

Phytochemicals as Potential DNA Polymerase β Inhibitors for Targeted Ovarian Cancer Therapy: An In-silico Approach

Anutosh Patra¹, Indranil Choudhuri², Prasenjit Paria³, Abhishek Samanta⁴, Kalyani Khanra², Anindita Chakraborty⁵ and Nandan Bhattacharyya^{2*}

¹Department of Physiology, Panskura Banamali College (Autonomous), P.O.-Panskura R.S., West Bengal, India.

²Department of Biotechnology, Panskura Banamali College (Autonomous), P.O.-Panskura R.S., West Bengal, India.

³Department of Biological Sciences, Indian Institute of Science Education and Research Kolkata (IISER K), Mohanpur, Nadia 741246, India.

⁴Department of Zoology, Panskura Banamali College, P.O., Panskura R.S., West Bengal, India.

⁵Department of Stress Biology, UGC DAE, Consortium for Scientific Research, Kolkata Centre, Sector III, LB-8, Bidhan Nagar, Kolkata, India.

<http://dx.doi.org/10.13005/bbra/3251>

(Received: 10 April 2024; accepted: 07 June 2024)

Ovarian cancer poses significant challenges due to limited treatment options and high mortality rates, necessitating innovative therapeutic strategies. Targeting DNA repair pathways, such as DNA polymerase β (Pol β), holds promise for improving treatment outcomes. This study aims to identify phytochemicals from the Super Natural database as natural inhibitors of Pol β activity to enhance ovarian cancer therapy efficacy, particularly when used in combination with damaging agents. Screening a library of 21,105 drug-like molecules alongside 800 compounds from the natural products collection (NatProd, a unique compound library) involved applying Lipinski's Rule of Five, the Golden Triangle rule, and Pfizer's rule. Following this, compounds predicted to exhibit carcinogenicity, toxicity, and mutagenicity were removed. The outcome of this rigorous screening process yielded 1,104 molecules eligible for structure-based virtual screening. Docking-based virtual screening using two servers was conducted on selected molecules, followed by computer simulations to assess their interaction dynamics and stability with Pol β . Molecular dynamics simulations further evaluated stability and interactions, considering energy, forces, and interaction scores. From these analyses, four promising Pol β inhibitors—SN00158342, SN00305418, SN00004251, and SN00341636—were identified, exhibiting favorable stability profiles, interactions. The binding energies for SN00158342, SN00305418, SN00004251, and SN00341636 were found to be -22.0327 ± 3.8493 , -15.9181 ± 4.5020 , -29.7465 ± 6.7833 and -27.3184 ± 5.1579 kcal/mol respectively. Utilizing these compounds alongside DNA-damaging agents presents a novel and potentially fruitful approach to improving ovarian cancer treatment outcomes. Overall, this study underscores the potential of phytochemicals as effective Pol β inhibitors, offering a promising avenue for enhancing ovarian cancer therapy effectiveness.

Keywords: DNA Polymerase beta; DNA-damaging agents; Inhibitors of DNA Polymerase beta; Molecular dynamic simulation; Ovarian epithelial carcinoma; Phytochemicals.

Human body undergoes continuous exposure to mutagens, which damage DNA. An array of repair systems fixes those damages¹. DNA

damage and repair have significant biological consequences on aging and diseases like cancers²⁻⁴. It was found that several cancers are linked to

*Corresponding author E-mail: bhattacharyya_nandan@rediffmail.com

the mutation of DNA repair proteins like DNA polymerase β (Pol β), which is up-regulated and/or mutated in several types of cancers, such as colon cancer, where the mutation rate reaches around 40%⁵. Mutations in Pol β are also highly linked to lung, breast, bladder, and esophageal cancers⁶. Khana *et al.* explored various Pol β mutations associated with ovarian cancers⁷⁻¹¹. Ovarian cancer is a significant health concern among Indian women, ranking as the third most common cancer and the eighth most prevalent overall in the country. Ovarian cancer accounts for 3.44% of all cancer cases and is also a leading cause of cancer-related deaths in Indian women, constituting 3.34% of all cancer-related deaths in the same year¹². Unfortunately, the prognosis for ovarian cancer is often poor due to late-stage diagnosis. Only 15% of cases are detected in Stage I, where the 5-year survival rate is a promising 94%¹³. The majority of cases, approximately 62%, are diagnosed in Stages III and IV, where the 5-year survival rate drops significantly to only 28%. These statistics highlight the urgent need for effective treatments. Currently, the standard approach for treating ovarian cancer involves platinum-based chemotherapy, which has remained unchanged for the past two decades¹⁴. This chemotherapy regimen, using drugs such as carboplatin or cisplatin along with paclitaxel, initially showed efficacy in most patients. However, approximately 80% of women experience relapse, primarily due to the development of platinum resistance¹³⁻¹⁶. To improve treatment efficacy and patient survival, it is crucial to understand the molecular mechanisms underlying this resistance. This challenge has prompted researchers to explore novel therapeutic strategies. The DNA damage response system is emerging as a popular inhibition target. A recent study ascribes excellent responses to chemotherapy by some cancer patients to defects in DNA repair¹⁷. Poly(ADP-ribose) polymerase (PARP) inhibitors are well known for treating BRCA1-deficient cancers¹⁸. Other DNA repair enzymes like glycosylases¹⁹, phosphodiesterases²⁰, and polymerases are often used as targets^{21,22}. Inhibitors of DNA repair enzymes work in conjunction with damaging agents like radiation to kill cells. Pol β is also an attractive target as it is the key enzyme of base excision repair pathways and also plays a

role in double-strand break repair via the alternative nonhomologous end-joining pathway²³⁻²⁶.

However, synthetic inhibitors of Pol β are often costly²⁷⁻³³. Therefore, in this study, our objective is to explore the possibilities of natural molecules as inhibitors of Pol β using an in-silico approach. The use of natural molecules offers a promising avenue for the development of cost-effective inhibitors for Pol β . These molecules will be identified through in-silico screening, which binds to the catalytic domain of Pol β . Such interactions could disrupt the DNA repair machinery, potentially working in conjunction with damaging agents to treat cancer.

SUBJECTS AND METHODS

Phytochemical database

The SuperNatural database (https://bioinf-applied.charite.de/supernatural_3/index.php) alongside 800 compounds from the natural products collection (NatProd, a unique compound library) were used in this study as a source of phytochemical structures. Supernatural is the largest repository of phytochemical information, containing information on over 449,058 phytochemical molecules, including their chemical structures, biological activities, and pharmacological properties³⁴.

Initial Screening of phytochemicals for drug-likeness

Drug-likeness of the phytochemicals was assessed using Lipinski's rule of five (RO5). RO5 is a set of criteria used to predict whether a compound is probably orally bioavailable and has good drug-like properties^{35,36}. Based on Lipinski's rule of five, the cut-off values were set as molecular weight \leq 500 Da, hydrogen bond donors \leq 5, hydrogen bond acceptors \leq 10, log P \leq 5, and the desired molecules were screened. Again, all the screened molecules were further screened for Absorption, Distribution, Metabolism and Excretion (ADME) properties using the ADMETLab2 server. (<https://admetmesh.scbdd.com/service/screening/index>) to predict the pharmacokinetics and toxicity properties. The molecules showing toxicity and mutagenic activity were removed from the library.

Protein model preparation

A three-dimensional (3D) structure of Pol β was prepared using homology modeling

as described previously³⁷. In brief, the nucleotide sequences of Pol α were translated in-silico using the ExPasy Translate tool (<https://www.expasy.org/>)³⁸. The translated amino acid sequences were used for template identification by employing BLASTP against the PDB (<http://www.rcsb.org/>), and the PDB structure 1bpx (human DNA polymerase α , 2.40 Å, X-ray diffraction) was selected as the template³⁹. The structure of DNA Pol α was constructed using Modeller v 9.11⁴⁰. The model was validated using the discrete optimized protein energy score (DOPE score), and the energy profile was plotted. Following model generation, primary structural analyses including target-template alignment, root-mean-square deviation (RMSD) value, and template modeling score (TM-score) were determined, and the sequence logo based on the conservation of the target-template alignment was generated using Chimera 1.16⁴¹.

Structure-Based Virtual Screening (VS)

For affinity-based screening for primary screening, an affinity-based approach was employed Using the EasyVS server with the background algorithm Vina(<http://biosig.unimelb.edu.au/easyvs/>, accessed on 25 September 2022)⁴². Three-dimensional structures of the DNA polymerase α proteins were used for further screening of target drug molecules. Chemical spaces of dimensions X: 3.672, Y: 12.247, and Z: 6.153 were subsequently prepared around the drugable site that was selected as the potential target site (Table 1). The best conformation with an affinity score lower than -6.5 and numerous hydrogen bonds e^{-7} were selected for further analysis. All the screen molecules were again re-docked using another server, DockThor(<https://www.dockthor.lncc.br/v2/>)⁴³.

Electrostatic complementary-based screening

Electrostatic complementary-based screening (EC) was performed using Flare Pro+ software to identify the types of interactions between the protein and the ligands. The screening process is necessary to narrow down the number of molecules to be tested in the next step⁴⁴. Following the initial screening, an evaluation copy of the software was requested, which is freely available for academic use for a period of one month. This step resulted in the shortlisting of approximately 30-40 molecules, representing approximately one-fifth of the original pool (Table 2).

Neural network for final screening drug molecules

Neural networking models, inspired by the intricacies of the human brain, are powerful tools for decoding complex biological phenomena^{45, 46}. In the context of separating phytochemical molecules, these models employ machine learning algorithms to categorize compounds based on their physicochemical properties. Critical parameters include van der Waals (vdW) energy, electrostatic energy, and hydrogen bonding strength, which influence molecular interactions and facilitate effective phytochemical separation. Van der Waals forces, which govern noncovalent interactions, maintain an optimal energy range of -47 to -35 kcal/mol. This balance ensures controlled and efficient phytochemical separation by harmonizing the attractive forces between the molecules. Electrostatic interactions, vital for stability, operate optimally within -25 to -15 kcal/mol, striking a delicate balance between attractive and repulsive forces to maintain specificity in separation. Hydrogen bonding strength, integral to molecular recognition, functions optimally between 8 and 12, aiding phytochemical isolation and contributing to the prevention of Pol α mutation and subsequent cancer development.

Molecular dynamics study

A molecular dynamics study was performed for 100 ns using the Amber ff19SB force field and the general AMBER force field (GAFF) to evaluate the binding stability of docking complexes^{47, 48}. A dodecahedral box of 12 Å was constructed around the protein-ligand complexes, and the box was dissolved in TIP3P water. The charges were neutralized by the addition of either Na⁺ or Cl⁻ ions at a molar concentration of 0.15 M. The systems were subjected to energy minimization at 300 K under a pressure of 1 bar. The systems were subsequently equilibrated for 20 ns, imposing positional restraints of 700 kJ/mol. The simulations were performed using the GPU-accelerated version of the OpenMM 7.6 engine and the 'Making it Rain' cloud-based molecular simulation notebook environment^{49, 50}. The trajectories generated during the MD simulations of the protein-ligand complexes were analyzed to calculate the values of RMSD, root-mean-square fluctuation (RMSF), radius of gyration (Rg), and hydrogen bonds (Fig. 3a, Fig. 3b, Fig. 3c).

Determination of Free Energies of Protein-Ligand Complexes

The binding free energies of the docked complexes were calculated using the mechanics/generalized Born surface area (MM/GBSA) approach⁵¹. The binding free energies (ΔG_{bind}) were calculated using the following equations^{52,53}.

$$\Delta G_{bind} = \Delta G_{complex} - (\Delta G_{receptor} + \Delta G_{ligand})$$

Where $\Delta G_{complex}$, $\Delta G_{receptor}$, and ΔG_{ligand} represent the free energy of the complex, receptor, and ligand, respectively in the following equations:

$$\Delta G = \Delta E_{gas} + \Delta G_{sol} - T\Delta S_{gas}$$

$$\Delta E_{gas} = \Delta E_{int} + \Delta E_{ELE} + \Delta E_{VDW}$$

$$\Delta G_{sol} = \Delta G_{GB} + \Delta G_{Surf}$$

Where ΔG represents free energy.

The energy in the gas phase (ΔE_{gas}) comprises the internal energy (ΔE_{int}), electrostatic interactions (ΔE_{ELE}), and van der Waals interactions (ΔE_{VDW}) energy terms. The solvation-free energy (ΔG_{sol}) comprises the polar energy (ΔG_{GB}) and nonpolar energy (ΔG_{Surf}) terms. $T\Delta S_{gas}$ represents the contribution of conformational entropy.

RESULTS

Screening of drug molecules based on pharmacokinetics

The SuperNatural library contains 449,058 natural compounds along with their structural

and physicochemical information³⁴. Based on Lipinski's rule of five, 21104 drug-like molecules were screened. These molecules, alongside 800 compounds from the natural products collection (NatProd, a unique compound library) molecules, were subjected to further screening by predicting their ADMET properties. Most of the compounds qualified for the Golden Triangle rule, Lipinski's rule of five, and Pfizer's rule. Molecules with predicted carcinogenicity, rat oral acute toxicity, AMES toxicity, and mutagenicity potentials were removed, and finally, 1,104 molecules were selected for structure-based virtual screening.

Structure prediction by homology modeling

In the model quality assessment, the predicted structures of DNA polymerase β from the amino acid sequences showed >96% favorable regions in the Ramachandran plots, with the QMEAN score <0.90 (>0.6), as determined by the MolProbity tool of the SWISS MODEL (Table 1).

Pocket Identification and Docking-Based Virtual Screening

Grid-based HECOMi finder (Ghecom) is a program for finding multi-scale pockets on protein surfaces using mathematical morphology^{34,54-55}. A total of fifteen pockets were identified in the structure of the DNA Pol β protein by the Ghecom algorithm. The largest pocket in the DNA Pol β protein had a volume of 3654.65 \AA^3 . The second largest pocket had a volume of 246.78 \AA^3 . The smallest pocket within the structure of the Pol β protein had a volume of 29.18 \AA^3 . In this study,

Table 1. Identification of Binding Pockets in Human DNA Pol β Protein and Grid Box Generation and total amino acid residues present in the binding pockets of DNA Pol β for Molecular Docking Studies

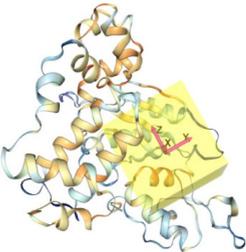
Protein Name	Coordinate of Docking Box	Structure	Amino Acid Residues
DNA polymerase β	X:3.672; Y:12.247; Z:6.153		P-151, K-167, K168, V-169, E-172, G-179, F-181, R-182, G-184, A-185, M-191, D-192, V-193, L-194, L-195, T-196, H-197, P-198, S-199, F-200, T-201, S-202, E-203, K-206, Q-207, P-208, K-209, L-210, L-211, H-212, Q-213, V-214, V-215, E-216, Q-217, L-218, D-226, T-227, K-230, G-231, E-232, T-233, K-234, F-235, G-237, V-238, L-241, P-242, R-253, I-255, D-256, I-257, R-258, L-259, I-260, P-261, K-262, D-263, Q-264, Y-265, Y-266, G-268, V-269, L-270, F-272, T-273, G-274

Table 2. Molecular Docking Parameters and Binding Characteristics

ID	Easy Vs Affinity	H bonds	Internal energy	Total energy	ThorDock Van der waal energy	Electrostatic energy	EC score	Amino acid involved in h bond	No of π - π bond	flare Aa in π - π bonds (bond distance)
SN00305418	-7.6	8	-39.107	-1.302	-9.785	-29.322	0.349	SER (A204), LYS (A206), LYS (A234), THR (A233), LEU (A259)	2	PHE(A200)(2.8), LYS(A234)(3.0)
SN00158342	-7.6	8	-35.5	-10.752	-4.305	-31.195	0.333	SER (A202), LYS (A206), GLU (A295)	2	LYS(A206)(2.6), TYR(A296)(3.1)
SN00004251	-7.1	12	-42.452	8.728	-16.872	-25.58	0.335	LYS (A206), GLU (A232), LEU (A259), SER (A204), THR (A201)	2	THY(A233) (3.7), PHE(A200)(3.0)
SN00341636	-7.4	9	-43.26	24.963	-20.089	-23.171	0.294	ASP (160), GLU (A329), VAL (A177), ARG (A152), ARG (A182)	0	

the second largest pocket was chosen which lies into the catalytic domain of the protein, were set as (X:3.672; Y:12.247; Z:6.153).

After the primary screening, approximately 1104 drug molecules were selected. The selected molecules were screened using docking-based virtual screening using the “EsayVS” server. Compounds with an affinity score of -6.5 with h-bond 7 protein-ligand interactions were selected. Totals of 135 molecules were finally selected against Pol β protein. These 135 molecules were further screened using the DockThor server.

Additionally, the selected molecules were further screened based on their EC using Flare v5.0.0, and compounds with EC scores > 0.25 were considered for further study.

Selection of drug molecules based on Neural Network

Based on 6 parameters, affinity score, internal energy, van der Waal force, EC scores, electrostatic interaction, and number of h

bonds, finally, seven molecules, SN00261400, SN00305418, SN00006989, SN00158342, SN00305418, SN00004251, and SN00341636 were selected from the pool of 135 compounds. As SN00006989 and SN00158342 were found to be structurally similar to the SN00261400, and SN00305418 was found to be structurally similar to the SN00305418, only four molecules, SN00305418, SN00158342, SN00004251, and SN00341636 were considered for the md simulation study (Table 3).

Analysis of Protein-Ligand Interactions

The intermolecular interactions between the selected molecules with the binding sites of the Pol β are analyzed from the best docked pose for each molecule (Fig. 1).

The SN00158342 molecule is primarily stabilized by four hydrogen bonds to Ser202, Glu295, and Lys206, with bond distances ranging from 1.9 to 2.3 Å. Additionally, it forms two van der Waals bonds with Lys206 and Tyr296, having bond distances of 2.6 Å and 3.1 Å, respectively.

Table 3. Selected drug molecules based on Neural Network

No.	1	2	3	4
Supernatural Id	SN00305418	SN00158342	SN00004251	SN00341636

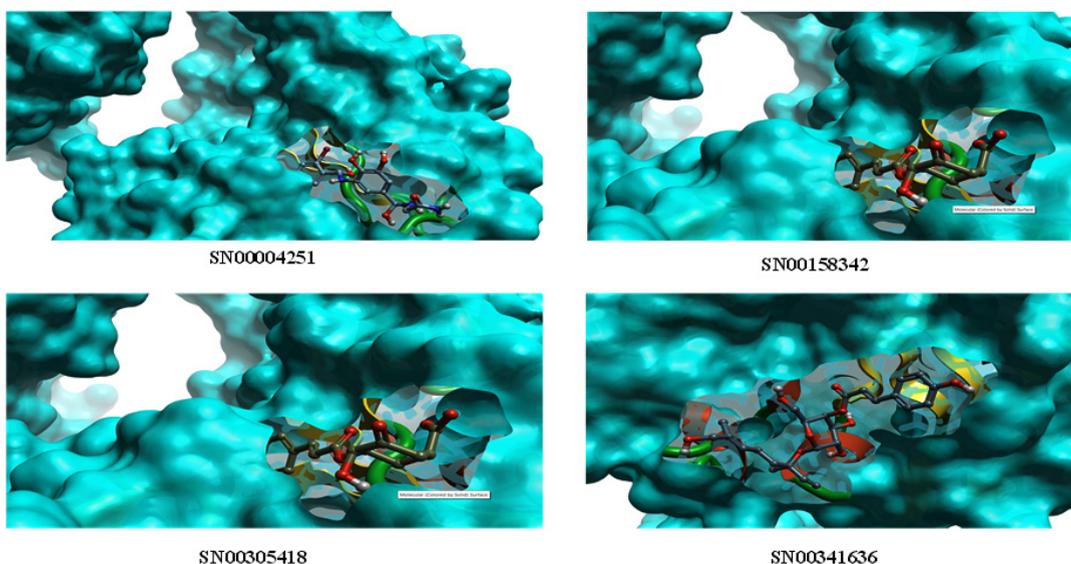


Fig. 1. Molecular docking of the compounds to the predicted binding pockets of the DNA polymerase beta enzyme

Similarly, SN00305418 is stabilized by six hydrogen bonds and two van der Waals bonds. The molecule forms hydrogen bonds with Ser204, Lys206, Thr233, Lys234, and Leu259, with bond distances ranging from 1.8 Å to 2.5 Å. It also creates two van der Waals bonds with Phe200 and Lys234, having bond distances of 2.8 Å and 3.0 Å, respectively.

The SN00004251 forms five hydrogen bonds with Thr201, Ser204, Glu232, and Leu259, with bond lengths ranging from 1.9 to 2.2 Å. Additionally, two van der Waals bonds are observed between the molecule and Phe200, and Thr233, with bond distances of 3.0 Å and 3.7 Å, respectively.

Unlike the previous three molecules, SN00341636 does not form van der Waals bonds

with the target protein. Instead, the ligand is stabilized by forming five hydrogen bonds with Arg152, Val177, Arg182, Glu329, and Pro330, with bond distances ranging from 1.9 Å to 2.4 Å (Fig.2).

MD Simulations of Protein-Ligand Complexes

MD simulation is a computational approach to predict and analyze the stabilities of protein-ligand complexes, and to study the atomic movements with respect to a macromolecule. The stabilities and behaviors of the protein-ligand complexes were analyzed in a dynamic environment based on the following parameters; (i) root-mean-square deviation (RMSD), (ii) root-mean-square fluctuation (RMSF), (iii) the radius of gyration (Rg), and (iv) molecular mechanics/generalized Born surface area (MM/GBSA) energy⁵²⁻⁵⁵.

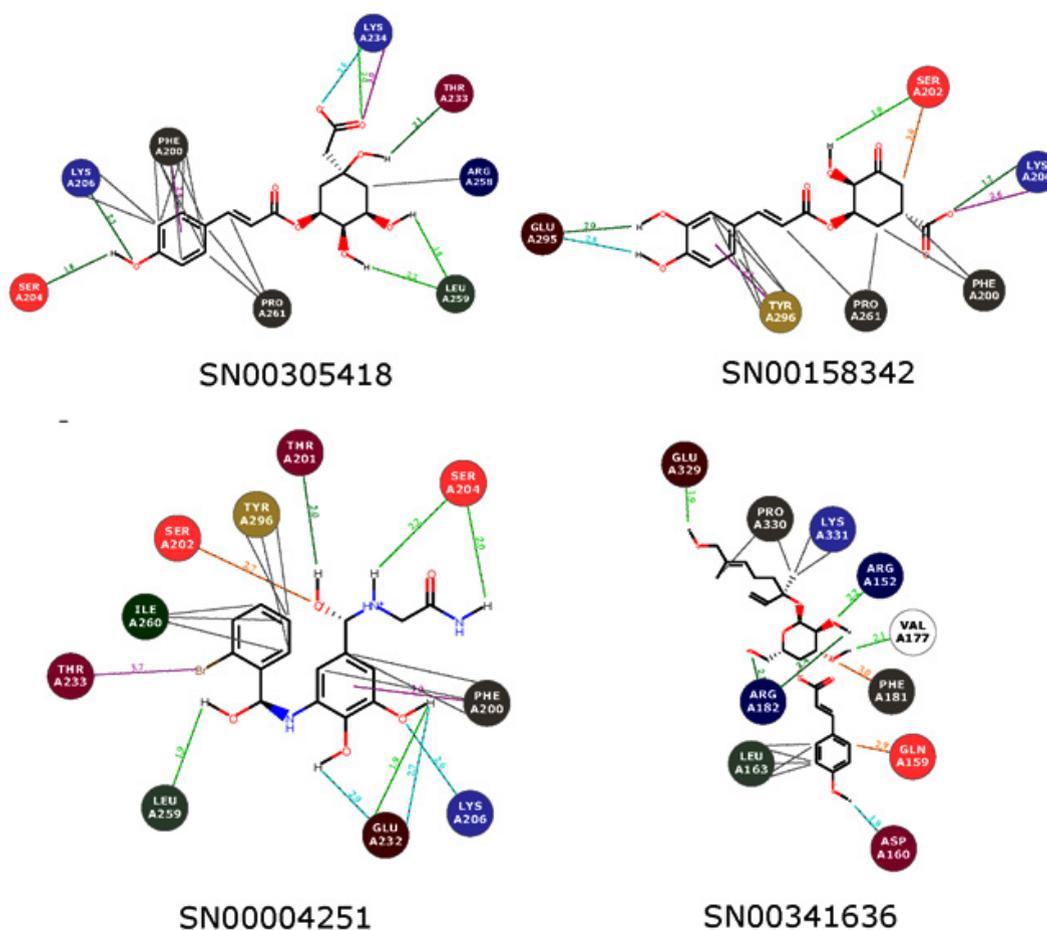


Fig. 2. Intramolecular interaction of the compounds with the predicted binding pocket of DNA Polymerase beta enzyme

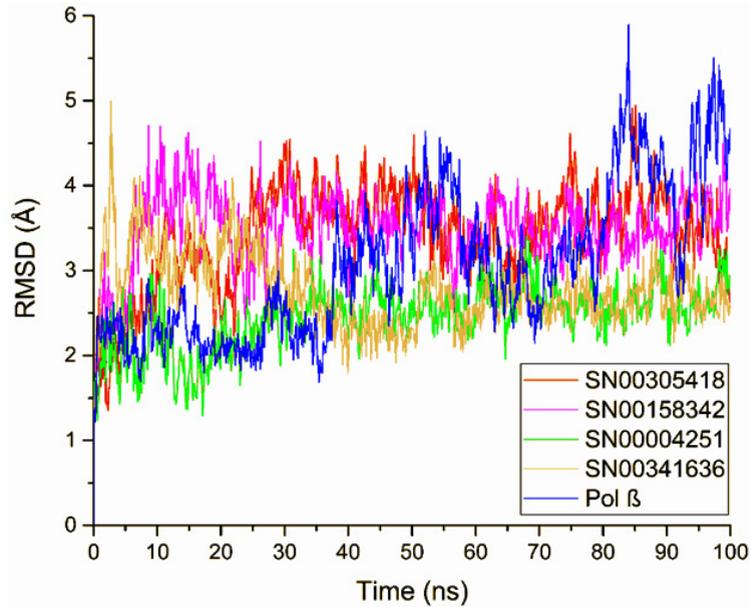


Fig. 3a. The line graph depicting the stability of various protein-ligand complexes, including DNA Polymerase β , during a 100 ns Molecular Dynamics (MD) simulation. RMSD values of the protein-ligand complexes fitted to the C α Backbone of target protein DNA Polymerase β . The X-axis represents the simulation duration in nanoseconds, while the Y-axis shows the Root Mean Square Deviation (RMSD) values in angstroms, measuring the deviation of each complex from a reference structure. Each colored line corresponds to a different complex, with the degree of line fluctuation indicating the stability of the complex: less fluctuation signifies a more stable complex, while more fluctuation suggests less stability. This graph assessing complex stability during the simulation, an important factor in effective drug design and protein engineering and identify a complex with minimal RMSD fluctuations, indicating that it maintains consistency with the target protein’s structure throughout the simulation.

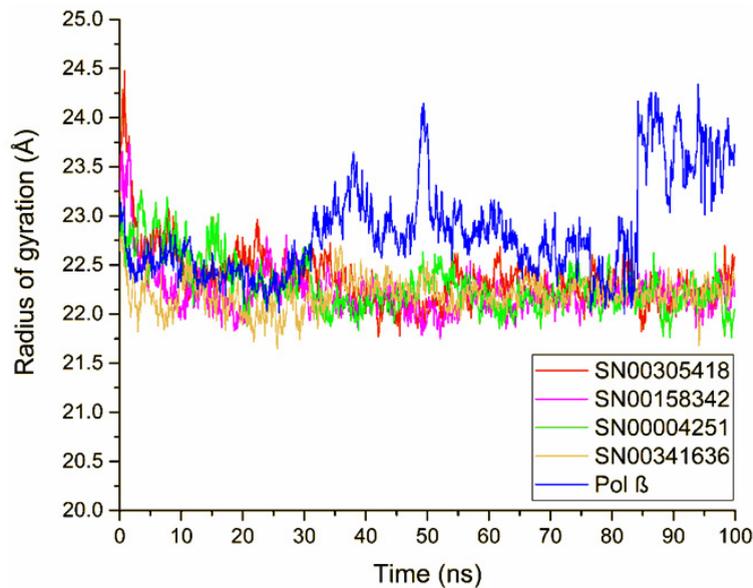


Fig. 3b. Analysis of the Rg values of the protein-ligand complexes. The x-axis depicts the duration of simulation, while the y-axis represents the deviations in Rg

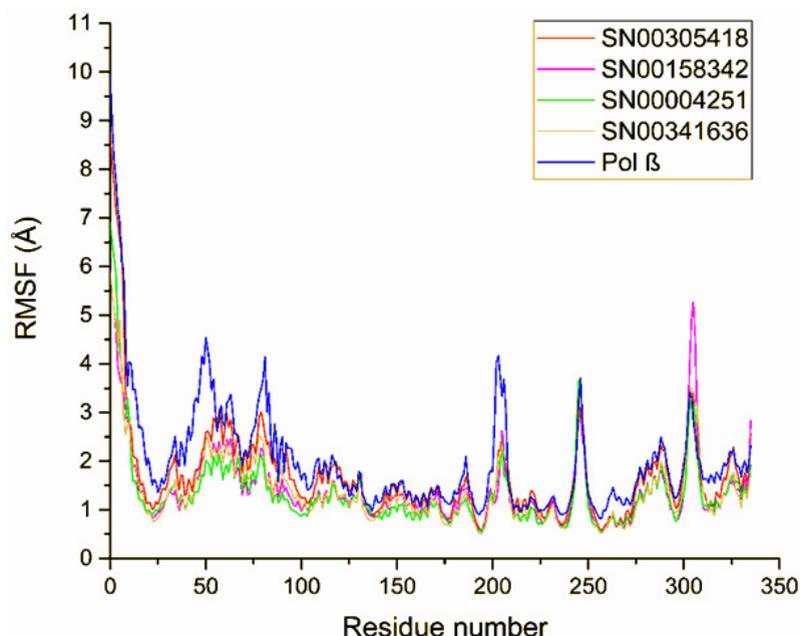


Fig. 3c. RMSF values of the C α backbone of Target protein DNA Polymerase β . The x-axis depicts the total number of residues, while the y-axis represents the RMSF in Å

RMSD Values of the C α Backbone of Target Proteins

The RMSD values of the four protein-ligand complexes were visualized by plotting RMSD values against time, which helps to understand the structural stability and integrity of the complexes. The trajectory analysis revealed that all the four compounds complex with Pol β remain stable throughout the 100 ns simulation. The average RMSD values of the 4 molecules selected against Pol β ranged from 2.449 to 4.781 Å. Compounds SN00004251, SN00341636 possess high stabilities during the simulation, with an RMSD fluctuation of 0.378 Å and 0.444 Å respectively.

The SN00305418, SN00158342 molecules found to be less stable during the simulation, with an RMSD fluctuation of 0.837 Å and 0.602 Å respectively. The overall RMSD fluctuation of protein is more in comparison with other four protein ligand complexes (Fig. 3a).

Rg Values of the C α Backbone of Target Proteins

The compactness of the protein-ligand complexes during the simulation was determined by measuring the values of Rg. The average values of Rg for the 4 compounds complexed with Pol

β ranged from 22.177 to 22.636 Å. Rg of Pol β fluctuates more in comparison with other four protein ligand complexes. Suddenly the Rg value of Pol β increase after 80 ns of simulation (Fig. 3b).

RMSF Values of the C α Backbone of Target Proteins

The average atomic mobility of the protein backbone during the MD simulations was measured using the values of RMSF. The average RMSF values of the 4 molecules complexed with Pol β ranged from 1.703 to 2.020 Å. Further analysis revealed that residues 203, 246 and 306 of Pol β underwent fluctuations.

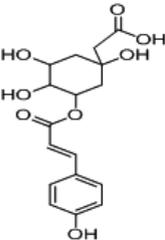
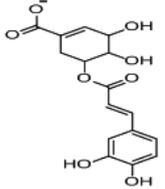
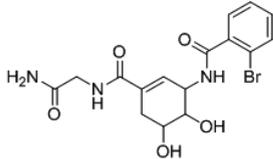
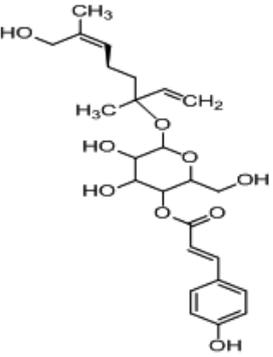
However, only two amino acid residues 246, and 306 underwent fluctuations when complexed with SN00305418, SN00158342, SN00004251 and SN00341636. Residues Arg152, Asp160, Val 177, Thr201, Ser204, Lys206, Lys234, Thr233, Leu259, Glu295 and Glu329 which were primarily involved in the formations of ligand-protein hydrogen bonds with four different chemical compounds, underwent minimal fluctuations. Similarly, the amino acid residues that mediated the formations of stacking interactions remained stable during the simulation (Fig. 3c).

Determination of Binding Free Energies of Protein-Ligand Complexes

The binding free energy represents the sum total of all the interaction energies, including the van der Waals energy, polar solvation energy, electrostatic energy, and solvent accessible surface

area SASA energy. The binding free energies of all the complexes were estimated using the MM/GBSA approach. The binding free energies of the four compounds complexed with Pol α ranged from -15.9181 to -29.7465 kcal/mol, of which SN00004251 (-29.7465 \pm 6.7833) had the lowest free energy of binding (Table 4).

Table 4. The two-dimensional structures and binding energies of four compounds against the target proteins

Molecular ID	Compound Name	Chemical Structure	Binding Energy
SN00305418	(E)-2-(1,3,4-trihydroxy-5-((3-(4-hydroxyphenyl)acryloyl)oxy)cyclohexyl)acetic acid		-15.9181 \pm 4.5020
SN00158342	(E)-5-((3-(3,4-dihydroxyphenyl)acryloyl)oxy)-3,4-dihydroxycyclohex-1-enecarboxylate		-22.0327 \pm 3.8493
SN00004251	N-(3-((2-amino-2-oxoethyl)carbamoyl)-5,6-dihydroxycyclohex-2-en-1-yl)-2-bromobenzamide		7465 \pm 6.7833
SN00341636	(E)-4,5-dihydroxy-6-(((Z)-8-hydroxy-3,7-dimethylocta-1,6-dien-3-yl)oxy)-2-(hydroxymethyl)tetrahydro-2H-pyran-3-yl 3-(4-hydroxyphenyl)acrylate		3184 \pm 5.1579

DISCUSSION

To find suitable inhibitors of Pol δ , we start screening with a large number of natural molecules. The methodology employed involved a multi-step approach, starting with the screening of a vast library of natural compounds, followed by refinement based on pharmacokinetics, ADMET properties, and virtual screening. The meticulous selection process, involving the removal of compounds with potential toxicity and mutagenicity, led to a focused set of 1,104 molecules for further structure-based virtual screening⁵⁶. The homology modelling of Pol δ demonstrated high-quality predicted structures, ensuring the reliability of subsequent analyses. The identification of key pockets on the protein surface using the Ghecom algorithm, especially the selection of the second-largest pocket within the catalytic domain, showcased a thoughtful approach to target selection. The subsequent virtual screening steps, involving two different servers ("EsayVS" and "DockThor"), as well as evaluation based on EC scores using Flare v5.0.0, highlighted the rigor applied to filter and prioritize potential drug candidates. The final selection, guided by a neural network considering six critical parameters, resulted in the identification of seven molecules for further investigation.

—A Molecular Dynamics (MD) study for protein-ligand interaction involves simulating the dynamic behavior of molecules over time, particularly focusing on how a protein and a ligand interact with each other at the atomic level. MD simulations rely on mathematical models called force fields to describe the interactions between atoms within the molecules. Force fields include terms for bonded interactions (bonds, angles, and dihedrals) and non-bonded interactions like van der Waals forces and electrostatic interactions. Throughout the simulation, various parameters can be monitored and analyzed to understand the protein-ligand interaction dynamics. These include the root-mean-square deviation (RMSD) to measure structural changes, root-mean-square fluctuation (RMSF) to assess flexibility, hydrogen bonding patterns, solvent accessible surface area (SASA), and others. The findings from the MD simulation are interpreted to gain insights into the molecular mechanisms underlying the protein-

ligand interaction. This could involve identifying key protein-ligand interactions, understanding conformational changes induced by ligand binding, or proposing strategies for rational drug design, which provide a powerful tool for studying protein-ligand interactions at an atomic level, offering insights that complement experimental techniques and guiding the design of novel therapeutics⁵⁷⁻⁵⁸. In this study, MD simulation was employed to simulate Pol δ -phytochemical interactions over a period of 100 ns. The root-mean-square deviation (RMSD), root-mean-square fluctuation (RMSF), radius of gyration (Rg), and molecular mechanics/generalized Born surface area (MM/GBSA) energy were calculated for each Pol δ -phytochemical complexes in order to determine the most stable binding. After considerations of all parameters, four phytochemicals SN00305418, SN00158342, SN00004251 and SN00341636 were considered as most suitable molecules for inhibiting Pol δ by binding its DNA binding domain.

DNA polymerase δ (Pol δ) plays a vital role in base excision repair in the nucleus and mitochondria⁵⁹. Pol δ also contributes to double strand break repair via an alternative nonhomologous end joining pathway²⁶. DNA Pol δ has two domains. A small 8 KDa domain involves DNA binding and dRP lyase activity. The large 31 KDa domain has dNTP selection and catalytic activity³⁸. DNA repair pathways are often considered as a target^{60,15,61}. Recent advances in research explore the opportunities to target Pol δ along with DNA damaging agents^{62,63,64}. In our previous study, we established a variant form of Pol δ , which has a deletion in its catalytic domain and lacks polymerase activity, making it more sensitive to gamma radiations³⁷. This variant retains its DNA binding domain intact. When a single-strand break or damaged base occurs, Pol δ binds to the DNA. However, as the variant lacks catalytic activity, it cannot repair the damage, leading to subsequent cell death via an apoptotic pathway. Because Pol δ has already bound to the DNA, other repair proteins like pol α or pol β cannot function. Similar results were found in another study⁶⁵. Bhattachayya and Banerjee found a variant of Pol δ with a deletion of 208–236 amino acid act as a dominant negative with wild type Pol δ ⁶⁶. The variant lost its catalytic activity, whereas its DNA binding activity domain is remains intact. Wang et al. found that silencing

DNA polymerase α enhances the radio-therapeutic sensitivity of human esophageal carcinoma cell lines³³. These results suggest that the lack of catalytic activity of DNA Pol α could lead a cell to undergo apoptosis following any single-strand break or damaged base. Therefore, in this study, we focus on a pocket within the catalytic domain of Pol α , specifically encompassing amino acids 151 to 274.

—DNA damage is a common occurrence arising from various factors such as exposure to radiation, chemicals, and errors during DNA replication. Cells have evolved intricate mechanisms to detect and repair DNA damage, ensuring the maintenance of genomic stability⁶⁷. The types of DNA damage encompass chemical modifications, physical lesions, and replication errors⁶⁸. Detection of DNA damage involves the action of DNA damage sensors like ATM, ATR, and DNA-PK⁶⁹. Cells employ diverse DNA repair mechanisms, including direct reversal, base excision repair (BER), nucleotide excision repair (NER), mismatch repair (MMR), homologous recombination (HR), non-homologous end joining (NHEJ), and translesion synthesis (TLS). Each repair pathway targets specific types of DNA lesions, ensuring the fidelity of the genetic material. Additionally, cell cycle checkpoints, governed by p53, play a crucial role in halting cell division to allow for repair or, if damage is irreparable, initiating programmed cell death through apoptosis⁶⁹. Post-repair surveillance mechanisms, orchestrated by p53, further contribute to the overall genomic integrity by regulating cell cycle progression and apoptosis based on the fidelity of the DNA repair processes. Understanding these DNA damage and repair mechanisms is pivotal for unravelling the complexity of cellular responses to genomic insults and holds significance for fields such as cancer research and therapeutic development⁷⁰.

Overall, the study's systematic approach, combining computational screening, molecular modeling, virtual screening, and MD simulations, contributes valuable information for the potential development of drugs targeting Pol α . The identification of specific compounds with favorable interactions and low binding free energies sets the stage for further experimental validation and optimization of these candidates for therapeutic purposes. The thorough analysis presented in

this study lays a solid foundation for future drug discovery efforts targeting DNA polymerase α .

CONCLUSION

Four phytochemicals were identified against Pol α targets in its catalytic domain. These compounds had high binding affinities and low free binding energies, as indicated by the results of extensive in-silico analyses. According to the in-silico study, The screened molecules demonstrate an ability to bind to the catalytic domain of Pol α . These molecules have shown promise for use as pharmaceuticals, potentially in conjunction with damaging agents, to treat cancer. The computational analyses conducted in this study has its own limitation, and therefore it will undergo experimental validation through both in vitro and in vivo studies in the future.

ACKNOWLEDGMENTS

The author Anutosh Patra expresses his thanks to UGC DAE, Consortium for Scientific Research, Kolkata Centre (UGC DAE CSR KC) and PanskuraBanamaliCollege (Autonomous). The author would like to acknowledge to Cresset group for their generous support in providing us with a Flare Pro+ Evaluation Copy.

Conflict of Interest

All the authors declare no conflict of Interest.

Funding Sources

This study was not supported by any funding agency. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Authors' contributions

AP conceptualized, designed the study, conducted the experiments, and prepared the manuscript. IC designed the study and prepared the manuscript. PP experiments the study. AS performed the artificial neural network implementation. KK revised the manuscript. AC conceptualized, designed and supervision. NB conceptualized, designed and supervised the study, and prepared the final manuscript.

Ethical approval

This is an in silico study and does not require ethical approval.

REFERENCES

1. Sancar A, Lindsey-Boltz LA, Ünsal-Kaçmaz K, Linn S. Molecular mechanisms of mammalian DNA repair and the DNA damage checkpoints. *Annual review of biochemistry*. 2004 Jul;73(1):39-85.
2. Nelson BC, Dizdaroglu M. Implications of DNA damage and DNA repair on human diseases. *Mutagenesis*. 2020 Jan;35(1):1-3.
3. Bernstein C, Prasad AR, Nfonso V, Bernstein H. DNA damage, DNA repair and cancer. Rijeka, Croatia: InTech; 2013 May 22.
4. Maynard S, Fang EF, Scheibye-Knudsen M, Croteau DL, Bohr VA. DNA damage, DNA repair, aging, and neurodegeneration. *Cold Spring Harbor perspectives in medicine*. 2015 Oct 1;5(10):a025130.
5. Donigan KA, Sun KW, Nemecek AA, Murphy DL, Cong X, Northrup V, Zelterman D, Sweasy JB. Human POLB gene is mutated in high percentage of colorectal tumors. *Journal of Biological Chemistry*. 2012 Jul 6;287(28):23830-9.
6. Starcevic D, Dalal S, Sweasy JB. Is there a link between DNA polymerase beta and cancer?. *Cell cycle*. 2004 Aug 21;3(8):996-9.
7. Khanra K, Panda K, Bhattacharya C, Mitra AK, Sarkar R, Banerjee S, Bhattacharyya N. Association between newly identified variant form of DNA polymerase beta Å208–304 and ovarian cancer. *Cancer Biomarkers*. 2012 Jan 1;11(4):155-60.
8. Khanra K, Bhattacharya C, Bhattacharyya N. Association of a newly identified variant of DNA polymerase beta (polá 63-123, 208-304) with the risk factor of ovarian carcinoma in India. *Asian Pacific Journal of Cancer Prevention*. 2012;13(5):1999-2002.
9. Khanra K, Chakraborty A, Bhattacharyya N. HeLa cells containing a truncated form of DNA polymerase beta are more sensitized to alkylating agents than to agents inducing oxidative stress. *Asian Pacific Journal of Cancer Prevention*. 2016;16(18):8177-86.
10. Khanra K, Panda K, Mitra AK, Sarkar R, Bhattacharya C, Bhattacharyya N. Exon 8-9 mutations of DNA polymerase ã in ovarian carcinoma patients from haldia, India. *Asian Pacific journal of cancer prevention*. 2012;13(8):4183-6.
11. Khanra K, Panda K, Bhattacharya C, Mitra AK, Sarkar R, Bhattacharyya N. Association of two polymorphisms of DNA polymerase beta in exon-9 and exon-11 with ovarian carcinoma in India. *Asian Pacific journal of cancer prevention*. 2012;13(4):1321-4.
12. Soerjomataram I, Lortet-Tieulent J, Parkin DM, Ferlay J, Mathers C, Forman D, Bray F. Global burden of cancer in 2008: a systematic analysis of disability-adjusted life-years in 12 world regions. *The Lancet*. 2012 Nov 24;380(9856):1840-50.
13. Danforth KN, Im TM, Whitlock EP. Addendum to screening for ovarian cancer: Evidence update for the US Preventive services task force reaffirmation recommendation statement. *AHRQ publication*. 2012 Apr(12-05165).
14. Damia G, Brogginì M. Platinum resistance in ovarian cancer: role of DNA repair. *Cancers*. 2019 Jan 20;11(1):119.
15. Gavande NS, VanderVere-Carozza PS, Hinshaw HD, Jalal SI, Sears CR, Pawelczak KS, Turchi JJ. DNA repair targeted therapy: The past or future of cancer treatment?. *Pharmacology & therapeutics*. 2016 Apr 1;160:65-83.
16. López-Camarillo C, Rincón DG, Ruiz-García E, Astudillo-de la Vega H, Marchat LA. DNA repair proteins as therapeutic targets in ovarian cancer. *Current Protein and Peptide Science*. 2019 Apr 1;20(4):316-23.
17. Wheeler DA, Takebe N, Hinoue T, Hoadley KA, Cardenas MF, Hamilton AM, Laird PW, Wang L, Johnson A, Dewal N, Miller V. Molecular features of cancers exhibiting exceptional responses to treatment. *Cancer cell*. 2021 Jan 11;39(1):38-53.
18. Lord CJ, Ashworth A. PARP inhibitors: Synthetic lethality in the clinic. *Science*. 2017 Mar 17;355(6330):1152-8.
19. Tahara YK, Auld D, Ji D, Beharry AA, Kietrys AM, Wilson DL, Jimenez M, King D, Nguyen Z, Kool ET. Potent and selective inhibitors of 8-oxoguanine DNA glycosylase. *Journal of the American Chemical Society*. 2018 Feb 14;140(6):2105-14.
20. Dorjsuren D, Kim D, Vyjayanti VN, Maloney DJ, Jadhav A, Wilson III DM, Simeonov A. Diverse small molecule inhibitors of human apurinic/apyrimidinic endonuclease APE1 identified from a screen of a large public collection.
21. Strittmatter T, Brockmann A, Pott M, Hantusch A, Brunner T, Marx A. Expanding the scope of human DNA polymerase ã and ã inhibitors. *ACS chemical biology*. 2014 Jan 17;9(1):282-90.
22. Gowda AP, Suo Z, Spratt TE. Honokiol inhibits DNA polymerases ã and ã and increases bleomycin sensitivity of human cancer cells. *Chemical research in toxicology*. 2017 Feb 20;30(2):715-25.
23. M Goellner E, Svilar D, H Almeida K, W Sobol R. Targeting DNA polymerase ss for therapeutic intervention. *Current molecular pharmacology*. 2012 Jan 1;5(1):68-87.

24. Jaiswal AS, Panda H, Law BK, Sharma J, Jani J, Hromas R, Narayan S. NSC666715 and its analogs inhibit strand-displacement activity of DNA polymerase α and potentiate temozolomide-induced DNA damage, senescence and apoptosis in colorectal cancer cells. *PLoS one*. 2015 May 1;10(5):e0123808.
25. Krokan HE, Bjørås M. Base excision repair. *Cold Spring Harbor perspectives in biology*. 2013 Apr 1;5(4):a012583.
26. Ray S, Breuer G, DeVeaux M, Zeltermann D, Bindra R, Sweasy JB. DNA polymerase beta participates in DNA End-joining. *Nucleic acids research*. 2018 Jan 9;46(1):242-55.
27. Yuhas SC, Lavery DJ, Lee H, Majumdar A, Greenberg MM. Selective inhibition of DNA polymerase α by a covalent inhibitor. *Journal of the American Chemical Society*. 2021 May 20;143(21):8099-107.
28. Yuhas SC, Majumdar A, Greenberg MM. Protein Domain Specific Covalent Inhibition of Human DNA Polymerase α . *ChemBioChem*. 2021 Aug 17;22(16):2619-23.
29. Chalmers AJ, Lakshman M, Chan N, Bristow RG. Poly (ADP-ribose) polymerase inhibition as a model for synthetic lethality in developing radiation oncology targets. *In Seminars in radiation oncology 2010 Oct 1 (Vol. 20, No. 4, pp. 274-281)*. WB Saunders.
30. Barakat KH, Gajewski MM, Tuszynski JA. DNA polymerase beta (Pol β) inhibitors: A comprehensive overview. *Drug discovery today*. 2012 Aug 1;17(15-16):913-20.
31. Patel P, Parmar I, Jha A, Shastri S. In-silico screening of phytochemicals as an approach against poly-ADP-ribose polymerase-1 in ovarian cancer. *Journal of Applied Pharmaceutical Science*. 2024 Jan 4;14(1):280-5.
32. Ali R, Alblihy A, Miligy IM, Alabdullah ML, Alsaleem M, Toss MS, Algethami M, Abdel-Fatah T, Moseley P, Chan S, Mongan NP. Molecular disruption of DNA polymerase α for platinum sensitisation and synthetic lethality in epithelial ovarian cancers. *Oncogene*. 2021 Apr 8;40(14):2496-508.
33. Wang Y, Chen X, Hu X, Zhang R, Du Y, Zang W, Dong Z, Zhao G. Enhancement of silencing DNA polymerase α on the radiotherapeutic sensitivity of human esophageal carcinoma cell lines. *Tumor Biology*. 2014 Oct;35:10067-74.
34. Gallo K, Kemmler E, Goede A, Becker F, Dunkel M, Preissner R, Banerjee P. SuperNatural 3.0—a database of natural products and natural product-based derivatives. *Nucleic Acids Research*. 2023 Jan 6;51(D1):D654-9.
35. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced drug delivery reviews*. 2012 Dec 1;64:4-17.
36. Lipinski CA. Lead-and drug-like compounds: the rule-of-five revolution. *Drug discovery today: Technologies*. 2004 Dec 1;1(4):337-41.
37. Patra A, Nag A, Chakraborty A, Bhattacharyya N. PA1 cells containing a truncated DNA polymerase α protein are more sensitive to gamma radiation. *Radiation Oncology Journal*. 2022 Mar;40(1):66.
38. Beard WA, Wilson SH. Structure and mechanism of DNA polymerase α . *Chemical reviews*. 2006 Feb 8;106(2):361-82.
39. Uddin R, Masood F, Azam SS, Wadood A. Identification of putative non-host essential genes and novel drug targets against *Acinetobacter baumannii* by in silico comparative genome analysis. *Microbial pathogenesis*. 2019 Mar 1;128:28-35.
40. Muhammed MT, Aki-Yalcin E. Up-to-Date Developments in Homology Modeling. *Applied Computer-Aided Drug Design: Models and Methods*. 2023 Dec 8:116.
41. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry*. 2004 Oct;25(13):1605-12.
42. Pires DE, Veloso WN, Myung Y, Rodrigues CH, Silk M, Rezende PM, Silva F, Xavier JS, Velloso JP, da Silveira CH, Ascher DB. EasyVS: a user-friendly web-based tool for molecule library selection and structure-based virtual screening. *Bioinformatics*. 2020 Jul 15;36(14):4200-2.
43. Guedes IA, Costa LS, Dos Santos KB, Karl AL, Rocha GK, Teixeira IM, Galheigo MM, Medeiros V, Krempser E, Custódio FL, Barbosa HJ. Drug design and repurposing with DockThor-VS web server focusing on SARS-CoV-2 therapeutic targets and their non-synonym variants. *Scientific Reports*. 2021 Mar 10;11(1):5543.
44. Taylor RD, Jewsbury PJ, Essex JW. FDS: flexible ligand and receptor docking with a continuum solvent model and soft core energy function. *Journal of computational chemistry*. 2003 Oct;24(13):1637-56.
45. Gurney K. Neural networks for perceptual processing: from simulation tools to theories. *Philosophical Transactions of the Royal Society B: Biological Sciences*. 2007 Mar 29;362(1479):339-53.
46. Abbott LF. Decoding neuronal firing and modelling neural networks. *Quarterly reviews of biophysics*. 1994 Aug;27(3):291-331.

47. Tian C, Kasavajhala K, Belfon KA, Raguette L, Huang H, Miguez AN, Bickel J, Wang Y, Pincay J, Wu Q, Simmerling C. ff19SB: Amino-acid-specific protein backbone parameters trained against quantum mechanics energy surfaces in solution. *Journal of chemical theory and computation*. 2019 Nov 12;16(1):528-52.
48. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. *Journal of computational chemistry*. 2004 Jul 15;25(9):1157-74.
49. Eastman P, Swails J, Chodera JD, McGibbon RT, Zhao Y, Beauchamp KA, Wang LP, Simmonett AC, Harrigan MP, Stern CD, Wiewiora RP. OpenMM 7: Rapid development of high performance algorithms for molecular dynamics. *PLoS computational biology*. 2017 Jul 26;13(7):e1005659.
50. Arantes PR, Polêto MD, Pedebos C, Ligabue-Braun R. Making it rain: cloud-based molecular simulations for everyone. *Journal of Chemical Information and Modeling*. 2021 Oct 1;61(10):4852-6.
51. Onufriev A, Bashford D, Case DA. Exploring protein native states and large scale conformational changes with a modified generalized born model. *Proteins: Structure, Function, and Bioinformatics*. 2004 May 1;55(2):383-94.
52. Tuccinardi T. What is the current value of MM/PBSA and MM/GBSA methods in drug discovery?. *Expert opinion on drug discovery*. 2021 Nov 2;16(11):1233-7.
53. Forouzesh N, Mishra N. An effective MM/GBSA protocol for absolute binding free energy calculations: A case study on SARS-CoV-2 spike protein and the human ACE2 receptor. *Molecules*. 2021 Apr 20;26(8):2383.
54. Mutlu O. Molecular modeling, structural analysis and identification of ligand binding sites of trypanothione reductase from *Leishmania mexicana*. *Journal of Vector Borne Diseases*. 2013 Jan 1;50(1):38-44.
55. Shukla R, Tripathi T. Molecular dynamics simulation of protein and protein–ligand complexes. *Computer-aided drug design*. 2020:133-61.
56. Paria P, Tassanakajon A. Identification of potential druggable targets and structure-based virtual screening for drug-like molecules against the shrimp pathogen *Enterocytozoon hepatopenaei*. *International Journal of Molecular Sciences*. 2023 Jan 11;24(2):1412.
57. Araki M, Matsumoto S, Bekker GJ, Isaka Y, Sagae Y, Kamiya N, Okuno Y. Exploring ligand binding pathways on proteins using hypersound-accelerated molecular dynamics. *Nature communications*. 2021 May 14;12(1):2793.
58. Sasidharan S, Gosu V, Tripathi T, Saudagar P. Molecular Dynamics Simulation to Study Protein Conformation and Ligand Interaction. In *Protein Folding Dynamics and Stability: Experimental and Computational Methods 2023* May 28 (pp. 107-127). Singapore: Springer Nature Singapore.
59. Prasad R, Çađlayan M, Dai DP, Nadalutti CA, Zhao ML, Gassman NR, Janoshazi AK, Stefanick DF, Horton JK, Krasich R, Longley MJ. DNA polymerase α : A missing link of the base excision repair machinery in mammalian mitochondria. *DNA repair*. 2017 Dec 1;60:77-88.
60. Wang M, Chen S, Ao D. Targeting DNA repair pathway in cancer: Mechanisms and clinical application. *MedComm*. 2021 Dec;2(4):654-91.
61. Huang R, Zhou PK. DNA damage repair: historical perspectives, mechanistic pathways and clinical translation for targeted cancer therapy. *Signal transduction and targeted therapy*. 2021 Jul 9;6(1):254.
62. Saha P, Mandal T, Talukdar AD, Kumar D, Kumar S, Tripathi PP, Wang QE, Srivastava AK. DNA polymerase ϵ : A potential pharmacological target for cancer therapy. *Journal of Cellular Physiology*. 2021 Jun;236(6):4106-20.
63. Biau J, Chautard E, Verrelle P, Dutreix M. Altering DNA repair to improve radiation therapy: specific and multiple pathway targeting. *Frontiers in oncology*. 2019 Oct 10;9:1009.
64. Jaiswal AS, Banerjee S, Aneja R, Sarkar FH, Ostrov DA, Narayan S. DNA polymerase α as a novel target for chemotherapeutic intervention of colorectal cancer. *PLoS One*. 2011 Feb 2;6(2):e16691.
65. Neijenhuis S, Verwijs-Janssen M, van den Broek LJ, Begg AC, Vens C. Targeted radiosensitization of cells expressing truncated DNA polymerase α . *Cancer research*. 2010 Nov 1;70(21):8706-14.
66. Bhattacharyya N, Banerjee S. A variant of DNA polymerase α acts as a dominant negative mutant. *Proceedings of the National Academy of Sciences*. 1997 Sep 16;94(19):10324-9.
67. Alhmoud JF, Woolley JF, Al Moustafa AE, Mallei MI. DNA damage/repair management in cancers. *Advances in Medical Biochemistry, Genomics, Physiology, and Pathology*. 2021 Dec 22:309-39.
68. Carusillo A, Mussolino C. DNA damage: from threat to treatment. *Cells*. 2020 Jul 10;9(7):1665.
69. Visconti R, Della Monica R, Grieco D. Cell cycle checkpoint in cancer: a therapeutically targetable double-edged sword. *Journal of Experimental & Clinical Cancer Research*. 2016 Dec;35:1-8.
70. Molinaro C, Martoriati A, Cailliau K. Proteins from the DNA damage response: Regulation, dysfunction, and anticancer strategies. *Cancers*. 2021 Jul 29;13(15):3819.