

Glide Docking, Autodock, Binding Free Energy and Drug-Likeness Studies for Prediction of Potential Inhibitors of Cyclin-Dependent Kinase 14 Protein in Wnt Signaling Pathway

Revanth Bathula ¹, Narasimha Muddagoni ¹, Goverdhan Lanka ¹, Mahender Dasari ¹, Sarita Rajender Potlapally ^{1,*}

¹ Molecular Modeling Laboratory, Department of Chemistry, Nizam College, Osmania University, Hyderabad, India

* Correspondence: saritarajender@gmail.com;

Scopus Author ID 55317928000

Received: 23.03.2021; Revised: 30.04.2021; Accepted: 5.05.2021; Published: 18.06.2021

Abstract: Cyclin-dependent kinase 14 plays an essential role in multiple cancers. Cyclin-dependent kinase 14 is a serine/threonine kinase and is a member of the cell division cycle 2(cdc2) related protein kinase family, which plays a key role in promoting Wnt signaling pathway of the cell cycle and its overexpression causes various human cancers. The 3D structure of cyclin-dependent kinase 14 was built using the homology-based modeling technique. The generated model is optimized by NAMD-VMD software. The quality of stabilized CDK14 protein was checked using Ramachandran plot and ProSA servers. The potential binding site region was recognized using SiteMap and manual correlation techniques from literature studies. The virtual screening was performed with the TOSLab database of 27253 output molecules against CDK14 protein using Glide docking to assess novel chemical entities. Their binding energies were calculated from PrimeMMGSA and AutoDock. The novel lead molecules have been prioritized based on efficient binding energies (from AutoDock and PrimeMMGSA), better glide scores, good bioavailability, and acceptable ADME properties. Thus, these are considered as CDK14 protein inhibitors for cancer therapeutics.

Keywords: multiple cancer; Wnt signaling pathway; NAMD-VMD; PrimeMMGSA; AutoDock and ADME properties.

© 2021 by the authors. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Cancer is currently rising rapidly worldwide, with a high rate of morbidity and mortality [1]. Cyclin-dependent kinases (CDKs) belong to the serine/threonine kinases family, regulates the cell cycle progression, transcription and cell differentiation via their association with cyclins [2]. The dysregulation of the CDK-cyclin complex is involved in various cancers [3]. Cyclin-dependent kinase-14(CDK14), also named PFTK1 is a member of CDKs. The activity of CDK14 depends on binding with its partners Cyclin Y and Cyclin D3, which are regulated in higher eukaryotic cell cycles [4,5]. Overexpression of CDK14/Cyclin Y complex causes dysregulation in the Wnt/ β -catenin signaling pathway resulting in hyperphosphorylation and activation of LRP6 receptors, thereby upregulating β -catenin. Eventually, overexpressed β -catenin migrates to the nucleus from the cytoplasm and then combines with TCF/LEF enhancing translational co-activators such as C-myc, Cyclin D1 and MMP9 genes, which participates in cancer cell cycle progression during the G2/M phase (Figure1) [6,7,8,9]. The β -

<https://biointerfaceresearch.com/>

catenin can affect the centrosomal activity localized at kinetochores, disturbs the dynamics of microtubules and misorientation of alignment of the mitotic spindles in mitosis [10].

Irregular activation of CDK14 protein has been involved in multiple cancers such as breast cancer [9], non-small cell lung cancer [11], pancreatic cancer [12], colon cancer [13], ovarian cancer [14,15] and gastric cancer [16].

The present study involves establishing the 3D structure of CDK14 protein, stabilized by the NAMD-VMD Tool. The quality of the 3D structure was validated using ProSA and Ramachandran plot. The glide docking is carried out at the active site of CDK14 protein using TOSLab database molecules. PrimeMMGBSA and AutoDock calculated the binding energies of protein-ligand complexes for further optimization of final hits, which were analyzed with pharmacokinetic properties. Therefore, novel lead molecules were designed, which bind to CDK14 protein at the active site to halt the progression of cancer cells within G2-M Phase.

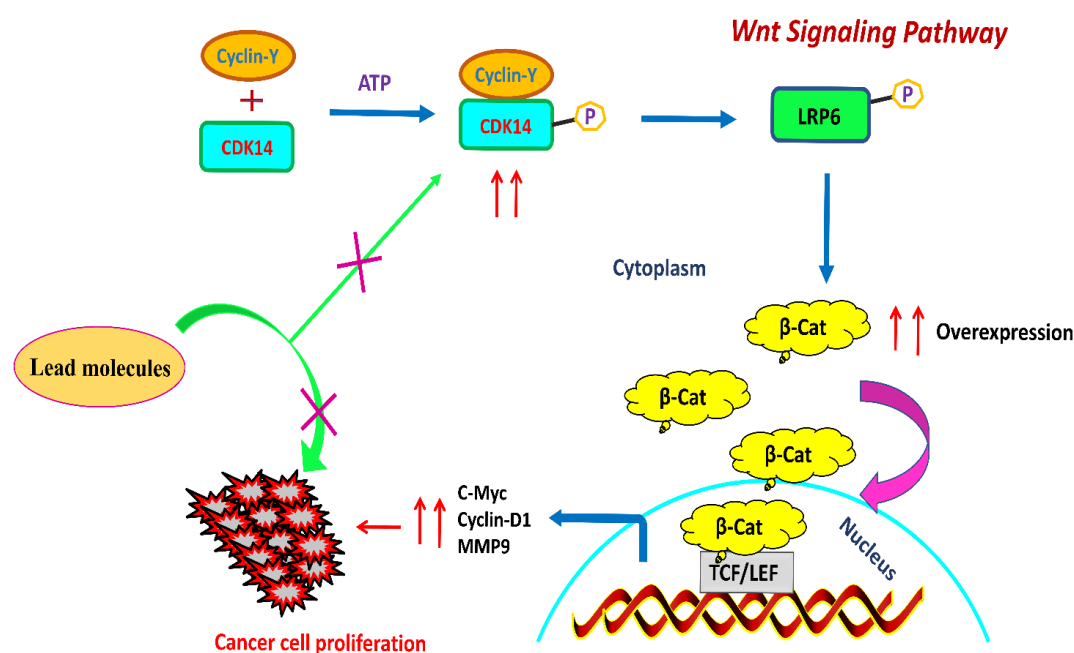


Figure 1. Biochemical pathway of CDK14 protein in cancer cell progression. Overexpression of CDK14 protein combined with Cyclin Y hyperphosphorylated the LRP6 activating Wnt/ β -catenin signaling pathway via increase the amount of β -catenin and then move from cytoplasm to nucleus. The β -catenin binds to TCF/LEF to activate transcription genes (C-Myc, Cyclin-D1 and MMP9), leading to cancer cell proliferation.

2. Materials and Methods

2.1. Homology modeling.

The experimental methods such as X-ray crystal and NMR structure are not available for CDK14 protein. Therefore, knowledge-based modeling was used to produce the 3D structure of CDK14 protein. The fasta sequence CDK14 was retrieved from UniprotKB with accession id O94921 with 469 amino acids [17]. This target fasta sequence was subjected to BlastP and J Pred4 servers to search for experimentally determined homologous template protein (PDB ID: 3MTL), based on parameters such as query coverage low and lowest E value [18,19]. The sequence alignment between CDK14 and 3MTL is carried out in ClustalX1.2 to define the similarity sequences of structural and functional regions. Modeller9.9 program was used to build the 20 homology modeled structures of CDK14 protein and the best model was selected based on MolPDF (molecular probability density function) for further studies [20].

2.2. Energy minimization of CDK14 protein.

The constructed 3D model had unfavorable bond distances, bond angles, and improper planarity of dihedral angles and hence is mandatory to reduce the potential energy of CDK14 protein [21]. NAMD-VMD (NanoScale Molecular Dynamic-Visual Molecular Dynamics) program was used for the refinement of CDK14 protein [22]. It worked with CHARMM (Chemistry at Harvard Molecular Mechanics) force field and visualized in the VMD tool. The modeled protein was solvated with water molecules in periodic boundary conditions and simulation was carried out within 100000-time steps [23]. The lowest energy state of CDK14 protein was analyzed by final RMSD trajectory files and monitored with RMSD values against time steps.

2.3. Validation.

It is necessary to evaluate the correctness of constructed 3D model protein in receptor-based drug design. Energy minimized 3D model protein was validated using ProSA (Protein structure analysis), Errat and Ramachandran plot. ProSA is a program that gives information about the quality of protein and energy level of amino acid residues [24]. Errat is a tool used for the evaluation of non-bonded atom-atom interactions in 3D modeled proteins. The stereochemical quality of the 3D model protein was evaluated by considering steric hindrance between phi (Φ) and psi (ψ) torsional angles of amino acid residues in the Ramachandran plot [25].

2.4. Protein preparation and Ligand preparation.

The homology modeled CDK14 protein is accurately optimized for molecular docking studies. The refinement of CDK14 protein was carried out by the protein preparation wizard in the maestro version (9.0.111) [26]. Protein preparation wizard involves adding missing hydrogen atoms, water molecules were deleted and correct bond orders were assigned at force field OPLS-2005. The energy minimization in impref module was terminated at the default constraint reaches the specific RMSD value of 0.30Å. LigPrep module of Schrodinger suite was used to refine the ligand molecules by submitting the TOSLab database molecules of data set 17643 and using forcefield OPLS-2005. Various ionic states, tautomeric states and stereochemistries generated from each input ligand molecule at pH 7.0+/-2.0 using Epik. The low-energy ring conformations produced up to five [27,28].

2.5. Identification of binding site.

Structure-based drug design involves the identification of binding sites of targeted protein for inhibiting cell progression. The binding site of CDK14 protein has been identified using SiteMap, literature studies and manual correlation techniques. SiteMap module of Schrodinger suite provides potential binding sites of CDK14 protein with their surface area of hydrogen bond acceptors, hydrogen bond donors, hydrophilic and hydrophobic regions [29]. The reported active site residues of the template were manually correlated to the CDK14 protein sequence inClustalX2.1 [30]. These residues are considered as active site residues of CDK14 protein for docking studies. The 3-dimensional grid was generated for favorable binding mode in glide docking. The 3-dimensional grid box was built with the active site amino acid residues of CDK14 protein using the receptor grid generation program in the glide tool [31].

2.6. Virtual screening using glide.

A virtual screening study is a good method to search lead molecules from large databases against the biological target protein in the drug discovery process. A TOSlab database 27253 output molecules were subjected to virtual screening by the filtering mode in HTVS, SP and XP docking at the active site of CDK14 protein using virtual screening workflow of glide tool [32]. 10% of molecules are filtered at each step and the best protein-ligand docked molecule was analyzed based on scoring functions and visualized using Discovery studio 3.5 [33].

2.7. Binding free energy calculation.

The XP docked output molecules are used to calculate the binding free energy of protein-ligand complexes using prime MMGBSA (molecular mechanics generalized born surface area) at forcefield OPLS-2005 [34]. Free energy of binding describes the affinity of ligand molecule with a protein. The binding free energy was calculated at binding poses of protein-ligand complexes as follows

$$\Delta G_{\text{Binding}} = \Delta G_{\text{complex}} - (\Delta G_{\text{protein}} + \Delta G_{\text{ligand}})$$

Where $\Delta G_{\text{Binding}}$ is the Minimized binding free energy; Whereas $\Delta G_{\text{complex}}$, $\Delta G_{\text{protein}}$ and ΔG_{ligand} represent the free energy of protein-inhibitor complex, protein, inhibitor, respectively.

2.8. AutoDock.

The program AutoDock4.2 is a computer-aided docking tool used to identify the binding energy of ligand molecules against the target protein. Thirteen XP output docked ligand molecule and CDK14 protein are prepared in PDBQT format files for docking. GPF and DPF files were generated by given grid and docking parameters. The conformations of hit molecules and binding energies of docked complexes were identified using Lamarckian genetic algorithm in AutoDock4.2 [35,36].

2.9. ADME properties.

Most of the new drugs have failed in clinical trials in drug development because of ADME characteristics, thereby increasing the time and cost. Qikprop tool is useful for predicting the accurate ADME properties of hit molecules so as to reduce the cost and avoid spending valuable time [37]. The pharmacokinetic and physicochemical properties were observed for more drug-likeness candidates. Final hit molecules were filtered using ADME properties such as Octanol-water partition coefficient, % Human oral absorption. The ligand molecules with acceptable pharmacokinetic and physicochemical properties were considered potential biological inhibitors of CDK14 protein [38].

3. Results and Discussion

3.1. Analysis of CDK14 protein structure and validation.

Experimentally, the 3-D model by NMR, X-ray crystallography studies of CDK14 protein was not reported. Hence, a model is constructed based on a homologous template using modeller9.9. Residue sequences of CDK14 were taken from UniportKB, consisting of 469

amino acid residues with accession ID O94921. The template 3MTL (CDK16) was chosen from BlastP and Jpred servers by submitting the fasta sequences of CDK14 protein, resulting in E-value ($5e-127$), sequence similarity(59%) and Query coverage (68%) are illustrated in Table 1 [39]. The conserved domain of CDK14 protein obtained from BlastP is showing binding site residues between135-420 displayed in Figure 2. Pairwise sequence alignment between CDK14 and CDK16 protein is carried out in ClusterX 1.2 and visualized with discovery studio 3.5 (Figure 3). Twenty homology models of CDK14 protein were generated using modeller9.9 [40]. The O94921.B99990009.pdb model of CDK14 protein having the lowest molpdf (molecular probability density function) value of 2458.40 is used for future studies. The modeled CDK14 protein consisting of steric errors is energy minimized using NAMD-VMD software applying CHARMM++ forcefield. The CDK14 protein was solvated in all directions in created periodic boundary conditions with a 10 Å layer of water molecules. Counter ions are added for the neutralizing system. The 100000 steps were run to minimize CDK14 protein at 1 atmospheric pressure and constant temperature. The whole minimization process was analyzed by RMSD value with 1969 time steps. Figure 4 shows the average RMSD value of 1.2 Å is considered the lowest energy of CDK14 protein at time stages (821-985fs) [41]. This resulting stabilized protein is used for further docking studies. The quality of 3D model of CDK14 protein was validated with the Ramachandran plot and ProSA servers. The Ramachandran plot revealed that 99% of residues fall within the most favorable region, indicating the good stereochemical model quality of CDK14 protein, as shown in figure 5. ProSA analysis is employed to check the quality of modeled CDK14 protein with respect to that of experimentally solved proteins. Figure 6a shows a z-score of the 3-D model of CDK14 protein as -6.61, denoting the overall model quality, falling within the region of z-scores of all determined proteins by X-ray and NMR techniques. Figure 6b shows the local model quality with most amino acid residues' energies falling in the negative region, revealing a good protein model quality. The secondary structure details of CDK14 protein were obtained from the PDBsum server showing 16 helices, 19 helix-helix interactions, 23 beta turns, and 6 gamma turns (Table2) [42]. The 3D model of CDK14 protein is visualized using the Pymol tool (Figure7) [43].

Table 1. BlastP and JPred servers are used for Suitable template recognition of CDK14 Protein

S. No	Database server	Template selection Parameters	E-Score	PDB-Code
1	NCBI Server	Sequence position	4e-127	3MTL
2	JPred	Secondary structure, solvent accessibility and coiled-coil regions of prediction	1e-101	3MTL

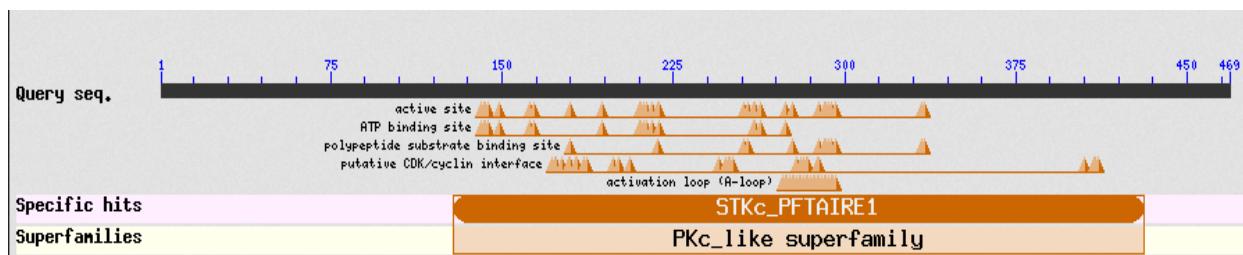


Figure 2. The conserved domain of CDK14 protein. The domain region showing active residues between 135-420.

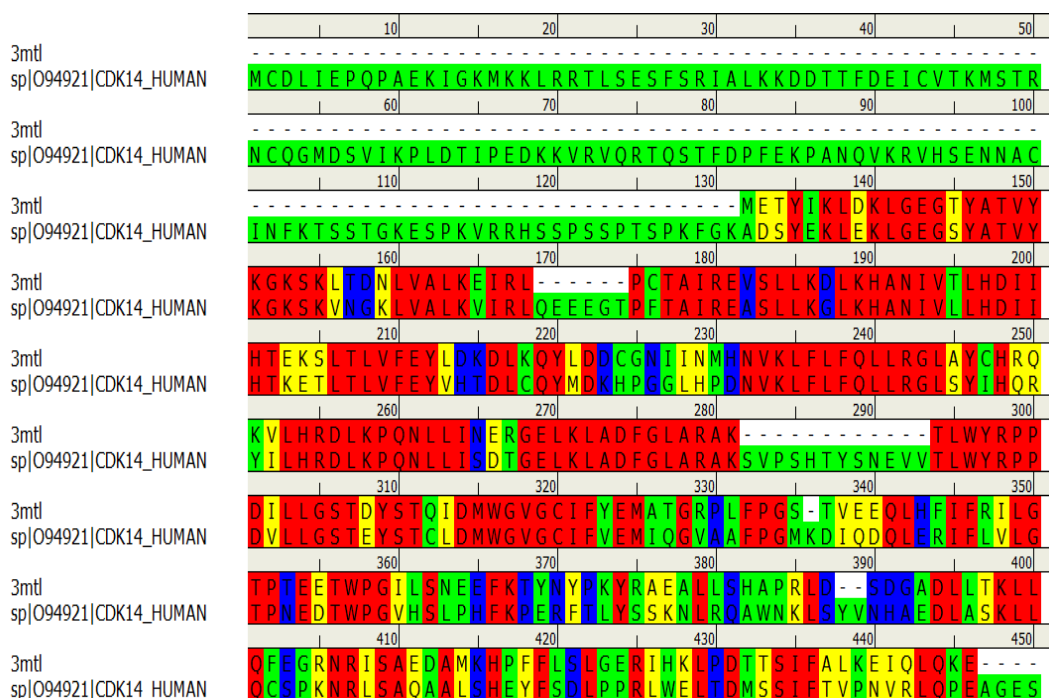


Figure 3. Amino acid sequences of CDK14 were aligned with template 3MTL in ClusterX1.2 and visualized in Discovery studio 3.5. Identical Residues shown in Red region, the yellow color representing the strong zone, blue color region indicating that the residues fall in the weak zone and light green color are shown in the alignment of unmatching residues.

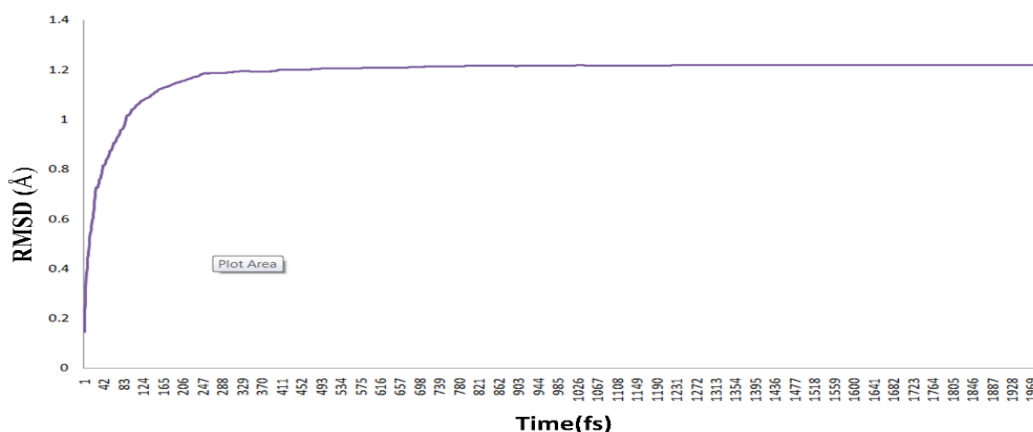


Figure 4. Graphical representation of different energy levels was observed in RMSD with various time steps. The CDK14 protein was stabilized at an average RMSD value is 1.2 Å with time frames 821-985 fs.

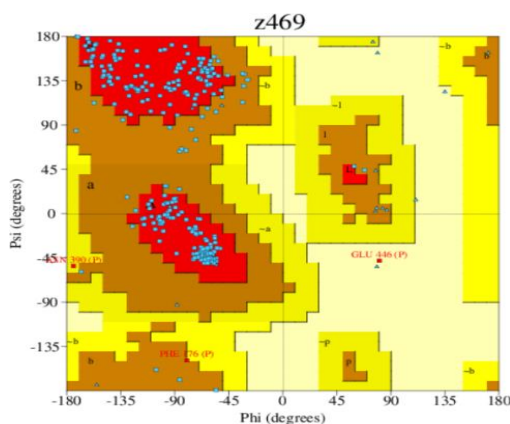


Figure 5. Stereochemical analysis of CDK14 protein in Ramachandran plot. The red color field represents the most favored region energetically, the brown color field indicates the additionally allowed region and the yellow area represents generously allowed regions. Ramachandran plot shows 98.9% of residues are present in the energetically allowed regions, indicating stereochemically stable protein.

plot statistics

	No of residues	%age
Most favoured regions [A,B,L]	259	89.6%
Additionally allowed regions [a,b,l,p]	27	9.3%
Generously allowed regions [~a,~b,~l,~p]	2	0.7%
Disallowed regions [XX]	1	0.3%
Nonglycine and non-proline residues	289	100%
End residues (excl.Gly and Pro)	2	
Glycine residues	20	
Proline residues	20	
Total no of residues	331	

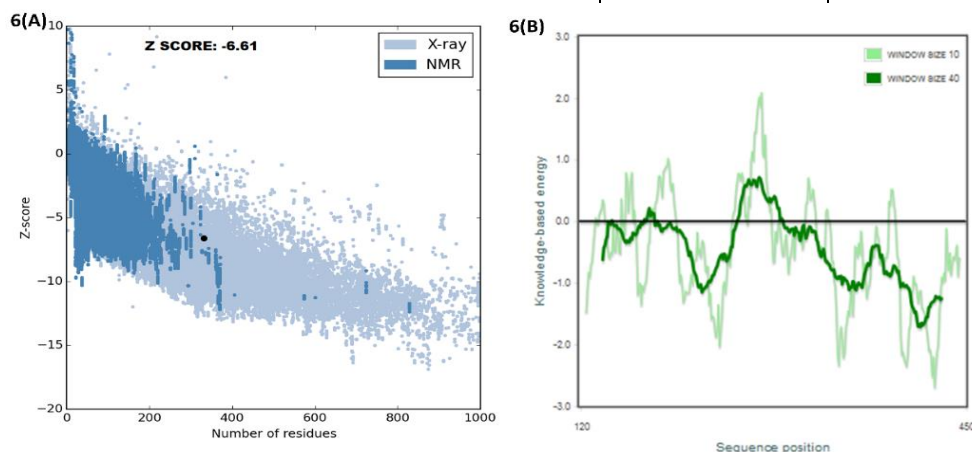


Figure 6. ProSA server was analyzed for quality of modeled CDK14 protein; **(6A)**. The z-score of CDK14 protein shown as a black spot is -6.61, indicating the good model quality of CDK14 protein falling in the experimentally determined region of proteins by NMR and X-ray studies; **(6B)**. The local model quality of CDK14 protein exhibited maximum residues in the negative region, indicating the good model quality of CDK14 protein

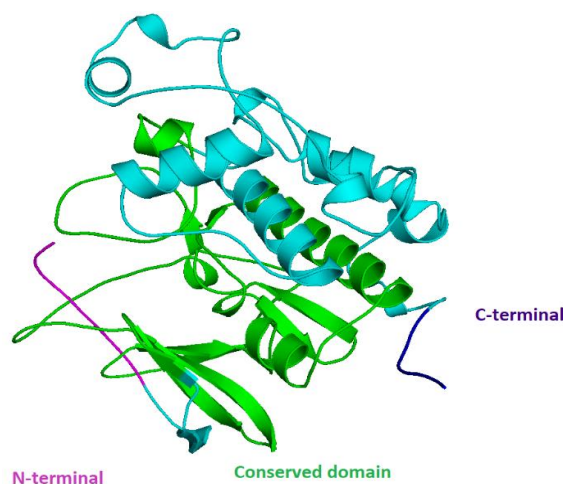


Figure 7. The 3- dimensional structure of CDK14 protein contains 16 helices, 19 helix-helix interacts, 23 beta turns, 6 gamma turns, C-terminal and N-terminal, shown in blue and magentas colors and conserved domain exhibited in green color. The 3D structure was visualized by Pymol software.

Table 2. The Secondary structure data of CDK14 protein recognized from the PDBsum server.

S. No	Start	End	No. of. residues	Length In Å	Sequence
1	Val212	Leu215	4	5.78	VSLL
2	Leu247	Asp253	7	10.94	LKQYLDD
3	Met260	Arg279	20	30.63	MHNVKLFQLLRGLAYCHR
6	Pro330	Leu333	4	6.53	PDIL
7	Gln342	Thr357	16	24.65	QIDMWGVGVGCIFYEMAT
8	Val367	Leu378	12	17.57	VEEQLHFIFRIL
10	Glu393	Thr397	5	8.41	EEFKT
11	Leu408	His411	4	6.74	LLSH
12	Ser417	Leu426	10	15.56	SDGADLLTKL

S. No	Start	End	No. of. residues	Length In Å	Sequence
14	Ala437	Ala440	4	6.30	AEDA
15	Pro444	Leu449	6	10.60	PFFLSL

3.2. Binding cavity recognition of CDK14 protein.

The identification of active site pockets is mandatory for lead optimization and virtual screening hits. A siteMap is a tool used to recognize the active sites of CDK14 protein. It provides graphical data and quantitative numbers that can be a guide to identify ligand molecules with enhanced potency in lead optimization [44]. The surfaces of the hydrophilic region, hydrophobic region, hydrogen bond acceptors and hydrogen-bond donors suitable for the nature of binding regions and their graphical surface region measurement in angstrom units (Å) obtained for SiteMap is illustrated in Table 3 (Figure 8) [45]. The active site residues of homologous template 3MTL are taken from pdbsum server by analyzing ligplot 2D diagram (Figure 9). These residues were manually correlated to CDK14 residues using ClusterX1.2 server, resulting in active site residues Leu191, Val199, Ala212, Val244, Phe260, Glu261, Tyr262, Asp266, Gln269, Gln310, Asn291, Leu313, Asn324 and Phe325 of CDK14 protein (Figure 10) [46]. These active residues were used to build a 3-dimensional grid box using the Glide tool of the Schrodinger suite.

Table 3. The active site binding regions and their volumes of the CDK14 protein are identified from the sitemap.

Cavity	Binding region	Volume
1	HBacceptor	741.818
2	HBdonar	1133.525
3	Hydrophilic	1889.974
4	Hydrophobic	166.138
5	Metal-binding	0.00
6	Surface	3065.139

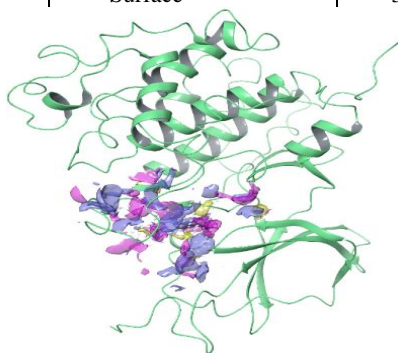


Figure 8. Putative binding site of CDK14 protein recognized from SiteMap. The Hydrogen acceptor region showed with magenta, hydrogen donor region indicated in the light blue, hydrophilic and hydrophobic field represents in red and yellow color respectively. Gray dots indicate the active site of CDK14 protein.

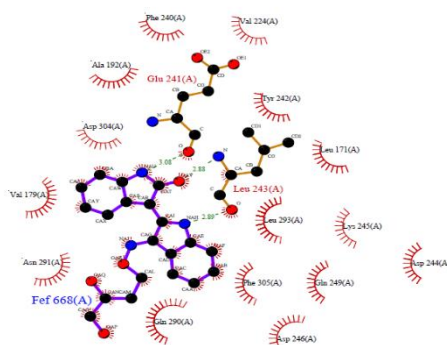


Figure 9. The active site residues of the 3MTL protein interacting with the ligand molecule to identify the active residues of the CDK14 protein.


```

191199
sp|094921|CDK14_HUMAN      INFKTSSTGKESPKVRRHSSPSSPTSPKFGKADSYEKLEK GEGSYATVY
3mtl                        -----METYIKLDK GEGTYATVY
                               :.* **:*:*****:*****

212244
sp|094921|CDK14_HUMAN      KGKSKVNGKLV ALKVIRLQEEETPFTAIREASLLKGLKHANI VLLHDII
3mtl                        KGKSKLTDNVALKEIRL -----PCTAIREVSLLKDLKHANIVTLHDII
                               *****:..:***** **          * *****.****.***** *****

260, 261, 262, 266, 269
sp|094921|CDK14_HUMAN      HTKETLTLV FEV VHT DL C Q YMDKHPGGLHPDNVKLFLFQLLRGLSYIHQR
3mtl                        HTEKSLTLVFEYLDKDLK QYLDDCGNIINMHNVKLFLFQLLRGLAYCHRQ
                               **:::*****:..** **:* . .: .*****:*****:* *::

310, 311, 313, 324, 325
sp|094921|CDK14_HUMAN      YILHRDLKPNL L ISDTGELKLA DF GLARAKSVPSHTYSNEVVT LWYRPP
3mtl                        KVLHRDLKPNL L INERGELKLA DF GLARAK -----TLWYRPP
                               :*****:..:*****:***** *****

sp|094921|CDK14_HUMAN      DVLLGSTEYSTCLDMWVGVCIFVEMIQGVAAFPGMKDIQDQLERIFLVLG
3mtl                        DILLGSTDYSTQIDMWVGVCIFYEMATGRPLFPGS -TVEEQLHFIFRILG
                               *:*:*****:*** :***** ** * . *** :::** ** :**

sp|094921|CDK14_HUMAN      TPNETWPGVHSLPHFKPERFTLYSSKNLRQAWNKLSYVNHAE DLASKLL
3mtl                        TPTEETWPGILSNEEFKTYNPKYRAEALLSHAPRLD --SDGADLLTKLL
                               **.*:*****: * .** .: . * : : * . :* . . . ** :***

sp|094921|CDK14_HUMAN      QCSPKNRLSAQAALSHEYFSDLPRLWEL TDMSSIFTVPNVLQPEAGES
3mtl                        QFEGRNRI SAEDAMKHPF LSLGERIHKLPD TTSIFALKEIQLQKE ----
                               * . :*:**:* *:* * * * :*:* :***: :::** *

```

Figure 10. Identification of active site residues of CDK14 protein obtained by aligning with 3MTL template protein using ClusterX1.2. The active residues of CDK14 and 3MTL are highlighted green color and magenta color, respectively and active residues number shown in cyan color.

3.3. Docking analysis and binding energies.

3.3.1. Docking using GLIDE.

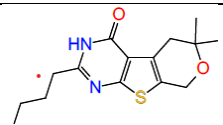
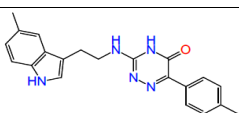
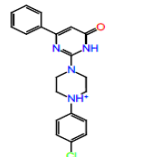
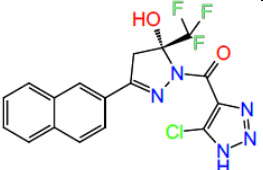
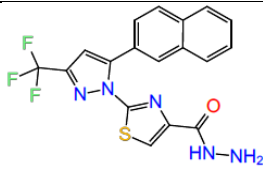
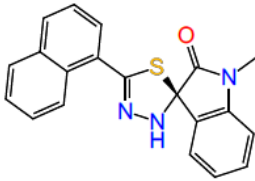
Virtual screening is done with the Glide tool to predict the lead molecules that selectively bind to the biologically active residues of CDK14 protein to inhibit the function of CDK14 protein during cell proliferation. A gridbox is generated with active residues of CDK14 protein using receptor grid generation in glide tool of Schrodinger suite, to obtain produces the good docking interaction at the created binding domain. TOSlab database of 17643 ligand molecules is optimized using ligprep, which gave rise to five low energy of 27253 ligand molecules for virtual screening. This process involves filtering by flexible docking through HTVS, SP, and XP mode, 10% of molecules filtered by HTVS docking mode gave rise to 1300 ligand molecules. 10 % of these ligand molecules were further screened in SP docking mode, which resulted in 130 molecules. 10% of these molecules were further filtered in XP docking mode, which generated 13 lowest energy conformers [47]. The final 13 molecules were prioritized based on glide score, glide energy and XP visualizer analysis of the protein-ligand interaction. The binding free energy of 13 XP output docked complexes were calculated using prime MMGBSA of the Schrodinger suite. The binding free energy explained the affinity of H-bond and pi-sigma interaction between target CDK14 protein and small ligand molecules. Table 4 shows six docked complexes observed in H-bond length below 3.2 suggested that the docked complexes have stable conformation. The binding free energy of docked complexes was falling in the range of -34.27 to -60.23, showing negative dG values indicating the formation of stable complexes [48,49].

3.3.2. Docking using AutoDock.

Autodock program was used for molecular docking for the XP out file of Glide Schrodinger suite of 13 hit molecules. The prioritization of lead molecules was analyzed based on binding energies(-7.16 to -8.75kcal/mol) and protein-ligand interaction.

Surprisingly in both Glide and AutoDock docking tools, the lead molecules have occupied the same binding cavity. Glide score, Glide energy, binding energy(AutoDock) and binding free energy (PrimeMMGBSA) indicate the accuracy of the interaction of protein-ligand molecules to optimize the novel lead molecules shown in Table 4 [50].

Table 4. Glide score, glide energy, binding energies (MM/GBSA and AutoDock) and interaction of the lead molecules with amino acid residues of CDK14 protein.

S.No	Structure	Glide score	Glide energy	PrimeMM-GBSA Complex energy (dG Bind)	Binding energy from AutoDock (kcal/Mol)	Docking complex (protein-ligand interactions)	Bond Distance (Å)
860371 (TOSLab)		-10.63	-33.10	-49.43	-7.16	ASP274:N-M1:O11 M1:H21- :VAL213:O Pi-Pi interactions: TYR212-M1	3.02 1.82
872480 (TOSLab)		-9.95	-41.79	-34.27	-7.20	M2:H28:VAL213:O M2:H29:GLN260:O Pi-Pi interactions: M2:-TYR212	1.64 1.96
858233 (TOSLab)		-9.46	-35.93	-51.63	-8.47	M3:H26: VAL213:O VAL213: N- M3:O14 Pi-sigma interactions: PHE210-M3: H42	2.32 2.81
808781 (TOSLab)		-9.24	-43.57	-60.23	-8.75	M4:H31:-HIS214:O M4:H33- GLN260:O ASP216N-M4:Cl19	1.77 1.64 3.19
23037 (TOSLab)		-9.22	-46.22	-59.73	-7.80	P:LYS258:NZ- M5:O16 P:LYS258:NZ- M5:N24 M5:H33- P:GLU143:O Pi-Sigma interactions: P:PHE210-M5:H40	2.30 3.13 2.06
836986 (TOSLab)		-9.05	-35.43	-48.87	-7.10	M6:H26- P:GLN260:O Pi-Sigma interactions: P:PHE210-M6:H36	2.11

3.4. Pharmacokinetic properties.

The pharmacokinetic properties of all docked molecules were predicted using the Qikprop of the Schrodinger suite. The six docked molecules comply with the Lipinski rule of five and Jorgensen rule of three and follow permissible ranges ages of ADME properties shown

in Table 5. These six molecules have drug-like properties and can be considered potentially novel lead molecules for drug design against CDK14 protein [51].

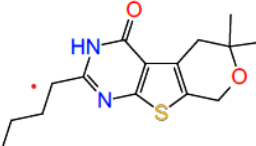
Table 5. The ADME properties of the best lead molecules were predicted from the QikProp of the Schrodinger suite.

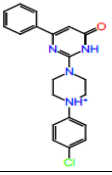
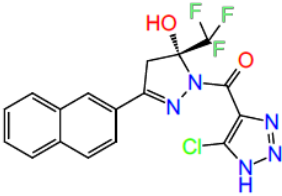
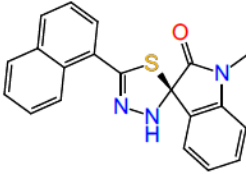
S. No	Mol.Wt	DonarHB	AcceptHB	Qplog Po/W	%Humanoral Absorption	Rule of three	Rule of five
860371 (TOSLab)	292.395	1	4.2	3.229	100%	0	0
872480 (TOSLab)	359.430	3	5	3.685	95.58%	1	0
858233 (TOSLab)	366.849	1	5	4.543	100%	1	0
808781 (TOSLab)	409.754	0	3.5	4.555	100%	1	0
23037 (TOSLab)	403.381	3	5.5	3.690	95.43%	1	0
836986 (TOSLab)	345.48	1	5	3.818	100%	0	0

3.4. Lead optimization.

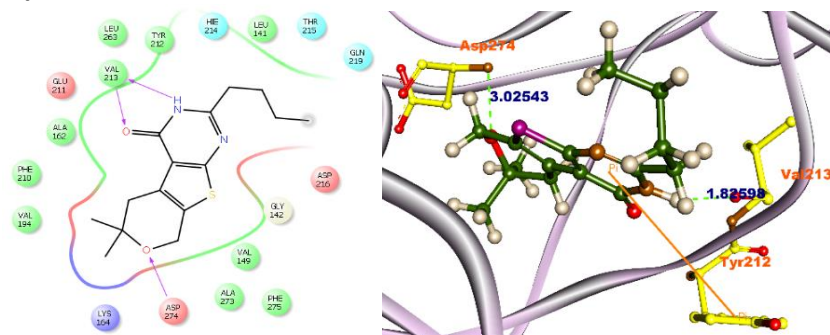
The novel leads molecules 860371, 858233, 808781 and 836986 were finalized based on good binding free energy (PrimeMMGBSA), least binding energy (AutoDock), and interaction protein-ligand complexes with a good percentage of human oral absorption. These four ligand molecules have 100% human oral absorption. The binding energy of lead molecules 860371, 858233, 808781 and 836986 obtained from prime MMGBSA and AutoDock as -49.43, -51.63, -60.23, -48.87 kcal/mol and -7.16, -8.75, -8.47, -7.10 kcal/mol, respectively confirms the formation of most stable protein-ligand complexes which are studied from both docking programs Glide and AutoDock. The 860371 and 858233 lead molecules have a common pyrimidin-4(3H)-one ring that shows specific binding interactions with Val213. The lead molecules 860371 and 858233 are also consistently binding with Val 213. The docked complex of 808781 lead molecules is forming H-bonds with His214 (1.7Å), Asp216 (3.1Å) and Gln260 (1.6Å). All of the above H-bond lengths are observed below 3.2 Å, indicating a lower bond length, with increasing the strength of binding interactions of CDK14 protein-ligand molecules [52]. The lead molecules 860371, 858233, 808781 and 836986 were optimized using binding free energy (MM/GBSA), binding energy (AutoDock), percent human oral absorption, and interaction protein-ligand molecules, as exhibited in Table 6. Figure 11 shows 3-D and 2-D structures of docked molecules 860371, 858233, 808781 and 836986 were visualized by Discovery studio 3.5. The docked molecules 860371, 858233, 808781 and 836986 superimposed at an active site of CDK14 protein revealed that they occupy the same regions of the binding cavity and interact with Val213, His214, Gln260, Phe210 and Tyr212 residues as shown in Figure 12 [53].

Table 6. The best-docked molecules were identified from the TOSLab database by structure-based virtual screening against the CDK14 protein.

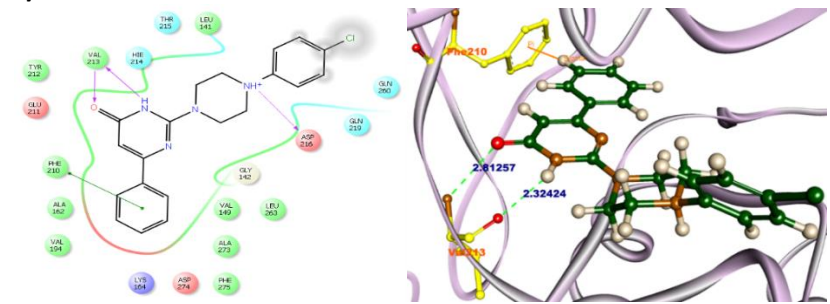
Compound ID	Structure	Protein-ligand binding free energy (dG)	%of Human oral absorption	Binding energy from Autodock (kcal/mol)	Interacting Amino acids
860371 (TOSLab)		-49.43	100%	-7.16	H-bonding Asp274 Val213 Pi-Pi Tyr212

Compound ID	Structure	Protein-ligand binding free energy (dG)	%of Human oral absorption	Binding energy from Autodock (kcal/mol)	Interacting Amino acids
858233 (TOSLab)		-51.63	100%	-8.47	H-bonding Val213 Pi-Sigma Phe210
808781 (TOSLab)		-60.23	100%	-8.75	H-bonding His214 Gln260 Asp216
836986 (TOSLab)		-48.87	100%	-7.10	H-bonding Gln260 Pi-sigma Phe210

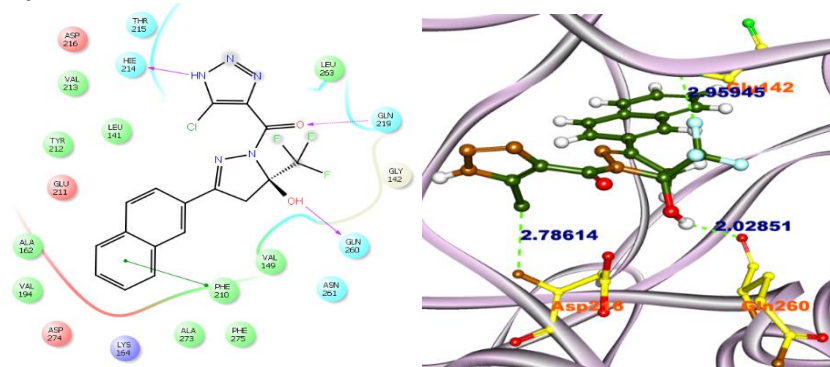
860371(TOSLab)



858233(TOSLab)



808781(TOSLab)



836986(TOSLab)

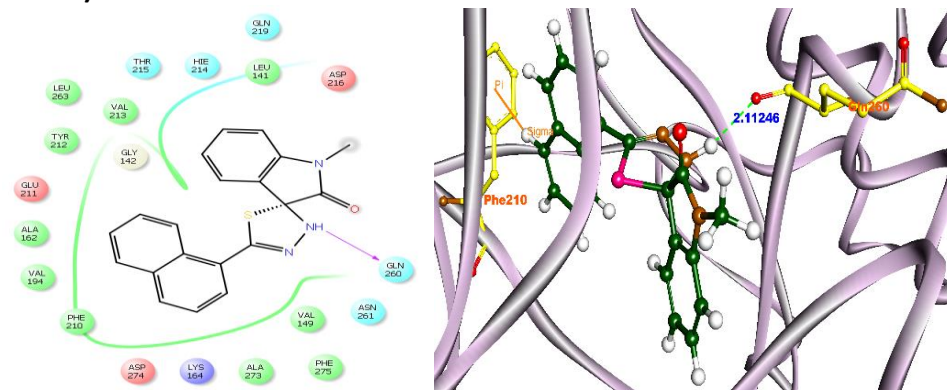


Figure 11. The 3-Dimensional docked poses of best lead molecules 860371, 858233, 808781 and 836986 with active amino acid residues of CDK14 protein are visualized by Discovery studio 3.5 and lead molecules to show that dark green, active residues indicate that yellow, H-binding interaction represents that light green, Pi-sigma, Pi-Cation and Pi-Pi interaction shows that orange color. 2-dimensional interactions of lead molecules were taken from the Schrodinger suite.

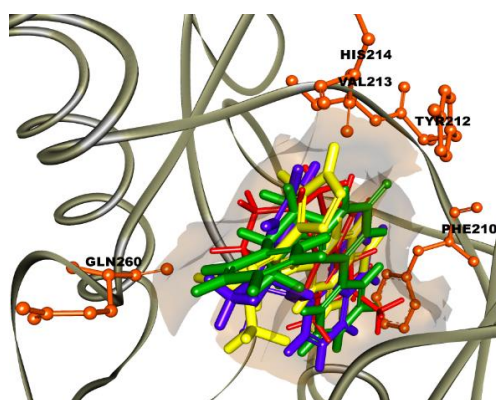


Figure 12. The Best docked molecules 860371(Red), 858233(Green), 808781(Yellow) and 836986(Blue) are superimposed at the active site region of CDK14 protein and active residues (Orange) as visualized using Discovery Studio3.5.

4. Conclusions

In this present study, novel lead molecules were identified against CDK14 protein as cancer therapeutics. The 3D structural details of CDK14 protein were evaluated using comparative modeling. NAMD-VMD simulations carry out the energy minimization. The binding sites were predicted by Sitemap and Ligplot analysis. The potential lead molecules were identified from the Toslab database by performing virtual screening at the binding site of CDK14 protein using Glide and AutoDock tools. The best-docked molecules 860371, 858233, 808781 and 836986, were shortlisted as final potential lead inhibitors based on glide score, % Human oral absorption, binding energies (PrimeMMGBSA and AutoDock) and then drug-likeness properties. The lead molecules 860371, 858233, 808781 and 836986 have 100% human oral absorption and also showing best binding energies (-49.43, -51.63, -60.23, -48.87 kcal/mol from prime MMGBSA and -7.16, -8.47, -8.75, -7.10 kcal/mol from AutoDock) and obeying permissible ADME properties. Both docking studies evidencing that lead molecules 860371, 858233, 808781 and 836986 are showing best binding interactions with Phe210, Tyr212, Val213, His214, Asp216 and Gln260 amino acid residues of CDK14 protein by forming H-bonds, Pi-Pi, Pi-Sigma, which are crucial for inhibition of the overexpression of CDK14 protein. Therefore, the lead candidates 860371, 858233, 808781 and 836986 can be considered as potential inhibitors against CDK14 protein for cancer treatment.

Acknowledgments

The authors also acknowledge the Principal and Head, Department of Chemistry, Nizam College, University College of Science, Osmania University, Hyderabad, for providing facilities to carry out this work.

Conflict of Interest

The authors declare no conflict of interest.

Funding information

The author Revanth Bathula thankful to the Council of Scientific and Industrial Research (CSIR) -INDIA, New Delhi, for providing financial support as SRF (file no: 09/132 (0846)/2015-EMR-I).

References

1. Ferlay, J.; Colombet, M.; Soerjomataram, I.; Mathers, C.; Parkin, D.M.; Piñeros, M.; Znaor, A.; Bray, F. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int. J. Cancer* **2019**, *144*, 1941-1953, <https://doi.org/10.1002/ijc.31937>.
2. Wood, D.J.; Endicott, J.A. Structural insights into the functional diversity of the CDK–cyclin family. *Open Biology* **2018**, *8*, 180112, <https://doi.org/10.1098/rsob.180112>.
3. Peyressatre, M.; Prével, C.; Pellerano, M.; Morris, M.C. Targeting Cyclin-Dependent Kinases in Human Cancers: From Small Molecules to Peptide Inhibitors. *Cancers (Basel)* **2015**, *7*, <https://doi.org/10.3390/cancers7010179>.
4. Kaldis, P.; Pagano, M. Wnt signaling in mitosis. *Dev. Cell* **2009**, *17*, 749-750, <https://doi.org/10.1016/j.devcel.2009.12.001>.
5. Shu, F.; Lv, S.; Qin, Y.; Ma, X.; Wang, X.; Peng, X.; Luo, Y.; Xu, B.-e.; Sun, X.; Wu, J. Functional characterization of human PFTK1 as a cyclin-dependent kinase. *Proceedings of the National Academy of Sciences* **2007**, *104*, 9248, <https://doi.org/10.1073/pnas.0703327104>.
6. Davidson, G.; Shen, J.; Huang, Y.-L.; Su, Y.; Karaulanov, E.; Bartscherer, K.; Hassler, C.; Stannek, P.; Boutros, M.; Niehrs, C. Cell cycle control of wnt receptor activation. *Dev. Cell* **2009**, *17*, 788-799, <https://doi.org/10.1016/j.devcel.2009.11.006>.
7. Baarsma, H.A.; Königshoff, M. ‘WNT-er is coming’: WNT signalling in chronic lung diseases. *Thorax* **2017**, *72*, 746-759, <https://doi.org/10.1136/thoraxjnl-2016-209753>.
8. Arce, L.; Yokoyama, N.N.; Waterman, M.L. Diversity of LEF/TCF action in development and disease. *Oncogene* **2006**, *25*, 7492-7504, <https://doi.org/10.1038/sj.onc.1210056>.
9. Gu, X.; Wang, Y.; Wang, H.; Ni, Q.; Zhang, C.; Zhu, J.; Huang, W.; Xu, P.; Mao, G.; Yang, S. Upregulated PFTK1 promotes tumor cell proliferation, migration, and invasion in breast cancer. *Med. Oncol.* **2015**, *32*, 195, <https://doi.org/10.1007/s12032-015-0641-8>.
10. Boras-Granic, K.; Wyslomerski, J.J. Wnt signaling in breast organogenesis. *Organogenesis* **2008**, *4*, 116-122, <https://doi.org/10.4161/org.4.2.5858>.
11. Liu, M.-h.; Shi, S.-m.; Li, K.; Chen, E.-q. Knockdown of PFTK1 Expression by RNAi Inhibits the Proliferation and Invasion of Human Non-Small Lung Adenocarcinoma Cells. *Oncology Research Featuring Preclinical and Clinical Cancer Therapeutics* **2016**, *24*, 181-187, <https://doi.org/10.3727/096504016X14635761799038>.
12. Zheng, L.; Zhou, Z.; He, Z. Knockdown of PFTK1 inhibits tumor cell proliferation, invasion and epithelial-to-mesenchymal transition in pancreatic cancer. *Int. J. Clin. Exp. Pathol.* **2015**, *8*, 14005-14012.
13. Zhu, J.; Liu, C.; Liu, F.; Wang, Y.; Zhu, M. Knockdown of PFTK1 Inhibits Proliferation, Invasion, and EMT in Colon Cancer Cells. *Oncology Research Featuring Preclinical and Clinical Cancer Therapeutics* **2016**, *24*, 137-144, <https://doi.org/10.3727/096504016X14611963142218>.

14. Ou-Yang, J.; Huang, L.-H.; Sun, X.-X. Cyclin-dependent kinase 14 promotes cell proliferation, migration and invasion in ovarian cancer by inhibiting Wnt signaling pathway. *Gynecol. Obstet. Invest.* **2017**, *82*, 230-239, <https://doi.org/10.1159/000447632>.
15. Zhang, W.; Liu, R.; Tang, C.; Xi, Q.; Lu, S.; Chen, W.; Zhu, L.; Cheng, J.; Chen, Y.; Wang, W.; Zhong, J.; Deng, Y. PFTK1 regulates cell proliferation, migration and invasion in epithelial ovarian cancer. *Int. J. Biol. Macromol.* **2016**, *85*, 405-416, <https://doi.org/10.1016/j.ijbiomac.2016.01.009>.
16. Yang, L.; Zhu, J.; Huang, H.; Yang, Q.; Cai, J.; Wang, Q.; Zhu, J.; Shao, M.; Xiao, J.; Cao, J.; Gu, X.; Zhang, S.; Wang, Y. PFTK1 Promotes Gastric Cancer Progression by Regulating Proliferation, Migration and Invasion. *PLoS One* **2015**, *10*, e0140451, <https://doi.org/10.1371/journal.pone.0140451>.
17. Morgat, A.; Lombardot, T.; Coudert, E.; Axelsen, K.; Neto, T.B.; Gehant, S.; Bansal, P.; Bolleman, J.; Gasteiger, E.; de Castro, E.; Baratin, D.; Pozzato, M.; Xenarios, I.; Poux, S.; Redaschi, N.; Bridge, A.; The UniProt, C. Enzyme annotation in UniProtKB using Rhea. *Bioinformatics* **2020**, *36*, 1896-1901, <https://doi.org/10.1093/bioinformatics/btz817>.
18. Hameduh, T.; Haddad, Y.; Adam, V.; Heger, Z. Homology modeling in the time of collective and artificial intelligence. *Computational and Structural Biotechnology Journal* **2020**, *18*, 3494-3506, <https://doi.org/10.1016/j.csbj.2020.11.007>.
19. Basak, N.; Krishnan, V.; Pandey, V.; Punjabi, M.; Hada, A.; Marathe, A.; Jolly, M.; Palaka, B.K.; Ampasala, D.R.; Sachdev, A. Expression profiling and in silico homology modeling of Inositol pentakisphosphate 2-kinase, a potential candidate gene for low phytate trait in soybean. *3 Biotech* **2020**, *10*, 268, <https://doi.org/10.1007/s13205-020-02260-y>.
20. Webb, B.; Sali, A. Protein Structure Modeling with MODELLER. *Structural Genomics:Methods Mol. Biol.* **2021**, *2199*, 239-255, https://doi.org/10.1007/978-1-0716-0892-0_14.
21. Chandler, P.G.; Broendum, S.S.; Riley, B.T.; Spence, M.A.; Jackson, C.J.; McGowan, S.; Buckle, A.M. Strategies for Increasing Protein Stability. *Methods Mol. Biol.* **2020**, *2073*, 163-181, https://doi.org/10.1007/978-1-4939-9869-2_10.
22. Phillips, J.C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **2005**, *26*, 1781-1802, <https://doi.org/10.1002/jcc.20289>.
23. Salsbury, F.R. Molecular dynamics simulations of protein dynamics and their relevance to drug discovery. *Curr. Opin. Pharm.* **2010**, *10*, 738-744, <https://doi.org/10.1016/j.coph.2010.09.016>.
24. Wiederstein, M.; Sippl, M.J. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res.* **2007**, *35*, W407-W410, <https://doi.org/10.1093/nar/gkm290>.
25. Rose, G.D. Ramachandran maps for side chains in globular proteins. *Proteins: Structure, Function, and Bioinformatics* **2019**, *87*, 357-364, <https://doi.org/10.1002/prot.25656>.
26. Protein preparation wizard, Version 3.3. New York (NY): Schrodinger, LLC; **2016**.
27. Madhavi Sastry, G.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput.-Aided Mol. Des.* **2013**, *27*, 221-234, <https://doi.org/10.1007/s10822-013-9644-8>.
28. LigPrep, version 3.3. New York, NY: Schrodinger, LLC; **2016**.
29. Anderson, A.C. The process of structure-based drug design. *Chem. Biol.* **2003**, *10*, 787-797, <https://doi.org/10.1016/j.chembiol.2003.09.002>.
30. Lanka, G.; Bathula, R.; Dasari, M.; Nakkala, S.; Bhargavi, M.; Somadi, G.; Potlapally, S.R. Structure-based identification of potential novel inhibitors targeting FAM3B (PANDER) causing type 2 diabetes mellitus through virtual screening. *J. Recept. Signal Transduct.* **2019**, *39*, 253-263, <https://doi.org/10.1080/10799893.2019.1660897>.
31. Glide, version 6.1. New York, NY: Schrodinger, LLC; **2016**.
32. Alogheli, H.; Olanders, G.; Schaal, W.; Brandt, P.; Karlén, A. Docking of Macrocycles: Comparing Rigid and Flexible Docking in Glide. *J. Chem. Inf. Model.* **2017**, *57*, 190-202, <https://doi.org/10.1021/acs.jcim.6b00443>.
33. Greenfield, D.A.; Schmidt, H.R.; Skiba, M.A.; Mandler, M.D.; Anderson, J.R.; Sliz, P.; Kruse, A.C. Virtual Screening for Ligand Discovery at the σ 1 Receptor. *ACS Med. Chem. Lett.* **2020**, *11*, 1555-1561, <https://doi.org/10.1021/acsmchemlett.9b00314>.
34. Prime. New York, NY: Schrodinger, LLC; **2016**.

35. Cosconati, S.; Forli, S.; Perryman, A.L.; Harris, R.; Goodsell, D.S.; Olson, A.J. Virtual screening with AutoDock: theory and practice. *Expert Opinion on Drug Discovery* **2010**, *5*, 597-607, <https://doi.org/10.1517/17460441.2010.484460>.
36. Bitencourt-Ferreira, G.; Pintro, V.O.; de Azevedo, W.F. Docking with AutoDock4. *Methods Mol. Biol.* **2019**, 2053,125-148, https://doi.org/10.1007/978-1-4939-9752-7_9.
37. QikProp. New York, NY: Schrodinger, LLC; **2016**.
38. Guo, W.; Li, Z.; Yuan, M.; Chen, G.; Li, Q.; Xu, H.; Yang, X. Molecular Insight into Stereoselective ADME Characteristics of C20-24 Epimeric Epoxides of Protopanaxadiol by Docking Analysis. *Biomolecules* **2020**, *10*, <https://doi.org/10.3390/biom10010112>.
39. Haddad, Y.; Adam, V.; Heger, Z. Ten quick tips for homology modeling of high-resolution protein 3D structures. *PLoS Comp. Biol.* **2020**, *16*, e1007449, <https://doi.org/10.1371/journal.pcbi.1007449>.
40. Isa, M.A. Homology modeling and molecular dynamic simulation of UDP-N-acetylmuramoyl-l-alanine-d-glutamate ligase (MurD) from Mycobacterium tuberculosis H37Rv using in silico approach. *Comput. Biol. Chem.* **2019**, *78*, 116-126, <https://doi.org/10.1016/j.compbiolchem.2018.11.002>.
41. Bathula, R.; Lanka, G.; Muddagoni, N.; Dasari, M.; Nakkala, S.; Bhargavi, M.; Somadi, G.; Sivan, S.K.; Rajender Potlapally, S. Identification of potential Aurora kinase-C protein inhibitors: an amalgamation of energy minimization, virtual screening, prime MMGBSA and AutoDock. *J. Biomol. Struct. Dyn.* **2020**, *38*, 2314-2325, <https://doi.org/10.1080/07391102.2019.1630318>.
42. Paxman, J.J.; Heras, B. Bioinformatics Tools and Resources for Analyzing Protein Structures. *Methods Mol. Biol.* **2017**, 1549, 209-220, https://doi.org/10.1007/978-1-4939-6740-7_16.
43. Schiffrin, B.; Radford, S.E.; Brockwell, D.J.; Calabrese, A.N. PyXlinkViewer: A flexible tool for visualization of protein chemical crosslinking data within the PyMOL molecular graphics system. *Protein Sci.* **2020**, *29*, 1851-1857, <https://doi.org/10.1002/pro.3902>.
44. Martin, D.R.; Dinpajoo, M.; Matyushov, D.V. Polarizability of the Active Site in Enzymatic Catalysis: Cytochrome c. *The Journal of Physical Chemistry B* **2019**, *123*, 10691-10699, <https://doi.org/10.1021/acs.jpbc.9b09236>.
45. Halgren, T.A. Identifying and Characterizing Binding Sites and Assessing Druggability. *J. Chem. Inf. Model.* **2009**, *49*, 377-389, <https://doi.org/10.1021/ci800324m>.
46. Bhargavi, M.; Vhora, N.; Lanka, G.; Somadi, G.; Kanth, S.S.; Jain, A.; Potlapally, S.R. Homology modelling and virtual screening to explore potent inhibitors for MAP2K3 protein. *Struct. Chem.* **2021**, *32*, 1039-1051, <https://doi.org/10.1007/s11224-020-01667-w>.
47. Rastelli, G.; Pinzi, L. Refinement and Rescoring of Virtual Screening Results. **2019**, *7*, <https://doi.org/10.3389/fchem.2019.00498>.
48. Rajagopal, K.; Arumugasamy, P.; Byran, G. In-silico Drug Design, ADMET Screening, MM-GBSA Binding Free Energy of Some Chalcone Substituted 9-Anilinoacridines as HER2 Inhibitors for Breast Cancer. *International Journal of Computational and Theoretical Chemistry* **2019**, *7*, 6, <https://doi.org/10.2174/2589977511666190912154817>.
49. Wang, Z.; Wang, X.; Li, Y.; Lei, T.; Wang, E.; Li, D.; Kang, Y.; Zhu, F.; Hou, T. farPPI: a webserver for accurate prediction of protein-ligand binding structures for small-molecule PPI inhibitors by MM/PB(GB)SA methods. *Bioinformatics* **2019**, *35*, 1777-1779, <https://doi.org/10.1093/bioinformatics/bty879>.
50. Bhargavi, M.; Sivan, S.K.; Potlapally, S.R. Identification of novel anti cancer agents by applying insilico methods for inhibition of TSPO protein. *Comput. Biol. Chem.* **2017**, *68*, 43-55, <https://doi.org/10.1016/j.compbiolchem.2016.12.016>.
51. Adinehbeigi, K.; Shaddel, M.; Khalili, S.; Zakeri, A. Suramin could block the activity of Arabinono-1, 4-lactone oxidase enzyme from Leishmania donovani: structure-based screening and molecular dynamics analyses. *Trans. R. Soc. Trop. Med. Hyg.* **2020**, *114*, 162-172, <https://doi.org/10.1093/trstmh/trz091>.
52. Wang, E.; Sun, H.; Wang, J.; Wang, Z.; Liu, H.; Zhang, J.Z.H.; Hou, T. End-Point Binding Free Energy Calculation with MM/PBSA and MM/GBSA: Strategies and Applications in Drug Design. *Chem. Rev.* **2019**, *119*, 9478-9508, <https://doi.org/10.1021/acs.chemrev>.
53. Liu, S.; Zhou, L.-H.; Wang, H.-Q.; Yao, Z.-B. Superimposing the 27 crystal protein/inhibitor complexes of β -secretase to calculate the binding affinities by the linear interaction energy method. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 6533-6537, <https://doi.org/10.1016/j.bmcl.2010.09.050>.