Role of ELA region in auto-activation of mutant KIT receptor: a molecular dynamics simulation insight

Rituraj Purohit*

Human Genetics Foundation, via Nizza 52, Torino I-10126, Italy; Bioinformatics Division, School of Bio Sciences and Technology, Vellore Institute of Technology University, Vellore, Tamil Nadu 632014, India

Communicated by Ramaswamy H. Sarma

(Received 3 October 2012; final version received 5 May 2013)

KIT receptor is the prime target in gastrointestinal stromal tumor (GISTs) therapy. Second generation inhibitor, Sunitinib, binds to an inactivated conformation of KIT receptor and stabilizes it in order to prevent tumor formation. Here, we investigated the dynamic behavior of wild type and mutant D816H KIT receptor, and emphasized the extended A-loop (EAL) region (805–850) by conducting molecular dynamics simulation (~100 ns). We analyzed different properties such as root mean square cutoff or deviation, root mean square fluctuation, radius of gyration, solvent-accessible surface area, hydrogen bonding network analysis, and essential dynamics. Apart from this, clustering and cross-correlation matrix approach was used to explore the conformational space of the wild type and mutant EAL region of KIT receptor. Molecular dynamics analysis indicated that mutation (D816H) was able to alter intramolecular hydrogen bonding pattern and affected the structural flexibility of EAL region. Moreover, flexible secondary elements, specially, coil and turns were dominated in EAL region of mutant KIT receptor during simulation. This phenomenon increased the movement of EAL region which in turn helped in shifting the equilibrium towards the active kinase conformation. Our atomic investigation of mutant KIT receptor which emphasized on EAL region provided a better insight into the understanding of Sunitinib resistance mechanism of KIT receptor and would help to discover new therapeutics for KIT-based resistant tumor cells in GIST therapy.

Keywords: A-loop region; KIT receptor; GIST; molecular dynamics simulation

Introduction

Gastrointestinal stromal tumors (GISTs) are supposed to arise from the interstitial cells of Cajal (ICC) or their precursors located throughout the muscular wall of gastrointestinal tract. Studies across the world show remarkably consistent annual incidences of 11–19.6 per million population (Chan et al., 2006; Corless, Fletcher, & Heinrich, 2004; Miettinen & Lasota, 2006a, 2006b; Nilsson et al., 2005; Tryggvason, Gislason, Magnusson, & Jonasson, 2005). Under normal physiological conditions, the stem cell factor (SCF) binds to receptor tyrosine kinase KIT and induces KIT homodimerization resulting in the phosphorylation of its tyrosine residues. The tyrosine-phosphorylated KIT receptor subsequently becomes a new docking site for signal transduction molecules and induces substrate binding and phosphorylation (Blechman, Lev, Givol, & Yarden, 1993). Thus, the interaction between SCF and the KIT receptor is essential for normal hematopoiesis, melanogenesis, gametogenesis, and the growth and differentiation of mast cells and ICCs (Fleischman, 1993; Huizinga et al., 1995). Oncogenic mutations in c-kit cause a constitutive phosphorylation of the KIT receptor which is independent of SCF binding, leading to a cascade of intracellular signalling events that contribute to the abnormal proliferation and survival of these neoplastic cells (Boisson, Feger, Guillosson, & Arock, 2000; Testa, 2008).

GISTs appear to arise from ICC or their precursors and knockout mice model of KIT failed to express ICC cells leading to the hypothesis that KIT is essential for the development of ICC cells (Chan et al., 2006; Nilsson et al., 2005). KIT has been shown to be the key oncogenic driver in GIST that is essential for growth and survival (Mol et al., 2003).

KIT mutations found in GIST cause constitutional activation of the receptor tyrosine kinase pathway in the absence of their ligands. Four different regions of KIT, namely, exon 9, exon 11, exon 13, and exon 17 are most often mutated in sporadic GISTs. The activation of KIT receptors occurs through the binding of extracellular...
ligand (e.g. SCF) that cause receptor dimerization. Alternatively, oncogenic mutations in KIT receptors can cause ligand-independent receptor activation. Mutations in the activation loop (A-loop) stabilize the active conformation of the kinase (Mol et al., 2003) but how, it is still unknown.

KIT is a ATP-binding protein and catalyzes the reaction of the transfer of γ-phosphate to specific tyrosine sites. KIT inhibitor targets the ATP-binding pocket and is subject to competition from high intracellular ATP concentrations. Out of the17 GIST therapeutics, only two KIT Inhibitor, Imatinib and Sunitinib have FDA approval (Corless, Barnett, & Heinrich, 2011). Imatinib and Sunitinib bind to amino acid residues within the ATP-binding pocket in order to inhibit the KIT kinase activity and prevent the cancerous condition. Multiple mutations in KIT have been described and specific mutations correlate with therapeutic response to Imatinib and Sunitinib (Liegler et al., 2008). In our study, we analyze a most deleterious mutant (D816H) which has been found to be resistant toward Imatinib as well as Sunitinib molecules and had been found at kinase A-loop at exon 17 (Corless et al., 2011; Gajiwala et al., 2009; Lasota et al., 2008; Liegl et al., 2008). The position D816 is important in KIT receptor and other mutation, such as D816V, found to affect the normal enzyme allosteric regulation (Laine, Auclair, & Tchertanov, 2012). Imatinib and Sunitinib are preferentially bound to the inactivate form of KIT. It was found that mutation D816H is responsible for auto-activation and favors an activated enzyme conformation that is resistant to Sunitinib (DiNitto et al., 2010; Gajiwala et al., 2009). Even though mutant D816H binds Sunitinib in an inactivated conformation and how it shifts the drug-favorable inactivated kinase conformation to the drug-insensitive active form in the presence of physiological ATP concentrations that results in loss of inhibition is still unknown. Structural studies (Gajiwala et al., 2009; Mol et al., 2003, 2004) show that KIT receptor can sample diverse conformations like the inactivated, auto-inhibited, and activated and these states of the conformation can be defined by specific orientation. The protein can be considered to be in equilibrium in these conformations with a shift to the activated form upon phosphorylation. Previous crystallography studies also indicated that Sunitinib binds to the ATP-binding pocket of kinases and inhibits the active form of the enzyme in an ATP-competitive manner (Gajiwala et al., 2009; Mol et al., 2003, 2004). Since the apo- and auto-inhibited form retain similar conformation in 3D space (Gajiwala et al., 2009), we use wild type and mutant KIT structures which are in auto-inhibited conformation. Studies highlight that the drug resistance exhibited by D816H protein is the result of a shift in equilibrium towards the active kinase conformation and an accelerated autophosphorylation of this mutant (Gajiwala et al., 2009). Previous studies also highlight that the conformation of binding pocket remains unchanged in mutant enzyme and major conformation alterations are observed in A-loop region (Gajiwala et al., 2009; Mol et al., 2003, 2004).

Biological function of a protein molecule depends upon its conformational flexibility at different levels (Kamaraj, Rajendran, Sethumadhavan, & Purohit, 2013; Kumar, Rajendran, Sethumadhavan, & Purohit, 2013; Purohit, Rajasekaran, Sudandiradoss, & Doss, 2008; Purohit, Rajendran, & Sethumadhavan, 2011a, 2011b; Purohit & Sethumadhavan, 2009; Rajendran, Purohit, & Sethumadhavan, 2012, Rajendran & Sethumadhavan, 2013). There are three objectives of our investigation. First, to investigate the complete conformational profile of wild type and mutant KIT protein in order to understand the drug resistance mechanism, second, to capture a detail picture of atomic behavior of extended A-loop (EAL) region of the KIT protein and its secondary structure arrangement during simulation time periods, and third, to understand that how KIT A-loop oncogenic mutation (D816H) alters the conformational equilibrium of the kinase toward the active form. It is apparent that our knowledge of the biological basis of GIST is not only driving our treatment of the disease, but also led to increase the survival for our patients.

In this study, we found that mutant KIT protein attains different conformation states, particularly, at EAL region when compared with wild type KIT. Secondary structure analysis, hydrogen bonds, and clustering analysis clearly show that the mutation D816H at KIT protein works as a gain-of-function mutation and helps in auto-activation. Molecular dynamics simulation (MDS) indicates more atomic flexibility in mutant KIT than in wild type.

Materials and methods

Datasets
We selected wild type KIT receptor (PDB ID, 3G0E) (Gajiwala et al., 2009) and one mutant (D816H) structures (PDB ID, 3G0F) (Gajiwala et al., 2009) from Brookhaven Protein Data Bank (Berman et al., 2002) for our investigation. Wild type and mutant structures were solved with 1.60 and 2.60 Å resolutions, respectively. The Structure of wildtype KIT receptor is indicated in Figure 1. Sunitinib (DrugBank ID, DB01268), a drug approved by FDA for GISTs was retrieved from DrugBank database as SMILES string and was submitted to CORINA (www.molecular-networks.com/onlinedemos/corinademo.html) for constructing the 3D structure of small molecule. The 2D structure of Sunitinib is shown in Figure 2.

Receptor–ligand interaction analysis
Receptor–ligand interaction studies were performed between wild type and mutant KIT and Sunitinib. In order to carry out the docking simulation, we used the
AutoDock 4.0 suite as molecular-docking tool (Morris et al., 2009). In this docking simulation, we used semi-flexible docking protocols in which the binding residues of the target protein were kept as flexible while the others were kept rigid. We removed crystallized water molecules and heteroatoms from KIT receptor PDB files. The Graphical User Interface program “AutoDock Tools” was used to prepare, run, and analyze the docking simulations. Kollman united atom charges, solvation parameters and polar hydrogens were added into the receptor PDB file for the preparation of protein in docking simulation. The geometry of KIT protein and Sunitinib structures were optimized via Steepest Descent method followed by Conjugate Gradient method with 1000 steps each of GROMACS 4.0.5 package (Hess, Kutzner, van der Spoel, & Lindahl, 2008; van der Spoel et al., 2005). Each minimization was carried out with GROMOS-96 (Baran & Mazerski, 2002) 43a1 force field. Bond lengths were constrained with LINCS (Hess, Bekker, Berendsen, & Fraaije, 1997) and SETTLE (Miyamoto & Kollman, 1992) for water. A time step of integration of 2 fs was used within the leap frog integration scheme. Wild type and mutant systems were solvated in, respectively, 20,782 and 21,877 explicit SPC water molecules in periodic rectangular simulation boxes. The box dimensions were chosen to provide at least a 10 Å buffer of water molecules around the solute. To neutralize and prepare the systems under a physiological ionic concentration (.15 M), five or four chloride ions were added by replacing water molecules. Restrained MDS and unrestrained production MDS setup were based on our previous studies (Purohit et al., 2011b; Rajendran et al., 2012; Rajendran & Sethumadhavan, 2013).

**Molecular dynamics simulations**

MDSs were performed using the GROMACS 4.0.5 (Hess et al., 2008; van der Spoel et al., 2005) and GROMOMS-96 (Baran & Mazerski, 2002) 43a1 force field. Bond lengths were constrained with LINCS (Hess, Bekker, Berendsen, & Fraaije, 1997) and SETTLE (Miyamoto & Kollman, 1992) for water. A time step of integration of 2 fs was used within the leap frog integration scheme. Wild type and mutant systems were solvated in, respectively, 20,782 and 21,877 explicit SPC water molecules in periodic rectangular simulation boxes. The box dimensions were chosen to provide at least a 10 Å buffer of water molecules around the solute. To neutralize and prepare the systems under a physiological ionic concentration (.15 M), five or four chloride ions were added by replacing water molecules. Restrained MDS and unrestrained production MDS setup were based on our previous studies (Purohit et al., 2011b; Rajendran et al., 2012; Rajendran & Sethumadhavan, 2013).

**Essential dynamics (ED)**

ED is a widely applied technique based on the principal component analysis of conformational ensembles that allows the identification of the most relevant or correlated motions of groups of residues of a protein along a trajectory generated by MDS (Amadei, Linssen, & Berendsen, 1993; Stepanova, 2007). The KIT protein coordinates were extracted at 16 ps intervals after 18 ns until end of each molecular dynamics (MD) trajectory for the simulation systems. The covariance matrix of C-α atoms was constructed and then diagonalized by using g_covar and g_anal with within the GROMACS package (Hess et al., 2008; van der Spoel et al., 2005) yielding a set of eigenvectors and their respective eigenvalues. Each eigenvector represents one single direction of collective motion, whereas the corresponding eigenvalue represents the amplitude of motion along that vector. The deviations
of the KIT protein C-α atoms from their time-averaged positions were obtained as averages over the eigenvectors representing 80% of the total mobility of the system (i.e. the total sum of eigenvalues for the system).

**Root mean square cutoff or deviation (RMSD) clustering**

In order to select a reduced set of representative models of the wild type and mutant EAL region of KIT receptor, RMSD conformational clustering was performed using the gromos method (Daura et al., 1999) implemented in GROMACS (g_cluster). In the gromos clustering algorithm, the conformation with the highest number of neighbors, identified within the chosen RMSD cutoff, is chosen as the center of the first cluster. All the neighbors of this conformation are removed from the ensemble of conformations. The center of the second cluster is then determined in the same way, and the procedure is repeated until each structure has been assigned to a cluster (Daura et al., 1999).

**Cross-correlation analysis**

The dynamic cross-correlation matrix (DCCM) $C_{ij}$ that reflects the fluctuations of the coordinates of Cα atoms was calculated to analyze the collective motions of wild type and mutant KIT. The cross-correlation coefficient $C_{ij}$, between atoms $i$ and $j$, is a measure of the correlated nature of their atomic fluctuations and is computed as follows: $C_{ij} = \langle \Delta r_i \times \Delta r_j \rangle / (\langle \Delta r_i^2 \rangle \times \langle \Delta r_j^2 \rangle)^{1/2}$ where $\Delta r_i$ and $\Delta r_j$ correspond to the atomic displacement vectors for atoms $i$ and $j$, respectively, and the angle brackets indicate time averages (Laberge & Yonetani, 2008). The elements $C(i,j)$ can be collected in matrix form and displayed as a 3D dynamical cross-correlation map (Swaminathan, Harte, & Beveridge, 1991).

The trajectory files were analyzed through the use of g_rms, g_gyrate, g_sas, and g_rmsf of GROMACS utilities in order to obtain the RMSD, radius of gyration (Rg), solvent-accessible surface area (SASA), and root mean square fluctuation (RMSF) values. The number of distinct hydrogen bonds formed by specific residues to other amino acids within the protein during the simulation (NH bond) was calculated using g_h bond. NH bond determined on the basis of donor–acceptor distance smaller than .35 nm and of donor–hydrogen-acceptor angle larger than 150°. Moreover, VMD (Humphrey, Dalke, & Schulten, 1996) and Coot (Emsley & Cowtan, 2004) packages were used for trajectory analysis and for the management of the simulation snap shot structures.

To that end, we compare RMSF of carbon-α, RMSD of backbone structure, Rg, and SASA of protein between the trajectories generated at 300 K to investigate the flexible nature of mutants. In order to generate the 3D backbone, RMSD, Rg, SASA, and RMSF of carbon-α and motion projection of the protein in phase space of the system were plotted for all three simulations using xmgrace program.

**Results and discussion**

Calculation of interaction energy is very important to understand the biological activity of most of the ligands. Overall, interaction energy of the complex mainly contributes to van der Waals and electrostatic interaction energy between KIT protein (wildtype and mutant) and Sunitinib. In the wild type and mutant complex, there were a similar contributions of van der Waals and electrostatic energy (Table 1). The total ligand receptor interaction energy of wild type KIT protein complex was $-8.25 \text{ kcal/mol}$ and that of mutant KIT protein was $-7.87 \text{ kcal/mol}$. It proved that the actual position of binding residues (Val603, Cys673, Glu671, and Cys809) were intact even after mutation and it was the reason of comparable interaction energy between wild type and mutant KIT with Sunitinib. Wild type complex showed two hydrogen bonds and 13 hydrophobic interactions with Sunitinib and, in a similar fashion, mutant KIT also formed two hydrogen bonds and 10 hydrophobic interactions with Sunitinib (Figure 3). Sunitinib binding pose also indicated similar kind of bonding behavior in wild type and mutant KIT receptor, it is depicted in Figure 3. We also superimposed the wild type and mutant complex to analyze the orientation of binding residues in surrounding inhibitor, it is depicted in Figure 4. This superimposed structure showed RMSD value of .35 Å. This structure also indicated the intact position of binding residues of mutant enzyme with the wild type (Figure 4). This analysis clearly showed that there is no conformational alteration of binding residues due to mutation (D816H) at EAL region. Our docking studies were also well supported by previous wet-lab experiments (Gajiwala et al., 2009; Mol et al., 2003, 2004) and

| KIT-ligand complex | Electrostatic energy (kcal/mol) | Van Der walls energy (kcal/mol) | Total interaction energy (kcal/mol) | Number of H-bonds | Number of hydrophobic contacts |
|-------------------|---------------------------------|---------------------------------|-----------------------------------|-------------------|-------------------------------|
| Native            | +.04                            | −5.89                           | −8.25                             | 2                 | 13                            |
| Mutant            | +.05                            | −5.52                           | −7.87                             | 2                 | 10                            |
it is proved that mutation does not affect its binding with Sunitinib. Moreover, Sunitinib docking modes were also found to be consistent with different conformers which were generated during MD simulation of wild type and mutant KIT and this is depicted in supporting information (S1). This observation provoked us to study about how the mutation D816H in KIT enzyme involved in auto-activation and produced resistance against Sunitinib without affecting its binding pocket.

Auto-activation mechanism of KIT mutant receptor is still unknown, even in mutant, the binding patterns of Sunitinib are same as that of the wild type structure but it showed considerable amount of resistance against inhibitors. Since inhibitor binding modes and docking energies were approximately similar in both wild type and mutant (Figure 4 and Table 1), we decided to simulate free enzyme. We explored the conformational behavior of KIT receptor by set up classical MD simula-
tions (100 ns). We explained how mutant is exhibits auto-inhibition conformation in order to become resistant towards its inhibitors. In this investigation, different dynamic properties of wild type and mutant KIT receptor were analyzed such as RMSD, RMSF, Rg, SASA, hydrogen bond network (H-bonds), secondary structure analysis, principle component, clustering, and DCCM analysis. In order to verify the system stabilized along the MD simulations, we have showed a plot of energy vs. time (Figure 5(a)).

We calculated the RMSD for the entire C-α atom from the initial structure, which was considered as a central criterion to measure the convergence of the protein system concerned (Figure 5(b)). In figure, wild type and mutant showed different starting points in RMSD trajectory at 0 ns and it was .16 nm in wild type and .23 nm in mutant graph. Until the time period of 2.8 ns, this difference in RMSD maintained and subsequently decreased from their starting structure. Between the periods of 2.8–18 ns, both wild type and mutant showed similar fashion of divination resulting in a backbone RMSD of ϵ.14–16 nm during the simulation. After the time periods of 18 ns, mutant structure showed sudden increment in RMSD and attains .31 nm at ϵ.45 ns, while the wild type structure exhibited less fluctuation and maintained its RMSD until the end of the simulation between .13 and .2 nm. From their starting structure, the wild type structure deviated minimum while the mutant exhibited maximum deviation. This magnitude of fluctuations together with a very small difference in the average RMSD values after the relaxation period led to the conclusion that the simulations produced stable trajectories, thus providing a suitable basis for further analysis.

We performed surface and geometry analysis of wild type and mutant KIT receptor by analyzing Rg and SASA graphs. We found that wild type and mutant showed similar fashion of fall in Rg values from starting structure to early time periods (until 20 ns). After 20 ns, mutant showed sudden increment in Rg compared with wild type and it was varied until 41 ns, while wild type structure maintained their Rg values until the end of the simulation (Figure 5(c)). SASA analysis showed a similar kind of fluctuation except between time periods of 20–40 ns (Figure 5(d)). Between these periods, mutant showed higher values of SASA as compared with wild type structure and it is well correlated with Rg values between the said periods. It indicated that there was a significant conformational transition occurred in mutant structure compared with wild type.

ED analysis were performed over the stable MD trajectories in order to identify possible relevant displacements of groups of residues and given emphasis to the direction and amplitude of the most dominant KIT motions. Moreover, it was organized to obtain a better view of dynamic mechanical properties of the investigated system. To validate the variation in structural flexibility behaviour and further support our MD simulation result, the large-scale collective motions of the wild type and mutant protein using ED analysis were determined via characterization of its phase space behavior. The eigenvectors of the covariance matrix were called its principle components. The change of particular trajectory along each eigenvector was obtained by this projection. The spectrum of the corresponding eigenvalues (Figure 6(a)) indicated that the fluctuation of the system was basically confined within the first two eigenvectors. The projection of trajectories obtained at 300 K onto the first two principal components (PC1, PC2) showed the motion of two proteins in phase space (Figure 6(a)). On these projections, we saw clusters of stable states. Two features were very apparent from these plots. Firstly, the clusters were well defined in wild type than in mutant. Secondly, mutant covered a larger region of phase space, particularly along PC2 plane than wild type and it is depicted in Figure 6(a). Our observation thus corroborated with the idea of higher flexibility of mutant than wild type at 300 K. The overall flexibility of two proteins had been calculated by the trace of the diagonalized covariance matrix of the Ca atomic positional fluctuations. We have obtained the following values for wild type protein 98.62 nm² and mutant protein 166.91 nm², thereby confirming the overall increased flexibility of mutant than wild type at 300 K.

In Figure 6(b), RMSF graph of wild type and mutant KIT protein showed more flexibility in mutant structure compared with wild type since crystal information was missing from the residues 695 to 751 in wild type structure and the residues 690–763 in mutant showed flat curve with respect to time (Figure 6(b)). Both terminals of mutant protein showed extra flexibility when compared with wild type protein. It is also evident that there was a significant fluctuation between residual positions 805–852 in mutant structure. This region is the EAL region of KIT receptor and is located near to ATP/inhibitor binding site. The importance of EAL region is noted in several experiments (Gajiwala et al., 2009; Laine et al., 2012; Mol et al., 2003, 2004). We assume that this region could play an important role in auto-activation mechanism of KIT receptor. These observations provoked us to investigate the dynamic behavior of EAL region. In order to unravel the concept, we performed several analyses on simulation trajectories with special emphasis on EAL region. Extra flexibility of region between 868 and 894 was also noted in mutant enzyme when compared to wild type.

RMSF graph (Figure 7(a)) of EAL region showed that there was more flexibility in mutant than in wild type protein. In Figure 7(a), the sharp peaks indicated by positions 821, 827, and 843 in mutant showed RMSF values...
of .25, .34, and .25 nm, respectively, while wild type showed .12, .17, and .09 nm, respectively. Protein flexibility is directly proportional to intramolecular NH bonds between amino acid residues (Purohit et al., 2011a; Rajendran et al., 2012). To understand the reason of extra flexibility of EAL region of mutant, we plotted a graph between the number of hydrogen bonds over time. This graph (Figure 7(b)) showed that mutant EAL residues have fewer tendencies to form hydrogen bonds with neighboring residues and showed more flexibility when compared with wild type EAL. The average number of hydrogen bonds in mutant was found to be 46 and that of
The compact conformation of wild type ELA region was well supported by less SASA values during simulation, and mutant ELA region showed more SASA values which supported its expanded conformation. These geometrical observations validated the extra flexibility of mutant ELA region when compared with wild type. Our analysis also endorses the wet-lab experiment performed by Zhang and co-workers (Zhang et al., 2010) who observed the extra SASA values in ELA region of mutant KIT as compared to wild type by solution-phase hydrogen/deuterium exchange experiments and dry-lab experiments performed by Laine, Chauvot de Beauchêne, Perahia, Auclair, and Tchertanov (2011). Moreover, our intramolecular NH bond analysis provided a reason for alteration in flexibility, Rg, and SASA values.

The secondary structure transitions of 100 ns MD simulations were monitored for EAL region by using DSSP (Figure 8), a program that uses hydrogen bond and geometric pattern recognition for secondary structure assignment (Kabsch & Sander, 1983). This graph provided information on the structural behavior of EAL region of KIT proteins obtained by the analysis of time-dependent secondary structure fluctuations. Figure 8 shows the secondary structural elements as a function of simulation time and it reveals that b-sheets, coil, bends, turns, and helix were observed in both wild type and mutant protein during simulation time period. First few residues positioned between 805 and 811 (indicated as 0–6 in Figure 8) show coil conformations in both wild type and mutant EAL. In wild type structure, position between 811 and 817 (indicated as 6–12 in Figure 8) appear as turns and bend conformations with few traces of coil conformation, while in mutant, apart from bend conformation, coil conformations are dominated over other conformations. In mutant, coil conformation remains dominated between positions of 817–823 (indicated as 12–18 in Figure 8) which were not visible in wild type structure. As compared with wild type, mutant show more structural transition between positions 837 and 847 (indicated as 32–42 in Figure 8). In this region,
after 40 ns, helix conformation was totally lost and turn conformations were dominated in mutant structure when compared with wild type structure. Between the positions 847 and 852 (indicated as 42–47 in Figure 8), mutant had not showed helix conformation, while this conformation was dominated in wild type EAL region. As we know, coil and turns conformations are more flexible elements than other elements in protein structure. These elements were found dominated in mutant when compared with wild type EAL region and this was the

Figure 7. Analysis of simulation trajectory of ELA region of wildtype and mutant KIT. (a) RMSF of the backbone CAs of C-\(\alpha\) atoms of EAL region of wildtype and mutant KIT structures vs. time. (b) Average number of protein–protein intermolecular hydrogen bonds in EAL region of wildtype and mutant KIT structures vs. time. (c) Rg of C-\(\alpha\) atoms vs. time. (d) SASA vs. time. The symbol coding scheme is as follows: wildtype (black) and mutant (red).
reason for more structural flexibility of mutant EAL region. Moreover, secondary structural studies performed by Laine et al. (2011) showed that there was less variation in secondary structure elements at EAL region of wild type and mutant (D816 V) KIT protein. Our secondary structure transition studies showed similar behavior at EAL region. Since the studies performed by Laine et al. (2011) have limited simulation time scale (50 ns), it was unable to explore the complete behavior of protein after 50 ns. In our 100 ns MSD, we observed that after the period of 50 ns, the mutant (D816H) structure significantly lost its β-sheets and was not able to regain it until the end of the simulation, and turn conformations were dominated during this time. This is also a main reason why mutant structure has more flexibility than wild type.

Clustering is one way of examining the data from an extensive sampling of a high-dimensional phase space, such as that obtained from MD calculation. To better describe the conformational flexibility of EAL region of the wild type and mutant KIT receptors, we clustered the sampled configurations in each trajectory on the basis of the method by Daura et al. (1999). We ranked each trajectory frame based on the number of neighboring frames that were within 1 Å RMSD. The top-ranked frame, together with all its neighbors, was then removed and the ranking process was repeated. Here, we have divided simulation trajectories into two slabs, first slab was considered between 0 and 50 ns and the second slab between 50 and 100 ns. The representative structures for a member of one of the most populated clusters of wild type and mutant EAL region were superimposed and shown in Figure 9. This analysis clearly indicated that except the residues between positions 820 and 830, other residues of the members of the wild type cluster were deviated in a similar fashion like members of mutant cluster but in the second-half of simulation, members of the wild type structure contained a rigid conformation and exhibited less structural deviation, while members of mutant structures acquired flexible conformation and showed more deviation to each other. EAL region of mutant showed loss of beta-sheets and loop structure when compared with wild type EAL region (Figure 9) and this loss in rigid secondary elements help in increasing flexibility behavior which was indicated in Figure 7(a). Moreover, during the first-half of simulation, most of the secondary structures (beta-sheet and loop) disappeared in mutant when compared with wild type and remained absent in the second-half of the simulation (Figure 8). Apart from RMSD graph, the equilibration is provided by a graph between number of cluster and simulation time and it is depicted in supporting information (S2). The premise is that convergence is achieved when the number of clusters no longer increases, as this means that the simulation has visited every substate. This analysis suggests that convergence is observed by ~70 ns and the curve has a comforting appearance of saturation.

The statistics of the cluster analysis were tabulated in Table 2 which revealed some relevant features of the EAL region dynamics, namely its noticeable flexibility,
which was attested by the significant amount of clusters obtained (Table 2). Mutant EAL region showed more number of clusters and less members in populated clusters when compared with wild type EAL region. This observation also indicated the flexible behavior of mutant structure compared with wild type. To capture a better picture of conformation behavior, a representation of the central conformations representative of the average structure of each clusters of EAL region of wild type and mutant KIT receptor in MD simulation was calculated and indicated in Figure 10(a–d). The differences observed in cluster analysis in conformation of wild type and mutant structure were well supported by alteration indicated in a number of secondary structure elements of EAL region (Figure 8). Wild type EAL region showed more number of rigid structure elements than mutant (Figure 8(a and b)). As it was also depicted in Figure 11(a) that C-α RMSF of the EAL region (position 821, 827 and 842) was most flexible in mutant structure and rigid in wild type structure, it was also implied in representative of the average structure of each clusters of EAL region with each time slab (Figure 10(a–d)). We also plotted cluster index of wild type and mutant for each of the time slab and it is showed in supporting information (S2). Maximum RMSD values of wild type EAL region was found to be .316 and .202 nm, while mutant EAL region exhibited .452 and .39 nm for time slab I and slab II, respectively. This observation further indicated that mutant EAL region exhibited a scattered conformation in 3D space and acquired more degree of RMSD than the wild type structure.

To gain further insight into the conformational changes of EAL region, we investigated the correlation between the motions of residues of EAL region in wild type and mutant trajectories. In the cross-correlation matrix, the $C_{ij}$ elements of the matrix are symmetrical about the diagonal. As these maps are normalized, the magnitude of the correlation can be quantified by calculating the cross-correlation coefficient between the atomic displacements, strong correlation in the movements of specific residues is displayed by the highly positive regions (red); whereas strong anticorrelated motion of the residues is represented by negative regions (dark blue). The higher the absolute cross-correlation value, the better the two atoms are correlated (or anticorrelated). If $C_{ij}=0$, the atomic motions are not correlated, and their movements are random when compared with each other. As a whole, few highly correlated motions are noticed, except for the diagonal describing the correlation of a residue with itself. The correlated motion of the mutant KIT protein is stronger than the wild type.

In the case of the mutant (Figure 11(b)), the mutation (D816H) not only induces an intrasubunit-correlated motion (indicated as square), but also intersubunit-correlated motion between the other regions (off-diagonal, indicated as circle). Additionally, an anticorrelated motion occurs between the region of 670–825 and c-terminal region.

In comparison with the mutant (Figure 11(b)), intersubunit- and intrasubunit-correlated motions were found weaken in wild type KIT. In addition, intersubunit exhibited random motion (white colour) in wild type KIT than

Table 2. Results of clustering of the EAL region of KIT receptor structural ensemble obtained from MD trajectories. Time slab I is between 0 and 50,000 ns and time slab II is between 50,000 and 100,000.

| Protein | Total number of clusters | Total number of members in most populated cluster |
|---------|--------------------------|-----------------------------------------------|
|         | Time slab I | Time slab II | Time slab I | Time slab II |
| Native  | 66          | 61           | 6477        | 14,811       |
| Mutant  | 433         | 372          | 650         | 1156         |

Figure 9. Superposition of the average structure of most dominant cluster of EAL region of wildtype and mutant KIT receptor in MD simulation. (a) Time slab I. (b) Time slab II. Wildtype in green color and mutant in red color. Time slab I is between 0–50 ns and time slab II is between 50–100 ns.
mutant. This result also suggested that the conformational changes in 3D space of the mutant were higher than the wild type.

This analysis also showed that the mutant KIT exhibited a good amount of correlated motion between inhibitor binding residues (603, 673, 671, and 809) and the residues of EAL region (805–850) than wild type KIT. Moreover, the DCCM analysis agrees with the RMSF and superimposition analyses, where increased flexibility at ELA region due to mutation was indicated.

Our study clearly showed that the mutation (D816H) in KIT receptor and the mechanism of auto-activation

Figure 10. Representation of the central conformations representative of the average structure of each clusters of EAL region of wildtype and mutant KIT receptor in MD simulation. (a) Wildtype conformations in time slab I, (b) wildtype conformations in time slab II, (c) mutant conformations in time slab I and (d) mutant conformations in time slab II. Time slab I is between 0–50 ns and time slab II is between 50–100 ns.

Figure 11. Cross-correlation matrix of the fluctuations of coordinates for Cα atoms around their mean positions during the last 15 ns of MD simulation: the extent of correlated motion and anticorrelated motions are color-coded, (a) wildtype KIT, (b) mutant KIT.
are complementary to each other. Extra flexibility of EAL region is responsible for auto-activation in mutant KIT receptor and it provided a new functionality of resistance toward Sunitinib without altering its binding in receptor cavity.

Conclusion

KIT is an essential target in GIST and to overcome drug resistance of this target is a major challenge in oncology. Reconnaissance MD provide an extensive exploration of the conformational space of wild type and mutant KIT proteins. In RMSD analysis, mutant structure showed more deviation than wild type. Wild type and mutant KIT structure exhibited comparable Rg and SASA values with respect to time. Apart from the terminals, EAL region of mutant structure was found to be more flexible than wild type in RMSF analysis. Dynamics behavior of EAL region was studied by organizing cluster and DCCM analysis. Secondary structure analysis of EAL region suggested that mutant EAL region acquired more flexible elements with time during simulation, while wild type EAL region maintained its rigid elements from starting to until the end of the simulation. Moreover, EAL region of mutant showed more number of clusters and scattered conformation in 3D space than wild type structure. Based on our analysis, we concluded that mutation (D816H) in KIT receptor was able to alter the atomic conformation especially at EAL region and the extra flexibility acquired by mutant EAL region helped in auto-activation of KIT receptor.

Supplementary material

The supplementary material for this paper is available online at http://dx.doi.org/10.1080/07391102.2013.803264.

Acknowledgments

I gratefully acknowledge the management of Vellore Institute of Technology University and Prof. Riccardo Zecchina from HuGeF-Torino for providing the facilities to carry out this work. I thank the anonymous reviewers for their helpful comments and critical reading of the manuscript.

References

Amadei, A., Linssen, A. B. M., & Berendsen, H. J. C. (1993). Essential dynamics of proteins. *Proteins: Structure, Function, and Genetics*, 17, 412–425.

Baran, M., & Mazerski, J. (2002). Molecular modelling of membrane sterols with the use of the GROMOS 96 force field. *Chemistry and Physics of Lipids*, 120, 21–31.

Berman, H. M., Battistuz, T., Bhat, T. N., Bluhm, W. F., Bourne, P. E., Burkhardt, K., ... Zardecki, C. (2002). The protein data bank. *Acta Crystallographica. Section D: Biological Crystallography*, 58, 899–907.

Blechman, J. M., Lev, S., Givol, D., & Yarden, Y. (1993). Structure-function analyses of the kit receptor for the steel factor. *Stem Cells*, 11, 12–21.

Boissan, M., Feger, F., Guillosson, J. J., & Arock, M. (2000). c-Kit and c-kit mutations in mastocytosis and other hematological diseases. *Journal of Leukocyte Biology*, 67, 135–148.

Chan, K. H., Chan, C. W., Chow, W. H., Kwan, W. K., Kong, C. K., Mak, K. F., ... Lau, L. K. (2006). Gastrointestinal stromal tumours in a cohort of Chinese patients in Hong Kong. *World Journal of Gastroenterology*, 12, 2223–2228.

Corless, C. L., Barnett, C. M., & Heinrich, M. C. (2011). Gastrointestinal stromal tumours: Origin and molecular oncology. *Nature Reviews Cancer*, 11, 865–878.

Corless, C. L., Fletcher, J. A., & Heinrich, M. C. (2004). Biology of gastrointestinal stromal tumours. *Journal of Clinical Oncology*, 22, 3813–3825.

Daura, X., Gademen, K., Jaun, B., Seebach, D., van Gunsteren, W. F., & Mark, A. E. (1999). Peptide folding: When simulation meets experiment. *Angewandte Chemie International Edition in English*, 38, 236–240.

DiNitto, J. P., Deshmukh, G. D., Zhang, Y., Jacques, S. L., Coli, R., Worrall, J. W., ... Wu, J. C. (2010). Function of activation loop tyrosine phosphorylation in the mechanism of c-Kit auto-activation and its implication in sunitinib resistance. *Journal of Biochemistry*, 147, 601–609.

Emsley, P., & Cowtan, K. (2004). Coot: Model-building tools for molecular graphics. *Acta Crystallographica. Section D: Biological Crystallography*, 60, 2126–2132.

Fleischman, R. A. (1994). From white spots to stem cells: The role of the Kit receptor in mammalian development. *Trends in Genetics*, 9, 285–290.

Gajiwala, K. S., Wu, J. C., Christensen, J., Deshmukh, G. D., Diehl, W., DiNitto, J. P., ... Demetri, G. D. (2009). KIT kinase mutants show unique mechanisms of drug resistance to imatinib and sunitinib in gastrointestinal stromal tumor patients. *Proceedings of the National Academy of Sciences of the USA*, 106, 1542–1547.

Hess, B., Bekker, H., Berendsen, H. J. C., & Fraaije, J. G. M. (1997). LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, 18, 1463–1472.

Hess, B., Kutzner, C., van der Spoel, D., & Lindahl, E. (2008). GROMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation. *Journal of Chemical Theory and Computation*, 4, 435–447.

Huizinga, J. D., Thuneberg, L., Klüppel, M., Malysz, J., Mikkelsen, H. B., & Bernstein, A. (1995). W/kit gene required for interstitial cells of Cajal and for intestinal pacemaker activity. *Nature*, 373, 347–349.

Humphrey, W., Dalke, A., & Schulten, K. (1996). VMD: Visual molecular dynamics. *Journal of Molecular Graphics*, 14, 33–38.

Kabsch, W., & Sander, C. (1983). Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, 22, 2577–2637.

Kamaraj, B., Rajendran, V., Sethumadhavan, R., & Purohit, R. (2013). Investigation of binding phenomenon of NSP3 and p130Cas mutants and their effect on cell signalling. *Cell Biochemistry and Biophysics*. doi:10.1007/s12013-013-9551-6

Kumar, A., Rajendran, V., Sethumadhavan, R., & Purohit, R. (2013). Relationship between a point mutation S97C in CK16 protein and its affect on ATP binding affinity. *Journal of Biomolecular Structure & Dynamics*. doi:10.1080/07391102.2013.770373
Mol, C. D., Lim, K. B., Sridhar, V., Zou, H., Chien, E. Y., Morris, G. M., Huey, R., Lindstrom, W., Sanner, M. F., Belew, R. K., Goodsell, D. S., & Olson, A. J. (2009). Autodock4: Eektor-linked perturbation of tertiary conformations and HbA concerted dynamics. Biophysics Journal, 94, 2737–2751.

Laine, E., Auclair, C., & Tchertanov, L. (2012). Allostoric communication across the native and mutated KIT receptor tyrosine kinase. PLoS Computational Biology, 8, e1002661.

Laine, E., Chauvot de Beauchêne, I., Perahia, D., Auclair, C., & Tchertanov, L. (2011). Mutation D816V alters the internal structure and dynamics of c-KIT receptor cytoplasmic region: Implications for dimerization and activation mechanisms. PLoS Computational Biology, 7, e1002068.

Lasota, J., Corless, C. L., Heinrich, M. C., Debic, Rycher, M., Sciot, R., Wardelmann, E., … Miettinen, M. (2008a). Clinicopathologic profile of gastrointestinal stromal tumors (GISTs) with primary KIT exon 13 or exon 17 mutations: A multicenter study on 54 cases. Modern Pathology, 21, 476–484.

Ligl, B., Kepten, I., Le, C., Zhu, M., Demetri, G. D., Heinrich, M. C., … Fletcher, J. A. (2008). Heterogeneity of kinase inhibitor resistance mechanisms in GIST. Journal of Pathology, 216, 64–74.

Miettinen, M., & Lasota, J. (2006a). Gastrointestinal stromal tumors: Review on morphology, molecular pathology, prognosis, and differential diagnosis. Archives of Pathology and Laboratory Medicine, 130, 1466–1478.

Miettinen, M., & Lasota, J. (2006b). Gastrointestinal stromal tumors: Pathology and prognosis at different sites. Seminars in Diagnostic Pathology, 23, 70–83.

Miyamoto, S., & Kollman, P. A. (1992). SETTLE: An analytical version of the shake and rattle. Journal of Computational Chemistry, 13, 952–962.

Mol, C. D., Dougan, D. R., Schneider, T. R., Skene, R. J., Kraus, M. L., Scheibe, D. N., … Wilson, K. P. (2004). Structural basis for the autoinhibition and STI-571 inhibition of c-Kit tyrosine kinase. Journal of Biological Chemistry, 279, 31655–31663.

Mol, C. D., Lim, K. B., Sridhar, V., Zou, H., Chien, E. Y., Sang, B.C., … McRee, D. E. (2003). Structure of a c-kit product complex reveals the basis for kinase transactivation. Journal of Biological Chemistry, 278, 31461–31464.

Morris, G. M., Huey, R., Lindstrom, W., Samner, M. F., Belov, R. K., Goodsell, D. S., & Olson, A. J. (2009). Autodock4 and AutoDockTools4: Automated docking with selective receptor flexibility. Journal of Computational Chemistry, 30, 2785–2791.

Nilsson, B., Bümminxg, P., Meis-Kindblom, J. M., Odén, A., Dörtok, A., Gustavsson, B., … Kindblom, L. G. (2005). Gastrointestinal stromal tumors: The incidence, prevalence, clinical course, and prognostication in the preimatinib mesylate era. Cancer, 103, 821–829.

Purohit, R., Rajasekaran, R., Sudandiradoss, C., George Priya Doss, C., Ramanathan, K., & Sethumadhavan, R. (2008). Studies on flexibility and binding affinity of Asp25 of HIV-1 protease mutants. International Journal of Biological Macromolecules, 42, 386–391.

Purohit, R., Rajendran, V., & Sethumadhavan, R. (2011a). Relationship between mutation of serineresidue at 315th position in M. tuberculosis catalase-peroxidase enzyme and isoniazid susceptibility: An in silico analysis. Journal of Molecular Modeling, 17, 869–877.

Purohit, R., Rajendran, V., & Sethumadhavan, R. (2011b). Studies on adaptability of binding residues and flap region of TMC-114 resistance HIV-1 protease mutants. Journal of Biomolecular Structure & Dynamics, 29, 137–152.

Purohit, R., & Sethumadhavan, R. (2009). Structural basis for the resilience of Darunavir (TMC114) resistance major flap mutations of HIV-1 protease. Interdisciplinary Sciences, 1, 320–328.

Rajendran, V., Purohit, R., & Sethumadhavan, R. (2012). In silico investigation of molecular mechanism of laminopathy cause by a point mutation (R482W) in lamin A/C protein. Amino Acids, 43, 603–615.

Rajendran, V., & Sethumadhavan, R. (2013). Drug resistance mechanism of PnCA in Mycobacterium tuberculosis. Journal of Biomolecular Structure & Dynamics. doi:10.1080/07391102.2012.759885

Schüttekopf, A. W., & van Aalten, D. M. F. (2004). PRODRG – a tool for high-throughput crystallography of protein-ligand complexes. Acta Crystallographica, D60, 1355–1363.

Stepanova, M. (2007). Dynamics of essential collective motions in proteins: Theory, Physical Review E, 76, 051918.

Swaminathan, S., Harte, W. E. Jr, & Beveridge, D. L. (1991). Investigation of domain structure in proteins via molecular dynamics simulation: Application to HIV-1 protease dimer. Journal of the American Chemical Society, 113, 2717–2721.

Testa, U. (2008). Membrane tyrosine kinase receptors are an important target for the therapy of acute myeloid leukemia. Current Cancer Therapy Reviews, 4, 31–49.

Tryggvason, G., Gislason, H. G., Magnusson, M. K., & Jonasson, J. G. (2005). Gastrointestinal stromal tumors in Iceland, 1990–2003: The Icelandic GIST study, a population-based incidence and pathologic risk stratification study. International Journal of Cancer, 117, 289–293.

van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., & Berendsen, H. J. (2005). GROMACS: Fast, flexible, and free. Journal of Computational Chemistry, 26, 1701–1718.

Zhang, H. M., Yu, X., Greig, M. J., Gaijwala, K. S., Wu, J. C., Diehl, W., … Marshall, A. G. (2010). Drug binding and resistance mechanism of KIT tyrosine kinase revealed by hydrogen/deuterium exchange FTICR mass spectrometry. Protein Science, 19, 703–715.