CHAPTER 4

Salient structural features of Human Lemur Tyrosine Kinase-3 (LMTK3) domain from Molecular dynamics simulation study

Salient structural features of Human Lemur Tyrosine Kinase-3 (LMTK3) domain from Molecular dynamics simulation study

4.1 Abstract

Estrogen receptor-α (ERα) positive breast cancer is considered to be one of the most common metastatic diseases. Estrogenic signalling in breast cancer is one of the most critical oncogenic pathways. Recently LMTK3 was identified as a potential oncogenic ERα regulator with a significant role in endocrine resistance. Therefore, targeting LMTK3 in breast cancer would control ERα modulation and may provide a better diagnostic development and a new therapeutic target to fight these resistant and aggressive tumours. Our objective is to understand the salient structural features of LMTK3 using molecular dynamics simulation. In this computational study, we modelled 3-D structure of LMTK3 domain using I-TASSER and studied conformational dynamics using molecular dynamics simulation. We used online computational tools and software to perform comprehensive investigation on the cavities, hydrophobicity, electrostatic potential, secondary structure topology and intra-molecular interactions in LMTK3. The LMTK3 structure was observed to be stable during MD simulation. We also predicted the probable binding cavities in LMTK3. In addition, we determined hydrophobic clusters and patches in LMTK3 which may be crucial for folding and stabilisation. Possible interaction sites in LMTK3 were then studied from electrostatic potential analysis based on positive and negative surfaces. From the secondary structure topology analysis, we noticed nine antiparallel β-sheets forming β-sheets topology and five hairpins were involved in forming the secondary structure. Our inferences from this study would be helpful in understanding the structure–function relationships of LMTK3. And from LMTK3-ATP interaction we determine key residues of the ATP binding site that may be helpful in designing potential ATP competitive inhibitors against LMTK3.

4.2. Introduction

Breast cancer is the commonest malignant dreadful disease that mainly affects women which leads to death [284]. Estrogen receptors-α (ERα) found to be expressed in more than 70 % of breast cancers cases [90]. Estrogen is a female hormone which is responsible for the development of breast cancer and its growth [285]. In the recent studies the estrogen signalling pathway has been highlighted to be the key target for the ERα positive breast cancer [8]. The current therapy to treat breast cancer includes anti-estrogen (tamoxifen), estrogen withdrawal (aromatase inhibitor) and direct targeting of the receptor (fulvestrant) [286]. But resistance to these therapies are common and ERα phosphorylation is mainly involved in endocrine resistance [287] by regulating the transcriptional activity and altering the stability of ERα [288, 289]. The prevalence of endocrine resistance in ERα positive breast cancer patients leads to increased levels of ERα, wherein protein kinase enzymes have been found to be the key target [290].

Protein kinase enzymes share a conserved catalytic core that generally consists of 250-300 amino acids [291] common with both serine/threonine and tyrosine protein kinase. The protein kinase domain has been classified into eleven conserved sub-domains that are folded into a small N-terminal lobe and a large C-terminal lobe. N-terminal lobe favours ATP binding which is glycine rich loop (P-loop) and the C-terminal lobe regulates the phosphorylation activity which is an activation loop (T-loop) [292].

LMTK3, a serine-threonine-tyrosine kinase [293-295] has been observed to be over expressed in more aggressive form of breast cancers and its level has been correlated with disease free survival [6,7]. In the recent past, LMTK3 was reported to regulate ERα activity through phosphorylation and protects ERα from proteasomal degradation and also triggers endocrine resistance [5] which in turn aids the breast cancer cells to undergo proliferation [85]. In addition it has also been reported that cytoplasmic elevation of LMTK3 in triple-negative breast cancer promotes breast cancer cell motility, migration and invasion through transcriptional activation of intigrins [9]. It is evident from these studies that LMTK3 to be a new therapeutic target. So research should be focussed on the structural characteristics of LMTK3 to overcome breast cancer metastasis and endocrine resistance [109]. Since the crystal structure of LMTK3

is not available, the breast cancer therapeutics research remains elusive. Thereby, in the literature, we see only a few structural studies carried out on LMTK3 [81].

In our computational study, an attempt has been made to study the salient structural features of human LMTK3 domain. Since the 3-D structure of LMTK3 is not available in protein data bank (PDB), we have modelled the 3-D structure and validated. In the present study we performed all-atom molecular dynamics (MD) simulations to explore the structural features of modelled LMTK3 domain. From the MD simulation we focused on: stability, flexibility, transition, intramolecular interaction (hydrogen bonds and hydrophobic contacts) of LMTK3. We conducted a rapid structural bioinformatics analysis of LMTK3 to obtain a better model of secondary structure topology mapping, interactions between residues, hydrophobic patches, electrostatic potential, and probable binding pockets of LMTK3. As a whole, our findings shed light on some of the significant insights into the structural features of LMTK3 domain which will enhance our understanding of the structure–function relationships, and may assist in further development of effective drugs for therapeutic targeting of LMTK3.

4.3. Materials and Methods

4.3.1. Structural modelling and validation of LMTK3 domain

We retrieved the amino acid sequence of LMTK3 kinase domain (279 amino acids, 133-411) from the UniprotKB database [296] (accession no: Q96Q04). As the 3-D structure of LMTK3 domain is not available, we modelled the 3-D structure from I-TASSER server [99] by submitting the amino acid (133-411 res) sequence. According to the CASP ranking, I-TASSER has been reported to be the best automated protein prediction server. We obtained five models. The best model was identified based on the C-score calculated from the relative clustering structural density and consensus significance. Template modelling (TM) score and root mean square deviation (RMSD) are used to evaluate the accuracy of the best model. Then Pfam domain analysis was carried out to confirm the kinase domain of LMTK3. And also to reassert our prediction we superimposed the LMTK3 model structure with a known protein kinase (PDK1, PDB ID: 3HRF) which is found to be implicated in the development and progression of

melanomas [297-299]. Then the best modelled structure of $LMTK3_{133-411}$ domain was validated by the ProSA server [300], verify 3-D [301,302], and the RAMPAGE server [303].

4.3.2. Sequence alignment of LMTK3 domain with other kinases

To determine important conserved regions in LMTK3 domain we carried out multiple sequence alignment of various kinases and compared with LMTK3 domain, using Clustal Omega online multiple sequence alignment tool [304]. In this tool seeded guide trees and HMM profile-profile techniques are used to get the sequence alignment [304].

4.3.3. Molecular Dynamics (MD) simulation of LMTK3 domain

The modelled structure of LMTK3 domain was subjected to MD simulation. The Particle Mesh Ewald Molecular Dynamics (PMEMD) [305] module of AMBER12 software package was used [122]. AMBER ff99SB force field [306] protein parameters were used for LMTK3. Using implicit solvation method topology and co-ordinate files were prepared in xleap module of AMBER package. The resultant structure was then pursued to energy minimization by using 500 steps of steepest descent and another 500 steps of conjugate gradient. The minimized structure was then subjected to 100 ps of MD using 2 fs time step for integration. During the heating dynamics, the system was gradually heated from 0 to 300 K and ensured slow relaxation of the built initial structure. In addition SHAKE constraints [165] using a geometrical tolerance of 5×10^{-4} Å was imposed on all the covalent bonds involving hydrogen atoms. Then MD was performed under constant pressure-temperature conditions (NPT) with temperature regulation achieved using Berendsen weak coupling method [166] (0.5 ps time constant for heat bath coupling and 0.2 ps pressure relaxation time). This was followed by another 40 ps of equilibration step. Finally for the analysis of structures and properties we carried out long NPT MD runs up to 25 ns using a heat bath coupling time constant of 1 ps.

4.3.4. Analysis

From MD simulation trajectory, we analyzed the Root Mean Square Deviation (RMSD) of the protein (LMTK3 domain) backbone atoms from the initial structure as

reference was used as a measure of structural differences. In order to check the compactness of LMTK3 the gyration radius of protein backbone atoms was also calculated. The B-factor and Root Mean Square Fluctuation (RMSF) of Cα atoms was calculated to display of structural flexibility. The solvent-accessible surface area (SASA) and Molecular Accessible Surface Area was calculated with a 1.4 Å probe radius from Surface racer software [307]. The secondary structure analysis was conducted using the Kabsch and Sander algorithm incorporated in the dictionary of secondary structure for protein (DSSP) [178] program during the course of MD simulation.

We also studied characteristics features of LMTK3 using standard computational tools that have been used in assessment of Molecular Dynamics and interaction properties of C-reactive protein [308]. To predict the pictographic view of the secondary structure, we used PDB-sum online server [250,252] which summarizes the secondary structural elements of LMTK3. From PDB-sum we depicted topology of LMTK3 and studied the effect and connectivity of the helices and strands. We have also identified and shown the significant hydrogen bonds stabilizing the LMTK3 structure. Further to determine hydrophobicity in LMTK3 we have used ProtScale online server, wherein amino acid propensity scale is utilised to get the hydrophobicity plot [309]. We then used ExPASy and Swiss Pdb viewer to perform hydrophobic cluster analysis (HCA) [310] and deciphered the hydrophobic patches respectively.

Using chimera [180], molecular visualization tool, we plotted electrostatic potential map for the equilibrated structure of LMTK3 with the option of coulombic surface colouring. Coulomb's law is valid for charges immersed in an infinite medium of a uniform dielectric [311]. While plotting the electrostatic potential map we used the coulombs law with dielectric constant 4 Å and distance from surface (off set) to be 1.4 Å. We also identified probable pockets in LMTK3 by using pocket detection tools such as LIGSITE^{csc} [312] and Surface racer [307]. In Surface racer the probe radius was set to 1.4 Å and in LIGSITE^{csc} server the grid space and radius of the probe was set to 3 Å and 1.5 Å.

4.3.5 LMTK3-ATP binding mechanism

We studied the ATP binding mechanism on LMTK3 using Autodock4.2 [211]. From Autodock tools, PDBQT files were generated for the ligand (ATP) and the macromolecule (LMTK3) for docking. Grid spacing was set in AutoGrid by selecting two important residues: Lys164 and Asp166 (Uniprot as reference) with a grid box size x = 60 Å, y = 60 = Å and z = 60 Å and grid spacing of 0.375 Å, provide search space and grid center was selected at dimensions x= -0.056, y= 0.028, z= -0.083used to dock ATP with the equilibrated structure of LMTK3 domain. AutoDock was used to dock ATP with LMTK3. To perform the protein-ligand docking, Lamarckian Genetic Algorithm (LGA) was used with default parameters. The best docking conformation having strong binding affinity value in kcal mol-1 was selected, and followed by the analysis of the binding site based on hydrogen bonding interaction and the hydrophobic interaction between ATP and LMTK3 using LigPlot+ [254]. The best docking complex of LMTK3-ATP was selected and used as an input for molecular dynamics simulations

4.4. Results and Discussion

4.4.1. Structural validation of LMTK3 domain

We modelled the structure of LMTK3 domain I-TASSER**.** The server generated five models using threading programs. The best model was identified based on the C-score (0.87) calculated from the comparative clustering structural density and consensus significance. The overall quality of model was determined by parameters like TM score of 0.83 ± 0.08 and RMSD value 4.3 ± 2.8 Å. Among the top ten templates proposed by I-TASSER server, we selected the protein (Insulin receptor, InsR: ID: 1irk) having highest identity percent (36%) with the LMTK3 domain. **Figure 4.1** shows the generated model of LMTK3 kinase with different key parts and **Figure 4.2** visualized the superposition between the modelled LMTK3 kinase domain and the resolved 3-D structure of insulin receptor. The RMSD of this superposition was 0.3 Å which confirms the high identity between this two proteins.

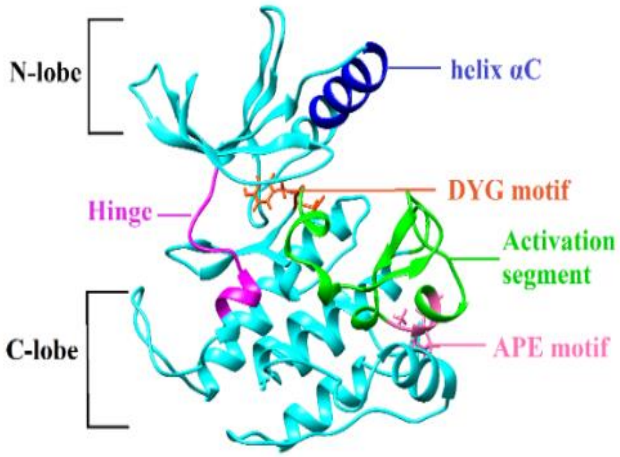

Figure 4.1. *3-D structure of LMTK3 domain obtained from I-TASSER*

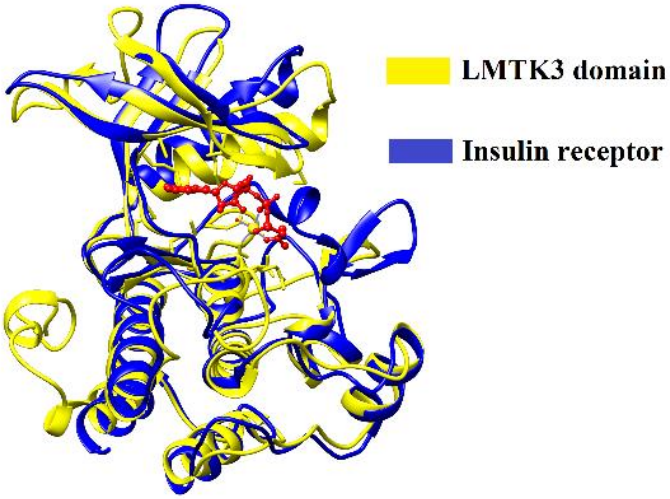

Figure 4.2. *Representation of superimposed LMTK3 domain model with Insulin receptor (InsR)*

Then we carried out structural validation for the model structures by constructing Ramachandran plot using RAMPAGE server. In the modeled structure of LMTK3, we noticed 87% of residues are in most favored region (243 residues), and 5.1% in disallowed region **(10 residues)** shown in **Figure 4.3** The overall quality and protein folding energy of the model structure was validated using ProSA server with the quality index represented by Z-score of -7.2 as shown in **Figure 4.4**. From the Pfam domain analysis we have confirmed the characteristic signature of protein kinase domain.

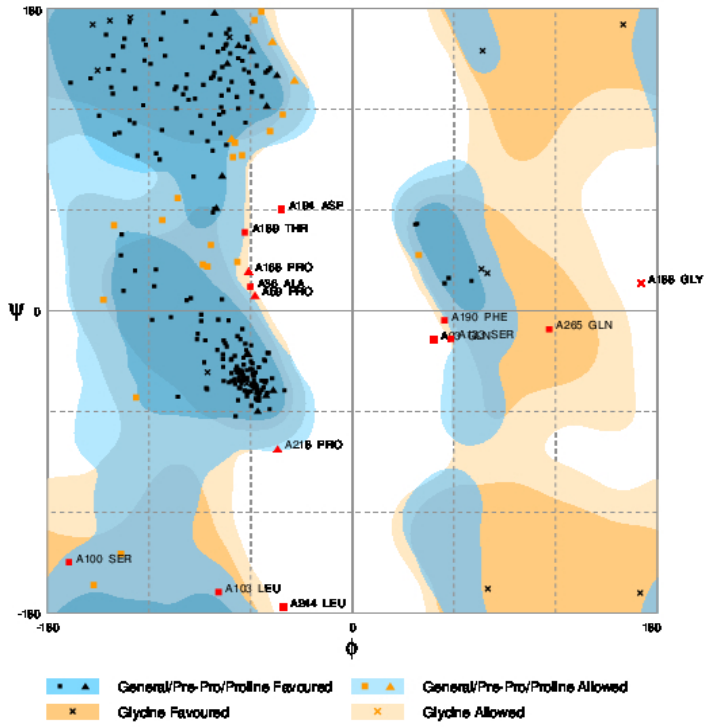

Figure 4.3. *Ramachandran plot of LMTK3 domain: 87% of residues are in most favored region (243 residues) and 5.1% in disallowed region*

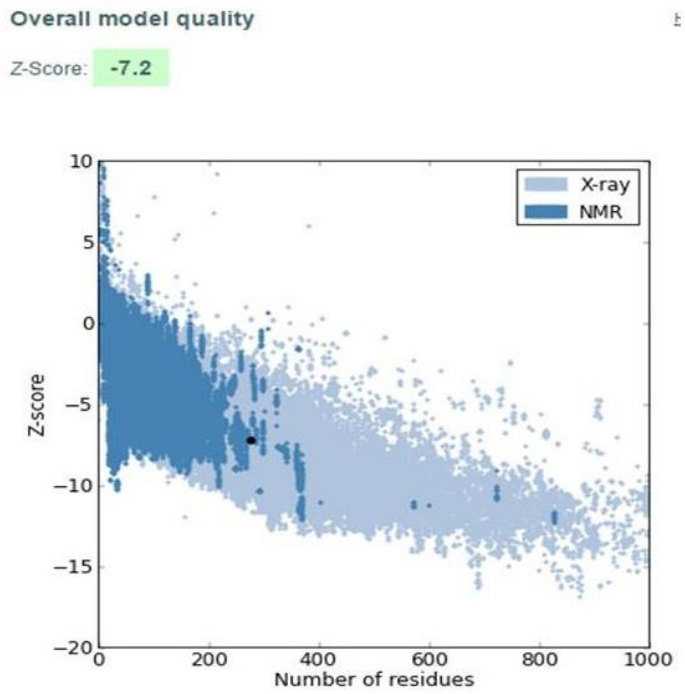

Figure 4.4. *Quality index for the model structure of LMTK3 domain by ProSA server. The index value was represented by Z-score*

Verify-3D server ascertains the accuracy of the predicted secondary structure (3D) model with its respective residues (1D) by assigning a structural class based on its location and environment. 77.78 % of the residues had an averages 3D-1D score >= 0.2 which confirms the accuracy of the predicted model to near correct to the actual structure. The verify 3D plot is shown in **Figure 4.5**

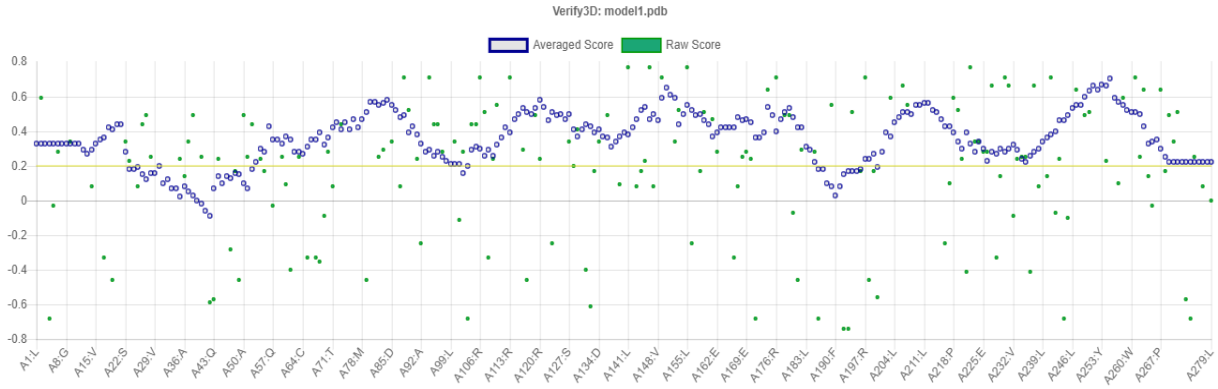

Figure 4.5. *Verify 3D plot of the modelled structure of LMTK3 domain*

We also superimposed LMTK3 model structure with a known protein kinase PDK1 (PDB ID: 3HRF) [297,298] which belong to AGC kinase group **(Figure 4.6)**. This also confirmed that the LMTK3 possesses almost the same characteristic architecture as PDK1 [299] protein kinase. From **Figure 4.7**, we notice that LMTK3 contains one P-loop, αC helix and an activation loop, which we can also see in PDK1.

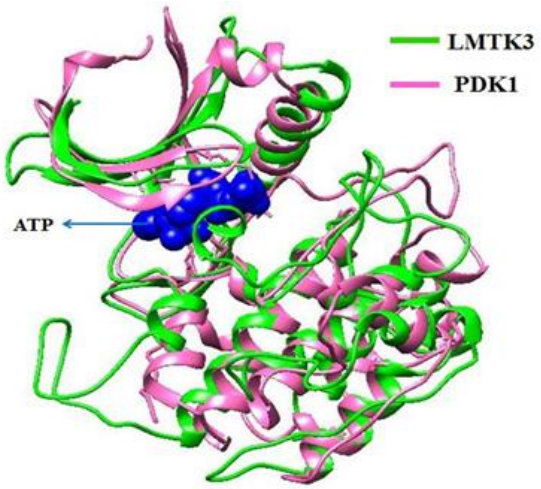

Figure 4.6 *Superimposed structure of modelled LMTK3 domain with PDK1 protein*

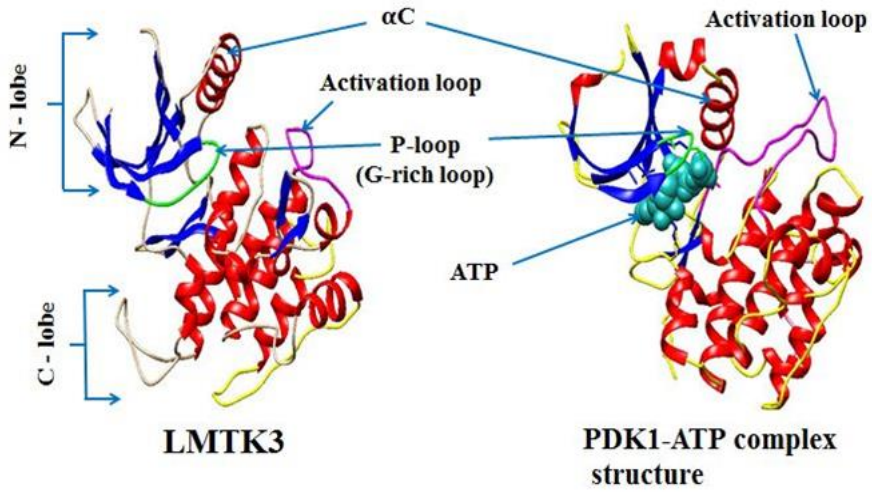

Figure 4.7. *Architecture of LMTK3 domain and PDK1 (pdb ID:3HRF) in cartoon representation. The three panels use the following color scheme: αC helix (red), P-loop: glycine rich (green), activation loop (magenta).*

4.4.2. Identification of conserved regions in LMTK3 domain using multiple sequence alignment

From the **Figure 4.8** we can see some residues in all the kinases kinases (ERK2, PKA, Aurora A, PDK1) including LMTK3 domain are conserved. Most importantly the yellow highlighted regions are important for the functional and catalytic activity of kinase. The residue 'D' (Asp) in the third row in the sequence is mentioned as active site (proton acceptor) among all the including LMTK3 domain based on PROSITE-ProRule annotation. Generally the activation loop of eukaryotic protein kinases start with DFG motif and end at APE/SPE motif. From the **Figure 4.8** we see that, LMTK3 also contain conserved APE motif as that of other kinases but instead of DFG motif, LMTK3 has DYG sequence.

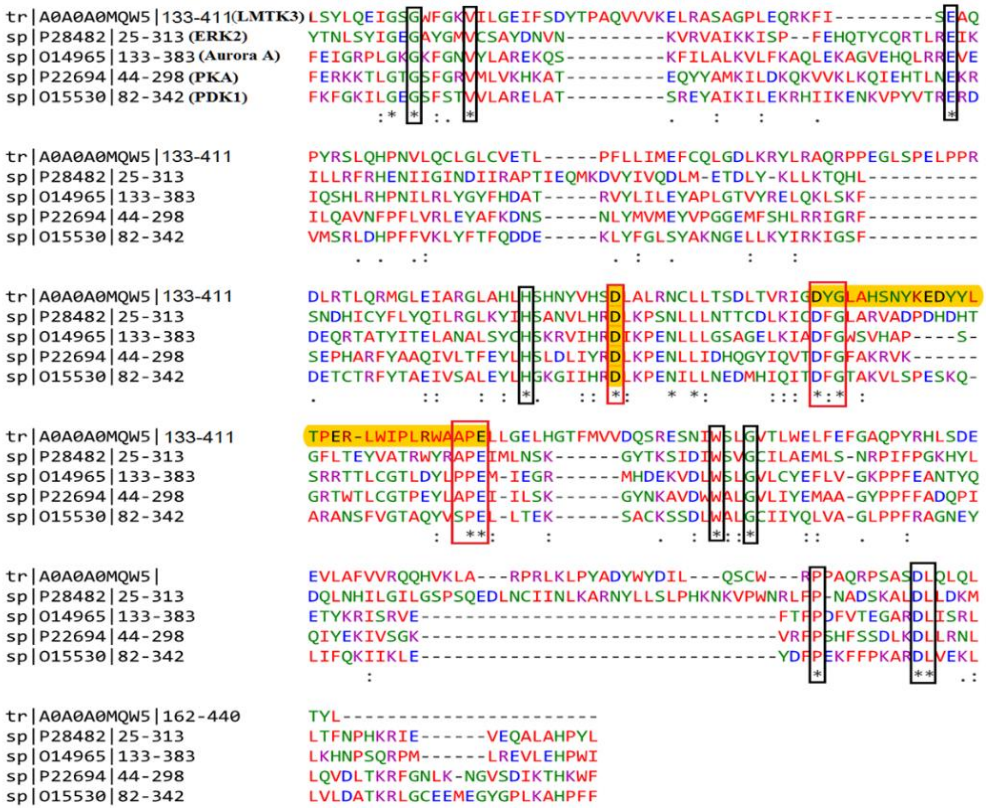

Figure 4.8. *Multiple sequence alignment of LMTK3 domain with different kinases*

Protein kinases exhibit active and inactive conformation and activated protein kinases able to phosphorylate its substrates [60]. In order to determine the conformation of LMTK3 domain we carried out structural comparison of LMTK3 domain with different active and inactive kinases (**Figure 4.9 & 4.10)**. From **Figure 4.9** we see in case of the active kinases, the activation loop protrude outside and can move freely, which may allow the phosphoryl transfer reaction [69-71]. We can see, as that of other active kinase the activation loop of LMTK3 is also protrude out and could move freely. Structural comparison analysis revealed the conformation of catalytic domain of LMTK3 to be active and DYG-in conformation.

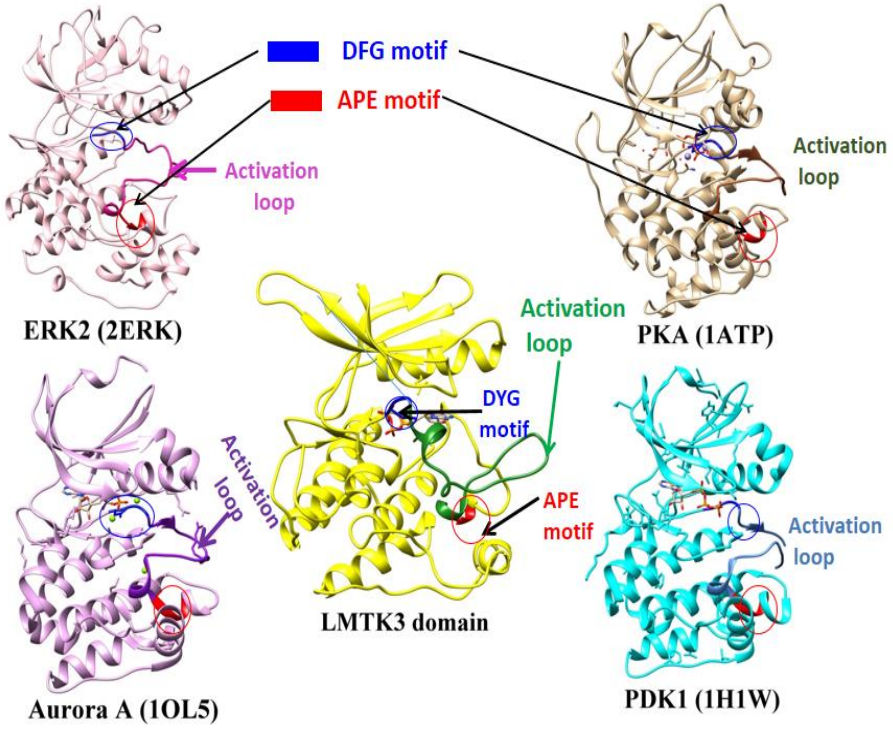

Figure 4.9. *Structural comparison of LMTK3 domain with different active kinases*

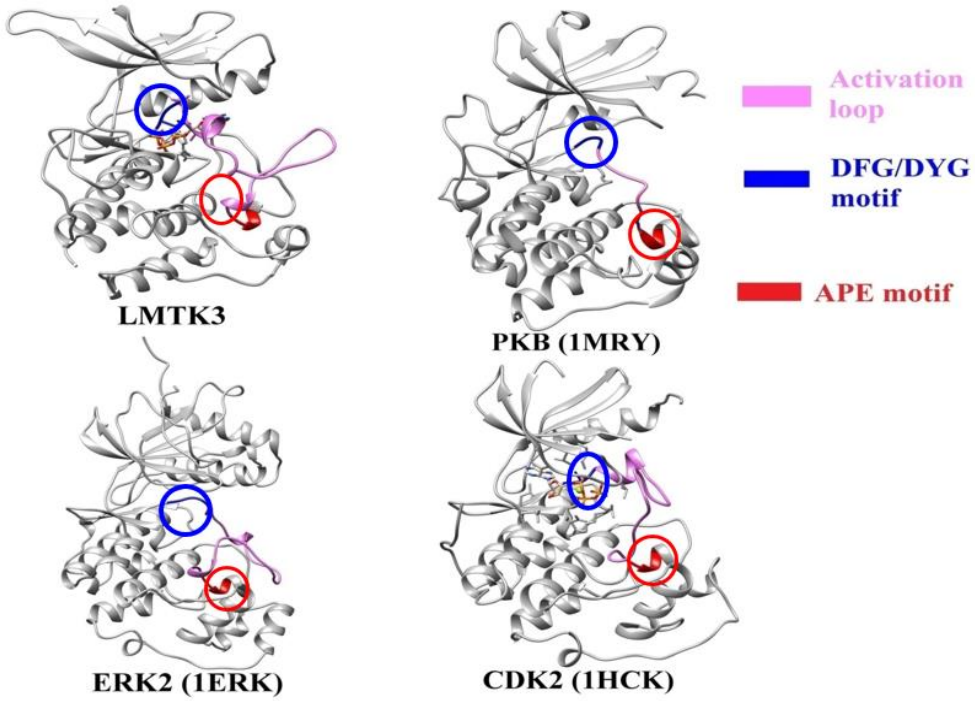

Figure 4.10. *Structural comparison of LMTK3 domain with different inactive kinases*

4.4.3. Structural characteristics features of LMTK3 domain

4.4.3.1. RMSD analysis: We have analyzed the conformational changes in LMTK3 during the time course of MD simulation. The equilibrated structure of LMTK3 was used as a reference structure for the RMSD analysis. The RMSD analysis was done based on C- α atoms showed initial equilibration up to 4 ns, following which the structure started converging after 5 ns and attained the stable conformation after 15 ns with RMSD near 4 Å in the production MD run as depicted in **Figure 4.11**. The RMSD plot thus confirmed the stability of LMTK3 with less deviation from the initial structure.

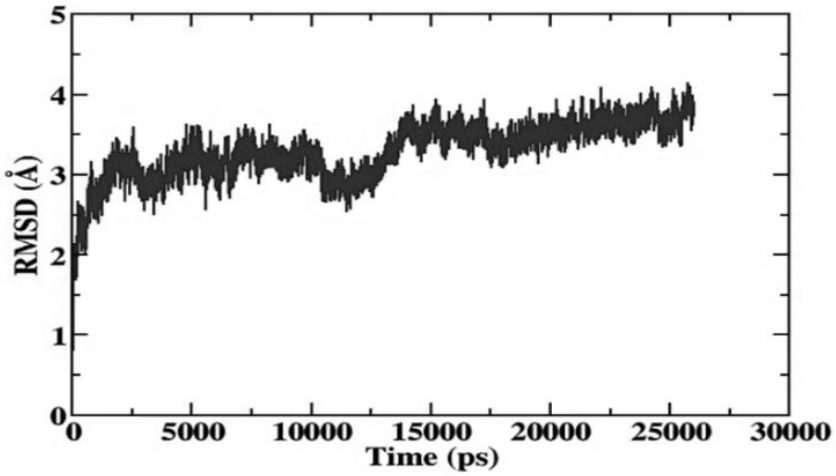

Figure 4.11. *RMSD plot of LMTK3 for Cα atoms as a function of time*

4.4.3.2 Radius of gyration: In order to check the compactness of LMTK3 protein we performed radius of gyration (Rg) analysis **Figure 4.12.** In **Figure 4.12** we can see that the Rg value for LMTK3 showed a drift between 19.3 Å and 18.3 Å and reached a stable conformation after 18.5 Å at the end of the simulation. From the plot we can infer that there is not much degree of compactness during the course of MD simulation, thereby confirming the folding of the protein structure.

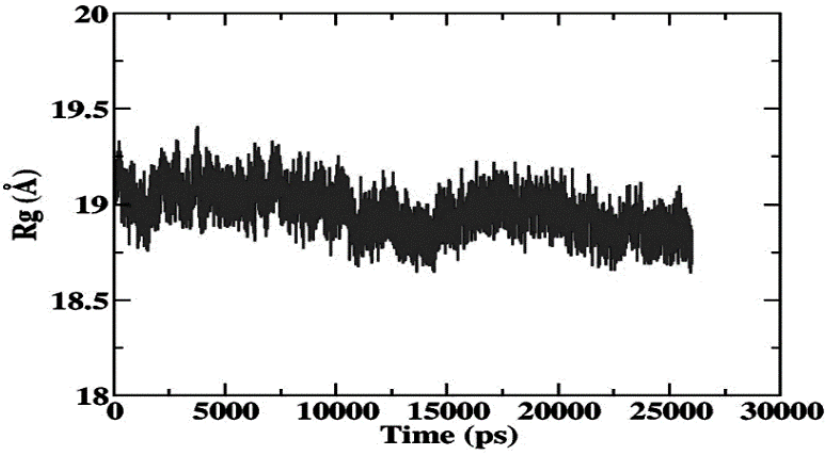

Figure 4.12. *Rg plot of LMTK3 as a function of time*

4.4.3.3 B-factor analysis and RMSF: In order to check the flexibility in LMTK3 we analyse B factor value and Root mean square fluctuation (RMSF) for the C-α atoms depicted in **Figure 4.13 (A)** and **(B)**. Both the plots showed that fluctuation prevailing in the loop region of protein with two high peaks around 750 $Å^2$ (residue index 19 to 27) and 1250 $Å^2$ (residue index 93 to 107) in B-factor plot **Figure 4.13 (A)** and two high peaks 3.8 $Å^2$ (residue index 18 to 29) and 5 $Å^2$ (residue index 93 to 107) in RMSF plot respectively shown in **Figure 4.13 (B)**. Both the plots confirmed that in LMTK3 the fluctuations are less in standard secondary structure regions than in the loop regions.

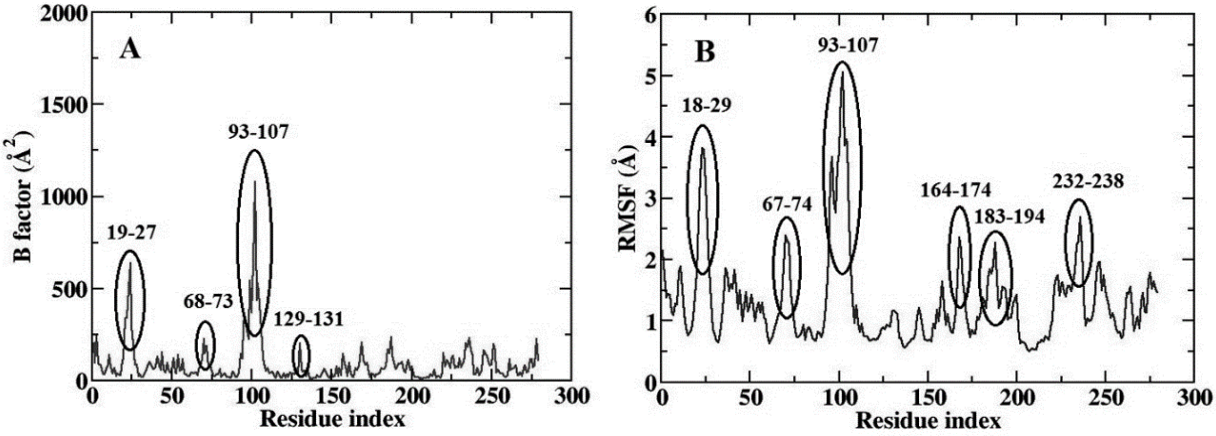

Figure 4.13. *B-factor and RMSF plot showing flexibility in loop region in LMTK3 as a function of residue index*

4.4.3.4. SASA analysis: Solvent accessible surface area (SASA) analysis was performed to explore the behavior of total SASA (polar and non-polar) shown in **Figure**

4.14. SASA results of around 13000Å² - 15000Å² showed that the accessibility was retained with less change during the simulation and confirm that residues were well exposed to the solvent. From the **Table 4.1** we can also infer that most of the residues are well exposed to the solvent.

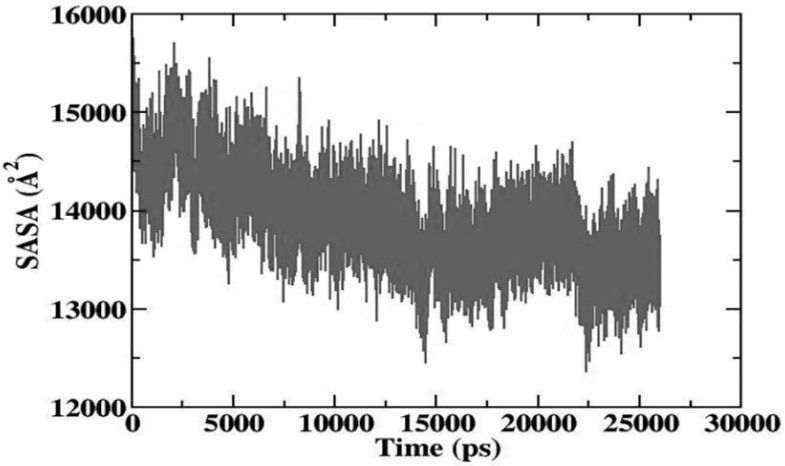

Figure 4.14. *SASA plot of LMTK3 as a function of time*

Table 4.1. Solvent accessible surface area and Molecular accessible surface area of the modelled LMTK3 domain structure determined using Surface Racer software

LMTK3	**Solvent Accessible Surface Area (Å)²**			**Molecular Accessible Surface Area (Å)²**		
	Total Area	**Polar Area**	**Non-Polar Area**	**Total Area**	**Polar Area**	**Non-Polar Area**
Final equilibrated Str. 25 ns	13279.78	6422.65	6857.13	12289.61	4946.80	7342.81

4.4.3.5. Secondary structure analysis: **Figure 4.15** shows secondary structure variation of each residue as a function of simulation time period. We see anti-parallel β sheets near N-terminal region (residue index 2-6, 13-20, 25-34, 65-68, 75-80, 142-148, and 166-170), we also observed few turns in N-terminal region. And there is a secondary structure transition from turns to $\pi 3_{14}$ helix in the region (residue index 200-240), and turns to 3_{10} helix in the region (residue index 40-60). Few turns also observed in C-terminal region (residue index 210-225, 235-245, 260-262, 279). The equilibrated

structure of LMTK3 with the details of secondary structure content was shown in **Figure 4.16**.

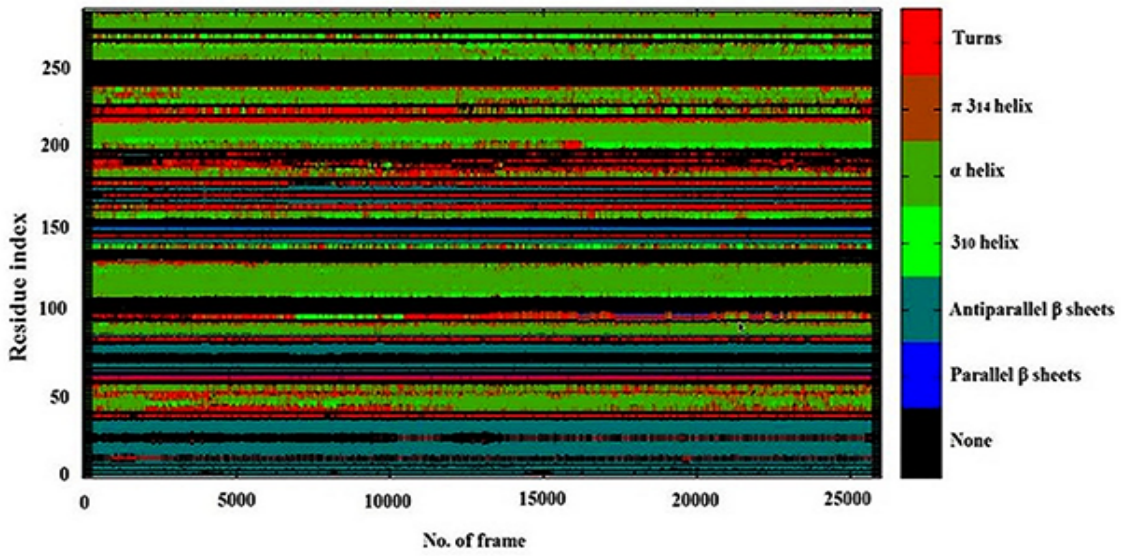

Figure 4.15. *DSSP plot: secondary structure variation of each residue as a function of simulation time period.*

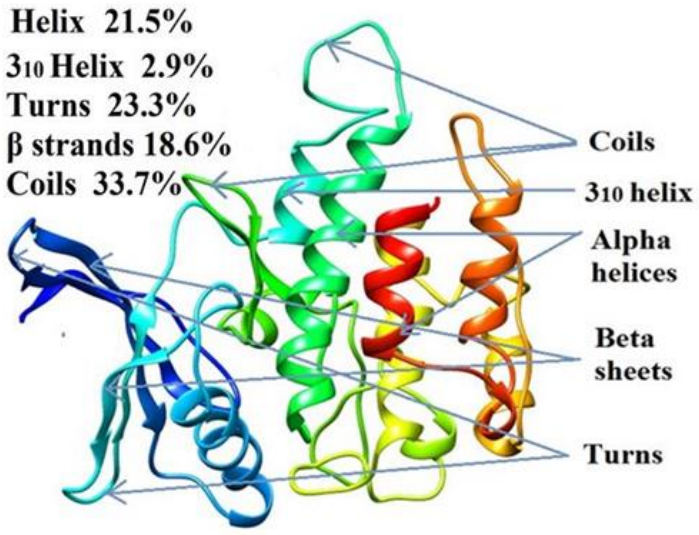

Figure 4.16. *Equilibrated structure of LMTK3 domain with the percentage of secondary structure content*

From the PDBsum server we obtained detailed secondary structure topology for equilibrated structure of LMTK3. **Figure 4.17 (A)** shows the secondary structure model of LMTK3 which includes β sheets, β hairpins, β strands, α helices, helix–helix

interactions and β turns. LMTK3 model structure exhibits five β hairpins which are involved in the secondary structure formation. **Figure 4.17 (B)** represents the topology of β sheet that shows the directions and connectivity of helices and strands in LMTK3. We observed nine β sheets to be involved in the formation of β sheet topology **Figure 4.17 (C)** shows the hydrogen bonding pattern in β sheet topology. Two to three intra-hydrogen bonds were found between the β sheet residues. And more number of residues were involved in the formation of inter-hydrogen bonds between the β sheets. These hydrogen bonds may play a vital role in the formation and stabilization of 3D structure of LMTK3.

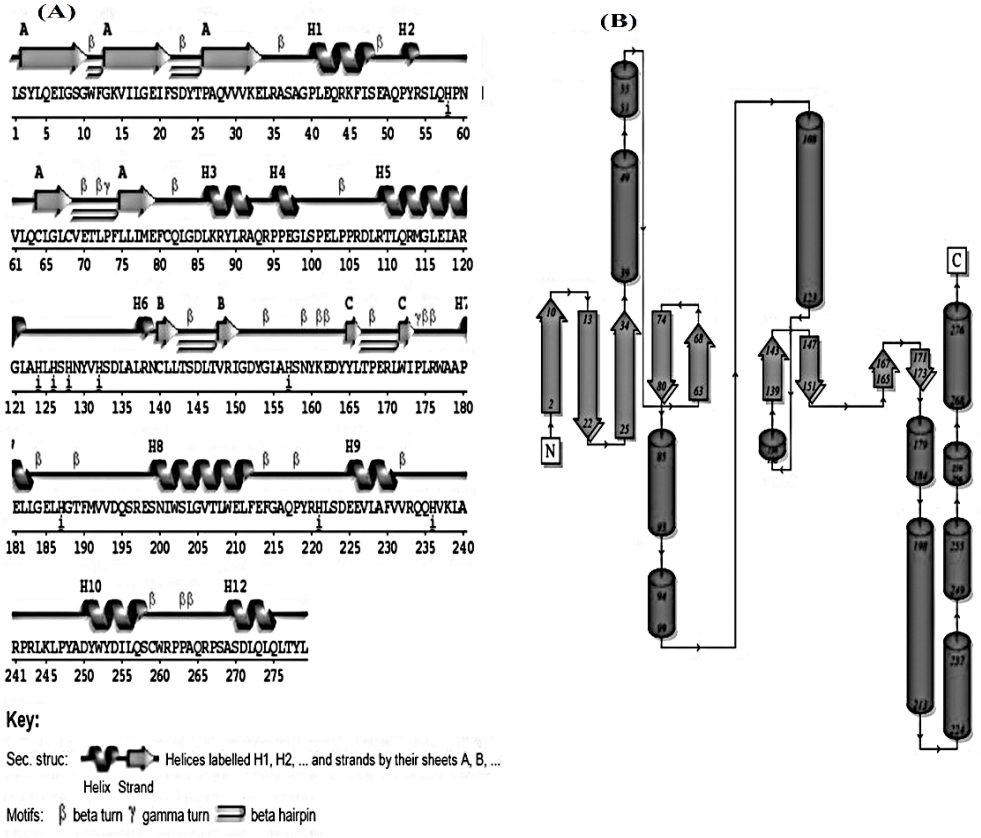

Figure 4.17 *Secondary structure topology, β sheet topology and hydrogen bond pattern of LMTK3 domain.* ***(A)*** *Secondary structural elements (β sheets, β hairpins, β turns, strands, helices, helix–helix interactions, and γ turns).* ***(B)*** *Topology diagram illustrating the β strands, represented by the large arrows (grey), join up side-by-side. The figure also illustrates the relative locations of theα helices (black cylinders). The small arrows specify the directionality of the protein chain, from the N- to the C-terminus.*

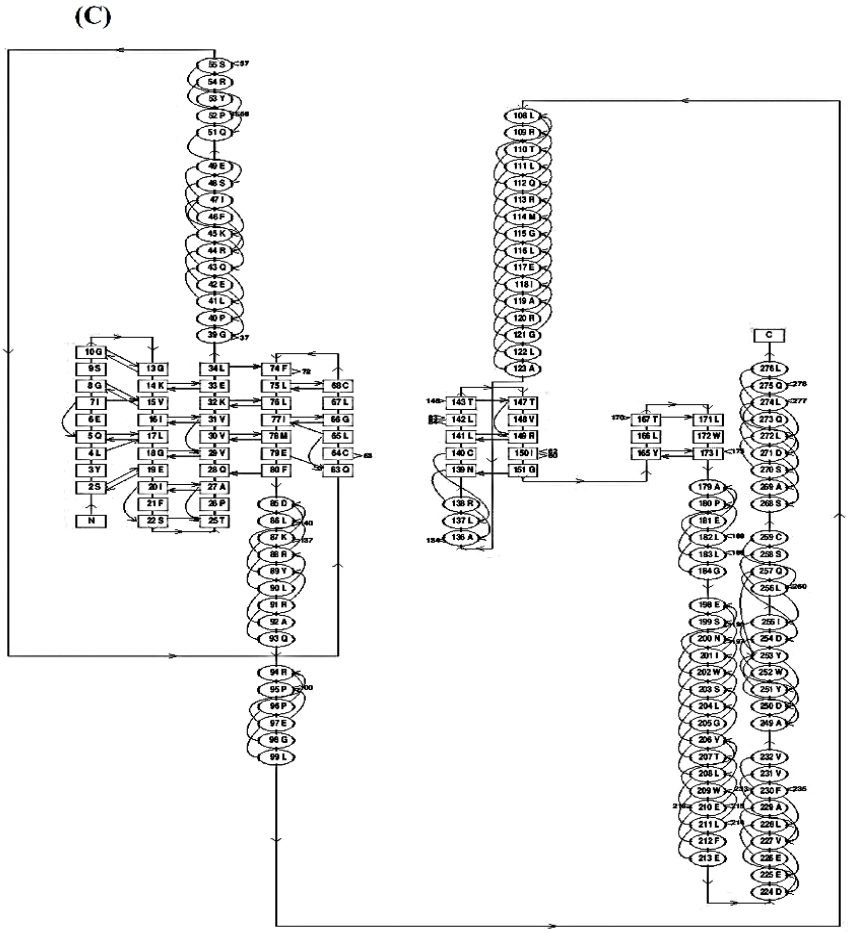

Figure 4.17 *(C) Intra- and inter-hydrogen bond pattern between residues in β sheet topology.*

4.4.3.6. Hydrophobic contact analysis and Hydrophobicity: Hydrophobic interactions are important for protein folding and thereby rendering the protein alive and biologically active [313]. The hydrophobic effect is also considered to be one of the major driving forces for the folding of globular proteins [314]. We analyze the number of hydrophobic contacts and hydrophobicity in LMTK3. Analysis of hydrophobic contacts from MD simulation reveals that the hydrophobic contact ranges between 44500 to 46300 in LMTK3 **(Figure 4.18)**, which stabilizes the LMTK3 in dynamics system during the course of MD simulation.

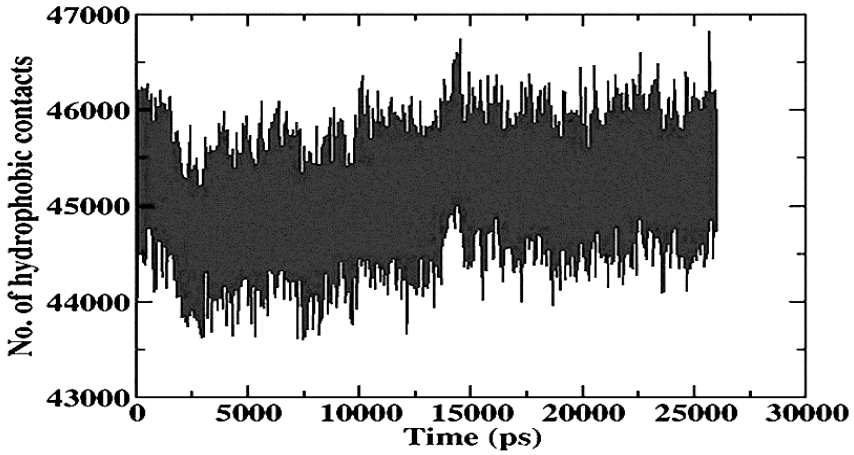

Figure 4.18. *Number of hydrophobic contacts in LMTK3 as a function of time*

We also analyzed hydrophobicity plot of the LMTK3 amino acid sequence (**Figure 4.19)** and hydrophobic cluster analysis (HCA) (**Figure 4.20**). Hydrophobicity plot analysis revealed that LMTK3 expresses low complexity on the segments with high-predicted flexibility (**Figure 4.19**). HCA demonstrated sequences with 2 D representation where the sets of adjacent hydrophobic amino acid residues were encircled and designated hydrophobic clusters. More than 30 hydrophobic clusters were observed in LMTK3 shown in **Figure 4.20**. The packing of residues in the hydrophobic area is critical for maintaining the stability of LMTK3.

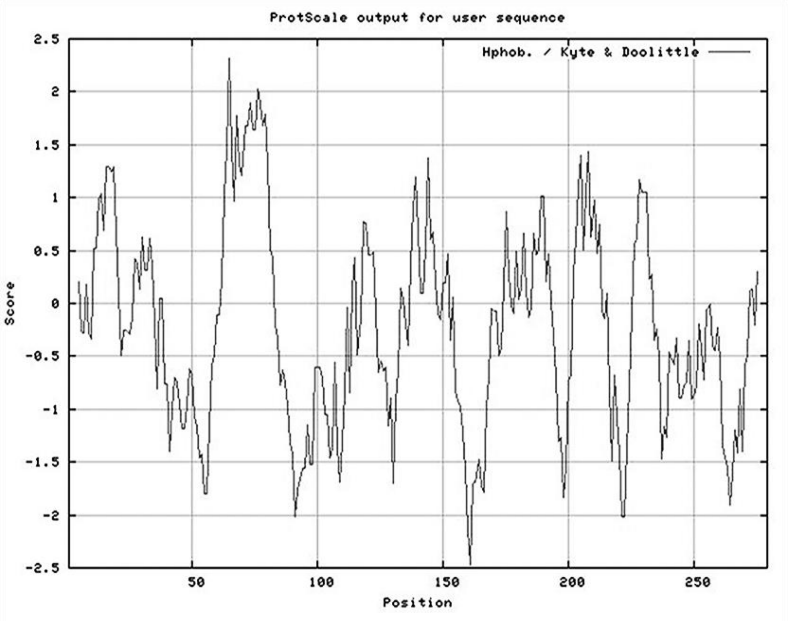

Figure 4.19*. Hydrophobic plot analysis of LMTK3 domain*

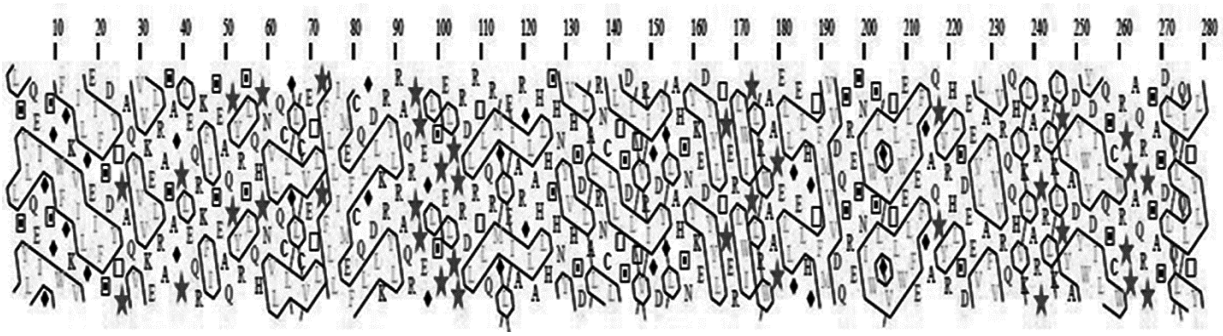

Figure 4.20. *Hydrophobic cluster analysis of LMTK3 domain*

We also detected numerous hydrophobic patches distributed on the surface of equilibrated structure of LMTK3. The hydrophobic patches are represented in different colours according to the area depicted in **Figure 4.21(A)** Hydrophobic patches dispersed on the surface of LMTK3 may be involve in the formation of interaction sites that will be responsible for the stability of LMTK3 structure. Thus hydrophobic interaction is considered to be one of the most important factors for protein folding and stability [315-317].

4.4.3.7. Electrostatic potential: Electrostatic interactions play a central role in protein structure and its function [318-320]. The electrostatic surface potential at a particular point on the surface of equilibrated structure of LMTK3 is coloured blue, red and white **Figure 4.21(B)**. The red symbolizes negative charge, blue is positive and white is neutral. Protein surface are coloured according to electrostatic projection at the level of -10 kcal/mol and +10 kcal/mol. The electrostatic potential study is crucial for the prediction of the possible interaction sites on a protein. For example, a DNA-binding protein will most likely have a pocket of positive charges, which helps it to bind DNA, as DNA is covered in negative charges due to the phosphates. If two proteins are said to interact with each other, then we can more certainly look for a pocket of negative charges (red) on one protein and a bump of positive charge (blue) on the other (or vice versa) which indicates the interaction site on a protein.

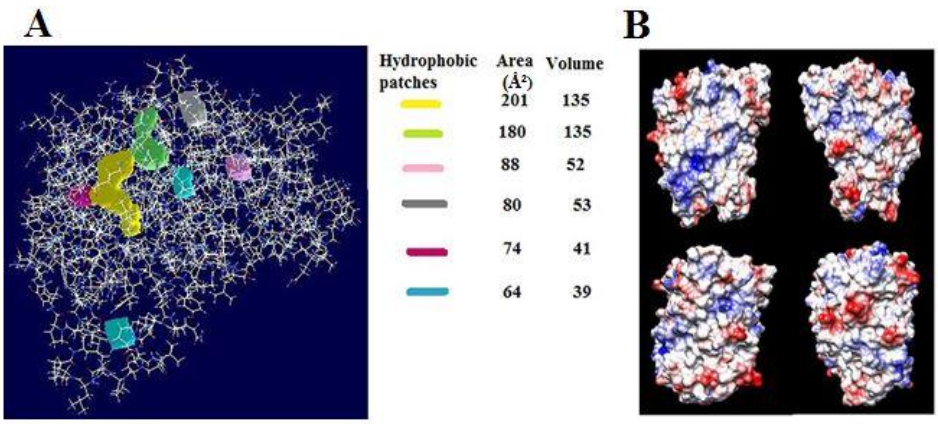

Figure 4.21. *Hydrophobic patches as determined with Swiss pdb viewer for equilibrated structure of LMTK3 domain. Surface representation of hydrophobic patches shown in different colours along with area. (B) Electrostatic potential distributions on the LMTK3 domain surface from -10.0 kcal mol^{-1} (red) to +10.0 kcal mol^{-1} (blue)*

4.4.3.8. Hydrogen bond analysis: Intra-molecular hydrogen bonds play a significant and an indispensable role in protein folding, protein structure and molecular recognition and protein stability [321]. For calculating the intra-molecular hydrogen bond, the cut-off for angle and distance was set to 120° and 3.5 Å. We see the number of intra-molecular hydrogen bonds from the conformational dynamics of LMTK3 to be oscillating around 110-150 (**Figure 4.22**) and stabilizing the LMTK3 structure. **Table 4.2** shows details of H-bond donor and acceptor between residues with particular frames and fractions.

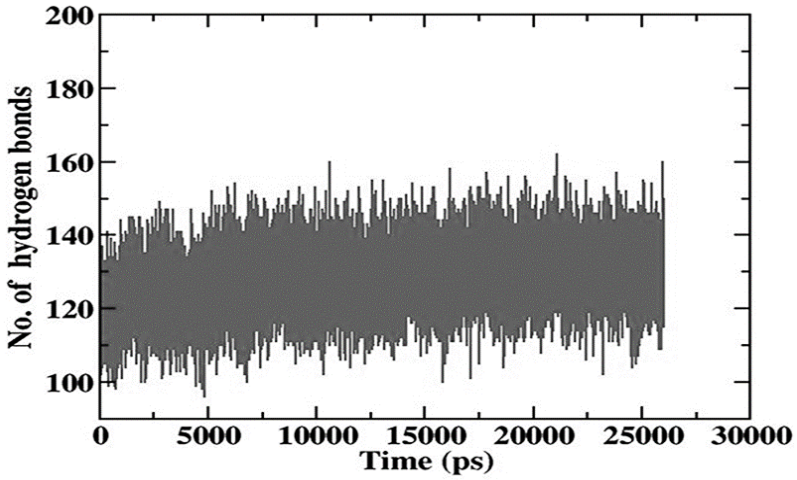

Figure 4.22. *Number of intra-molecular hydrogen bonds in LMTK3 domain as a function of time*

Table 4.2. Hydrogen bond interaction in LMTK3 domain

#Acceptor	DonorH	Donor	Frames	Frac	AvgDist	AvgAng
ASP_163@OD2	ARG_138@HH21	ARG_138@NH2	22159	0.8523	2.7915	151.8802
ARG_149@O	LEU_141@H	LEU_141@N	21284	0.8186	2.8437	161.2398
GLU_181@OE2	ARG_266@HH22	ARG_266@NH2	21154	0.8136	2.781	161.4618
ARG_261@O	ARG_266@HE	ARG_266@NE	20897	0.8037	2.8295	155.4723
ASP_85@OD1	ARG_88@HH11	ARG_88@NH1	20638	0.7938	2.7937	152.9027
ALA_240@O	TRP_209@HE1	TRP_209@NE1	20592	0.792	2.8448	159.8777
GLU_198@O	ARG_266@HH21	ARG_266@NH2	20145	0.7748	2.8366	158.7638
LEU_122@O	HIE_126@H	HIE_126@N	19782	0.7608	2.8448	158.1573
ASP_134@O	ASN_139@HD21	ASN_139@ND2	19631	0.755	2.8359	158.6635
ILE_16@O	VAL_31@H	VAL_31@N	19577	0.753	2.8622	162.0283
LEU_76@O	LYS_32@H	LYS_32@N	19404	0.7463	2.8648	163.2152
ASP_163@OD1	GLN_217@HE21	GLN_217@NE2	18895	0.7267	2.8021	161.568
SER_203@O	THR_207@HG1	THR_207@OG1	18732	0.7205	2.7908	160.5925
GLN_82@OE1	GLN_82@H	GLN_82@N	18547	0.7133	2.8479	152.2982
ILE_77@O	LEU_65@H	LEU_65@N	18496	0.7114	2.8597	160.9261
MET_78@O	VAL_30@H	VAL_30@N	18192	0.6997	2.8621	160.8774
GLU_210@O	GLY_215@H	GLY_215@N	18145	0.6979	2.8655	158.4749
CYS_140@O	LEU_86@H	LEU_86@N	17940	0.69	2.8696	161.5917
GLU_213@OE1	ARG_243@HE	ARG_243@NE	17739	0.6823	2.798	159.9088
GLN_5@O	LEU_17@H	LEU_17@N	17685	0.6802	2.8649	162.9525
ASP_85@OD1	ARG_88@HE	ARG_88@NE	17329	0.6665	2.8401	151.6119
ASP_85@OD2	TYR_160@HH	TYR_160@OH	17304	0.6655	2.759	160.1909
ILE_150@O	LEU_62@H	LEU_62@N	17176	0.6606	2.8611	157.1847
TYR_89@O	GLN_93@H	GLN_93@N	16797	0.646	2.8511	159.1208
GLY_215@O	ARG_91@HH11	ARG_91@NH1	16464	0.6332	2.8445	155.2627
ASP_134@OD1	ARG_138@HH11	ARG_138@NH1	16397	0.6307	2.7697	162.2504
VAL_61@O	ARG_149@HH11	ARG_149@NH1	15958	0.6138	2.8254	152.1672
GLU_181@OE1	HIE_236@HE2	HIE_236@NE2	15941	0.6131	2.8157	161.3485
GLU_213@OE1	ARG_243@H	ARG_243@N	15781	0.607	2.8395	164.9036
GLN_112@OE1	ARG_109@HH21	ARG_109@NH2	15758	0.6061	2.8229	155.0396
HIE_128@O	TYR_130@H	TYR_130@N	15574	0.599	2.8292	148.3753
ALA_136@O	ASN_139@H	ASN_139@N	15373	0.5913	2.8779	152.3522
VAL_31@O	ILE_16@H	ILE_16@N	15307	0.5887	2.8838	162.4897
LEU_75@O	CYS_68@H	CYS_68@N	15126	0.5818	2.8794	159.7774
LEU_17@O	LEU_4@H	LEU_4@N	14795	0.569	2.8689	158.8718
GLU_181@OE1	ARG_266@HH12	ARG_266@NH1	14790	0.5688	2.8241	159.4602
GLU_213@OE2	ARG_243@HH11	ARG_243@NH1	14674	0.5644	2.8162	160.2034
TYR_165@O	ILE_173@H	ILE_173@N	14600	0.5615	2.8794	159.7593
GLY_84@O	LEU_142@H	LEU_142@N	14365	0.5525	2.8833	161.0033

ALA_27@O	ILE_20@H	ILE_20@N	14063	0.5409	2.8896	163.0985
ASP_134@OD2	TYR_164@HH	TYR_164@OH	14051	0.5404	2.7366	162.3549
ARG_44@O	SER_48@HG	SER_48@OG	14021	0.5393	2.7877	159.4544
LEU_204@O	LEU_208@H	LEU_208@N	13946	0.5364	2.8897	158.5641
VAL_30@O	MET_78@H	MET_78@N	13867	0.5333	2.8891	162.3567
VAL_29@O	GLY_18@H	GLY_18@N	13862	0.5332	2.8782	157.7378
LEU_279@OXT	ARG_109@HH22	ARG_109@NH2	13676	0.526	2.8185	155.28
GLY_18@O	VAL_29@H	VAL_29@N	13635	0.5244	2.888	164.1551
LEU_141@O	ARG_149@H	ARG_149@N	13584	0.5225	2.8906	161.7878
GLY_8@O	VAL_15@H	VAL_15@N	13539	0.5207	2.8726	157.1051
TYR_153@O	LEU_155@H	LEU_155@N	13452	0.5174	2.861	147.0696
ILE_20@O	ALA_27@H	ALA_27@N	13422	0.5162	2.8789	160.2085
LEU_279@O	ARG_109@HH22	ARG_109@NH2	13295	0.5113	2.8198	155.0324
ASN_60@O	ASN_60@HD21	ASN_60@ND2	13244	0.5094	2.816	150.6624

4.4.3.9. Binding cavity: In proteins, the identification of ligand binding is critical for a range of applications which includes molecular docking, *de novo* drug design and structural identification and comparison of functional sites [322]. Here we made an attempt to identify probable binding pockets in LMTK3. Binding pocket analysis of LMTK3 was done using LIGSITE^{csc} online server and Surface racer software. From LIGSITE^{csc} server we predicted five binding pockets denoted as spheres shown in **Figure 4.23**. And information regarding the coordinates and number of grids in clusters of all the five pockets was mentioned in **Table 4.3,** and the list of 34 residues involved in forming five binding pockets listed in **Table 4.4.** The determination of probable cavities in LMTK3 may be helpful in designing the inhibitors against LMTK3.

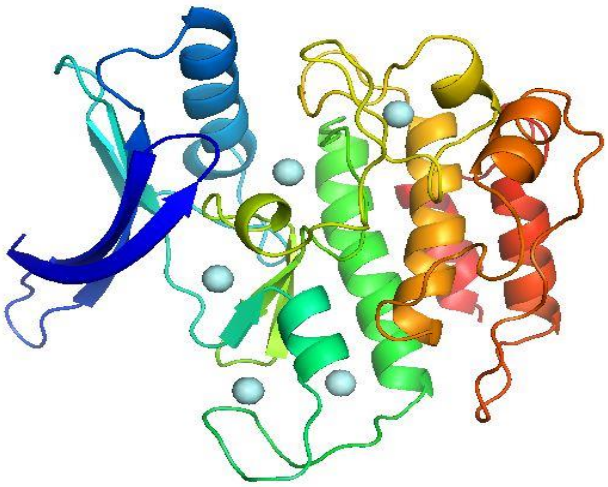

Figure 4.23. *Five binding pockets in LMTK3 domain denoted as spheres obtained LIGSITE^{csc} serve*

Table 4.3. Details of the five binding Pockets (determined using LIGSITE^{csc} server) in the model structure of LMTK3 domain

No. of pocket	No. of grids in clusters	Coordinates
1	107	43.477, 24.439, 47.501
2	74	30.638, 38.375, 34.829
3	54	40.152,32.379, 37.936
4	29	23.162, 31.330, 39.388
5	29	41.921, 20.468, 58.285

Table 4.4. List of residues along with the residue index involved in the formation of five binding pockets in human LMTK3 domain structure as determined by LIGSITE^{csc} server

Residue	Residue index	Residue	Residue index
ILE	7	LEU	111
LYS	45	MET	114
GLU	49	HIS	132
LEU	83	SER	133
GLY	84	ASP	134
ASP	85	ASN	139
ARG	88	LEU	141
TYR	89	LEU	146

LEU	90	ASP	152
ALA	92	TYR	153
GLN	93	LEU	155
ARG	94	LEU	182
GLY	98	PHE	190
LEU	99	GLN	195
SER	100	SER	196
ARG	106	SER	199
THR	110	ASN	200

4.4.4. ATP binding pocket of LMTK3 domain

The ATP binding cavity in LMTK3 was determined by performing molecular docking in Autodock4.2. We obtained ten conformations of LMTK3-ATP docked complex. Among the ten conformations, the best conformation with a least binding energy (-1.71 kcal mol^{-1}) was chosen **as shown in Figure 4.24 (A)**. Using the Ligplot+ tool we have analysed the hydrogen bonding and hydrophobic interaction present in LMTK3-ATP complex as shown in **Figure 4.24 (B)**. From the **Figure 4.24 B**, we see residues Asp284 and Lys177 forming two hydrogen bonds with ATP while His264, His289, Ser290, Tyr185 forming one hydrogen bond with ATP and other residues involved in hydrophobic interaction.

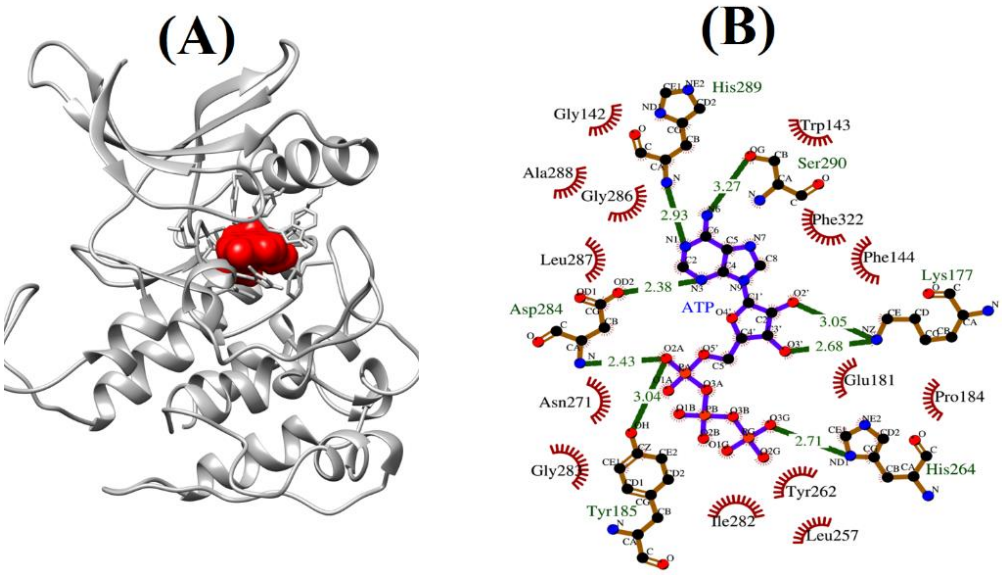

Figure 4.24. *LMTK3-ATP interaction (A) LMTK3-ATP complex with least binding free energy (-1.71 kcal* mol^{-1}*), ATP is shown in red sphere form (B) Ligplot+ analysis: Residues of LMTK3 involved in the formation of hydrogen bonds (green colour) and*

hydrophobic contacts are represented by an arc with spokes radiating towards the ATP atoms they contact.

4.5. Conclusions

In this study, we have modelled the 3D structure of human LMTK3 domain. The modelled structure of LMTK3 was observed to show similar architecture as that of well-studied protein kinase. From molecular dynamics simulation study, we inferred the predicted model structure of LMTK3 to have stable conformation with proper folding, flexible loop regions, secondary structure transition such as turns to $\pi 3_{14}$ helix and turns to 3_{10} helix. The stability of LMTK3 structure can also been seen from the solvent accessibility analysis and hydrophobic contact analysis. Our computational analysis establishes that LMTK3 has diverse structural topology along with electrostatic potential, hydrophobic patches and clusters. Since LMTK3 has been recently identified as therapeutic target for ERα positive breast cancer, the salient structural properties of LMTK3 dictated in this work may be useful to understand the mechanism of the structure–function relationships of LMTK3. Lastly, the prediction of probable cavities and interaction properties represent a platform for *in silico* drug design studies as a potential method for the designing of potential inhibitors of LMTK3.