

Human Rab8b Protein as a Cancer Target - An *In Silico* Study

Aboubakr HA¹, Lavanya SP¹, Thirupathi M¹, Rohini R¹, Sarita RP², Uma V^{1*}

¹Department of Chemistry, University College of Science, Osmania University, Hyderabad, Telangana, India

²Department of Chemistry, Nizam College, Osmania University, Hyderabad, Telangana, India

Abstract

Testicular cancer develops in one or both of the testicles in young men. Rab8b is a member of the Rab small G protein family, participates in intracellular trafficking events at the site of the adherence junction dynamics in the testis. Overexpression of Rab8b and loss of functioning adherence junction accelerates the tumorigenesis in testis. In the present work, the computer aided high throughput virtual screening studies are applied to identify potent leads for human Rab8b protein. The homology model of Rab8b of 207 amino acid residues chain length was evaluated based on the crystal structure of an appropriate template, and reveals the presence of 6 α -helices and 6 β -strands. The energy of the generated model was minimized and the model was validated using ProSA PROCHECK and ERRAT server tools. The active site was identified using the computational binding site prediction tools like CASTp, efindsite servers and Sitemap of Schrödinger, which show that the residues (GLU33 to GLN60) are important for binding. The molecular interactions of Rab8b with its natural substrate Rabin 8, were examined by *in silico* protein-protein docking studies using patchDock tool, and the results were corroborated with the active site identified from the computational tools. Virtual screening studies were carried out with ligand databases using Glide docking program of Schrödinger suite. The ligands obtained from XP docking with high Glide score and Glide energy, were subjected to QikProp module to predict their ADME properties. These ligands, based on the pharmacokinetic properties, which are new entities, were considered as novel potent inhibitors in cancer therapy.

Keywords: Rab8b; Testicular cancer; Virtual screening; Protein-protein docking; ADME

Introduction

The uncontrolled proliferation and growth of the cells leads to cancer [1,2]. Testicular cancer affects young men, and has one of the highest cure rates of all cancers with an average five-year survival rate of 95% [3]. Rabs are signaling proteins of approximately 20 kDa, constituting the largest family of monomeric GTPases that localize on the cytoplasmic surfaces of distinct membrane-bound organelles [4]. Their function in diverse intracellular pathways such as vesicular trafficking, polarized transport of proteins, and cell movement depends on their ability to shuttle between GTP (an active form, membrane-associated)-bound and GDP (an inactive form, cytosol-associated)-bound conformations. Rab8b is overexpressed in testicular cancer, which is implicated in testis cancer progression [5]. In this study, Rab8b, which represents a novel class of cellular modulators that affects both initiation and progression of tumor cells in *Homo sapiens*, is treated as a novel target for design of new leads against testicular cancer.

Role of Rab8b protein

Rab8b locates in the cell tissue of testis, and participates in adherence junction dynamics [5]. Cell junctions regulate the small molecule trafficking between cells, the organization of cells into tissues, and the adherence of cells to each other and the extracellular matrix [6]. Dysfunction junctions in testis, due to mutation of proteins are implicated in several diseases including cancer [7]. Figure 1 shows a mechanism responsible for adherence junction in the testis mediated by Rab8b protein. Activation/inactivation cycle in Rab proteins plays a critical role in the trafficking of vesicles between various intracellular compartments. Two regulatory proteins, Guanine nucleotide exchange factor (GEF) and GTPase-Activating Protein (GAP), are included in the switch between an active GTP-bound form and an inactive GDP-bound form of Rab8b [8]. Rabin8 (Rab interacting protein) [9] is Rab8bGEF which activates the target protein from the GDP-bound state (inactive

form) to the GTP-bound state (active form). The figure represents the role of Rab8b protein in the testis. In the GTP-bound conformation, Rab8b is likely to deliver endosome-bound and de novo synthesized proteins to the site of the adherence junction. It is also free to interact with an effector molecule, in turn eliciting a cascade of downstream signaling events that regulate junction dynamics. Malfunction in the system of vesicular traffic is responsible for abnormal biological behaviour of carcinous cells [10]. In the present study, the process of activation of Rab8b is targeted to identify antagonists of the target using computational techniques. The present study is aimed at identifying new drug candidates which are potential antagonists against the Rab8b protein, by using *in silico* approaches.

Methodology

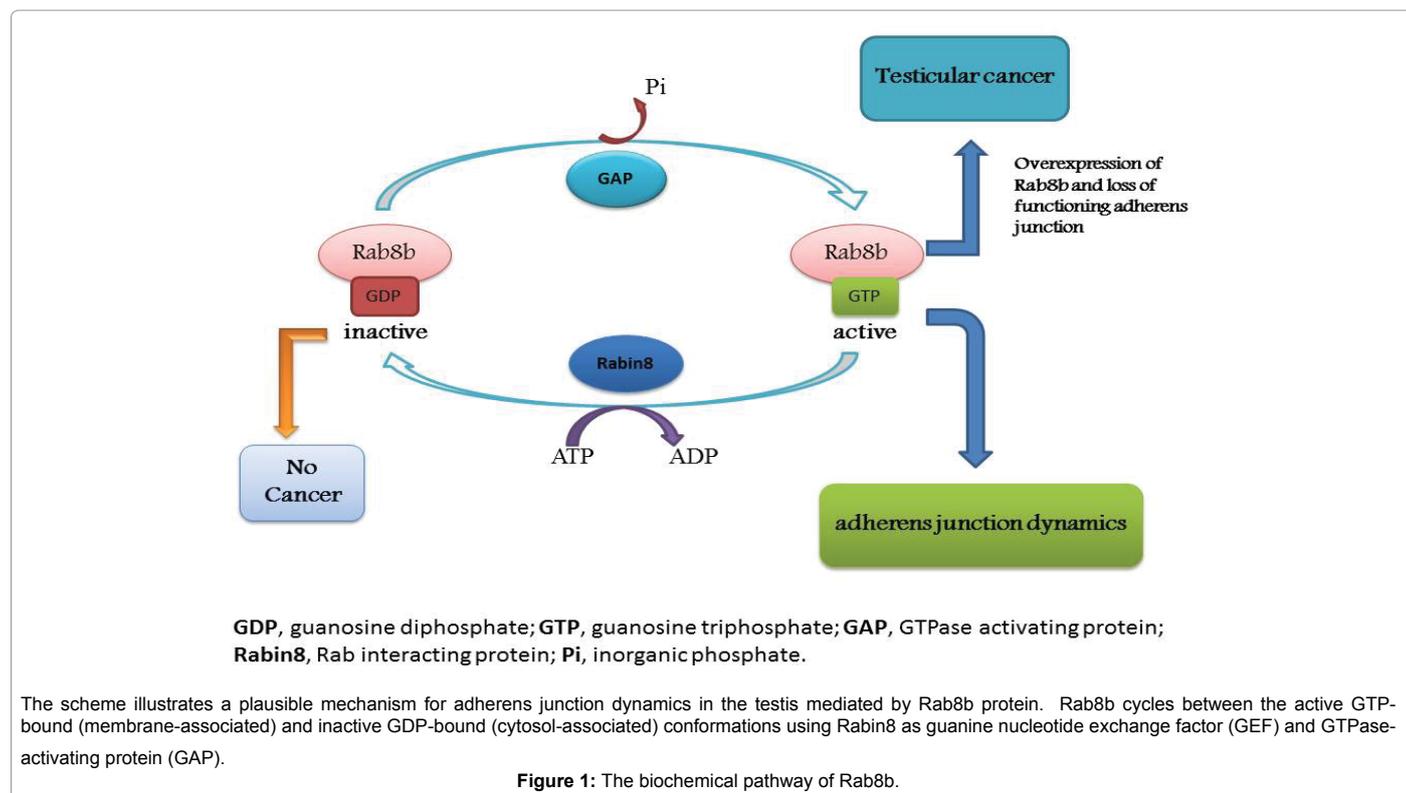
The prediction of a protein structure from its sequence is the basic goal of protein modeling. Herein, we describe the 3D model of a protein which is required to understand how the protein performs its function. The experimental methods such as X-ray crystallography and nuclear magnetic resonance (NMR) are used to determine the protein structure at high resolution [11]. In the absence of these methods, a variety of homology modeling techniques have been developed, which provide reliable models of proteins [12]. The homology models have been useful in drug design projects and allow taking key decisions in compound optimization [13].

***Corresponding author:** Uma Vuruputuri, Department of Chemistry, University College of Science, Osmania University, Hyderabad-500 007, Telangana, India, Tel: +919866104824; E-mail: vuma@osmania.ac.in (or) vuma1957@gmail.com

Received June 23, 2016; Accepted July 05, 2016; Published July 11, 2016

Citation: Aboubakr HA, Lavanya SP, Thirupathi M, Rohini R, Sarita RP, et al. (2016) Human Rab8b Protein as a Cancer Target - An *In Silico* Study. J Comput Sci Syst Biol 9: 132-149. doi:10.4172/jcsb.1000231

Copyright: © 2016 Abdelmonsef AH, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.



3D model generation

The crystal structure for the Rab8b protein is not available. Therefore, understanding of comparative homology modeling method is required to generate the 3D model of target protein. FASTA sequence of human Rab8b is retrieved from the Uniprot Swiss-Prot/TrEMBL database [14] of the Expert Protein Analysis System (ExPASy), a molecular biology server with accession code Q92930 having 207 amino acid residues. The templates for modeling are identified by subjecting the FASTA sequence to comparison methods, such as Basic Local Alignment Tool (BLASTp) program, Phyre2, JPred3, and domain fishing servers [15-18]. The template is selected depending on the criteria of sequence identity and a statistical measure of E-value obtained from the sequences. The pairwise sequence alignment of the protein Rab8b, with the selected template protein (PDB ID: 2BCG) is performed with ClustalW server [19], using the Gonnet matrix [20], to define the structurally conserved and similar regions between the two proteins. The 3D models of the Rab8b are generated by a protein structure modeling program, Modeller 9.11. Twenty-five homology models are built, and the model with the lowest modeller objective function value is selected for structure refinement in the structurally variable regions [21,22]. The RMSD value for the Rab8b protein and its template is calculated in the Swiss-Pdb Viewer (SPDV 4.1.0) [23] by superimposing it on the template structure to assess the reliability of the generated 3D model.

Model validation

The protein preparation wizard in Maestro 9.11 (Maestro 9.11 Schrodinger, LLC, New York, NY), is used in the energy minimization of generated 3D structure of Rab8b. Bond orders are assigned to residues of the protein, hydrogen bonds added and the default RMSD (Root Mean Square Deviation) value of 0.30 Å was assigned using

OPLS_2005 (Optimized potentials for liquid simulations) force field [24,25]. The quality of the generated model is evaluated by a series of tests for its internal consistency and reliability. The quality of the refined and energy minimized 3D model of Rab8b protein is performed by ProSA, PROCHECK and ERRAT server tools [26-28]. The quality of the modeled structure of Rab8b protein is evaluated using ProSA, by checking the potential errors. PROCHECK, uses the Ramachandran 2D contour plot between Ψ and Φ torsion angles of each amino acid residue, to predict the stereochemical quality of the generated structure of the protein Rab8b. Additionally, the ERRAT is a protein structure verification algorithm, describes the overall quality factor of Rab8b protein, and a normally acceptable range for a high quality model being above 50%.

Active site prediction

Identification of binding pockets is an important factor in understanding the function of a protein which is a key step in the process of structure based drug design. Active site (binding pockets) is the region of the protein that is responsible for its activity [29,30]. The active site of Rab8b protein is predicted using Sitemap of Schrodinger suite, CASTp and efindsite servers [31-34]. The Grid generated around the active site region, of size $80 \times 80 \times 80 \text{ \AA}^3$, is used for further virtual screening studies [35].

Protein-protein docking studies

Protein-protein interactions (PPIs) play a central role in various aspects of biological processes such as signal transduction, transport, cell regulation and gene expression controls [36]. These interactions form stable protein-protein complexes that are essential to perform their biological functions. The Rab8/Rabin8 interactions are examined by an *in silico* protein-protein docking studies using patchDock sever Beta 1.3 version [37], and the results are corroborated with the

binding pockets identified from computational active site prediction techniques. The Rabin8/Rab8b inter-molecular interactions are visualized in Accelrys Discovery Visualizer 3.5. The solvent accessible surface area (SASA) describes the potential binding region in the protein structure [38]. The SASA for the Rab8b protein-its natural substrate (Rabin8) complex, before and after docking, is analyzed by using Accelrys Discovery Studio Visualizer [39].

Virtual screening studies

The most important application of docking is the virtual screening, in which most interesting and promising molecules are selected from an existing database for further research [40]. The 3D structure of Rab8b protein is used as the target protein for virtual screening to identify the potential small molecule inhibitors of the target protein.

Ligand preparation: Schrödinger's LigPrep [41] is a tool used to generate the 3D structures from 2D representations, search for tautomers, steric isomers, and ionization states of different ligand molecules, followed by geometry optimization of ligands based upon the OPLS-2005 force field [42]. The process of ligand preparation involves conserving the specified chiralities to generate minimum 4 stereoisomers per ligand, using default conditions at pH 7.0 ± 2.0, to obtain several conformers. Four low energy conformers are generated per ligand. The ligands obtained from the databases, are used for further computational analyses in docking studies.

Computational docking: Docking is a procedure applied to predict the preferable binding orientation between protein active site and a small molecule, forming a stable complex [43]. The virtual screening program of the protein Rab8b using databases, at the active site grid, is performed by three scoring protocols of docking with Glide, namely the High Throughput Virtual Screening (HTVS), the Standard Precision (SP) and the Extra Precision (XP) [44]. A grid is created in the active site domain of Rab8b using Glide module of Schrödinger suite to perform virtual screening. The docked ligands are ranked and prioritized based on Glide scoring function which uses OPLS-2005 force field. The solvent accessible surface area (SASA) is calculated typically using a probe sphere of a given radius to the surface of the molecule. The probe radius is set to 1.40 Å and 240 grid points per atom [45] using Accelrys DS Visualizer 3.5 software and visualized.

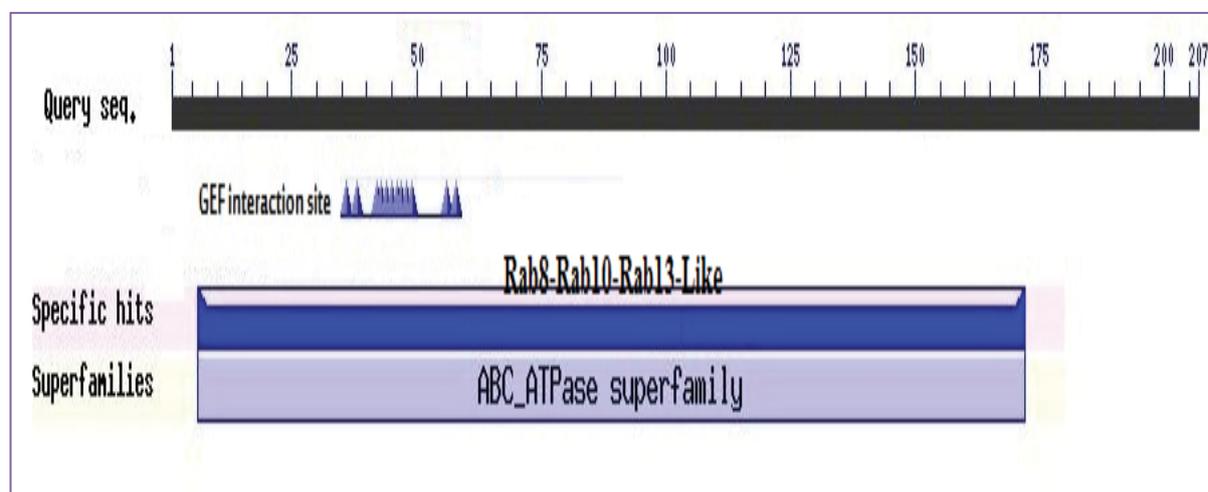
ADME

Understanding the molecular properties of absorption, distribution, metabolism, and excretion (ADME of drug candidate) is an important step in drug discovery process, to predict which specific chemical compounds are best suited for treating a disease, and to optimize the balance of properties necessary to convert the leads into good drug candidates [46,47]. The lead molecules obtained from screening are then tested *in silico*, for their pharmacokinetic properties and their human oral absorption values using QikProp module of Schrodinger software suite [48].

Results and Discussion

Structural analysis of Rab8b protein

Computational structure predictions are important techniques because they can provide the 3D information about huge number of experimentally undetermined structures for proteins. In this work, Rab8b, the target protein 3D model was determined by a comparative homology modeling program, Modeller 9.11. The amino acid sequence of Rab8b protein was retrieved from Uniprot KB [49]. Figure 2 represents the conserved domains of Rab8b. The amino acid residues from 17 to 158 are important for GTP/Mg²⁺ binding. Meanwhile, the amino acid residues from 33 to 60 are important for putative GEF interaction site. The protein template structure was identified in PDB with acceptable similarities in sequence, fold, secondary structure and domain, using various servers such as BLAST, JPred3, Phyre2, and Domain Fishing respectively. All the results are shown with their corresponding E-values in Table 1. A low E-value indicates a high protein specificity and sequence identity [50,51]. 2BCG was selected as a template depending on maximum identity, statistical E-value and Query coverage. The sequence alignment of Rab8b with template 2BCG using ClustalW, reveals a sequence identity of 51.47% as depicted in Figure 3. The conserved residues are shown in green colour, strongly similar residues in yellow colour and weakly similar residues in pink colour. Homology models of Rab8b, using Modeller 9.11, were built. 25 models were initially considered based on comparative protein structure modeling protocols, satisfying the spatial restraints in terms of probable density function [52]. The modeller objective function

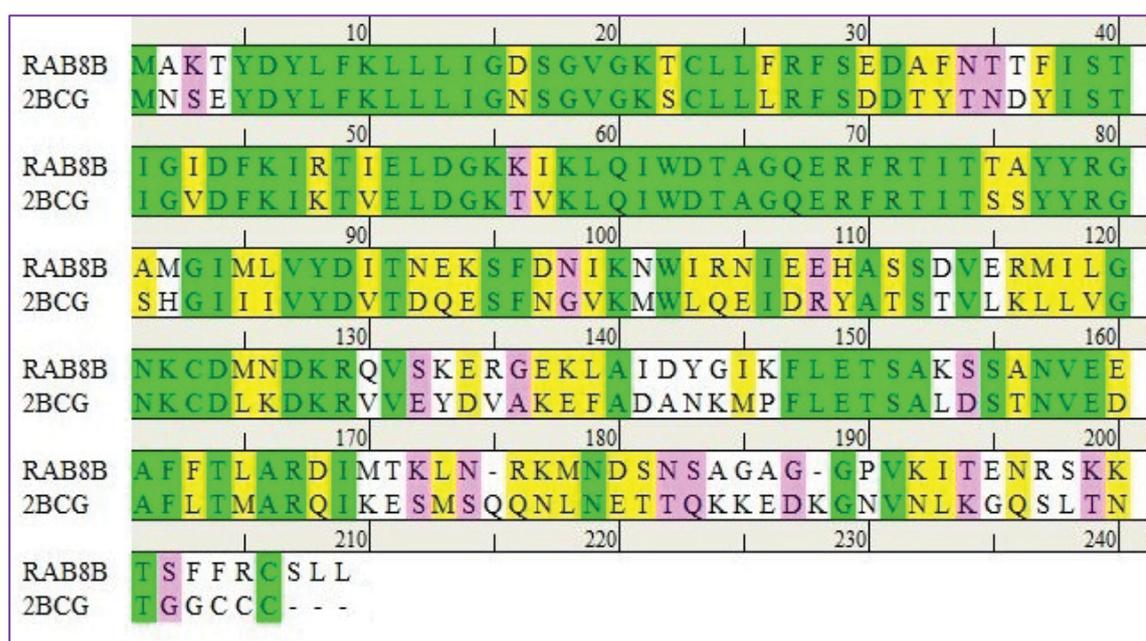


Basic Local Alignment search tool illustrates that the amino acid residues from 33 to 60 are important for putative GEF interaction site. **Figure 2:** The conserved domains of Rab8b protein.

Name of server	Parameters for template selection	E-Values	PDB code
NCBI-Blast	Sequence similarity	2.00E-72	2BCG
Phyre2	Protein fold recognition(threading)	2.00E-25	2BCG
Jpred3	Secondary structure prediction	7.00E-57	2BCG
Domain Fishing	Domain similarity	4.00E-32	2BCG

The protein with PDB ID: 4110, is selected as a template based on the E- values, by using the BLAST, Phyre2, Jpred3 and Domain Fishing servers.

Table 1: Template search results from various servers (E-value) for the Rab8b protein.



The conserved (identical) residues are shown in green colour, strongly similar residues with yellow colour and weakly similar residues with pink colour. Pairwise alignment was carried out using ClustalW server.

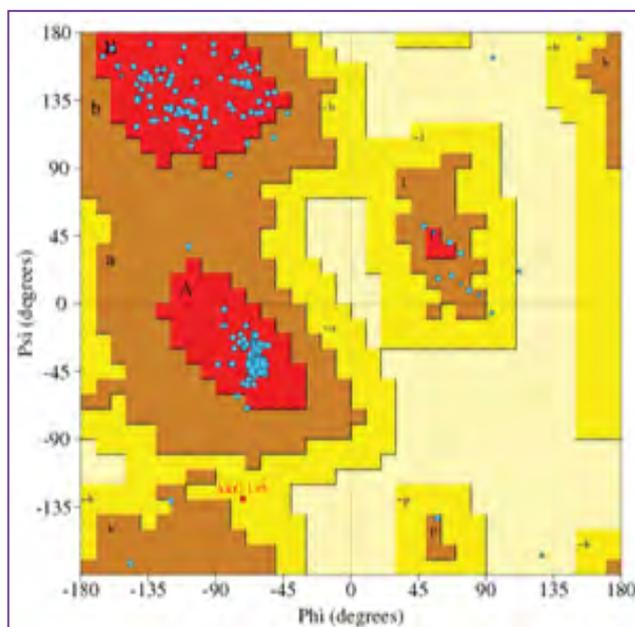
Figure 3: Sequence alignment of Rab8b protein with template PDB ID 2BCG.

values for these models range from 960 to 1209, and the model with the lowest probability density function (960) was selected for further refinement. Protein loop modeling (“Build Loop” module) was carried out using Swiss-Pdb Viewer by applying GROMOS96 Force-Field 4.1.0. The RMSD value of Rab8b protein with its template protein 2BCG is 0.31 Å, calculated using the Swiss-Pdb Viewer (SPDBV_4.1.0, the permissible range of RMSD for any protein being ≤ 2 Å) to assess the reliability of the generated model. The low overall RMSD reflects high structural conservation making it a good structure for further use of in silico study [53]. Energy minimization of 3D structure, is vital for protein stability, and was carried out by using protein preparation wizard in Schrödinger suite. The total energy of the protein after protein minimization was -1132 k.cal/mol. The quality of the refined and energy minimized 3D model of Rab8b protein was performed by PROCHECK, ProSA, and Errat analysis. In the Ramachandran 2D contour plot of Rab8b protein, as given in Figure 4a, shows 173 residues out of 207 (91.1%) and 16 residues (8.4%) fall in the most favored and additionally allowed regions respectively, indicating a good and acceptable quality model. In ProSA, the Z-score (dark spot) of Rab8b protein (-5.41) falls within the range of the values observed for the experimentally determined protein structures of similar chain lengths. The protein Z-score is within the range of the Z-scores for the PDB proteins whose structures are determined experimentally by

NMR (dark blue region) and X-ray crystallography (light blue region), which is indicative of a good quality model (Figure 4b). Figure 4c shows local model quality by plotting energies as a function against amino acid sequence position. Finally, the overall quality factor value for the structure of the protein, Rab8b, is 80.80%, which is calculated using Errat program (Figure 4d). The generated 3D model was visualized using pyMOL 1.3 software [54], which reveals six α -helices and six β -strands (Figure 5), N-terminal indicates the starting residue and C-terminal the end residue. Tables 2 and 3 show the number and the sequence of amino acid residues forming the alpha helices and beta strands in the Rab8b protein, which are identified using PDB sum server (<http://www.ebi.ac.uk/pdbsum/>) [55]. The stability of the protein structure comes from the non-covalent bonds that hold the amino acid residues together in the Rab8b protein structure [56]. Table 4 shows the amino acid residues that are involved in the formation of the salt bridges, Pi-Pi, Pi-cation, and Pi-sigma interactions. The salt bridges arise from the anionic carboxylate (RCOO⁻) of ASP31, ASP53, GLU68, GLU149 and GLU159 with the cationic ammonium (RN⁺H3) of ARG27, ARG167, LYS21, ARG129 and ARG48 respectively.

Active site identification

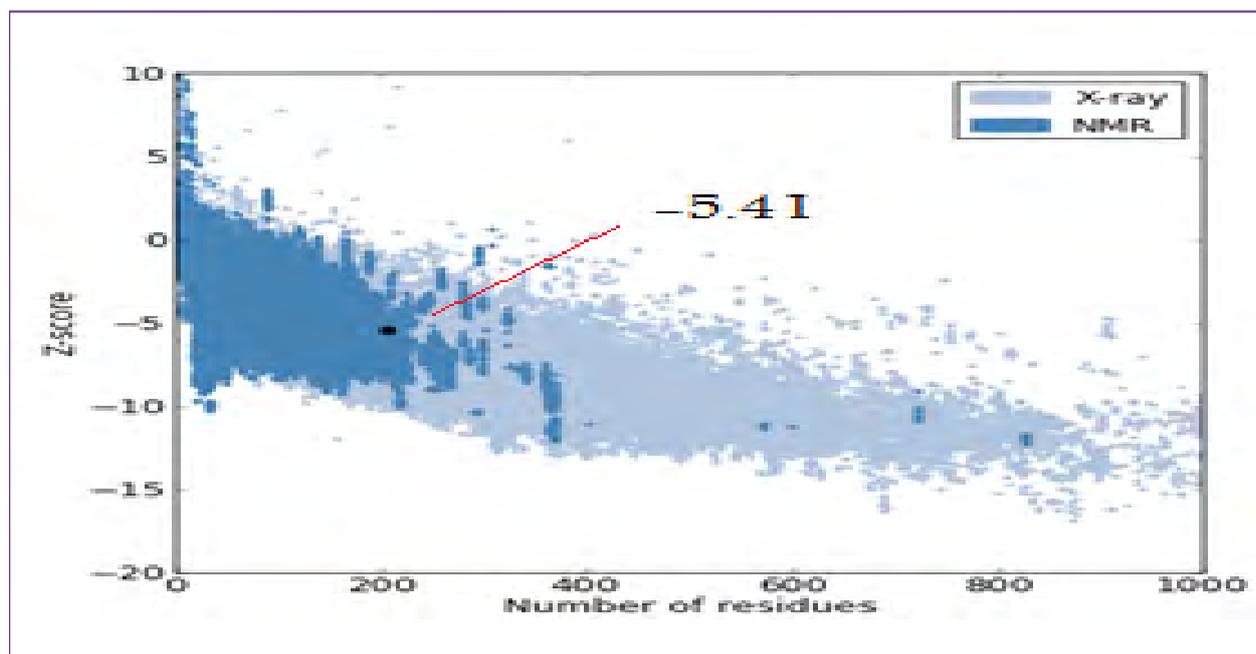
The active site residues were identified by using in silico binding site prediction tools like CAST-p, efindsite, Sitemap (Sitemap



	No. of residues	%-tage
Most favored regions [A,B,L]	173	91.1%
Additional allowed regions [a,b,l,p]	16	8.4%
Generously allowed regions [~a,~b,~l,~p]	1	0.5%
Disallowed regions [XX]	0	0.0%
Non-glycine and non-proline residues	190	100%
End-residues (excl. Gly and Pro)	2	
Glycine residues	14	
Proline residues	1	
Total number of residues	207	

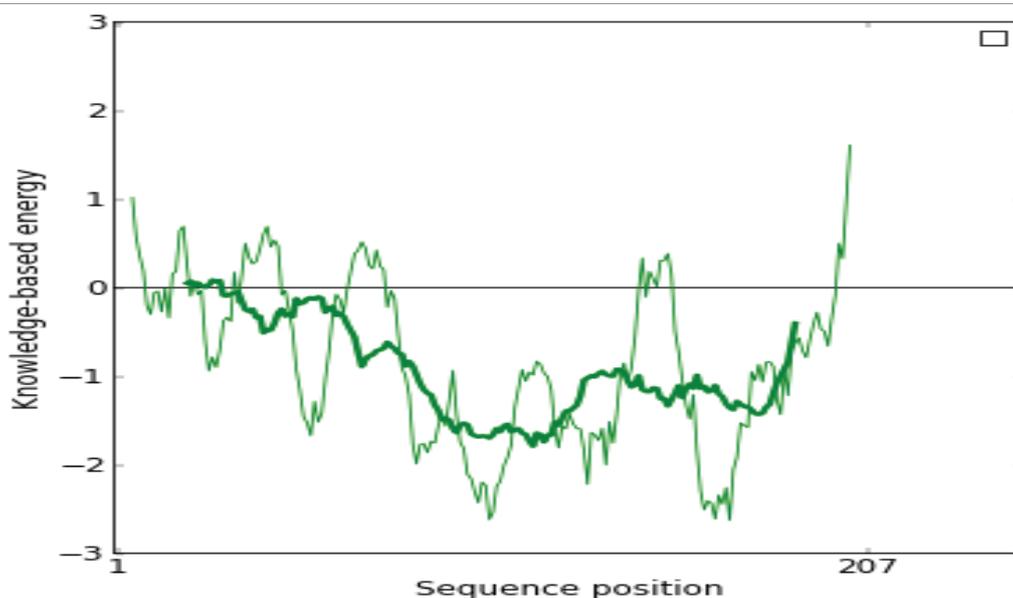
Ramachandran plot obtained by Structural Analysis and Verification Server (SAVS). The red area represents the most favorable region of amino acid residues; the yellow region is additionally allowed and the generously allowed residues are in the light yellow region. The Ramachandran plot of protein represents 91.1% of amino acid residues falling in the allowed region indicating a good local quality model.

Figure 4a: Stereochemical analysis of Rab8b using Ramachandran plot.



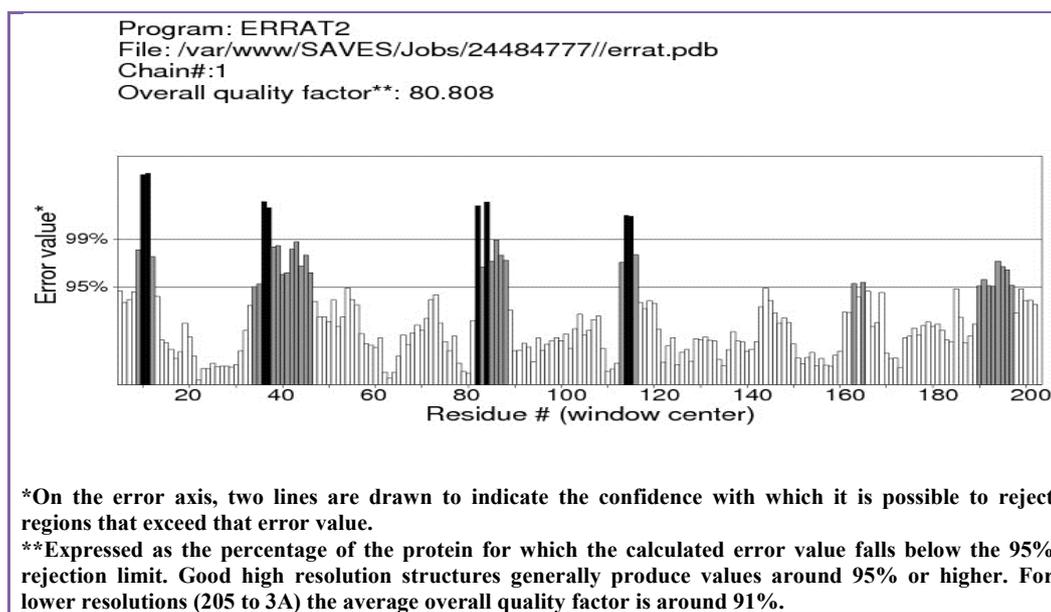
The protein Z-score (-5.41) falls in the range of the Z-score for PDB proteins whose structures are determined by NMR (dark blue region) and X-ray crystallography (light blue region), indicating good quality model.

Figure 4b: The local model quality of the target protein.



The ProSA of the model shows maximum residues with energy in the negative region.

Figure 4c: ProSA energy plot of Rab8b protein



The Errat evaluation method of Rab8b protein gave 80.80 % as the overall quality factor which is considered to be good to use this model. In the Errat, the incorrect regions are shown in black, and the correct regions are shown in gray.

Figure 4d: The overall quality value of Rab8b protein using ERRAT program

Schrodinger), and patchDock servers (Figures 6-9 and Tables 5-8). The results obtained show that the amino acid regions ranging from GLU33 to GLN60 are considered to be the active site of Rab8b and are important for binding to the ligand databases or the natural substrate. The molecular interactions between Rab8b and its natural substrate Rabin8 were examined by an *in silico* protein-protein docking studies which show that the residues ASN34, THR36, PHE37, ILE38, SER39, THR40, ILE41, and ASP44 in Rab8b interact with GLY227, LYS223, GLU219, GLN215, LYS220 and GLU224 in Rabin8 respectively, as represented in Figure 9. The results were compared with the active

site identified from computational active site prediction techniques, which show that amino acid residues GLU33 to GLN60 are important for Rab8b binding to the ligand molecules and the substrate. In the docked complex of Rab8b-Rabin8, twelve hydrogen bonds were observed, which are responsible for the stability of the structure. The solvent accessible surface area (SASA) for the target protein-its natural substrate (Rab8b-Rabin8) complex, before and after docking (Figure 10), was calculated using Accelrys Discovery Studio Visualizer. It was observed from these studies that the residues GLU33 to GLN60 have lower SASA values after docking as compared to that before docking,

No.	Start	End	No. residue	Length (Å)	Sequence
1.	LYS21	SER29	9	13.63	KTCLLFRFS
2.	THR75	TYR78	4	7.33	TAYY
3.	GLU93	HIS109	16	24.48	EKSFDIKNWIRNIEEH
4.	LYS133	ASP142	10	15.45	KERGEKLAID
5.	VAL158	LYS172	15	23.12	VEEAFFTLARDIMTK
6.	LYS176	ALA185	8	13.94	KMND SAGA

The amino acid sequences forming six alpha helices in the modeled Rab8b protein.

Table 2: Secondary structure details of the Rab8b protein-the α -helices identified using PDBsum server.

No.	Start	End	No. residue	Length (Å)	Sequence
1.	TYR7	GLY15	9	26.02	YLFKLLIG
2.	PHE45	LEU52	8	20.74	FKIRTIEL
3.	LYS55	TRP62	8	21.59	KKIKLQIW
4.	GLY83	ASP89	7	19.87	GIMLVYD
5.	GLU115	ASN121	7	19.02	ERMILGN
6.	LYS146	GLU149	4	10.54	KFLE

The amino acid sequences forming six beta strands in the modeled Rab8b protein.

Table 3: Secondary structure details of the Rab8b protein-the β -strands, identified using PDBsum server.

Non-covalent bonds	No.	Amino Acid Residues	Distance (Å)
Salt bridges	1	ARG27:NH1....ASP31:OD1	3.79
	2	ARG27:NH2....ASP31:CG	3.41
	3	ARG167:NH1....ASP53:OD1	3.94
	4	ARG167:NH2....ASP53:OD1	3.7
	5	ARG167:NH1....ASP53:OD2	3.38
	6	LYS21....GLU68	2.9
	7	ARG129:NH1....GLU149:OD2	3.5
	8	ARG129:NH2....GLU149:OE1	2.66
	9	ARG129:NH1....GLU149:OE2	2.71
	10	ARG48....GLU159	3.99
π - π interactions	1	PHE28....PHE162	4.28
	2	PHE9....TYR7	6.19
	3	HIS109....TYR78	4.71
π -cation interactions	1	TYR62....LYS10:NZ	4.53
	2	PHE28....ARG148:NH1	5.49
	3	TRP102....ARG71:NH2	6.96
	4	PHE26....LYS46:NZ	6.99
π -sigma interactions	1	TRP62....LYS10:CE	3.7
	2	TYR88....PHE96:CD1	3.96

The amino acid residues forming the non-covalent bonds and their distance in Angstrom unit. Five pairs of amino acids are involved in the formation of salt bridges, three for π - π interactions, four for π -cation interactions, and two in π -sigma interactions in Rab8b protein.

Table 4: The amino acid residues forming salt bridges, π - π , π -cation, and π -sigma interactions in Rab8b protein.

Site No.	Volume (Å) ³	Amino Acid Residues
Site 1	930.1	SER17-PHE26, PHE33-ILE43, ASP63-PHE70, LYS122-ASP124, SER151-LYS153
Site 2	45.9	LEU25, PHE26, SER29, GLU30, ILE43, LYS46
Site 3	68.5	MET1, ALA2, THR4, ASP5, THR49-GLU51, LYS56
Site 4	26.7	LEU52, ASP53, ARG167, MET170

The four binding cavities identified in CASTp server were based on the hydrophobicity and the ligand binding site prediction, respectively.

Table 5: Active site regions generated in the Rab8b protein using CAST-p server.

Site	Amino Acid Residues
Interfacial residues	PHE37-SER39, ILE41-ILE43, PHE45, ILE47, TRP62, GLN67, ARG69-THR74, ALA76-TYR77, ARG79, HIS109, PHE202
Hydrogen bonds	PHE37-SER39, ILE41-ILE43, PHE45, ILE47, GLN67, ARG69-THR74, ALA76-TYR77, ARG79, HIS109
Hydrophobic interactions	PHE37-SER39, ILE41, ILE43, PHE45, ILE47, PHE70, ILE73, ALA76

The putative binding residues such as interfacial residues, hydrogen bonds and hydrophobic interactions were identified in efindsite sever.

Table 6: Putative binding residues generated in the Rab8b protein using efindsite sever.

Site	Residues	Volume (Å) ³
Sitemap_site_1	LEU13-PHE26, PHE33-ASP44, TRP62-GLU68, MET85-LEU119, GLU159-ILE169	303
Sitemap_site_2	ILE41-ARG48, LEU59-THR64, ALA76-VAL87, HIS109-ILE119, SER151-ILE169	131
Sitemap_site_3	TYR77-MET82, SER111-GLU115, ILE169, ASN181	169

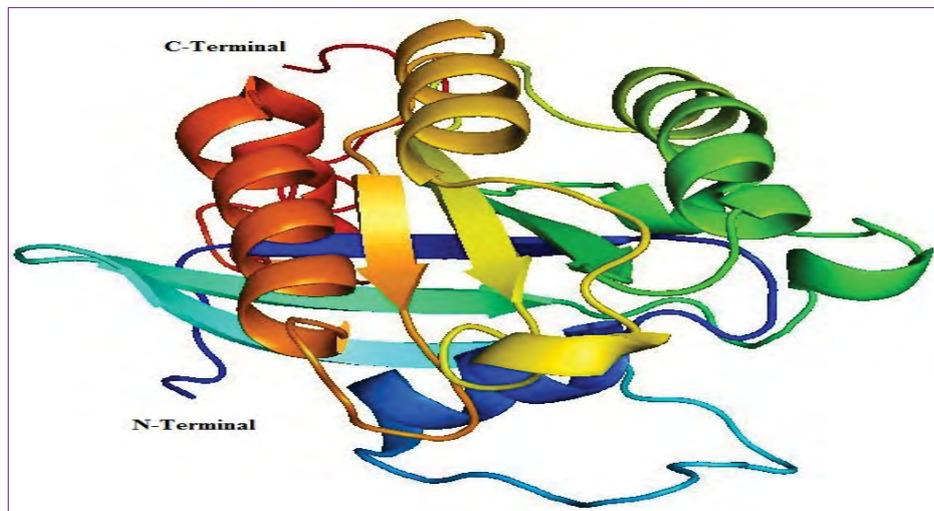
The three binding cavities were identified in Sitemap module and the corresponding cavity volumes are shown in angstrom units.

Table 7: Active site regions generated in the Rab8b protein using Sitemap module.

	Residues of Binding	Distance of H-bonds (Å)
H-bonds	THR36:OG1-C:GLY227:O	2.91
	ILE38:N-C:LYS223:O	2.88
	SER39:N-C:GLU219:O	2.97
	THR40:N-C:GLU219:O	2.53
	ASP44:N-C:GLN215:OE1	2.89
	C:GLU219:N:-SER39:OG	2.94
	C:GLU219:N:-ILE41:O	2.60
	C:LYS220:N:-SER39:OG	1.87
	C:LYS220:NZ:-ASN34:OD1	2.01
	C:LYS223:N:-ILE38:O	1.71
	C:LYS223:NZ:-PHE37:O	1.66
	C:GLU224:N:-ILE38:O	2.59

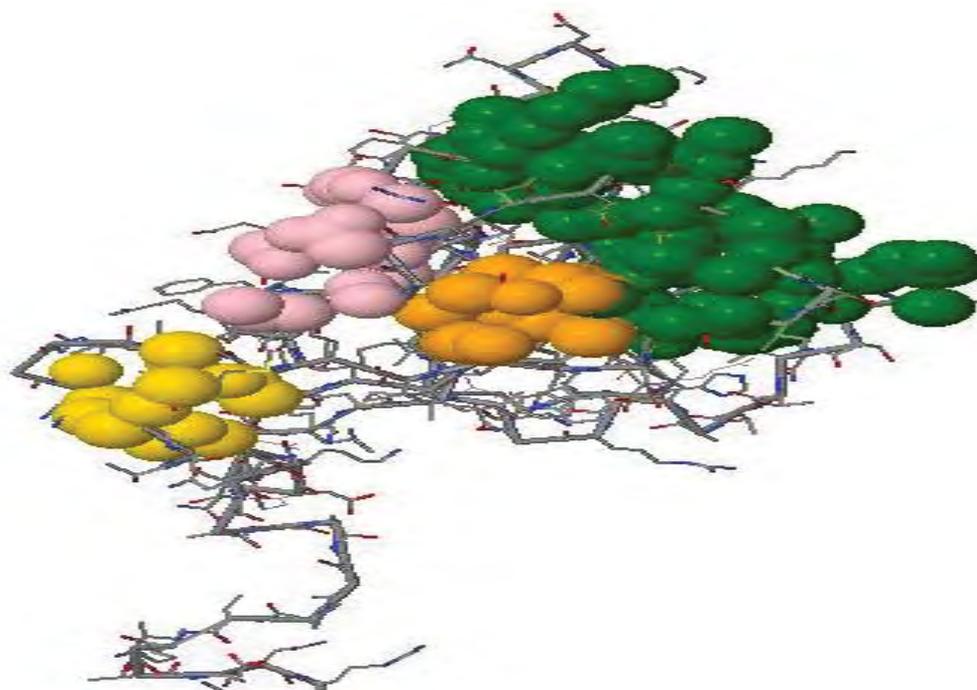
The ASN34, THR36, PHE37, ILE38, SER39, THR40 and ASP44 residues in Rab8b protein bind with the natural substrate (Rabin8), identified using PatchDock server.

Table 8: Intermolecular interactions in the docked complex of the Rab8b and Rabin8.



The 3D structure of the Rab8b protein has 6 α -helices and 6 β -strands, obtained from the Modeller 9.11 program.

Figure 5: The three-dimensional structure of the Rab8b protein.



The active site regions of Rab8b are represented in the figure. Site 1 in Rab8b is represented in green, site 2 in yellow, site 3 in violet and site 4 in orange colour.

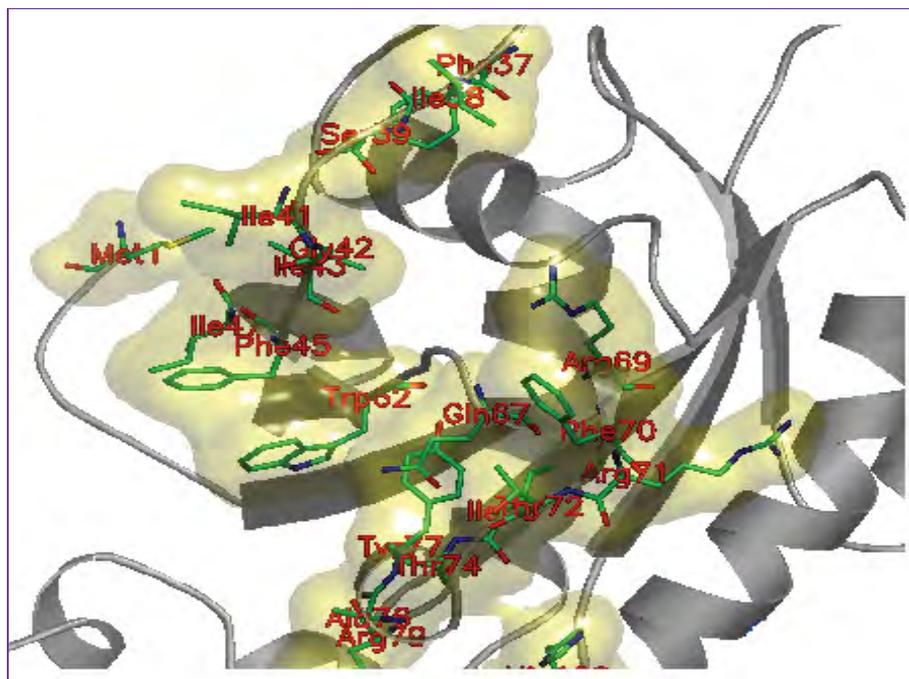
Figure 6: Active site of Rab8b obtained from the CAST-p server.

which shows that the acid residues are involved in the formation of the complex (Figure 10). A cubic grid was generated at the active site region with $80 \text{ \AA} \times 80 \text{ \AA} \times 80 \text{ \AA}$ dimensions (Figure 11) for further studies.

Virtual screening

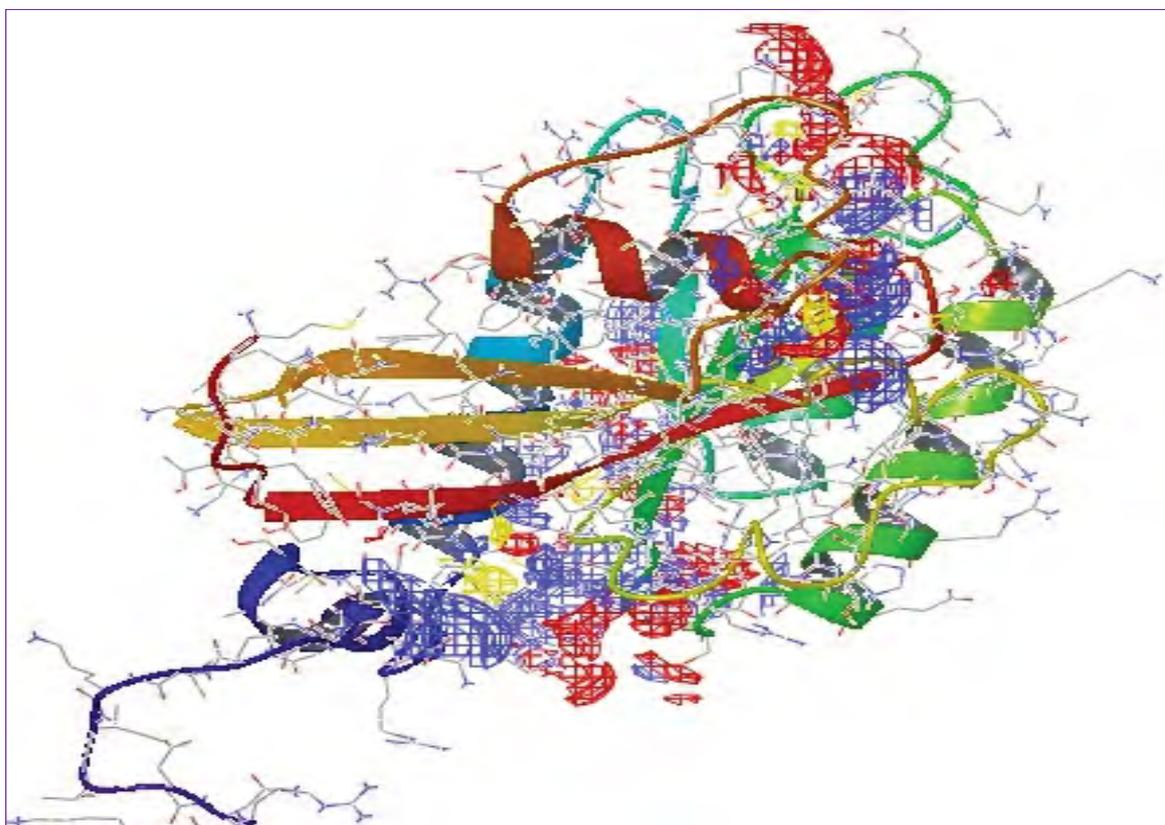
A set of 64,369 conformers were generated by the ligand preparation suite of Schrodinger software by Screening stock and Asinex BioDesign databases. A total of 64,369 prepared ligands were docked flexibly in HTVS/SP/XP mode in Glide docking program of Schrodinger. 10% of the molecules were docked in HTVS mode after flexible docking which

generated 6436 poses. 10% of these conformers were docked flexibly in SP mode which further generated 643 poses. Keeping 10% of 643 poses, 64 ligand molecules were obtained from SP and were docked flexibly in XP mode. All the docked complexes are observed to show a Glide score in the range -10.13 to -7.34. The samples of 10 best docked molecules were ranked depending on the glide score, the glide energy and hydrogen bonding interactions as shown in Table 9. The docked complexes were visualized using Discovery Studio Visualizer 3.5 software. Figure 12 illustrates the docked complexes of Rab8b protein with ligand molecules L1-L10. The top ten hits exhibit major interactions such as π - π , π -cation, hydrogen bonding and hydrophobic



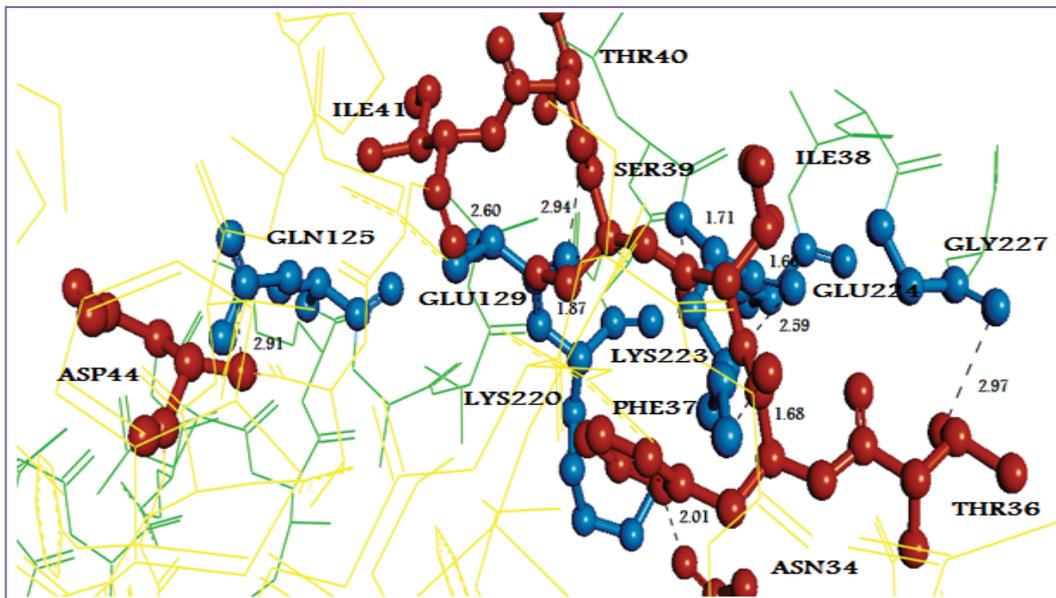
The active site regions in Rab8b obtained from efindsite sever are represented in green colour.

Figure 7: Active site regions of Rab8b obtained from the efindsite sever.



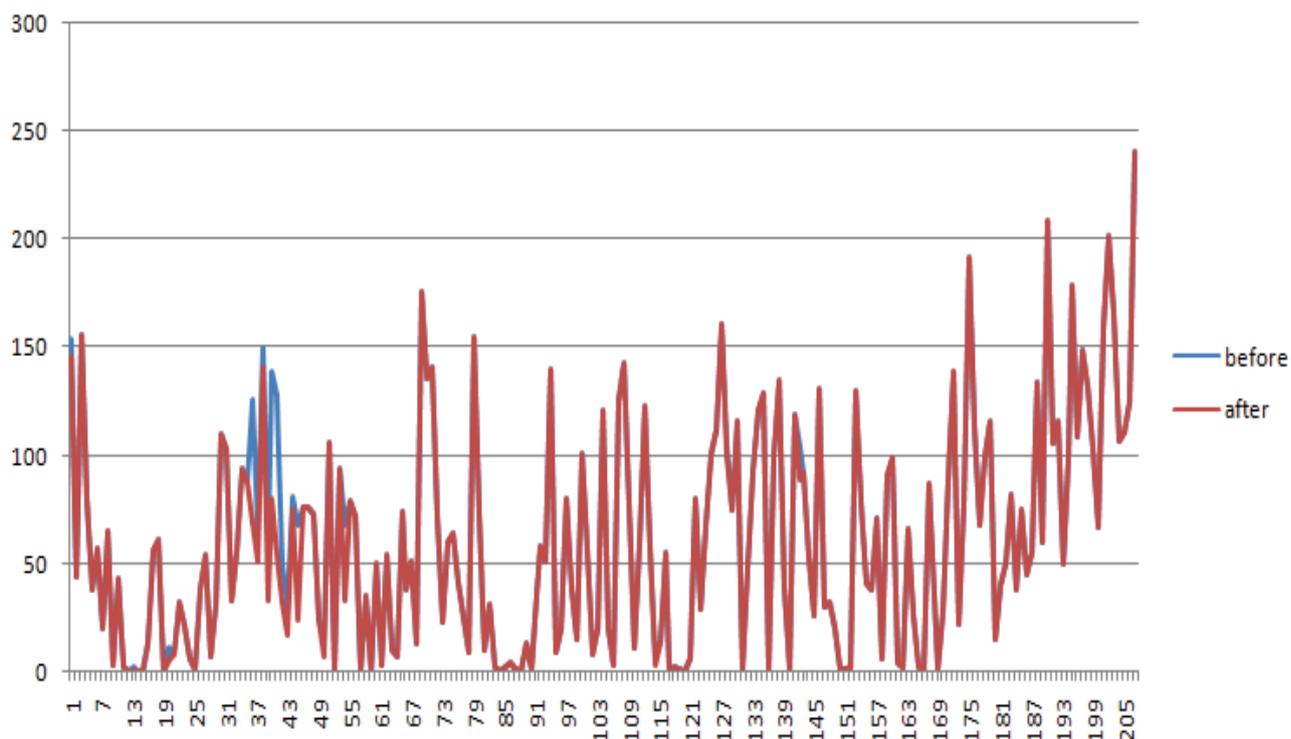
The active site regions obtained from Site map module in Schrodinger illustrates. The hydrogen bond acceptor region shown in red, H-bond donor regions in blue, and hydrophobic pockets in yellow.

Figure 8: Active site of Rab8b obtained from the Sitemap module.



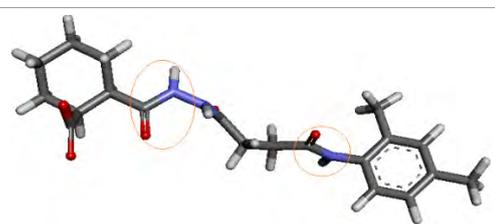
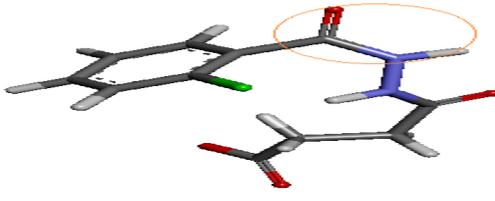
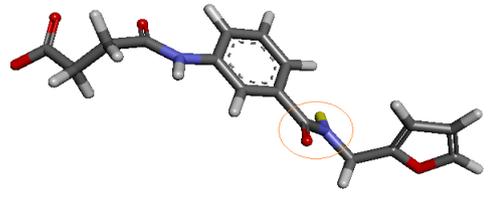
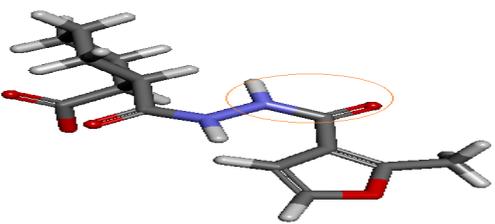
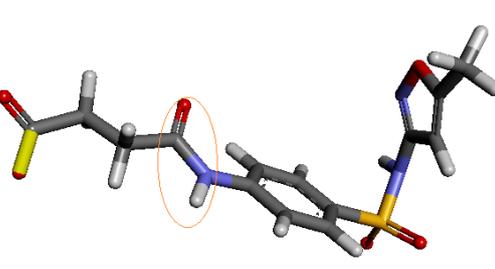
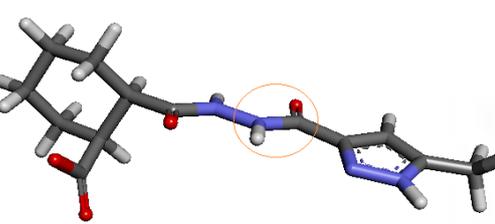
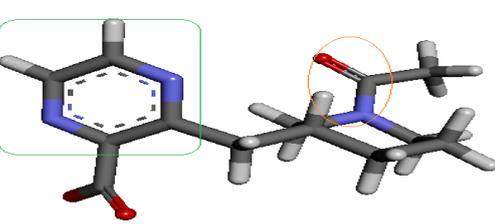
The Rab8b protein (yellow lines) interacts with Rabin8 (green). The specific binding residues are shown in orange ball & sticks and labeled (ASN34, THR36, PHE37, ILE38, SER39, THR40, ILE41, and ASP44) in Rab8b protein. The residues of Rabin8 are represented in cyan colour. The hydrogen bond interactions are represented in black colour.

Figure 9: The protein-protein docking of Rab8b protein and natural substrate Rabin 8.



The data shows significant SASA variations before and after docking. Brown colored peaks represent after docking of Rab8b protein solvent accessibility, blue colored peaks represent before docking. The decreased SASA values of the protein active site residues are GLU33 to GLN60 after docking is shown in brown color and the residues before docking are in blue lines.

Figure 10: The surface accessibility of the Rab8b before and after docking with its natural substrate Rabin 8.

S. No.	Structure	Glide score	Glide energy (K.Cal/mol)	Docked complex (amino acid –ligand atom) interactions	Bond Distance (Å)
L1		-10.13	-38.40	Hydrogen bonds ILE47 [N---O] LYS58[NZ---O] PHE45[O---HN]	2.87 2.44 1.91
L2		-9.98	-42.28	Hydrogen bonds SER39 [N---O] PHE37 [O---HN] LYS58 [NZ---O] LYS58 [NZ---OC] Pi-cation L2 ---LYS21:NZ	2.71 2.35 2.41 2.23 5.86
L3		-9.44	-31.98	Hydrogen bonds ASP44 [N---O] LYS58 [NZ---O] PHE45 [O---HN] Pi-Pi interactions L3---PHE45	2.91 2.61 1.98 5.15
L4		-9.38	-27.44	Hydrogen bonds ILE47 [N---OC] LYS58 [NZ---O] Pi-Pi interactions L4 ---PHE45	2.97 2.47 5.19
L5		-9.03	-31.66	Hydrogen bonds PHE45 [N---O] LYS46 [NZ---N] LYS46 [NZ---O] LYS58 [NZ---O] PHE45 [O---HN] Pi-Pi interactions PHE45 - :L5 Pi-cation L5 ---LYS46:NZ	2.87 2.73 2.83 2.25 2.42 4.29 5.89
L6		-8.80	-30.32	Hydrogen bonds ILE47 [N---O] LYS58 [NZ---O] Pi-Pi interactions L6 ---PHE45	2.93 2.48 4.33
L7		-8.24	-28.38	Hydrogen bonds ILE47 [N---O] LYS58 [NZ---O]	2.89 2.24

L8		-8.12	-28.6	Hydrogen bonds LYS46 [NZ---O] LYS58 [NZ---O]	2.40 2.38
L9		-7.94	-32.71	Hydrogen bonds LYS58 [NZ---OC]	2.30
L10		-7.65	-28.54	Hydrogen bonds LYS46 [NZ---OC] LYS58 [NZ---OC] Pi-cation L10---LYS58:NZ	2.45 2.27 6.99

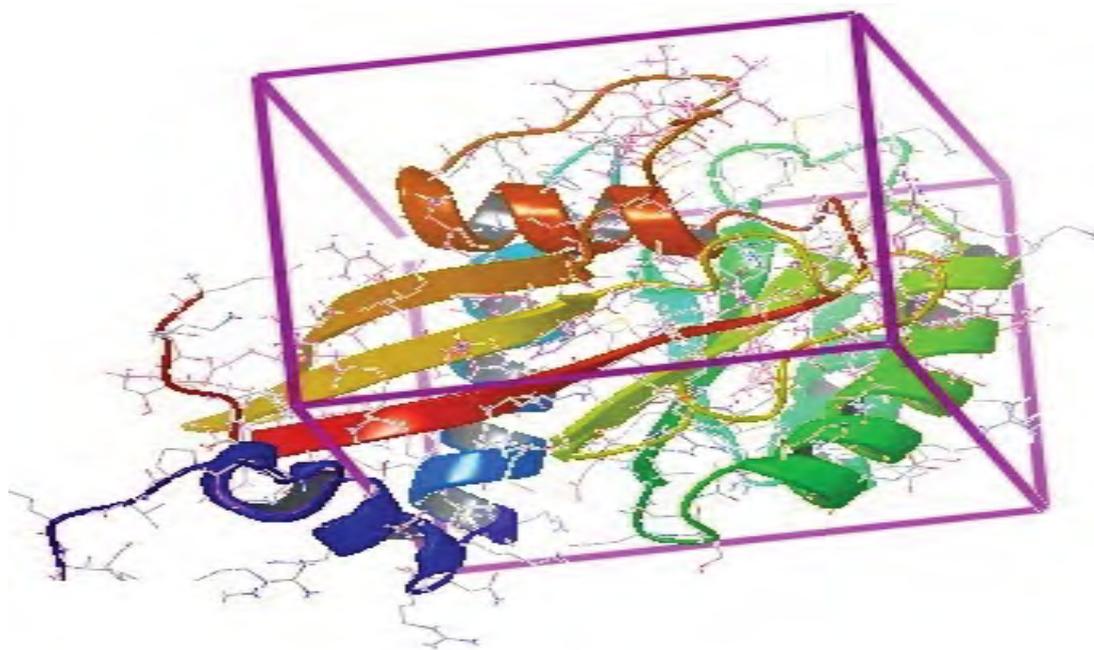
The screening process carried out with HTVS, SP and XP docking modes, gave an output of 68 molecules and are analyzed. The 10 molecules (L1 to L10) with the best docking score are represented with docking interactions in the table showing H-bonding, Pi-Pi, Pi-sigma, and Pi-cation interactions. Carboxamide group is represented in orange cycle while the pyrimidine is in green squares.

Table 9: Structures and glide scores of top scoring ligands against Rab8b protein.

S. No.	Stars	CNS	M.Wt (g/mol)	SASA (Å ²)	Volume (Å ³)	Donor HB	Acceptor HB	QPlog Po/w	QPPCaco (nm/sec)	QPlog BB	Percent Human Oral Absorption	Role of Five	Role of Three
L1	0	-2	389.45	721.94	1273.69	2.5	8	2.96	28.7	-2.06	70.39	0	0
L2	0	-2	270.67	462.25	797.56	1.5	5.5	1.35	19.46	-1.44	57.93	0	1
L3	0	-2	316.31	597.96	1012.34	3	7.5	1.91	50.26	-1.65	68.6	0	0
L4	0	-2	294.3	537.28	912.17	1.5	6	1.94	65.82	-1.22	70.86	0	0
L5	0	-2	353.34	620.78	1053.26	3	10.5	0.07	3.67	-2.92	37.51	0	1
L6	0	-2	294.31	578.84	960.51	1.5	5.5	1.8	11.05	-2.18	56.16	0	1
L7	0	-2	263.29	509.57	875.45	1	7	0.75	37.08	-1.09	59.44	0	0
L8	0	-2	277.32	552.97	953.8	1	7	1.2	41.79	-1.07	63.01	0	0
L9	0	-2	327.39	605.33	1044.07	1	8.5	1.54	40.36	-1.54	64.72	0	0
L10	0	-2	277.32	553.13	945.91	1	7	1.03	25.33	-1.32	58.1	0	0

The pharmacokinetic properties of the ligand molecules (L1-L10) which form docked complexes with Rab8b protein are evaluated by QikProp module and are presented in the table. The permissible ranges are as follows: CNS: -2 (inactive), +2 (active); Mol wt.: (130-725); Donor HB: (0.0-6.0); Acceptor HB: (2.0-20.0); QPlogPo/w: (-2.0 to 6.5); QPlogBB: (-3.0 to -1.2); %Human oral absorption: >80% high, <25% low; Rule of three (3); Rule of five (4).

Table 10: Pharmacokinetic properties of the docked ligand molecules obtained from virtual screening.



The grid generated in the active site regions of Rab8b with the dimensions 80Å ×80 Å ×80Å using the receptor grid generation module of Glide module for virtual screening studies.

Figure 11: Receptor grid generation of Rab8b by using Schrodinger suite.

interactions. The ligand molecules L1-L10 are docked flexibly to the protein and Rab8b being held rigid in the docking process, as default condition in glide, the structure basically remains fixed upon binding to these ligands. The intermolecular hydrogen bonding, Pi-Pi, Pi-sigma and Pi-cation interactions of docked complexes are given in Table 9. The ligand molecules based on their isosteric structural similarity share a common scaffold and binding mode, and their experimental binding data are shown in Figure 12 and Table 9. The carboxamide group in all the ligands except L9, is a common functional group which forms hydrogen bonds with LYS46, ILE47 and LYS58 of the Rab8b protein. The pyrimidine moiety in L7, L8, L9 and L10, binds to LYS46, ILE47 and LYS58 of Rab8b. The virtual screening analysis indicates that the amino acid residues PHE37, SER39, ASP44, PHE45, LYS46, ILE47 and LYS58 are involved in hydrogen bonds, π - π , π -sigma, and π -cation interactions with L1-L10 ligand molecules. It is observed that all the ligand molecules L1-L10 show hydrogen bond interactions with LYS58 in the groove of the binding site. The highly reactive amino group of lysine 58 interacts with carboxylate anion (RCOO⁻) of all the 10 ligand molecules. The SASA calculations are performed for Rab8b protein before and after docking with the ligand molecules and a sample case with L5 is shown in Figure 13. The SASA values of Rab8b for the amino acid residues, which are involved in bond formation (PHE37, SER39, ASP44, PHE45, LYS46, ILE47 and LYS58) and other spatially nearby residues in the binding site, is observed to decrease after docking, as compared to that before docking. The decrease in SASA values shows that these residues are involved in the formation of bonds with the ligand molecules.

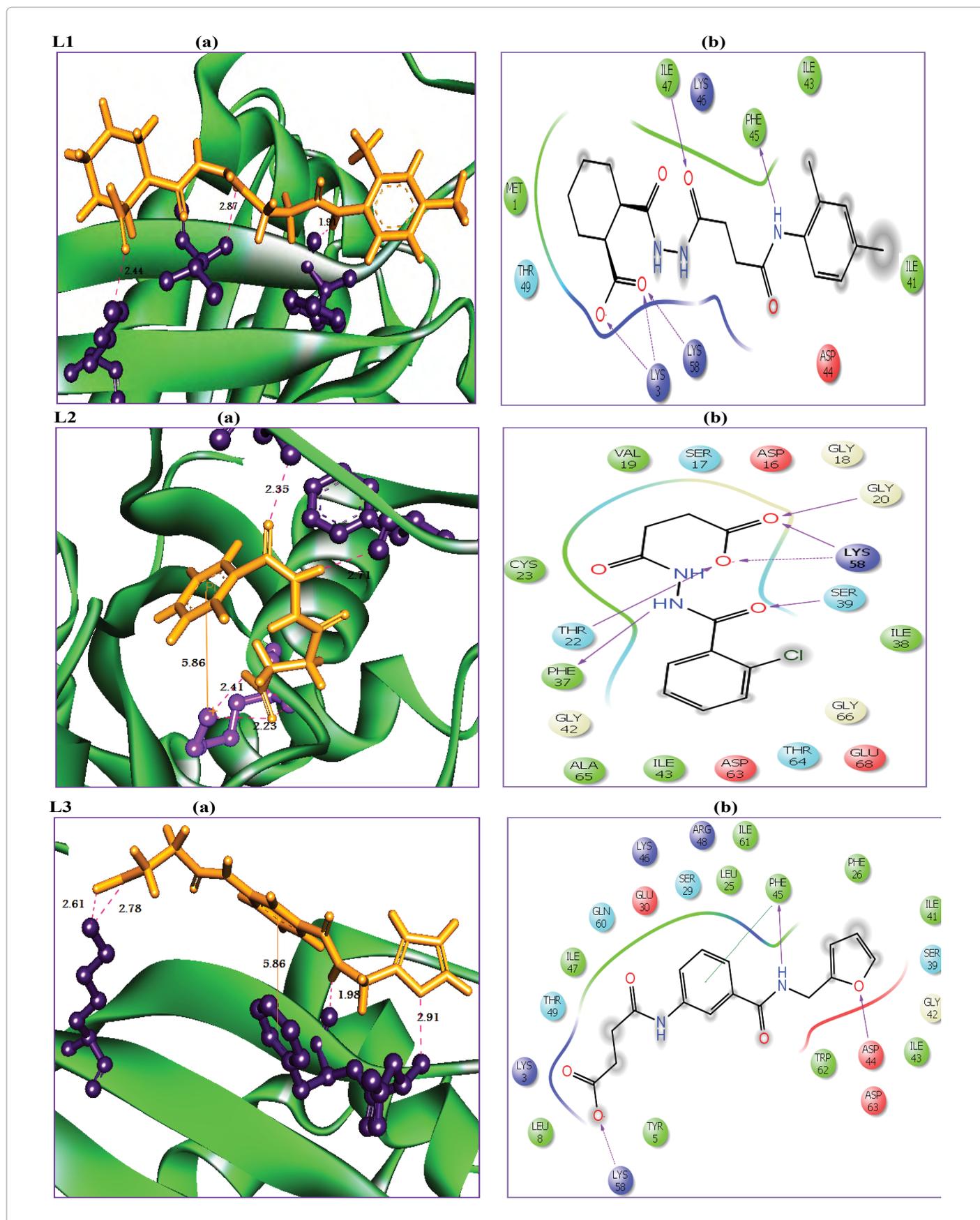
ADME analysis

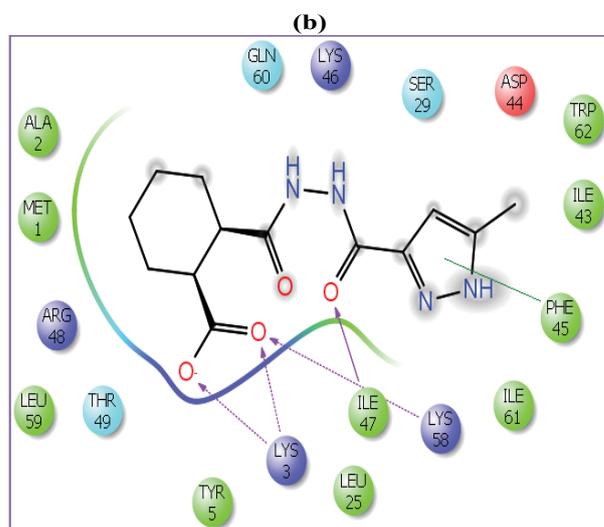
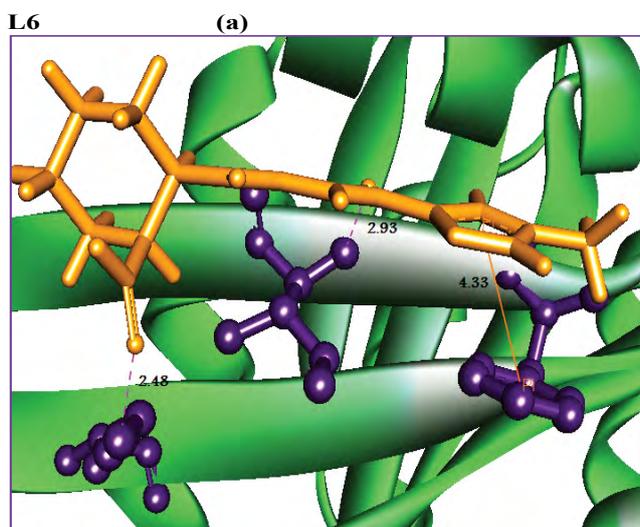
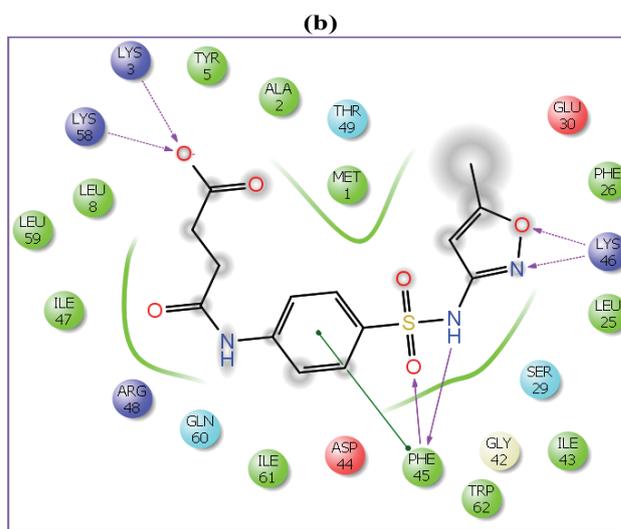
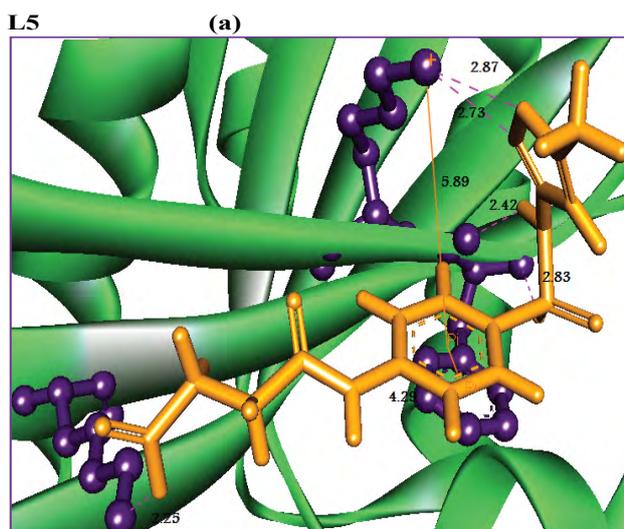
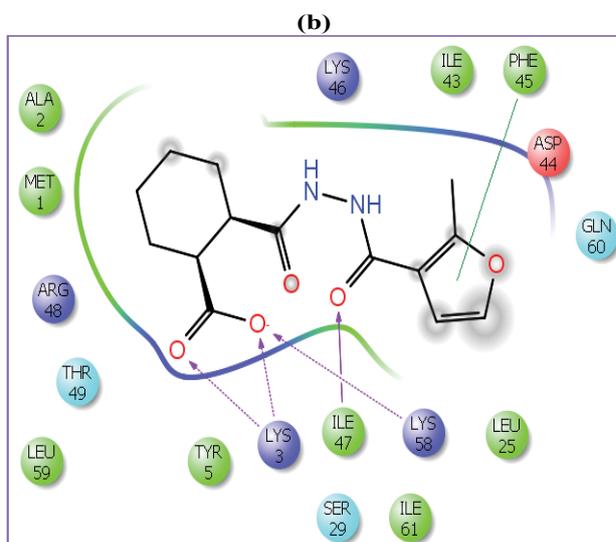
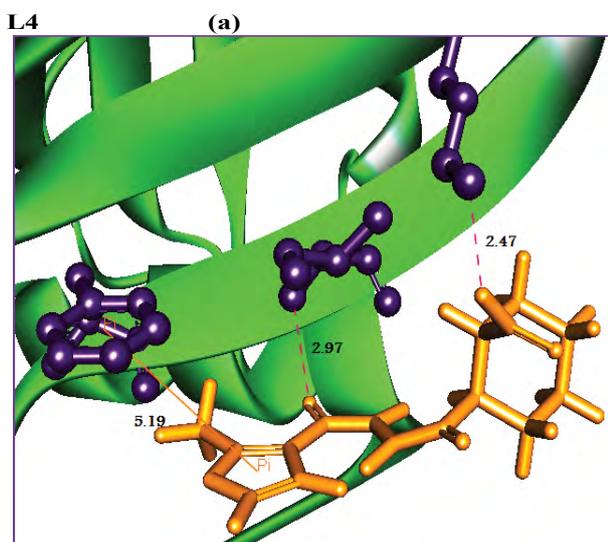
The ligands with the ADME properties in the permissible range, are considered as potent lead molecules. The predicted ADME properties

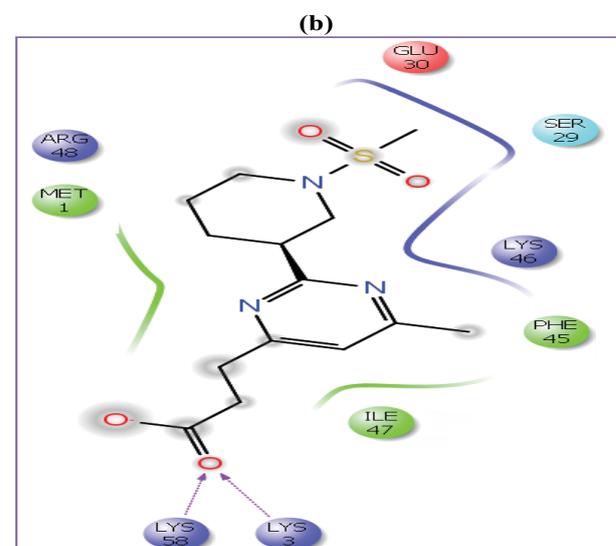
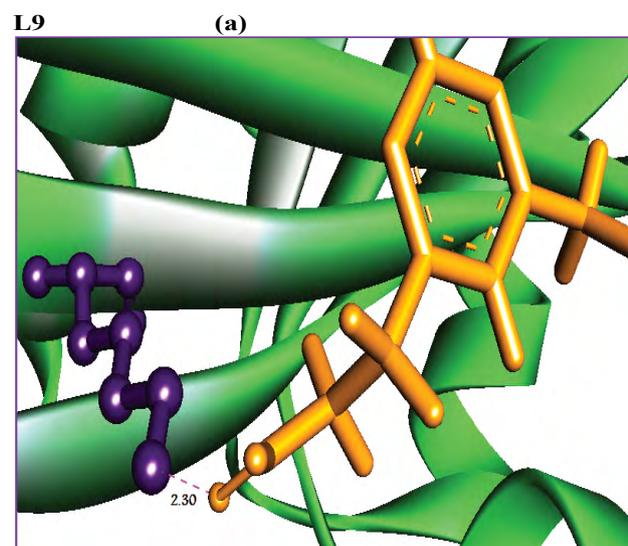
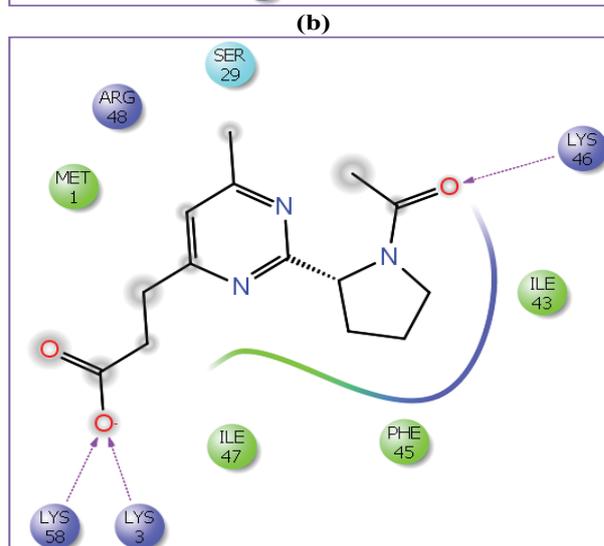
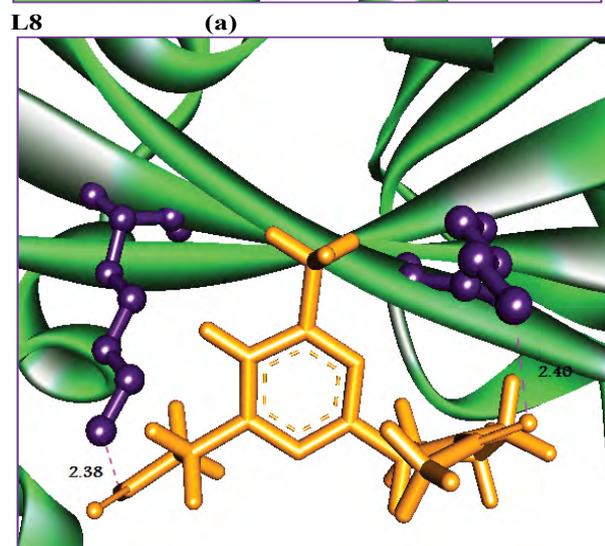
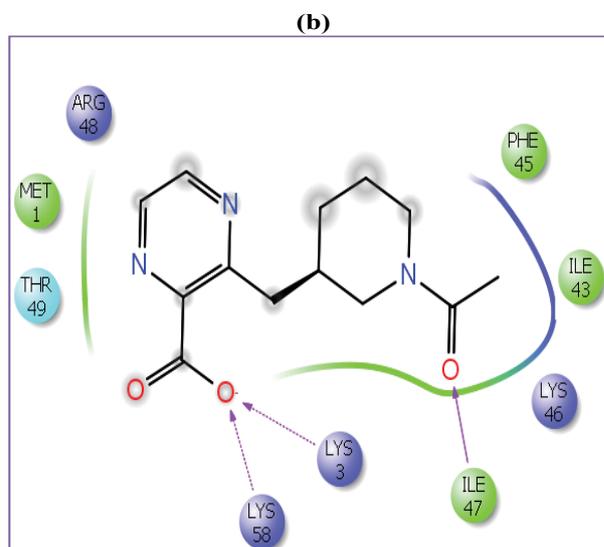
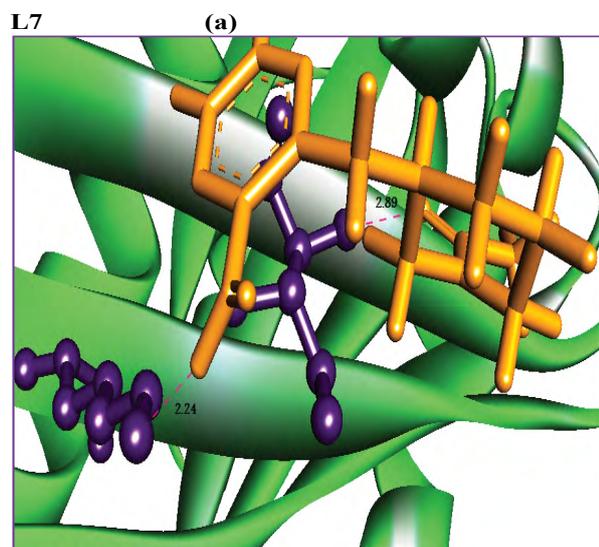
of these ligands are listed in Table 10. The molecules L1-L10 have acceptable range of stars (0-5), which fall in the range of 95% of the existing drug molecules. All molecules are within the acceptable range of a QlogP of n-Octanol/ water value (i.e., <5). The screened molecules show good percentage of human absorption, and obey the Lipinski's, Rule of Five (RoF) and Jorgenson's, Rule of Three (RoT). These results suggest that the prioritized molecules are with drug like properties and can be considered as novel leads that can inhibit Rab8b protein. In summary, we presented in this study, the structural information on Rab8b protein which may be a novel drug target for studies aimed to developing inhibitors of cancer, and the details of the screened ligands which have drug like properties.

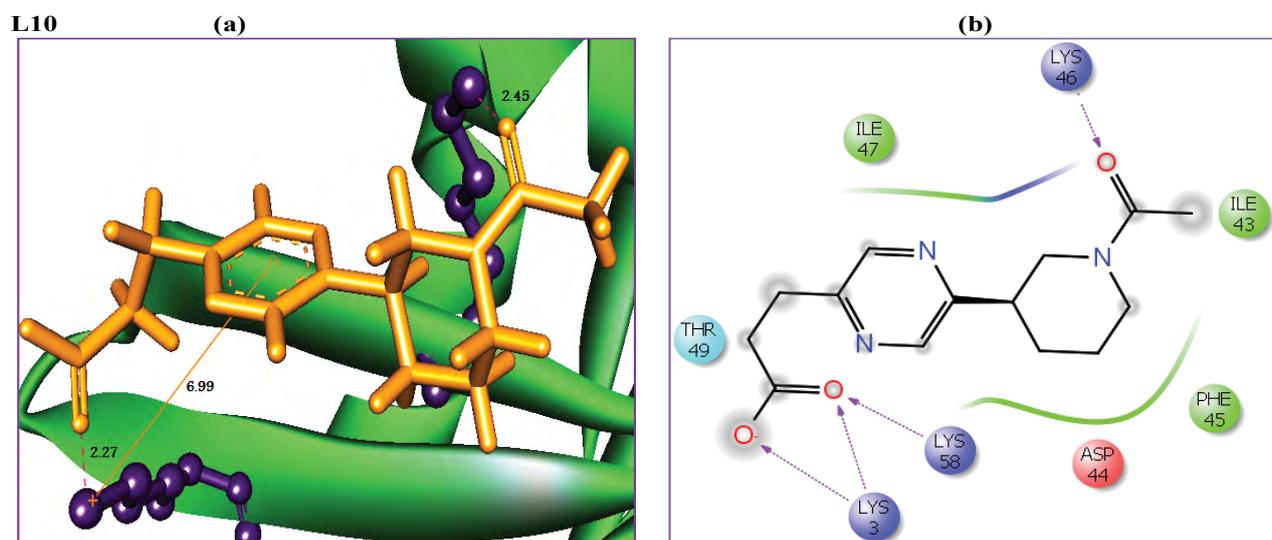
Conclusion

The Rab proteins are implicated in multiple aspects of tumour progression, and represent new targets for discovery of anticancer therapies. We have targeted Rab8b, a new member of the human Rab family, overexpressed in testis carcinoma. Our study provides the 3D structure of Rab8b with 207 amino acid residues through homology modeling approach. The energy minimization and other validation procedures were carried out to check the stability of the target protein 3D model. Our docking study with ligand databases shows that the carboxamide group and pyrimidine moiety in the ligands to be commonly participated in the formation of hydrogen bonds with LYS46, ILE47 and LYS58 of Rab8b protein. In addition, we observed that the amino acid residues PHE37, SER39, ASP44, PHE45, LYS46, ILE47 and LYS58 play a major role in protein-ligand interactions. Using virtual screening studies, we identified ligand molecules with better docking score, glide energy and acceptable ADME properties, as potential inhibitors of Rab8b. The present work helps in the identification of ligand molecules as drug candidates for cancer therapy.



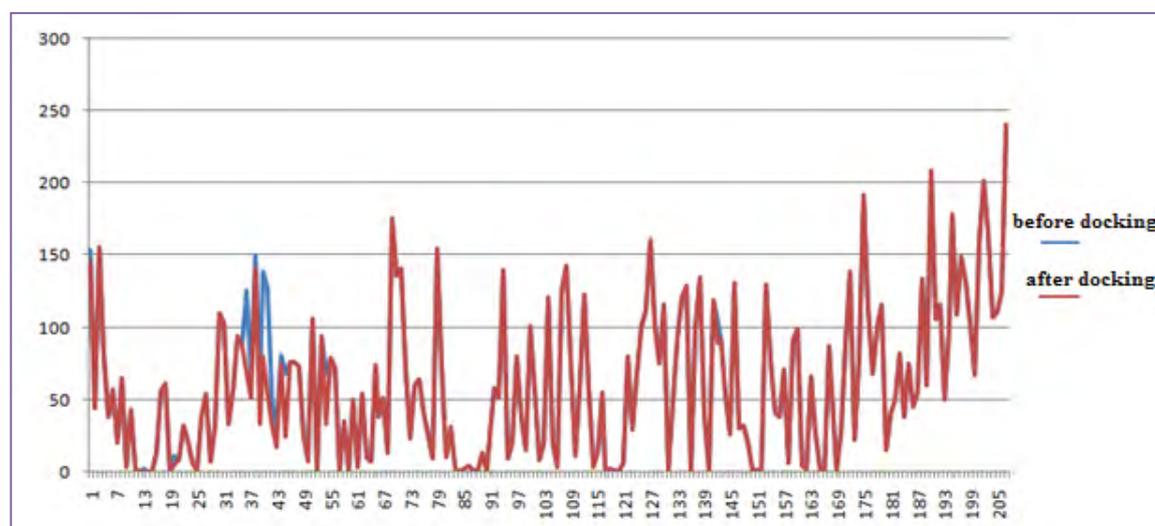






(a) The binding residues of Rab8b protein are shown in violet ball & sticks with labels, the hydrogen bonds are represented by pink dotted lines, and π - π in orange lines. The ligand molecules are shown in orange stick model.
 (b) The amino acids are shown in 3 letter code, H-bonds in pink lines, and the π - π in green colour.

Figure 12: The docked structures of top-ten hits against Rab8b protein.



The Solvent Accessible Surface Area value of Rab8b protein (before docking) is represented in blue colour peaks and the protein-ligand (L5) complex (after docking) is represented in red colour peaks. X-axis represents the amino acid residues and Y-axis represents SASA value.

Figure 13: SASA values of Rab8b protein and Rab8b protein-ligand (L5) complex.

References

- Evan GI, Vousden KH (2001) Proliferation, cell cycle and apoptosis in cancer. *Nature* 411: 342-348.
- Scott AM, Wolchok DJ, Old JL (2012) Antibody therapy of cancer. *Nat Rev Cancer* 12: 278-287.
- Siegel R, Ma J, Zou Z, Jemal A (2014) Cancer statistics. *CA Cancer J Clin* 64: 9-29.
- David C, Juanru G, William AG (1997) RAB GTPases expressed in human melanoma cells. *Biochim Biophys Acta* 1355: 1-6.
- Lau AS, Mruk DD (2003) Rab8b GTPase and Junction Dynamics in the Testis. *Endocrinology* 144: 1549-1563.
- Mitchell RS, Kumar V, Abbas AK, Fausto N (2007) Box on morphology of squamous cell carcinoma. *Robbins Basic Pathology*. 8th edn. Philadelphia, Saunders, USA.
- Wai WC, Dye DE, Coombe DR (2012) The Role of Immunoglobulin Superfamily Cell Adhesion Molecules in Cancer Metastasis. *Int J Cell Bio* 2012: 340296.
- Stenmark H, Olkkonen VM (2001) The Rab GTPase family. *Genome Bio* 2: 3007.1-3007.7.
- Asher M (2007) A spate of Rab8. *Nat Rev Mol Cell Bio* 8: 600-601.
- He H, Dai F, Yu L, She X, Zhao Y, et al. (2002) Identification and Characterization of Nine Novel Human Small GTPases Showing Variable Expressions in Liver Cancer Tissues. *Gene Expr* 10: 231-242.
- Sasin JM, Bujnicki JM (2004) COLORADO3D, a web server for the visual analysis of protein structures. *Nucleic Acids Res* 32: 586-589.

12. Elmar K, Sander BN, Gert V. Homology modeling. Structural Bioinformatics. ISBN 0-471-20199-5.
13. Liao C, Sitzmann M, Pugliese A, Nicklaus MC (2011) Software and resources for computational medicinal chemistry. *Future Med Chem* 3: 1057-1085.
14. Boeckmann B, Bairoch A, Apweiler R, Blatter MC, Estreicher A, et al. (2003) The Swiss-Prot Protein knowledgebase and its supplement TrEMBL. *Nucleic Acids Res* 31: 365-370.
15. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389-3402.
16. Kelley LA, Sternberg MJE (2009) Protein structure prediction on the web: a case study using the Phyre server. *Nat Protoc* 4: 363-371.
17. Cuff JA, Clamp ME, Siddiqui AS, Finlay M, Barton GJ (1998) Jpred: a consensus secondary structure prediction server. *Bioinformatics* 14: 892-893.
18. Contreras-Moreira B, Bates PA (2002) Domain Fishing: a first step in Protein comparative modeling. *Bioinformatics* 18: 1141-1142.
19. Myers EW, Miller W (1988) Optimal alignments in linear space. *Comput Appl Biosci* 4: 11-17.
20. Gonnet GH, Cohen MA, Benner SA (1992) Exhaustive matching of the entire protein sequence database. *Science* 256: 1443-1445.
21. Martí-Renom MA, Stuart AC, Fiser A, Sánchez R, Melo F, et al. (2000) Comparative protein structure modeling of genes and genomes. *Annual Review of Biophysics and Biomolecular Structure* 29: 291-325.
22. Sali A, Blundell TL (1993) Comparative protein modeling by satisfaction of spatial restraints. *J Mol Biol* 234: 779-815.
23. Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-pdb viewer: an environment for comparative protein modeling. *Electrophoresis* 18: 2714-2723.
24. Jorgensen WL, Maxwell DS, Tirado-Rives J (1996) Development and testing of the OPLS all-atom force field on conformational energetic and properties of organic liquids. *J Am Chem Soc* 118: 11225-11236.
25. Salam NK, Nuti R, Sherman W (2009) Novel method for generating structure based pharmacophores using energetic analysis. *J Chem Inf Model* 49: 2356-2368.
26. Wiederstein M, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res* 35: 407-441.
27. Laskowski RA, Rullmann JA, MacArthur MW, Kaptein R, Thornton JM (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structure solved by NMR. *J Biomol NMR* 8: 477-486.
28. Colovos C, Yeates TO (1993) Verification of protein structures-patterns of non-bonded atomic interactions. *Protein Sci* 2: 1511-1519.
29. Powers R, Copeland JC, Germer K, Mercier KA, Ramanathan V, et al. (2006) Comparison of protein active site structures for functional annotation of proteins and drug design. *Proteins* 65: 124-135.
30. Ubersax JA, Ferrell JE (2007) Mechanisms of specificity in protein phosphorylation. *Nat Rev Mol Cell Bio* 8: 530-541.
31. Halgren T (2007) New method for fast and accurate binding-site identification and analysis. *Chem Biol Drug Des* 69: 148-168.
32. Dundas J, Ouyang Z, Tseng J, Binkowski A, Turpaz Y, et al. (2006) CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res* 34: 116-118.
33. Brylinski M, Feinstein WP (2013) efindsite: Improved prediction of ligand binding sites in protein models using meta-threading, machine learning and auxiliary ligands. *J Comput Aided Mol Des* 27: 551-567.
34. Feinstein WP, Brylinski M (2014) efindsite: Enhanced fingerprint-based virtual screening against predicted ligand binding sites in protein models. *Mol Info* 33: 135-150
35. Kawatkar S, Wang H, Czerminski R, Joseph-McCarthy D (2009) Virtual fragment screening: an exploration of various docking and scoring protocols for fragments using Glide. *J Comput Aided Mol Des* 23: 527-539.
36. Vajda S, Vakser IA, Sternberg MJE, Janin J (2002) Modeling of Protein Interactions in Genomes. *Proteins* 47: 444-446.
37. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ (2005) PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res* 33: 363-367.
38. Amid MR, Mehdi S, Hamid P, Sayed AM (2008) Impact of residue accessible surface area on the prediction of protein secondary structures. *BMC Bioinformatics* 9: 1-11.
39. Accelrys Discovery Studio Visualizer (2010) v 3.5. San Diego: Accelrys Software Inc., San Diego.
40. Manju K, Jat RK, Gunjan P, Anju G (2012) Review on introduction to molecular docking software technique in medicinal chemistry. *Int J Drug Res Tech* 2: 189-197.
41. Chen IJ, Foloppe NJ (2010) Drug-like bioactive structures and conformational coverage with the LigPrep/ ConfGen suite: comparison to programs MOE and catalyst. *J Chem Inf Model* 50: 822-839.
42. Shivakumar D, Williams J, Wu Y, Damm W, Shelley J, et al. (2010) Prediction of absolute solvation free energies using molecular dynamics free energy perturbation and the OPLS force field. *J Chem Theory Comput* 6: 1509-1519.
43. Lengauer T, Rarey M (1996) Computational methods for biomolecular docking. *Curr Opin Struct Biol* 6: 402-406.
44. Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, et al. (2006) Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J Med Chem* 49: 6177-6196.
45. Konstantin VK, Frank T, Timo S, Wolfgang W (2011) Derivatives of Molecular Surface Area and Volume: Simple and Exact Analytical Formulas. *J Comput Chem* 32: 2647-2653.
46. Selić HE, Beresford AP, Tarbit MH (2002) The emerging importance of predictive ADME simulations in drug discovery. *Drug Discov Today* 7: 109-116.
47. Matter H, Baringhaus KH, Naumann T, Klabunde T, Pirard B (2001) Computational approaches towards the rational design of drug-like compound libraries. *Comb Chem High Throughput Screen* 4: 453-475.
48. Loakimidis L, Thoukydidis L, Mirza A, Naeem S, Reynisson J (2008) Benchmarking the reliability of Qikprop, Correlation between experimental and predicted values. *QSAR & Combinat Sci* 27: 445-446.
49. Uniprot Consortium (2013) Update on activates at the Universal Protein Resource (UniProt). *Nucleic Acids Res* 41: 43-47.
50. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic Local Alignment Search Tool. *J Mol Biol* 215: 403-410.
51. Altschul SF, Madden TL, Schaffer A, Zhang J, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389-3402.
52. Sali A, Pottertone L, Yuan F, Van VH, Karplus M (1995) Evaluation of comparative protein modeling by MODELLER. *Proteins* 23: 318-326.
53. Umamaheswari A, Pradhan D, Hemanthkumar M (2010) Virtual screening for potential inhibitors of homology modeled Leptospira interrogans MurD Ligase. *J Chem Biol* 3: 175-187.
54. The PyMOL Molecular Graphics System (2010) Version 1.3 Schrödinger, LLC, New York, USA.
55. Laskowski RA, Hutchinson EG, Michie AD, Wallace AC, Jones ML, et al. (1997) PDBsum: a Web-based database of summaries and analyses of all PDB structures. *Trends Biochem Sci* 22: 488-490.
56. Bosshard HR, Marti DN, Jelesarov I (2004) Protein stabilization by salt bridges: concepts experimental approaches and clarification of some misunderstandings. *Journal of Molecular Recognition* 17: 1-16.