IDENTIFYING NEW 9-ANILINOACRIDINE-BASED PARP1 INHIBITORS USING TEXT MINING AND INTEGRATED MOLECULAR MODELING APPROACHES

Baliwada Aparna[a], Kalirajan Rajagopal[b], Potlapati Varakumar[c], Kannan Raman[d], Gowramma Byran[e], Srikanth Jupudi[f]

Article History: Received: 08.06.2022 Revised: 07.07.2022 Accepted: 02.08.2022

Abstract: The PARP research on cancer and ischemia is advancing at a breakneck pace. Olaparib, Rucaparib, Niraparib, and Talazoparib are the four PARP1 inhibitors currently on the market, according to the FDA. All of these compounds are non-selective PARP1 inhibitors. Novel and selective PARP1 inhibitors are desperately needed right now. A small molecule database (Specs SC) was used to find new selective lead inhibitors of PARP1 in this study. The 9-anilinoacridine scaffold is a new fragment that is employed as a PARP1 inhibitor and anti-proliferative drug. Thus, 21 compounds containing 9-Anilinoacridine fragments were discovered and virtually tested in the binding site of target protein PARP1 based on text mining studies. In molecular dynamics (MD) simulations, compounds with high docking scores were employed. The anticipated binding energies were compared to known PARP1 inhibitors using free energy calculations. Docking study revealed that among all 21 compounds 1v showed highest g score. Prime MMGBSA analysis gave the relative binding energies of 1v. The essential amino acid interactions of these newly discovered hits in the binding pocket were also studied in depth in order to gain a better understanding of the structural properties required for next-generation PARP1 inhibitors. Thus, we identified novel 9-Anilinoacridine-based hits against the PARP1 enzyme using a mix of text-mining and integrated molecular modelling techniques.

Keywords: PARP1 inhibitors; virtual screening; text mining; molecular docking; molecular dynamics simulations; Pharmacophore; 9-Anilino acridine derivatives.

[a]. Department of Pharmaceutical Chemistry, JSS College of Pharmacy (JSS Academy of Higher Education & Research), Ooty, The Nilgiris, Tamil Nadu, India.
[b]. Department of Pharmaceutical Chemistry, JSS College of Pharmacy (JSS Academy of Higher Education & Research), Ooty, The Nilgiris, Tamil Nadu, India.
[c]. Department of Pharmaceutical Chemistry, JSS College of Pharmacy (JSS Academy of Higher Education & Research), Ooty, The Nilgiris, Tamil Nadu, India.
[d]. Department of Pharmaceutical Chemistry, JSS College of Pharmacy (JSS Academy of Higher Education & Research), Ooty, The Nilgiris, Tamil Nadu, India.
[e]. Department of Pharmaceutical Chemistry, JSS College of Pharmacy (JSS Academy of Higher Education & Research), Ooty, The Nilgiris, Tamil Nadu, India.
[f]. Department of Pharmaceutical Chemistry, JSS College of Pharmacy (JSS Academy of Higher Education & Research), Ooty, The Nilgiris, Tamil Nadu, India.

DOI: 10.31838/ecb/2022.11.05.004

INTRODUCTION

According to the American Cancer Society's projections, there will be an estimated: A new diagnosis of ovarian cancer will be given to about 19,880 women. Ovarian cancer will claim the lives of about 12,810 women[1]. The sixth most common cancer among women, ovarian cancer kills more people than any other malignancy of the female reproductive system. The BRCA1 gene was discovered to be a hereditary breast cancer risk factor in 1990. It was sequenced in 1994, and it has subsequently been demonstrated to participate in the intricate HR pathway of DNA repair[2]. By the time they reach the age of 70, women who have heterozygous detrimental mutations within the BRCA1 gene have a cumulative chance of having breast cancer of 57% and ovarian cancer of 40% [3]. Women with a harmful BRCA2 mutation had a 49 percent chance of getting breast cancer by the time they were 70 and an 18 percent chance of getting ovarian cancer[4]. The wild-type allele of the BRCA gene is lost in tumors that develop in BRCA mutation carriers, which exclusively express the altered allele[5]. As a result, the BRCA-dependence repair mechanisms in tumor cells are disabled. DNA repair mechanisms are controlled by BRCA1/2 and poly-ADP ribose polymerase (PARP). Breast and ovarian cancers arise in the presence of BRCA insufficiency due to the accumulation of mutations brought about by error-prone DNA repair. Recent clinical trials have examined the effectiveness of using PARP inhibition to treat cancers that lack the BRCA1/2 gene. BRCA-deficient cancer cells are 1000 times more susceptible to single-agent PARP inhibition because of the different synthetic lethality of the underlying deficiency in BRCA-dependent DNA repair[6][7]. The use of PARP inhibitors in BRCA-deficient malignancies serves as an excellent illustration quality, which states that functional inhibition of two proteins results in cell death whereas blockage of either, protein alone does not. A cell will live as long as PARP or BRCA function is still present. Because these non-cancerous

cells have one functioning copy of BRCA, blocking PARP should not have an impact on them. However, losing both functions makes life impossible. This idea was supported in vitro and in mice models of BRCA-deficient malignancies by preclinical research, which prompted the clinical development of PARP inhibitors for these genetic cohorts of patients[8]. PARP1 is encoded by the ADP ribosyl transferase-1 gene [9] and plays its part by stopping the death of cells. The PARP1 enzyme is involved in various biological processes, including cell formation, inflammation, stress responses, cancer, etc. Molecular targeting must be prioritized in contemporary drug design methodologies. One of the targets exploited in the development of anticancer medicines is PARP1. When DNA is broken, PARP1 attaches to the damage site and catalyses PAR production, which recruits host DNA repair proteins (Figure 1). PARP1 AutoPARylation results in electrostatic instability and separation from the site of DNA damage, allowing protein repair to locate the lesion[10][11]. The suppression of excessive PARP1 action is the focus of anticancer medication development targeting this enzyme. The influence of radiation and chemotherapeutics is enhanced when PARP-1 is inhibited[12]. PARP1 inhibitors have antiproliferative activity, particularly in BRCA-deficient cells[13]. The discovery of new lead chemicals from small molecule data banks that show promise in various biovariouslems has piqued interest in this subject in recent years[14][15][16][17]. The development of several beneficial biologically active chemicals, notably antiproliferative medicines, has piqued interest in 9-anilino acridines. Olaparib, an FDA-approved anticancer medication, is currently on the market and is used to treat a variety of malignancies, including myelogenous leukemia[18]. Another anticancer medication Palbociclib was recently licensed by the FDA for the treatment of metastatic breast cancer[19]. Ponatinib, which contains the piperazine ring, has been developed piperazine for treating chronic myeloid leukemia[20]. Moreover, one of the currently used FDA-approved PARP1 in FDA-approved parib also has an antiproliferative property. By swiftly screening a vast database with keywords, text mining aids in the discovery of molecules of interest[21]. Thus, in the current study, a text mining study was considered for identifying the 9-Anilino acridine derivatives from a large database. To do this, 5687 compounds from the Schrodinger database were converted into IUPAC text filing the Maestro's Protein Preparation. Using Maestro's Protein Preparation text mining software, this text file was screened for compounds that included the phrase "9-anilino acridines" and 26 molecules were discovered. Our rigorous virtual screening technique was then applied to uncover new hits against PARP1 using filtered 9-anilino acridines- based compounds.

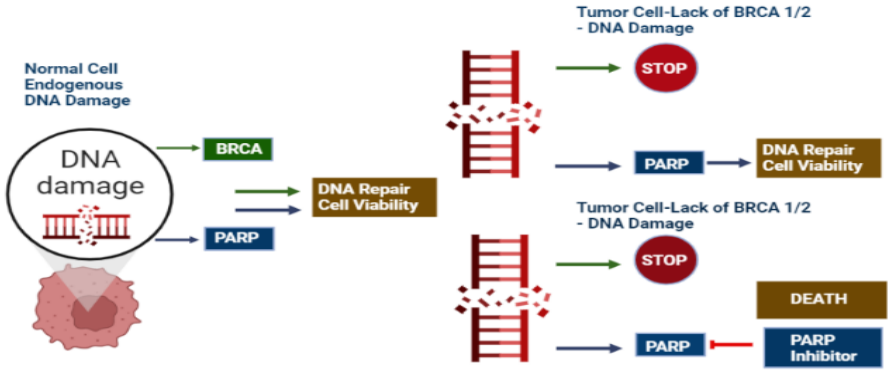

Figure 1. Mechanism of Action of PARP inhibitors (PARP is necessary for single-strand break repair, while BRCA1/2 is necessary for double-strand break repair. Single-strand breaks become double-strand breaks when PARP is inhibited. The build-up of double-strand breaks overwhelms DNA repair systems in the absence of a functioning BRCA 1/2. Cell death is brought on by persistent DNA damage[22].).

Materials and methods

Ligand preparation

Using the Maestro molecular modeling suite's LigPrep module (LigPrep, Schrodinger, LLC, New York, NY, 2019) and the OPLS3 forcefield (Banks et al., 2005), the 2D structure of selected 26 ligands were transformed to the signature. It might be challenging to be aware of the proper protonation states of the ligand in physiological situations.

The Maestro molecular modeling suite's Epik module (Shelley et al., 2007) was used to accurately predict the protonation states of filtered compounds from text-mining studies that also produced stereoisomers of substances at neutral pH.

Protein preparation

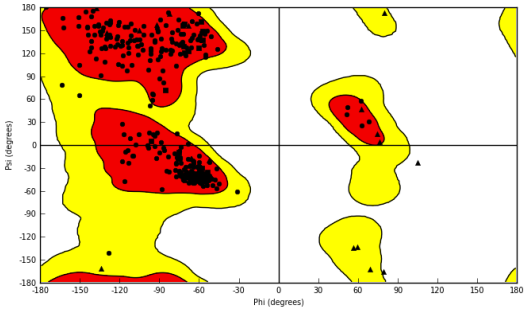

Figure 2. Ramachandran Plot (99% of residues in allowed region of Ramachandran plot[23])

The Protein Data Bank (PDB) (http://www.rcsb.org) this target contains 355 residues; we chose to use the PARP1 target (PDB ID: 2IOG). Maestro's Protein Preparation module was used to pre-process the complete structure using the default settings. The Maestro Prime module replaces the backbone's missing side chains, atoms, and loops. 99% of residues in allowed region of Ramachandran plot[23] (Figure 2). The influence of water molecules at the catalytic domain was modelled using water molecules inside a sphere with a radius of 5.0 from the center of mass of the ligand. For protonation states, structural optimization, and minimization, the PROPKA and OPLS3 (Optimized Potentials for Liquid Simulations) molecular force field with RMSD of crystallographic heavy atoms kept at 0.30 Å. were utilized (Protein Preparation, Version 2.5, Schrodinger, LLC, New York, 2019).

Molecular docking

Using the Glide module of Schrödinger suite 2020-2 in the XP (Extra precision) mode, docking techniques anticipate ligands binding in the target protein's binding pocket with the most proper conformation. The Glide was utilized to execute molecular docking between the identified ligands and the PARP1 target to achieve postures with high docking scores. The conformations obtained during docking are scored using the Glide scoring function in Glide (Raha & Merz, 2004). Ten bioactive conformers were created for each ligand, with residues positioned in the binding pocket inside ten co-crystallized ligands. To generate the top-docking poses, all ligands were docked in the PARP1 binding pocket using the Maestro molecular modeling method's grid-based docking program XP visualizer (Friesner et al., 2004; Jacobson et al., 2004).

Molecular dynamics (MD) simulations

MD simulations are required to determine the correct binding matches for investigated complexes during docking experiments (Hou, Wang, Li, & Wang, 2011). The top-docking complexes were employed with the Desmond v 4.9 for MD simulations. To explore the conformational stability of identified hit compounds at the PARP1 binding site, MD simulations are used to evaluate structural and dynamical changes throughout the protein-ligand interactions (Desmond, Version 4.9). The orthorhombic single-point charge (SPC) water models are used to solve the complex structures (Berendsen, Postma, van Gunsteren, & Hermans, 1981). The systems were neutralised with counter ions Na+ (75 atoms, +75 charge), Cl- (72 atoms, -72 Charge). On periodic boundary conditions, the system was established with a cut-off of 10 for Lennard-Jones interactions (Friesner et al., 2004). A 2.0 fs time step was employed in the integration steps. The temperature and pressure of the systems were controlled at 310 K and 1.01325 bar, respectively, using the Nose-Hoover thermostat (Hoover, 1985) and Martyna-Tobias Klein procedures (Ma & Tuckerman, 2010).

The system is minimized and equilibrated before the production run. Then, for each system, 100-ns MD simulations were run.

Molecular mechanics/generalized born surface area (MM/GBSA)

For protein-ligand complexes, MM/GBSA was utilized to calculate the binding free energy. MM/GBSA computations for the analyzed systems were performed using the Maestro Prime module (Jacobson et al., 2004). The MD simulations of each selected hit compound at every 10 ps were utilized to generate the trajectories of protein ligand complexes (Desmond, Version 4.9). The VSGB solvation system (Li et al., 2011), which is a practical solvation parameterization, and the OPLS-3 forcefield (Banks et al., 2005) was employed for protein flexibility. One hundred trajectories were employed during the entire production MD simulations (100-ns), and average values were calculated in the binding free energy predictions of selected hits using MM/GBSA calculations.

An enrichment analysis was carried out by combining the decoy set obtained from the Schrodinger database with 10 active compounds (pIC50 > 8.3) in order to verify the usefulness of the created model. In order to accurately rank the compounds and evaluate if it is possible to distinguish between active and inactive molecules, the robust initial enhancement (RIE) and enrichment factor of the decoy set molecules were determined.

Additionally, a training set was used to conduct a Y-scrambling or randomization test to assess the validity of the developed DDHRR 1 model. While the dependent variable (activity data) was randomly assigned throughout the investigation, all independent variables (molecular descriptors) remained constant during the model construction. After 10 Y scrambling tests were conducted, and the models with low R2 and Q2 values were deemed to be trustworthy (ESI, Table S3).

The developed hypothesis DDHRR 1's prediction quality and systematic errors were assessed using the mean absolute error (MAE)-based criteria method and the DTC software lab's (https:// dtclab.webs.com/sofwafre-tools) Xternal Validation plus 1.1 tool. The MAE-based criterion, which categorizes the prediction quality of the model into "good," "bad," and "moderate," based on their evaluation parameters such as standard deviation (SD), MAE, and absolute error, was used to calculate all of the external validation parameters of the hypothesis (AE). The definition of the standard deviation of absolute error (sAE) is as follows: The predictions that do not meet either of the aforementioned criteria are classified as (a) bad predictions (MAE + 3 sAE > 0.25 training set range and MAE > 0.15 training set range); (b) good predictions (MAE + 3 sAE # 0.2 training set range and MAE # 0.1 training set range); and (c) moderate predictions. The MAE-based criterion technique was used to reject nearly 5% of the ligands with high AE values in order to completely rule out the potential of outlier predictions. While the predictivity of the model was punished using 100% MAE-based criteria, about 95% of compounds were used to identify the external data set of badly forecasted data. Additionally, using the Xternal validation plus 1.1 tool, absolute value means of the average absolute prediction errors (AAE), average prediction errors (AE), mean of positive prediction errors (MPE), and number of positive prediction errors (nPE) were used to test for the presence of systematic error prediction. If any one or more of the prediction errors listed above follow the guidelines established by the MAE-based criteria technique (ESI, Table S4), the results produced by the current approach are deemed adequate.

RESULTS AND DISCUSSION

The binding pocket of the PARP1 target was evaluated with 5467 (9-anilinoacridine) compounds from the Specs SC

database (Figure S4, Supplementary material). As a result, we investigated 9-Anilino acridine compounds and used Olaparib as a positive control (Figure 3).
The two structures were superimposed at the enzyme's binding site to compare the fingerprints of the detected hit molecule and Olaparib in the PARP1 binding pocket (Figure S5, Supplementary material). Tyr896 and Tyr907 were the most prevalent residues found in Olaparib and the hit compound 1v (Figure 3).

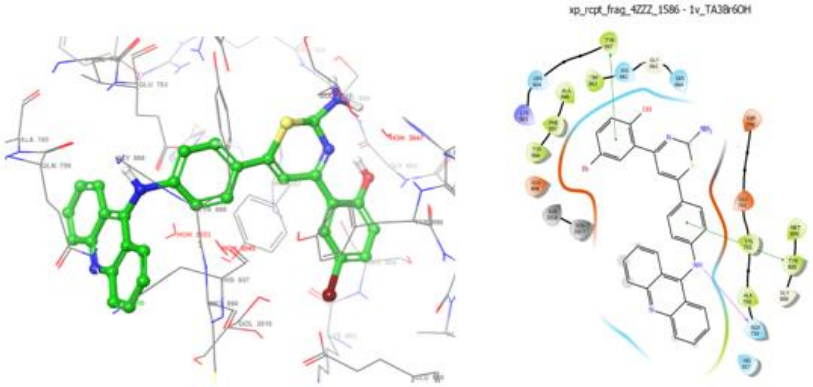

Figure 3. 1v 3D interaction and 2D interaction images

This study used text mining and integrated molecular modeling to find new 9-Anilino acridine-based PARP1 inhibitors. Figure 1 depicts the entire process employed in the current study We picked the top-10 compounds with high docking scores as hit molecules based on docking scores within 5467 compounds, and MD simulations were performed for these complexes. One of the FDA-approved medications (Olaparib) was also included in MD simulations to compare the binding free energy.
The 1v hit compound in the PARP1 binding pocket was subjected to 100 ns MD simulations, with the associated coordinates gathered separately in the trajectory file. Table 1 summarises the results of ligand docking and post-processing binding free energy analyses using MD simulation trajectories. Figure S1 (Supplementary material) shows Ca atoms' root-mean square deviations (RMSDs) in the studied systems. After 20-ns of MD simulations, all the examined systems have relaxed and do not show significant variations at the RMSD values. All of the systems evaluated had an average RMSD of less than 3.0.
The RMSDs of the examined ligand were also explored throughout the MD simulations. Figure 4 shows the LigFitLig RMSD plot (Supplementary material). The RMSD of the examined molecule based on its starting conformation during the simulation is depicted in this plot. As a result, the RMSD number represents the internal oscillations of the ligand's nonhydrogen atoms.
When the protein-ligand complex is first aligned as a reference, the RMSD plot of LigFit Prot reveals the RMSD of ligands. The RMSD of the compounds' heavy atoms (i.e., non-hydrogens) (i.e., structural stabilities) was then computed. During the simulations, at the binding site). The LigFit Protector is a LigFit Protector. Figure 4 shows the RMSD plot (Supplementary material). The average RMSD values for the corresponding variables are shown in Table 2: Throughout the MD simulations, most of the investigated compounds have modest rotating motions. Place of binding. RMSF values were also calculated during simulations to investigate the influence of defined hits on the mobility of the target protein's backbone atoms. The complex examines produced RMSF of polypeptide backbone atoms of each amino acid residue to determine the target structure's fluctuation zones. In MD simulations, high RMSF values in the RMSF plot indicate highly mobile areas. Figure 4 shows the RMSF graph of 1v hit compound.
Peaks on this graph show the protein residues that fluctuated the most during simulations. The tails (N- and C-terminals) of proteins are frequently found to be more variable than any other protein part. Loop domains fluctuate more than secondary protein structures, which have more consistent profiles throughout the simulations. The associations that emerge in the chosen path over 15% of the simulation duration (0.00–100.00 ns) are shown in Figures 4. Crucial interactions are formed via hydrogen bonds with Gln759 and Met890, and p-p stacking interactions are formed with Tyr889. For selected hit molecules 1v that have a similar MM/GBSA score to Olaparib. It is worth noting that Gln759 was identified as a shared essential residue in both Olaparib and 1v.
The timeline depiction of interactions and contacts is shown in Figures 5 and 6. The top panel shows the total number of specific interactions between the ligand and the protein during MD simulations. In contrast, the bottom panel shows which amino acid residues at the target protein interact with the examined ligand. Many residues allow more than one direct interaction with the ligand, characterized by a darker orange color on the map, according to the scale to the right of the image.
1v's protein-ligand interactions graph demonstrates that Ala762, Gln759, Tyr889, and Met890 have stable nonbonded chemical connections throughout the simulations. While the protein-ligand interactions graph of 1v demonstrates that Ala762, Gln759, Tyr889, and Met890 establish regular nonbonded chemical contacts (Figure 5), Olaparib uses the following residues: Gly863, Arg878, Tyr896, Ser904, and Tyr907 to form comparable interactions fig6. We also tested three other known PARP1 inhibitors with their co-crystallized PARP1 complexes using the same MD simulation parameters as the hit compounds and Olaparib. While Olaparib has a slightly higher average interaction energy score than the discovered hit compounds, the MM/GBSA data suggest that some of the selected hits (e.g., 1z, 1v, 1d) have similar or better average interaction energy values than the other three FDA

approved medications. 1v and Olaparib have docking scores of 6.11 kcal/mol and 14.92 kcal/mol, respectively. Based on docking scores, the discovered 9-Anilino acridine derivative 1v is more selective than Olaparib. Olaparib is a PARP1 inhibitor approved by the FDA as cancer targeted therapy. The 9-Anilino acridine ring is also included in the structure of this chemical.
We used MetaCore/MetaDrug (MC/MD) platform to test the "cancer-QSAR" model to predict the therapeutic activity of the hit chemical 1v. The model's sensitivity, specificity, accuracy, and Matthews Correlation Coefficient (MCC) values were 0.89, 0.83, 0.86, and 0.72, respectively.

Pharmacophore Generation and Validation

On the basis of 10 active compounds with pIC50 values greater than 8.3, the best five-point model DDHRR 1 was developed. Based on the values obtained for the cross-validated coefficient (R2) of 0.9076, survival score of 5.639, Pearson-correlation coefficient (r) of 0.9454, regression coefficient of 0.89, lower significance level of variance ratio (p) of 4.271017, and highest variance ratio (F) value of 83.5, the pharmacophore hypothesis DDHRR 1 was deemed to be the best model (Supplementary data). Two hydrophobic, two ring, and a hydrogen bond donor make up the DDHRR 1 model. Angles and distances are key characteristics for MMP-9's inhibitory function, as demonstrated by the alignment of hypothesis DDHRR 1 over actives (10) and inactives (7). Additionally, a higher value of QSAR model stability (0.709 on a maximum scale of 1), an SD of 0.3473, an RMSE of 0.48, and a Pearson correlation coefficient (r) of 0.9454 showed a better degree of condence in the hypothesis model DDHRR 1. The constructed model was relevant in the training set, as evidenced by the high coefficient of determination (R2) of 0.9076. The higher cross-validated correlation-coefficient (Q2) of 0.8170 PLS-4 (ESI, Table S1) further demonstrated the validity and capability of the hypothesis model in experimental activity prediction of test chemicals. The results of comparing the experimental activity and phase-predicted activities of the molecules in the test and training sets are shown in Figs. 3a and b, respectively. The training and test scatter plots show a moderate difference between expected and experimental activity.

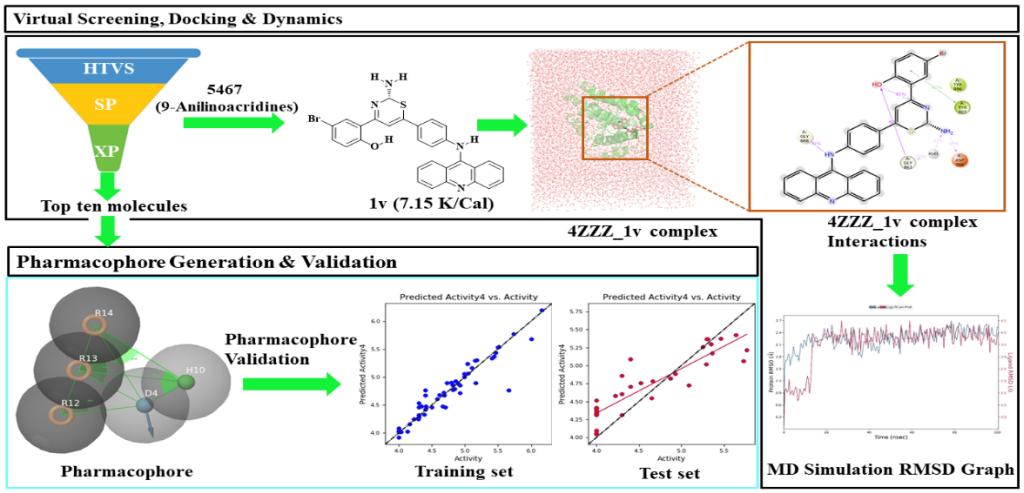

Figure 4. Docking protocol, Molecular dynamics and Pharmacophore generation and validation

Table 1: Glide docking scores

Compound	Glide score	Lipophilic EvdW	Phob En	H Bond	XP Electro	Low MW	XP Penalties	Rot Penal
1v	-7.15	-6.39	0	-0.39	0.05	0	0	0.09
1z	-7.01	-6.13	-0.17	-0.43	-0.04	0	0	0.12
1d	-6.17	-6.2	0	-0.66	0.06	0	0	0.16
1t	-6.12	-5.52	-0.2	-0.15	0.13	0	0	0.09
1e	-5.9	-6.12	0	-0.47	0.02	0	0	0.16
1h	-5.79	-6.24	0	-0.28	0.06	0	0	0.15
1p	-5.78	-5.98	0	-0.62	-0.07	0	0	0.17
1y	-5.75	-5.86	0	0	-0.02	0	0.32	0.14
1i	-5.72	-5.82	-0.15	-0.58	-0.17	0	0.13	0.15
1j	-5.65	-5.87	0	-0.43	0.22	0	0	0.15
1b	-5.58	-6	0	-0.33	0.11	0	0	0.12
1c	-5.5	-5.88	-0.2	-0.22	-0.04	0	0	0.12
1f	-5.4	-6.11	0	-0.05	0.06	0	0	0.16
1o	-5.24	-5.87	0	-0.21	0.14	0	0	0.17
1m	-5.15	-5.53	-0.25	-0.1	-0.1	0	0	0.15
1g	-4.87	-5.43	-0.2	-0.01	0.08	0	0	0.15

1l	-4.7	-5.45	-0.26	-0.04	0.15	0	0.18	0.14
1a	-4.59	-5.46	-0.26	0	0.15	0	0.32	0.12
1k	-4.4	-3.1	-0.15	-0.32	-0.35	0	0.08	0.15
1q	-4.03	-3.09	0	-0.48	-0.17	0	0	0.13
1w	-3.75	-5.25	0	-0.54	0.07	0	0.07	0.13
1n	-3.62	-4.47	0	0	-0.05	0	0.28	0.15
1s	-2.4	-2.38	0	-0.48	0.26	0	0.06	0.14
1x	-1.15	-1.3	0	-0.21	-0.33	0	1.57	0.12
1u	-0.92	-1.59	0	-0.29	-0.1	0	1.15	0.09
1r	-0.7	-2.16	0	-0.72	0.05	0	2.28	0.14
Olaparib (Std)	-13.523	-54.374	0	-2.8	0.08	0	2.34	0.16

Table 2: Binding free energy calculation using Prime/MM-GBSA approach. (The relative binding-free energies (kcal/mol) obtained by Prime MM–GBSA, where MMGBSA dG Bind = Complex − Receptor – Ligand)

Comp	ΔG bind (Kcal/mol)	ΔG bind Coulomb	ΔG bind covalent	ΔG bind Vander	ΔG bind H Bond	ΔG bind Lipophilic
1v	-66.49171639	-4.18893	-1.35374	-58.8862	-0.01384	-27.4339
1z	-60.00616749	-0.81833	1.303844	-53.4898	-0.86131	-29.7363
1d	-62.48823687	-28.2791	-0.34761	-55.0559	-0.90108	-31.2744
1t	-62.44412201	-10.2143	-2.43759	-57.2534	-1.08165	-28.0787
1e	-58.72047618	-19.5877	4.497234	-53.0051	-0.93736	-31.7566
1h	-68.13400645	-27.0246	3.543555	-54.3662	-0.98075	-31.3607
1p	-83.89780101	-57.7245	15.83091	-49.7239	-1.21923	-31.685
1y	-64.68929947	-28.2915	4.605725	-50.8881	-1.08584	-30.9985
1i	-63.32624045	-33.23	9.525944	-60.8497	-1.71956	-31.3147
1j	-60.27032405	-17.9346	-0.03958	-48.133	-0.55517	-29.9884
1b	-67.59953803	-17.1804	2.866804	-54.3901	-0.40983	-29.5697
1c	-53.8791456	-29.8598	7.656486	-50.764	-0.81732	-28.5608
1f	-63.52570406	-22.8444	4.71572	-51.9496	-0.87411	-30.634
1o	-52.66252519	-16.9291	12.34218	-56.756	-0.1564	-32.0143
1m	-76.74147769	-59.6281	16.20851	-47.9299	-3.49017	-29.7028
1g	-54.00356059	-29.0223	6.274753	-47.5316	-1.42172	-30.3041
1l	-68.21018403	-24.329	4.25761	-54.3782	-1.43291	-29.625
1a	-58.84485589	-26.5697	4.611916	-56.1414	-1.2882	-29.036
1k	-45.14636481	-20.2512	1.806595	-41.7393	0.305122	-26.577
1q	-74.83119429	-15.8321	2.750097	-58.0127	-1.0481	-29.8893
1w	-65.20701327	-33.6139	12.24531	-68.0795	-0.70432	-31.0574
1n	-65.13163358	-55.1128	28.93047	-61.3352	-5.75397	-32.1806
1s	-40.68418876	-7.79577	5.760213	-35.3374	-0.39227	-15.5054
1x	-25.33605535	5.257303	1.394554	-37.4145	0.653675	-12.3614
1u	-50.95537589	-34.7084	-3.71413	-28.2942	-0.20059	-13.1211
1r	-29.21809815	-11.0596	-2.04769	-31.3845	-0.7604	-14.0775

Table 3. In silico ADMET screening for proposed compounds (1a-z)

Compound	Mol. Wt.	Donor HB	Accpt HB	QPlog HERG	metab	QPlog Khsa	%Human Oral Absorption	Rule of Five
1v	553.475	4	4.25	-8.869	5	1.284	1	2
1z	543.469	4	4.25	-8.681	5	1.366	1	2
1d	474.579	4	4.25	-8.841	5	1.142	1	1
1t	502.633	3	4.25	-9.213	5	1.562	1	2
1e	472.606	3	3.5	-8.872	5	1.538	1	1
1h	537.476	3	3.5	-8.899	4	1.519	1	2
1p	553.475	4	4.2	-8.827	5	1.279	1	2
1y	504.605	4	5	-8.772	6	1.18	1	2
1i	472.606	3	3.5	-8.865	5	1.536	1	1
1j	503.577	3	4.25	-8.889	5	1.354	1	2
1b	507.051	3	3.5	-8.715	5	1.635	1	2
1c	515.632	3	5	-8.879	6	1.456	1	2
1f	490.578	5	5	-8.775	6	0.917	1	0

1o	476.57	3	3.5	-8.839	4	1.417	1	1
1m	488.606	3	4.2	-8.859	5	1.405	1	1
1g	537.476	3	3.5	-8.945	4	1.515	1	2
1l	488.606	3	4.25	-8.946	5	1.429	1	1
1a	502.633	3	4.25	-9.082	5	1.559	1	2
1k	488.606	3	4.25	-8.937	5	1.413	1	1
1q	472.606	3	3.5	-8.815	5	1.515	1	1
1w	548.658	3	5.75	-8.512	7	1.419	1	2
1n	486.633	3	3.5	-8.997	5	1.660	1	1
1s	537.476	3	3.5	-8.891	4	1.516	1	2
1x	476.57	2	3.5	-8.884	4	1.42	1	1
1u	527.47	3	3.5	-8.761	4	1.603	1	1
1r	553.475	4	4.25	-8.728	5	1.258	1	2
Recommended values	130-725	0– 6	2-20	-2-6.5	1 – 8	-1.5 – 1.5	>80% is high <25% is poor	max 4

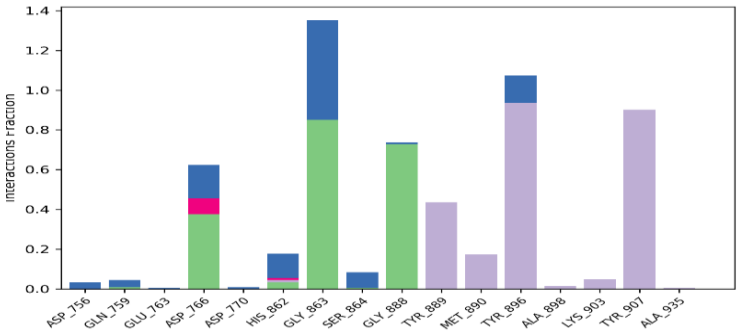

Figure 5. 2D Protein-ligand interactions graph for the selected hit compound Mol-1v.

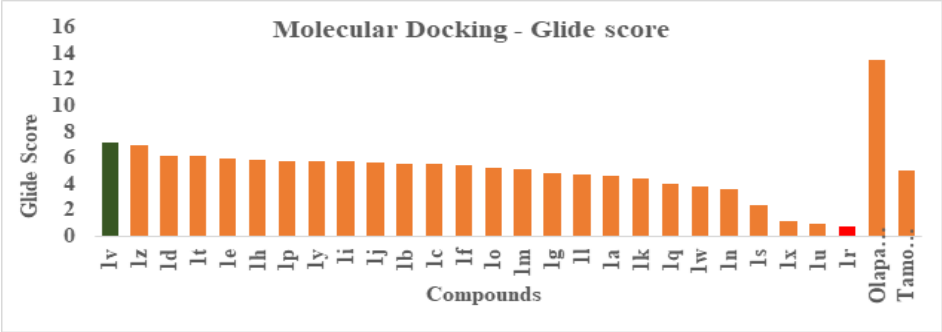

Figure 6 Glide scores The highest glide score 1v (-7.15) green bar, least glide score 1r (0.7) red bar and Standard Olaparib (-13.523)

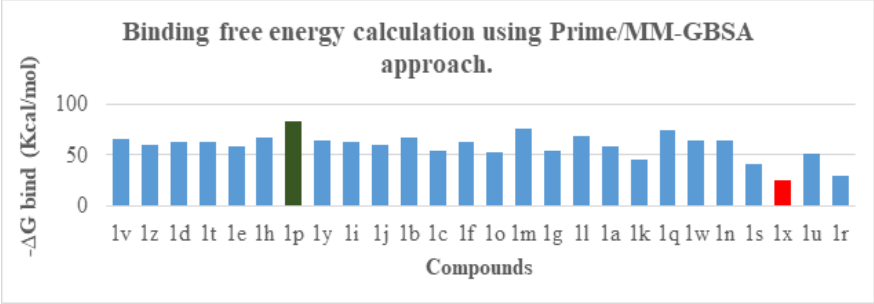

Figure 7. MMGBSA scores The highest glide score 1v (-66.49171639) green bar, least glide score 1r (-29.21809815) red bar.

CONCLUSION

We used advanced text mining and coupled molecular modeling techniques to find 9-Anilino acridine based new PARP1 small molecule inhibitors in this study. This study chose a total of 5467 compounds containing 9-Anilino acridine from the Specs-SC small molecule database. These compounds were screened via molecular docking against the PARP1 target, and the top-docking pose of 1 hit was subsequently employed in lengthy (100 ns) MD simulation. MM/GBSA binding free energy calculations were performed on all ligands. The FDA-approved PARP1 inhibitors were likewise subjected to the same process. we took advantage of text mining approaches in this study, which can reduce the total number of compounds in small molecule large databases to manageable numbers for further analysis in a short period. The collected statistical findings suggest that the model has high accuracy and predictive capacity. With normalized values between 0 (inactive) and 1 (active), the model forecasts therapeutic activity potentials of tested substances (active). 1v has a therapeutic activity potential of 0.60 according to the Cancer-QSAR model in MC/MD. This chemical was further tested in 26 different toxicity models, and it was discovered that the compound has no toxicity in any of them. Finally, we looked at finding compounds that could be employed to block PARP1 activity during this stringent combination screening. This novel scaffold could pave the way for developing small PARP1 inhibitors.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the authors.

REFERENCES

i. Ovarian Cancer Statistics | How Common is Ovarian Cancer. https://www.cancer.org/cancer/ovarian-cancer/about/key-statistics.html#references (accessed Jun. 2022).

ii. Tutt, A.; Ashworth, A. The relationship between the roles of BRCA genes in DNA repair and cancer predisposition. Trends Mol. Med. **2002**, *8*(12), 571–576.

iii. Chen, S.; Parmigiani, G. Meta-Analysis of BRCA1 and BRCA2 Penetrance. J. Clin. Oncol. **2007**, *25*(11), 1329.

iv. Annunziata, C.M.; Bates, S.E. PARP inhibitors in BRCA1/BRCA2 germline mutation carriers with ovarian and breast cancer. F1000 Biol. Rep. **2010**, *2*(1).

v. Lim, E.; Vaillant, F.; Wu, D.; Forrest, N.C.; Pal, B.; Hart, A.H.; Asselin-Labat, M.L.; Gyorki, D.E.; Ward, T.; Partanen, A.; Feleppa, F.; Huschtscha, L.I.; Thorne, H.J.; Fox, S.B.; Yan, M.; French, J.D.; Brown, M.A.; Smyth, G.K.; Visvader, J.E.; Lindeman, G.J. Aberrant luminal progenitors as the candidate target population for basal tumor development in BRCA1 mutation carriers. Nat. Med. **2009**, *15*(8), 907–913.

vi. Kaelin, W.G. The Concept of Synthetic Lethality in the Context of Anticancer Therapy. Nat. Rev. Cancer 2005 59 **2005**, *5*(9), 689–698.

vii. Iglehart, J.D.; Silver, D.P. Synthetic lethality--a new direction in cancer-drug development. N. Engl. J. Med. **2009**, *361*(2), 189–191.

viii. Lord, C.J.; Ashworth, A. Targeted therapy for cancer using PARP inhibitors. Curr. Opin. Pharmacol. **2008**, *8*(4), 363–369.

ix. Mégnin-Chanet, F.; Bollet, M.A.; Hall, J. Targeting poly(ADP-ribose) polymerase activity for cancer therapy. Cell. Mol. Life Sci. **2010**, *67*(21), 3649–3662.

x. Kawaichi, M.; Oka, J.; Zhang, J.; Ueda, K.; Hayaishi, O. Properties of poly(ADP-ribose) synthetase and ADP-ribosyl histone splitting enzyme. Princess Takamatsu Symp. **1983**, *13*, 121–128.

xi. Satoh, M.S.; Lindahl, T. Role of poly(ADP-ribose) formation in DNA repair. Nature **1992**, *356*(6367), 356–358.

xii. Michels, J.; Vitale, I.; Saparbaev, M.; Castedo, M.; Kroemer, G. Predictive biomarkers for cancer therapy with PARP inhibitors. Oncogene **2014**, *33*(30), 3894–3907.

xiii. Chen, A. PARP inhibitors: its role in treatment of cancer. Chin. J. Cancer **2011**, *30*(7), 463.

xiv. Home - ClinicalTrials.gov. https://www.clinicaltrials.gov/ (accessed Jun. 2022).

xv. Salmas, R.E.; Senturk, M.; Yurtsever, M.; Durdagi, S. Discovering novel carbonic anhydrase type IX (CA IX) inhibitors from seven million compounds using virtual screening and in vitro analysis. J. Enzyme Inhib. Med. Chem. **2016**, *31*(3), 425–453.

xvi. Salmas, R.E.; Unlu, A.; Yurtsever, M.; Noskov, S.Y.; Durdagi, S. In silico investigation of PARP-1 catalytic domains in holo and apo states for the design of high-affinity PARP-1 inhibitors. J. Enzyme Inhib. Med. Chem. **2016**, *31*(1), 112–120.

xvii. Mirza, S.B.; Salmas, R.E.; Fatmi, M.Q.; Durdagi, S. Virtual screening of eighteen million compounds against dengue virus: Combined molecular docking and molecular dynamics simulations study. J. Mol. Graph. Model. **2016**, *66*, 99–107.

xviii. Audeh, M.W.; Penson, R.T.; Friedlander, M.; Powell, B.; Bell-McGuinn, K.M.; Scott, C.; Weitzel, J.N.; Carmichael, J.; Tutt, A. Phase II trial of the oral PARP inhibitor olaparib (AZD2281) in BRCA-deficient advanced ovarian cancer. https://doi.org/10.1200/jco.2009.27.15_suppl.5500 **2009**, *27*(15_suppl), 5500–5500.

xix. Morikawa, A.; Henry, N.L. Palbociclib for the Treatment of Estrogen Receptor-Positive, HER2-Negative Metastatic Breast Cancer. Clin. Cancer Res. **2015**, *21*(16), 3591–3596.

xx. Pavlovsky, C.; Chan, O.; Talati, C.; Pinilla-Ibarz, J. Ponatinib in the treatment of chronic myeloid leukemia and philadelphia chromosome positive acute lymphoblastic leukemia. Future Oncol. **2019**, *15*(3), 257–269.

xxi. Krallinger, M.; Erhardt, R.A.A.; Valencia, A. Text-mining approaches in molecular biology and biomedicine. Drug Discov. Today **2005**, *10*(6), 439–445.

[xxii] Murai, J.; Huang, S.Y.N.; Das, B.B.; Renaud, A.; Zhang, Y.; Doroshow, J.H.; Ji, J.; Takeda, S.; Pommier, Y. Differential trapping of PARP1 and PARP2 by clinical PARP inhibitors. Cancer Res. **2012**, *72*(21), 5588.

[xxiii] Lovell, S.C.; Davis, I.W.; Arendall, W.B.; Bakker, P.I.W. De; Word, J.M.; Prisant, M.G.; Richardson, J.S.; Richardson, D.C. Structure validation by Cα geometry: φ,ψ and Cβ deviation. Proteins Struct. Funct. Genet. **2003**, *50*(3), 437–450.