

Research Article

Homology Modeling and Virtual Screening Studies of Antigen MLAA-42 Protein: Identification of Novel Drug Candidates against Leukemia—An *In Silico* Approach

Ihsan A. Shehadi,¹ Huda R. M. Rashdan,² and Aboubakr H. Abdelmonsef ³

¹Chemistry Department, Faculty of Science, University of Sharjah, Sharjah 27272, UAE

²Chemistry of Natural and Microbial Products Department, Pharmaceutical and Drug Industries Research Division, National Research Centre, Dokki 12622, Cairo, Egypt

³Chemistry Department, Faculty of Science, South Valley University, Qena 83523, Egypt

Correspondence should be addressed to Aboubakr H. Abdelmonsef; aboubakr.ahmed@sci.svu.edu.eg

Received 22 July 2019; Revised 2 December 2019; Accepted 13 January 2020; Published 15 March 2020

Academic Editor: Prem Chapagain

Copyright © 2020 Ihsan A. Shehadi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Monocytic leukemia-associated antigen-42 (MLAA-42) is associated with excessive cell division and progression of leukemia. Thus, human MLAA-42 is considered as a promising target for designing of new lead molecules for leukemia treatment. Herein, the 3D model of the target was generated by homology modeling technique. The model was then evaluated using various cheminformatics servers. Moreover, the virtual screening studies were performed to explore the possible binding patterns of ligand molecules to MLAA's active site pocket. Thirteen ligand molecules from the ChemBank™ database were identified as they showed good binding affinities, scaffold diversity, and preferential ADME properties which may act as potent drug candidates against leukemia. The study provides the way to identify novel therapeutics with optimal efficacy, targeting MLAA-42.

1. Introduction

Cancer is a result of a progressive accumulation of epigenetic changes and genetic aberrations, which lead to uncontrolled accumulation of blood cells [1]. Leukemia is an incurable disease, characterized by the unrestrained proliferation of abnormal white blood cells, readily spread to the entire body through blood stream and lymphatic system [2]. It starts in blood-forming organs and directs the formation of abnormal blood cells which can divide to produce copies of themselves [3]. Acute monocytic leukemia (AMoL) is an irremediable disease in older adults [4]. The present need is to identify novel and more effective therapies for AMoL to alleviate the suffering of patients.

MLAA-42 is one of immunogenic antigens and a novel-identified monocytic leukemia-associated antigen which is overexpressed in acute monocytic leukemia. In an attempt to uncover the mechanism of MLAA-42 overexpression, it significantly works as an elongation factor which plays an

important role in polypeptide elongation in leukemia as shown in the biochemical pathway (Figure 1) [5]. Furthermore, MLAA-42 plays a critical role in the transformation of normal cells to malignant phenotype through binding with GTP bound protein leading to unrestrained proliferation of the abnormal white blood cells which lead to AMoL. In the meantime, inhibition of proliferation of bone marrow cells, which takes place if there is no binding between MLAA-42 and GTP, is caused directly to prevent formation of AMoL [5].

Heterocyclic compounds have been found to gain extensive interest due to their considerable wide range applications in medicinal, pharmaceutical, and pharmacological activities especially as anticancer agents [6, 7]. In the present study, computer aided virtual screening studies were carried out using ChemBank™ database to report new molecular entities which can be used as potent inhibitors against MLAA-42 protein. This study warrants further investigations to indicate that MLAA-42 is a novel target for

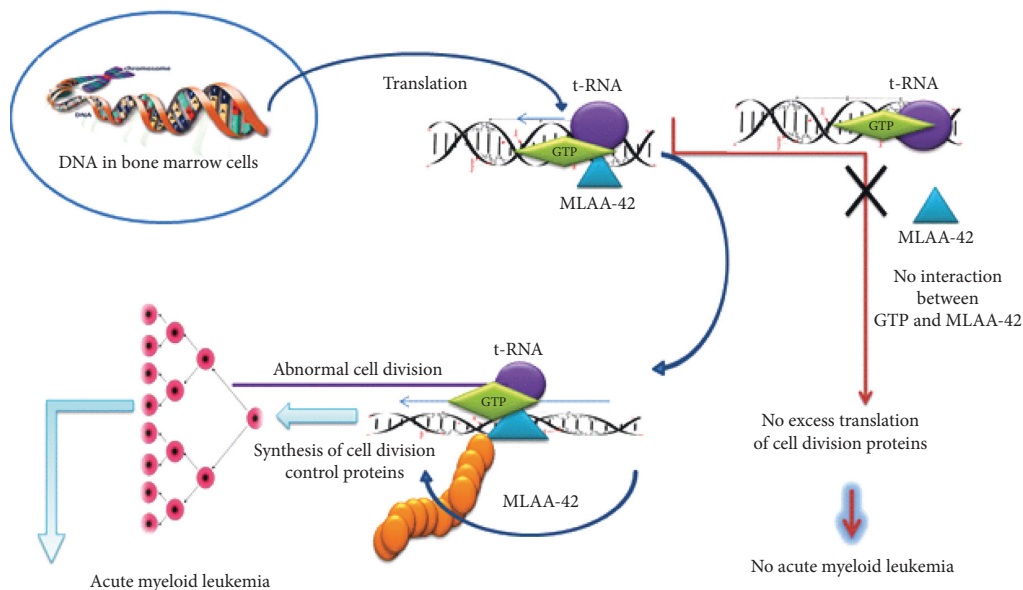


FIGURE 1: Biochemical pathway of MLAA-42 protein involves transformation of normal cells to malignant phenotype through binding with GTP bound protein leading to unrestrained proliferation of the abnormal white blood cells which lead to AMOL.

designing newer drug-like molecules for leukemia treatment, using in silico approaches.

2. Results and Discussion

2.1. Generation of the Three-Dimensional Structure of MLAA-42. For the case where the target MLAA-42 is not experimentally determined by means of X-ray crystallography and nuclear magnetic resonance (NMR), the computational-based techniques are used for 3D structure prediction [8]. The FASTA sequence of the target protein (ID: Q6W6M8) with 151 amino acid residues was identified from Universal Protein Resource “UniProtKB” database (<https://www.uniprot.org/uniprot/Q6W6M8>) [9]. The template was recognized by subjecting the obtained sequence to BLAST, Jpred3, and Domain Fishing servers, and Table 1 represents the resulted maximum E -value. The 1N0V “Crystal structure of elongation factor 2” was selected as a template protein, on the basis of parameters such as query cover (77%), maximum identity (63%), and statistical E -value (8×10^{-47}) with the target [10]. ClustalW server was used to perform the alignment of both elongation factor proteins MLAA-42 and 1N0V (541-720 aa), as shown in Figure 2. The conserved amino acid residues are elucidated as (*) (39.3%), highly similar residues as (:), and weakly similar residues as (.) (51.7%). The individual percentages, resulted from the server, refer to the identity and similarity of the target residue sequence with its template sequence, respectively.

Twenty-five models of MLAA-42 were generated by using Modeller 9.11 software, depending on satisfying spatial restraints in terms of probability density function [11]. The 3D model with fewest restraint violations, lowest probability density function (855.33), and acceptable geometry was selected for further validation [12]. Figure 3 represents the 3D model of the target, which consists of five alpha-helices and six beta-strands. The secondary structures of both 1N0V

TABLE 1: The template protein selection.

Server name	Parameters for template selection	E -values	PDB code
NCBI-Blast	Sequence similarity	$9e^{-47}$	1N0V
Jpred3	Secondary structure prediction	$6e^{-40}$	1N0V
Domain fishing	Domain similarity	$3e^{-40}$	1N0V

1N0V was selected as a template protein using different servers, based on query cover, maximum identity, and statistical E -value.

and MLAA-42, which represent the amino acid sequences of helices and strands in both proteins, are represented in Figures 4 and 5 and (Figure S1 and Tables S1 and S2) in the supplementary data section, using PDBsum server [13]. To obtain a stable confirmation, energy minimization and protein loop modeling were performed by GROMOS96 Force-Field set, using Swiss-Pdb Viewer [14]. To get insight of the superimposition between the proteins MLAA-42 and 1N0V, the root-mean-square deviation RMSD for the backbone of MLAA-42 and 1N0V is 0.51 \AA (within the permissible range i.e. $\leq 2 \text{ \AA}$), which indicates close homology and ensures reliability of the model, as declared in Figure 6.

2.2. Physicochemical Properties and the Abundance of Amino Acids in MLAA-42. The physicochemical properties of MLAA-42 were predicted using ProtParam tool [15]. The protein sequence has 151 amino acid residues with molecular weight of 16.871 kDa. Furthermore, the most abundant amino acid residues are LEU, SER, GLU, ALA, VAL, LYS, GLY, and ASP, respectively, in high percentages in the target protein, as shown in Figure 7. Lucine has the highest abundance (9.9%), and methionine has the lowest abundance (0.7%). The physicochemical parameters predicted a

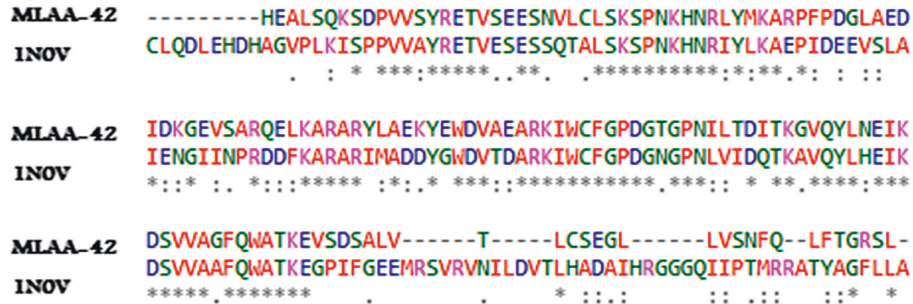


FIGURE 2: Pairwise sequence alignment of MLAA-42 with 1NOV. The conserved amino acid residues are elucidated as (*), highly similar residues as (:), and weakly similar residues as (.)

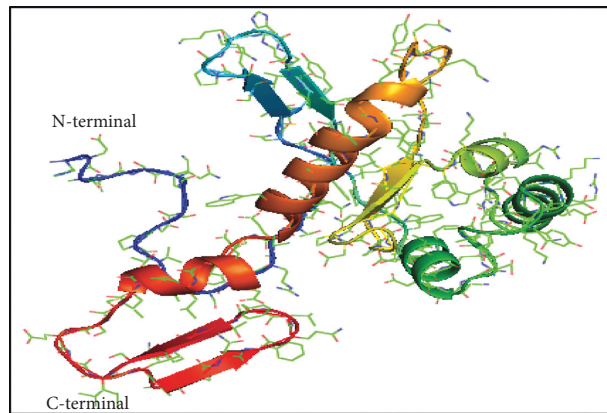


FIGURE 3: 3D structure of MLAA-42 protein. It consists of five α -helices and six β -strands.

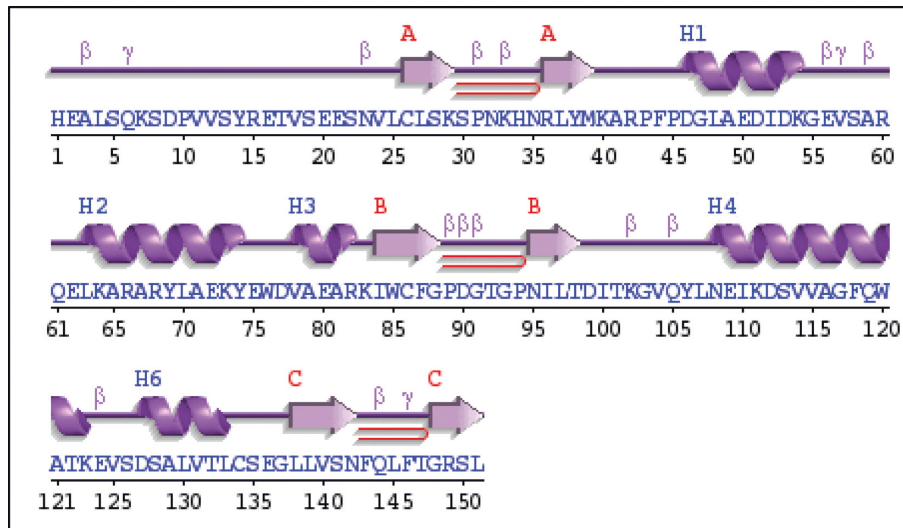


FIGURE 4: The secondary structure of the target MLAA-42 obtained by the PDBsum server.

negatively charged protein as the result of the high number of negatively charged residues (aspartic acid 6.0% and glutamic acid 8.6%) in contrast with positively charged ones (arginine 5.3% and lysine 7.3%). The atomic composition of MLAA-42 is 2364, with 747 carbon (C), 1176 hydrogen (H), 204 nitrogen (N), 233 oxygen (O), and 4 sulfur (S). In addition, the protein is acidic, with an isoelectric point (pI)

of 5.47. The estimated half-life of the target showed that it can remain intact without being degraded for 3.5 h in humans, 10 min in yeast, and less than 10 h in *E. coli*, and its extinction coefficient is $24075 \text{ M}^{-1} \text{ cm}^{-1}$. Finally, the generated aliphatic index was 82.65, with grand average of hydrophobicity (GRAVY) of -0.426 and an instability index of 34.55.

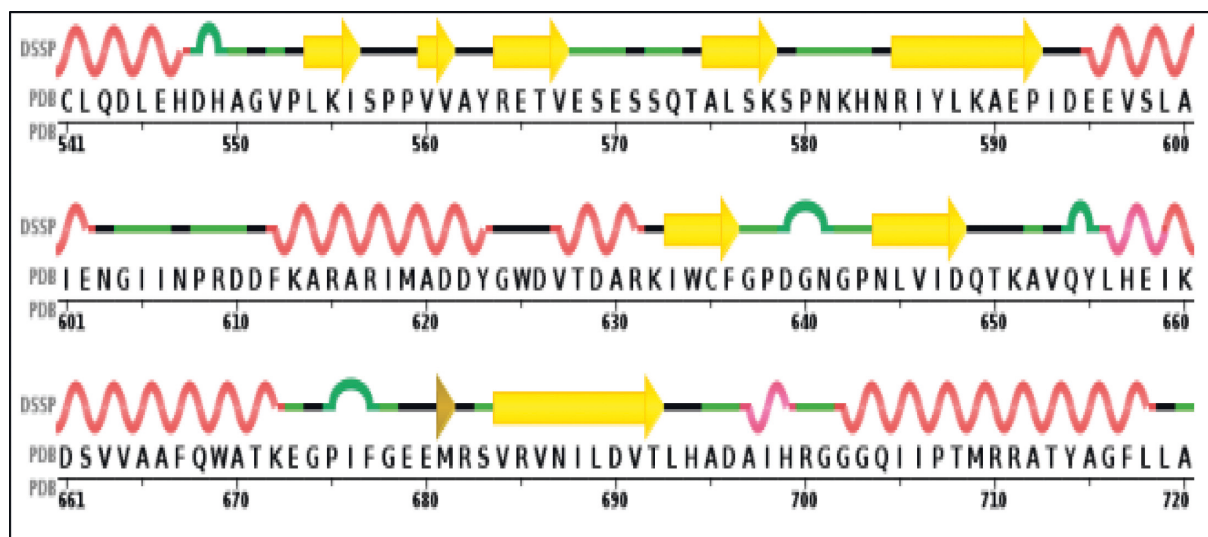


FIGURE 5: The secondary structure of the template 1N0V (541-720 aa) using the PDBsum server.

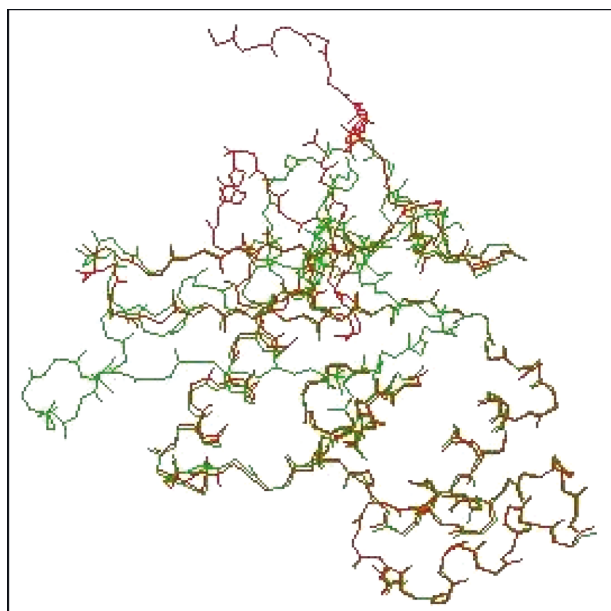


FIGURE 6: Superimposition of MLLA-42 (red) on template 1N0V (green) in the Swiss-PDB viewer.

2.3. Validation of the Model and Active Site Analysis. The 3D model was validated by different tools such as PROCHECK (Ramachandran plot) and ProSA to check its structural integrity. Ramachandran plot (Figure 8) of the modeled protein represents 87.3% (117 aa) of the total residues in the most favored regions and 11.9% (16 aa) in additionally allowed regions, indicating a good quality model.

Moreover, ProSA Z-score (dark spot) is -2.85 , which falls within the values range of the known proteins determined by X-ray (light blue) and NMR (dark blue) (see Figure S2, in the supplementary data section), predicting a reasonable quality model. Figure S3, in the supplementary data section, shows the 3D model quality by plotting energies (X-axis) as a function against residue sequence

position (Y-axis). The figure shows that most of residues of MLLA-42 have negative energies, which indicates an acceptable model. In view of the facts mentioned above about the unavailability of 3D structure of MLLA-42 in protein data bank PDB, thus, there are no previous data obtained for the binding regions of the target. Therefore, to overcome the lack of experimental data, herein, CASTp and SiteMap cheminformatics tools were used to identify the binding site cavities of the target (Figures S4 and S5 and Tables S3 and S4), in the supplementary data section, respectively. CASTp server gives information about hydrophobic pocket regions; in addition, it measures cavity volume and area. The results declared four sites, site 1 has volume of 40.9 \AA^3 , site 2 with volume of 28.2 \AA^3 , site 3 with volume of 50.7 \AA^3 , and site 4 has volume of 105.4 \AA^3 . Among the four different cavities obtained from CASTp, the amino acids from PHE44 to ILE96 are consistently present in all sites. SiteMap gives information about the hydrophilic and hydrophobic regions present on the protein surface. It identified two binding cavities on MLLA-42 protein. The site 1 with volume 1222.79 \AA^3 and size of 4.38 \AA and site 2 has volume of 46.30 \AA^3 and size of 32 \AA . The active sites predicted by SiteMap are from TYR38 to ILE100. In conclusion, these computational prediction techniques give information about the binding site regions that are composed of TYR38 to ILE100 residues are responsible for MLLA-42 docking with ligand molecules. The docking interactions show that the residues TYR38, LYS40, THR89, ILE96, ASP99, and ILE100 are observed to be crucial to identify the antagonists against MLLA-42 protein. Figure S6, in the supplementary data section, represents the grid with dimensions $80 \text{ \AA} \times 80 \text{ \AA} \times 80 \text{ \AA}$ around the binding site, to perform further studies.

2.4. Molecular Docking Analysis and ADMET Profile. The compounds with heterocyclic core structures are pharmaceutically important class of compounds because of their diverse range of biological importance such as anticancer

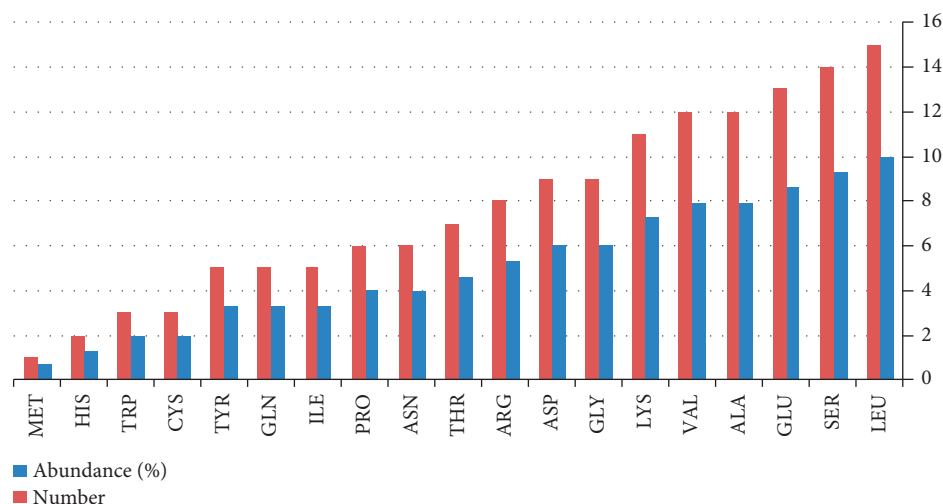


FIGURE 7: The abundance of amino acids in MAAA-42. Graphical representation of the abundance of 20 amino acids present in the target. Leucine has the highest abundance (9.9%), and methionine has the lowest abundance (0.7%).

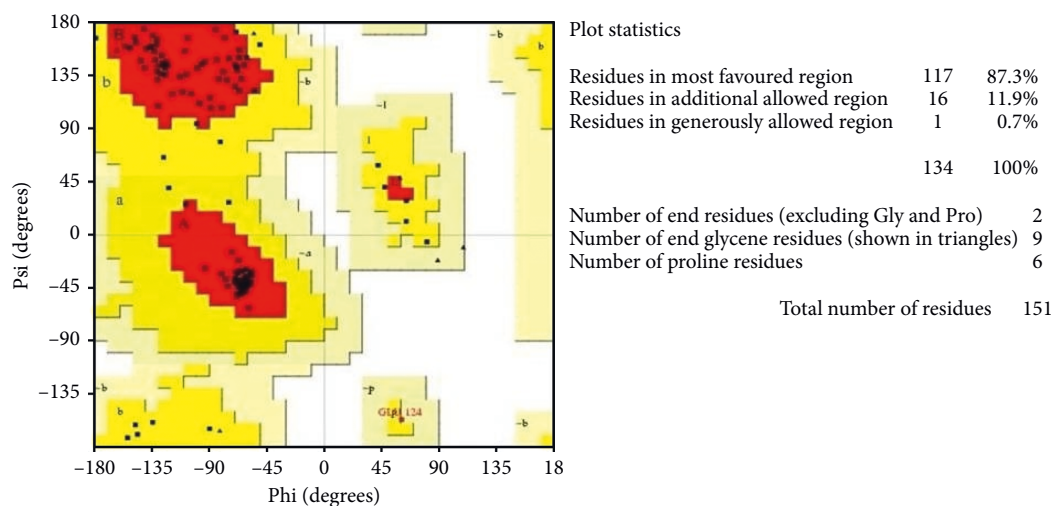


FIGURE 8: Stereochemical analysis of MAAA-42. The red region declares the most favorable area of residues; the yellow region is additionally allowed; and generously allowed residues in the light-yellow region. The RC plot declares 99.2% of residues falling in the allowed region.

activity against cancer cell lines [6, 7]. In view of the facts mentioned above and as a part of our efforts to identify new anticancer drug candidates [16], a ChemBank™ library of 2344 small molecules with an anticancer activity was downloaded [17] for specific docking. Computer-aided screening protocol at the binding regions of the target was performed using virtual screening workflow docking program of Glide, Schrodinger suite. The glide docking program identifies specific structural motifs and provides exceptionally large contribution to enhance binding affinity [18]. The docking approach was run in a flexible docking mode which automatically generates different confirmations for each ligand molecule. The total isomeric structures (13,354) obtained after ligand preparation was subjected to high throughput virtual screening (HTVS) mode and 10% hit structures (1335) obtained as output, followed by 10% of them, 133, screened in standard precision (SP) mode, and

finally the 10% of the resultant were screened in the Extra precision (XP) mode to obtain 13 hits. The docked compounds were prioritized according to their binding energies (in range of -12.19 to -9.22 Kcal/mol) after docking to the active site pocket of the target protein, as tabulated in Table 2.

A sample of six best docking interactions between the ligands and protein, arranged rankwise based on their binding energies, is depicted in Figure 9 and Table 2. The top six ligand-protein complexes were visually checked in Discovery Studio Visualizer (for 3D representation) and Maestro (for 2D representation), which exhibited the noncovalent interactions such as hydrogen bonding and pi-pi stacking [19, 20]. In 3D representation, the binding residues of MAAA-42 are shown in yellow-colored stick model, the ligands in purple one, H-bonds are in black dotted lines, and pi-pi interactions are in orange lines

TABLE 2: The chemical structures of top six scoring ligand molecules against MLLA-42 protein.

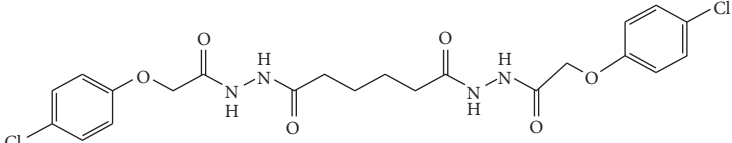
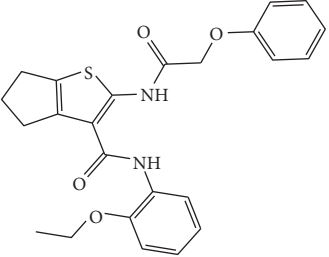
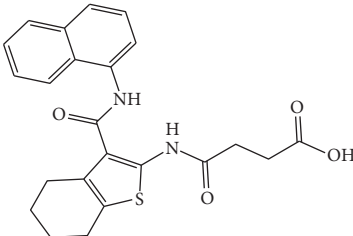
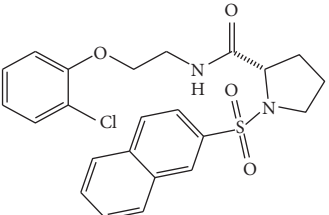
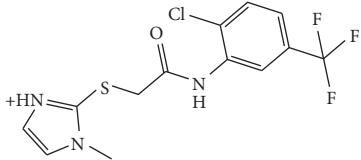
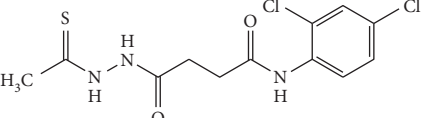
S. no.	Structure and IUPAC name of molecule	Glide energy (KJ/Mol)	Binding energy (Kcal/mol)	The interacting amino acid residues	Bond distance (Å)
1	 <p>6-[N'-(2-(4-Chloro-phenoxy)-acetyl)-hydrazino]-6-oxo-hexanoic acid N'-(2-(4-chloro-phenoxy)-acetyl)-hydrazide</p>	-63.59	-12.19	H-bonds LYS40: N - L1 THR98: N - L1 ILE100: N - L1 LYS40: O - L1 ILE96: O - L1 THR98: O - L1	2.91 2.50 2.67 2.32 2.18 2.15
2	 <p>2-[2-(1-Methyl-buta-1,3-dienyloxy)-acetylamino]-5,6-dihydro-4H-cyclopenta[b]thiophene-3-carboxylic acid (2-ethoxy-phenyl)-amide</p>	-50.27	-11.83	H-bonds TYR38: N - L2 ILE100: N - L2 TYR38: O - L2 π - π interactions TRP85-L2	2.89 2.99 1.94 4.95
3	 <p>N-[3-(Naphthalen-1-ylcarbamoyl)-4,5,6,7-tetrahydro-benzo[b]thiophen-2-yl]-succinamic acid</p>	-42.22	-11.61	H-bonds ILE100: H - L3 THR98: O - L3 TYR38: O - L3	1.68 2.36 2.00
4	 <p>1-(Naphthalene-2-sulfonyl)-pyrrolidine-2-carboxylic acid [2-(2-chloro-phenoxy)-ethyl]-amid</p>	-47.52	-11.30	H-bonds LYS40: N - L4 THR98: N - L4 THR98: O - L4	2.95 2.65 2.26
5	 <p>2-[(2-Chloro-5-trifluoromethyl-phenylcarbamoyl)-methylsulfanyl]-3-methyl-3H-imidazol-1-ium</p>	-39.53	-11.12	H-bonds LYS40: N - L5 THR98: O - L5 ASP99: OD1-L5 π - π interactions TRP85-L5	2.67 2.13 2.79 4.87
6	 <p>N-(2,4-Dichloro-phenyl)-3-(N'-thioacetylhydrazine-carbonyl)-propionamide</p>	-40.65	-11.10	H-bonds ILE100: N - L6 TYR38: O - L6 THR98: OG1-L6	2.69 2.30 2.29

Table represents top six ligand molecules with binding energies in range of -12.19 to -9.22 Kcal/mol, hydrogen bonds, and pi-interactions to MLLA-42's active site pocket.

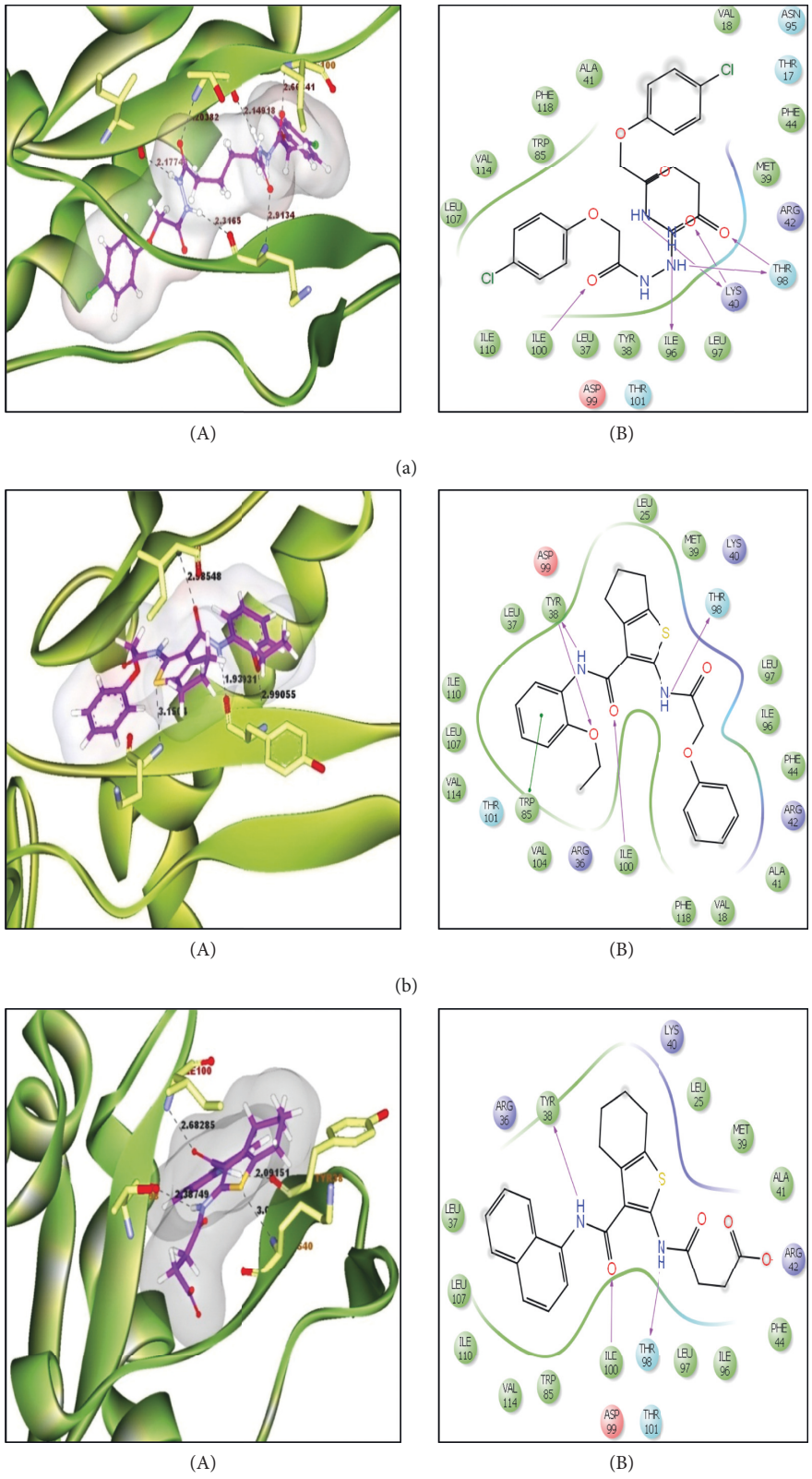


FIGURE 9: Continued.

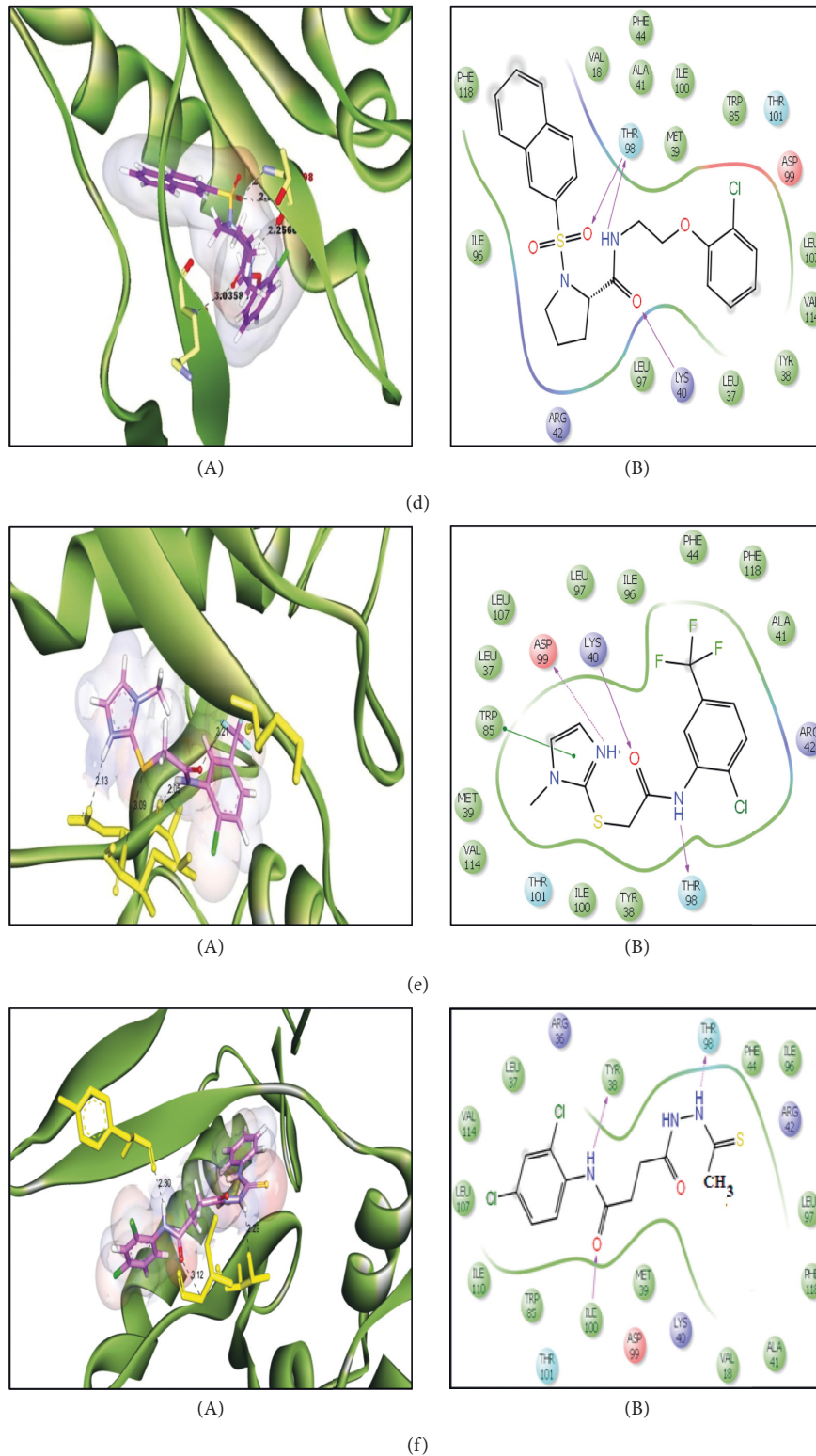


FIGURE 9: Interactions between MAA-42 with ligand molecules (a)–(f). (A) Three-dimensional docked poses: the binding residues of MAA-42 are represented in yellow stick, the ligand molecules are shown in the purple stick model, the hydrogen bonds are shown in black dotted lines, and pi-pi interactions in orange lines. (B) Two-dimensional docked poses: the residues are shown in a 3-letter code, hydrogen bonds in pink lines, and pi-pi interactions in green lines.

(Figure 9). While in 2D representation, the residues are shown in a 3-letter code, the hydrogen bonds are represented in pink lines, and pi-interactions are shown in green lines. The ligand molecules with heterocyclic moieties such as thiophene, pyrazole, and pyrrolidine and amide groups such as carboxamide (-CONH-) and sulphonamide (-SO₂NH-) in ligands L1-L6 are observed to be common pharmacophore groups which interact with the active site region (TYR38, LYS40, TRP 85, ILE96, THR98, ASP99, and ILE100) of MLAA-42 through a network of hydrogen bonds and pi-interactions, as represented in Table 2. The carboxamide group, which is present in all ligand molecules, formed intermolecular hydrogen bonds with the residues TYR38, LYS40, TRP 85, ILE96, THR98, ASP99, and ILE100. In addition, the sulphonamide group, which is present in ligand L4, formed a hydrogen bond interaction (N---H---O) with the residue THR98 at the distance of 2.65 Å. Furthermore, the tryptophan (TRP85) formed π - π interactions (face to face) with the benzene ring of ligand L2 on its side chain plus the pyrazole ring of ligand L5, at the distance of 4.95 and 4.87 Å, respectively. The intermolecular hydrogen bonds and pi-interactions add more stability to the docked complexes [21]. The other docked compounds to the target are included in supplementary data section as Figure S7.

SASA values of MLAA-42, before and after docking, were analysed using Discovery Studio Visualizer (see Figure S8, in the supplementary data section). The decrease in SASA values confirms that the amino acid residues (TYR38 to ILE100) are involved in bonds formation with ligand molecules (a)–(f).

The ADMET properties for the resulted ligand molecules were determined in silico by using the QikProp module of the Schrödinger suite. The molecules with agreeable ADME properties are considered as new novel drug candidates, as shown in Table 3. Interestingly, all the compounds have a molecular weight in the acceptable range of 349.7 to 511.3 Daltons (less than 725 Dalton). BBB⁺ value describes the ability of the compounds to cross the blood-brain barrier, which is in the permissible ranges for all compounds. They had <10 hydrogen bond acceptors and <5 hydrogen bond donors, and log *P* values of <5. These properties are in the reasonable range of Lipinski's rule of five (LORF) [22]. The human oral absorption, partition coefficient (QPlogPo/w), and water solubility (QPlogS) values are in the agreeable range defined for human use, which indicates their possible ability to be drug candidates, as shown in Table 3. Also, the values show that the compounds can be absorbed by the human intestines. Furthermore, all molecules (a)–(f) displayed negative AMES toxicity test which means that the ligands are non-mutagenic. Also, they displayed negative carcinogenicity test. These ligands can be considered as highly potent inhibitors against leukemia.

3. Materials and Methods

3.1. 3D Homology Model of MLAA-42. The protein's three-dimensional structure is required to understand its function [23]. Herein, an in silico homology modeling generates a three-dimensional (3D) for an unknown structure of protein

(target) depending on one or more proteins of known structures (templates) as reported earlier by Aboubakr et al. in 2016 [24, 25]. The generation of (3D) structure of MLAA-42 was carried out using comparative modeling techniques. The amino acid sequence of MLAA-42 is retrieved from the UniProtKB database in FASTA format. The resulted sequence is submitted to BLASTp, Jpred3, and Domain fishing servers, to identify suitable templates [26–28]. E-value is identified the similarity between the target and template proteins [29]. The template with maximum identity is selected for generating the 3D structure of the target. The pairwise sequence alignment of MLAA-42 with its template is performed using the ClustalW tool to recognize the structurally conserved regions [30]. ClustalW server uses the Gonnet matrix algorithm to make certain the presence of conserved motifs and to minimize the atomic gaps. The % identity of alignment is measured as a ratio number of the identical residues in the alignment and the total number of residues of the target. The % similarity is measured as a ratio of the total number of identical and similar residues to the total number of residues of the target. The 3D structure of the target is generated using Modeler 9.11. Then, the model is energy minimized using SPDBV and Impref in Schrödinger to obtain a stable conformation [31]. The RMSD value, which is calculated using the Swiss-Pdb viewer, is used to find the best superimposition for protein structures. The lower value (≤ 2 Å) means the best alignment between the structures.

3.2. 3D Structure Validation. Subsequently, quality of the generated model is evaluated by computational protocols such as PROCHECK and Protein Structural Analysis (ProSA) [32, 33]. PROCHECK, a program that relies on Ramachandran plots for structure verification, figures out the stereochemical quality of the model. Furthermore, ProSA is applied to check for energy criteria in comparison with a large set of known protein structures with similar size [34].

3.3. Binding Site Prediction and Ligand Preparation. It is the hydrophobic cavities on the surface of a protein, which is responsible for its specificity [35]. It is theoretically determined by means of tools based on recent theoretical and algorithmic results of computational geometry such as CAST-p and SiteMap of Schrödinger suite [36, 37]. These computational prediction tools analytically furnished the area and volume of cavities. The grid is created around the binding cavity to perform further structure-based virtual screening studies [38]. *LigPrep* module of Schrödinger suite is used to generate various conformers of small molecule depending upon its structural features [39].

3.4. In silico Molecular Docking, SASA, and ADME Analysis. Molecular docking protocol is used to predict the preferred orientation of a ligand molecule to the target protein to form a stable complex [40, 41]. The docking accuracy is determined by finding how closely the binding confirmation with

TABLE 3: ADME properties of the top docked ligand molecules.

Molecule	CNS	M. (Wt)	SASA	Volume	Donor HB	Accept HB	QPlogPo (w)	QPPCaco	QPlogBB	Percent human oral absorption	#NandO	Rule of five	Rule of three
L1	-2	511.3	905.7	1554.4	1	8	4.7	64.5	-2.8	74.0	10	1	1
L2	0	436.5	751.0	1348.2	1	5	5.5	2264.9	-0.5	100	6	1	2
L3	-2	422.4	740.5	1298.5	2	6	4.4	48.1	-1.7	83.1	6	0	1
L4	-1	458.9	755.6	1350.7	1	8	3.9	883.3	-0.6	100	6	0	0
L5	1	349.7	564.0	962.4	1	4	4.1	2206.4	-0.2	100	4	0	0
L6	-1	411.3	708.3	1210.2	3	7	3.9	538.7	-0.8	100	6	0	1

The agreeable ranges are as follows: CNS: -2 (inactive), +2 (active). Mol wt: (130–725). Volume: (500–2000). Donor HB: (0.0–6.0). Accept HB: (2.0–20.0). QPlogPo/w: (-2.0 to 6.5). QPPCaco: <25 poor, >500 great. QPlogBB: (-3.0 to -1.2). % human oral absorption: >80% high, <25% low. Rule of three (3). Rule of five (4).

the lowest energy of the cocrystallized ligand molecule predicted by the object scoring function; G-score (Glide score) resembles an experimental binding modes determined by X-ray crystallographic technique [42]. Once the 3D model of MLAA-42 is validated and the target binding pockets are defined, the docking-based virtual screening study is performed using the Glide docking tool incorporated in the Schrödinger package by Maestro [43]. The virtual screening approach is performed through hierarchical flexible docking methods: High Throughput Virtual Screening (HTVS), Standard Precision (SP), and Extra Precision (XP). The ligand molecules are prioritized on the basis of the docking score, docking energy in each step, and by default 10% of the molecules selected and considered for the next hierarchical step [18]. Solvent Accessible Surface Area (SASA) of the target, before and after docking, is computed using Accelrys Discovery Studio Visualizer 3.5 software. Prediction of drug-like profiles, such as physico-chemical, pharmacokinetic, and safety of the compounds, is performed using the QikProp module [44].

4. Conclusion

In this work, the computer-aided drug design protocols were used to identify novel leads against MLAA-42 protein. The homology model of the target was evaluated by homology modeling techniques. In silico molecular docking were also adopted to identify the lead compounds. The resulted molecules with heteroscaffolds and amide groups (-CONH- and -SO₂NH-) exhibited better estimated binding energy values and agreeable pharmacokinetic properties and were ranked as potent drug-like candidates against the target protein. Hence, MLAA-42 has emerged as a therapeutic target for treatment of leukemia carcinoma.

Abbreviations

MLAA-42:	Monocytic leukemia-associated antigen-42
UniProt	Universe protein resource crystal structure of
1NOV:	elongation factor 2
3D:	Three-dimensional
GEF:	Guanine nucleotide exchange factors
GAP:	GTPase-activating proteins
BLASTp:	Basic alignment tool program
Jpred3:	Java prediction V 3.0

Impef:	Impact refinement module
ProSA:	Protein structure analysis
Castp:	Computed atlas of surface topography of protein
RMSD:	Root-mean-square deviation
SPDV:	Swiss PDB viewer
OPLS:	Optimized potentials for liquid simulations
PPIs:	Protein-protein interactions
SASA:	Solvent accessible surface area
ADME:	Adsorption, distribution, metabolic, and excretion.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare no conflicts of interest.

Authors' Contributions

I. S, A.H.A, and H. R. M. R designed the study. I. S. carried out the homology modeling of the target and evaluated the 3D model. A. H. A performed the virtual screening approach and wrote the preliminary draft of the manuscript. H. R. M. R was responsible for identification of ADME profiles of the molecules and wrote the corresponding section. All authors contributed to the revision of the drafts and agreed on the final version to be submitted.

Acknowledgments

The authors thank Prof. V. Uma, Molecular Modeling Research Lab., UCS, Osmania University, India, for the help and support.

Supplementary Materials

Figure S1: the secondary structure of MLAA-42. The 3D model of the target protein consists of five helices as red cylinders and six strands as pink arrows. Figure S2: The 3D model quality of MLAA-42. The score (-2.85) is falling in the range of PDB proteins identified by NMR (dark-blue region) and X-ray crystallography (light-blue region). Figure S3: ProSA of MLAA-42 protein. The plot shows local model

quality by plotting energies as a function against amino acid sequence position. Figure S4: binding site of MLAA-42 from the CAST-p server. Figure S5: binding site of MLAA-42 by SiteMap module. Figure S6: receptor grid generation by Schrodinger Suite. The grid generated with $80 \text{ \AA} \times 80 \text{ \AA} \times 80 \text{ \AA}$ dimensions using receptor grid generation of Glide module. Figure S7: two-dimensional interactions between MLAA-42 with other ligand molecules. Figure S8: the surface accessibility of MLAA-42 and docked complex (L1). Red-colored peaks represent SASA value of MLAA-42 (after docking), blue-colored peaks represent before docking. X-axis represents the amino acid residues and Y-axis represents SASA value. Tables S1 and S2: the secondary structure details of MLAA-42 protein (S1) the α -helices in MLAA-42 protein. (S2) the β -strands in MLAA-42 protein. Table S3: binding site of the MLAA-42 by CAST-p server. Table S4: binding site of MLAA-42 by SiteMap module. (*Supplementary Materials*)

References

- [1] F. Bray and B. Møller, "Predicting the future burden of cancer," *Nature Reviews Cancer*, vol. 6, no. 1, pp. 63–74, 2006.
- [2] I. S. Liran, S. Zandi, A. Mitchel et al., "Identification of pre-leukemic hematopoietic stem cells in acute leukemia," *Nature*, vol. 506, no. 7488, pp. 328–333, 2014.
- [3] J. W. Vardiman, J. Thiele, D. A. Arber et al., "The 2008 revision of the World Health Organization (WHO) classification of myeloid neoplasms and acute leukemia: rationale and important changes," *Blood*, vol. 114, no. 5, pp. 937–951, 2009.
- [4] H. D. Klepin and L. Balducci, "Acute myelogenous leukemia in older adults," *The Oncologist*, vol. 14, no. 3, pp. 222–232, 2009.
- [5] G. Chen, W. Zhang, X. Cao, F. Li, X. Liu, and L. Yao, "Serological identification of immunogenic antigens in acute monocytic leukemia," *Leukemia Research*, vol. 29, no. 5, pp. 503–509, 2005.
- [6] K. M. Amin, M. M. Kamel, M. M. Anwar, M. Khedr, and Y. M. Syam, "Synthesis, biological evaluation and molecular docking of novel series of spiro [(2H,3H) quinazoline-2,1'-cyclohexan]-4(1H)-one derivatives as anti-inflammatory and analgesic agents," *European Journal of Medicinal Chemistry*, vol. 45, no. 6, pp. 2117–2131, 2010.
- [7] M. Kamel, I. Nabih, K. Z. Gadalla, and M. M. Ashour, "New pyrazolinylanilinoquinolines and other related products of possible antibacterial activity," *Pharmazie*, vol. 41, no. 1, pp. 55–56, 1986.
- [8] J. Augen, "The evolving role of information technology in the drug discovery process," *Drug Discovery Today*, vol. 7, no. 5, pp. 315–323, 2002.
- [9] B. Boeckmann, A. Bairoch, R. Apweiler et al., "The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003," *Nucleic Acids Research*, vol. 31, no. 1, pp. 365–370, 2003.
- [10] C. S. Oehmen and D. J. Baxter, "ScalaBLAST 2.0: rapid and robust BLAST calculations on multiprocessor systems," *Bioinformatics*, vol. 29, no. 6, pp. 797–798, 2013.
- [11] A. Šali, L. Potterton, F. Yuan, H. V. Vlijmen, and M. Karplus, "Evaluation of comparative protein modeling by MODELLER," *Proteins: Structure, Function, and Genetics*, vol. 23, no. 3, pp. 318–326, 1995.
- [12] M. A. Martí-Renom, A. C. Stuart, A. Fiser, R. Sánchez, F. Melo, and A. Šali, "Comparative protein structure modeling of genes and genomes," *Annual Review of Biophysics and Biomolecular Structure*, vol. 29, no. 1, pp. 291–325, 2000.
- [13] R. A. Laskowski, "PDBsum: summaries and analyses of PDB structures," *Nucleic Acids Research*, vol. 29, no. 1, pp. 221–222, 2001.
- [14] W. R. P. Scott, P. H. Hünenberger, I. G. Tironi et al., "The GROMOS biomolecular simulation program package," *The Journal of Physical Chemistry A*, vol. 103, no. 19, pp. 3596–3607, 1999.
- [15] E. Gasteiger, C. Hoogland, A. Gattiker et al., "Protein identification and analysis tools on the ExPASy server," in *The Proteomics Protocols Handbook*, J. M. Walker, Ed., Humana Press, Totowa, NJ, USA, pp. 571–607, 2005.
- [16] M. E. Awatef and H. A. Aboubakr, "Synthesis, characterization and in silico molecular docking studies of novel chromene derivatives as Rab23 inhibitors," *Egyptian Journal of Chemistry*, vol. 20, no. 9, 2019.
- [17] K. P. Seiler, G. A. George, M. P. Happ et al., "ChemBank: a small-molecule screening and cheminformatics resource database," *Nucleic Acids Research*, vol. 36, pp. D351–D359, 2007.
- [18] R. A. Frienser, J. Y. Banks, R. B. Murphy et al., "Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening," *Journal of Medicinal Chemistry*, vol. 47, no. 7, pp. 1739–1749, 2004.
- [19] Accelrys Software Inc., *Accelrys Discovery Studio Visualizer V 3.5*, Accelrys Software Inc., San Diego, CA, USA, 2010.
- [20] Schrödinger, *Maestro Version 9.1*, Schrodinger, LLC, New York, NY, USA, 2010.
- [21] H. R. Bosshard, D. N. Marti, and I. Jelesarov, "Protein stabilization by salt bridges: concepts, experimental approaches and clarification of some misunderstandings," *Journal of Molecular Recognition*, vol. 17, no. 1, pp. 1–16, 2004.
- [22] C. A. Lipinski, F. Lombardo, B. W. Dominy, and P. J. Feeney, "Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings IPII of original article: S0169-409X (96)00423-1. The article was originally published in advanced drug delivery reviews 23 (1997) 3-25. 1," *Advanced Drug Delivery Reviews*, vol. 46, no. 1–3, pp. 3–26, 2001.
- [23] C. L. Mills, P. J. Beuning, and M. J. Ondrechen, "Biochemical function predictions for protein structures of unknown or uncertain function," *Computational and Structural Biotechnology Journal*, vol. 13, pp. 182–191, 2015.
- [24] A. H. Abdelmonsef, R. Dulapalli, T. Dasari, L. S. Padmarao, T. Mukkera, and U. Vuruputuri, "Identification of novel antagonists for Rab38 protein by homology modeling and virtual screening," *Combinatorial Chemistry & High Throughput Screening*, vol. 19, no. 10, pp. 875–892, 2016.
- [25] A. Haredi Abdelmonsef, "Computer-aided identification of lung cancer inhibitors through homology modeling and virtual screening," *Egyptian Journal of Medical Human Genetics*, vol. 20, no. 1, 2019.
- [26] S. F. Altschul, T. L. Madden, A. A. Schäffer et al., "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs," *Nucleic Acids Research*, vol. 25, no. 17, pp. 3389–3402, 1997.
- [27] J. A. Cuff, M. E. Clamp, A. S. Siddiqui, M. Finlay, and G. J. Barton, "Jpred: a consensus secondary structure prediction server," *Bioinformatics*, vol. 14, no. 10, pp. 892–893, 1998.
- [28] B. Contreras-Moreira and P. A. Bates, "Domain fishing: a first step in protein comparative modelling," *Bioinformatics*, vol. 18, no. 8, pp. 1141–1142, 2002.

- [29] J. Shi, T. L. Blundell, and K. Mizuguchi, "Fugue: sequence-structure homology recognition using environment-specific substitution tables and structure-dependent gap penalties 11 edited by B. Honig," *Journal of Molecular Biology*, vol. 310, no. 1, pp. 243–257, 2001.
- [30] J. D. Thompson, D. G. Higgins, and T. J. Gibson, "Clustal W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice," *Nucleic Acids Research*, vol. 22, no. 22, pp. 4673–4680, 1994.
- [31] N. Guex and M. C. Peitsch, "Swiss-model and the swiss-Pdb viewer: an environment for comparative protein modeling," *Electrophoresis*, vol. 18, no. 15, pp. 2714–2723, 1997.
- [32] R. A. Laskowski, M. W. MacArthur, D. S. Moss, and J. M. Thornton, "Procheck: a program to check the stereochemical quality of protein structures," *Journal of Applied Crystallography*, vol. 26, no. 2, pp. 283–291, 1993.
- [33] M. Wiederstein and M. J. Sippl, "ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins," *Nucleic Acids Research*, vol. 35, pp. W407–W410, 2007.
- [34] M. J. Sippl, "Recognition of errors in three-dimensional structures of proteins," *Proteins: Structure, Function, and Genetics*, vol. 17, no. 4, pp. 355–362, 1993.
- [35] A. D. Sol, H. Fujihashi, D. Amoros, and R. Nussinov, "Residue centrality, functionally important residues, and active site shape: analysis of enzyme and non-enzyme families," *Protein Science*, vol. 15, no. 9, pp. 2120–2128, 2006.
- [36] J. Dundas, Z. Ouyang, J. Tseng, A. Binkowski, Y. Turpaz, and J. Liang, "CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues," *Nucleic Acids Research*, vol. 34, pp. W116–W118, 2006.
- [37] T. Halgren, "New method for fast and accurate binding-site identification and analysis," *Chemical Biology & Drug Design*, vol. 69, no. 2, pp. 146–148, 2007.
- [38] S. Kawatkar, H. Wang, R. Czerminski, and D. Joseph-McCarthy, "Virtual fragment screening: an exploration of various docking and scoring protocols for fragments using glide," *Journal of Computer-Aided Molecular Design*, vol. 23, no. 8, pp. 527–539, 2009.
- [39] Schrödinger, *Ligprep, Version 2.5*, Schrödinger, LLC, New York, NY, USA, 2010.
- [40] M. El-Naggar, M. E. Mohamed, A. M. Mosallam, W. Salem, H. R. Rashdan, and A. H. Abdelmonsef, "Synthesis, characterization, antibacterial activity, and computer-aided design of novel quinazolin-2,4-dione derivatives as potential inhibitors against *Vibrio cholerae*," *Evolutionary Bioinformatics*, vol. 16, pp. 1–13, 2020.
- [41] A. H. Abdelmonsef and A. M. Mosallam, "Synthesis, in vitro biological evaluation and in silico docking studies of new quinazolin-2,4-dione analogues as possible anticarcinoma agents," *Journal of Heterocyclic Chemistry*, pp. 1–18, 2020.
- [42] R. Kalirajan, S. Sankar, S. Jubie, and B. Gowramma, "Molecular docking studies and in-silico ADMET screening of some novel oxazine substituted 9-anilinoacridines as topoisomerase II inhibitors," *Indian Journal of Pharmaceutical Education and Research*, vol. 51, no. 1, pp. 110–115, 2017.
- [43] Schrödinger, *Glide, Version 5.6*, Schrödinger, LLC, New York, NY, USA, 2010.
- [44] Schrödinger, *QikProp Version 3.3*, Schrodinger, LLC, New York, NY, USA, 2010.