Article

Structural and Dynamic Characterizations Highlight the Deleterious Role of SULT1A1 R213H Polymorphism in Substrate Binding

Raju Dash [1] , **Md. Chayan Ali** [2] , **Nayan Dash** [3] , **Md. Abul Kalam Azad** [2] , **S. M. Zahid Hosen** [4] , **Md. Abdul Hannan** [1,5] **and Il Soo Moon** [1,*]

[1]   Department of Anatomy, Dongguk University College of Medicine, Gyeongju 38066, Korea;
      rajudash.bgctub@gmail.com (R.D.); hannanbau@gmail.com (M.A.H.)
[2]   Department of Biotechnology and Genetic Engineering, Islamic University, Kushtia 7003, Bangladesh;
      chayanali7@gmail.com (M.C.A.); akazadbtge@hotmail.com (M.A.K.A.)
[3]   Department of Computer Science and Engineering, BGC Trust University, Bangladesh,
      Chittagong 4381, Bangladesh; dashnayan.py@gmail.com
[4]   Pancreatic Research Group, South Western Sydney Clinical School, University of New South Wales,
      and Ingham Institute for Applied Medical Research, Liverpool, NSW 2170, Australia;
      s.hosen@student.unsw.edu.au
[5]   Department of Biochemistry and Molecular Biology, Bangladesh Agricultural University,
      Mymensingh 2202, Bangladesh
[*]   Correspondence: moonis@dongguk.ac.kr

check for updates

Abstract: Sulfotransferase 1A1 (SULT1A1) is responsible for catalyzing various types of endogenous and exogenous compounds. Accumulating data indicates that the polymorphism rs9282861 (R213H) is responsible for inefficient enzymatic activity and associated with cancer progression. To characterize the detailed functional consequences of this mutation behind the loss-of-function of SULT1A1, the present study deployed molecular dynamics simulation to get insights into changes in the conformation and binding energy. The dynamics scenario of SULT1A1 in both wild and mutated types as well as with and without ligand showed that R213H induced local conformational changes, especially in the substrate-binding loop rather than impairing overall stability of the protein structure. The higher conformational changes were observed in the loop3 (residues, 235–263), turning loop conformation to A-helix and B-bridge, which ultimately disrupted the plasticity of the active site. This alteration reduced the binding site volume and hydrophobicity to decrease the binding affinity of the enzyme to substrates, which was highlighted by the MM-PBSA binding energy analysis. These findings highlight the key insights of structural consequences caused by R213H mutation, which would enrich the understanding regarding the role of SULT1A1 mutation in cancer development and also xenobiotics management to individuals in the different treatment stages.

Keywords: binding site; cancer; molecular dynamics simulation; R213H; rs9282861; SULT1A1

1. Introduction

Various endogenous and exogenous compounds in the body are metabolized by the sulfonation process, which is generally initiated by sulfotransferases [1,2]. Among the members in the sulfotransferase family, SULT1 isoform SULT1A1 (Sulfotransferase under the Family of 1A Member 1) is the most abundant and distributed throughout different tissues [3,4]. Compounds including, drugs, steroid hormones, neurotransmitters, heterocyclic aromatic amines (HAAs), and some polyaromatic hydrocarbons (PAHs) [4–8] generally undergoes the sulfonation process [9,10], where the addition

of sulfonate (SO_3^-) group enhances the substrate hydrophilic potentiality or polarity and generates less toxic metabolites, and finally facilitates them for excretion [5,11]. In some cases, sulfonation of exogenous compounds turns them into active mutagens and carcinogens [12–14], although it sometimes acts as a chemical defense mechanism against xenobiotics [15].

Recently, mutation at 213 positions of SULT1A1 exon 7 (rs9282861) [16,17] was reported by multiple studies to be involved in breast cancer progression [18], especially among the postmenopausal Asian women [19,20]. A meta-analysis showed that polymorphic rs9282861 has a significant association with bladder cancer [21], while a small scale case-control study found a significant association with the prevalence of intestinal cancer [22]. The significant association was also found for developing head and neck cancer [23], lung cancer [11,17,24], and also accountable for acute leukemia in infants and acute myeloid leukemia [25]. Patients having rs9282861 mutation, which caused the change of an amino acid from arginine to histidine at the 213th position (R213H), may have a higher susceptibility for developing esophageal cancer, due to the lower enzymatic activity and thermal stability that affect the procarcinogens metabolisms [4,8,26,27], including various steroid hormones, neurotransmitters, and non-opiateanalgesics [28,29]. Furthermore, the substitution occurred with amino acid having the same side chain polarity, although the mutation was reported to reduce the binding affinity of various substrates, as well as their pharmacokinetic properties [27,30,31].

Despite the pathological role of this rs9282861 mutation in SULT1A1, the contribution of this substitution in the alteration of substrate-binding pocket and reduction in enzymatic activity is not clearly understood and need to be clarified in detail at molecular levels. In the present study, we performed a detailed analysis of the role of SULT1A1 R213H in substrate binding by using long-term molecular dynamics simulations. Here we show the atomic conduct of conformational change in regard to destabilization, secondary spatial structures and folding schemes of wild and R213H SULT1A1.

2. Results

2.1. Architecture of SULT1A1 and Stability of Simulation Systems

The SULT1A1 is a sphere-shaped protein, which has a single α/β domain and α-helix surrounded five β-sheets at the central region (Figure 1a). The β-sheets are responsible for the substrate binding and are conserved in cytosolic sulfotransferase [32]. The dimerization site relies on the N terminal end of sulfotransferase and hypothesized as the source of substrate inhibition [33]. The catalytic activity of SULT1A1 depends on the Glu[83], Asp[134], Glu[246], and Asp[263] residues [2,34]. The mentioned structure contains a connected molecule named *p*-nitrophenol (PNP). The motif, [257]RKGMAGDWKXXFT[269], which is similar to the P-loop observed in ATP and GTP binding proteins, is supposedly essential for the PAPS binding. The 5′-phosphate positioning is important for the organizing co-factor, thus transferring sulfonate group to the substrate [5,17]. The [45]TYPKSGTT[52] loop is termed as phosphosulfate binding (PSB) loops of SULT1A1 and responsible for binding 5′-phosphate group of PAP. There are three substrate binding loops of SULT1A1 named as Loop1 to Loop3 with residues of 81–90, 145–154, and 235–263, respectively. The substrate binding site of SULT1A1 is hydrophobic and flexible, permitting this enzyme to embrace various structures, thus it can interact with a variety of L-shaped uncharged substrates including, aromatics (diiodothyronine), tiny aromatics (PNP), and fused ring compounds [2,31,35,36]. The mutation of our interest, R213H is located between Loop2 and 3, and its contribution to SULT1A1 dysfunction was systematically analyzed by 100 ns molecular dynamics simulation by considering both wild and mutant type structures in the presence and absence of substrate (Figure 1b,c).
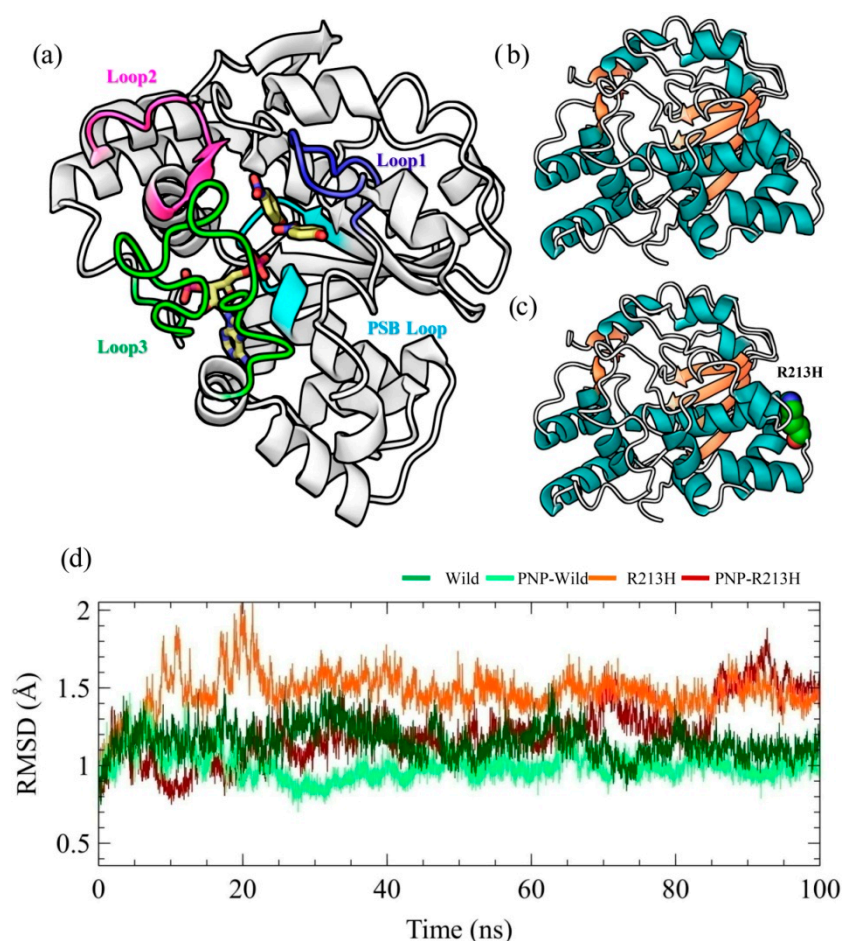
Figure 1. Substrate binding motifs in SULT1A1 (**a**), along with wild (**b**) and mutated (R213H) (**c**) type structures. Protein thermodynamics stability during the simulation was evaluated through root-mean-square deviations (RMSDs) for wild, R213H, PNP-Wild complex, and PNP-R213H complex, by considering backbone atoms (C, Cα, and N) of protein (**d**). Here, dark green, light green, orange, and dark red colors denote wild, PNP-Wild, R213H, and PNP-R213H, respectively.

The root-mean-square deviation (RMSD) of the protein was analyzed in all systems for the original description of simulation performance and stability, reflecting the general thermodynamic stability of the system (Figure 1d) [37]. The RMSD represents the conformational stability of the protein structure during the course of the simulation, where the larger deviations indicate more flexibility of the protein structure [38]. According to the RMSD plot, all wild and mutant type structures with and without ligand achieved equilibrium after 10 ns and remained stable afterward, which allows an appropriate framework for further analyses [39,40]. In both wild and mutant types, conformational flexibility of the SULT1A1 was seen to reduce by the binding of ligand, as revealed from the RMSD calculation (Figure 1d). Although it has been ascertained that many proteins undergo side-chain flexibilities upon the binding of ligands [41], the findings derived from RMSD calculation are consistent with the previous experimental results that ligand binding decreases backbone flexibility of SULT1A1 [42–46]. Furthermore, the RMSD patterns were changed due to the mutation in both apo and holo form, indicating that mutation induced the changes in overall structure, in terms of flexibility and ligand stability.

2.2. Effects of Mutation on Conformation Stability

To observe the outcome of the mutation on inclusive protein stability, the radius of gyration (Rg) was calculated, which indicates the mean-square mass-weighted root range of a set of atoms that

shared the mass centre [47]. The higher Rg values indicate loose packing of the protein structure, which means conformation that is more flexible. Additionally, the solvent accessible surface area (SASA) was calculated, where the lower values of SASA indicates the contracted nature of protein [48]. According to the Rg analysis, no significant difference was found between the wild and mutated protein, and a similar effect was also observed in case of SASA profile (Figure 2). However, binding of the ligand showed the effect in protein conformation, where the mutated PNP-R213H complex showed the higher Rg value for around 18.13–18.5 Å and the lower Rg value was observed for the PNP-Wild complex around 18.1–18.48 Å (Figure 2a). In the case of the wild type, it was clearly seen that the flexibility of the protein was reduced after binding of the ligand, thus increasing the compactness of protein folding. In contrast, the mutant type retained similar flexibility patterns to its apo form. The similar fashion of structural deviation in wild and mutant types was also observed in the SASA calculation (Figure 2b), where binding of ligand reduced the total SASA values. The PNP-R213H complex showed a higher SASA value than the wild type complex. At the initial stage, R213H complex was seen to fluctuate but after 7 ns, it remained stable and continued to 50 ns. Afterward, the SASA values decreased rapidly at 60 ns and again remained stable at 70 ns to the rest of the simulation period.
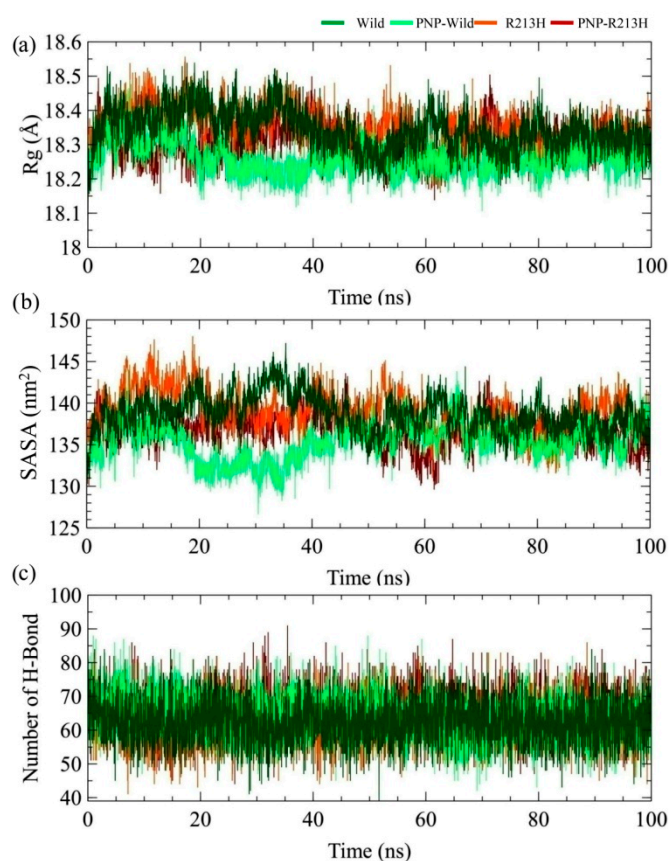

Figure 2. Stability of the protein during the simulation by means of radius of gyration (**a**), solvent accessible surface area (SASA; (**b**)), and number of intra-residue hydrogen bond (**c**) for both wild and mutated type structures. Here, dark green, light green, orange, and dark red colors denotes wild, PNP-Wild, R213H, and PNP-R213H, respectively.

The PNP-Wild complex, on the other side, showed less SASA values about 126.5 nm^2 at 31 ns and its retained stability at 40 ns, and also persisted stable through the simulations. The overall flexibility of the protein can be modulated by the number of the internal hydrogen bond, which was calculated and shown in Figure 2c. However, no significant difference was observed in the total hydrogen bonding (Table S1). Overall, Rg and SASA calculation denoted the mutation-induced flexibility in the protein–ligand complex, which was consistent with the RMSD analysis.

2.3. Effects of Mutation in Protein Dynamics

Correlative motions play a crucial part in recognizing and interacting bio-macromolecular system, which can be achieved from the covariance equation of molecular fluctuation generated through the molecular dynamics (MD) simulation trajectory. The correlative motions of different simulation systems were analyzed by the dynamic cross correlation matrix (DCCM). The DCCM fluctuations are depicted in Figure 3, which graphically displays time-correlated information among the residues of the protein. The deeper color intensity highlights more positive or negative correlated motions between the structures. The red color indicates strong positive correlation whereas; the blue color indicates a negative correlation. The R213H protein and PNP-R213H complex exhibited different movements compared to wild protein and PNP-Wild complex, respectively. Higher positive and negative correlations were observed in wild type, especially within the inter loop residues of the active site, that is, between loop3 and loop2 (Figure 3a). Loop3 also showed negative correlation with the residues located in loop1, however, residues contained in the loop1 were further anti-correlated with the PSB loops containing residues. Besides, the mutation showed only the negative, but insignificant, correlations between the residues of PSB loop and loop1 (Figure 3b). Interestingly, these correlations were lost in the case of ligand binding conferring the stable protein-ligand complex; instead, the complex generated random motions within the loops (Figure 3c). The ligand bounded R213H protein exhibited the same correlations as in unbounded forms. Moreover, a few but new types of correlations were observed between loop1 and loop3 (Figure 3d), which was different from the wild complex.

In addition, principle component analysis (PCA) was performed to observe whether these correlations resulted in any significant flexibility of the protein. The PCA can be applied to any system and permits to study the influence of any varying parameters, by lessening the collective motions complexity [49–52], which is associated with the phase space behavior related to protein functions and stability. Therefore, it is frequently used to characterize the conformational variances involved in the protein folding [53,54], opening and closing mechanisms of ion channels and other proteins [55–58], as well as the conformational dynamics induced by the mutations [46,50,59,60] and the ligand binding [61]. The motion of the C_α atoms was described by the eigenvectors of the covariance matrix, which were calculated in PCA analysis and resulted in first PCA of 22.72%, 23.03%, 16.48%, and 21.05% for wild, R213H, wild-PNP, and R213H-PNP systems, respectively (Figure S1). Compared to the wild and wild-PNP systems, the apo and holo structure in the mutated form showed unusual patterns on to the phase space of the foremost two principal components (PC1 and PC2), describing the significant alternation in the protein conformation. Furthermore, the highest PC1 in the wild apo structure described that wild apo structure was undergone conformational changes, while its ligand binding state, i.e., binding with ligand reduced the conformational stability (Figure 4a). A closure looks to the local residue mobility, in terms of PC1, ensured that the conformational changes were induced in the active site loops, which was decreased binding of ligand and in mutated forms (Figure 4b). In mutated form, the holo structure increased the movement in loop3, while reduced in PSB loop and loop1 and 2, compared to both wild type structures (Figure 4b). These results are correlated with the finding from DCCM analysis and confirmed that the residues in the active site endured conformational changes due to the mutation.
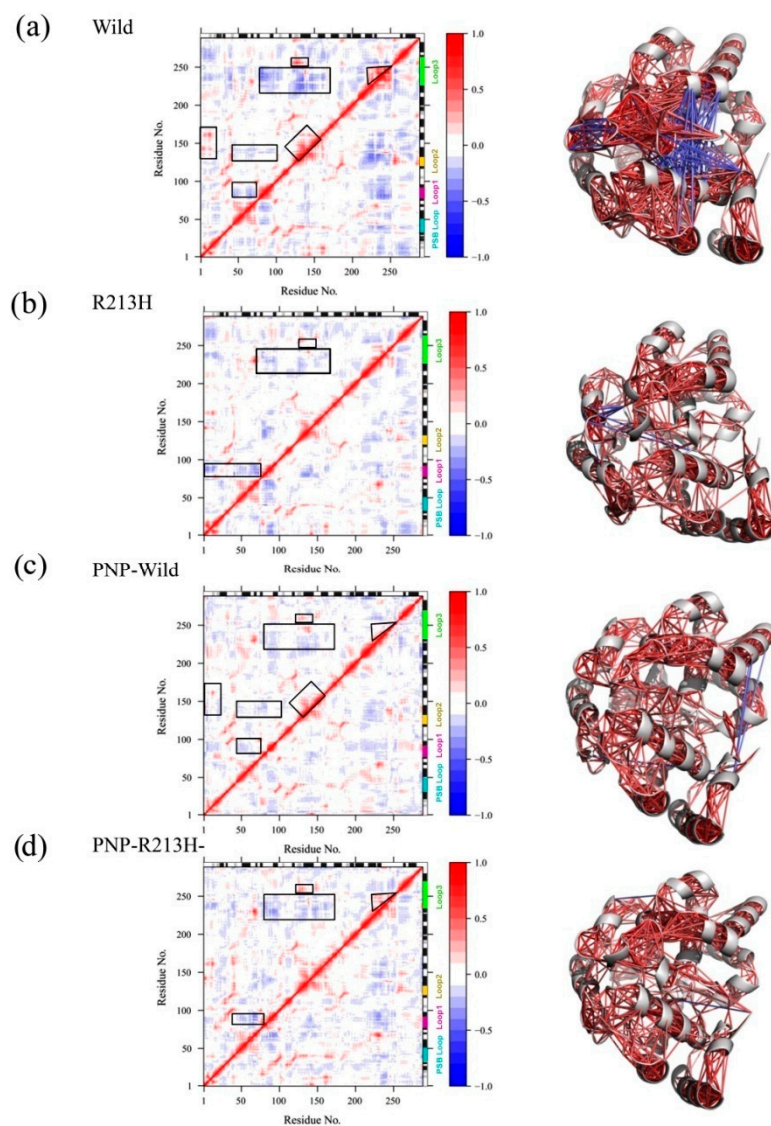
Figure 3. The residual correlative motion in protein structures was visualized by the dynamic cross correlation map analysis, for wild (**a**), PNP-Wild (**b**), R213H (**c**), and PNP-R213H (**d**). Here, strong correlative motions (±4 to ±1) were depicted in three dimensional cartoon model on the right side of each panel. In all cases, correlations were represented in the color spectrum, from red to blue. The positive correlation between two residues, that is, both residues move the same direction, is represented as red color. On the other hand, residues having opposite motion are denoted as the anti-correlative motion, marked as blue color.
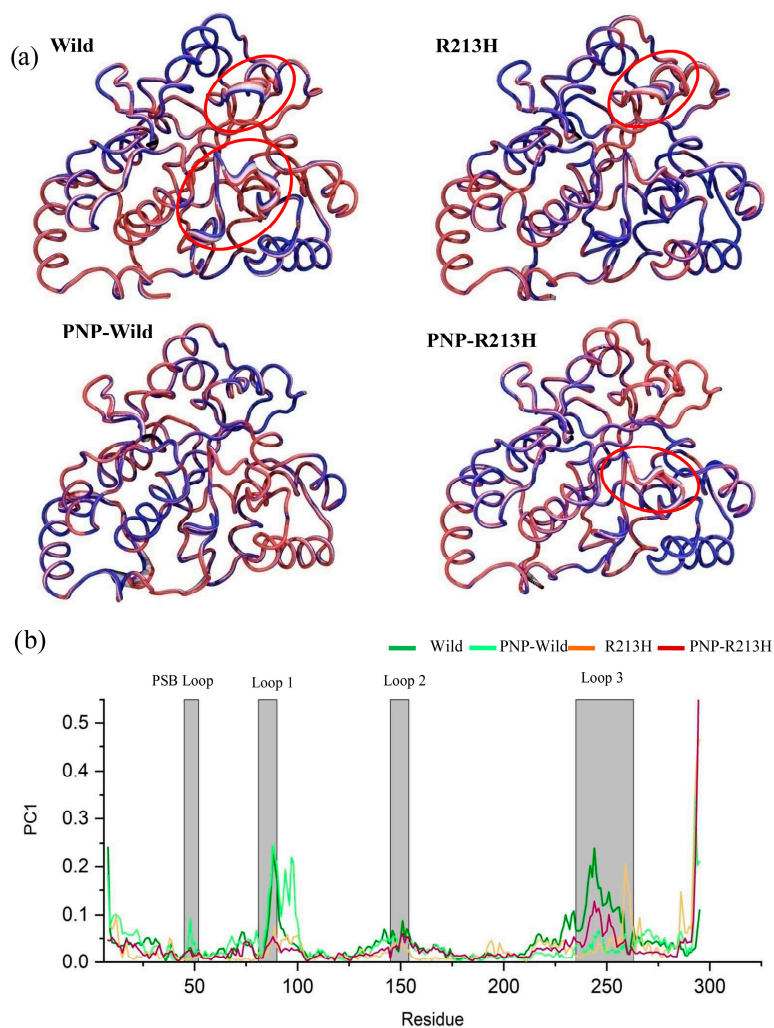
Figure 4. Changes of protein dynamics due to mutation by means of principle component analysis. Here, principle component 1 is represented in the tube structures of the three dimensional structure of wild, PNP-Wild, R213H, and PNP-R213H, respectively (**a**). The widen tube means flexibility, highlighted by the red circle. In addition, residue based mobility plot showing the variation in the mobility in PC1 for different simulation systems, where dark green, light green, orange, and dark red colors denote wild, PNP-Wild, R213H, and PNP-R213H, respectively (**b**).

With the purpose of observing the dynamics changes in the active site, Define Secondary Structure of Proteins (DSSP) analysis was performed to analyze the structural stability in the loops, shown in Figure 5, which provides detailed information about the loss and gain of the secondary structure all through the simulation. As shown in the results, the secondary structure content was stably maintained by all residues in both wild and mutant types, except the residues located in the loop3. The residues located 235–263, showed diverse conformations throughout the simulation in different systems. In wild type, these residues, including 255–263, showed 3-helix conformation over time (Figure S2). However, this conformation was changed to A-helix after 20 ns of simulation due to the mutation (Figure S2c). Furthermore, mutation reduced the bend conformation of the loops, which seemed to increase in case of wild complex structure (Figure S2c). In both apo and holo structures, mutation induced the persistency of B-bridge conformation in the loop3 (Figure S2) denoting that mutation increased the rigidity of the loops that may damage the plasticity of the active site.
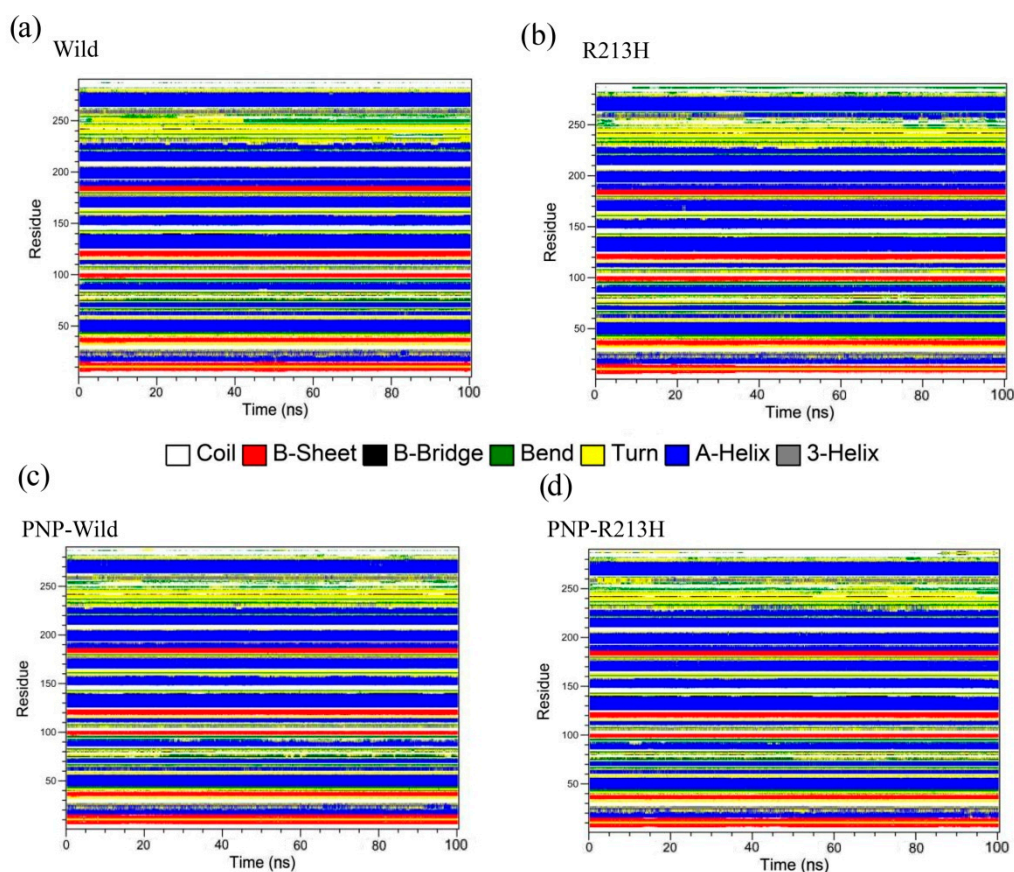
Figure 5. Evolution of the secondary structure elements of wild (**a**), R213H (**b**), PNP-Wild (**c**), and PNP-R213H (**d**) during molecular dynamics simulations based on DSSP classification.

2.4. Effects of Mutation in Active Site and Ligand Binding

Before the detailed investigation, the residue fluctuation profile due to the mutation was considered, thus RMSF analysis was performed on simulation trajectories (Figure 6). The RMSF calculation represents the dynamics of the backbone atoms, where higher values designate greater flexibility through the MD simulation. The highest peaks on this graph show the protein regions, which were mostly fluctuated throughout the simulation period. The protein in all simulation systems showed RMSF within a range of 0.2–9.8 Å. As can be seen in Figure 6, the overall RMSF value of the residues of wild and mutated proteins was almost the same in the PSB loop, loop1, and loop2, however, the difference was observed in loop3, indicating a major binding site of the ligand. Evidently, R213H mutation induced higher fluctuation in the loop3 segment compared to the wild type, which was reduced by ligand binding. For instant, binding of ligand showed less fluctuation compared to wild type, while the mutation reduced more than that of wild type.
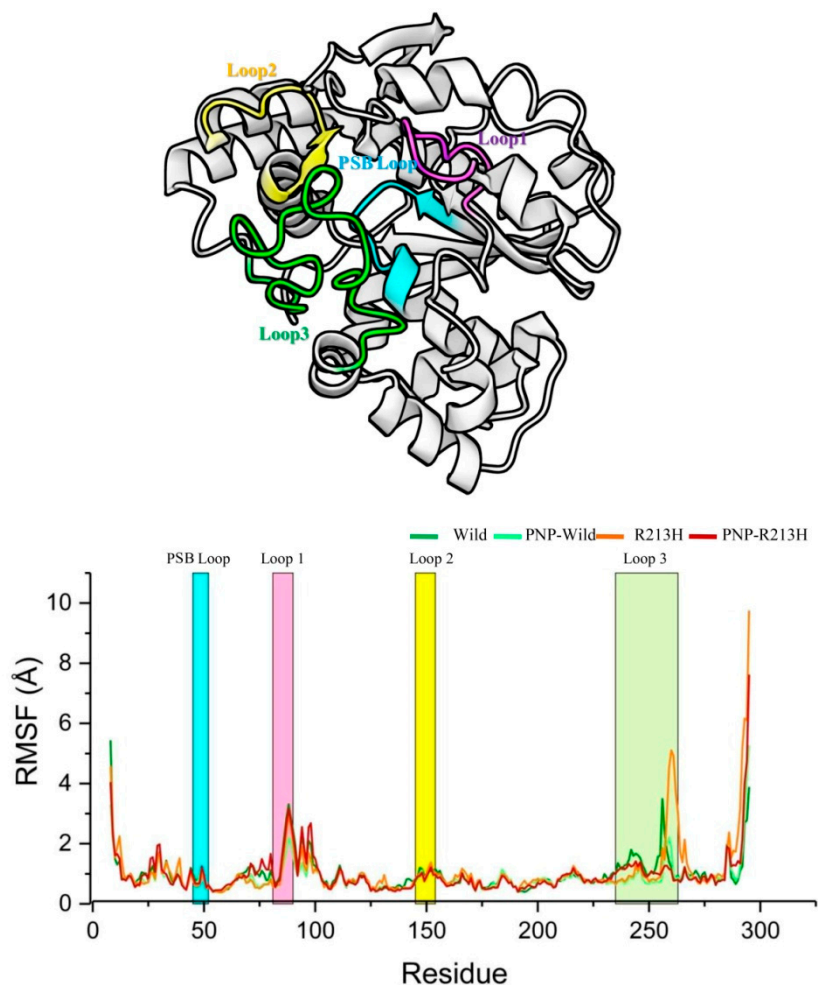
Figure 6. The root mean square fluctuation (RMSF) of the Cα atoms for the wild protein, R213H, PNP-Wild complex, and PNP-R213H complex. In this graph, dark green line represents wild protein, whereas light green, orange and dark red colors denote wild, PNP-Wild, R213H, and PNP-R213H, respectively. The horizontal *x*-axis bar region denotes different loop regions.

Accordingly, the ligand RMSD and intermolecular interaction profiling was considered to analyze the effect of rigid behavior of loop3 on ligand binding, where the plot showed that mutation reduced the ligand flexibility in the active site, while ligand was seemed to be very flexible in the pocket of native structure, as revealed from the wild type complex (Figure 7a). Likewise, the intermolecular hydrogen bonding analysis showed that the number of hydrogen bonding partners, specifically, residues from the active site was decreased in the mutated structure; the compound only showed maximum interaction with His108 (Figure 7c). Surprisingly, mutation also reduced the total contact with the active site residue (Figure 7e), and induced a notable change in protein–ligand interaction network. Specifically, the total number of ligand-mediated contact was reduced in the mutant complex in the region of loop1 and loop3 (Figure 7e), however, mutation induced the total contact with the residues in loop2, including Ala146, Val148, and His149. Besides the ligand also maintained the contact with the residues from PSB loop such as Pro47 and Lys48, whether these inter molecular contacts were favorable to binding affinity or not, additionally binding energy calculated by MM-PBSA method was conducted on resulted 4000 snapshots from 100 ns simulation [62]. MM-PBSA binding energy results were depicted against time for wild-PNP and PNP-R213H, which is shown in Figure 7b. The average binding energy for wild-PNP complex was 110 kJ/mol, while 60 kJ/mol for mutant types (Table 1). Previous studies based on biochemical pharmacogenetic assays presented that, R213H caused very low thermostable sulfonation activity to 4 mM 4-nitrophenol, while the wild type exerted a higher

binding affinity [63,64]. Consistent with previous findings, the MM-PBSA analysis thus revealed that the mutation reduced the binding affinity to the PNP, as a result of changing intermolecular contacts, which might contribute to low sulfonation activity.
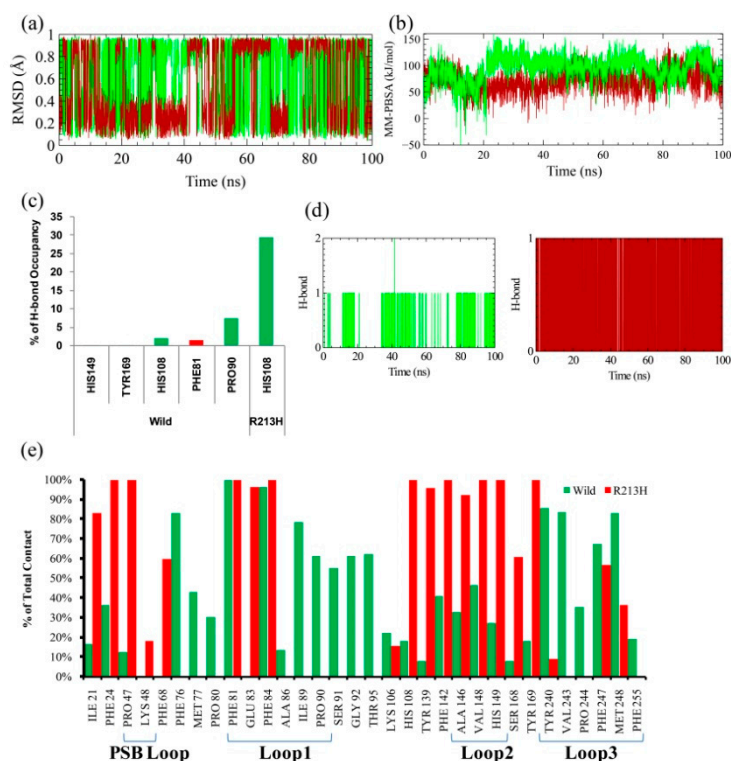

Figure 7. Ligand RMSD in both wild (green line) and mutated (dark red line), which means the degree of ligand flexibility in substrate binding site (**a**). MM-PBSA binding energy analysis calculated from the trajectory of 100 ns molecular dynamics simulation with AMBER14 force field, using YASARA dynamics software (**b**). The total number of intermolecular hydrogen bonds (**c**) and their occupancy with their respective donor and acceptor (**d**). Total intermolecular contact analysis (**e**). Here, green color represents the residues of wild protein with their contact percentage where red color indicates R213H protein contact percentage.

Table 1. Change of protein–ligand stability by means of MM-PBSA binding energy analysis, and the active site characteristics.

Systems	MM-PBSA [$] $\Delta G_{binding}$	CASTp [&]		MM-PBSA [#] $\Delta G_{binding}$				
	(kJ/mol)	Volume	SASA	ΔE_{vdW} (kJ/mol)	ΔE_{elec} (kJ/mol)	ΔG_{pol} (kJ/mol)	ΔG_{nonpo} (kJ/mol)	$\Delta G_{Binding}$ (kJ/mol)
Wild		453.31	810.58	–	–	–	–	–
R213H		596.88	1020.92	–	–	–	–	–
Wild-PNP	110 ± 1.03	336.87	537.45	−80.18 ± 0.27	−8.23± 0.54	77.80 ± 0.36	−89.84 ± 0.30	−120.42 ± 0.68
R213H-PNP	60 ± 2.46	236.66	607.83	−73.25 ± 0.27	−24.06 ± 0.54	66.27 ± 0.56	−83.99 ± 0.34	−115.00 ± 0.62

[$] Calculated based on the trajectories obtained 100 ns molecular dynamics simulation by YASARA dynamics software using default MM-PBSA binding energy calculation protocol. [&] Calculated by CASTpver 3.0 server (http://sts.bioe.uic.edu/castp/index.html). [#] Calculated based on the trajectories obtained 25 ns molecular dynamics simulation by GROMACS software using the g_mmpbsa tool.

2.5. Insights Into Substrate Binding

Since the mutation disrupted significant intermolecular interactions between the loop residues and the ligand, the conformational variability of these loops was mapped onto the free energy landscape by considering the radius of gyration of loop1 and loop3 as the reaction coordinates. The disperse area with color depth from light to dark red color indicates the conformational variability of the loops,

where the darker area represents the conformation with the lowest energy minima. The degree of dispersion describes the conformational flexibility, where the larger area denotes more flexibility. As can be seen in the Figure 8, the conformation of the loops with the lowest energy was dispersed through the projection, where the two largest resulting clusters were seen to shift towards the larger Rg value, describing the loops flexibilities during the simulation (Figure 8a). On the other hand, the energy minima of R213H was seen to shift towards the lower Rg values, in the case of both loops, demonstrating the loss of loop flexibility (Figure 8b). In the case of ligand binding, wild type showed normal distribution in the phase behavior, moreover concentrated, which explains loop contribution to the ligand binding (Figure 8c). Consequently, mutant reduced the conformation flexibility more than the wild type in case of ligand binding, where the flexibility of loop3 was seen to be reduced, while the distribution energy minima were shifted toward the larger value of loop1 (Figure 8d).
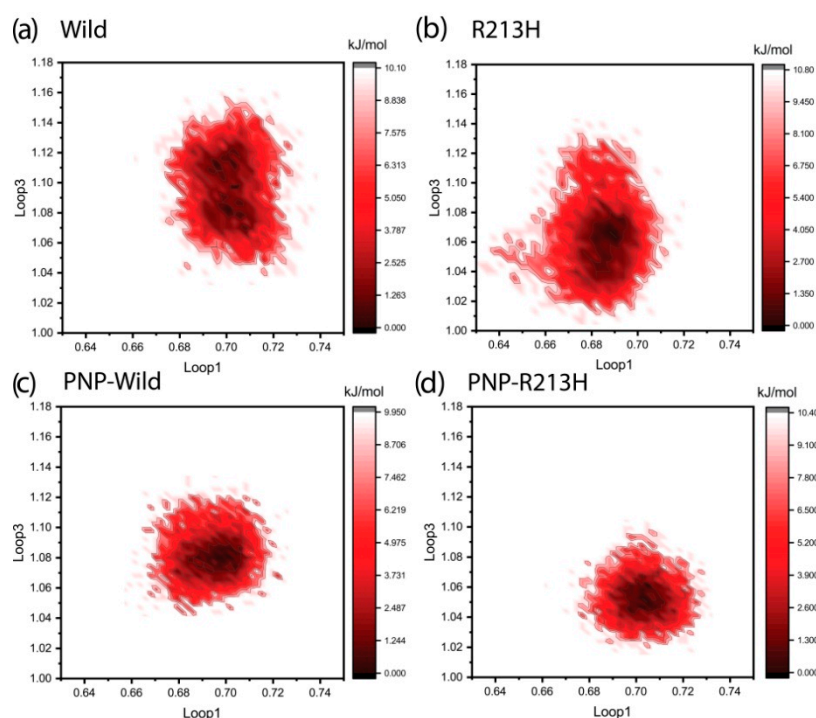

Figure 8. Energy landscapes of different simulation systems including Wild (**a**), R213H (**b**), PNP-Wild (**c**), and PNP-R213H (**d**) respectively. Here the reaction coordinates are the Rg (nm) of loop1 (residue 81–90) and loop3 (residues 235–263) as x and y axis, respectively. The color scheme to the right of each plot is given in units of kJ/mol.

In order to understand the effect of mutation in the substrate binding, the conformer having the lowest energy was further isolated from the simulation trajectories of holo protein based on the additional free energy landscape analysis. The free energy landscape was thus constructed based on the RMSD and Rg time-dependent changes to the original structure, which demonstrates the effect of mutation on global conformational space of SULT1A1 (Figure S3). The energy minima in the figure were represented from red to blue color, where the more concentrated blue zones described the more stable conformation with global minimum energy. On the other hand, the shape and size of the area display the conformational stability of the protein. As can be seen in Figure S3, the distribution of global energy minima was changed due to the mutation, although both wild and mutant types exhibited Gibbs free energy varying from 0 to 11 kcal/mol. Since SULT1A1 in both wild and mutant types exhibited stable conformation in the lowest energy level of 0 kcal/mol, the respective conformation was visualized and analyzed for intermolecular interaction analysis (Figure S3).

Interestingly, significant fluctuations were noted between wild and mutant forms in the protein–ligand interactions (Figure 9). In wild complex, the ligand formed multiple hydrophobic

interactions with the residues from loop1 and loop3, including Phe81, Phe84, Val243, and Met248 (Figure 9a). On the other hand, it showed only hydrophobic interaction with Phe24 residue and hydrogen bond with His108 in case of mutant type (Figure 9b), which indicates that mutation might alter the active site of the protein, therefore the volume and total SASA of the ligand binding pocket was calculated with the help of CASTp server, where the higher SASA is more pronounced for hydrophilicity, shown in Table 1. The analysis displayed that mutation increased the active site volume and SASA of apo SULT1A1, however, the volume was reduced in complex with ligand, in both cases, SASA was still increased (Table 1).
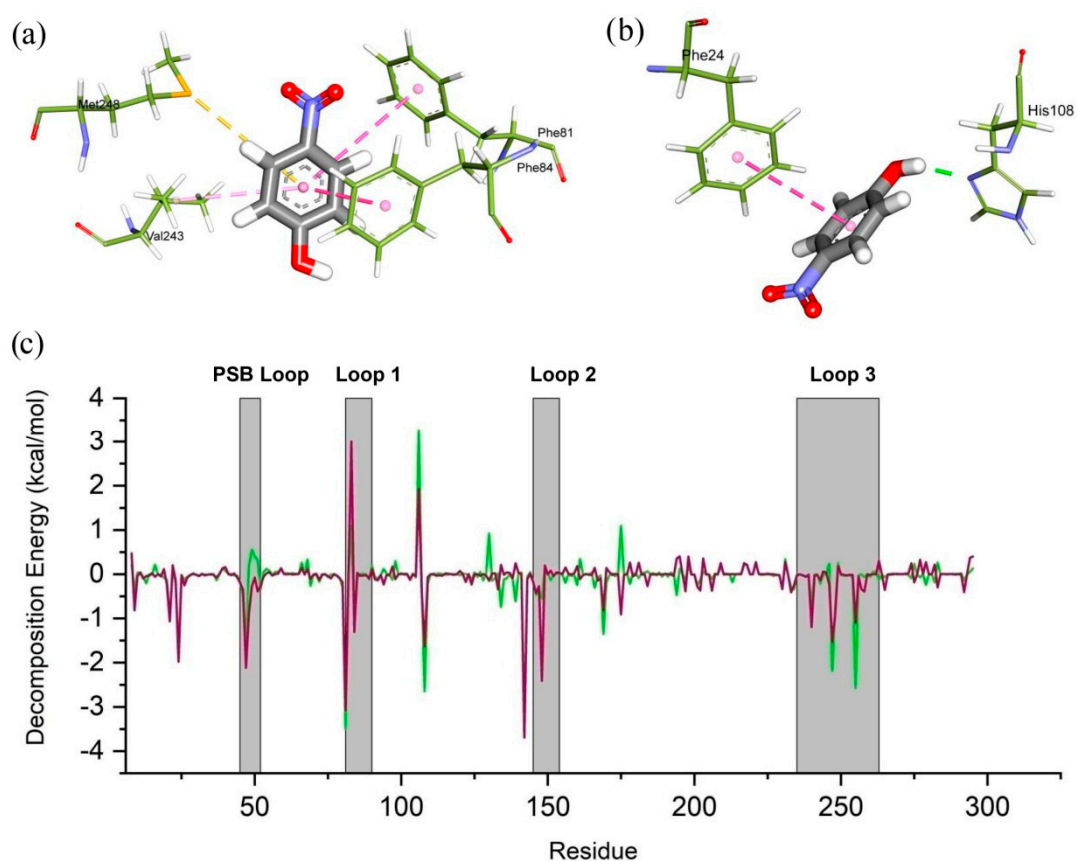

Figure 9. The intermolecular non bonded interaction pattern is rendered for wild (**a**) and mutated (**b**) complex derived from the free energy landscape. Here, pink color denotes Pi-alkyl bond, yellow for pi-cation, and green for hydrogen bond, respectively. Per-residue decomposition analysis (**c**). Here, green color denotes PNP-Wild protein whereas, magenta color represents PNP-R213H protein.

Moreover, per-residue decomposition analysis, which describes the contribution of each resides in total binding energy, was conducted to elucidate the mutation-induced effect on ligand binding, based on the supplementary 25 ns MD simulation by GROMACS software. Noticeably, the mutation caused significant changes in the binding energy contribution, where it was represented that, mutation reduced the favorable contribution of the residues from loop1 and loop3, especially Phe247, Phe255 and Phe81, and Phe84 respectively, while it increased the unfavorable contributions of Leu82, Glue83, Ala86, and Tyr240 (Figure 8c). Furthermore, the mutation also showed a reduction in the total MM-PBSA binding energy than the wild type, supporting the previous MM-PBSA analysis by YASARA dynamics calculation. These interpretations were consistent with fluctuated notion that conformational flexibility of loops in the active site plays a vital role in the ligand binding, which eventually disrupted by R213H.

3. Discussion

Sulfonation has been recognized as the vital metabolic reaction for metabolizing not only various endogenous compounds but also bioinactivation of a variety of xenobiotics [10,65–67] through the action of cytosolic sulfotransferase (SULT). The genetic polymorphisms of *SULT1A1* [26,68–70] caused the theatrical changes in the protein function. In this study, we considered SULT1A1 R213H, which is associated with various cancers. To look for key structural consequences offered by the mutation for this deleterious behavior, a series of comparative analyses were accumulated in this study based on computational analysis.

The molecular dynamics simulation was thus conducted to reveal the dynamic changes of the protein due to the mutation in both apo and holo form, where the initial characterization by RMSD showed that mutations induced different transitional pathways from the initial to the final stage of simulation. Consistent with the RMSD profile, the Rg and SASA calculations also confirmed the mutation induced conformational flexibility in both apo and holo form of SULTA1, where mutated types increase the dimension and solvent accessibility of the protein. In dynamics condition, protein undergoes various conformational changes for particular function, where residual communication in terms of correlative motions serves a vital role in recognizing and binding of various biological macromolecules, and this communication is usually disrupted by the mutation [71]. The correlative motion in SULTA1 based on the DCCM analysis showed that mutation reduced the correlative motions particularly in the loop1 and loop3 of active site as a results structure flexibility is lost in those positions, which was clearly reflected through collective motion analysis by PCA. Furthermore, mutation also increased the correlated motions in mutated-PNP complex compared to the other systems, which indicates the functional disruption of SULTA1. These consequences were further supported by RMSF analysis, where the effect was revealed in the loop1 and loop3 regions.

The active site of SULT1A1 is plastic, which is maintained by the aforementioned three loops, can process various conformational changes to adopt a verity of aromatic compounds [2,31,35,36]. Furthermore, the binding of these aromatic substances to the active site is highly modulated by Phe142 and Phe81, which creates an entry portal that allows only the catalytic site to bind linear substrates [72]. On the other hand, Val148, Phe247, and Met248 in the loop3 maintained interactions with nitrogen containing groups of the substrate followed by water bridge and van der Waals bonds [31].

Previous study showed that substitution of His213 makes the protein more thermolabile than the wild type [73], where the crystal studies presented that the mutation influence both structural stability and substrate regulation of SULT1A1 [31]. In this regard, Lu and colleagues identified that the substrate binding affinity and kinetics are susceptible to the loss of secondary structures stability [74], which is achieved by the mutation. Perhaps, it has been presumed that the mutation might influence the PAPS binding to PSB loop by changing the interaction of Tyr62-Arg213 [30], thus contribute to loss of enzymatic activity. Interestingly, total contact analysis showed that mutation reduced total number of interactions with the ligand during the simulation, which was also revealed in the analysis of the most native structure with lowest energy minima, calculated from the free energy landscape. Furthermore, it appeared that the active site volume and total hydrophobicity were reduced (as SASA is increased) due to the mutation; thus, the ligand flexibility in the active site was hindered, which was necessary for maintaining maximal interactions with the pocket residues. Besides, the increase of hydrophilicity in the active site resulted in unfavorable binding of the ligand, which might affect ligand binding properties as well as selectivity.

Overall, the PAPS site of cytosolic SULTs is characterized by the conserved residues, however, the core hydrophobic substrate binding site exhibits wide range of substrate specificity, which allows sulfonation for a specific substrate [2]. Accordingly, the uncharged phenolic compounds are preferred by SULT1A1, where the mutation R213H disorders the plasticity of substrates binding loops and increased the hydrophilicity of the pocket, affecting the sulfonation process, and thus contributes in cancer progression.

4. Materials and Methods

4.1. Preparation of the Simulation System

The structure of SULT1A1 has been extracted from the protein database (http://www.rcsb.org/pdb) [75] in three-dimensional crystal frameworks with a protein data bank (PDB) ID of 1LS6 [31], having the molecular mass of 34.165 kDa. The structure was initially prepared by adding bond orders, hydrogen and charges; and also refined by removing water molecules and optimizing it at neutral pH. The structure was further fixed by correcting some amine derivatives, thiol groups, hydroxyl groups, protonated glutamic acids, aspartic acids, and histidines. In order to adjust the heavy atom of the structure, energy minimization has been applied by using optimized potentials for liquid simulation (OLPS3) force field until the RMSD reached to 0.30 Å. Since the structure already contained the histidine residue at the position of 213, the wild type structure (H213R) was further constructed by computational mutagenesis methodology, using Schrödinger Mutate Residues script 2017-1 suite (LLC, New York, NY, USA) [76]. Additional refinement by short molecular dynamics simulation was performed to improve the structure quality. For that, 500 ps of molecular dynamics simulation was conducted by using Yet Another Model Building and Energy Refinement force field (YAMBER3) force field [77], followed by maintaining 298 K temperature at pH 7.4. The simulation was performed in the transferable intermolecular potential3 points (TIP3P) water model with a density of 0.997. Based on the lowest energy, the final representative structure was considered for further study. For ligand binding characterization, the native co-crystal ligand, *p*-nitrophenol (PNP) was only considered and kept in protein structure. Accordingly, four individual simulation systems were developed, including wild, wild-PNP, R213H, and R213H-PNP.

4.2. Molecular Dynamics Simulation

The consequence of mutation on SULT1A1 protein structure and stability was analyzed by molecular dynamics simulation. The YASARA dynamic software (YASARA Biosciences GmBH, Vienna, Austria) [78] was used to analyze the dynamics of the wild and R213H type structure in the presence and absence of ligand, according to the previous protocols [59,62,79,80]. The process was carried out by cleaning each structure, followed by optimizing the hydrogen bond network. A cubic simulation cell was then constructed with a periodic boundary condition. The assisted model building with energy refinement14 (AMBER14) force field [78,81,82] was used to input the atoms of every single complex. This force field was preferred because, in regards to knowledge-based interactive potentials of the homology models, it was optimized for structure prediction, refinement, and energy minimization. The ligand was parameterized by AutoSMILES [83] algorithm, which used combined, AM1BCC [84] and General AMBER Force Field (GAFF) [81] for typing atomic charges and bond orders to unknown organic molecules. The TIP3P solvation system was used with 0.997 g/L density for solvating the simulation cell, and acid dissociation constant values (pKa) were calculated for titratable amino acids in the protein [85]. The TIP3P solvent system was known to give the best experimental output with a combination of AMBER14 force field [86–89]. The pH of the system was maintained at 7.4 to mimic the physiological conditions, and accordingly, the protonation states of each amino acid residue were determined in combination with H-bonding network and SCWRL algorithm, which employs dead-end elimination and graph theory [90]. Furthermore, the solvation system was supplemented with Na^+ and Cl^- ions [91]. Followed by energy minimization through the steepest descent approach, the conformational stress of the system was reduced by applying simulated annealing protocol. In order to explain the long-range electrostatic interactions, the Ewald particle mesh (PME) was used, where the distance cut-off was set to 8 Å. Finally, 100 ns molecular dynamics simulation was performed with Berendsen thermostat at a time step interval of 2.00 fs, together with a multiple time step algorithm [92,93]. The pressure was set constant and trajectories were accumulated with an interval of 25 ps. A preinstalled macro (md_run.mcr) within the YASARA suite was used for all simulation phases, while the results was characterized using YASARA suite

along with VMD (Version 1.9.3, Theoretical and Computational Biophysics Group, Urbana, IL, USA, 2016) [94] and DSSP (EMBL, Heidelberg, Germany) [95,96] tools. The corresponding trajectories were used in different evaluative procedures to analyze structural insights and stability through root mean square deviation (RMSD), radius of gyration (Rg), solvent accessible surface area (SASA), number of hydrogen bonds, dynamic cross correlation matrix (DCCM), root mean square fluctuation (RMSF), and secondary structure element (SSE) analysis.

Initially, the resultant trajectories were subjected to evaluate binding energy by MM-PBSA calculation, for that integral YASARA binging energy Macro was used. The method of calculation follows the theory of nuclear physics [97]; that is, the better binding is indicated by positive energy [62,98]. The equation is as follows,

$$\text{Binding Energy} = E_{potRecept} + E_{solvRecept} + E_{potLigand} + E_{solvLigand} - E_{potComplex} - E_{solvComplex} \quad (1)$$

where, YASARA built in Macros was used to calculate MM-PBSA binding energy, using AMBER 14 as a force field, where the more positive energies indicate the better binding [99]. The approach is similar to MM/PBSA, reported by [100].

4.3. Dynamic Cross-Correlation Map (DCCM) Analysis

The Bio3D [101] package integrated with R program was used to calculate residue-residue dynamic cross-correlations to determine how mutations influence the inner dynamics of protein conformations. The Cα-coefficient of cross-correlation was calculated by the average structure. Bio3D "DCCM" derives co-variation of the matrices internally from the coordinates provided and calculates Pearson's co-variance matrices correlation coefficients, calling on "cov2dccm". Based on the following equation, a cross-correlation ratio, C_{ij}, has been considered for the Cα electrons [102]:

$$C_{ij} = \frac{\langle \Delta r_i \cdot \Delta r_j \rangle}{\left\{ \langle \Delta r_i \rangle^2 \langle \Delta r_i \rangle^2 \right\}^{1/2}} \quad (2)$$

where, the average location of ith and jth residues were represented by Δr_i and Δr_j. The time average was represented by the angular bracket. The DCCM values were between -1 and $+1$, where the positive correlation was represented by the positive value and negative correlation represented by negative values.

4.4. Principal Component Analysis (PCA)

Generally, PCA is used to understand the dynamics of any biological system. This technique has been used to achieve diagonal eigenvectors and the equivalent eigenvalues based on the calculation and diagonalization of the covariant atomic fluctuation equation [103,104]. This is also called by their eigenvectors are principal components (PCs), which supply data with their corresponding vectors on movement and displacement of the atoms. The mathematical details have been described previously [105,106].

4.5. Free Energy Landscape Analysis

The free energy landscape maps all the probable macromolecular conformational modifications, by emphasizing their respective energy levels to detect the location of the interaction molecules in the system [107,108]. In the evaluation of free energy landscape, the stability of protein was defined by Gibb's free energy calculation. It also represents the various structural nature of protein's structure-function correlation. In this study, the wild, and mutant protein free energy landscapes were analyzed by the following equations:

$$G_i = -K_B T \ln (N_i / N_{max}) \quad (3)$$

where, Boltzmann's constant is signified through K_B, T stands for temperature (300 K), population bin I is presented by N_i and the most inhabited population bin was depicted by N_{max}. As the smallest provability, an unnatural barrier scale was set for the bin without any populations. A color code model was used to display different energy levels.

4.6. Per Residue Energy Decomposition Analysis

In order to reveal the contribution of active site residue in the binding energy in both wild and mutated system, per residue energy decomposition analysis was performed by using the g_mmpbsa tool, as a function of APBS and GROMACS, thus independent MD simulations were individually performed for each of the complexes [109]. The lowest energy structures of protein–ligand complex derived from the free energy landscape analysis in both wild and mutated type were subjected for additional 25 ns molecular dynamics simulation by GROMACS software, according to the previously described method by Kumari et al. [109], where the protein was parameterized by the AMBER ff99SB force field [110], and the AM1-BCC and GAFF were considered for the van der Waals and bonded parameters for the ligand (See Supplementary file 1 for details). For the calculation, a series of 500 snapshots taken from 25 ns of the total simulation, with intervals of 50 ps [111]. The following equation has been used to calculate the binding free energy [112–114]:

$$\Delta G_{bind} = G_{complex} - (G_{protien} + G_{ligand}) \text{ (e1)},$$
$$\Delta G_{bind} = \Delta E_{MM} + \Delta G_{sol} - T\Delta S \text{ (e2)},$$
$$\Delta E_{MM} = \Delta E_{elec} + \Delta E_{vdW} \text{ (e3)},$$
$$\Delta G_{sol} = \Delta G_{pol} + \Delta G_{npol} \text{ (e4)},$$
$$\Delta G_{npol} = \gamma \times {}_{SASA} + b \text{ (e5)},$$

$$(4)$$

where, the total binding free energy is represented by ΔG_{bind}, the energies of protein–ligand complex, protein (SULT1A1) and the ligand, PNP were depicted by $G_{complex}$, $G_{protein}$, and G_{ligand}, respectively. Binding free energy can also be calculated using the equations e2–e5, where ΔG_{sol} denoted as the solvation energy, ΔG_{pol} and ΔG_{npol} represents polar and nonpolar solvation energy components, respectively. The electrostatic interaction energy was represented by ΔE_{elec} and van der Waals interaction energy represented by ΔE_{vdW}. In equation (e5) γ and b are the experimental parameters, SASA stands for the solvent accessible surface area. Polar solvation effects have been computed by standard Poisson–Boltzmann equation, whereas SASA was used to calculate non-polar solvation energies through the following values γ (0.0226778 kJ.mol^{-1}·Å^{-2}) and b (3.84928 kJ.mol^{-1}) [115].

5. Conclusions

The present study addressed the alteration of structural properties of SULT1A1 due to the R213H mutation, and also its effect on the substrates binding. A variety of structural characterization methods were applied followed by molecular dynamics simulations in this study, where the results highlighted significant structural differences in wild and mutant SULT1A1 proteins in terms of their native and ligand bounded states. Essentially, this mutation caused the actual conformational alteration in substrate binding loops, where the loop3 was mostly influenced. This manipulation changed the plasticity and intrinsic properties of the active site, which reduced not only the ligand flexibility in the pocket but also decreased the binding affinity, as well as the substrates recognition capabilities. Together with previous findings on SULT1A1 R213H, this study hereby contributes to a better understanding regarding the role of SULT1A1 mutation in cancer development.

References

1.   Huang, X.; Cao, M.; Wang, L.; Wu, S.; Liu, X.; Li, H.; Zhang, H.; Wang, R.Y.; Sun, X.; Wei, C.; et al. Expression of sulfotransferase SULT1A1 in cancer cells predicts susceptibility to the novel anticancer agent NSC-743380. *Oncotarget* **2015**, *6*, 345–354. [CrossRef] [PubMed]

2.   Gamage, N.; Barnett, A.; Hempel, N.; Duggleby, R.G.; Windmill, K.F.; Martin, J.L.; McManus, M.E. Human sulfotransferases and their role in chemical metabolism. *Toxicol. Sci.* **2005**, *90*, 5–22. [CrossRef] [PubMed]

3.   Glatt, H. Sulfotransferases in the bioactivation of xenobiotics. *Chem. Biol. Inter.* **2000**, *129*, 141–170. [CrossRef]

4.   Glatt, H.; Boeing, H.; Engelke, C.E.; Ma, L.; Kuhlow, A.; Pabel, U.; Pomplun, D.; Teubner, W.; Meinl, W. Human cytosolic sulphotransferases: Genetics, characteristics, toxicological aspects. *Mutat. Res.* **2001**, *482*, 27–40. [CrossRef]

5.   Barnett, A.C.; Tsvetanov, S.; Gamage, N.; Martin, J.L.; Duggleby, R.G.; McManus, M.E. Active site mutations and substrate inhibition in human sulfotransferase 1A1 and 1A3. *J. Biol. Chem.* **2004**, *279*, 18799–18805. [CrossRef] [PubMed]

6.   Berger, I.; Guttman, C.; Amar, D.; Zarivach, R.; Aharoni, A. The molecular basis for the broad substrate specificity of human sulfotransferase 1A1. *PLoS ONE* **2011**, *6*, e26794. [CrossRef]

7.   Pachouri, S.S.; Sobti, R.C.; Kaur, P.; Singh, J.; Gupta, S. Impact of polymorphism in sulfotransferase gene on the risk of lung cancer. *Cancer Genet. Cytogenet.* **2006**, *171*, 39–43. [CrossRef]

8.   Li, K.; Ren, Y.W.; Wan, Y.; Yin, Z.H.; Wu, W.; Zhou, B.S. SULT1A1 Arg213His polymorphism and susceptibility of environment-related cancers: A meta analysis of 5915 cases and 7900 controls. *Mol. Biol. Rep.* **2012**, *39*, 2597–2605. [CrossRef]

9.   Glatt, H. Sulfation and sulfotransferases 4: Bioactivation of mutagens via sulfation. *FASEB J.* **1997**, *11*, 314–321. [CrossRef]

10. Weinshilboum, R.M.; Otterness, D.M.; Aksoy, I.A.; Wood, T.C.; Her, C.; Raftogianis, R.B. Sulfation and sulfotransferases 1: Sulfotransferase molecular biology: cDNAs and genes. *FASEB J.* **1997**, *11*, 3–14. [CrossRef]

11. Liang, G.; Miao, X.; Zhou, Y.; Tan, W.; Lin, D. A functional polymorphism in the SULT1A1 gene (G638A) is associated with risk of lung cancer in relation to tobacco smoking. *Carcinogenesis* **2004**, *25*, 773–778. [CrossRef] [PubMed]

12. Chou, H.; Lang, N.; Kadlubar, F. Metabolic activation of N-hydroxyarylamines and N-hydroxy heterocyclic amines by human sulfotransferase (s). *Cancer Res.* **1995**, *55*, 525–529. [PubMed]

13. Yamazoe, Y.; Nagata, K.; Yoshinari, K.; Fujita, K.; Shiraga, T.; Iwasaki, K. Sulfotransferase catalyzing sulfation of heterocyclic amines. *Cancer Lett.* **1999**, *143*, 103–107. [CrossRef]

14. Banoglu, E. Current status of the cytosolic sulfotransferases in the metabolic activation of promutagens and procarcinogens. *Curr. Drug Metab.* **2000**, *1*, 1–30. [CrossRef] [PubMed]

15. Nimmagadda, D.; Cherala, G.; Ghatta, S. Cytosolic sulfotransferases. *Indian J. Exp. Biol.* **2006**, *44*, 171–182.

16. Wang, Z.; Fu, Y.; Tang, C.; Lu, S.; Chu, W.-m. SULT1A1 R213H polymorphism and breast cancer risk: A meta-analysis based on 8454 cases and 11,800 controls. *Breast Cancer Res. Treat.* **2010**, *122*, 193–198. [CrossRef] [PubMed]

17. Tasnim, T.; Al-Mamun, M.M.A.; Nahid, N.A.; Islam, M.R.; Apu, M.N.H.; Bushra, M.U.; Rabbi, S.N.I.; Nahar, Z.; Chowdhury, J.A.; Ahmed, M.U. Genetic variants of SULT1A1 and XRCC1 genes and risk of lung cancer in Bangladeshi population. *Tumor Biol.* **2017**, *39*, 1010428317729270. [CrossRef]

18. Forat-Yazdi, M.; Jafari, M.; Kargar, S.; Abolbaghaei, S.M.; Nasiri, R.; Farahnak, S.; Foroughi, E.; Neamatzadeh, H. Association between SULT1A1 Arg213His (Rs9282861) polymorphism and risk of breast cancer: A systematic review and meta-analysis. *J. Res. Health Sci.* **2017**, *17*, e00396.

19. Jiang, Y.; Zhou, L.; Yan, T.; Shen, Z.; Shao, Z.; Lu, J. Association of sulfotransferase SULT1A1 with breast cancer risk: A meta-analysis of case-control studies with subgroups of ethnic and menopausal statue. *J. Exp. Clin. Cancer Res.* **2010**, *29*, 101. [CrossRef]

20. Francis, A.M.; Ramya, R.; Ganesan, N.; Kumarasamy, P.; Paul, S.F.; Munirajan, A.; Divya, M. Breast cancer susceptibility genes in estrogen metabolizing pathway in a southern Indian population. *Meta Gene* **2019**, *19*, 225–234. [CrossRef]

21. Li, W.; Gu, M. SULT1A1 Arg213His polymorphism is associated with bladder cancer risk: A meta-analysis. *Med. Sci. Monit.* **2014**, *20*, 1590. [PubMed]

22. Boccia, S.; Persiani, R.; La Torre, G.; Rausei, S.; Arzani, D.; Gianfagna, F.; Romano-Spica, V.; D'Ugo, D.; Ricciardi, G. Sulfotransferase 1A1 polymorphism and gastric cancer risk: A pilot case-control study. *Cancer Lett.* **2005**, *229*, 235–243. [CrossRef] [PubMed]

23. Boccia, S.; Cadoni, G.; La Torre, G.; Arzani, D.; Volante, M.; Cattel, C.; Gianfagna, F.; Paludetti, G.; Almadori, G.; Ricciardi, G. A case–control study investigating the role of sulfotransferase 1A1 polymorphism in head and neck cancer. *J. Cancer Res. Clin. Oncol.* **2006**, *132*, 466. [CrossRef] [PubMed]

24. Wang, Y.; Spitz, M.R.; Tsou, A.M.-H.; Zhang, K.; Makan, N.; Wu, X. Sulfotransferase (SULT) 1A1 polymorphism as a predisposition factor for lung cancer: A case-control analysis. *Lung Cancer* **2002**, *35*, 137–142. [CrossRef]

25. Lopes, B.A.; Emerenciano, M.; Gonçalves, B.A.A.; Vieira, T.M.; Rossini, A.; Pombo-de-Oliveira, M.S. Polymorphisms in CYP1B1, CYP3A5, GSTT1, and SULT1A1 are associated with early age acute leukemia. *PLoS ONE* **2015**, *10*, e0127308. [CrossRef] [PubMed]

26. Nagata, K.; Yamazoe, Y. Pharmacogenetics of sulfotransferase. *Ann. Rev. Pharmacol. Toxicol.* **2000**, *40*, 159–176. [CrossRef]

27. Nagar, S.; Walther, S.; Blanchard, R.L. Sulfotransferase (SULT) 1A1 polymorphic variants* 1,* 2, and* 3 are associated with altered enzymatic activity, cellular phenotype, and protein degradation. *Mol. Pharmacol.* **2006**, *69*, 2084–2092. [CrossRef]

28. Bertolini, A.; Ferrari, A.; Ottani, A.; Guerzoni, S.; Tacchi, R.; Leone, S. Paracetamol: New vistas of an old drug. *CNS Drug Rev.* **2006**, *12*, 250–275. [CrossRef]

29. Rasool, M.I.; Bairam, A.F.; Gohal, S.A.; El Daibani, A.A.; Alherz, F.A.; Abunnaja, M.S.; Alatwi, E.S.; Kurogi, K.; Liu, M.-C. Effects of the human SULT1A1 polymorphisms on the sulfation of acetaminophen, O-desmethylnaproxen, and tapentadol. *Pharmacol. Rep.* **2019**, *71*, 257–265. [CrossRef]

30. Wang, L.-Q.; James, M.O. Inhibition of sulfotransferases by xenobiotics. *Curr. Drug Metab.* **2006**, *7*, 83–104. [CrossRef]

31. Gamage, N.U.; Duggleby, R.G.; Barnett, A.C.; Tresillian, M.; Latham, C.F.; Liyou, N.E.; McManus, M.E.; Martin, J.L. Structure of a human carcinogen-converting enzyme, SULT1A1 Structural and kinetic implications of substrate inhibition. *J. Biol. Chem.* **2003**, *278*, 7655–7662. [CrossRef] [PubMed]

32. Bidwell, L.M.; McManus, M.E.; Gaedigk, A.; Kakuta, Y.; Negishi, M.; Pedersen, L.; Martin, J.L. Crystal structure of human catecholamine sulfotransferase. *J. Mol. Boil.* **1999**, *293*, 521–530. [CrossRef] [PubMed]

33. Shevtsov, S.; Petrotchenko, E.V.; Pedersen, L.C.; Negishi, M. Crystallographic analysis of a hydroxylated polychlorinated biphenyl (OH-PCB) bound to the catalytic estrogen binding site of human estrogen sulfotransferase. *Environ. Health Perspect.* **2003**, *111*, 884–888. [CrossRef] [PubMed]

34. Chen, G.; Rabjohn, P.A.; York, J.L.; Wooldridge, C.; Zhang, D.; Falany, C.N.; Radominska-Pandya, A. Carboxyl residues in the active site of human phenol sulfotransferase (SULT1A1). *Biochemistry* **2000**, *39*, 16000–16007. [CrossRef] [PubMed]

35. Duggleby, R.G.; Gaedigk, A.; Mcmanus, M.E. Structural characterization of human aryl sulphotransferases. *Biochem. J.* **1999**, *337*, 337–343.

36. Dong, D.; Ako, R.; Wu, B. Crystal structures of human sulfotransferases: Insights into the mechanisms of action and substrate selectivity. *Expert Opin. Drug Metabol. Toxicol.* **2012**, *8*, 635–646. [CrossRef] [PubMed]

37. Thirumal, D.K.; Iyer, S.; Christy, J.P.; Siva, R.; Tayubi, I.A.; George, C.P.D.; Zayed, H. A comparative computational approach toward pharmacological chaperones (NN-DNJ and ambroxol) on N370S and L444P mutations causing Gaucher's disease. *Adv. Protein Chem. Struct. Biol.* **2019**, *114*, 315–339.

38. Hannan, M.; Dash, R.; Sohag, A.A.M.; Moon, I.S. Deciphering Molecular Mechanism of the Neuropharmacological Action of Fucosterol through Integrated System Pharmacology and InSilico Analysis. *Mar. Drugs* **2019**, *17*, 639. [CrossRef]

39. Thirumal, D.K.; Judith, E.; Priyadharshini, J.C.; Siva, R.; Tayubi, I.A.; Chakraborty, C.; George, C.P.D.; Zayed, H. Computational and modeling approaches to understand the impact of the Fabry's disease causing mutation (D92Y) on the interaction with pharmacological chaperone 1-deoxygalactonojirimycin (DGJ). *Adv. Protein Chem. Struct. Biol.* **2019**, *114*, 341–407.

40. Kumar, D.T.; Emerald, L.J.; Doss, C.G.P.; Sneha, P.; Siva, R.; Jebaraj, W.C.E.; Zayed, H. Computational approach to unravel the impact of missense mutations of proteins (D2HGDH and IDH2) causing D-2-hydroxyglutaric aciduria 2. *Metab. Brain Dis.* **2018**, *33*, 1699–1710. [CrossRef]

41. Tang, X.; Wang, Z.; Lei, T.; Zhou, W.; Chang, S.; Li, D. Importance of protein flexibility on molecular recognition: Modeling binding mechanisms of aminopyrazine inhibitors to Nek2. *Phys. Chem. Chem. Phys.* **2018**, *20*, 5591–5605. [CrossRef] [PubMed]

42. Li, Y.; Li, X.; Dong, Z. Concerted dynamic motions of an FABP4 model and its ligands revealed by microsecond molecular dynamics simulations. *Biochemistry* **2014**, *53*, 6409–6417. [CrossRef] [PubMed]

43. Hodsdon, M.E.; Cistola, D.P. Ligand binding alters the backbone mobility of intestinal fatty acid-binding protein as monitored by 15N NMR relaxation and 1H exchange. *Biochemistry* **1997**, *36*, 2278–2290. [CrossRef] [PubMed]

44. Lu, J.; Lin, C.L.; Tang, C.; Ponder, J.W.; Kao, J.L.; Cistola, D.P.; Li, E. Binding of retinol induces changes in rat cellular retinol-binding protein II conformation and backbone dynamics. *J. Mol. Boil.* **2000**, *300*, 619–632. [CrossRef] [PubMed]

45. Franzoni, L.; Lucke, C.; Perez, C.; Cavazzini, D.; Rademacher, M.; Ludwig, C.; Spisni, A.; Rossi, G.L.; Ruterjans, H. Structure and backbone dynamics of Apo- and holo-cellular retinol-binding protein in solution. *J. Boil. Chem.* **2002**, *277*, 21983–21997. [CrossRef] [PubMed]

46. Laine, J.M.; Amat, M.; Morgan, B.R.; Royer, W.E., Jr.; Massi, F. Insight into the allosteric mechanism of Scapharca dimeric hemoglobin. *Biochemistry* **2014**, *53*, 7199–7210. [CrossRef]

47. Nemaysh, V.; Luthra, P.M. Computational analysis revealing that K634 and T681 mutations modulate the 3D-structure of PDGFR-β and lead to sunitinib resistance. *RSC Adv.* **2017**, *7*, 37612–37626. [CrossRef]

48. Kamaraj, B.; Purohit, R. In silico screening and molecular dynamics simulation of disease-associated nsSNP in TYRP1 gene and its structural consequences in OCA3. *Biomed. Res. Int.* **2013**, *2013*, 697051. [CrossRef]

49. Kitao, A.; Hirata, F.; Gō, N. The effects of solvent on the conformation and the collective motions of protein: Normal mode analysis and molecular dynamics simulations of melittin in water and in vacuum. *Chem. Phys.* **1991**, *158*, 447–472. [CrossRef]

50. Amadei, A.; Linssen, A.B.; Berendsen, H.J. Essential dynamics of proteins. *Proteins Struct. Funct. Bioinform.* **1993**, *17*, 412–425. [CrossRef]

51. Kitao, A.; Go, N. Investigating protein dynamics in collective coordinate space. *Curr. Opin. Struct. Boil.* **1999**, *9*, 164–169. [CrossRef]

52. García, A.E. Large-amplitude nonlinear motions in proteins. *Phys. Rev. Lett.* **1992**, *68*, 2696. [CrossRef] [PubMed]

53. Maisuradze, G.G.; Liwo, A.; Scheraga, H.A. Principal component analysis for protein folding dynamics. *J. Mol. Boil.* **2009**, *385*, 312–329. [CrossRef] [PubMed]

54. Maisuradze, G.G.; Liwo, A.; Scheraga, H.A. Relation between free energy landscapes of proteins and dynamics. *J. Chem. Theory Comput.* **2010**, *6*, 583–595. [CrossRef] [PubMed]

55. Grottesi, A.; Domene, C.; Hall, B.; Sansom, M.S. Conformational dynamics of M2 helices in KirBac channels: Helix flexibility in relation to gating via molecular dynamics simulations. *Biochemistry* **2005**, *44*, 14586–14594. [CrossRef]

56. Yang, L.; Song, G.; Carriquiry, A.; Jernigan, R.L. Close correspondence between the motions from principal component analysis of multiple HIV-1 protease structures and elastic network modes. *Structure* **2008**, *16*, 321–330. [CrossRef]

57. Lou, H.; Cukier, R.I. Molecular dynamics of apo-adenylate kinase: A principal component analysis. *J. Phys. Chem. B* **2006**, *110*, 12796–12808. [CrossRef]

58. Spellmon, N.; Sun, X.; Sirinupong, N.; Edwards, B.; Li, C.; Yang, Z. Molecular dynamics simulation reveals correlated inter-lobe motion in protein lysine methyltransferase SMYD2. *PLoS ONE* **2015**, *10*, e0145758. [CrossRef]

59. Hosen, S.Z.; Dash, R.; Junaid, M.; Mitra, S.; Absar, N. Identification and structural characterization of deleterious non-synonymous single nucleotide polymorphisms in the human SKP2 gene. *Comput. Boil. Chem.* **2019**, *79*, 127–136. [CrossRef]

60. Agrahari, A.K.; George, P.D.C.; Siva, R.; Magesh, R.; Zayed, H. Molecular insights of the G2019S substitution in LRRK2 kinase domain associated with Parkinson's disease: A molecular dynamics simulation approach. *J. Theor. Boil.* **2019**, *469*, 163–171. [CrossRef]

61. Kim, H.J.; Choi, M.Y.; Kim, H.J.; Llinás, M. Conformational dynamics and ligand binding in the multi-domain protein PDC109. *PLoS ONE* **2010**, *5*, e9180. [CrossRef] [PubMed]

62. Dash, R.; Junaid, M.; Mitra, S.; Arifuzzaman, M.; Hosen, S.Z. Structure-based identification of potent VEGFR-2 inhibitors from in vivo metabolites of a herbal ingredient. *J. Mol. Model.* **2019**, *25*, 98. [CrossRef] [PubMed]

63. Raftogianis, R.B.; Wood, T.C.; Weinshilboum, R.M. Human phenol sulfotransferases SULT1A2 and SULT1A1: Genetic polymorphisms, allozyme properties, and human liver genotype–phenotype correlations. *Biochem. Pharmacol.* **1999**, *58*, 605–616. [CrossRef]

64. Raftogianis, R.B.; Wood, T.C.; Otterness, D.M.; Van Loon, J.A.; Weinshilboum, R.M. Phenol sulfotransferase pharmacogenetics in humans: Association of commonSULT1A1alleles with TS PST phenotype. *Biochem. Biophys. Res. Commun.* **1997**, *239*, 298–304. [CrossRef]

65. Ketterer, B.; Mulder, G.J. Glutathione conjugation. In *Conjugation Reactions in Drug Metabolism: An Integrated Approach*; Taylor & Francis London: London, UK, 1990; pp. 307–364.

66. Coughtrie, M. Sulfation through the looking glass—recent advances in sulfotransferase research for the curious. *Pharmacogn. J.* **2002**, *2*, 297. [CrossRef]

67. Strott, C.A. Sulfonation and molecular action. *Endocr. Rev.* **2002**, *23*, 703–732. [CrossRef]

68. Shimada, T. Xenobiotic-metabolizing enzymes involved in activation and detoxification of carcinogenic polycyclic aromatic hydrocarbons. *Drug Metab. Pharmacokinet.* **2006**, *21*, 257–276. [CrossRef]

69. Hebbring, S.J.; Adjei, A.A.; Baer, J.L.; Jenkins, G.D.; Zhang, J.; Cunningham, J.M.; Schaid, D.J.; Weinshilboum, R.M.; Thibodeau, S.N. Human SULT1A1 gene: Copy number differences and functional implications. *Hum. Mol. Genet.* **2006**, *16*, 463–470. [CrossRef]

70. Hildebrandt, M.; Carrington, D.; Thomae, B.; Eckloff, B.; Schaid, D.J.; Yee, V.; Weinshilboum, R.M.; Wieben, E.D. Genetic diversity and function in the human cytosolic sulfotransferases. *Pharmacogn. J.* **2007**, *7*, 133. [CrossRef]

71. Kormos, B.L.; Baranger, A.M.; Beveridge, D.L. Do collective atomic fluctuations account for cooperative effects? Molecular dynamics studies of the U1A− RNA complex. *J. Am. Chem. Soc.* **2006**, *128*, 8992–8993. [CrossRef]

72. Hempel, N.; Gamage, N.; Martin, J.L.; McManus, M.E. Human cytosolic sulfotransferase SULT1A1. *Int. J. Biochem. Cell Boil.* **2007**, *39*, 685–689. [CrossRef] [PubMed]

73. Ozawa, S.; Shimizu, M.; Katoh, T.; Miyajima, A.; Ohno, Y.; Matsumoto, Y.; Fukuoka, M.; Tang, Y.M.; Lang, N.P.; Kadlubar, F.F. Sulfating-activity and stability of cDNA-expressed allozymes of human phenol sulfotransferase, ST1A3*1 ((213)Arg) and ST1A3*2 ((213)His), both of which exist in Japanese as well as Caucasians. *J. Biochem.* **1999**, *126*, 271–277. [CrossRef] [PubMed]

74. Lu, L.-Y.; Chiang, H.-P.; Chen, W.-T.; Yang, Y.-S. Dimerization is responsible for the structural stability of human sulfotransferase 1A1. *Drug Metab. Dispos.* **2009**, *37*, 1083–1088. [CrossRef] [PubMed]

75. Rose, P.W.; Beran, B.; Bi, C.; Bluhm, W.F.; Dimitropoulos, D.; Goodsell, D.S.; Prlić, A.; Quesada, M.; Quinn, G.B.; Westbrook, J.D. The RCSB Protein Data Bank: Redesigned web site and web services. *Nucleic Acids Res.* **2010**, *39*, D392–D401. [CrossRef] [PubMed]

76. Dash, R.; Junaid, M.; Islam, N.; Akash, C.; Forhad, M.; Khan, M.; Arifuzzaman, M.; Khatun, M.; Zahid Hosen, S.M. Molecular insight and binding pattern analysis of Shikonin as a potential VEGFR-2 inhibitor. *Curr. Enzyme Inhib.* **2017**, *13*, 235–244. [CrossRef]

77. Krieger, E.; Darden, T.; Nabuurs, S.B.; Finkelstein, A.; Vriend, G. Making optimal use of empirical energy functions: Force-field parameterization in crystal space. *Proteins Struct. Funct. Bioinform.* **2004**, *57*, 678–683. [CrossRef]

78. Land, H.; Humble, M.S. YASARA: A tool to obtain structural guidance in biocatalytic investigations. In *Protein Engineering*; Springer: Berlin, Germany, 2018; pp. 43–67.

79. Mitra, S.; Dash, R. Structural dynamics and quantum mechanical aspects of shikonin derivatives as CREBBP bromodomain inhibitors. *J. Mol. Graphics Modell.* **2018**, *83*, 42–52. [CrossRef]

80. Dash, R.; Das, R.; Junaid, M.; Akash, M.F.C.; Islam, A.; Hosen, S.Z. In silico-based vaccine design against Ebola virus glycoprotein. *Adv.Appl. Bioinform. Chem. AABC* **2017**, *10*, 11. [CrossRef]

81. Wang, J.; Wolf, R.M.; Caldwell, J.W.; Kollman, P.A.; Case, D.A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157–1174. [CrossRef]

82. Dickson, C.J.; Madej, B.D.; Skjevik, Å.A.; Betz, R.M.; Teigen, K.; Gould, I.R.; Walker, R.C. Lipid14: The amber lipid force field. *J. Chem. Theory Comput.* **2014**, *10*, 865–879. [CrossRef]

83. Stewart, J.J. MOPAC: A semiempirical molecular orbital program. *J. Comput. Aided Mol. Des.* **1990**, *4*, 1–103. [CrossRef] [PubMed]

84. Jakalian, A.; Jack, D.B.; Bayly, C.I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.* **2002**, *23*, 1623–1641. [CrossRef] [PubMed]

85. Harrach, M.F.; Drossel, B. Structure and dynamics of TIP3P, TIP4P, and TIP5P water near smooth and atomistic walls of different hydroaffinity. *J. Chem. Phys.* **2014**, *140*, 174501. [CrossRef] [PubMed]

86. Shaik, M.S.; Liem, S.Y.; Popelier, P.L. Properties of liquid water from a systematic refinement of a high-rank multipolar electrostatic potential. *J. Chem. Phys.* **2010**, *132*, 174504. [CrossRef] [PubMed]

87. Mark, P.; Nilsson, L. Structure and Dynamics of the TIP3P, SPC, and SPC/E Water Models at 298 K. *J. Phys. Chem. A* **2001**, *105*, 9954–9960. [CrossRef]

88. Florová, P.; Sklenovský, P.; Banáš, P.; Otyepka, M. Explicit water models affect the specific solvation and dynamics of unfolded peptides while the conformational behavior and flexibility of folded peptides remain intact. *J. Chem. Theory Comput.* **2010**, *6*, 3569–3579. [CrossRef]

89. Fennell, C.J.; Kehoe, C.W.; Dill, K.A. Modeling aqueous solvation with semi-explicit assembly. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 3234–3239. [CrossRef]

90. Krieger, E.; Dunbrack, R.L.; Hooft, R.W.; Krieger, B. Assignment of protonation states in proteins and ligands: Combining pK a prediction with hydrogen bonding network optimization. In *Computational Drug Discovery and Design*; Springer: Berlin, Germany, 2012; pp. 405–421.

91. Krieger, E.; Nielsen, J.E.; Spronk, C.A.; Vriend, G. Fast empirical pKa prediction by Ewald summation. *J. Mol. Gr. Model.* **2006**, *25*, 481–486. [CrossRef]

92. Krieger, E.; Vriend, G. New ways to boost molecular dynamics simulations. *J. Comput. Chem.* **2015**, *36*, 996–1007. [CrossRef]

93. Essmann, U.; Perera, L.; Berkowitz, M.L.; Darden, T.; Lee, H.; Pedersen, L.G. A smooth particle mesh Ewald method. *J. Chem. Phys.* **1995**, *103*, 8577–8593. [CrossRef]

94. Stukowski, A. Visualization and analysis of atomistic simulation data with OVITO–the Open Visualization Tool. *Modell. Simul. Mater. Sci. Eng.* **2009**, *18*, 015012. [CrossRef]

95. Frishman, D.; Argos, P. Knowledge-based protein secondary structure assignment. *Proteins Struct. Funct. Bioinform.* **1995**, *23*, 566–579. [CrossRef] [PubMed]

96. Kabsch, W.; Sander, C. DSSP: Definition of secondary structure of proteins given a set of 3D coordinates. *Biopolymers* **1983**, *22*, 2577–2637. [CrossRef] [PubMed]

97. Blatt, J.M.; Weisskopf, V.F. *Theoretical Nuclear Physics*; Courier Corporation: North Chelmsford, MA, USA, 1991.

98. Lovering, A.L.; Lee, S.S.; Kim, Y.-W.; Withers, S.G.; Strynadka, N.C. Mechanistic and structural analysis of a family 31 a-glycosidase and its glycosyl-enzyme intermediate. *J. Biol. Chem.* **2004**, *280*, 2105–2115. [CrossRef] [PubMed]

99. Srinivasan, E.; Rajasekaran, R. Computational investigation of curcumin, a natural polyphenol that inhibits the destabilization and the aggregation of human SOD1 mutant (Ala4Val). *RSC Adv.* **2016**, *6*, 102744–102753. [CrossRef]

100. Swanson, J.M.; Henchman, R.H.; McCammon, J.A. Revisiting free energy calculations: A theoretical connection to MM/PBSA and direct calculation of the association free energy. *Biophys. J.* **2004**, *86*, 67–74. [CrossRef]

101. Grant, B.J.; Rodrigues, A.P.; ElSawy, K.M.; McCammon, J.A.; Caves, L.S. Bio3d: An R package for the comparative analysis of protein structures. *Bioinformatics* **2006**, *22*, 2695–2696. [CrossRef]

102. Ichiye, T.; Karplus, M. Collective motions in proteins: A covariance analysis of atomic fluctuations in molecular dynamics and normal mode simulations. *Proteins Struct. Funct. Bioinform.* **1991**, *11*, 205–217. [CrossRef]

103. Tripathi, S.; Srivastava, G.; Sharma, A. Molecular dynamics simulation and free energy landscape methods in probing L215H, L217R and L225M βI-tubulin mutations causing paclitaxel resistance in cancer cells. *Biochem. Biophys. Res. Commun.* **2016**, *476*, 273–279. [CrossRef]

104. Zhang, K.; Zhang, Y.-Q.; Ai, W.-B.; Hu, Q.-T.; Zhang, Q.-J.; Wan, L.-Y.; Wang, X.-L.; Liu, C.-B.; Wu, J.-F. Hes1, an important gene for activation of hepatic stellate cells, is regulated by Notch1 and TGF-β/BMP signaling. *World J. Gastroenterol. WJG* **2015**, *21*, 878. [CrossRef]

105. Shlens, J. A tutorial on principal component analysis. *arXiv* **2014**, arXiv:1404.1100.

106. Salmas, R.E.; Yurtsever, M.; Durdagi, S. Investigation of inhibition mechanism of chemokine receptor CCR5 by micro-second molecular dynamics simulations. *Sci. Rep.* **2015**, *5*. [CrossRef] [PubMed]

107. Frauenfelder, H.; Sligar, S.G.; Wolynes, P.G. The energy landscapes and motions of proteins. *Science* **1991**, *254*, 1598–1603. [CrossRef] [PubMed]

108. Lange, O.F.; Grubmüller, H. Generalized correlation for biomolecular dynamics. *Proteins Struct. Funct. Bioinform.* **2006**, *62*, 1053–1061. [CrossRef]

109. Kumari, R.; Kumar, R.; Consortium, O.S.D.D.; Lynn, A. g_mmpbsa A GROMACS tool for high-throughput MM-PBSA calculations. *J. Chem. Inf. Model.* **2014**, *54*, 1951–1962. [CrossRef]

110. Maier, J.A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K.E.; Simmerling, C. ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **2015**, *11*, 3696–3713. [CrossRef]

111. Kollman, P.A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W. Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models. *Acc. Chem. Res.* **2000**, *33*, 889–897. [CrossRef]

112. Leonis, G.; Steinbrecher, T.; Papadopoulos, M.G. A contribution to the drug resistance mechanism of Darunavir, Amprenavir, Indinavir, and Saquinavir complexes with HIV-1 protease due to flap mutation I50V: A systematic MM–PBSA and thermodynamic integration study. *J. Chem. Inf. Model.* **2013**, *53*, 2141–2153. [CrossRef]

113. Chen, J.; Wang, X.; Zhu, T.; Zhang, Q.; Zhang, J.Z. A comparative insight into amprenavir resistance of mutations V32I, G48V, I50V, I54V, and I84V in HIV-1 protease based on thermodynamic integration and MM-PBSA methods. *J. Chem. Inf. Model.* **2015**, *55*, 1903–1913. [CrossRef]

114. Musyoka, T.M.; Kanzi, A.M.; Lobb, K.A.; Bishop, Ö.T. Structure based docking and molecular dynamic studies of plasmodial cysteine proteases against a South African natural compound and its analogs. *Sci. Rep.* **2016**, *6*, 23690. [CrossRef]

115. Moonrin, N.; Songtawee, N.; Rattanabunyong, S.; Chunsrivirot, S.; Mokmak, W.; Tongsima, S.; Choowongkomon, K. Understanding the molecular basis of EGFR kinase domain/MIG-6 peptide recognition complex using computational analyses. *BMC Bioinform.* **2015**, *16*, 103. [CrossRef] [PubMed]