

RESEARCH

Open Access



Computer-aided identification of lung cancer inhibitors through homology modeling and virtual screening

Aboubakr Haredi Abdelmonsef 

Abstract

Background: Lung cancer is the most often event cancer around the world and the first leading cause of cancer death in human beings. Rab39a protein is implicated in vesicular trafficking and fusion of phagosomes with lysosomes. Rab39a is overexpressed in lung cancer, which converts normal cells to abnormal cells that reproduce quickly, and resists programmed cell death that usually kills aberrant cells.

Aim: In the present study, the structure-based drug discovery approach is applied to identify new lead structures as cancer drug candidates against Rab39a.

Methods: A valid three-dimensional (3D) model of Rab39a generation, the prediction of protein–protein interactions (Rab39a/DENND5B) and active site identification were achieved by computational techniques.

Results: Our studies suggest that the amino acid residues from PHE28 to LYS63 are important for binding with the ligand molecules. Subsequently, the virtual screening study was carried out with ligand databases against the active site of Rab39a.

Conclusion: The ligand molecules with hetero amine moieties and amide group (-CONH-) have shown good value of docking score and agreeable ADME properties, so they were prioritized as potential inhibitors of Rab39a protein. Hence, Rab39a has emerged as a therapeutic target for drug development towards lung cancer.

Keywords: Lung cancer, Rab39a, Homology modeling, Virtual screening, ADME

Background

Cancer is a genetic disease caused by many changes within the DNA of a cell, which lead to conversion of normal cells to malignant cells [1]. Lung cancer is the first leading cause of cancer death in Homo sapiens [2]. Rab proteins are key mediators of regulation of vesicle transport pathway, which is fundamental to all cells [3]. Rab family proteins are active in the macromolecular transport across plasma

membrane along exo- and endocytic pathways. However, the best-characterized function of these GTPases is in membrane fusion, where they have an active role in tethering the compartments that are going to be fused [4]. Rab39a protein is expressed ubiquitously in the cells of human beings. Human Rab39a controls the regulation of Golgi-associated vesicular transport along the endocytosis pathway [5]. Cancer is associated with the functional

Correspondence: aboubakr.ahmed@sci.svu.edu.eg

Department of Chemistry, Faculty of Science, South Valley University, Qena 83523, Egypt

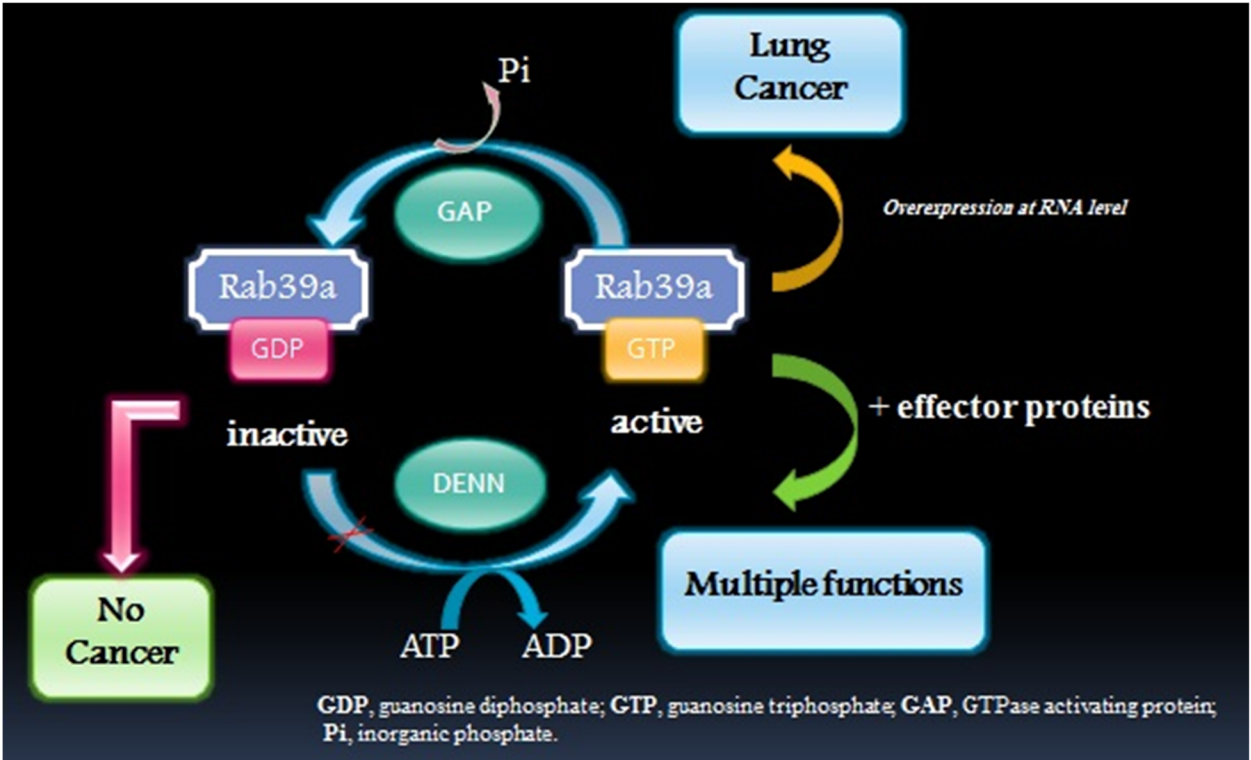


Fig. 1 Biochemical pathway of Rab39a: The human Rab39a cycle between the active GTP-bound (membrane-associated) and inactive GDP-bound (cytosol-associated) conformations using DENN as guanine nucleotide exchange factor (GEF) and GTPase-activating protein (GAP). Rab39a is involved in regulation of cellular endocytosis pathway along Golgi-associate vesicular transport. Alteration in Rab39a has been implicated in causing lung cancer [5, 8, 9]

impairments of Rab signalling pathways [6]. An alteration in the regulation of the vesicular traffic is responsible for tumour progression. The present work is based on the application of computational approaches for identifying new antagonists against Rab39a, which may be used as lung cancer therapeutics.

Role of Rab39a protein

The role of RabGTPases, in regulating membrane trafficking between organelles, is in the recruitment of

effector proteins [7]. The functional cycle of Rab39a is shown in Fig. 1, which illustrates that the target protein is upregulated (activation) by GEF (guanine nucleotide exchange factors) family proteins and downregulated (deactivation) by GAPs (GTPase-activating proteins) [10, 11]. DENND5B acts as Rab39aGEF for facilitating the activation of the target [8]. The active form of Rab39a is associated with the trafficking of and interaction between endosomal compartments. Aberration in Rab39a and/or its effector function is implicated in progression

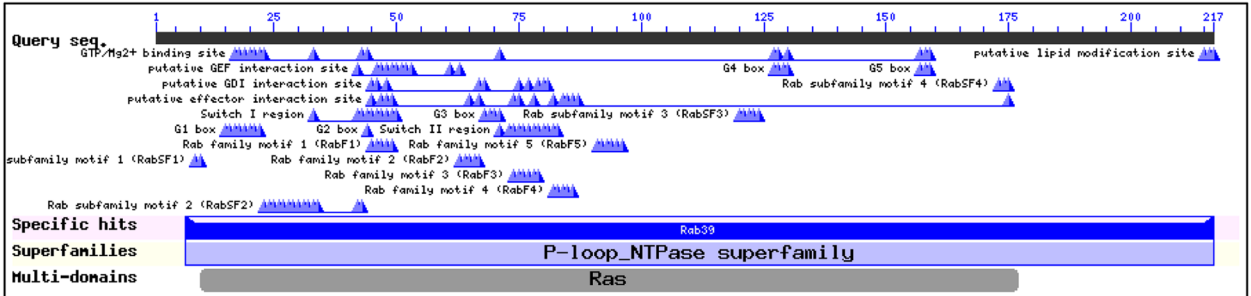


Fig. 2 The conserved domains of Rab39a protein: Basic local alignment search tool [12] illustrates that the residues from 28 to 54 are important for putative GEF interaction site

Table 1 Template search results for Rab39a

Name of server	Parameters for template selection	<i>E</i> values	PDB code
BLAST	Sequence similarity	5e–46	4D0L
Phyre2	Protein fold recognition (threading)	2e–27	4D0L
Jpred4	Secondary structure prediction	3e–35	4D0L

of lung cancer [5, 9]. In this study, the activation pathway of Rab39a is targeted to identify the antagonists against lung carcinoma, using computational techniques.

Methods

Homology modelling

The 3D structure of the target protein is generated by a homology modeling technique. The amino acid sequence of Rab39a (ID: Q14964) is retrieved from UniProt database (<https://www.uniprot.org/uniprot/Q14964>). The identification of suitable homologous template structures for the modeling of the target protein is carried out using different server tools, namely, BLASTp (Basic Alignment Tool program), Phyre2 (Protein Homology/analogy Recognition Engine V 2.0), JPred4 (Java Prediction V 4.0) and Domain Fishing servers [12–15]. The sequence identity

and the statistical measure *E* value are the parameters used for the selection of the homologous template. The sequence alignment is performed by subjecting the amino acid sequences of the target and template proteins to ClustalW server tool [16], to identify the structurally conserved and similar regions, applying protein weight matrix GONNET [17]. Some models of the target protein are generated using a protein structure modeling program, Modeller 9.11 [18], and the model with the lowest modeler objective function is selected for further validation.

Model validation

The generated 3D model of Rab39a is refined by Ramachandran plot (RC) and protein structure analysis (ProSA) servers. RC plot statistics gives the Phi (ϕ) versus Psi (Ψ) angle distribution of residues of the protein, to predict its stereochemical validity [19]. The quality of 3D model is further evaluated using ProSA server [20].

Binding regions in Rab39a

The hydrophobic cavities in the protein that act as an active site are important for binding a variety of ligands and groups [21].

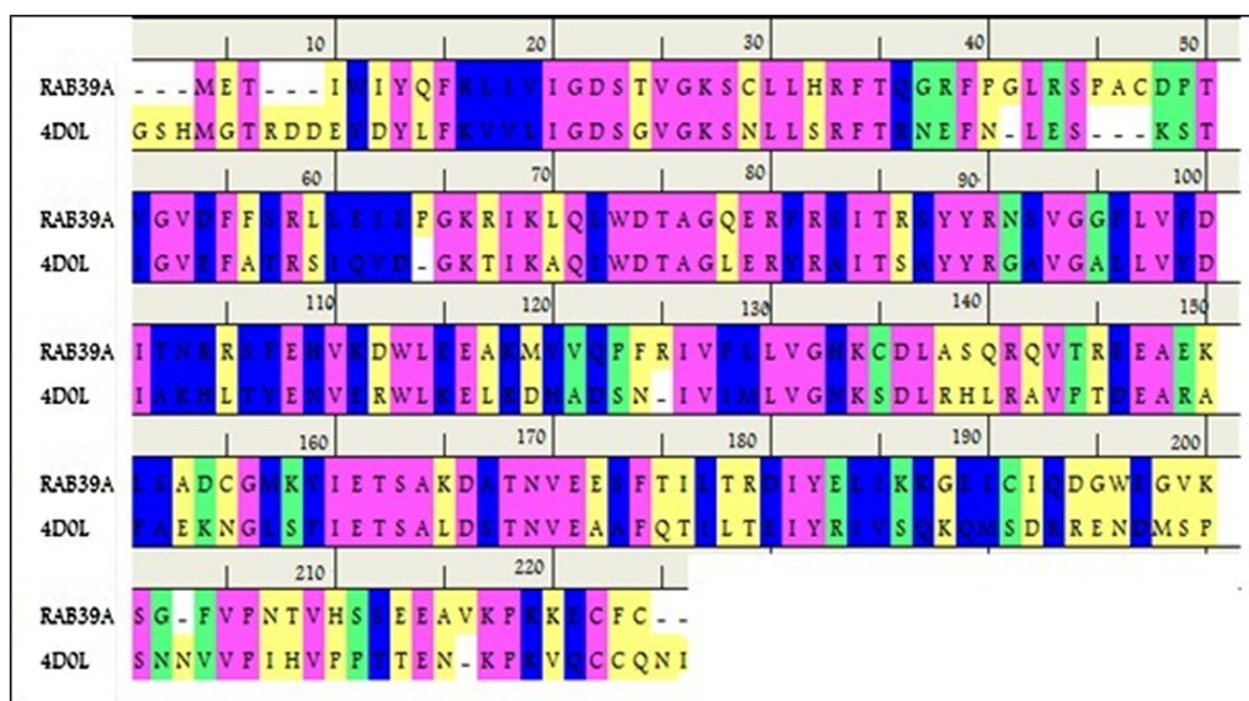


Fig. 3 Sequence alignment of Rab39a with template protein: The pairwise sequence alignment between two proteins was carried out using ClustalW [16]. The conserved residues (96.45%) are represented in pink color, while highly and weakly similar residues (3.55%) in blue and green color, respectively

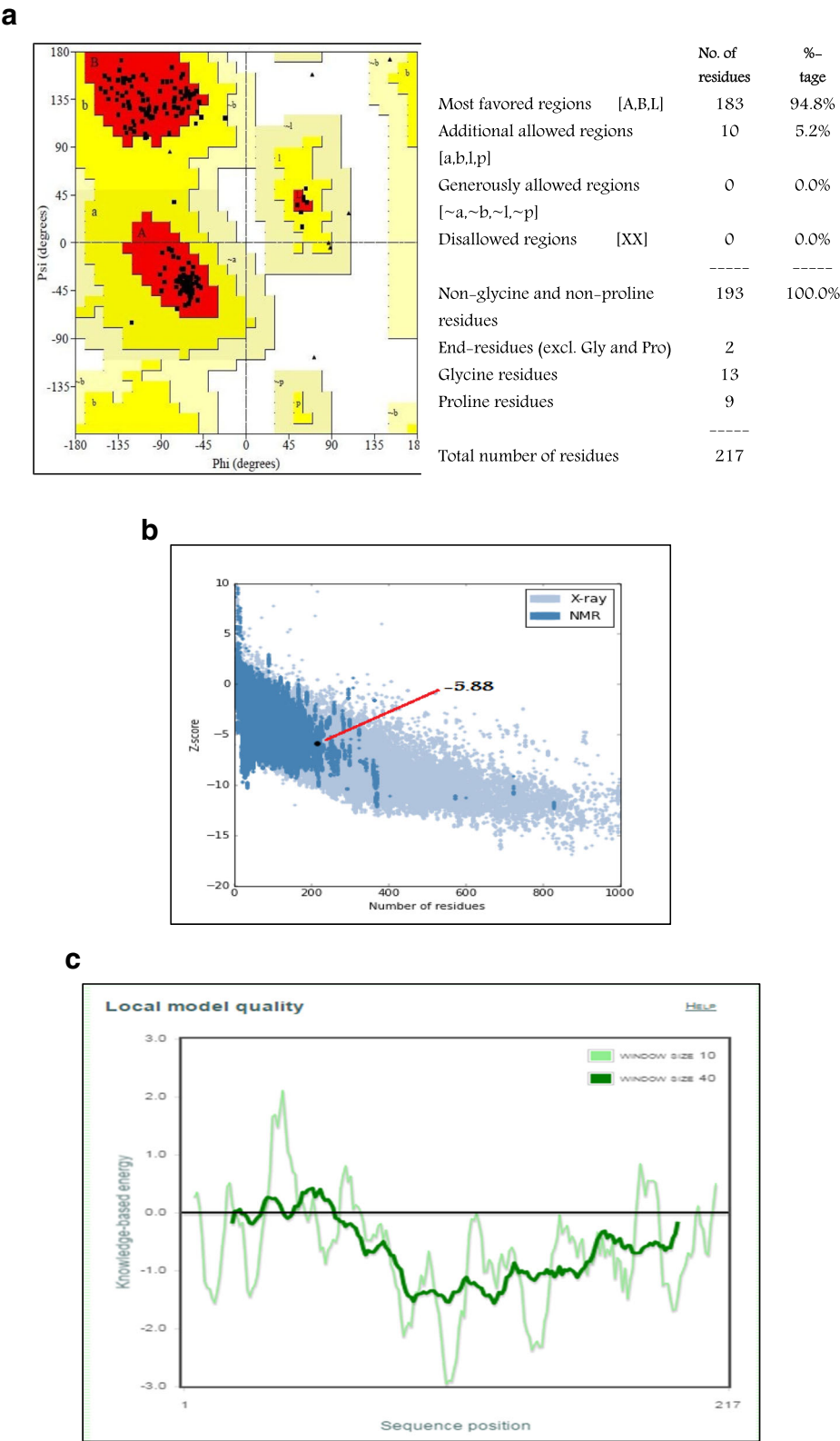


Fig. 4 (See legend on next page.)

(See figure on previous page.)

Fig. 4 a The stereochemical analysis of Rab39a: The red area represents most favorable region of residues, yellow region is additionally allowed and light yellow region is generously allowed. The RC plot [19] represents 94.8% of residues in most favored region indicating a good quality model. **b** The local model quality of Rab39a: The protein score (−5.88) falls in the range of PDB proteins submitted by NMR (dark blue region) and X-ray crystallography (light blue region), indicating a good quality model [19]. **c** ProSA energy plot of Rab39a: The graph shows local model quality by plotting energies as a function against amino acid sequence position. The ProSA [20] of the model shows maximum residues in the negative energy region. The negative ProSA energies of Rab39a indicate a reliable arrangement of residues in the 3D model

Binding site identification

The binding cavities are identified using computational tools such as Computed Atlas of Surface Topography of protein (CASTp) [22] and Sitemap module [23] of Schrodinger. CASTp identifies the binding sites, their respective volumes and areas. In addition, sitemap predicts the characteristics of ligand binding sites.

Protein–protein docking

Protein–protein interactions (PPIs) are critical for all cellular pathways and signal transduction [24]. The Rab39a/DENND5B interactions are examined by in silico protein–protein docking studies using patch-Dock sever Beta V 1.3 [25], and the results are corroborated with the binding pockets identified from computational prediction tools. Accelrys Discovery

Studio Visualizer 3.5 [26] is used to visualize the Rab39a/DENND5B intermolecular interactions.

Virtual screening

Virtual screening study is considered as an important concept to identify new drug-like compounds [27].

Protein optimization

The energy minimization is a basic process to prepare chemically correct and fully optimized protein structure [28]. The process is carried out using protein preparation wizard in Schrodinger suite, which is performed using an all-atom Impact Refinement (Impref) (Impact v 5.0, Schrodinger, NY), to adjust steric clashes and to remove water molecules [29].

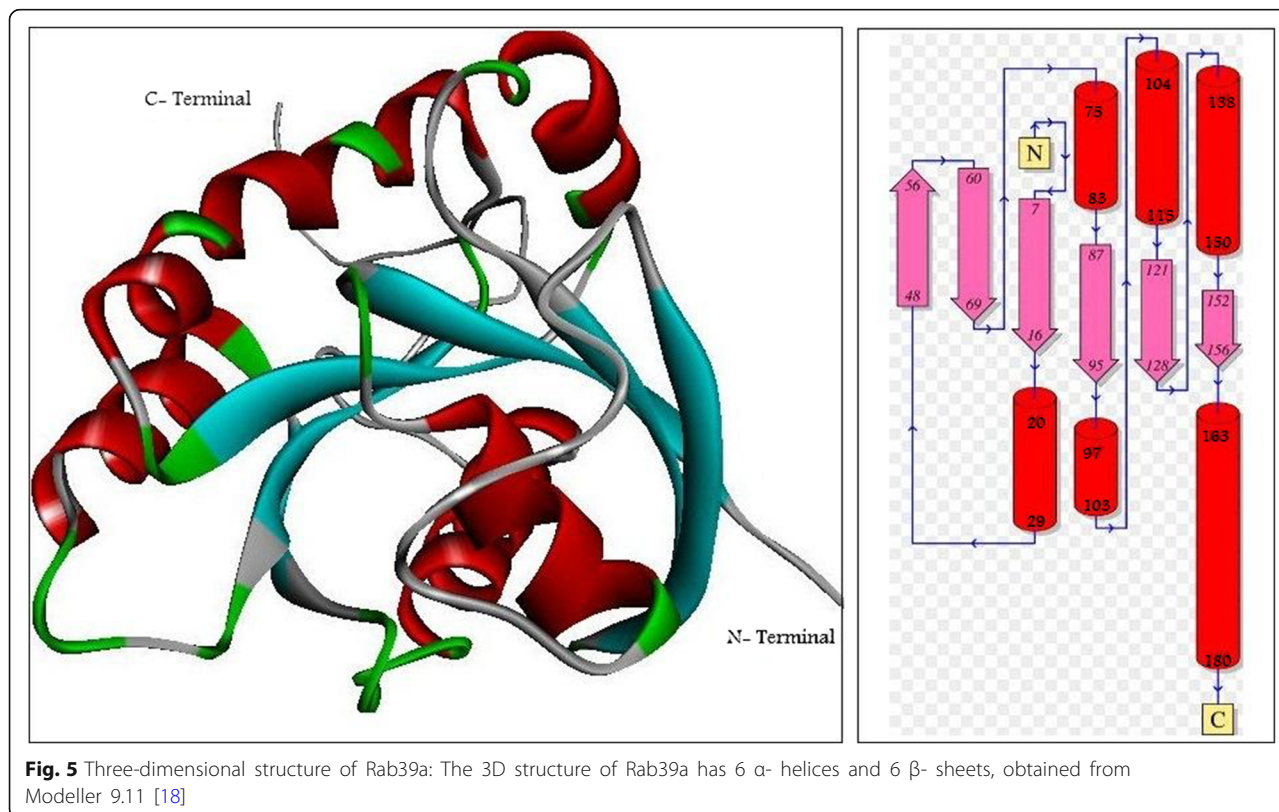


Table 2 Secondary structure details of the Rab39a: the α -helices in Rab39a

No.	Start	End	No. residue	Length(A°)	Amino acid sequence
1	Lys21	Phe28	8	12.24	KSCLLHRF
2	Arg76	Tyr82	7	11.55	RSITRSY
3	Arg98	Glu102	5	8.38	RRSFE
4	Val104	Tyr114	11	16.53	VKDWLEEKMY
5	Arg139	Cys149	11	16.99	REEAEKLSADC
6	Val164	Ile179	16	25.01	VEESFTILTRDIYELI

The amino acid sequences forming the alpha helices in the modeled Rab39a

Ligand preparation

The ligands are subjected to Schrödinger's Ligprep module [30], to produce different conformers, based on their ionization states and stereochemistry. Partial atomic charges are computed by using the OPLS-2005 (optimized potentials for liquid simulations) force field [31].

Docking and solvent accessible surface area (SASA)

Molecular docking is a computational tool which helps in the prediction of binding abilities between Rab39a and ligands [32]. The ability of protein to interact with small molecules controls a significant part of the protein dynamics, which may inhibit their biological function [33]. The potential binding regions in the target are described using solvent accessible surface area (SASA) [34]. The SASA values of Rab39a and ligand molecules are analysed before and after docking, using Discovery Studio Visualizer 3.5 [26].

ADME (adsorption, distribution, metabolic and excretion)

Structures with unfavorable ADME have been reported to be a reason for failure of the drug candidate molecules in drug design [35]. The pharmacokinetic properties of the ligands are calculated by subjecting them to QikProp [36] module (QikProp version 3.3) [37].

Table 3 Secondary structure details of the Rab39a: the β -sheets in Rab39a

No.	Start	End	No. residue	Length(A°)	Amino acid sequence
1	Tyr7	Gly15	9	13.56	YQFRLVIG
2	Asp48	Glu55	8	12.22	DFFSLRLE
3	Arg61	Asp68	8	12.25	RIKLQLWD
4	Gly88	Asp94	7	11.50	GGFLVFD
5	Val121	His127	7	11.35	VFLLVGH
6	Lys152	Thr156	5	8.27	KYIET

The amino acid sequences forming the beta sheets in the modeled Rab39a

Results

In the current study, the homology modeling technique for evaluating the Rab39a structure was applied, and the virtual screening of ligand molecules against Rab39a was performed to identify novel inhibitors which may help in cancer therapy.

Template identification and sequence alignment

The homology modeling technique was applied to generate 3D structure of the target, based on template. The FASTA sequence of Rab39a (217 amino acid residues) was retrieved from UniProtKB [38]. The plausible template was identified by submitting the FASTA sequence to BLAST, Phyre2, Jpred4 (Java Prediction V 4.0) and Domain Fishing server tools. BLAST illustrates the residues ranging from PHE28 to LYS63 to be important for putative GEF interaction site. The residues from SER17 to SER157 are responsible for GTP/Mg²⁺ binding, as presented in Fig. 2. For many Rab family proteins, Mg²⁺ ion is necessary for guanine nucleotide binding and GTP hydrolysis [39]. Phyre2 was used to identify the templates with analogue fold structures. The secondary structure prediction of Rab39a was analysed using Jpred4. The results are shown with their corresponding *E* values in Table 1. Domain fishing tool shows the P-loop NTPase fold (G¹⁵XXXXGKS²²), which is the most prevalent domain of several distinct nucleotide-binding protein folds. The homologous protein sequence obtained for Rab39a with low *E* value was considered as a template [40]. A low *E* value represents a high-sequence identity. 4D0L was selected depending on the maximum identity (95%), statistical *E* value (2×10^{-57}) and query coverage (93%) with that of Rab39a. The results of ClustalW show that 96.40% are conserved and 3.60% are strongly/weakly similar residues as represented in Fig. 3.

3D model building

Thirty models were initially built with a modeler objective function ranging from 1085 to 1260. The model with the lowest probability density function (1085) was selected for further refinement.

Structural refinement and validation

The structure of Rab39a was subjected to Swiss-Pdb Viewer for loop modeling and energy minimization, by applying GROMOS96 Force-Field 4.1.0. The total energy calculated for the refined protein model was $-1716 \text{ kcal mol}^{-1}$. The stereochemical validity of Rab39a was assessed using the RC plot, as shown in Fig. 4a. The plot statistics shows 94.8% (183aa) of the total residues in the most favored regions and 5.2% residues (10aa) in

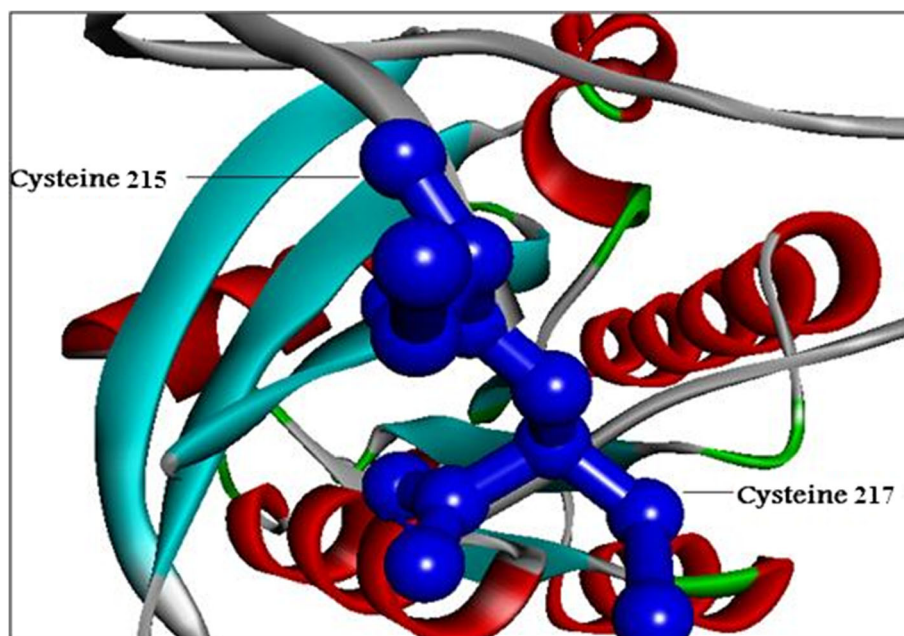


Fig. 6 The double-cysteine residues at terminal-C in Rab39a: The double cysteines 215 and 217 are represented in blue ball and stick model that help in the identification and stability of Rab39a [45]

additionally allowed regions, which implies that the generated 3D model is a good quality model. The ProSA Z-score value is compared with that for the proteins determined by X-ray crystallography (light blue) and NMR spectroscopy (dark blue) (Fig. 4b). The ProSA server results show a Z-score value of -5.88 for Rab39a structure. The score is within the range of the PDB proteins evaluated by X-ray and NMR, predicting a good quality model. The residues with negative ProSA energies are consistently good in quality. Most of residues of Rab39a have negative energies, as shown in Fig. 4c, indicating an acceptable model.

Structural features of Rab39a

The 3D model was visualized using PyMOL 1.3 software [41], which shows that the model consists of six α - helices and six β - sheets (Fig. 5). N-terminal indicates the starting residue, and C-terminal indicates the end residue. Similar computational protocol for the secondary structure details was reported earlier by Aboubakr et al. 2016 [42]. Tables 2 and 3 show sequences of helices and sheets in the modeled Rab39a. The residues are held together by intramolecular hydrogen bonds, salt bridges, π - π , π -cation and π -sigma interactions. The structure shows two

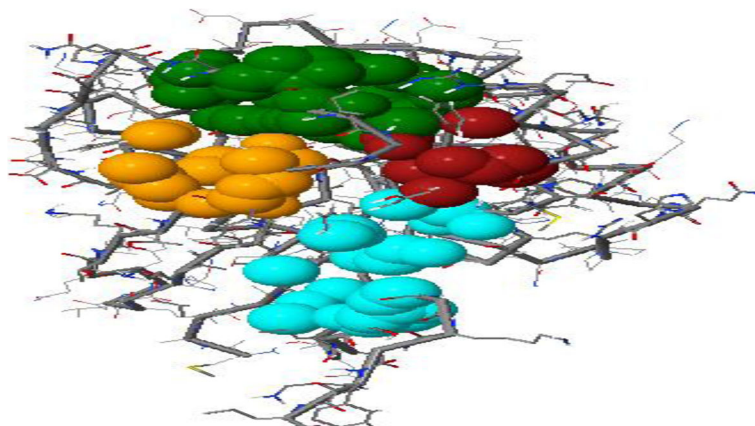


Fig. 7 The binding cavities of Rab39a obtained from CASTp server [22]: Site1 is represented in green, site 2 in orange, site 3 in cyan and site 4 in pink color

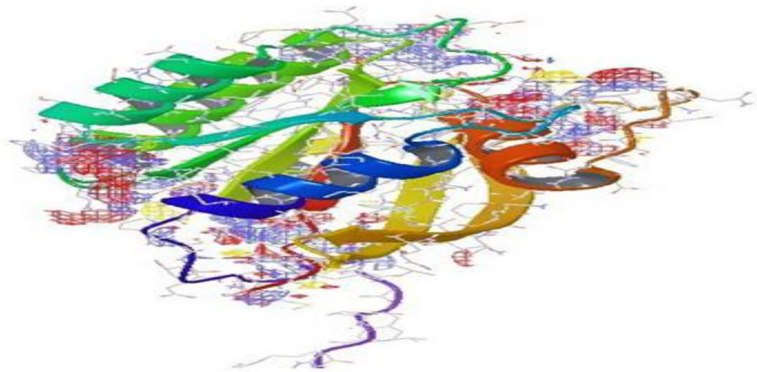


Fig. 8 Active site of Rab39a obtained from Sitemap module [23]: The hydrogen bond acceptor region is indicated in red, H-bond donor regions in blue and hydrophobic pockets in yellow color

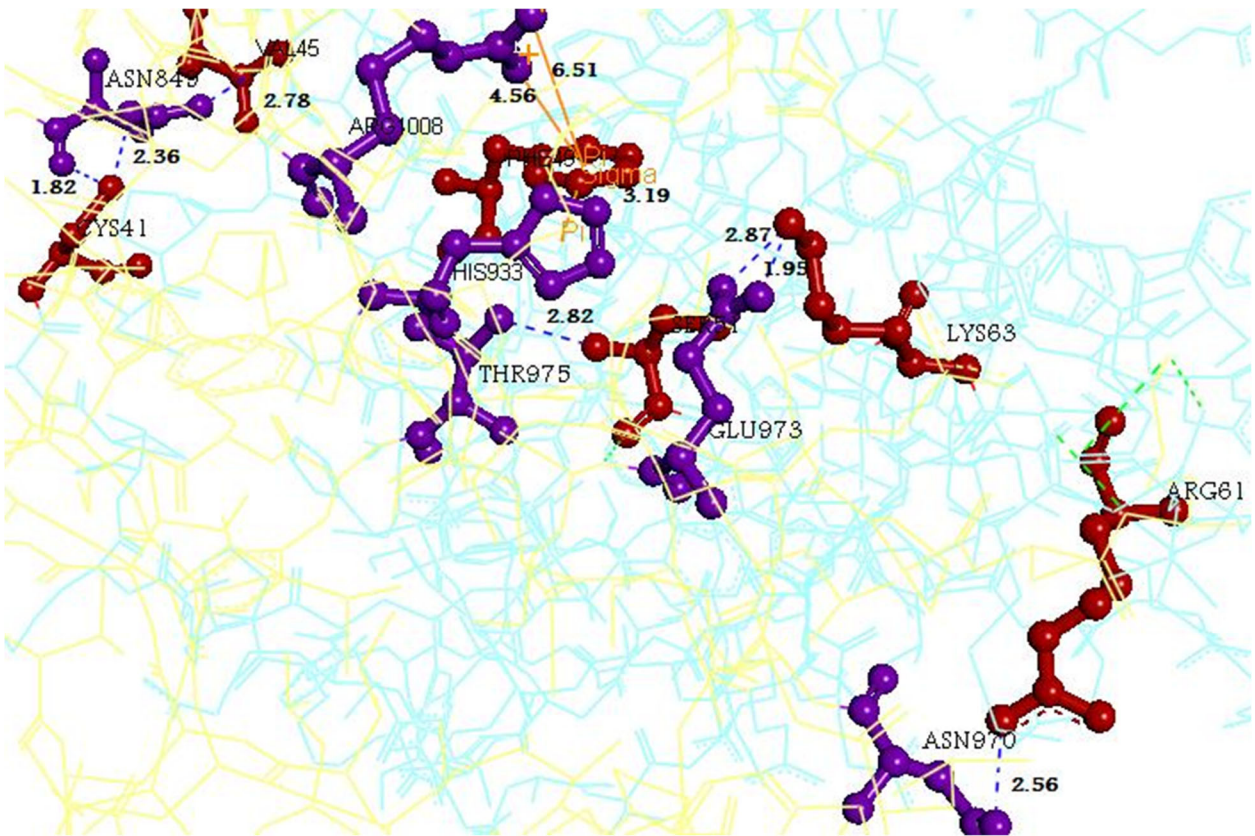


Fig. 9 The protein–protein interaction between Rab39a and natural substrate [25]: The Rab39a (yellow) interacts with DENND5B (green). The specific binding residues of Rab39a are shown in red-labeled sticks (CYS41, VAL45, PHE49, SER51, ARG61 and LYS63) in Rab39a. The residues of DENND5B are represented in violet color, while the H-bond interactions in blue color and π interactions in orange lines.

Table 4 Intermolecular interactions between Rab39a and natural substrate

	Binding residues	Distance of H-bonds (Å)
H-bonds	CYS41:SG - B:ASN849:O	1.82
	CYS41:SG - B:ASN849:OD1	2.36
	SER51:N - B:THR975:OG1	2.87
	SER51:OG - B:GLU973:OE2	1.87
	ARG61:NH2 - B:ASN970:OD1	2.56
	LYS63:NZ - B:GLU973:OE1	1.96
	LYS63:NZ - B:GLU973:OE2	2.82
	ASN849:ND2 - :VAL45:O	2.78
Pi-cation	PHE49 - B:ARG1008:NH1	6.51
	PHE49 - B:ARG1008:NH2	4.56
Pi-sigma	B:HIS933 - :PHE49:CE2	3.19

CYS41, VAL45, PHE49, SER51, ARG61 and LYS63 residues in Rab39a bind with natural substrate DENND5B, analysed using Discovery studio

pairs of amino acids that form π - π , ten π -cation and one π -sigma interactions, which add stability to the protein structure [43].

The Rab-specific regions

The highly variable C-terminal end of Rab proteins is the main structural determinant for their specific targeting of distinct cellular membranes [44]. The two C-terminal cysteine residues, found as XXCXC combination, form the Rab prenylation motifs with isoprenyl group, to facilitate the attachment to cell membrane [45]. The presence of double-cysteine motif helps in the folding and stability of target Rab39a, as shown in Fig. 6.

Binding site prediction

The binding sites were predicted by computational techniques as illustrated in Figs. 7 and 8. The results declare that PHE28 to LYS63 are important for binding with ligands. The aim of the PPI study of Rab39a with its natural substrate DENND5B is to find the nature of association of the two interacting proteins. The molecular interactions of two proteins were examined using patchDock sever Beta V 1.3 (Fig. 9, Table 4), which shows CYS41, VAL45, PHE49, SER51, ARG61 and LYS63 of Rab39a interact with ASN849, ASN849, HIS933, ARG1008, THR975, ASN970 and GLU973 of DENND5B, respectively. The results were corroborated and compared with that identified from computational prediction tools, which show that residues ranging from PHE28 to LYS63 are important for Rab39a binding.

Structure-based virtual screening, SASA and ADME analysis

Structure-based virtual screening technique helps in identifying new molecular entities (NMEs) which may act as effective and selective drug candidates. The binding site region of the 3D structure and ligands prepared from Sigma Tim Tec database were used in a screening study. A grid was generated around the binding region with $70\text{\AA} \times 70\text{\AA} \times 70\text{\AA}$ dimensions (Fig. 10) for specific docking [46]. The study was performed using workflow docking program of Glide incorporated in the Schrodinger package by Maestro, to predict the best binding modes. The ligands, 12690, were submitted to Ligprep module of Schrodinger suite. Four low-energy conformers were generated per ligand. The output of 17673

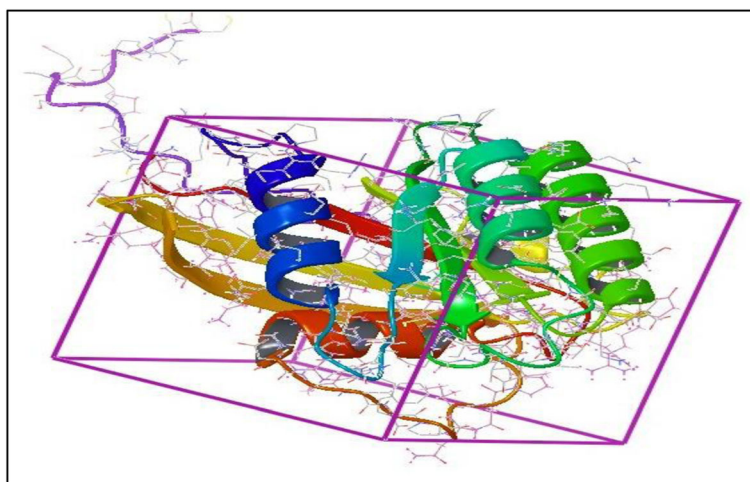
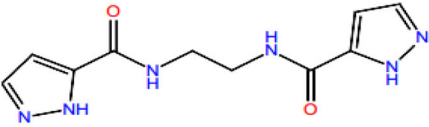
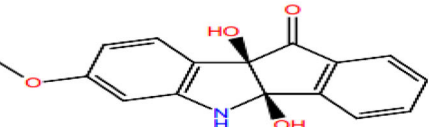
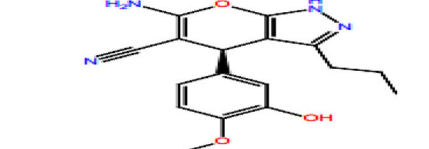
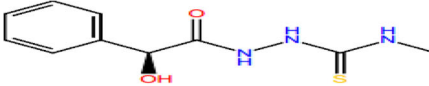
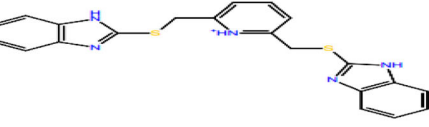
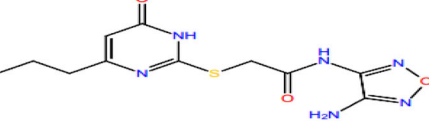
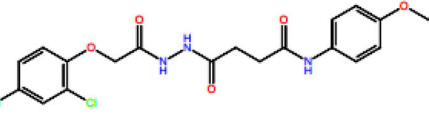
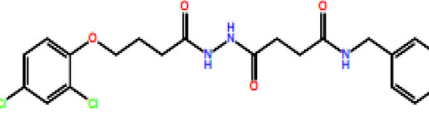


Fig. 10 Grid generation around binding site region: The grid generated with $70\text{\AA} \times 70\text{\AA} \times 70\text{\AA}$ dimensions using receptor grid generation of Glide module [46]

Table 5 The ligand structures and the intermolecular interactions with Rab39a protein

S. No	Structure	Glide score	Glide energy (K.Cal/mol)	Docked complex (amino acid–ligand atom)interactions	Distance (Å°)
M1		-7.31	-33.34	Hydrogen bonds PHE49:H - :M1 SER51:H - :M1 PHE49:O- :M1 SER51:O- :M1 Pi-Pi interaction PHE50 - :M1	2.09 1.75 2.16 1.95 4.96
M2		-7.08	-20.93	Hydrogen bonds SER51:H - :M2 SER51:O- :M2 SER51:O- :M2	1.85 2.00 1.86
M3		-6.93	-29.85	Hydrogen bonds GLN30:HE21- :M3 SER51:H - :M3 SER51:O- :M3	2.26 2.32 1.62
M4		-6.73	-29.35	Hydrogen bonds SER51:H - :M4 SER51:O- :M4 THR29:O- :M4	2.10 1.97 2.26
M5		-6.67	-47.55	Hydrogen bonds PHE28:O- :M5 SER51:O- :M5 ASP48:OD1- :M5 Pi-cation ARG27:NH1- :M5 ARG52:NH2- :M5	2.09 1.84 2.41 6.11 5.08
M6		-6.64	-35.55	Hydrogen bonds THR29:O- :M6 SER51:H - :M6 Pi-cation ARG52:NH1- :M6 ARG52:NH2- :M6	1.80 2.11 4.16 3.51
M7		-6.60	-50.39	Hydrogen bonds VAL47:O- :M7 SER51:O- :M7	1.91 2.06
M8		-6.52	-50.92	Hydrogen bonds THR29:O- :M8 SER51:O- :M8 Pi-sigma PHE50:HA- :M8	2.08 2.31 2.88

The binding interactions of docked ligands M1 to M8 with Rab39a show H-bonding, Pi–Pi, Pi-cation and Pi-sigma interactions

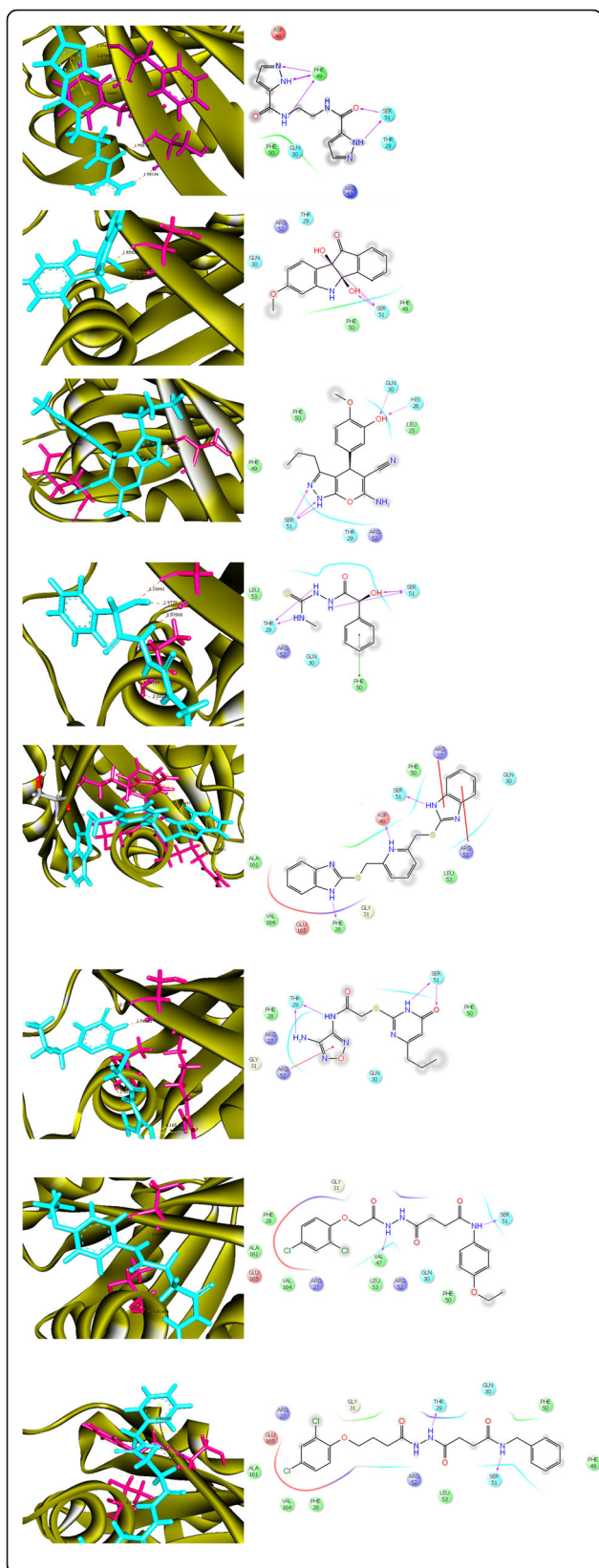
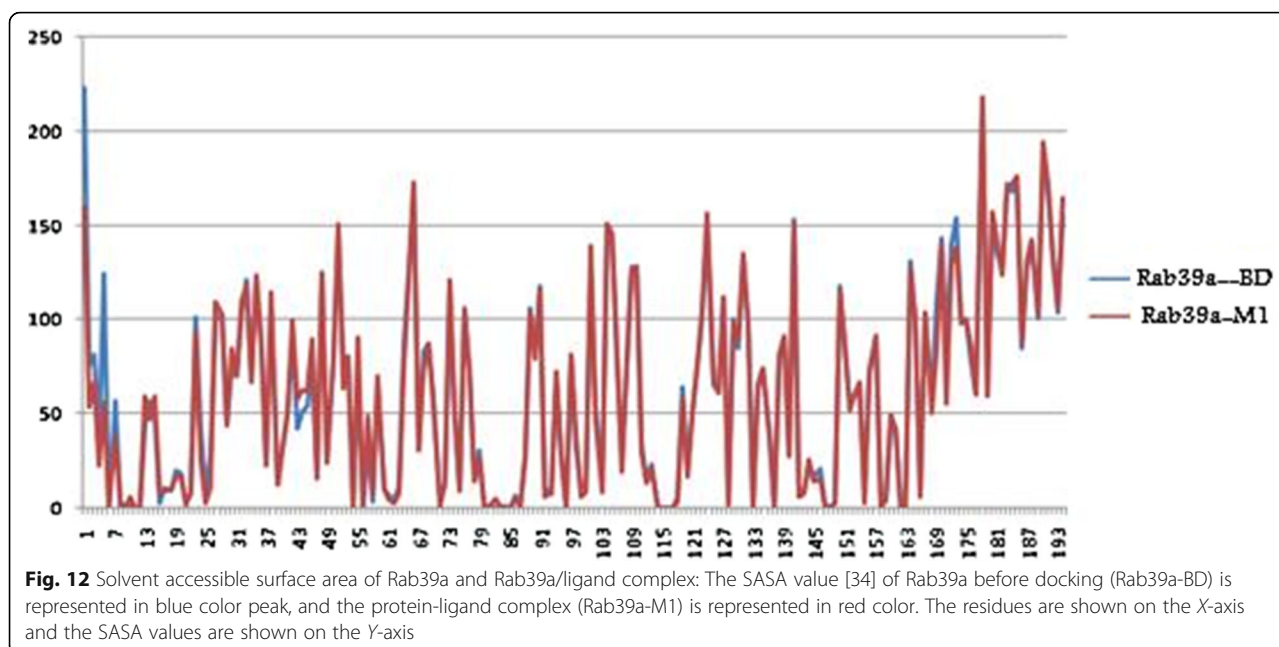


Fig. 11 Docked structures of top eight hits with Rab39a (i) three-dimensional [26] and (ii) two-dimensional depicting the interactions between Rab39a and ligand molecules: (i) Rab39a protein is represented in gold color, binding residues in pink stick, and the ligands in cyan stick model. The H-bonds are in brown, and π interactions in yellow color. (ii) The residues are shown in a three-letter code, H-bonds in pink lines, and π - π in green color

isomeric structures were obtained (Table 5). The generated structures were subjected to HTVS mode and 10% ligands (1767) obtained as output, followed by 10% of the structures, 176, screened in SP mode, and then the 10% of the resultant SP docking mode structures were screened in XP mode and finally 17 ligand-protein docked complex structures were obtained as an output of screening. These structures can be considered as new molecular entities and can be used against Rab39a.

Discussion

A total of 17 docked complexes obtained show a Glide score in the range of -7.31 to -6.52 . A sample of eight best docked ligand molecules, arranged rank wise based on their docking score and docking energy, is shown in Table 5. The interactions of Rab39a and ligands are shown in Fig. 11. The docked complexes were visualized using Discovery Studio Visualizer 3.5. The binding residues of Rab39a are shown in a pink-colored stick model, and the ligands are represented in a cyan-colored stick model. The hydrogen bonds are represented by brown lines, and all π interactions are shown in yellow lines (Fig. 11). The ligand molecules with hetero amines (indazole, pyrazole, oxadiazole, pyridine and pyrimidine) moieties, amide ($-\text{CONH}-$) and ($-\text{OH}$) groups in ligands M1–M8 are observed to be common pharmacophore groups which interact with the residues PHE28, THR29, GLN30, ARG32, GLY35, ARG37, ALA40, VAL47, ASP48, PHE49, PHE50, SER51, ARG52 and LYS63 of the target protein through hydrogen bonds interactions as represented in Table 5. The residues PHE49, PHE50 and ARG52 form π - π interactions with ligand M1. In addition, the residues ARG27, ARG52 and PHE50 form π -cation and π -sigma interaction with M5, M6 and M8, respectively, as illustrated in Table 5. The results showed that the compounds M5 and M6 form π -cation interactions with the target proteins through ARG27 and ARG52. This is due to arginine containing positively charged guanidinium group $\text{H}_2\text{N}-\text{C}(=\text{NH})-\text{NH}_2$ that is involved in forming π -cation interaction with hetero amine moieties in compounds. These non-covalent interactions add more stability to the protein-ligand complex. Figure 12 is an example for calculations of Rab39a (before docking) and Rab39a-M1 complex (after docking). The decrease in SASA values confirms that residues (PHE28 to LYS63) are involved in the formation of



the bonds with ligands. The screened ligands with permissible ADME properties are considered as novel drug-like molecules. The pharmacokinetics properties of eight ligand molecules are tabulated in Table 6. They have stars ranging from 0 to 5, which fall in the range of 95% of existing drug molecules. In addition, they obey the Lipinski rule of five and the Jorgenson rule of three, without showing any deviation, suggesting that they are with drug-like properties which can be considered as potent antagonists against Rab39a.

Conclusion

The experimental studies reveal that the overexpression of Rab signaling proteins promote the cellular functions that become overly active in cancerous cells. Rab39a is involved in regulation of cellular endocytosis pathway. Alteration in Rab39a and its associated regulatory proteins has been implicated in causing

lung carcinoma. Therefore, Rab39a is treated as a novel target for design of new molecular entities (NMEs), as lung cancer therapeutics by application of in silico technique. The homology model of Rab39a was evaluated by comparative modeling program. The energy of the generated 3D model was minimized and validated using Ramachandran plot and ProSA. The binding sites were identified using computational prediction tools, which show that PHE28 to LYS63 are important for binding. Computer-aided virtual screening was carried out to find novel ligands with better glide scores. The hetero amine moieties are important pharmacophore groups which may be useful for designing new anticancer leads against the target protein. The ligand molecules show admissible ADME properties, which may be used as cancer therapeutics. Our study helps in identification of potential antagonists against lung cancer.

Table 6 ADME properties of the top docked molecules

S. no.	Stars	CNS	M.Wt	DHB	AHB	QLog Po/w	QLog BB	Percent human oral absorption	Rule of five	Rule of three
M1	0	-2	389.46	1	9.75	1.95	-1.08	85.36	0	0
M2	2	-2	216.21	5	7.5	-2.18	-1.45	29.13	0	1
M3	5	-2	590.66	2	11	4.34	-2.15	76.72	1	1
M4	0	-2	444.48	2	11	2.37	-1.65	84.58	0	0
M5	1	-2	409.49	3	5.5	4.86	-1.01	100.00	0	1
M6	0	-2	428.42	3	13	0.44	-2.76	30.03	1	1
M7	0	-2	406.44	3	6.75	3.50	-1.35	93.35	0	1
M8	0	-2	348.41	3	7.2	2.50	-1.24	91.19	0	1

The permissible ranges are as follows: stars: (0-5); CNS: -2 (inactive) +2 (active); M.Wt.: (130-725); donor HB: (0.0-6.0); accept HB: (2.0-20.0); QLogPo/w: (-2.0 to 6.5); QLogBB: (-3.0 to -1.2); percent human oral absorption: > 80% high, < 25% low; rule of three (3); rule of five (4)

Abbreviations

3D: Three-dimensional; ADME: Adsorption, distribution, metabolic and excretion; BLASTp: Basic Alignment Tool program; CasTp: Computed Atlas of Surface Topography of protein; GAP: GTPase-activating proteins; GEF: Guanine nucleotide exchange factors; Jpred4: Java Prediction V 4.0; NMEs: New molecular entities; OPLS: Optimized potentials for liquid simulations; PPIs: Protein–protein interactions; ProSA: Protein structure analysis; SASA: Solvent accessible surface area; UniProt: Universe Protein Resource

Acknowledgements

I thank prof. V. Uma, Molecular Modeling Research Lab., UCS, Osmania University, India, for the help and support.

Author's contributions

I confirm that I carried out all the work and no other co-workers were involved. The author read and approved the final manuscript.

Authors' information

Dr. Aoubakr Haredi Abdelmonsef

- Assistant Professor (Lecturer) at Chemistry Department, Faculty of Science, South Valley University, Qena, Egypt.
- Ph.D. in Chemistry (Bioinformatics and drug design), Department of Chemistry, University College of Science, Osmania University, Hyderabad, Telangana, India.
- M.Sc., Organic Chemistry, Faculty of Science, South Valley University, Qena, Egypt.
- B. Sc., Chemistry, Faculty of Science, South Valley University, Qena, Egypt.

Funding

No fund

Availability of data and materials

- The amino acid sequence of Rab39a (ID: Q14964) is retrieved from UniProt database (<http://www.uniprot.org/uniprot/>).
- The identification of suitable homologous template structures for the modeling of the target protein is carried out using different server tools namely, BLASTp (Basic Alignment Tool program) <https://blast.ncbi.nlm.nih.gov/Blast.cgi>, Phyre2 (Protein Homology/analogy Recognition Engine V 2.0) <http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>, JPred4 (Java Prediction V 4.0) <http://www.compbio.dundee.ac.uk/jpred/index.html>, and Domain Fishing servers.
- The sequence alignment is performed by subjecting the amino acid sequences of the target and template proteins to ClustalW server tool <https://www.genome.jp/tools-bin/clustalw>.
- The models of the target protein are generated using a protein structure modeling program, Modeller 9.11.
- The 3D model was visualized using pymol 1.3 software <https://pymol.informer.com/1.3/>.
- The generated 3D model of Rab39a is refined by Ramachandran plot (RC) <http://servicesn.mbi.ucla.edu/SAVES/> and protein structure analysis (ProSA) <https://prosa.services.came.sbg.ac.at/prosa.php> servers.
- The binding cavities are identified using computational tools such as Computed Atlas of Surface Topography of protein (CASTp) <http://sts.bioe.uic.edu/castp/index.html?1bxw> and Sitemap module of Schrodinger <https://www.schrodinger.com/>.
- The Rab39a/DENND5B interactions are examined by in silico protein–protein docking studies using patchDock sever Beta V 1.3 <http://bioinfo3d.cs.tau.ac.il/PatchDock/php.php>.
- Accelrys Discovery Studio Visualizer 3.5 is used to visualize the Rab39a/DENND5B intermolecular interactions.
- The energy minimization is carried out using protein preparation wizard in Schrodinger suite, which is performed using an all-atom Impact Refinement (Impref) (Impact v 5.0, Schrodinger, NY) <https://www.schrodinger.com/>.
- The ligands prepared from Sigma Tim Tec database, were used in screening study.
- The ligands are subjected to Schrödinger's Ligprep module, to produce different conformers <https://www.schrodinger.com/>.

– The pharmacokinetic properties of the ligands are calculated by subjecting them to QikProp module (QikProp version 3.3) <https://www.schrodinger.com/>.

– The dataset used during the current study are available in the [Sigma Tim Tec database] repository, [<https://www.timtec.net/>].

– All data generated or analysed during this study are included in this published article [and its supplementary information files].

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Competing interests

The author declares that he has no competing interests.

Received: 20 May 2019 Accepted: 21 July 2019

Published online: 15 August 2019

References

1. Bray F, Moller B (2006) Predicting the future burden of cancer. *Nat Rev Cancer* 6:63–74
2. Ahmed I, Lobo DN (2009) Malignant tumors of the liver. *Surgery* 27(1):30–37
3. Harald S (2009) Rab GTPases as coordinators of vesicle traffic. *Nat Rev Mol Cell Biol* 10:513–525
4. Novick P, Brennwald P (1993) The diversity of Rab proteins in vesicle transport. *Cell* 75:597–601
5. Chin T, Han Y, Yang M, Zhang W, Li N, Wan T, Guo J, Cao X (2003) Rab39, a novel Golgi-associated Rab GTPase from human dendritic cells involved in cellular endocytosis. *Biochem Biophys Res Commun* 303(4):1111–1120
6. Subramani D, Alahari SK (2010) Integrin-mediated function of RabGTPases in cancer progression. *MolCancer* 9:312
7. Alex HH, Peter JN (2011) Role of RabGTPases in membrane traffic and cell physiology. *PhysiolRev* 91(1):119–149
8. Shin-ichiro Y, Andreas G, Andrea L, Daniel JR, Francis AB (2010) Family-wide characterization of the DENN domain Rab GDP-GTP exchange factors. *J Cell Biol* 191(2):367–381
9. Liang BU, Guanchao J, Fan Y, Jun L, Jun W (2011) Highly expressed SLC35F2 in non-small cell lung cancer is associated with pathological staging. *Mol Med Rep* 4:1289–1293
10. Francis AB (2013) RabGTPases and membrane identity: causal or inconsequential? *J Cell Biol* 202(2):191–199
11. Barr F, Lambright DG (2010) Rab GEFs and GAPs. *Curr Opin Cell Biol* 22:461–470
12. Altschul SF, Grish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410
13. Kelley LA, Sternberg MJE (2009) Protein structure prediction on the web: a case study using the Phyre server. *Nat Protoc* 4:363–371
14. Christian C, Jonathan DB, Geoffrey JB (2008) The Jpred4 secondary structure prediction server. *Nucleic Acids Res* 36:197–201
15. Contreras-Moreira B, Bates PA (2002) Domain fishing: a first step in protein comparative modeling. *Bioinformatics* 18:1141–1142
16. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) ClustalW and ClustalX version 2. *Bioinformatics* 23:2947–2948
17. Gonnet GH, Cohen MA, Benner SA (1992) Exhaustive matching of the entire protein sequence database. *Science* 256(5062):1443–1445
18. Fiser A, Sali A (2003) Modeller: generation and refinement of homology-based protein structure models. *Method Enzymol* 374:461–491
19. Rob WWF, Chris S, Gerrit V (1997) Objectively judging the quality of a protein structure from a Ramachandran plot. *Bioinformatics* 13(4):425–430
20. Markus W, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res* 35:407–410
21. Antonio DS, Hirotomo F, Dolores A, Ruth N (2006) Residue centrality, functionally important residues, and active site shape: analysis of enzyme and non-enzyme families. *Protein Sci* 15(9):2120–2128
22. Dundas J, Ouyang Z, Binkowski A, Tupraz Y, Liang J (2006) CASTp: computed atlas of surface topography of proteins with structural and

- topographical mapping of functionally annotated residues. *Nucleic Acids Res* 34:116–118
23. Halgren TA (2007) New method for fast and accurate binding-site identification and analysis. *Chem Biol Drug Des* 69:146–148
 24. Janin J, Wodak SJ (2002) Protein modules and protein-protein interaction. *Adv Protein Chemistry* 61:1–8
 25. 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
 26. Accelrys Discovery Studio Visualizer v 3.5. San Diego: Accelrys Software Inc, San Diego: 2010.
 27. Reddy AS, Pati SP, Kumar PP, Pradeep HN, Sastry GN (2007) Virtual screening in drug discovery - a computational perspective. *Curr Protein Pep Sci* 8(4): 329–351
 28. Dibrov A, Myal Y, Lyegue E (2009) Computational modeling of protein interactions: energy minimization for the refinement and scoring of association decoys. *ActaBiotheoretica* 57(4):419–428
 29. 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(45):11225–11236
 30. Schrödinger Suite (2010) Ligprep, version 2.5. Schrödinger, LLC, New York
 31. Shivakumar D, Williams J, Wu Y, Damm W, Shelley J, Sherman W (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
 32. Lengauer T, Rarey M (1996) Computational methods for biomolecular docking. *Curr Opin Stru Biol* 6(3):402–406
 33. Kitchen DB, Decornez H, Furr JR, Bajorath J (2004) Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat Rev Drug discov* 3(11):935–949
 34. Amir 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
 35. McInnes C (2007) Virtual screening strategies in drug discovery. *Curr Opin Chem Biol* 11(5):494–502
 36. Loakimidis L, Thoukydidis L, Mirza A, Naeem S, Reynisson J (2008) Benchmarking the reliability of QikProp, correlation between experimental and predicted values. *QSAR Comb Sci* 27(4):445–446
 37. Schrödinger Suite (2010) QikProp, version 3.3. Schrödinger, LLC, New York
 38. Gasteiger E, Gattiker A, Hooogland C (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* 31:3784–3788
 39. John JD, Zhongyuan Z, Joseph LC, David GL (1999) Structural basis of activation and GTP hydrolysis in Rab proteins. *Structure* 7(4):413–423
 40. Altsch SF, Madden TL, Schaffer A, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25(17):3389–3402
 41. The PyMOL molecular graphics system, Version 1.3 Schrödinger, LLC, NY, 2010.
 42. Aboubakr HA, Ramasree D, Thirupathi D, Lavanya SP, Thirupathi M, Uma V (2016) Comb Chem High Throughput Screen 19(10):875–892
 43. Bosshard HR, Mart DN, Jelesarov I (2004) Protein stabilization by salt bridges: concepts experimental approaches and clarification of some misunderstandings. *J Mol Recognit* 17:1–16
 44. Chavrier P, Gorvel JP, Stelzer E, Simons K, Gruenberg J, Zerial M (1991) Hypervariable C-terminal domain of Rab proteins acts as a targeting signal. *Nature* 353(6346):769–772
 45. Jose B, Pereira L, Miguel CS (2000) The mammalian Rab family of small GTPases: definition of family and subfamily sequence motifs suggests a mechanism for functional specificity in the Ras superfamily. *J Mol Biol* 301:1077–1087
 46. Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, Banks JL (2004) Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J Med Chem* 47:4750–4759

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)