Role of p38 and ERK MAP kinase in proliferation of erythroid progenitors in response to stimulation by soluble and membrane isoforms of stem cell factor. Blood.***100***,1287–1293.

44. Kitamura,Y; Hirota,S.(2004). Kit as a human oncogenic tyrosine kinase, Cell. Mol. Life Sci. ***61*** ,2924–2931.

45. Kitayama,H; Kanakura,Y; Furitsu,T; Tsujimura,T; Oritani,K; Ikeda,H; Sugahara,H; Mitsui,H; Kanayama,Y; Kitamura,Y; Matsuzawa,Y.(1995).Constitutively activating mutations of c-Kit receptor tyrosine kinase confer factor-independent growth and tumorigenicity of factor-dependent hematopoietic cell lines, Blood.***85*** ,790–798.

46. Kozawa, O; Blume-Jensen, P ; Heldin, CH ; Rönnstrand, L. (1997). Involvement of phosphatidylinositol 3'-kinase in stemcell- factor-induced phospholipase D activation and arachidonic acid release. Eur. J. Biochem. ***248***,149–155.

47. Kozlowski, M; Larose, L; Lee, F ; Le, DM; Rottapel , R; Siminovitch, KA. (1998). SHP-1 binds and negatively modulates the c-Kit receptor by interaction with tyrosine 569 in the c-Kit juxtamembrane domain. Mol. Cell. Biol. ***18***, 2089–2099.

48. Krystal, GW; Hines, SJ; Organ,CP.(1996).Autocrine growth of small cell lung cancer mediated by coexpression of c-Kit and stem cell factor, Cancer Res. ***56***,370–376.

49. Kurosawa,K; Miyazawa,K; Gotoh,A; Katagiri,T; Nishimaki,J; Ashman,L. K. et al. (1996). Immobilized anti-KIT monoclonal antibody induces ligand-independent dimerization and activation of Steel factor receptor: biologic similarity with membrane-bound form of Steel factor rather than its soluble form. Blood.***87***,2235–2243.

50. Larkin, MA ; Blackshields, G ; Brown,NP ; Chenna, R ; McGettigan, PA ; McWilliam, H; Valentin , F ; Wallace, IM ; Wilm, A ; Lopez, R ; Thompson, JD ; Gibson, TJ ; Higgins, DG.(2007). ClustalW and ClustalX version 2. Bioinformatics ***23***(21),2947-2948.

51. Lemmon,MA; Pinchasi,D; Zhou,M; Lax,I; Schlessinger,J. (1997).Kit receptor dimerization is driven by bivalent binding of stem cell factor. J. Biol. Chem. ***272***,6311–6317.

52. Lennartsson, J ; Blume-Jensen, P ; Hermanson, M ; Ponten, E ; Carlberg, M ; Ronnstrand, L.(1999). Phosphorylation of Shc by Src family kinases is necessary for stem cell factor receptor/c-Kit mediated activation of the Ras/ MAP kinase pathway and c-fos induction. Oncogene.***18***, 5546–5553.

53. Lennartsson,J; Jelacic,T; Linnekin,D; Shivakrupa,R. (2005). Normal and oncogenic forms of the receptor tyrosine kinase Kit.Stem Cells.***23***, 16–43.

54. Lewis, TS; Shapiro, PS; Ahn , NG. (1998).Signal transduction through MAP kinase cascades. Adv. Cancer Res. ***74***,49–139.

55. Linnekin, D; DeBerry, CS ; Mou, S.(1997). Lyn associates with the juxtamembrane region of c-Kit and is activated by stem cell factor in hematopoietic cell lines and normal progenitor cells. J. Biol. Chem. ***272***, 27450–27455.

56. Lorenz,U ; Bergemann, AD ; Steinberg, HN ; Flanagan, JG ; Li, X ;Galli, SJ. et al. (1996). Genetic analysis reveals cell type-specific regulation of receptor tyrosine kinase c-Kit by the protein tyrosine phosphatase SHP1. J. Exp. Med. ***184***,1111–1126.

57. Lyman, SD; Jacobsen SE. (1998) c-Kit ligand and Flt3 ligand: stem/ progenitor cell factors with overlapping yet distinct activities. Blood ***91***, 1101–1134.

58. Miyazawa,K; Williams,DA; Gotoh,A; Nishimaki,J; Broxmeyer,HE; Toyama,K. (1995).Membrane-bound Steel factor induces more persistent tyrosine kinase activation and longer life span of c-Kit gene-encoded protein than its soluble form. Blood.***85***,641–649.

59. Mol, CD; Dougan, DR; Schneider, TR; Skene, RJ; Kraus, ML; Scheibe, DN; Snell, GP; Zou, H; Sang, BC; Wilson, WP. (2004a). Structural basis for the autoinhibition and STI-571 inhibition of c-Kit tyrosine kinase. J. Biol. Chem. ***279***, 31655–31663.

60. Mol, CD; Fabbro, D; Hosfield, DJ. (2004b). insights into the conformational selectivity of STI-571 and related kinase inhibitors, Curr. Opin. Drug Discov. Dev.***7***, 639–648.

61. Mol, CD; Lim, KN; Sridhar, V; Zou, H; Chien, EY; Sang, BC; Nowakowski, J; Kassel, DB; Cronin, CN; McRee, DE. (2003). Structure of a c-Kit product complex reveals the basis for kinase transactivation, J. Biol. Chem. ***278***, 31461–31464.

62. Nolen, B;Taylor, S; Ghosh, G. (2004) Regulation of protein kinases; controlling activity through activation segment conformation. Mol Cell. ***15***, 661–675.

63. Pawson,T. (2004). Specificity in signal transduction: from phosphotyrosine-SH2 domain interactions to complex cellular systems. Cell.***116***,191–203.

64. Phillips, JC; Braun, R; Wang, W; Gumbart, J; Tajkhorshid, E; Villa, E; Chipot, C; Skeel, RD; Kale, L; Schulten, K. (2005). Scalable molecular dynamics with NAMD. Journal of Computational Chemistry. ***26***,1781-1802.

65. Philo, JS; Wen,J; Schwartz,MG; Mendiaz,EA; Langley,KE. (1996).Human stem cell factor dimer forms a complex with two molecules of the extracellular domain of its receptor. Kit. J. Biol. Chem. ***271***,6895–6902.

66. Price, DJ; Rivnay, B ; Fu, Y ; Jiang, S ; Avraham, S ; Avraham, H.(1997). Direct association of Csk homologous kinase (CHK) with the diphosphorylated site Tyr568/570 of the activated c-KIT in megakaryocytes. J. Biol. Chem. ***272***, 5915–5920.

67. Qiu , RG; Chen, J; Kirn D., McCormick, F; Symons, M. (1995).An essential role for Rac in Ras transformation. Nature. ***374***: 457–459.

68. Radosevic, N ; Winterstein, D ; Keller, J.R ; Neubauer, H ; Pfeffer, K ; Linnekin, D.(2004). JAK2 contributes to the intrinsic capacity of primary hematopoietic cells to respond to stem cell factor. Exp. Hematol. ***32***, 149–156.

69. Reith, AD; Ellis, C; Lyman, SD; Anderson, DM; Williams, DE; Bernstein, A; Pawson T.(1991) Signal transduction by normal isoforms and W mutant variants of the Kit receptor tyrosine kinase. EMBO J. ***10***, 2451–2459.

70. Reya, T; Morrison, SJ; Clarke, MF; Weissman, IL.(2001) Stem cells, cancer, and cancer stem cells. Nature . ***414***, 105-111.

71. Rodriguez-Viciana, P ; Warne, PH ; Khwaja, A ; Marte , BM ; Pappin , D; Das, P; Waterfield, MD; Ridley, A; Downward, J. (1997).Role of phosphoinositide 3-OH kinase in cell transformation and control of the actin cytoskeleton by Ras. Cell. ***89***,457–467.

72. Roskoski Jr R. (2005).Signaling by Kit protein-tyrosine kinase—The stem cell factor receptor. Biochem Biophys Res Comm. ***337***, 1–13.

73. Rottapel, R., Reedijk, M., Williams, D.E., Lyman, S.D., Anderson, D.M., Pawson, T., Bernstein, A., 1991. The Steel/W transduction pathway: Kit autophosphorylation and its association with a unique subset of cytoplasmic signaling proteins is induced by the Steel factor. Mol. Cell. Biol. 11, 3043–3051.

74. Serve , H; Hsu, YC; Besmer, P.(1994). Tyrosine residue 719 of the c-Kit receptor is essential for binding of the P85 subunit of Phosphatidylinositol (PI) 3- Kinase and for c-Kit-associated PI 3-Kinase activity in COS-1 cells. J. Biol. Chem. ***269***, 6026–6030.

75. Sette, C ; Bevilacqua, A ; Geremia, R; Rossi, P.(1998). Involvement of phospholipase Cγ1 in mouse egg activation induced by a truncated form of the c-Kit tyrosine kinase present in spermatozoa. J. Cell Biol. ***142***, 1063–1074.

76. Shultz , LD; Schweitzer, PA ; Rajan, TV ; Yi, T ; Ihle, JN ; Matthews,RJ. et al. (1993) .Mutations at the murine motheaten locus are within the hematopoietic cell protein-tyrosine phosphatase (Hcph) gene. Cell .***73***, 1445–1454.

77. Tang, B ; Mano, H ; Yi, T ; Ihle, JN.(1994). Tec kinase associates with c-Kit and is tyrosine phosphorylated and activated following stem cell factor binding.Mol. Cell. Biol. ***14***, 8432–8437.

78. Thommes, K; Lennartsson, J ; Carlberg, M; Ronnstrand, L.(1999). Identification of Tyr-703 and Tyr-936 as the primary association sites for Grb2 and Grb7 in the c-Kit/stem cell factor receptor. Biochem. J. ***341***, 211–216.

79. Till, JE; McCulloch, EA . (1961). A direct measurement of the radiation sensitivity of normal mouse bone marrow cells. Radiat. Res. ***14***, 1419–1430.

80. Timokhina, I; Kissel, H ; Stella, G ; Besmer, P.(1998). Kit signaling through PI 3-kinase and Src kinase pathways: an essential role for Rac1 and JNK activation in mast cell proliferation. EMBO J. ***17***, 6250–6262.

81. Ueda, S ; Mizuki, M ; Ikeda, H ; Tsujimura, T ; Matsumura, I ; Nakano, K ; Daino, H; Honda, Zi; Z., Sonoyama, J ; Shibayama, H ; Sugahara, H ; Machii, T ; Kanakura, Y.(2002). Critical roles of c-Kit tyrosine residues 567 and 719 in stem cell factor-induced chemotaxis: contribution of src family kinase and PI3-kinase on calcium mobilization and cell migration. Blood. ***99***, 3342–3349.

82. Voytyuk, O; Lennartsson, J ; Mogi, A ; Caruana, G ; Courtneidge, S ; Ashman, LK ; Ronnstrand, L.(2003). Src family kinases are involved in the differential signaling from two splice forms of c-Kit. J. Biol. Chem***. 278***, 9159–9166.

83. Wormald, S ; Hilton, DJ. (2004). Inhibitors of cytokine signal transduction. J. Biol. Chem. ***279***, 821–824

84. Yarden, Y; Kuang, WJ; Yang-Feng, T; Coussens, L; Munemitsu, S; Dull, TJ; Chen, E; Schlessinger, J; Francke, U;Ullrich, A. (1987). Human proto-oncogene c-Kit: a new cell surface receptor tyrosine kinase for an unidentified ligand. EMBO J. ***6***,3341–3351.

85. Yi, T ; Ihle, JN.(1993). Association of hematopoietic cell phosphatase with c-Kit after stimulation with c-Kit ligand. Mol. Cell. Biol. ***13***, 3350–3358.

86. Young, SM ; Cambareri, AC ; Ashman, LK.(2006). Role of c-Kit expression level and phosphatidylinositol 3-kinase activation in survival and proliferative responses of early myeloid cells. Cell. Signal. ***18***, 608–620.

87. Zhang, Z; Zhang, R; Joachimiak, A; Schlessinger, J; Kong, XP. (2000). Crystal structure of human stem cell factor: implication for stem cell factor receptor dimerization and activation, Proc. Natl. Acad. Sci. USA. ***97***, 7732–7737.

9. APPENDIX

The University of Illinois at Urbana-Champaign has created its molecular dynamics software, NAMD, developed by the Theoretical Biophysics Group ("TBG") at Illinois' Beckman Institute available free of charge for non-commercial use by individuals, academic or research institutions and corporations for in-house business purposes only, upon completion and submission of the online registration form presented when attempting to download NAMD at the web site http://www.ks.uiuc.edu/Research/namd/.In order to run any MD simulation, NAMD requires at least four things:

• a Protein Data Bank (pdb) file
• a Protein Structure File (psf)
• a force field parameter file
• a configuration file.

A) PDB Files

The term PDB can refer to the Protein Data Bank (http://www.rcsb.org/pdb/),to a data file provided there, or to any file following the PDB format. Files in the PDB include information such as the name of the compound, the species and tissue from which is was obtained, authorship, revision history, journal citation, references, amino acid sequence, stoichiometry, secondary structure locations, crystal lattice and symmetry group, and finally the ATOM and HETATM records containing the coordinates of the protein and any waters, ions, or other heterogeneous atoms in the crystal. Some PDB files include multiple sets of coordinates for some or all atoms. Due to the limits of x-ray crystallography and NMR structure analysis, the coordinates of hydrogen atoms are not included in the PDB. NAMD and VMD ignore everything in a PDB file except for the ATOM and HETATM records, and when writing PDB files the ATOM record type is used for all atoms in the system, including solvent and ions.

B) PSF Files

A PSF file, also called a protein structure file, contains all of the molecule specific information needed to apply a particular force field to a molecular system. The CHARMM force field is divided into a topology file, which is needed to generate the PSF file, and a parameter file, which supplies specific numerical values for the generic CHARMM potential function. The topology file defines the atom types used in the force field; the atom names, types, bonds, and partial charges of each residue type; and any patches necessary to link or otherwise mutate these basic residues. The parameter file provides a mapping between bonded and nonbonded interactions involving the various combinations of atom types found in the topology file and specific spring constants and similar parameters for all of the bond, angle, dihedral, improper, and van der Waals terms in the CHARMM potential function. The PSF file contains six main sections of

interest: atoms, bonds, angles, dihedrals, impropers (dihedral force terms used to maintain planarity), and cross-terms.

C) Topology Files

A CHARMM forcefield topology file contains all of the information needed to convert a list of residue names into a complete PSF structure file. It also contains internal coordinates that allow the automatic assignment of coordinates to hydrogens and other atoms missing from a crystal PDB file. The current versions of the CHARMM forcefield are CHARMM22 for proteins and CHARMM27 for lipids and nucleic acids including CMAP correction to proteins. The individual topology files are named, respectively, top_all22_prot_cmap.inp, top_all27_lipid.rtf, and top_all27_na.rtf. To enable computation on hybrid systems, combinations are also provided, named top_all27_na_lipid.rtf, top_all27_prot_lipid.rtf, and top_all27_prot_na.rtf which can all be found in the CHARMM31 release. While the tools used with NAMD allow multiple topology and parameter files to be used simultaneously, it is preferable to use these pre-combined files.

D) Parameter Files

A CHARMM forcefield parameter file contains all of the numerical constants needed to evaluate forces and energies, given a PSF structure file and atomic coordinates. The parameter file is closely tied to the topology file that was used to generate the PSF file, and the two are typically distributed together and given matching names. The current versions of the CHARMM forcefield are CHARMM22 for proteins and CHARMM27 for lipids and nucleic acids including CMAP correction to proteins. The individual parameter files are named, respectively, par_all22_prot_cmap.inp, par_all27_lipid.prm, and par_all27_na.prm. To enable hybrid systems, combinations are also provided, named par_all27_na_lipid.prm, par_all27_prot_lipid.prm, and par_all27_prot_na.prm which can all be found in the CHARMM31 release. While the tools used with NAMD allow multiple topology and parameter files to be used simultaneously, it is preferable to use these pre-combined files.

E) NAMD Configuration Files

The NAMD configuration file (also called a config file, .conf file, or .namd file) is given to NAMD on the command line and specifies virtually everything about the simulation to be done. The only exceptions are details relating to the parallel execution environment, which vary between platforms. Therefore, the config file should be portable between machines, platforms, or numbers of processors in a run, as long as the referenced input files are available. As a convenience, on startup NAMD will switch to the directory that contains the config file, so that all file paths are relative to that directory. NAMD also accepts multiple config files on the command line, but changes directories before parsing each file, so it is best to keep everything in the same directory when using multiple config files.