





Virtual screening and multilevel precision-based prioritisation of natural inhibitors targeting the ATPase domain of human DNA topoisomerase II alpha

Barathi Lenin, Sharada Ramasubramanyan, Umashankar Vetrivel & Srujana Chitipothu


To cite this article: Barathi Lenin, Sharada Ramasubramanyan, Umashankar Vetrivel & Srujana Chitipothu (2023) Virtual screening and multilevel precision-based prioritisation of natural inhibitors targeting the ATPase domain of human DNA topoisomerase II alpha, Journal of Biomolecular Structure and Dynamics, 41:24, 15177-15195, DOI: [10.1080/07391102.2023.2187234](https://doi.org/10.1080/07391102.2023.2187234)

To link to this article: <https://doi.org/10.1080/07391102.2023.2187234>

 View supplementary material [↗](#)

 Published online: 10 Mar 2023.

 Submit your article to this journal [↗](#)

 Article views: 280

 View related articles [↗](#)

 View Crossmark data [↗](#)

 Citing articles: 1 View citing articles [↗](#)



Virtual screening and multilevel precision-based prioritisation of natural inhibitors targeting the ATPase domain of human DNA topoisomerase II alpha

Barathi Lenin^a, Sharada Ramasubramanyan^b, Umashankar Vetrivel^{a,c,a,#} and Srujana Chitipothu^{a,d}

^aCentre for Bioinformatics, Kamalnayan Bajaj Institute for Research in Vision and Ophthalmology, Vision Research Foundation, Chennai, Tamil Nadu, India; ^bRS Mehta Jain Department of Biochemistry and Cell Biology, Kamalnayan Bajaj Institute for Research in Vision and Ophthalmology, Vision Research Foundation, Chennai, Tamil Nadu, India; ^cNational Institute of Traditional Medicine, Indian Council of Medical Research, Belagavi, Karnataka, India; ^dCentral Research Instrumentation Facility, Kamalnayan Bajaj Institute for Research in Vision and Ophthalmology, Vision Research Foundation, Chennai, Tamil Nadu, India

Communicated by Ramaswamy H. Sarma

ABSTRACT

Human DNA topoisomerase II alpha (hTopII α) is a classic chemotherapeutic drug target. The existing hTopII α poisons cause numerous side effects such as the development of cardiotoxicity, secondary malignancies, and multidrug resistance. The use of catalytic inhibitors targeting the ATP-binding cavity of the enzyme is considered a safer alternative due to the less deleterious mechanism of action. Hence, in this study, we carried out high throughput structure-based virtual screening of the NPASS natural product database against the ATPase domain of hTopII α and identified the five best ligand hits. This was followed by comprehensive validation through molecular dynamics simulations, binding free energy calculation and ADMET analysis. On stringent multilevel prioritization, we identified promising natural product catalytic inhibitors that showed high binding affinity and stability within the ligand-binding cavity and may serve as ideal hits for anticancer drug development.

Abbreviations: DNA: Deoxyribonucleic acid; ATP: Adenosine triphosphate; TopII: Eukaryotic DNA Topoisomerase II; NBD: Nucleotide-Binding Domain; TD: Transducer Domain; hTopII α : Human DNA Topoisomerase II alpha; ADMET: Absorption, Distribution, Metabolism, Excretion and Toxicity; NP: Natural Product; CI: Catalytic Inhibitor; PDB: Protein Data Bank; AMP-PNP: Adenylyl-imidodiphosphate; SIFt: Structural Interaction Fingerprints; NPASS: Natural Product Activity and Species Source; SBVS: Structure-based virtual screening; GLIDE: Grid-based Ligand Docking with Energetics; HTVS: High Throughput Virtual Screening; SP: Standard Precision; XP: Extra Precision; MM/GBSA: Molecular Mechanics/Generalized Born Surface Area; PLIP: Protein-Ligand Interaction Profiler; MD: Molecular dynamics; RMSD: Root Mean Square Deviation; RMSF: Root Mean Square Fluctuation; H-bond: Hydrogen bond; PL: Protein-ligand; PAINS: Pan-assay interference compounds; bRO5: Beyond Rule of Five; MMP: Mitochondrial Membrane Potential; GI: Gastrointestinal

ARTICLE HISTORY

Received 15 November 2022
Accepted 25 February 2023

KEYWORDS




Human DNA topoisomerase II alpha; nucleotide-binding domain; high throughput virtual screening (HTVS); molecular dynamics simulations; ADMET

1. Introduction


Topoisomerases are DNA-binding enzymes that are ubiquitous in maintaining genome topology in all living organisms. They bring about the alteration of DNA topology, through active site tyrosine-mediated, transient, single or double-stranded cleavage to fulfil the numerous demands of cellular metabolism (Nitiss, 2009a). The indispensable nature of these enzymes has long been recognised as evidenced by the vast number of studies detailing their classification, distribution, and repertoire of functions. Four structurally and functionally distinct subfamilies of topoisomerases have been elucidated, namely types IA, IB, IIA, and IIB. Type II topoisomerase enzymes catalyze ATP-dependent double-stranded breakage of DNA as opposed to the ATP-independent, single-stranded cleavage moderated by type I enzymes. The type IIA subfamily is of

clinical significance, as it encompasses the bacterial enzymes, DNA gyrase, and topoisomerase IV, along with the eukaryotic DNA topoisomerases (Champoux, 2001; McClendon & Osheroff, 2007; J. C. Wang, 2002). Consequently, the critical role played by the aforementioned enzymes in microbial pathogenesis and cancer metabolism has established them as classical drug targets, with several antibiotics and anticancer agents in clinical use (Bax et al., 2010; Delgado et al., 2018; Liang et al., 2019; Nitiss, 2009b; Pommier et al., 2010).

Eukaryotic DNA topoisomerase II (TopII) is a homodimer, delineated into three domains (Figure 1)—the N-terminal ATPase domain, the central, DNA-binding cleavage-religation domain, and the non-conserved C-terminal domain. The ATPase domain binds one ATP molecule per chain, both of which are hydrolysed to supply the energy required to power

CONTACT Umashankar vetrivel  vumashankar@gmail.com; Srujana Chitipothu  drsrujana@snmail.org  Centre for Bioinformatics, Kamalnayan Bajaj Institute for Research in Vision and Ophthalmology, Vision Research Foundation, Chennai, Tamil Nadu India.

[#]Present address: National Institute for Research in Tuberculosis, Indian Council of Medical Research, Chetpet, Chennai, Tamil Nadu, India.

 Supplemental data for this article can be accessed online at <https://doi.org/10.1080/07391102.2023.2187234>.

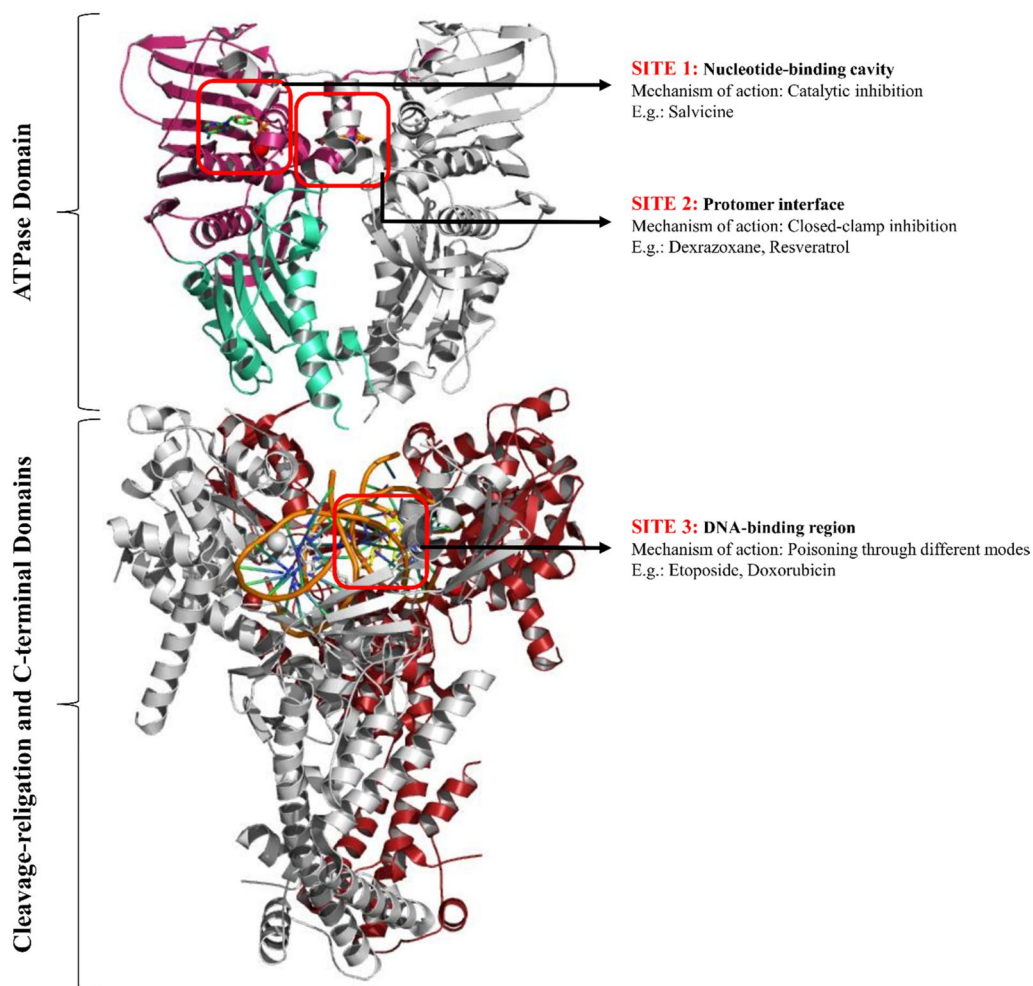


Figure 1. Schematic representation of domains and drug-target hotspots in human TopII (chain A in colour and chain B in grey): nucleotide-binding domain of ATPase domain (magenta), ATP (green), Mg^{2+} (red), transducer domain (green), resveratrol (orange), DNA (orange backbone and coloured bases), DNA-binding cleavage-religation domain (crimson), C-terminal domain (crimson).

the enzyme. This domain is further divided into two regions, namely, the nucleotide-binding domain (NBD) and the transducer domain (TD). The residues of the ligand-binding cavity of NBD, along with the Gln-Thr-Lys (QTK) loop from the TD, are highly conserved which denotes their significant role in ATP-binding and catalysis. The cleavage-religation domain binds with the DNA substrate and modulates double-stranded cleavage, along with the C-terminal domain. Together, the different domains work in synergy through the coordinated opening and closing of the dimer interface regions known as 'gates' to transport an intact DNA segment (T-segment) through the enzyme-linked, cleaved segment. This process is carried out iteratively and is driven by the hydrolysis of two ATPs per enzymatic cycle. In this way, the interconversion of different topological states of DNA takes place whenever required during processes such as DNA replication, chromosome regulation and dynamics, and transcription (Champoux, 2001; Pommier et al., 2016; Schoeffler & Berger, 2008; J. C. Wang, 2002; Wei et al., 2005).

In vertebrates, the two distinct type IIA isoenzymes are TopII alpha and beta, which are encoded by different genes and characterised by well-defined cellular roles and expression levels. Of the two, TopII alpha is the predominant isoform essential for cellular viability and its expression is

elevated at the S and G2/M phases of the cell cycle. It plays a significant role in the progression of DNA replication as well as the condensation, segregation and maintenance of chromosomes. Its activity has also proven to be vital for developing embryos and proliferating cells. The beta isoform, however, is not crucial for cell survival and shows basal expression throughout the cell cycle. It is found in all tissues but is elevated in the heart, lungs, and quiescent cells. Though it plays an accessory role during mitosis, it is known to be highly effective in DNA repair and transcription. It has also proven to be essential for neural development and spermatogenesis (Akimitsu et al., 2003; Austin et al., 2018; Bollimpelli et al., 2017; Capranico et al., 1992; Lee & Berger, 2019; Turley et al., 1997; Woessner et al., 1991). Both isoforms are observed to be significantly upregulated in various cancer conditions, thereby placing them among the principal antineoplastic targets in humans. Higher expression of human TopII alpha is associated with several aggressive cancers such as breast, prostate, lung and colorectal cancers, as well as gynaecological malignancies (Coss et al., 2009; Depowski et al., 2000; Hou et al., 2017; Schaefer-Klein et al., 2015; van der Zee et al., 1994; Zhao et al., 2020). Whereas, overexpression of beta enzyme is seen in aggressive glioblastoma, slow-growing, and solid tumours in humans. This

isoform has also been identified as the leading cause of the development of secondary malignancies (Austin et al., 2018; Azarova et al., 2007; Gao et al., 1999; Kenig et al., 2016; Vávrová & Šimůnek, 2012).

The wide range of chemotherapeutic drugs targeting the different stages of the TopII enzymatic cycle is categorised into poisons and catalytic inhibitors (CIs), based on the mechanism of action (Figure 1). TopII poisons operate through multiple modes, such as the prevention of DNA re-ligation by cleavage complex trapping (podophyllotoxins like etoposide, and teniposide) or DNA intercalation (anthracyclines like doxorubicin and daunorubicin). The resulting genotoxic stress due to the accumulation of DNA breaks and trapped cleavage complexes triggers apoptosis of both malignant and healthy cells expressing TopII. Therefore, although TopII poisons are potent anticancer drugs, the chief drawback is their detrimental mechanism of inhibition besides other deleterious side effects such as the development of cardiomyopathy, drug resistance, and therapy-induced metastasis (Hevener et al., 2018; Nitiss, 2009b; Pommier et al., 2010).

On the other hand, CIs interfere with TopII action without causing genotoxicity. CIs targeting the ATPase domain affect enzymatic activity in two primary ways. The first group consists of dexrazoxane (ICRF-187) and related molecules that arrest the ATPase domain in closed-clamp conformation and effectively prevent the passage of the T-segment of DNA. However, dexrazoxane is currently withdrawn from clinical use owing to poor efficacy. The other group is the ATP-site inhibitors that target the nucleotide-binding cavity to prevent ATP binding and subsequent hydrolysis, causing a decrease in enzyme turnover. Over the years, several classes of purine analogues, small molecules, and natural products have been predicted to exhibit ATP-site enzymatic inhibition (Bergant et al., 2018). Nevertheless, none of the ATP-site inhibitors have been approved for therapeutic use at present.

Therefore, due to the unfulfilled requirement for effective CIs targeting the ATPase domain, we have undertaken an *in silico* study for identifying novel natural product (NP) hits against human TopII alpha (hTopII α). We carried out high throughput virtual screening (HTVS) of the NPASS database and potential CIs of hTopII α were meticulously prioritised based on energetic parameters and significant interactions. Furthermore, molecular dynamics simulations of the selected receptor-ligand complexes were performed for understanding the stability of the hits followed by a detailed ADMET (Absorption, Distribution, Metabolism, Excretion and Toxicity) analysis. The methodology of the study and rationale for the identification of putative natural product CIs of hTopII α are discussed in the following sections.

2. Methods

2.1. Structure refinement of the ATPase domain of hTopII α

The structure of the ATPase domain of hTopII α (PDB ID: 1ZXM), co-crystallized with the non-hydrolyzable ATP analogue AMP-PNP (adenylyl-imidodiphosphate), was retrieved from RCSB Protein Data Bank (X-ray crystallographic resolution = 1.87 Å

(Wei et al., 2005). The Protein Preparation Wizard of the Schrödinger suite v.2019-3 (Schrödinger LLC, 2019) was utilised for optimising the receptor by generating bond orders, adding hydrogen atoms, and deleting water molecules. Missing atoms and residues were modelled using the Prime module. Residue protonation states and heteroatom ionization states at physiological pH were generated through PROPKA and Epik modules (Shelley et al., 2007), respectively. Finally, energy minimization using the OPLS_2005 force field was carried out for obtaining a refined receptor at a minimum energy conformation. Additionally, the residue interactions of AMP-PNP in chain A of the prepared protein were analysed through visual inspection and using the Structural Interaction Fingerprints (SIFt) panel (Deng et al., 2004).

2.2. Preparation of NPASS ligand dataset

In our quest to identify novel ATP-site CIs of hTopII α , we turned towards Natural Product Activity and Species Source (NPASS) v.1.0, a curated, open-source NP database (Zeng et al., 2018). The entire structure dataset (in SDF format) was downloaded and prepared for docking using the LigPrep panel of the Schrödinger suite. LigPrep offers a variety of parameters for generating high-quality, geometry-optimised structures of small molecules. Initially, all-atom three-dimensional structures of the ligands were prepared from the two-dimensional input, following which ionization states were generated at pH 7.0 ± 2.0 using the Epik module. The ligands were then desalted, and tautomers and stereoisomers were generated. Finally, the OPLS_2005 force field was used for energy minimization of the entire dataset of optimised ligands.

2.3. Preparation of reference catalytic inhibitor salvicine

At present, there is a lack of CI-bound crystallographic structures of the ATPase domain of hTopII α . Therefore, in this study, the compound salvicine was used as a reference CI for docking and comparing binding affinities of screened ligands (Hu et al., 2006). The structure of salvicine was retrieved from the PubChem database (PubChem CID 10359290) in SDF format and optimised for docking through the LigPrep panel.

2.4. Virtual screening of NPASS ligand dataset and salvicine against ATPase domain of hTopII α

Structure-based virtual screening (SBVS) is a powerful *in silico* method for rapidly predicting the binding affinity of small molecule libraries with a selected target. It has aided in significantly reducing the time, expenditure and human effort involved in the identification of potential lead molecules for drug design based on parameters such as docking score and binding mode (Maia et al., 2020). In this study, the GLIDE (Grid-based Ligand Docking with Energetics) program of the Schrödinger suite was utilised for the HTVS of the NPASS ligand dataset against the NBD in chain A of ATPase domain of hTopII α (Halgren et al., 2004). GLIDE efficiently employs hierarchical filters such as HTVS, SP, and XP for the successive filtration of docked ligands resulting in a refined pool of

molecules with high binding affinity and optimal binding conformation. Prior to the screening, all heteroatoms were stripped from the receptor.

2.4.1. Receptor grid generation

SIFt analysis and visual inspection revealed critical interactions between the co-crystallized ligand and catalytic site residues in chain A of 1ZXM. These residues were selected for constructing a receptor grid for performing docking calculations. The van der Waals scaling and partial charge cut-off for non-polar atoms of the receptor were set to 1.0 and 0.25 respectively during grid generation.

2.4.2. Virtual screening protocol

For the screening of the NPASS dataset, the potential for non-polar atoms of ligands was softened by setting the van der Waals scaling factor to 0.80 with a partial charge cut-off of 0.15. The ligands were flexibly docked and nitrogen inversions and ring conformations were sampled. Virtual screening was first carried out in HTVS mode following which the top 50% of unique hits were passed on to Standard Precision (SP) docking. The top 25% of unique hits obtained from this step were carried forward to Extra Precision (XP) mode, for which XP descriptor information was recorded. Per-residue interaction scores were written for all residues interacting with the ligands during docking. The top hundred unique hits shortlisted based on XP GScore were carried over to the next step. Finally, the reference inhibitor, salvicine, was also docked in XP mode.

2.5. Molecular Mechanics/Generalized Born Surface Area calculations

Molecular Mechanics-Generalized Born Surface Area (MM/GBSA) is a parameter calculated for assessing the relative binding free energies of a set of ligands with respect to the receptor (E. Wang et al., 2019). The ΔG_{bind} scores for the top hundred hits as well as the hTop11 α -salvicine docked complex from the previous step were calculated using the Prime MM-GBSA panel, as follows:

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

This equation can be decomposed into individual energy terms using the formula:

$$\Delta G_{\text{bind}} = \Delta H - T\Delta S = \Delta E_{\text{MM}} + \Delta G_{\text{SOL}} - T\Delta S$$

where ΔE_{MM} represents the gas-phase interaction energy given as a sum of internal, electrostatic, and van der Waals energies; ΔG_{SOL} represents the free energy of solvation given as a sum of polar and non-polar energy terms and $T\Delta S$ represents the entropy change.

The VSGB solvation model and OPLS_2005 force field were used in the MM/GBSA calculations.

2.6. Selection of best five NP hits

The top hundred NP hits obtained after SBVS and MM/GBSA scoring were manually inspected and endogenous molecules found in the human body (such as ATP analogues, metabolic precursors and metabolites), as well as non-endogenous aliphatic molecules, were removed. The remaining ligands were then stringently scrutinised and reranked based on XP GScore, HBond score, binding free energy, penalties, and critical residue interactions. Out of these, only the best five complexes were selected for further steps such as the assessment of physicochemical properties and stability analysis through MD simulations.

The protein-ligand interactions present in the selected docked complexes were examined using Protein-Ligand Interaction Profiler (PLIP) (Salentin et al., 2015) and the structures were rendered using Pymol v.2.3.1.

2.7. Evaluation of physicochemical properties, pharmacokinetics, and toxicity parameters of salvicine and selected NP hits

The physicochemical and pharmacokinetic properties of ligands together with the toxicity profiles are important criteria that aid in determining the suitability of screened hits to be carried forward for *in vitro* testing and subsequent clinical development. These parameters are crucial for predicting the behaviour of a ligand within a biological system, such as bio-availability, potency toward the desired target and induction of undesirable reactions such as toxicity or promiscuity.

Salvicine and the selected NP hits were assessed using various cheminformatic tools. The parent chemical groups of the ligands were identified using Classyfire, an online tool that computes chemical taxonomy based on structural features (Djombou Feunang et al., 2016). The physicochemical and pharmacokinetic properties were analysed by the Qikprop module of the Schrödinger suite. Furthermore, the webserver SwissADME was utilised for computing additional pharmacokinetic and medicinal chemistry parameters (Daina et al., 2017). The ProTox-II server, an open-source *in silico* toxicity prediction tool, was used for computing the acute toxicity class, various toxicological endpoints and toxicological pathway reactivities based on the structures of the ligands (Banerjee et al., 2018).

2.8. Molecular dynamics simulations of six docked complexes

Molecular dynamics (MD) simulations serve as a robust tool for studying protein-ligand interactions of complexes obtained from SBVS and aid in observing local fluctuations and deviations in the ligand-binding cavity (Alonso et al., 2006). The stability and interaction network of salvicine and the five NP hits within the binding cavity of hTop11 α were evaluated through 100 ns MD simulations.

Condensed phase all-atom MD simulations of the hTop11 α -ligand complexes were carried out using the Desmond package, using identical protocols (Bowers et al., 2006). Each simulation

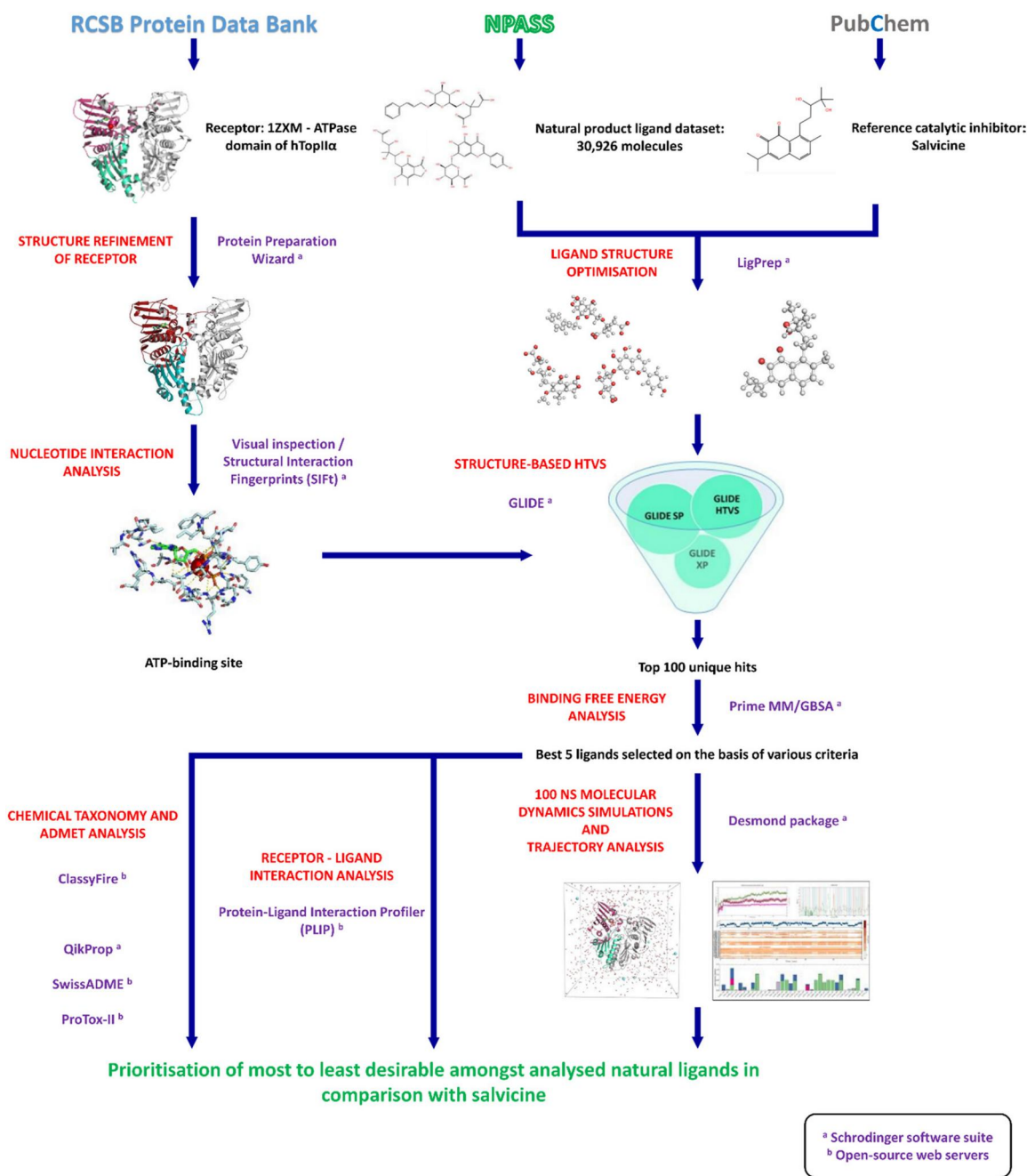


Figure 2. Schematic representation of study methodology with details of software and tools used.

system was set up through the System Builder panel by placing a docked complex in an orthorhombic box at a buffer distance of 10 Å from the box boundary. The system was then solvated by the TIP3P water model and neutralised by the addition of an appropriate number of Na⁺/Cl⁻ ions. Additional ions were added for simulating physiological salt concentrations of 0.15 M. The electroneutral system was brought to a relaxed state by energy minimization using the LBFGS algorithm, for a maximum of 2000 iterations until convergence with a threshold of 1.0 kcal/mol/Å, followed by an additional extensive default relaxation protocol. Short-range Coulomb interactions were treated by the cut-off method with a radius of 9 Å. Finally, an NPT ensemble was used for the production simulation run of 100 ns timescale. The simulation was carried out at a temperature of

300 K and pressure of 1.0 bar using Nose-Hoover chain thermostat (Tuckerman et al., 1992) and the Martyna-Tobias-Klein barostat (Martyna et al., 1994) respectively. Simulation trajectory data was recorded at an interval of 4.8 ps.

On completion, the MD simulation trajectories were subjected to exhaustive analyses using the Simulation Event Analysis and Simulation Interaction Diagram tools. Root Mean Square Deviation (RMSD) and Root Mean Square Fluctuation (RMSF) of receptors and ligands, the radius of gyration, and interaction network parameters were computed. The Simulation Quality Analysis panel was used to assess the overall quality of the simulations.

The methodology of the entire study has been summarised in Figure 2.

3. Results

3.1. Key interactions and structural insights on the ATPase domain of hTopII α

The target region of hTopII α , the ATPase domain, was a homodimer composed of two chains A and B of 400 amino acids each, ranging from residues 29 to 428. In each chain, the NBD comprised residues 29 to 264, followed by the TD, from residues 265 to 428. The prominent characteristic of the NBD was the Bergerat motif, which was a conserved structural feature common to ATP-binding proteins of the GHKL superfamily (Bergerat et al., 1997; Wei et al., 2005)

Critical inspection and SIFT analysis of the chain A catalytic site of 1ZXM revealed crucial protein-ligand interactions which were grouped based on the three distinct structural moieties of the co-crystallized ATP analogue, namely the adenine moiety, ribose moiety, and triphosphate tail which was co-ordinately bonded to Mg²⁺. Multiple water molecules were also present within this site and hence it was surmised that water bridges contributed significantly to ligand binding. The adenine moiety was strongly stabilised by a hydrogen bond (H-bond) with ASN 120, which along with THR 215 and TYR 34 (from chain B) participated in water-mediated interactions as well. Moreover, ILE 88, ASN 91, ASN 95, and LYS 123 and the side chain of ILE 125 were capable of water-mediated and hydrophobic interactions respectively.

The major interactions of the ribose moiety were through SER 149 which was within H-bond distance of both ribose hydroxy groups. Additionally, it was noted that both SER 149 and TYR 34 (B) might interact through water bridges and the side-chains of ILE 141 and PHE 142 were close enough to engage in non-polar interactions with ribose (Janežič et al., 2017; Pogorelčnik et al., 2015).

The H-bond network of the triphosphate groups was identified to be most crucial for enzymatic function. The triphosphate moiety interacted extensively with a pocket of the catalytic site formed by a glycine-rich sequence given by GXXGXG and comprised residues 161–166. Known as the Walker A motif, it is a greatly conserved sequence found in ATP-binding enzymes. In addition to this motif, ASN 91 formed an H-bond with the α -phosphate, whereas SER 148 and ASN 150 formed H-bonds with the β -phosphate. The residue LYS 168 which flanked the motif, was hydrogen-bonded to the O5' atom of the ligand and also formed a stable salt bridge with the α -phosphate. Most importantly, an Mg²⁺ cation formed an octahedral coordination complex with the oxygen atoms of the three phosphates, the carbonyl oxygen of ASN 91, and two intimately placed water molecules. One of these conserved water molecules also showed a water-mediated interaction with GLU 87 and this residue was believed to hold the water molecule in place to promote nucleophilic attack during ATP hydrolysis.

Furthermore, it has been established that the QTK loop from the TD, which consisted of GLN 376, THR 377, and LYS 378 here, played a critical role in enzyme turnover (Bendsen et al., 2009). The strong salt bridge formed between the side-chain amino nitrogen of LYS 378 and γ -phosphate of ATP will be hydrolysed to release the energy required for the catalysis

besides the opening of the N-gate. GLN 376 was also situated near enough to be hydrogen-bonded to the γ -phosphate.

Taking these observations into account, the receptor grid was generated using residues considered critical for ligand recognition namely, GLU 87, ASN 91, ASN 95, ASN 120, SER 148, SER 149, ASN 150, ARG 162, ASN 163, GLY 164, TYR 165, GLY 166, ALA 167, LYS 168, GLN 376, THR 377, and LYS 378. We evaluated the virtual screening results based on interactions with these key residues in the catalytic cavity.

3.2. Structure-based virtual screening of NPASS ligand dataset against ATPase domain of hTopII α

3.2.1. Rationale behind SBVS in the absence of Mg²⁺ cation

Based on the assumption that Mg²⁺ was essential for ligand binding, earlier studies had docked potential CIs of hTopII α in the presence of the cation (Bergant et al., 2018). However, we carried out high throughput SBVS after removing the Mg²⁺ cation from the catalytic site. The rationale behind this step is detailed below.

The catalytic cycle of hTopII utilises four Mg²⁺ cations, out of which two are involved in the two-metal-ion DNA cleavage mechanism at the cleavage-religation domain and the remaining two at the ATPase domain. As described earlier, each of the latter two Mg²⁺ cations binds an ATP in either chain of the ATPase domain. The triphosphate moiety of ATP exhibited strong tridentate coordination with the Mg²⁺ cation which held the γ -phosphate in place to facilitate hydrolysis. Though this illustrated the crucial role of the cofactor, the question arose whether a cation was required for the stabilisation of the CIs competing for the ATP-binding site. One of the primary reasons to support a cofactor-free docking environment was that there existed a greater probability of the native ligand binding to the NBD as an ATP-Mg substrate rather than a cation-free entity. This was experimentally demonstrated by Osheroff. N where it was observed that the rate of ATP-Mg complex formation was directly proportional to ATPase activity, showing that the ligand binds as an ATP-Mg substrate in the catalytic site (Deweese & Osheroff, 2010; Osheroff, 1987). Moreover, it is known that over 75% of cellular ATP exists in the form of Mg-complexed ATP and attained a biologically functional conformation by binding to the cofactor (Gout et al., 2014; Weston, 2009; Wolf & Cittadini, 2003). This meant that the cofactor in the catalytic site was not a component of the enzyme but rather came along with the nucleotide. Lastly, we sought to identify ligands that exhibited a strong affinity towards the binding site without the energetic contributions or stabilisation of a cofactor, which may prove to be more desirable.

3.2.2. Five NP hits selected from screened NPASS database

There are several open-source NP databases available now but the NPASS database was chosen for this study due to its many merits. Other repositories mostly contained a limited number of entries from specific biological domains, whereas, NPASS v.1.0 comprised 30,926 NPs (at the time of the study)

with diverse sources of origin across all domains of life from micro-organisms to plants and animals. Moreover, the curated, richly annotated entries are hosted on a well-maintained server, with a readily available structure dataset and meta-data for download.

The NPASS v.1.0 ligand dataset of 30,926 ligands was prepared to yield 46,458 structures in total, and subjected to HTVS. Following this, 8,726 ligands (top 50% of unique HTVS output) were carried forward to SP docking, from which 2,179 (top 25% of unique SP output) were selected for XP docking. Subsequently, the top hundred hits ranked based on XP docking score were subjected to MM/GBSA analysis. These hundred were manually inspected, undesirable ligands were removed and the remaining ones were reranked to give five top hits namely, NPC29607, NPC470359, NPC43211, NPC186418 and NPC132921. The chemical structures of these five hits along with salvicine are illustrated in [Supplementary Figure 1\(a\)–\(f\)](#). The binding modes of salvicine, the chosen NP ligands and ATP (modelled from PDB ID: 1ZXM as a reference for binding orientation) within the nucleotide-binding site are depicted in [Figure 3\(a\)–\(g\)](#).

The empirical formula, NPASS and PubChem database identifiers, chemical class and natural sources of origin of all six small molecules are given in [Table 1](#). The Classyfire tool computed the chemical taxonomy and the parent chemical group of each of the six ligands. The server output showing the hierarchical taxonomies of the ligands is displayed in [Supplementary Table 1](#).

The reference compound salvicine is a semi-synthetic diterpene quinone of the naphthoquinone class, derived from a phytocompound found in *Salvia prionitis*, a Chinese medicinal herb. It was seen to exhibit ATP-competitive anti-tumor activity against hTopII α and has currently advanced to clinical trials in China (Hu et al., 2006; Xu et al., 2012; Zhang et al., 1999). Hence, for this reason, it was chosen as the reference inhibitor for this virtual screening study.

NPC29607 was identified as longumoside B, which is a naturally occurring glycoside present in the Indian long pepper plant, *Piper longum* (Jiang et al., 2013). NPC470359 was a fungal phthalide named penicacid C, belonging to a class of mycophenolic acid derivatives, which was isolated from the marine *Penicillium* species SOF07 (Chen et al., 2012). NPC43211, NPC186418 and NPC132921 were identified as scutellarin, caffeic acid 4-O-glucoside, also known as linocaffein, and caffeoylmalic acid respectively, which have widespread occurrence in the plant kingdom (Budzianowski, 1990; Klosterman & Muggli, 1959; L. Wang & Ma, 2018). Scutellarin is a well-known flavonoid widely studied for its therapeutic properties. Linocaffein is a phenolic glycoside and caffeoylmalic acid belongs to the coumaric acid derivatives class. Both these ligands contain a hydroxycinnamic acid moiety and are being investigated for their nutraceutical effects on human ailments. It should also be noted that salvicine and all selected compounds contained aromatic substructures, with all but caffeoylmalic acid containing an additional heterocyclic or carbocyclic moiety as well.

The results from virtual screening, binding free energy analysis and PLIP analysis of the six docked complexes are listed in [Table 2](#). The docking scores and binding free energies

revealed that all five selected NP ligands scored significantly better than salvicine. The docking score of salvicine was only -6.732 kcal/mol whereas the values for the NP hits ranged from -17.072 to -14.484 kcal/mol. Longumoside B notably exhibited the highest docking score of -17.072 kcal/mol. It was followed by penicacid C (-14.924), scutellarin (-14.818), linocaffein (-14.623) and caffeoylmalic acid (-14.484). The dG bind of salvicine was -46.786 kcal/mol whereas, for the NP ligands, the binding free energies ranged between -81.777 and -62.640 kcal/mol. The contribution of the HBond component of the GLIDE docking score, which is a measure of the energetic contribution of H-bond formation, was also lower for salvicine than the NP hits. These results are indicative of relatively stronger binding affinities exhibited by the NP ligands within the catalytic site than salvicine.

As for the binding mode of salvicine, the exact pose has not been determined by crystallographic methods and previous studies have proposed different modes using a variety of tools (Radaeva et al., 2020). From these studies, we observed that some residues were repeatedly shown to be involved in the recognition of salvicine within the NBD, which correlated with the interacting residues in our docked complex as well. In the hTopII α -salvicine complex, the naphthalene group of salvicine partially coincided with the region occupied by the ribose of ATP and a terminal isopropyl group extended towards the QTK loop. H-bond formation with ASN 91, SER 148, ASN 150, ARG 162 and GLY 164 was detected.

On the other hand, the formation of extensive H-bond networks was observed in all the hTopII α -NP complexes. Unlike salvicine, all NP complexes showed the presence of at least one salt bridge as well. In this scenario, salt bridges have been defined as ionic bonding formed between the proton donated by LYS 168 or LYS 378 and negatively charged moieties from the NP ligands.

The binding mode of longumoside B was such that the trans-cinnamic acid group, β -D-glucopyranose ring and methyl pentanedioic acid group aligned with the positions of the adenine, ribose and triphosphate moieties of ATP respectively. The hydroxy and carbonyl moieties showed extensive H-bond interaction with the Walker A motif and salt bridges were formed with both LYS 168 and LYS 378.

In the case of penicacid C, the isobenzofuranone (phthalide) group fell roughly between the region occupied by the adenine and ribose moieties in the native/co-crystallized ligand and the position of the carboxylic acid group coincided with the phosphate tail. Scutellarin was the largest among the selected NP hits and its binding region extended well beyond that of the native ligand. The flavonoid fragment extended beyond the adenine binding region and the glucuronic acid moiety coincided with the position of the triphosphate tail. Both formed H-bonds with the Walker A motif and salt bridge formation was observed between penicacid C and LYS 378 as well as scutellarin with LYS 168. In linocaffein, the carboxyl moiety of the hydroxycinnamic acid fragment was the only region that coincided with the binding orientation of the native ligand. Caffeoylmalic acid was oriented such that the double-ring moiety coincided with the position of the adenine of the native ligand.

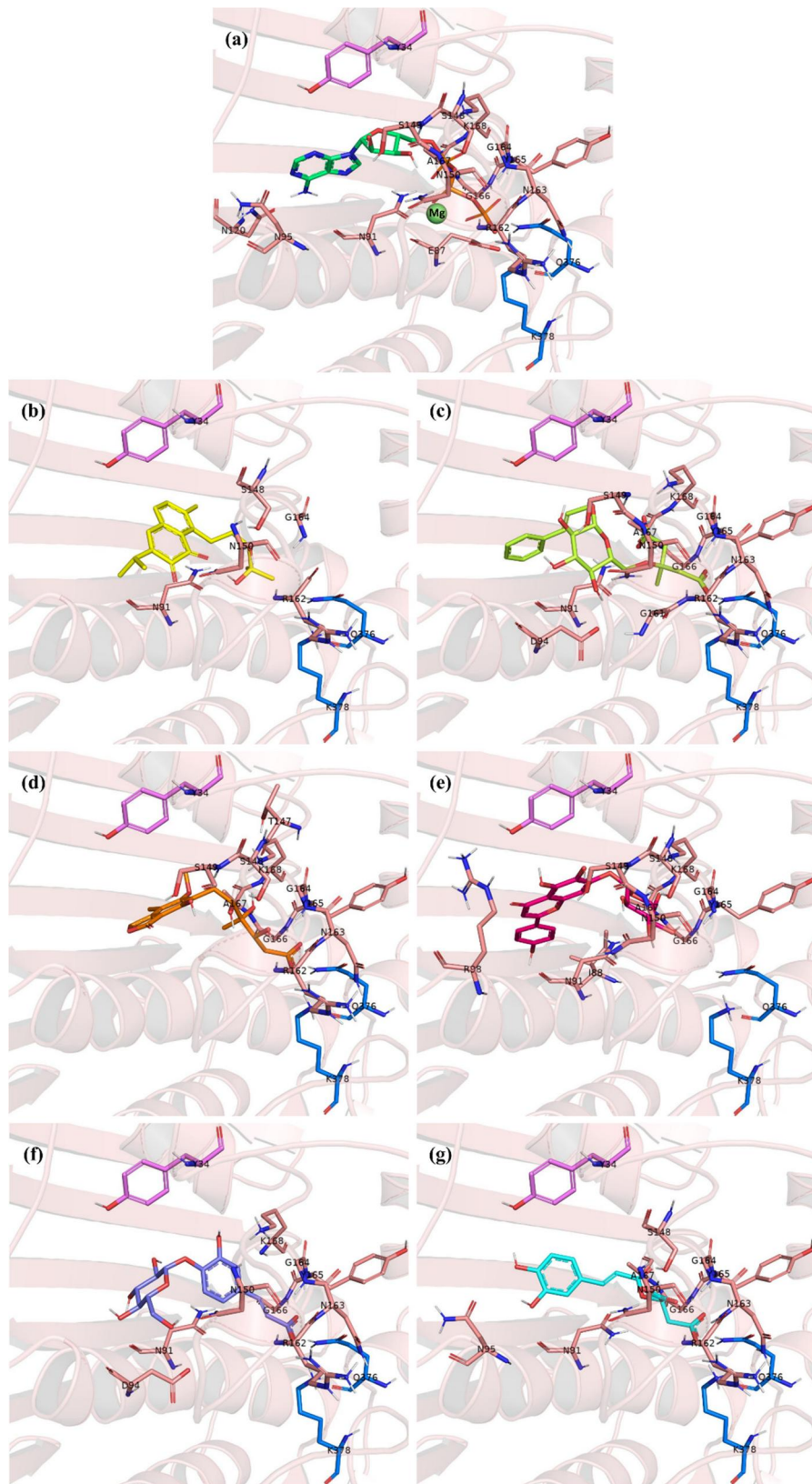


Figure 3. Binding mode of ligands within the NBD of hTopII α : (a) native ligand ATP (with Mg²⁺), (b) reference inhibitor salvicine, (c) Longumoside B, (d) Penicid C, (e) Scutellarin, (e) Linocaffeine and (f) Caffeoylmalic acid (all chain A residues in salmon except Q376 and K378 in blue and chain B residue Y34 in purple).

3.2.3. Analysis of molecular dynamics simulations of the six docked complexes

The MD simulations yielded interesting results, where the formation and stability of various interactions exhibited by

all six ligands were meticulously evaluated (Figures 4–9a–c and Supplementary Figure 2a–f).

The RMSD of the backbone atoms of the receptor protein indicated each structure's stability under simulated conditions

Table 1. Characteristic features of the reference inhibitor salvicine and five chosen NP ligands.

Ligand type	Compound name	Empirical formula	NPASS ID	PubChem CID	Parent chemical group	Source of origin
Reference catalytic inhibitor	Salvicine	C ₂₀ H ₂₆ O ₄	–	10359290	Naphthoquinones	Semi-synthetic derivative of quinone compound from <i>Salvia prionitis</i> , a Chinese medicinal plant
Screened natural product	Longumoside B	C ₂₁ H ₂₈ O ₁₀	NPC29607	71579641	Fatty acyl glycosides of mono- and disaccharides	<i>Piper longum</i> , the Indian long pepper plant
	Penicacid C	C ₁₇ H ₂₂ O ₉	NPC470359	59053798	Phthalides	<i>Penicillium</i> sp. SOF07, a marine-derived fungus
	Scutellarin	C ₂₁ H ₁₈ O ₁₂	NPC43211	185617	Flavonoid-7-O-glucuronides	Multiple plant species
	Linocaffein Caffeoylmalic acid	C ₁₅ H ₁₈ O ₉ C ₁₃ H ₁₂ O ₈	NPC186418 NPC132921	14825516 6124299	Phenolic glycosides Coumaric acids and derivatives	Multiple plant species Multiple plant species

Table 2. Virtual screening and MM/GBSA results along with residue interactions identified in the six docked complexes using the PLIP server.

Ligand name	Glide docking score	XP GScore (kcal/mol)	XP HBond (kcal/mol)	MM/GBSA dG bind (kcal/mol)	Hydrogen bonds	Salt bridges	Hydrophobic interactions
Salvicine	–6.732	–6.732	–2.105	–46.786	ASN 91, SER 148, ASN 150, ARG 162, GLY 164	–	ASP 94, ARG 98, ILE 141, <u>PHE 142</u> , ALA 167
Longumoside B	–17.072	–17.072	–6.324	–79.315	ASN 91, ASP 94, SER 149, ASN 150, GLY 161, ARG 162, ASN 163, GLY 164, TYR 165, GLY 166, ALA 167, GLN 376	LYS 168, LYS 378	ASN 91, ILE 125, ILE 141, <u>PHE 142</u> , <u>THE 215</u>
Penicacid C	–14.907	–14.924	–7.059	–81.777	THR 147, SER 148, SER 149, ARG 162, ASN 163, GLY 164, TYR 165, GLY 166, ALA 167, LYS 168, GLN 376	LYS 378	ASN 91, PHE 142
Scutellarin	–14.814	–14.818	–6.794	–62.640	ILE 88, ASN 91, ARG 98, SER 148, SER 149, <u>ASN 150</u> , <u>GLY 164</u> , TYR 165, GLY 166, ALA 167, LYS 168	<u>LYS 168</u>	ILE 88, ASN 91, PHE 142, THR 215, ILE 217
Linocaffein	–14.623	–14.623	–6.191	–72.406	ASN 91, ASP 94, ASN 150, ARG 162, ASN 163, GLY 164, TYR 165, GLY 166, LYS 168, GLN 376	LYS 378	–
Caffeoylmalic acid	–14.484	–14.484	–5.165	–66.276	ASN 91, ASN 95, SER 148, ASN 150, ARG 162, ASN 163, GLY 164, TYR 165, GLY 166, ALA 167, GLN 376	LYS 168, LYS 378	ASN 91, ILE 125, ILE 141, PHE 142, ALA 167

Note: Residues participating in multiple interactions are underlined and crucial LYS 378 interactions are in bold.

whereas the radius of gyration indicated the degree of compactness of the receptor. For all complexes, the RMSD of protein backbones was around 3 Å and they also exhibited a stable Rg in the range of 2.7–2.8 Å throughout the simulation timescales. The RMSF of the protein backbone atoms were computed to examine the local shifts in the receptor. While the residues in the loop and terminal regions showed large fluctuations as expected, the rest remained relatively stable, especially the catalytic site residues showed minimal fluctuations (<1.5 Å) in all complexes. The internal RMSD of all six ligands was maintained around 2 Å or under, which showed that they underwent mild internal fluctuations with respect to the input binding mode.

In the case of salvicine, visual inspection of the simulation trajectory showed that the ligand remained within the cavity

throughout and did not undergo major internal conformational changes (as seen from ligand internal RMSD). However, the Lig fit Prot RMSD (which indicated the RMSD of the ligand from complex superimposed on protein backbone RMSD) fluctuated throughout the simulation timescale which indicated poor stability as a whole within the binding site. This was supported by the Protein-Ligand contacts histogram and timeline where it was observed that salvicine failed to maintain all of its strong H-bonds with ASN 91, SER 148 and ASN 150 from the docked complex. Moreover, hydrophobic contacts and water bridges were the primary interactions between the ligand and protein, with negligible H-bond formation. The number and nature of interactions as well as interacting residues across different time points in the simulation were observed to vary throughout. It can be

concluded that a combination of multiple transient and mostly weak interactions kept salvicine from diffusing away from the NBD.

The top-scoring NP hit longumoside B exhibited the best interaction profile and a high degree of stability when compared to salvicine and other NPs. Trajectory inspection showed that it was tightly bound to the protein predominantly by numerous stable H-bonds followed by limited hydrophobic interactions and water bridges. The Lig fit Lig RMSD showed two distinct shifts which were mirrored by the Lig fit Prot RMSD. That is, the ligand seemed to have readjusted its binding orientation such that the trans-cinnamic acid moiety moved from its original position to form a stable π - π stacking with PHE 142. This change was reflected in the RMSD increase from \sim 20 ns to 60 ns. Around 60 ns the Lig fit Lig and in turn, Lig fit Prot had further increased showing more changes in the ligand binding mode, which remained stable till the end of the simulation. This was primarily due to the movement of the glucopyranoside ring. All H-bonds but one, observed in the docked complex were maintained extremely well throughout the simulation timescale. Instead

of SER 149, an H-bond was formed with SER 148 during the simulation. Most importantly, the two salt bridges with LYS 168 and LYS 378 were present for over 95% of the simulation timescale. The crucial QTK loop was strongly bound by forming an H-bond with GLN 376 as well.

Trajectory analysis of penicacid C revealed it was similar to salvicine such that the ligand did not undergo internal conformation changes but exhibited a poor interaction profile (as evidenced by Lig fit Lig, Lig fit Prot RMSD and PL contact graphs). While the H-bonds with SER 148, SER 149 and Walker motif residues were relatively well-maintained, the weak H-bond with GLN 376 and salt bridge with LYS 378 were reduced to water-mediated interactions. Moreover, recurring intramolecular H-bonds formed between the ligand hydroxy groups were identified as major deterrents of intramolecular H-bond formation. On the whole, the ligand seemed to be stabilised by five H-bonds on average and transient water bridges.

Scutellarin did neither undergo internal conformational changes nor did move much from its initial binding orientation (as evidenced by Lig fit Lig, Lig fit Prot RMSD graphs).

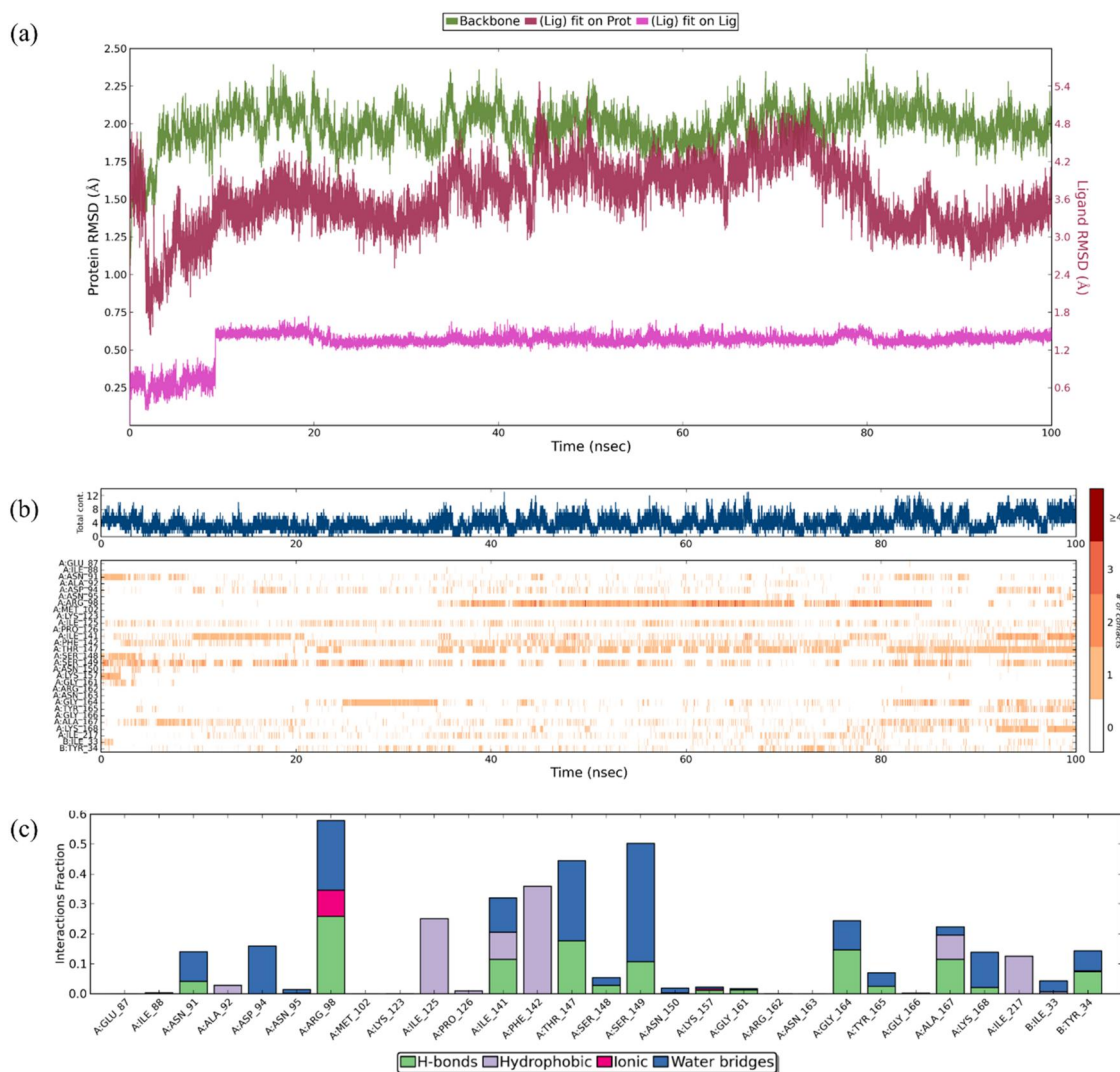


Figure 4. Results of 100 ns MD simulation of the hToll α -salvicine complex: (a) protein-ligand RMSD timeline, (b) protein-ligand contacts timeline and (c) protein-ligand contacts histogram.

An exceptionally strong H-bond as well as water bridges with ASN 91, and an extremely stable salt bridge with LYS 168 were observed. A moderately stable H-bond with SER 149 was seen and none of the others from the docked complex was maintained well. Surprisingly, a highly secure H-bond was introduced during the simulation between the phenolic hydroxy group and ASN 120. Also, ARG 62 established a couple of relatively stable H-bonds, which were not observed in the docked complex. On the whole, scutellarin remained firmly bound within the cavity through ASN 91, ASN 120, and LYS 168 and a mixture of fairly strong H-bonds and shifting water bridges.

Linocaffein formed stable intermolecular interactions mainly in the form of H-bonds followed by water-mediated interactions. Lig fit Prot shows a mild disturbance between around 65 to 73 ns, when the glucopyranoside hydroxyl group was temporarily attracted by the LYS 157 side chain, but the ligand remained otherwise stable. Compared to the docked complex, the ASN 91 and Walker motif H-bonds were conserved and supported by water bridges. Most importantly, the LYS 378 salt bridge was exceptionally strong

along with the GLN 376 H-bond, which held the QTK loop in place. On average, about 6 H-bonds were present at any point in the simulated duration.

Finally, caffeoylmalic acid was noted to be relatively stable (Lig fit Prot RMSD) and like linocaffein, was tightly bound through H-bonds and several water bridges during the simulation. The Walker motif H-bonds as seen in the docked complex, with residues ARG 162, ASN 163, TYR 165 and GLY 166 were preserved while the others mostly existed as transient water-mediated interactions. The QTK loop was well held steady by a robust GLN 376 H-bond and LYS 378 salt bridge.

3.3. Physicochemical properties and ADMET analysis of salvicine and five NP hits

The relevant physicochemical and ADMET properties computed for all six small molecules by Qikprop and SwissADME are presented in Tables 3 and 4.

QPlogPo/w and QPlogS were within the preferred range for all six ligands, which indicated good lipophilicity and solubility. QPPCaco indicated passive entry across Caco-2

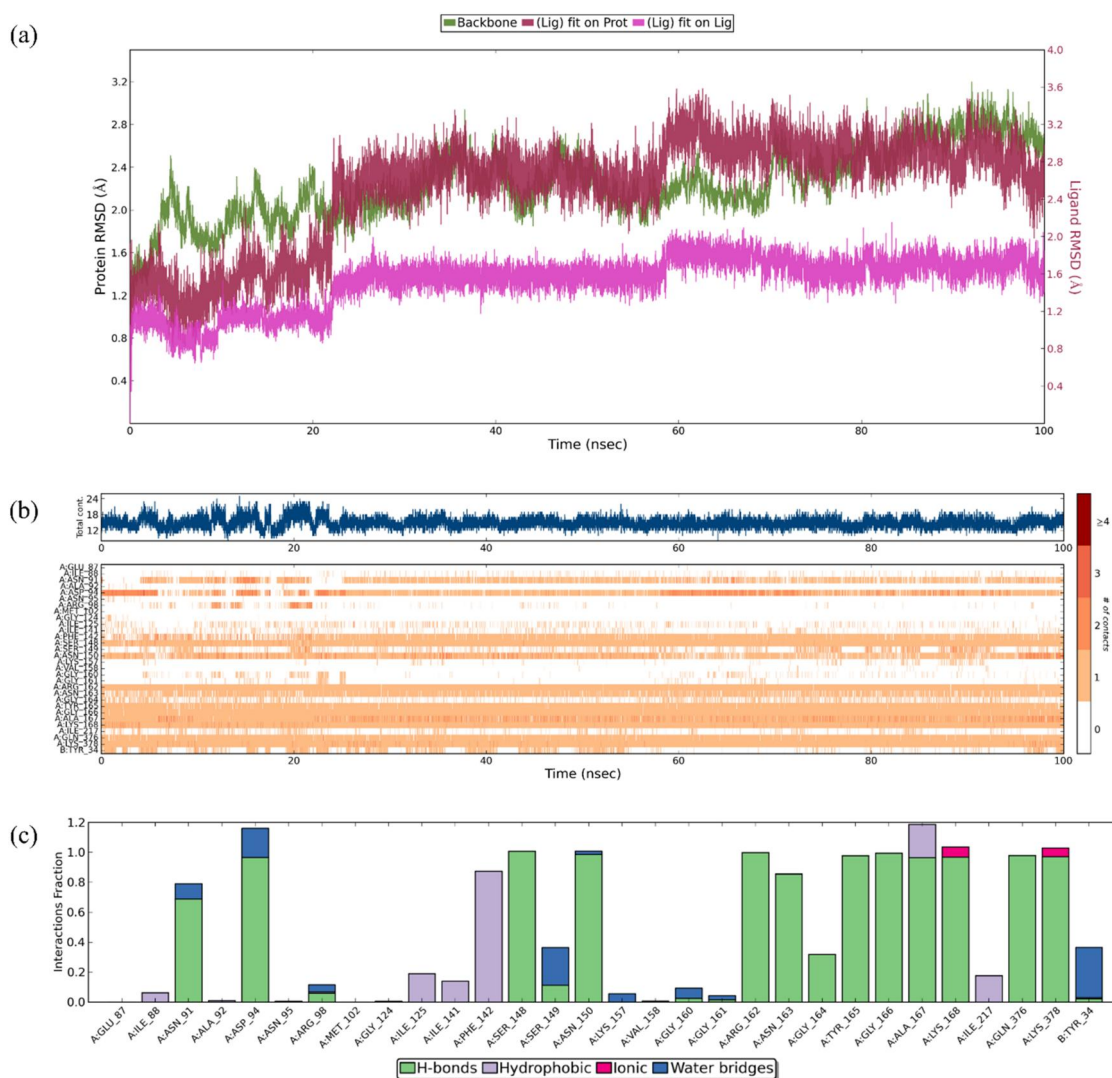


Figure 5. Results of 100 ns MD simulation of the hTopII α -longumoside B complex: (a) protein-ligand RMSD timeline, (b) protein-ligand contacts timeline and (c) protein-ligand contacts histogram.

cells which represented the gut-blood barrier, for which salvicine exhibited high GI absorption whereas the rest showed low GI absorption. SwissADME qualitatively indicated high GI absorption for salvicine, borderline absorption for caffeoylmalic acid and low absorption for the other NPs. Salvicine also showed high oral absorption followed by penicacid C and longumoside B which were predicted to have moderate absorption. The other NP hits exhibited poor oral absorption.

SwissADME computed longumoside B, penicacid C and scutellarin as potential P-glycoprotein substrates. All NP hits were observed to be non-inhibitors of cytochrome P450 isoforms whereas salvicine was predicted as a CYP2C19 inhibitor.

QPlogHERG predicted the likeliness of HERG K^+ channel inhibition which is a marker of cardiotoxicity and other adverse reactions. While all ligands scored in the recommended range, the predicted value for salvicine (-4.228) was close to the cut-off value (-5).

Only salvicine, longumoside B and penicacid C fell within the suggested range for QPlogBB. QPPMDCK, which indicated oral absorption as well as passive diffusion across the blood-brain barrier (BBB) was predicted to be excellent for salvicine

while the NP hits showed low MDCK cell permeability. The SwissADME prediction also supported the Qikprop results and indicated that only salvicine was BBB permeant. Moreover, salvicine was more likely to express CNS activity as well whereas the NP hits were predicted as inactive. All six ligands scored within the prescribed range for QPlogKhsa, which was indicative of human serum albumin binding which affects the pharmacokinetics.

Lipinski's rule of five (RO5) and Jorgensen's rule of three (RO3) define optimal physicochemical properties for orally bioavailable drugs. Similarly, the leadlikeness property predicted by SwissADME provides parameters for small molecules which can be optimised to become oral drugs. Salvicine, longumoside B, penicacid C and caffeoylmalic acid qualified the RO5 parameters. Salvicine was the only molecule which qualified RO3 and showed lead-like behaviour as well. The synthetic accessibility of all six ligands was predicted to be easy to moderate. It is a quantitative parameter supplied by SwissADME generated based on the principle that if fragments present in the screened molecules are frequently found in synthetic libraries, then they must be easier

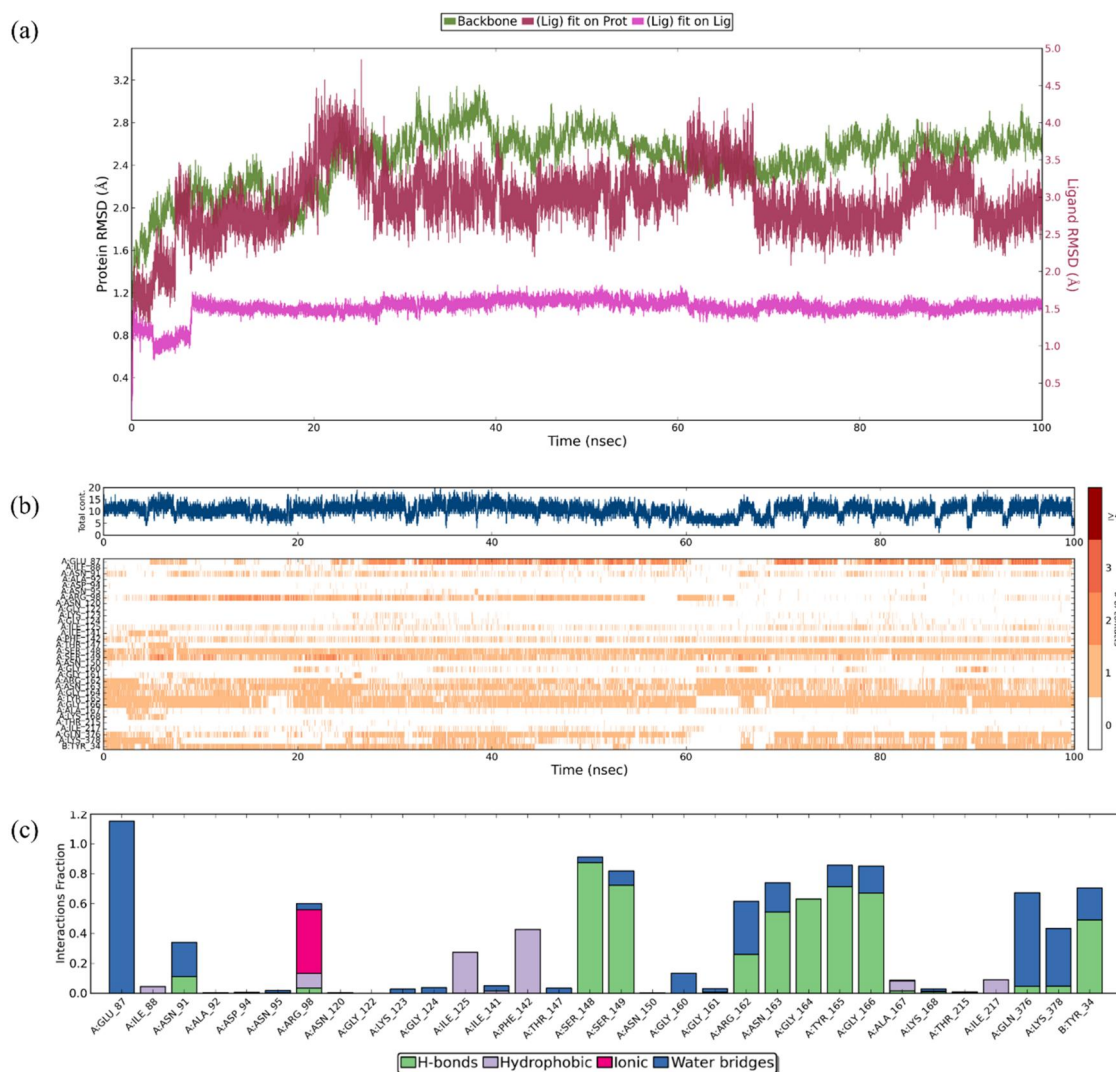


Figure 6. Results of 100 ns MD simulation of the hTopII α -penicacid C complex: (a) protein-ligand RMSD timeline, (b) protein-ligand contacts timeline and (c) protein-ligand contacts histogram.

to synthesise. Though all five NP hits are naturally occurring, synthetic accessibility is a marker for prioritising molecules as it is indicative of future synthetic potential if they will be considered for downstream biomedical applications.

Pan-assay interference compounds (PAINS) and Brenk's structural alert filters are used to identify promiscuous substructures which demonstrate non-specific activity and produce undesirable side effects. These filters are critical for screening and eliminating molecules with unwanted fragments from a pool of potential leads before beginning *in vitro* testing. Salvicine showed two PAINS alerts, namely imine_one_A, quinone_D and one Brenk alert, which was a diketo_group. Scutellarin contained a catechol_A, linocaffein contained a michael_acceptor_1 and caffeoylmalic acid had one each of both fragments. Longumoside B and penicacid C were free from undesirable fragments.

The acute oral toxicity module of ProTox-II (Table 5) predicted all ligands as less toxic, with the predicted classes being 4 for salvicine and penicacid C, and 5 for the other NP hits (the prediction accuracy for these calculations was estimated to be between 67.38% to 69.26%). The median lethal

dose (LD₅₀) for the class 4 ligands was significantly lower than the rest, predicted as 1600 mg/kg for salvicine and 352 mg/kg for penicacid C (lowest of all ligands). The LD₅₀ for the rest ranged from 3800 mg/kg to 5000 mg/kg. Salvicine was predicted to be a potential carcinogen and active for the MMP pathway, with possible mutagenic behaviour. Amongst the NP ligands, longumoside B and penicacid C were found to be inactive for all the toxicity endpoints, with only the former receiving high confidence scores for all parameters. Scutellarin was also predicted to be a potential carcinogen (low confidence score). Of all the ligands, both linocaffein and caffeoylmalic acid were classified as immunotoxic (high confidence scores).

4. Discussion

Catalytic inhibition of hTopII α has long been proposed as a viable alternative to the enzymatic poisoning caused by neoplastic drugs targeting the DNA-binding region. Though such inhibitors are currently unavailable for medical use, several promising molecules are under evaluation across various

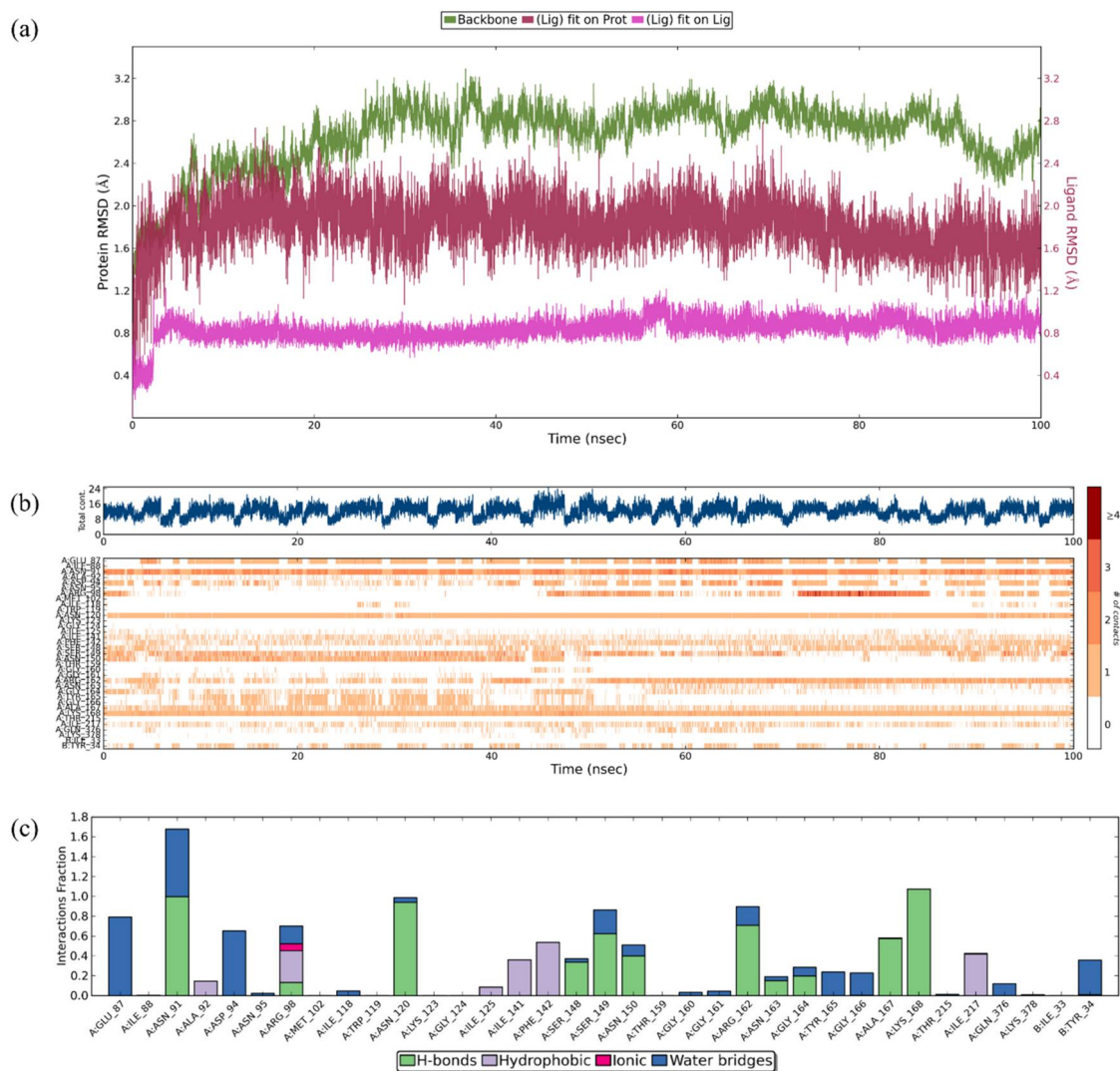


Figure 7. Results of 100 ns MD simulation of the hTopII α -scutellarin complex: (a) protein-ligand RMSD timeline, (b) protein-ligand contacts timeline and (c) protein-ligand contacts histogram.

stages of pre-clinical and clinical trials, some of which are NPs (Hevener et al., 2018; Liang et al., 2019). NPs have been used as novel precursors for drug development for decades and continue to inspire newer therapeutic scaffolds (Atanasov et al., 2015, Atanasov et al., 2021; Newman & Cragg, 2020). In recent times, several studies have utilised *in silico* tools for the screening of large NP databases in the hopes of identifying such novel lead molecules (Chikhale et al., 2021; Naik et al., 2020; Radaeva et al., 2020; Umashankar et al., 2021).

In this study, the ATPase domain of hTopII α (PDB ID: 1ZXM) was our chosen receptor, in which the NBDs of both protomers exhibited a similar ligand binding pattern. We targeted only chain A of the receptor and screened the NP database, NPASS. This was done due to the similarity in interaction fingerprints of both chains and most importantly it has been proven that the binding of a single ATP to the NBD of chain A was sufficient for inducing enzymatic cycling, by triggering dimerization and capture of the T-segment (Skouboe et al., 2003; Wei et al., 2005).

HTVS was carried out after removing the cofactor Mg $^{2+}$ since it was most likely to be brought in by the native ligand ATP. We supposed that strong NP binders which functioned independently of metal coordination stabilisation contributed by the Mg $^{2+}$ cation would exhibit better inhibitor potential. Therefore, we selected salvicine as a reference inhibitor for a fair comparison of binding affinity and binding mode, rather than the native ligand ATP.

Following HTVS, the top hundred hits were subjected to MM/GBSA analysis. As an important step in the stringent selection of best hits, critical residues involved in the recognition of ATP and especially the crucial LYS 378 (chosen through binding site analysis) were used for hit prioritisation. The LYS 378 was given importance since an ideal catalytic lead molecule should form a strong non-hydrolyzable bond with this residue and hold the QTK loop in place, and limit the movement of the TD. However, it should be noted that, though this trait was preferable, any strong binder which remained tightly bound to the NBD and blocked the cavity from subsequent ATP binding was also suitable as this would

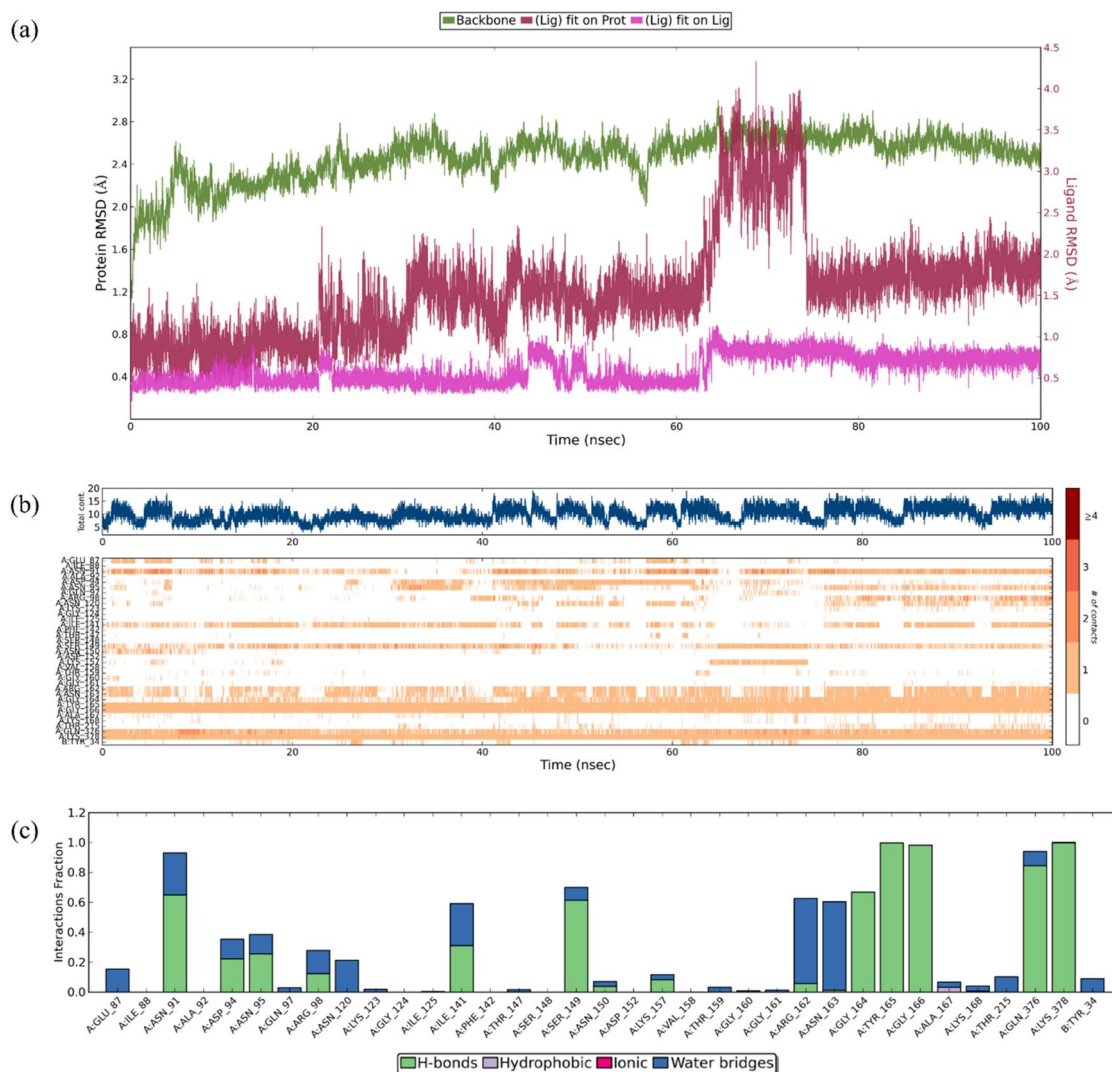


Figure 8. Results of 100 ns MD simulation of the hTopII α -linocaffeine complex: (a) protein-ligand RMSD timeline, (b) protein-ligand contacts timeline and (c) protein-ligand contacts histogram.

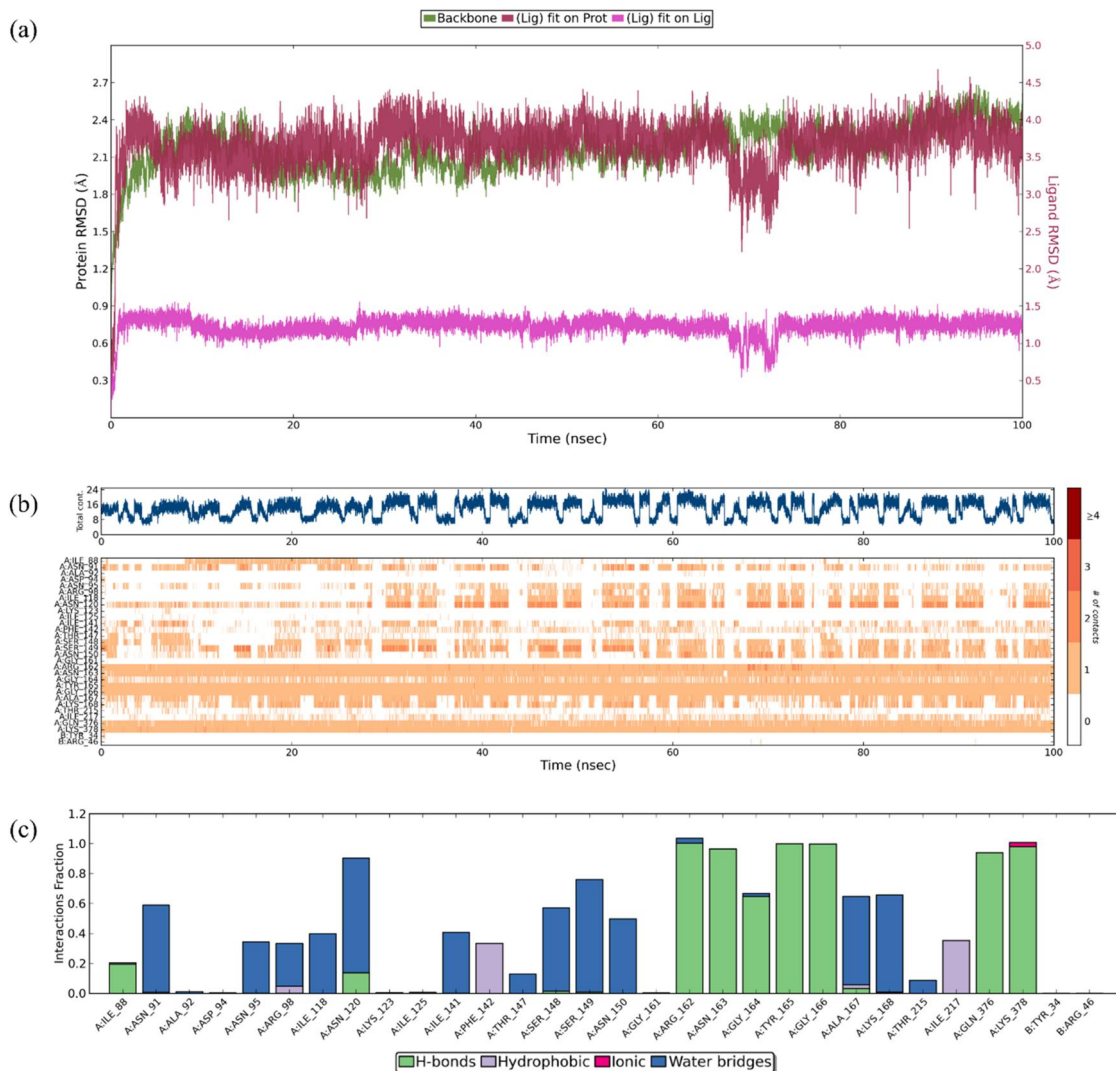


Figure 9. Results of 100 ns MD simulation of the hTOP11 α -caffeoylmalic acid complex: (a) protein-ligand RMSD timeline, (b) protein-ligand contacts timeline and (c) protein-ligand contacts histogram.

Table 3. Physicochemical and ADMET properties of salvicine and five NP ligands computed by Qikprop.

Ligand name	Salvicine	Longumoside B	Penicacid C	Scutellarin	Linocaffeine	Caffeoylmalic acid
Molecular weight (Da)	330.423	440.446	370.355	462.366	342.302	296.233
QPlogPo/w ^a	2.937	0.98	0.915	-0.328	-1.215	0.435
QPlogS ^b	-3.71	-1.883	-2.174	-2.844	-1.618	-2.233
QPPCaco ^c	1547.804	1.719	8.556	0.722	1.995	0.408
Human Oral Absorption ^d	3	1	2	1	1	1
% HOA ^e	100	36.896	48.986	0	12.238	22.523
RO5 violations ^f	0	0	0	2	1	0
RO3 violations ^g	0	2	2	2	1	1
QPlogHERG ^h	-4.228	-0.989	-1.901	-3.424	-2.546	-1.231
QPlogBB ⁱ	-0.616	-2.681	-2.391	-3.522	-3.093	-3.187
QPPMDCK ^j	793.251	0.822	3.663	0.253	0.759	0.174
CNS ^k	0	-2	-2	-2	-2	-2
QPlogKhsa ^l	0.079	-1.129	-0.702	-0.836	-1.192	-1.005

^aPredicted octanol/water partition coefficient: -2.0–6.5.

^bPredicted aqueous solubility: -6.5–0.5.

^cPredicted apparent Caco-2 cell permeability in nm/sec: <25 poor, >500 great.

^dPredicted qualitative human oral absorption: 1, 2, or 3 for low, medium, or high.

^ePredicted human oral absorption on a 0 to 100% scale: >80% is high, <25% is poor.

^fNo. of violations of Lipinski's rule of five, given by: Molecular weight < 500 Da, QPlogPo/w < 5, H-bond donors \leq 5, H-bond acceptors \leq 10.

^gNo. of violations of Jorgensen's rule of three, given by: QPlogS > -5.7, QPPCaco > 22 nm/s, No. of Primary Metabolites < 7.

^hPredicted IC₅₀ value for blockage of HERG K⁺ channel: concern below -5.

ⁱPredicted brain/blood partition coefficient: -3.0–1.2.

^jPredicted apparent MDCK cell permeability in nm/sec: <25 poor, >500 great.

^kPredicted central nervous system activity: -2 (inactive), +2 (active).

^lPrediction of binding to human serum albumin: -1.5–1.5.

Table 4. ADMET properties of salvicine and five NP ligands computed by SwissADME.

Ligand name	Salvicine	Longumoside B	Penicacid C	Scutellarin	Linocaffein	Caffeoylmalic acid
GI absorption	High	Low	Low	Low	Low	High
BBB permeant	Yes	No	No	No	No	No
P-gp substrate	No	Yes	Yes	Yes	No	No
CYP1A2 inhibitor	No	No	No	No	No	No
CYP2C19 inhibitor	Yes	No	No	No	No	No
CYP2C9 inhibitor	No	No	No	No	No	No
CYP2D6 inhibitor	No	No	No	No	No	No
CYP3A4 inhibitor	No	No	No	No	No	No
PAINS alerts	2 (imine_one_A, quinone_D)	0	0	1 (catechol_A)	0	1 (catechol_A)
Brenk alerts	1 (diketo_group)	0	0	1 (catechol)	1 (michael_acceptor_1)	2 (catechol, michael_acceptor_1)
Leadlikeness violations ^a	0	2	1	1	0	0
Synthetic Accessibility ^b	4.13	5.24	4.14	5.12	4.4	3.15

^aNo. of violations of leadlikeness rules, given by: $250 \leq \text{Molecular weight} \leq 350$, $\text{XLOGP} \leq 3.5$, No. of rotatable bonds ≤ 7 .

^bEase of synthesis: 1 (very easy) to 10 (very difficult).

Table 5. ProTox II results for salvicine and five NP ligands.

Properties	Prediction					
	Salvicine 1600 mg/kg	Longumoside B 5000 mg/kg	Penicacid C 352 mg/kg	Scutellarin 5000 mg/kg	Linocaffein 4000 mg/kg	Caffeoylmalic acid 3800 mg/kg
Predicted LD50 ^a	4	5	4	5	5	5
Predicted Toxicity Class ^a	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive
Hepatotoxicity ^b	Active	Inactive	Inactive	Active	Inactive	Inactive
Carcinogenicity ^c	Inactive	Inactive	Inactive	Inactive	Active	Active
Immunotoxicity ^c	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive
Mutagenicity ^c	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive
Cytotoxicity ^c	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive
Aryl hydrocarbon Receptor (AhR) ^d	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive
Androgen Receptor (AR) ^d	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive
Androgen Receptor Ligand Binding Domain (AR-LBD) ^d	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive
Aromatase ^d	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive
Estrogen Receptor Alpha (ER) ^d	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive
Estrogen Receptor Ligand Binding Domain (ER-LBD) ^d	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive
Peroxisome Proliferator Activated Receptor Gamma (PPAR-Gamma) ^d	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive
Nuclear factor (erythroid-derived 2)-like 2 /antioxidant responsive element (nrf2/ARE) ^e	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive
Heat shock factor response element (HSE) ^e	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive
Mitochondrial Membrane Potential (MMP) ^e	Active	Inactive	Inactive	Inactive	Inactive	Inactive
Phosphoprotein (Tumor Suppressor) p53 ^e	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive
ATPase family AAA domain-containing protein 5 (ATAD5) ^e	Inactive	Inactive	Inactive	Inactive	Inactive	Inactive

Note: Colour coding of predicted properties is given as green for inactive, red for active, light colours for low confidence score and dark colours for high confidence score.

^aBased on GHS categories for the median lethal dose.

^bOrgan toxicity.

^cToxicity endpoints.

^dTox21-Nuclear receptor signalling pathways.

^eTox21-Stress response pathways.

effectively prevent enzymatic cycling as well. Moreover, the slow turnover would in turn influence the DNA-dependent metabolic processes thus affecting the proliferation rate of the cells. In this regard, five NP hits—longumoside B, penicacid C, scutellarin, linocaffein and caffeoylmalic acid, were selected for subsequent analysis.

The XP docking scores, binding free energy results and binding mode analysis indicated salvicine having the lowest binding affinity and forming the least number of H-bonds, thus making the NPs better candidates. 100 ns MD simulations of the five NP and salvicine docked complexes were carried out for gaining insights into the stability and interaction pattern exhibited by them within the NBD. MD trajectory analysis confirmed that salvicine was indeed the weakest in forming stable bonds in the NBD. In the case of interactions as well as overall structural parameters during

the simulations, longumoside B and scutellarin fared best among the NP hits.

ADMET analysis of all six small molecules revealed that salvicine had higher potency of oral bioavailability and absorption through the GI tract. But it was also predicted to be a CYP2C19 inhibitor, which was undesirable behaviour in GI-absorbent leads. On the other hand, NPs often fall beyond (Lipinski's) rule of five space (bRO5) and break conventional rules of bioavailability (Doak et al., 2014). Despite this, it was observed that longumoside B and penicacid C, the more desirable NPs, fulfilled RO5 criteria. Salvicine was also predicted to have possible CNS activity and was predicted as BBB permeant, which are undesirable features in non-neurologic drug leads. Such qualities were however absent in the NPs. Most importantly, salvicine possessed unsafe and unwanted substructures, indicated by the PAINS and Brenk

filters, thus making it a less-than-satisfactory drug precursor. Once again longumoside B and penicacid C surpassed this quality of salvicine and turned out to be free from promiscuous fragments. Longumoside B also qualified all toxicological endpoints with high confidence scores, closely followed by penicacid C.

Most tools predicting drug-likeness and leadlikeness prioritise orally bioavailable drugs. However, such small molecules are not always desirable owing to the other complications arising from small size, as exhibited by salvicine. NPs, such as the screened NPASS molecules, though unappealing from an oral bioavailability perspective are nonetheless more stable binders as evidenced by the virtual screening and MD simulation results. In recent times, the bRO5 space has gained prominence since an unexplored space of diverse NPs is found there. The bioavailability of such bRO5 NPs can be enhanced by improving formulations, using targeted delivery mechanisms or using other modes of administration. Though such small molecules bring along many complexities, it is worth exploring such a chemical space due to the increasing requirement for novel therapeutic precursor molecules (Caron et al., 2021; Doak et al., 2014).

At the end of this study, the multilevel analysis of the various characteristics of the chosen natural inhibitors revealed that longumoside B and penicacid C proved to be the best hits against the NBD of hTopII α . Both these NPs showed an easy to moderate range of synthetic accessibility, which is suitable for any eventual downstream clinical applications. However, in terms of natural occurrence, penicacid C is currently only found in a rare species of fungi available in the South China Sea. Whereas, longumoside B is readily available for isolation from an abundantly found and commonly cultivated plant, *Piper longum*, making it a more suitable drug development candidate. Therefore, we conclude that longumoside B is a novel CI candidate for the ATPase domain of hTopII α . We suggest further *in vitro* validation to be carried out for understanding the putative therapeutic role of longumoside B as an anticancer precursor molecule targeting hTopII α .

5. Conclusion

hTopII α is a well-known, primary anticancer target for which several inhibitors have already been and continue to be developed for the sake of improving therapeutic and prognostic outcomes. Some of the widely used inhibitors of hTopII α have NP origins, for example, etoposide and teniposide are plant-derived podophyllotoxin analogues whereas daunorubicin isolated from *Streptomyces* species and its derivative doxorubicin are anthracyclines. Due to this reason, a natural product database was screened in this study for the identification of novel NP precursors that will serve as starting points for developing new catalytic inhibitors of hTopII α . Following multilevel prioritisation of HTVS results and MD simulations, Longumoside B, a fatty acyl glycoside found in *Piper longum*, was selected as a novel NP catalytic inhibitor of the ATPase domain of hTopII α . Experimental verification of this study outcome is necessary to support results

observed *in silico*. This would corroborate the potential of Longumoside B as a prospective inhibitor of hTopII α and also pave the way for the development of a new generation of NP-derived anticancer agents.

Disclosure statement

The authors declare that there are no conflicts of interest.

Funding

The authors gratefully acknowledge the funding provided by the Indian Council of Medical Research, New Delhi (ICMR grant No. ISRM/12(25)/2019) for carrying out this study

References

- Akimitsu, N., Adachi, N., Hirai, H., Hossain, M. S., Hamamoto, H., Kobayashi, M., Aratani, Y., Koyama, H., & Sekimizu, K. (2003). Enforced cytokinesis without complete nuclear division in embryonic cells depleting the activity of DNA topoisomerase II α . *Genes to Cells: Devoted to Molecular & Cellular Mechanisms*, 8(4), 393–402. <https://doi.org/10.1046/j.1365-2443.2003.00643.x>
- Alonso, H., Bliznyuk, A. A., & Gready, J. E. (2006). Combining docking and molecular dynamic simulations in drug design. *Medicinal Research Reviews*, 26(5), 531–568. <https://doi.org/10.1002/med.20067>
- Atanasov, A. G., Waltenberger, B., Pferschy-Wenzig, E.-M., Linder, T., Wawrosch, C., Uhrin, P., Temml, V., Wang, L., Schwaiger, S., Heiss, E. H., Rollinger, J. M., Schuster, D., Breuss, J. M., Bochkov, V., Mihovilovic, M. D., Kopp, B., Bauer, R., Dirsch, V. M., & Stuppner, H. (2015). Discovery and resupply of pharmacologically active plant-derived natural products: A review. *Biotechnology Advances*, 33(8), 1582–1614. <https://doi.org/10.1016/j.biotechadv.2015.08.001>
- Atanasov, A. G., Zotchev, S. B., Dirsch, V. M., & Supuran, C. T., the International Natural Product Sciences Taskforce. (2021). Natural products in drug discovery: Advances and opportunities. *Nature Reviews Drug Discovery*, 20(3), 200–216. <https://doi.org/10.1038/s41573-020-00114-z>
- Austin, C. A., Lee, K. C., Swan, R. L., Khazeem, M. M., Manville, C. M., Cridland, P., Treumann, A., Porter, A., Morris, N. J., & Cowell, I. G. (2018). TOP2B: The first thirty years. *International Journal of Molecular Sciences*, 19(9), 2765. <https://doi.org/10.3390/ijms19092765>
- Azarova, A. M., Lyu, Y. L., Lin, C.-P., Tsai, Y.-C., Lau, J. Y.-N., Wang, J. C., & Liu, L. F. (2007). Roles of DNA topoisomerase II isozymes in chemotherapy and secondary malignancies. *Proceedings of the National Academy of Sciences of the United States of America*, 104(26), 11014–11019. <https://doi.org/10.1073/pnas.0704002104>
- Banerjee, P., Eckert, A. O., Schrey, A. K., & Preissner, R. (2018). ProTox-II: A webserver for the prediction of toxicity of chemicals. *Nucleic Acids Research*, 46(W1), W257–W263. <https://doi.org/10.1093/nar/gky318>
- Bax, B. D., Chan, P. F., Eggleston, D. S., Fosberry, A., Gentry, D. R., Gorrec, F., Giordano, I., Hann, M. M., Hennessy, A., Hibbs, M., Huang, J., Jones, E., Jones, J., Brown, K. K., Lewis, C. J., May, E. W., Saunders, M. R., Singh, O., Spitzfaden, C. E., ... Gwynn, M. N. (2010). Type IIA topoisomerase inhibition by a new class of antibacterial agents. *Nature*, 466(7309), 935–940. <https://doi.org/10.1038/nature09197>
- Bendsen, S., Oestergaard, V. H., Skouboe, C., Brinch, M., Knudsen, B. R., & Andersen, A. H. (2009). The QTK loop is essential for the communication between the N-terminal ATPase domain and the central cleavage – ligation region in human topoisomerase II α . *Biochemistry*, 48(27), 6508–6515. <https://doi.org/10.1021/bi9005978>
- Bergant, K., Janezic, M., & Perdih, A. (2018). Bioassays and *in silico* methods in the identification of human DNA topoisomerase II α inhibitors. *Current Medicinal Chemistry*, 25(28), 3286–3318. <https://doi.org/10.2174/0929867325666180306165725>
- Bergerat, A., De Massy, B., Gadelle, D., Varoutas, P. C., Nicolas, A., & Forterre, P. (1997). An atypical topoisomerase II from archaea with

- implications for meiotic recombination. *Nature*, 386(6623), 414–417. <https://doi.org/10.1038/386414a0>
- Bollimpelli, V. S., Dholaniya, P. S., & Kondapi, A. K. (2017). Topoisomerase II β and its role in different biological contexts. *Archives of Biochemistry and Biophysics*, 633, 78–84. <https://doi.org/10.1016/j.abb.2017.06.021>
- Bowers, K. J., Chow, D. E., Xu, H., Dror, R. O., Eastwood, M. P., Gregersen, B. A., Klepeis, J. L., Kolossvary, I., Moraes, M. A., Sacerdoti, F. D., Salmon, J. K., Shan, Y., & Shaw, D. E. (2006). *Scalable algorithms for molecular dynamics simulations on commodity clusters*. ACM/IEEE SC 2006 Conference (SC'06), 43–43. <https://doi.org/10.1109/SC.2006.54>
- Budzianowski, J. (1990). Caffeoylmalic and two pyrrole acids from *Parietaria officinalis*. *Phytochemistry*, 29(10), 3299–3301. [https://doi.org/10.1016/0031-9422\(90\)80203-S](https://doi.org/10.1016/0031-9422(90)80203-S)
- Capranico, G., Tinelli, S., Austin, C. A., Fisher, M. L., & Zunino, F. (1992). Different patterns of gene expression of topoisomerase II isoforms in differentiated tissues during murine development. *Biochimica et Biophysica Acta*, 1132(1), 43–48. [https://doi.org/10.1016/0167-4781\(92\)90050-A](https://doi.org/10.1016/0167-4781(92)90050-A)
- Caron, G., Kihlberg, J., Goetz, G., Ratkova, E., Poongavanam, V., & Ermondi, G. (2021). Steering new drug discovery campaigns: Permeability, solubility, and physicochemical properties in the bRo5 chemical space. *ACS Medicinal Chemistry Letters*, 12(1), 13–23. <https://doi.org/10.1021/acsmchemlett.0c00581>
- Champoux, J. J. (2001). DNA topoisomerases: Structure, function, and mechanism. *Annual Review of Biochemistry*, 70(1), 369–413. <https://doi.org/10.1146/annurev.biochem.70.1.369>
- Chen, Z., Zheng, Z., Huang, H., Song, Y., Zhang, X., Ma, J., Wang, B., Zhang, C., & Ju, J. (2012). Penicacids A-C, three new mycophenolic acid derivatives and immunosuppressive activities from the marine-derived fungus *Penicillium* sp. SOF07. *Bioorganic & Medicinal Chemistry Letters*, 22(9), 3332–3335. <https://doi.org/10.1016/j.bmcl.2012.02.106>
- Chikhale, R. V., Gupta, V. K., Eldesoky, G. E., Wabaidur, S. M., Patil, S. A., & Islam, M. A. (2021). Identification of potential anti-TMPRSS2 natural products through homology modelling, virtual screening and molecular dynamics simulation studies. *Journal of Biomolecular Structure and Dynamics*, 39(17), 6660–6675. <https://doi.org/10.1080/07391102.2020.1798813>
- Coss, A., Tosetto, M., Fox, E. J., Sapetto-Rebow, B., Gorman, S., Kennedy, B. N., Lloyd, A. T., Hyland, J. M., O'Donoghue, D. P., Sheahan, K., Leahy, D. T., Mulcahy, H. E., & O'Sullivan, J. N. (2009). Increased topoisomerase II α expression in colorectal cancer is associated with advanced disease and chemotherapeutic resistance via inhibition of apoptosis. *Cancer Letters*, 276(2), 228–238. <https://doi.org/10.1016/j.canlet.2008.11.018>
- Daina, A., Michielin, O., & Zoete, V. (2017). SwissADME: A free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Scientific Reports*, 7(1), 42717. <https://doi.org/10.1038/srep42717>
- Delgado, J. L., Hsieh, C.-M., Chan, N.-L., & Hiasa, H. (2018). Topoisomerases as anticancer targets. *The Biochemical Journal*, 475(2), 373–398. <https://doi.org/10.1042/BCJ20160583>
- Deng, Z., Chuaqui, C., & Singh, J. (2004). Structural Interaction Fingerprint (SIFt): A novel method for analyzing three-dimensional protein–ligand binding interactions. *Journal of Medicinal Chemistry*, 47(2), 337–344. <https://doi.org/10.1021/jm030331x>
- Depowski, P. L., Rosenthal, S. I., Brien, T. P., Stylos, S., Johnson, R. L., & Ross, J. S. (2000). Topoisomerase II α expression in breast cancer: Correlation with outcome variables. *Modern Pathology*, 13(5), 542–547. <https://doi.org/10.1038/modpathol.3880094>
- Deweese, J. E., & Osheroff, N. (2010). The use of divalent metal ions by type II topoisomerases. *Metalomics: Integrated Biometal Science*, 2(7), 450–459. <https://doi.org/10.1039/c003759a>
- Djombou Feunang, Y., Eisner, R., Knox, C., Chepelev, L., Hastings, J., Owen, G., Fahy, E., Steinbeck, C., Subramanian, S., Bolton, E., Greiner, R., & Wishart, D. S. (2016). ClassyFire: Automated chemical classification with a comprehensive, computable taxonomy. *Journal of Cheminformatics*, 8(1), 61. <https://doi.org/10.1186/s13321-016-0174-y>
- Doak, B. C., Over, B., Giordanetto, F., & Kihlberg, J. (2014). Oral druggable space beyond the rule of 5: Insights from drugs and clinical candidates. *Chemistry & Biology*, 21(9), 1115–1142. <https://doi.org/10.1016/j.chembiol.2014.08.013>
- Gao, H., Huang, K.-C., Yamasaki, E. F., Chan, K. K., Chohan, L., & Snapka, R. M. (1999). XK469, a selective topoisomerase II β poison. *Proceedings of the National Academy of Sciences of the United States of America*, 96(21), 12168–12173. <https://doi.org/10.1073/pnas.96.21.12168>
- Gout, E., Rébeillé, F., Douce, R., & Bligny, R. (2014). Interplay of Mg 2+, ADP, and ATP in the cytosol and mitochondria: Unravelling the role of Mg 2+ in cell respiration. *Proceedings of the National Academy of Sciences of the United States of America*, 111(43), E4560–E4567. <https://doi.org/10.1073/pnas.1406251111>
- Halgren, T. A., Murphy, R. B., Friesner, R. A., Beard, H. S., Frye, L. L., Pollard, W. T., & Banks, J. L. (2004). Glide: A new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *Journal of Medicinal Chemistry*, 47(7), 1750–1759. <https://doi.org/10.1021/jm030644s>
- Hevener, K., Verstak, T. A., Lutat, K. E., Riggsbee, D. L., & Mooney, J. W. (2018). Recent developments in topoisomerase-targeted cancer chemotherapy. *Acta Pharmaceutica Sinica B*, 8(6), 844–861. <https://doi.org/10.1016/j.apsb.2018.07.008>
- Hou, G.-X., Liu, P., Yang, J., & Wen, S. (2017). Mining expression and prognosis of topoisomerase isoforms in non-small-cell lung cancer by using OncoPrint and Kaplan–Meier plotter. *PLoS One*, 12(3), e0174515. <https://doi.org/10.1371/journal.pone.0174515>
- Hu, C.-X., Zuo, Z.-L., Xiong, B., Ma, J.-G., Geng, M.-Y., Lin, L.-P., Jiang, H.-L., & Ding, J. (2006). Salvicine functions as novel topoisomerase II poison by binding to ATP pocket. *Molecular Pharmacology*, 70(5), 1593–1601. <https://doi.org/10.1124/mol.106.027714>
- Janežič, M., Pogorelčnik, B., Brvar, M., Solmajer, T., & Perdih, A. (2017). 3-substituted-1H-indazoles as catalytic inhibitors of the human DNA topoisomerase II α . *ChemistrySelect*, 2(1), 480–488. <https://doi.org/10.1002/slct.201601554>
- Jiang, Z.-Y., Liu, W.-F., Zhang, X.-M., Luo, J., Ma, Y.-B., & Chen, J.-J. (2013). Anti-HBV active constituents from *Piper longum*. *Bioorganic & Medicinal Chemistry Letters*, 23(7), 2123–2127. <https://doi.org/10.1016/j.bmcl.2013.01.118>
- Kenig, S., Faoro, V., Bourkoura, E., Podergajs, N., Ius, T., Vindigni, M., Skrap, M., Lah, T., Cesselli, D., Storici, P., & Vindigni, A. (2016). Topoisomerase II β mediates the resistance of glioblastoma stem cells to replication stress-inducing drugs. *Cancer Cell International*, 16(1), 58. <https://doi.org/10.1186/s12935-016-0339-9>
- Klosterman, H. J., & Muggli, R. Z. (1959). The glucosides of flaxseed. *Journal of the American Chemical Society*, 81(9), 2188–2191. https://doi.org/10.1021/JA01518A043/ASSET/JA01518A043.FP.PNG_V03
- Lee, J. H., & Berger, J. M. (2019). Cell cycle-dependent control and roles of DNA topoisomerase II. *Genes*, 10(11), 859. <https://doi.org/10.3390/genes10110859>
- Liang, X., Wu, Q., Luan, S., Yin, Z., He, C., Yin, L., Zou, Y., Yuan, Z., Li, L., Song, X., He, M., Lv, C., & Zhang, W. (2019). A comprehensive review of topoisomerase inhibitors as anticancer agents in the past decade. *European Journal of Medicinal Chemistry*, 171, 129–168. <https://doi.org/10.1016/j.ejmech.2019.03.034>
- Maia, E. H. B., Assis, L. C., de Oliveira, T. A., da Silva, A. M., & Taranto, A. G. (2020). Structure-based virtual screening: From classical to artificial intelligence. *Frontiers in Chemistry*, 8, 343. <https://doi.org/10.3389/fchem.2020.00343>
- Martyna, G. J., Tobias, D. J., & Klein, M. L. (1994). Constant pressure molecular dynamics algorithms. *The Journal of Chemical Physics*, 101(5), 4177–4189. <https://doi.org/10.1063/1.467468>
- McClendon, A. K., & Osheroff, N. (2007). DNA topoisomerase II, genotoxicity, and cancer. *Mutation Research*, 623(1–2), 83–97. <https://doi.org/10.1016/j.mrfmmm.2007.06.009>
- Naik, B., Gupta, N., Ojha, R., Singh, S., Prajapati, V. K., & Prusty, D. (2020). High throughput virtual screening reveals SARS-CoV-2 multi-target binding natural compounds to lead instant therapy for COVID-19 treatment. *International Journal of Biological Macromolecules*, 160, 1–17. <https://doi.org/10.1016/j.ijbiomac.2020.05.184>

- Newman, D. J., & Cragg, G. M. (2020). Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *Journal of Natural Products*, 83(3), 770–803. <https://doi.org/10.1021/acs.jnatprod.9b01285>
- Nitiss, J. L. (2009a). DNA topoisomerase II and its growing repertoire of biological functions. *Nature Reviews. Cancer*, 9(5), 327–337. <https://doi.org/10.1038/nrc2608>
- Nitiss, J. L. (2009b). Targeting DNA topoisomerase II in cancer chemotherapy. *Nature Reviews. Cancer*, 9(5), 338–350. <https://doi.org/10.1038/nrc2607>
- Osheroff, N. (1987). Role of the divalent cation in topoisomerase II mediated reactions. *Biochemistry*, 26(20), 6402–6406. <https://doi.org/10.1021/bi00394a015>
- Pogorelnik, B., Brvar, M., Žegura, B., Filipič, M., Solmajer, T., & Perdih, A. (2015). Discovery of mono- and disubstituted 1H-pyrazolo[3,4]pyrimidines and 9H-purines as catalytic inhibitors of human DNA topoisomerase II α . *ChemMedChem*, 10(2), 345–359. <https://doi.org/10.1002/cmdc.201402459>
- Pommier, Y., Leo, E., Zhang, H., & Marchand, C. (2010). DNA topoisomerases and their poisoning by anticancer and antibacterial drugs. *Chemistry & Biology*, 17(5), 421–433. <https://doi.org/10.1016/j.chembiol.2010.04.012>
- Pommier, Y., Sun, Y., Huang, S. N., & Nitiss, J. L. (2016). Roles of eukaryotic topoisomerases in transcription, replication and genomic stability. *Nature Reviews. Molecular Cell Biology*, 17(11), 703–721. <https://doi.org/10.1038/nrm.2016.111>
- Radaeva, M., Dong, X., & Cherkasov, A. (2020). The use of methods of computer-aided drug discovery in the development of topoisomerase II inhibitors: Applications and future directions. *Journal of Chemical Information and Modeling*, 60(8), 3703–3721. <https://doi.org/10.1021/acs.jcim.0c00325>
- Salentin, S., Schreiber, S., Haupt, V. J., Adasme, M. F., & Schroeder, M. (2015). PLIP: Fully automated protein–ligand interaction profiler. *Nucleic Acids Research*, 43(W1), W443–W447. <https://doi.org/10.1093/nar/gkv315>
- Schaefer-Klein, J. L., Murphy, S. J., Johnson, S. H., Vasmatzis, G., & Kovtun, I. V. (2015). Topoisomerase 2 alpha cooperates with androgen receptor to contribute to prostate cancer progression. *PLoS One*, 10(11), e0142327. <https://doi.org/10.1371/journal.pone.0142327>
- Schoeffler, A. J., & Berger, J. M. (2008). DNA topoisomerases: Harnessing and constraining energy to govern chromosome topology. *Quarterly Reviews of Biophysics*, 41(1), 41–101. <https://doi.org/10.1017/S003358350800468X>
- Schrödinger LLC. (2019). *Schrödinger Suite 2019-3: MacroModel, Glide, SiteMap, Maestro*.
- Shelley, J. C., Chollet, A., Frye, L. L., Greenwood, J. R., Timlin, M. R., & Uchimaya, M. (2007). Epik: A software program for pK_a prediction and protonation state generation for drug-like molecules. *Journal of Computer-Aided Molecular Design*, 21(12), 681–691. <https://doi.org/10.1007/s10822-007-9133-z>
- Skouboe, C., Bjergbaek, L., Oestergaard, V. H., Larsen, M. K., Knudsen, B. R., & Andersen, A. H. (2003). A human topoisomerase II α heterodimer with only one ATP binding site can go through successive catalytic cycles. *The Journal of Biological Chemistry*, 278(8), 5768–5774. <https://doi.org/10.1074/jbc.M210332200>
- Tuckerman, M., Berne, B. J., & Martyna, G. J. (1992). Reversible multiple time scale molecular dynamics. *The Journal of Chemical Physics*, 97(3), 1990–2001. <https://doi.org/10.1063/1.463137>
- Turley, H., Comley, M., Houlbrook, S., Nozaki, N., Kikuchi, A., Hickson, I., Gatter, K., & Harris, A. (1997). The distribution and expression of the two isoforms of DNA topoisomerase II in normal and neoplastic human tissues. *British Journal of Cancer*, 75(9), 1340–1346. <https://doi.org/10.1038/bjc.1997.227>
- Umashankar, V., Deshpande, S. H., Hegde, H. V., Singh, I., & Chattopadhyay, D. (2021). Phytochemical moieties from Indian traditional medicine for targeting dual hotspots on SARS-CoV-2 spike protein: An integrative in-silico approach. *Frontiers in Medicine*, 8, 672629. <https://doi.org/10.3389/fmed.2021.672629>
- van der Zee, A. G. J., de Vries, E. G. E., Hollema, H., Kaye, S. B., Brown, R., & Keith, W. N. (1994). Molecular analysis of the topoisomerase II α gene and its expression in human ovarian cancer. *Annals of Oncology*, 5(1), 75–81. <https://doi.org/10.1093/oxfordjournals.annonc.a058700>
- Vávrová, A., & Šimůnek, T. (2012). DNA topoisomerase II β : A player in regulation of gene expression and cell differentiation. *The International Journal of Biochemistry & Cell Biology*, 44(6), 834–837. <https://doi.org/10.1016/j.biocel.2012.03.005>
- Wang, E., Sun, H., Wang, J., Wang, Z., Liu, H., Zhang, J. Z. H., & Hou, T. (2019). End-point binding free energy calculation with MM/PBSA and MM/GBSA: Strategies and applications in drug design. *Chemical Reviews*, 119(16), 9478–9508. <https://doi.org/10.1021/acs.chemrev.9b00055>
- Wang, J. C. (2002). Cellular roles of DNA topoisomerases: A molecular perspective. *Nature Reviews. Molecular Cell Biology*, 3(6), 430–440. <https://doi.org/10.1038/nrm831>
- Wang, L., & Ma, Q. (2018). Clinical benefits and pharmacology of scutellarin: A comprehensive review. *Pharmacology & Therapeutics*, 190, 105–127. <https://doi.org/10.1016/j.pharmthera.2018.05.006>
- Wei, H., Ruthenburg, A. J., Bechis, S. K., & Verdine, G. L. (2005). Nucleotide-dependent domain movement in the ATPase domain of a human type IIA DNA topoisomerase. *The Journal of Biological Chemistry*, 280(44), 37041–37047. <https://doi.org/10.1074/jbc.M506520200>
- Weston, J. (2009). *Biochemistry of magnesium*. In *PATAI'S chemistry of functional groups*. John Wiley & Sons, Ltd. <https://doi.org/10.1002/9780470682531.pat0407>
- Woessner, R. D., Mattern, M. R., Mirabelli, C. K., Johnson, R. K., & Drake, F. H. (1991). Proliferation- and cell cycle-dependent differences in expression of the 170 kilodalton and 180 kilodalton forms of topoisomerase II in NIH-3T3 cells. *Cell Growth & Differentiation : The Molecular Biology Journal of the American Association for Cancer Research*, 2(4), 209–214. <http://www.ncbi.nlm.nih.gov/pubmed/1651102>
- Wolf, F. I., & Cittadini, A. (2003). Chemistry and biochemistry of magnesium. *Molecular Aspects of Medicine*, 24(1–3), 3–9. [https://doi.org/10.1016/S0098-2997\(02\)00087-0](https://doi.org/10.1016/S0098-2997(02)00087-0)
- Xu, B., Ding, J., Chen, K.-X., Miao, Z.-H., Huang, H., Liu, H., & Luo, X.-M. (2012). Advances in cancer chemotherapeutic drug research in china. In *Recent advances in cancer research and therapy* (1st ed., Vol. 54920515, pp. 287–350). Elsevier. <https://doi.org/10.1016/B978-0-12-397833-2.00012-1>
- Zeng, X., Zhang, P., He, W., Qin, C., Chen, S., Tao, L., Wang, Y., Tan, Y., Gao, D., Wang, B., Chen, Z., Chen, W., Jiang, Y. Y., & Chen, Y. Z. (2018). NPASS: Natural product activity and species source database for natural product research, discovery and tool development. *Nucleic Acids Research*, 46(D1), D1217–D1222. <https://doi.org/10.1093/nar/gkx1026>
- Zhang, J. S., Ding, J., Tang, Q. M., Li, M., Zhao, M., Lu, L. J., Chen, L. J., & Yuan, S. T. (1999). Synthesis and antitumor activity of novel diterpene-quinone salvicine and the analogs. *Bioorganic and Medicinal Chemistry Letters*, 9(18), 2731–2736. [https://doi.org/10.1016/S0960-894X\(99\)00472-2](https://doi.org/10.1016/S0960-894X(99)00472-2)
- Zhao, Q., Li, H., Zhu, L., Hu, S., Xi, X., Liu, Y., Liu, J., & Zhong, T. (2020). Bioinformatics analysis shows that TOP2A functions as a key candidate gene in the progression of cervical cancer. *Biomedical Reports*, 13(4), 1–10. <https://doi.org/10.3892/br.2020.1328>