

Homology Modeling, Virtual Screening, Prime-MMGBSA, AutoDock-Identification of Inhibitors of FGR Protein

Narasimha Muddagoni ¹ , Revanth Bathula 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, saritarajender@osmania.ac.in;

Scopus Author ID 55317928000

Received: 11.11.2020; Revised: 4.12.2020; Accepted: 5.12.2020; Published: 8.12.2020

Abstract: Carcinogenesis is a multi-stage process in which damage to a cell's genetic material changes the cell from normal to malignant. Tyrosine-protein kinase FGR is a protein, member of the Src family kinases (SFks), nonreceptor tyrosine kinases involved in regulating various signaling pathways that promote cell proliferation and migration. FGR protein is also called Gardner-Rasheed Feline Sarcoma viral (v-fgr) oncogene homolog. FGR, FGR protein has an aberrant expression upregulated and activated by the tumor necrosis factor activation (TNF), enhancing the activity of FGR by phosphorylation and activation, causing ovarian cancer. In the present study, 3D structure of FGR protein is built by using comparative homology modeling techniques using MODELLER9.9 program. Energy minimization of protein is done by NAMD-VMD software. The quality of the protein is evaluated with ProSA, Verify 3D and Ramchandran Plot validated tools. The active site of protein is generated using SiteMap and literature Studies. In the present study of research, FGR protein was subjected to virtual screening with TOSLab ligand molecules database in the Schrodinger suite, to result in 12 lead molecules prioritized based on docking score, binding free energy and ADME properties. These lead molecules are considered as a new policy against ovarian cancer.

Keywords: homology modeling; NAMD-VMD:Schrodinger suite; prime-MMGBSA; AutoDock; virtual screening.

© 2020 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 the most common human genetic disease. Cancer is the second leading cause of death globally. Ovarian cancer is 7th leading cause of cancer-related death of humans [1] is a most commonly causing cancer in African countries. The overexpression of FGR protein result causes ovarian cancer [2,3], FGR protein belongs to the Src family of protein tyrosine kinases (SFks), Src family kinases are nonreceptor tyrosine kinases [4]. Cell migration is carried out by FGR protein triggered by adhesion to $\beta 2$ integrin in the signal transduction pathway [5]. The normal function of SFks proteins regulates a wide range of cellular events as well as specialized functions such as immune responses, cell adhesion cell movement [6], $\beta 2$ integrins phosphorylates and activates polymorphonuclear neutrophils (PMNs) as well as FGR. In contrast, TNF involved in this pathway further enhances phosphorylation of FGR protein, which results in increased activation of FGR protein predominantly causing ovarian cancer progression is shown in Figure 1. Present work involves computer-based homology modeling technique to build 3D structure of FGR. The 3D model of generated protein was

validated using NAMD-VMD software, validated to check the quality of FGR protein using various validation servers, such as ProSA, Ramchandran plot, and verify 3D. The active site of protein is predicted from SiteMap. TOSLab ligand data bank was used as small ligand molecules for docking. The protein is prepared in a protein preparation wizard. Its active site is identified and subjected to virtual screening in Schrodinger suite and AutoDock to identify new lead molecules, prioritized based on docking score, binding energies, and ADME properties. These lead molecules are considered as potential inhibitors of FGR protein.

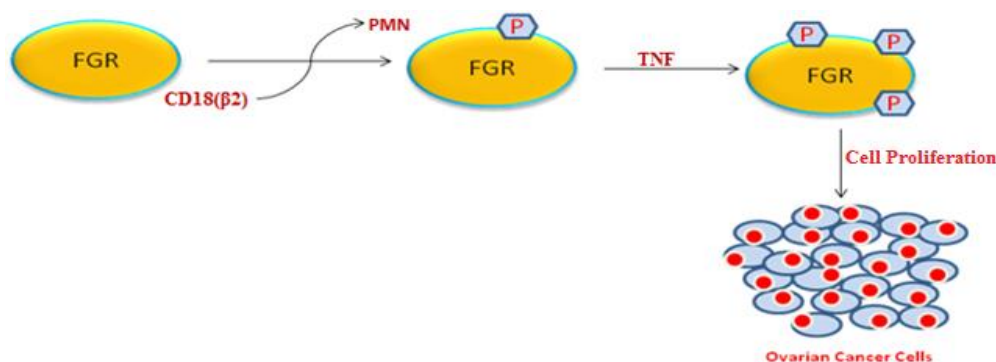


Figure 1. Schematic diagram of FGR protein causes ovarian cancer. CD18 ($\beta 2$) integrin is responsible for activation and phosphorylation of FGR protein in which polymorphonuclear neutrophil (PMN), phosphorylation enhanced by the tumor necrosis factor (TNF), leading to cell proliferation.

2. Materials and Methods

2.1. Homology modeling.

The term homology modeling, also called comparative modeling or sometimes template-based modeling (TBM), refers to modeling a 3D structure of a protein using a known experimental structure of a homologous template. Template selection is based on the query coverage and statistical measure of E-value [7]. The template selection was made using different servers, such as BLAST and Jpred [8,9], followed by sequence alignment between template and target protein operated in clustalW tool [10]. The 3D structure is built for FGR protein using modeler 9.9 [11]. Twenty models were generated. One with the least modeler objective function value of protein is selected for further future work.

2.2. Energy minimization of FGR protein.

The 3D model of FGR protein was energy minimized by using Visual Molecular Dynamics-Nano Scale Molecular Dynamics (VMD-NAMD)[12], FGR protein contains abnormal steric clashes, produced by improper bond angles, the bond length between amino acid residues of the protein, The initial 3d model of FGR was subjected for refinement in simulations environment with CHARMM, AMBER as a force field and solvated in a water box under a Periodic boundary environment with a size of 2.4 Å, to minimize the energy of the protein [13].

2.3. Validation of protein.

Quality of energy minimized protein was validated with different server tools, such as ProSA, Ramchandran plot, and Errat. To evaluate the quality of the protein, ProSA is a tool widely used to check 3D models of protein structures for potential errors, and the z- score indicates the quality of protein [14]. Ramachandran plot helps to visualize energetically

allowed regions for backbone dihedral angles ψ against ϕ of amino acid residues in protein structure [15]. ERRAT plot analyzes the statistics of non-bonded interactions between different atom types and comparison with statistics from highly refined structures.

2.4. Protein preparation.

The protein preparation wizard is used for the optimization of the FGR Protein structure [16]. Protein is taken as a raw stage in which hydrogen atoms are missing with incorrect bond orders. In protein preparation, wizard hydrogens are added, bond order is assigned unwanted Water molecules are removed. The energy minimization is carried out using protein preparation wizard at OPLS 2005 force field, until a stage where a Root Mean Square Deviation (RMSD) of 0.3 Å is reached [17].

2.5. Ligand preparation.

Ligand preparation was carried out by the LigPrep module of Schrodinger suite in Maestro 9.0.111 [18]. Ligand molecules are taken from the TOSLab ligand databank, containing 17643 small molecules. These ligand molecules were subjected to LigPrep, resulting in low energetic conformers for the virtual screening in the LigPrep module of the Schrodinger suite. Each ligand molecule generates different conformers based on ionic state, stereochemistry, and ring conformations in the module. Those with low energy output conformers are used for virtual screening to identify lead molecules [19].

2.6. Active site identification.

The active site region of the protein is a crucial step in structure-based drug design. Identification of the binding region is based on the Schrodinger suite's literature studies and sitemap tool [20]. The ligand N-(5-Chloro-1,3-Benzodioxol-4-Yl)-7-[2-(4-Methylpiperazin-1-Yl)ethoxy]-5-(Tetrahydro-2h-Pyran-4-Yloxy) quinazolin-4-Amine, binding to Amino acid residues of a template (2H8H) protein were taken as active site residues and are manually correlated to obtain active site residues of FGR protein through the pair sequence alignment in ClustalX2.1 [21]. Grid is generated from Glide tool using active site amino acid residues of the FGR protein, with dimensions of 32×32×32Å and 80×80×80Å. This box is used as a favorable binding cavity for screening studies [22]. SiteMap is a fast, accurate, and practical binding site identification tool; it generates site score, which is useful in identifying effective binding sites on the surface of the protein. Evaluation of the SiteMap results helps locate the binding sites with a high degree of accuracy, showing hydrogen bond donors, the hydrogen bond acceptors, hydrophobic and hydrophilic regions. Glide tool in Schrodinger suite was used for molecular docking studies [23].

2.7. Virtual screening.

Virtual screening is a prominent tool for the identification of novel ligand molecule discovery. Grid generated at the active region of FGR protein is used for docking using ligand molecules for the inhibition activity of FGR protein. This process proceeds through virtual screening, including protein preparation wizard, ligand preparation, and docking [24]. This process involves three filtration stages, 1) bulk LigPrep out file and protein is passed through the first stage, HTVS (High Throughput Virtual Screening) docking mode in the docking funnel results in the thoroughness of the final torsional refinement; 2) The conformers are

passed through the next stage SP (standard precision) docking mode where ligands are flexible, filters unstable ligands, and at the third stage performs XP (extra precision) docking mode. Under each stage, 10% of molecules are filtered [25,26], and finally, the best-docked molecules were prioritized by different scoring functions.

2.8. Prime MMGBSA.

The binding free energy of FGR protein and ligand complex was carried out in Prime MMGBSA (Molecular Mechanics Generalized Born Surface Area). The binding free energies of receptor-ligand complexes were calculated using the Prime (27) module of Schrodinger suite with the OPLS_2005 force field. The binding free energy of a ligand (L) to a protein (P) to form the complex (PL) is obtained as the difference.

$$\Delta G_{\text{binding}} = \Delta G_{(\text{complex})} - \Delta G_{(\text{Protein})} - \Delta G_{(\text{Ligand})}$$

Where G_{Binding} is the binding free energy; whereas G_{complex} , G_{protein} and G_{ligand} represents the free energy of complex, protein, ligand, respectively.

2.9. ADME properties.

Pharmacokinetic properties of the resulting screened XP out a file of top docked ligand molecule are identified in QikProp module of Schrodinger suite [28]. The identified docked lead ligand molecules obtained by docking obey the ADME properties. The ADME properties of new lead molecules are essential for the development of an effective druggable molecule. The ligands' ADME properties were all in permissible ranges and are considered new potential inhibitors of the FGR protein.

2.10. AutoDock.

AutoDock 1.5.6 is one of the most cited docking software in the research community. It is an excellent tool to recognize docking affinity for individual ligand molecules identified with protein [29]. The screened XP out file ligand molecules obtained from the Schrodinger suite was converted to PDBQT file format. Grid and docking parameter files were given as GPF files and DPF files in AutoDock. Different conformations of ligands were docked at the active region of the FGR protein following the Lamarckian type of genetic algorithm.

3. Results and Discussion

3.1. Structure evaluation and validation.

The 3D structure of FGR protein is not reported in PDB (protein data bank) by NMR spectroscopy and X crystallography experimental studies [30]. The 3D structure of FGR protein was built using the comparative homology model technique, based on the sequence similarity, E-value, and identity of the known structure of the template (2H8H) protein [31]. The amino acid sequence of FGR protein is retrieved from UniProt with accession ID P09769 and submitted to BLAST, and Jpred servers. These servers give the known proteins as templates with similar structures, resulting in the best match template protein as 2H8H show in Table 1.

Pairwise sequence alignment of target FGR protein with structurally similar 2H8H template was carried out in clustalX1.2. The alignment between FGR protein and template

2H8H protein reveals the identical and conserved residues in FGR protein. The sequence-structure relationship is visualized with discovery studio3.5, shown in Figure 2.

Table 1. Template selection from Blast and J Pred.

S.No	Name of the protein database search server	Parameters considered for template selection	E Score Value	Protein PDB Code
1	BLAST	Sequence similarity	3e-166	2H8HA
2	Jpred3	Three-state (a-helix, b-strand, and coil) prediction of secondary structure and solvent accessibility	e-131	2H8HA

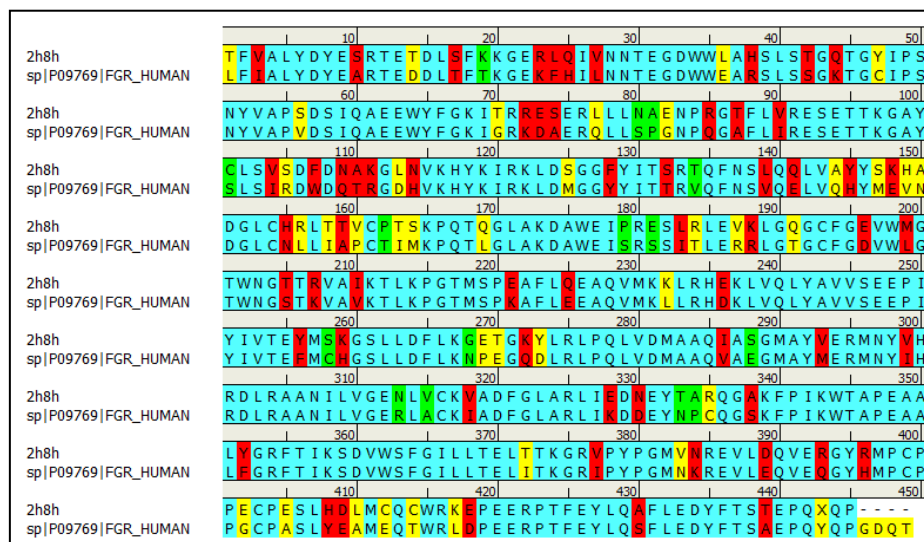


Figure 2. Pairwise sequence alignment of FGR protein with template 2H8HA. Identical residues are shown in cyan color, strongly similar residues with red color and weakly similar residues with yellow color, non-matching similar residues with a green color.

The alignment file is further used as an input file to MODELLER 9.9 [32] to build homology models. Twenty models of FGR protein were generated. The model with the least probability density function model (2646.37 KJ/mol) is taken for further studies. PyMOL, shown in Figure 3, visualize the 3D model of FGR protein.

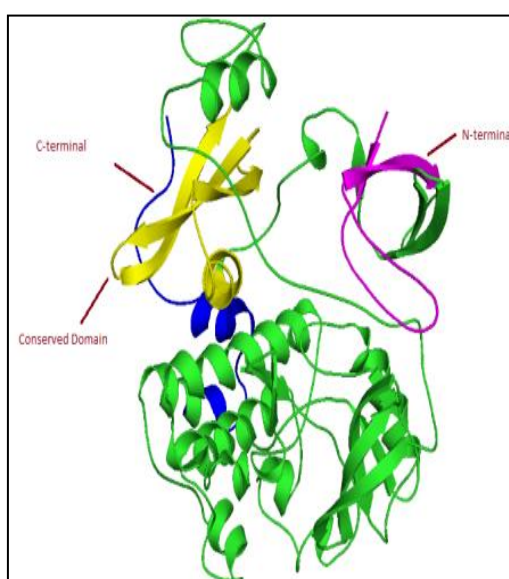


Figure 3. The energy minimized 3-dimensional structure of FGR protein is visualized using PyMOL software. The 3-dimensional protein contains 19-helices, 14 helix-helix interacts, 47 turns, and the N-terminal and C-terminal are exhibited with pink and blue color, respectively. The conserved domain is indicated in yellow color.

Energy minimization of protein was carried out using the NAMD-VMD tool to reduce the steric clashes in FGR protein. Solvation of protein is carried out in all directions with a 10 Å layer of water. Energy minimization was performed in order to stabilize the protein in NAMD_2.9_Win32-multicore package of NAMD-VMD module, which uses AMBER, CHARMM force fields by applying simulations. Stabilized protein was analyzed based on RMSD value 3D structure of FGR protein exhibited RMSD value of around 1.0-1.1 Å with time stages (1072–1828Ps) shown in Figure 4.

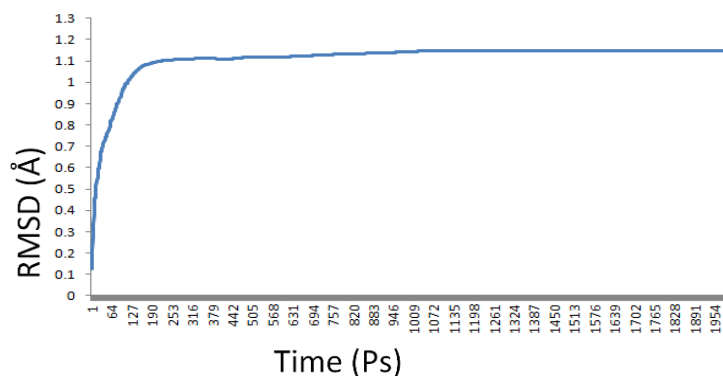


Figure 4. Energy minimization of FGR protein carried out using NAMD-VMD software. The graph displays the RMSD values of various energy states at different time steps. The FGR protein was stabilized at a constant RMSD value of 1.1 Å with time steps 1072–1828 Ps.

The FGR structure was stabilized by removing bad interaction and steric clashes, hydrogen atoms were added, water molecules are removed. This stabilized protein analyzed based on RMSD value 1.1 Å, considered for molecular docking. This process proceeds in the Schrodinger suite's protein preparation wizard module, and the optimized structure was used for further work.

The stabilized protein structure is validated with different standard validation protocols such as ProSA, Ramchandran plot, and verify 3D. ProSA is generating two models, i) overall model quality and ii) local model quality. Overall model quality z-score value -10.76 indicates the protein's quality by comparing with the experimentally determined proteins deposited in protein data bank by X-ray crystallography (light blue) and NMR spectroscopy (dark blue) shown in Figure 5a.

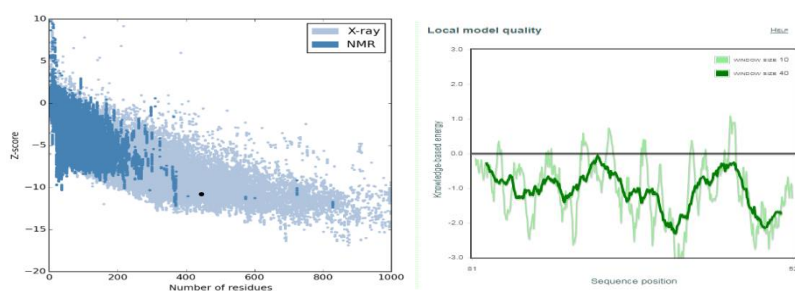


Figure 5. (a) The z-score of the FGR protein obtained from ProSA gives the value as -10.76. The black spot corresponds to the good quality of FGR protein. 5(b) ProSA analysis of the local model quality showed all the amino acid residues in the negative region of indicating a good model quality of the 3D model of FGR protein.

Local model quality of FGR protein has its amino acid residues falling below the baseline in the negative region, shown in Figure 5b, indicating the generated model is highly reliable. Ramchandran plot of FGR protein revealed 93% of amino acid residues in the

favorable region, 5% residues in the additionally allowed region, indicating the good stereochemical quality of FGR protein, shown in Figure 6.

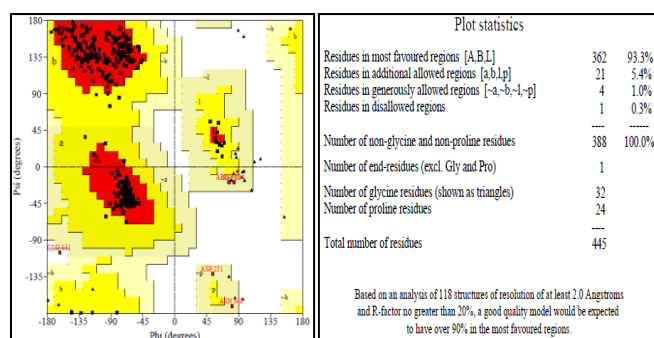


Figure 6. Ramachandran plot of FGR protein. The red area indicates the most favored amino acid residues; the brown area indicates the additional allowed regions, and the yellow area indicates the generously allowed regions. The Ramachandran plot of protein has 98.4% of residues in the favored region, indicating a good quality protein model.

The verify 3D is a well-organized web-based server to correct the 3d model by estimating compatibility with its own 1D amino acid sequence to its 3D profile. Verify3D plot of FGR protein reveals that 96.63% of residues had an average 3D-1D score ≥ 0.2 within the permeable ranges shown in Figure 7. The Protein 3D model is visualized by PyMOL software. FGR protein consists of 17 helices, 15 helix-helix interactions, 51 beta turns, and 3 gamma turns [33] shown in Table 2.

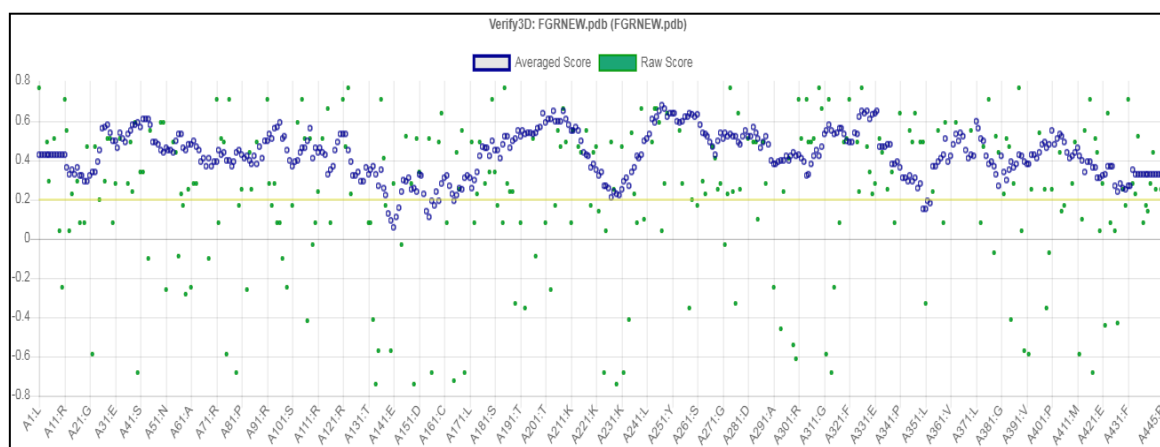


Figure 7. Verify 3d plot of FGR. The graph showing that 96.63% of residues of FGR had an averaged 3D-1D score within permeable ranges. At least 80% of amino acids have scored ≥ 0.2 in the 3D-1D profile.

Table 2. Secondary structure information on the FGR protein, identified by PDBSum server.

S.No	Start	End	Number of residues	Length in Å	Sequence of α helices
1	Arg155	Leu162	8	12.10	RRESERLL
2	Leu223	Lys232	10	15.95	LQQLVAYYSK
3	Pro304	Lys316	13	20.32	PEAFLQEAQVMKK
4	Leu346	Leu350	5	8.08	LLDFL
5	Gly352	Gly355	4	5.60	GETG
6	Leu360	Arg379	20	29.97	LPQLVDMAAQIASGMAYVER
7	Asn414	Thr417	4	6.43	NEYT
8	Pro431	Tyr436	6	8.46	PEAALY
9	Ile441	Leu455	15	22.74	IKSDVWSFGILLTEL
10	Asn468	Glu476	9	15.10	NREVLQVE
11	Leu360	Cys498	10	15.47	ESLHDLMCQC
12	Asn414	Glu517	9Å	13.75	FEYLQAFLE

3.2. Active site identification of FGR protein.

The active site plays a pivotal role in structure-based drug design. The active site residues of target FGR protein are identified from template protein by comparing its residues with active site residues of template protein (2H8H), Active site residues of template were obtained from Lig plot shown in Figure 8.

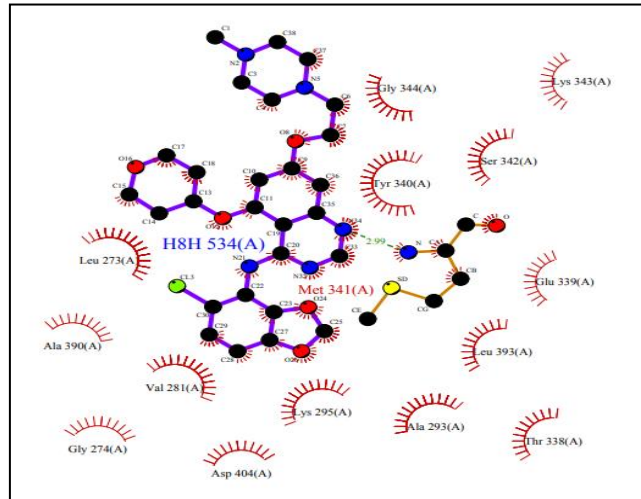


Figure 8. The Ligplot diagram exhibiting the template residues (2H8H), interacting with ligand compound 1-Tert-Butyl-3-(Naphthalen-1-Yl)-1h-Pyrazolo [3, 4- D] pyrimidin-4-Amine . All these amino acid residues were considered for the identification of an active site of FGR protein.

The amino acid residues of template 2H8H interacting with ligand molecule ligand are N-(5-Chloro-1,3-Benzodioxol-4-Yl)-7-[2-(4-Methylpiperazin-1-Yl)ethoxy]-5-(Tetrahydro-2h-Pyran-4-Yloxy) quinazolin-4-Amine, were Leu273, Gly274, Val281, Ala293, Lys295, Thr338, Glu339, Tyr340, Leu393, Ser342, Lys343, Gly344, Ala390, Leu393. These residues of 2H8H template were manually correlated to an amino acid sequence of FGR protein by the ClustalX1.2, which resulted in Leu269, Gly270, Val277, Ala289, Lys291, Thr334, Glu335, Tyr336, Leu337, Ser338, Lys339, Gly340 as active site residues, shown in Figure 9.

CLUSTAL 2.1 multiple sequence alignment

```

2h8h      CLSVSDFDNAKGLNVKHYKIRKLDSGGFYITSRTQFNSLQQLVAYYSKHDGLCHRLTTV
sp|P09769|FGR_HUMAN SLSIRDWDQTRGDHVKHYKIRKLDMGGYYITTRVQFNSVQELVQHMEVNDGLCNLLIAP
.*.*.*:*:*:*.*:*:*:*:*:*:*:*:*:*:*:*.*:*:*.*:*:*:*:*:*:*:*:*:*:*:*:*

2h8h      CPTS KPQTQGLAKDAWEIPRESLRLEVK 269, 270 LGQGC FGEV 277 WMGTWNGTTRV 289, 291 AIK TLKPGTMSP
sp|P09769|FGR_HUMAN CTIMK PQT LGLAKDAWEISRSSITLERR 277 LGTGCFGD 277 WLTWNGSTKVA 277 YK TLKPGTMSP
* . *** ** ** ** ** ** ** ** * . * : * : * * * : * : * : * : * : * : * : * : * : * : * : * : *

2h8h      EAF LQEAQVMKKLRHEKLVQLYAVVSEEPYIV 334 335 336 337 338 339 340 TEYMSKGSLLDFLKGETGKYLR LRPQLV
sp|P09769|FGR_HUMAN KAF LEEAQVMKLLRHKLVQLYAVVSEEPYIV 334 335 336 337 338 339 340 TEFMCHGSL LDFLKNPEGDLRLPQLV
.*.*.*:*:*:*.*:*:*:*:*:*:*:*:*:*:*:*.*:*:*.*:*:*:*:*:*:*.*:*:*:*.*:*

2h8h      DMAAQIASGMAYVERMNYVHRDLRAANILVGENLVCKVADFLGRLARLIEDNEY TARQGAKF
sp|P09769|FGR_HUMAN DMAAQVAEGMAYMERMNYIHRDLRAANILVGERLACKIADFLGRLARLIKDEYNPCQGSKF
*****.*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*
  
```

Figure 9. Prediction of active site residues of FGR protein from an alignment of the template 2H8H with FGR protein. Highlighted colors are manually correlated between a template and the target protein. FGR protein residues are shown in pink color. Template residues are shown in yellow color, and active residue numbers are shown in cyan color.

The conserved domain of FGR protein is taken from the BLAST server shown in Figure 10. <https://biointerfaceresearch.com/>

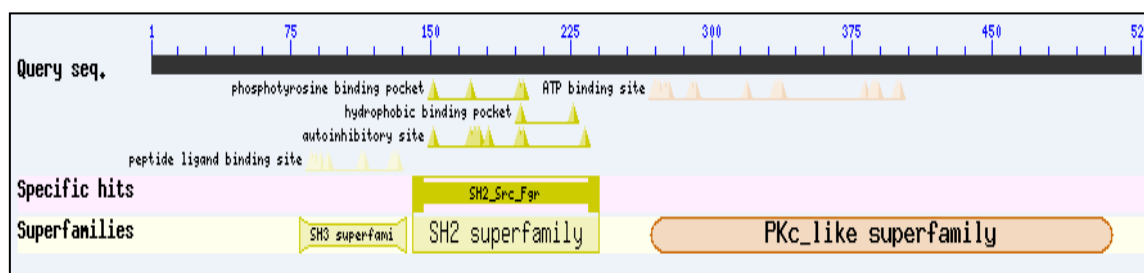


Figure 10. Conserved domain of FGR protein obtained from blast server. Query sequence submitted to BLAST server. The active site region is found to be between 150 to 230 amino acid residues of FGR protein.

The putative binding receptor of FGR protein was identified in the SiteMap module in Schrodinger suite using the OPLS_2005 force field. It also provides the nature of the cavities such as hydrogen bond acceptors, hydrogen bond donors, hydrophilic region, and hydrophobic region, depicting the nature of the active site binding region [34] shown in Figure 11, SiteMap also gives the volume of binding regions and binding modes illustrated shown in Table 3. Three Dimensional grid box was generated from Glide tool of Schrodinger software, using active amino acid residues of FGR protein which were Leu269, Gly270, Val277, Ala289, Lys291, Thr334, Glu335, Tyr336, Leu337, Ser338, Lys339, Gly340, with the box size 80×80×80 Å and the centroid of 32 Å x 32 Å x 32 Å dimension of FGR protein shown in Figure 12.

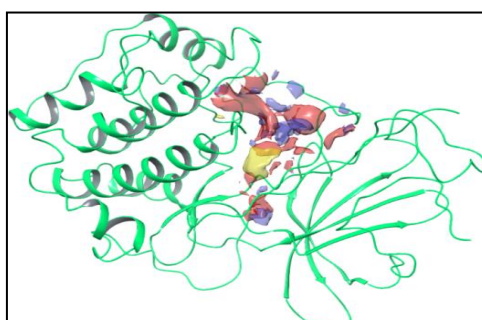


Figure 11. The active site domain of the FGR protein is identified from sitemap in Schrodinger tool Red color depicts the hydrogen donor; light blue indicates hydrogen acceptors, hydrophilic and hydrophobic are showed spring green and yellow respectively.

Table 3. Putative active site binding region of FGR protein was identified from the SiteMap tool of Schrodinger suite.

Cavity	Binding region	Volume(Å)
1	HB acceptor	521.102
2	HB donor	844.393
3	Hydrophilic	1362.17
4	Hydrophobic	245.94
5	Metal binding	0.000
6	Surface	2132.05

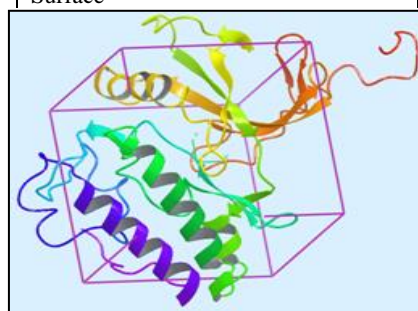
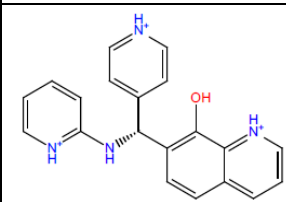
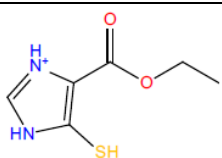


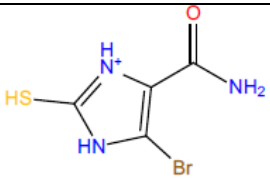
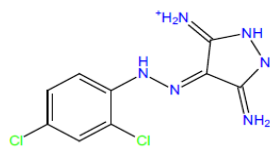
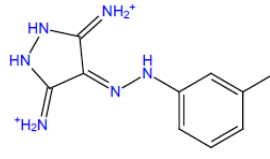
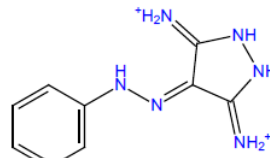
Figure 12. Grid is generated from Glide module of Schrodinger software, FGR active residues used for Grid generation.

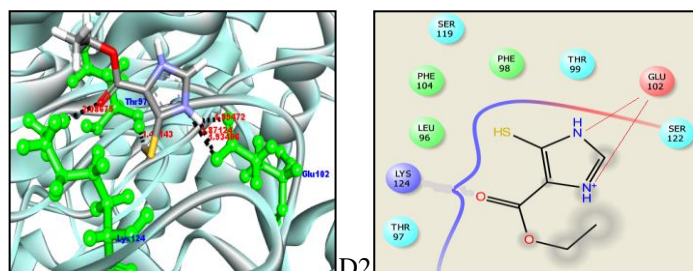
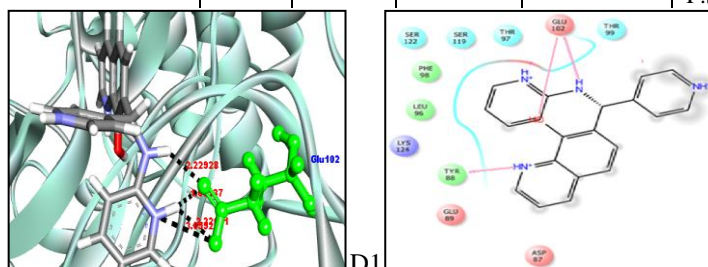
3.3. Docking studies.

Identification of potential ligand molecules against FGR protein is based on the structure-based virtual screening. It is a computer-based technique. It analyzes the binding affinity and interactions between protein-ligand complexes [35]. A library of chemical structures of compounds, TOSLab database containing 17643 Ligand molecules, were subjected to ligPrep to give 27253 low energy conformers as out file [36]. This out file is used as an input database for virtual screening in maestro 9.0.111, Schrodinger suite. TOSLab ligand library is used for the docking with a Grid of protein, built in the Glide module of Schrodinger suite. Various filtration methods of screening such as HTVS, SP and XP modes were used. In HTVS (high throughput virtual screening) mode, 1232 ligand molecules were produced. These molecules were passed through the SP (Standard precession) mode, which produced 123 (10%) molecules, which were further passed through the XP (extra precession) mode giving 12 (10%) best-docked molecules as a result of screening as XP outfile, shown in Table 4. The 3d binding orientations of lead structures were prioritized based on scoring functions as D1, D2, D3, D4, D5, and D6, which are visualized in discovery studio3.5, and then corresponding 2d interaction diagrams from Schrodinger suite are shown in Figure 13. Estimated binding free energy of docked molecules was obtained from Prime-MMGBSA. Binding strength was determined by the complex binding free energy, and the lowest binding free energy conformers had had the highest binding affinity. PrimeMMGBSA reveals stable docked complexes, revealing binding free energies, ranging between -25.63 to -17.03 dG (kcal/mol). The prioritized lead molecules obtained from the Schrodinger suite consistently are interacting with Glu102, THR97, and SER122 residues H-bond interactions [37].

Table 4. The final lead molecules are arranged in priority order with their, Glide score, Glide energy, prime-MMGBSA, AutoDock and interactions, bond distance.

S.No	Structure	Glide score	Glide energy	PrimeM M-GBSA complex energy(d G bind)	Binding Energy from AutoDock(Kcal/mol)	Protein-ligand interactions	Bond Distance (Å)
D1		-4.47	-29.75	-24.14	-6.3	P:GLU102:OE2 R1:N41 :D1:H24 P:GLU102:OE1 :D1:H42 P:GLU102:OE1 :D1:H42 P:GLU102:OE2	- 3.05 2.22 1.60 2.32 - - -
D2		-4.87	-19.14	-20.09	-4.3	P:GLU102:OE2 :D2:N3 P:LYS124:HZ1 :D2:O6 :D2:H11 P:GLU102:OE1 :2:H11 P:GLU102:OE2 :D2:H13 P:THR97:O	- 2.87 2.08 - 1.99 1.93 - 1.49 - - - -

S.No	Structure	Glide score	Glide energy	PrimeM M-GBSA complex energy(d G bind)	Binding Energy from AutoDock(Kcal/mol)	Protein-ligand interactions	Bond Distance (A)
D3		-5.11	-13.53	-18.03	-4.1	P:GLU102:OE2 :D3:N3 :D3:H10 P:GLU102:OE1 :D3:H10 P:GLU102:OE2 :D3:H11 P:THR97:O :D3:H13 P:THR97:OG1	- 3.04 - 1.75 - 2.15 - 1.82 - 1.99 - - -
D4		-4.33	-20.98	17.87	-3.8	:D4:H16 P:GLU102:OE1 :D4:H16 P:GLU102:OE2 :D4:H17 P:SER122:OG :D4:H20 P:GLU102:OE1	2.22 2.12 1.78 1.74 - - -
D5		-4.27	-22.05	-17.03	-3.5	P:ASP87:OD2 :D5:N29 P:GLU89:OE2 :D5:N27 P:THR97:HG1 :D5:F11 :D5:H19 P:GLU89:OE1 :D5:H20 P:GLU89:OE2 :D5:H21 P:ASP87:OD2	- 2.68 - 3.05 - 2.22 - 2.05 - 2.13 - 1.66 - - - - -
D6		-4.6	-20.10	-16.45	-3.1	P:GLU102:OE2 :D6:N26 :D6:H14 P:GLU102:OE1 :D6:H17 P:GLU102:OE1 :D6:H17 P:GLU102:OE2 :D6:H18 P:SER122:OG :D6:H27 P:SER122:OG	- 2.76 - 1.57 - 2.42 - 1.74 - 2.28 - 1.78 - - - - -



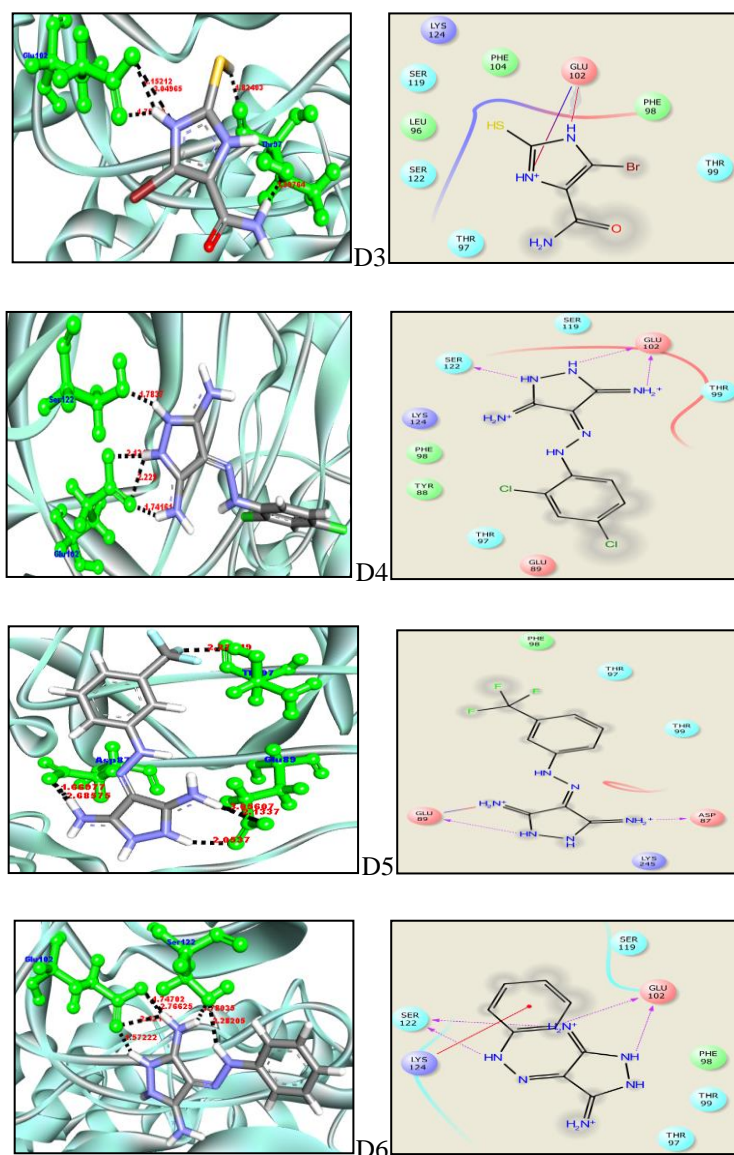


Figure 13. The 3d molecular interactions of lead molecules D1, D2, D3, D4, D5, and D6 with active site residues of FGR protein are identified from the discovery studio3.5 version. The active site amino acid residues interacting are shown in ball and stick green color and the protein backbone as cyan ribbon. The ligand compounds are in the stick model. The hydrogen bonds between lead compounds and active amino acid residues are shown in the black dotted line.

3.4. Auto docking.

Molecular docking was carried out by using AutoDock [38]. The 19 XP outfile molecules were docked using Autodock with FGR protein. The resulting docked complexes were prioritized based on their binding energies (-4.1 to -6.31 kcal/mol).

Based on GLIDE and AutoDock docking, it was observed that top prioritized ligand molecules D1-D6 were exhibiting good binding free energy negative dG values concerning protein-ligand complexes shown in Table 4.

3.5. ADME properties.

Pharmaceutically relevant properties of selected lead molecules were identified by the Qikprop module of the Schrodinger suite. It is used to estimate ADME properties of lead molecules, good human oral absorption. The Lipinski rule of five and Jorgensen rule of three

are shown in Table 5, which are helpful to identify potential lead molecules against FGR protein for cancer therapy [39].

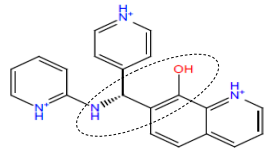
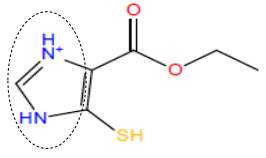
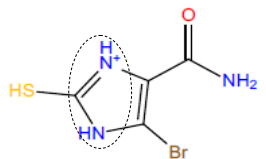
Table 5. The pharmacokinetic properties of lead docked complexes obtained from QikProp.

S.No	M.wt	Donor HB	Acceptor HB	Q plog Po/w	% of human Oral Absorption	Rule of three	Rule of five
D1	328.3	2.0	4.7	3.785	100	0	0
D2	172.2	0.0	1.5	2.158	93.9	0	0
D3	222.0	2.8	3.0	0.561	75.5	0	0
D4	271.1	5.0	4.0	1.061	73.3	0	0
D5	270.2	5.0	4.0	1.072	72.1	0	0
D6	202.2	5.0	4.0	0.173	66.1	0	0

3.6. Justification of lead molecules.

The final top lead molecules, which were observed with good binding energies from the MM/GBSA approach, also show a good % of human oral absorption, as shown in Table 6. The six molecules D1-D6 had binding energies between -17.03 to -25.63 kcal/mol with 100% to 66% human oral absorption. The ligand molecules D2 and D3 consist of imidazole structural moiety, and D4, D5, D6 molecules consist of pyrazolidine moiety. All ligands are consistently binding with GLU102, THR97, ASP87, SER122 amino acids and forming Hydrogen bonding, which increases the strength of the binding interaction between protein and ligand molecule. The best-docked complexes D1-D3 are prioritized based on binding free energy, % of human oral absorption, and binding energy from AutoDock. These molecules have effective binding free energy compared to other moieties, Quinoline 8-hydroxy ring of D1 molecule hydrogen atom of OH functional group is forming hydrogen bonding with oxygen (OE1) atom of GLU102 residue which is responsible inhibition of the activity of FGR protein at the active site, which increases the strength of the binding interaction between protein and ligand molecule.

Table 6. Structural based virtual screening of TOSLab database resulting in best-docked compounds against FGR protein.

Compound	Ligand Molecules	Protein-Ligand binding free energy (dG)	% of human oral absorption	Binding energy from AutoDock (kcal/mol)	Interacting amino acids
D1		-24.14	100.0	-6.3	H-bonding Glu-102
D2		-20.09	93.97	-4.3	H-bonding GLU102 THR97
D3		-18.09	75.59	-4.1	H-bonding GLU102 THR97

The best-docked complexes were optimized based on binding free energies, bioavailability, binding energy from Auto Dock, and interacting amino acids were provided in Table 6. These ligand molecules show druggable property and can be considered a potentially

viable molecule that inhibits the biological function of FGR protein. The Superimposed binding pattern of D1, D2, D3, ligand molecules show that they are occupying similar positions in the binding cavity of FGR and are consistently binding to Glu102 and Thr97 amino acid residues as shown in Figure 14.

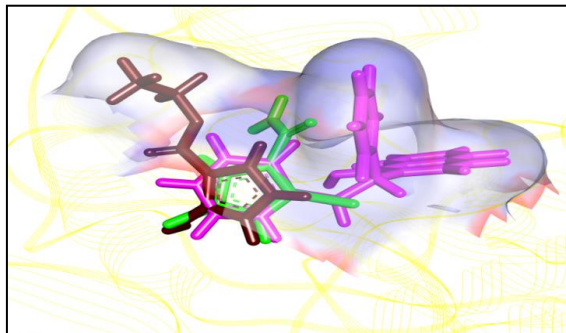


Figure 14. The superimposition of top 3 docked ligands into the FGR active site, lead molecules, and active residues as visualized by Discovery Studio3.5.

4. Conclusions

The work's major target involves identifying potential molecules for inhibiting FGR protein as cancer therapeutics. The homology modeling technique built the structure of FGR protein, FGR protein containing 529 amino acid residues is validated by the homology modeling technique, minimized by the NAMD-VMD software, and the structure is validated by the homology modeling technique different validation protocols. The active site is identified from the SiteMap tool. The active site reveals that the amino acid residues GLU102, THR97, SER122, GLU89, are important in binding with lead ligand molecules, TOSLab ligand database was selected to dock against FGR target protein for hit identification. The final 12 hit molecules were identified. The top 6 leads are D1 to D6, which were prioritized based on GLIDE score, binding free energies, and hydrogen bonding interactions. D1, D2, and D3 molecules were prioritized based on bioavailability, binding free energies, and binding energy from prime MMGBSA and AutoDock. Docking revealed that GLU102 amino acid of FGR protein was binding with the hydrogen atom of OH functional group of Quinoline 8-hydroxy moiety of D1 resulting in increased strength of the binding interaction Protein-ligand complex. It shows 100% human oral absorption. Hence, we conclude that all the lead molecules have a competent range of binding free energies, percentage of human oral absorption, and permissible range of ADME properties may be considered to develop new lead molecules against inhibition FGR protein cancer therapeutics.

Funding

Muddagoni Narasimha is thankful to the University Grants Commission (UGC) India for financial support under the junior research fellowship (JRF) (Roll NO: 123647, REF. NO:21/12/2014(ii) EU-V: Sr, NO. 2121410086).

Acknowledgments

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

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Momenimovahed, Z.; Tiznobaik, A.; Taheri, S.; Salehiniya, H. Ovarian cancer in the world: epidemiology and risk factors. *Int J Womens Health* **2019**, *11*, 287-299, <https://doi.org/10.2147/IJWH.S197604>.
2. Berton, G.; Fumagalli, L.; Laudanna, C.; Sorio, C. Beta 2 integrin-dependent protein tyrosine phosphorylation and activation of the FGR protein tyrosine kinase in human neutrophils. *Journal of Cell Biology* **1994**, *126*, 1111-1121, <https://doi.org/10.1083/jcb.126.4.1111>.
3. Kim, H.-S.; Han, H.D.; Armaiz-Pena, G.N.; Stone, R.L.; Nam, E.J.; Lee, J.-W.; Shahzad, M.M.K.; Nick, A.M.; Lee, S.J.; Roh, J.-W.; Nishimura, M.; Mangala, L.S.; Bottsford-Miller, J.; Gallick, G.E.; Lopez-Berestein, G.; Sood, A.K. Functional Roles of Src and Fgr in Ovarian Carcinoma. *Clinical Cancer Research* **2011**, *17*, 1713-1721, <https://doi.org/10.1158/1078-0432.CCR-10-2081>.
4. Siveen, K.S.; Prabhu, K.S.; Achkar, I.W.; Kuttikrishnan, S.; Shyam, S.; Khan, A.Q.; Merhi, M.; Dermime, S.; Uddin, S. Role of Non Receptor Tyrosine Kinases in Hematological Malignances and its Targeting by Natural Products. *Molecular Cancer* **2018**, *17*, <https://doi.org/10.1186/s12943-018-0788-y>.
5. Erpel, T.; Courtneidge, S.A. Src family protein tyrosine kinases and cellular signal transduction pathways. *Current Opinion in Cell Biology* **1995**, *7*, 176-182, [https://doi.org/10.1016/0955-0674\(95\)80025-5](https://doi.org/10.1016/0955-0674(95)80025-5).
6. Huveneers, S.; Danen, E.H.J. Adhesion signaling – crosstalk between integrins, Src and Rho. *Journal of Cell Science* **2009**, *122*, 1059, <https://doi.org/10.1242/jcs.039446>.
7. Fiser, A. Template-based protein structure modeling. *Methods Mol Biol* **2010**, *673*, 73-94, https://doi.org/10.1007/978-1-60761-842-3_6.
8. 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. *Structural Chemistry* **2020**, <https://doi.org/10.1007/s11224-020-01667-w>.
9. MacGowan, S.A.; Madeira, F.; Britto-Borges, T.; Warowny, M.; Drozdetskiy, A.; Procter, J.B.; Barton, G.J. The Dundee Resource for Sequence Analysis and Structure Prediction. *Protein Science* **2020**, *29*, 277-297, <https://doi.org/10.1002/pro.3783>.
10. Dong, S.; Sun, J.; Mao, Z.; Wang, L.; Lu, Y.-L.; Li, J. A guideline for homology modeling of the proteins from newly discovered betacoronavirus, 2019 novel coronavirus (2019-nCoV). *Journal of Medical Virology* **2020**, *92*, 1542-1548, <https://doi.org/10.1002/jmv.25768>.
11. 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, <http://doi.org/10.1016/j.combiochem>.
12. 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, <http://doi.org/10.1002/jcc.20289>.
13. Sneha, P.; Thirumal Kumar, D.; Saini, S.; Kajal, K.; Magesh, R.; Siva, R.; George Priya Doss, C. Analyzing the Effect of V66M Mutation in BDNF in Causing Mood Disorders: A Computational Approach. *Advances in protein chemistry and structural biology* **2017**, *108*, 85-103, <https://doi.org/10.1016/bs.apcsb.2017.01.006>.
14. 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, <http://doi.org/10.1093/nar/gkm290>.
15. Laskowski, R.A.; Rullmann, J.A.C.; MacArthur, M.W.; Kaptein, R.; Thornton, J.M. AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR. *Journal of Biomolecular NMR* **1996**, *8*, 477-486, <https://doi.org/10.1007/BF00228148>.
16. Schrodinger, L.L.C. *Protein preparation, Sitemap, Ligprep, Glide 5.6, and Prime modules*. New York. **2010**.
17. Salam, N.K.; Nuti, R.; Sherman, W. Novel Method for Generating Structure-Based Pharmacophores Using Energetic Analysis. *Journal of Chemical Information and Modeling* **2009**, *49*, 2356-2368, <http://doi.org/10.1021/ci900212v>.
18. Dizdaroglu, Y.; Albay, C.; Arslan, T.; Ece, A.; Turkoglu, E.A.; Efe, A.; Senturk, M.; Supuran, C.T.; Ekinici, D. Design, synthesis and molecular modelling studies of some pyrazole derivatives as carbonic anhydrase inhibitors. *J Enzyme Inhib Med Chem* **2020**, *35*, 289-297, <https://doi.org/10.1080/14756366.2019.1695791>.
19. Rajender, P.S.; Vasavi, M.; Vuruputuri, U. Identification of novel selective antagonists for cyclin C by homology modeling and virtual screening. *International Journal of Biological Macromolecules* **2011**, *48*, 292-300, <http://doi.org/10.1016/j.ijbiomac.2010.11.015>.
20. 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>.

21. 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. *Journal of Biomolecular Structure and Dynamics* **2020**, *38*, 2314-2325, <https://doi.org/10.1080/07391102.2019.1630318>.
22. Isa, M.A.; Majumdar, R.S.; Haider, S. In silico docking and molecular dynamics simulation of 3-dehydroquinase synthase (DHQS) from *Mycobacterium tuberculosis*. *Journal of Molecular Modeling* **2018**, *24*, <http://doi.org/10.1007/s00894-018-3637-4>.
23. Fan, J.; Fu, A.; Zhang, L. Progress in molecular docking. *Quantitative Biology* **2019**, *7*, 83-89, <http://doi.org/10.1007/s40484-019-0172-y>.
24. Shoichet, B.K. Virtual screening of chemical libraries. *Nature* **2004**, *432*, 862-865, <http://doi.org/10.1038/nature03197>.
25. Polishchuk, P.; Kutlushina, A.; Bashirova, D.; Mokshyna, O.; Madzhidov, T. Virtual Screening Using Pharmacophore Models Retrieved from Molecular Dynamic Simulations. *International Journal of Molecular Sciences* **2019**, *20*, <https://doi.org/10.3390/ijms20235834>.
26. Kaushik, A.C.; Kumar, S.; Wei, D.Q.; Sahi, S. Structure Based Virtual Screening Studies to Identify Novel Potential Compounds for GPR142 and Their Relative Dynamic Analysis for Study of Type 2 Diabetes. *Front Chem* **2018**, *6*, <https://doi.org/10.3389/fchem.2018.00023>.
27. 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. *Chemical Reviews* **2019**, *119*, 9478-9508, <https://doi.org/10.1021/acs.chemrev>.
28. Vardhan, P.; Reddy, V.; Reddy, C.; Suganya, R. Identification of potent lead molecules for furin receptor through HTVS and molecular docking. *International Journal of Pharmaceutical and Clinical Research* **2016**, *8*, 1548-1551.
29. Morris, G.M.; Green, L.G.; Radić, Z.; Taylor, P.; Sharpless, K.B.; Olson, A.J.; Grynspan, F. Automated Docking with Protein Flexibility in the Design of Femtomolar “Click Chemistry” Inhibitors of Acetylcholinesterase. *Journal of Chemical Information and Modeling* **2013**, *53*, 898-906, <http://doi.org/10.1021/ci300545a>.
30. Haddad, Y.; Adam, V.; Heger, Z. Ten quick tips for homology modeling of high-resolution protein 3D structures. *PLOS Computational Biology* **2020**, *16*, <https://doi.org/10.1371/journal.pcbi.1007449>.
31. Dhandare, B.C.; Rather, M.A.; Bhosale, B.P.; Pawar, R.; Guttula, P.K.; Pagarkar, A.U. Molecular modeling, docking and dynamic simulations of growth hormone receptor (GHR) of *Labeo rohita*. *Journal of Biomolecular Structure and Dynamics* **2020**, 1-14, <https://doi.org/10.1080/07391102.2020.1844063>.
32. Webb, B.; Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Current protocols in bioinformatics* **2016**, *54*, 5.6.1-5.6.37, <https://doi.org/10.1002/cpbi.3>.
33. Laskowski, R.A.; Jabłońska, J.; Pravda, L.; Vařeková, R.S.; Thornton, J.M. PDBsum: Structural summaries of PDB entries. *Protein Science* **2018**, *27*, 129-134, <https://doi.org/10.1002/pro.3289>.
34. Qawoogha, S.S.; Shahiwala, A. Identification of potential anticancer phytochemicals against colorectal cancer by structure-based docking studies. *Journal of Receptors and Signal Transduction* **2020**, *40*, 67-76, <https://doi.org/10.1080/10799893.2020.1715431>.
35. Fan, H.; Irwin, J.J.; Sali, A. Virtual ligand screening against comparative protein structure models. *Methods Mol Biol* **2012**, *819*, 105-126, https://doi.org/10.1007/978-1-61779-465-0_8.
36. Dasgupta, Y.; Golovine, K.; Nieborowska-Skorska, M.; Luo, L.; Matlawska-Wasowska, K.; Mullighan, C.G.; Skorski, T. Drugging DNA repair to target T-ALL cells. *Leukemia & Lymphoma* **2018**, *59*, 1746-1749, <https://doi.org/10.1080/10428194.2017.1397662>.
37. Mahadevi, A.S.; Sastry, G.N. Cooperativity in Noncovalent Interactions. *Chemical Reviews* **2016**, *116*, 2775-2825, <http://doi.org/10.1021/cr500344e>.
38. Jananie, R.K.; Priya, V.; Vijayalakshmi, K. Secondary metabolites of *Cynodon dactylon* as an antagonist to angiotensin II type1 receptor: Novel in silico drug targeting approach for diabetic retinopathy. *Journal of pharmacology & pharmacotherapeutics* **2012**, *3*, 20-25, <https://doi.org/10.4103/0976-500x.92497>.
39. Clark, R.D. Predicting mammalian metabolism and toxicity of pesticides in silico. *Pest Management Science* **2018**, *74*, 1992-2003, <https://doi.org/10.1002/ps.4935>.
40. Clark, R.D. Predicting mammalian metabolism and toxicity of pesticides in silico. *Pest Management Science* **2018**, *74*, 1992-2003, <http://doi.org/10.1002/ps.4935>.