Molecular Docking, Validation, Dynamics Simulations, and Pharmacokinetic Prediction of Phytochemicals Isolated From *Croton dichogamus* Against the HIV-1 Reverse Transcriptase

Bioinformatics and Biology Insights
Volume 16: 1–20
© The Author(s) 2022
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/11779322221125605
SAGE

Ermias Mergia Terefe[1] and Arabinda Ghosh[2]

[1]Department of Pharmacology and Pharmacognosy, School of Pharmacy and Health Sciences, United States International University-Africa, Nairobi, Kenya. [2]Microbiology Division, Department of Botany, Gauhati University, Guwahati, India.

ABSTRACT: The human immunodeficiency virus (HIV) infection and the associated acquired immune deficiency syndrome (AIDS) remain global challenges even after decades of successful treatment, with eastern and southern Africa still bearing the highest burden of disease. Following a thorough computational study, we report top 10 phytochemicals isolated from *Croton dichogamus* as potent reverse transcriptase inhibitors. The pentacyclic triterpenoid, aleuritolic acid (L12) has displayed best docking pose with binding energy of -8.48 kcal/mol and Ki of 0.61 μM making it superior in binding efficiency when compared to all docked compounds including the FDA-approved drugs. Other phytochemicals such as crotoxide A, crothalimene A, crotodichogamoin B and crotonolide E have also displayed strong binding energies. These compounds could further be investigated as potential antiretroviral medication.

KEYWORDS: HIV, *Croton dichogamus,* molecular docking, molecular dynamic simulation

RECEIVED: July 24, 2022. **ACCEPTED:** August 24, 2022.

TYPE: Original Research Article

FUNDING: The author(s) received no financial support for the research, authorship, and/or publication of this article.

DECLARATION OF CONFLICTING INTERESTS: The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CORRESPONDING AUTHOR: Ermias Mergia Terefe, Department of Pharmacology and Pharmacognosy, School of Pharmacy and Health Sciences, United States International University-Africa, P. O. BOX 14634 - 00800 Nairobi, Kenya. Emails: eterefe@usiu.ac.ke; mergiae@gmail.com

Introduction

Globally more than 75.7 million people have been infected and more than 33 million died since the HIV pandemic. Currently, 38 million people are living with the virus, out of which 54% (20.7 million) live in southern and East African region. By the end of 2020, the global death by the epidemic has reached more than half a million, where more than 300,000 of these deaths occurred in southern and East African region. In 2020, over 1.7 million people were infected by the virus.[1] In southern and East African region, more than 25% (5 million) people do not have access to the treatment.[1]

Phytochemicals isolated from natural products are important sources of lead compounds for the treatment of HIV/AIDS and other viral diseases. A number of natural products were investigated as promising phytochemicals for treatment of HIV including baicalin (a flavonoid),[2] calanolides (coumarins),[3] betulinic acid (a triterpene),[4,5] polycitone A (an alkaloid),[6] lithospermic acid, sulphated polysaccharides, cyanovirin-N,[7] pokeweed antiviral protein[8] and alpha-trichobitacin (proteins)[9]

The *Croton* genus has gained attention of many researchers for their potential source of bioactive compounds against HIV. Among the *Croton* genus, *Croton dichogamus* has wide ethnomedicinal use in East African countries including treatment of fever, stomach illness, respiratory disorders, malaria, impotence and infertility.[10-14] The objective of this study was to elucidate and analyze the binding mode of phytochemicals isolated from *C. dichogamus* in the active site of reverse transcriptase (PDB ID: **1REV**) and to compare results with the FDA-approved antiretroviral drugs such as delaviridine (DLV), zidovudine (AZT), nevirapine (NVP) and abacavir (ABC). The docking study gives a new insight into the investigation of molecular interaction between the phytochemical compounds and allows us subsequently to select one or more best active compounds that could be synthesized for an *in vitro* test in a future study.

Materials and Methods

Computational tools

Docking studies were performed using AutoDock 4.2.6 (http://autodock.scripps.edu/).[15-17] Discovery Studio Visualizer 3.1 Studio (Version 3.5, Accelrys Software Inc. Accelrys, San Diego, CA, USA) was used to visualize the protein ligand interactions.[18]

Ligand selection and preparation

The two dimensional structures of 24 phytochemicals isolated from *Croton dichogamus* were obtained from literature and chemical structure databases (Table S3 in Supplementary).[19,20] The ChemDraw 19.1 software was used to draw the 2D structures of compounds. For purpose of comparison, the structures of 4 FDA-approved drugs (abacavir, delaviridine, nevirapine and zidovudine) were obtained from pubchem database. The structures for ligands (phytochemicals and ARV's) were optimized for energy minimization using the MMFF94 force field[21] and were subsequently converted to protein data bank (pdb) format using Discovery studio.[18]

Creative Commons Non Commercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (https://creativecommons.org/licenses/by-nc/4.0/) which permits non-commercial use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).

BeiGene Exhibit 1138.001
BeiGene v. Pharmacyclics
PGR2024-00009

In silico drug likeness and ADMET prediction

The ligands were screened on the basis of Lipinski's "Rule of five," such as MW < 500, log P < 5, hydrogen bond donors < 5 and hydrogen bond acceptors < 10[22] using SWISS ADME web based application.[21,23,24] Typical ADME prediction methods that involve aqueous solubility (PlogS), PlogBB (blood/brain), logHIA (intestinal barrier), PCaco-2 (cell permeability), log-Pgp (substrate/non-Inhibitor), PlogS (aqueous solubility) and LogPapp (cell permeability), CYP inhibition were calculated using ADMET SAR Toolbox and SWISS ADME, a free web tool used to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules.[25]

Toxicity prediction that involve AMES toxicity, carcinogenicity, fish toxicity, *Tetrahymena Pyriformis* toxicity (TPT), honey bee toxicity, mutagenicity, tumerogenicity, reproductive effect and irritation were calculated using admetSAR, a comprehensive source and free web tool for assessment of chemical toxicity properties.[26] AMES toxicity test, is bacterial reverse mutation assay to detect frame-shift mutations or base-pair substitutions invitro which may be detected by exposure of histidine-dependent strains of *Salmonella typhimurium* to a test compound. *Tetrahymena pyriformis* toxicity is often used as a toxic endpoint.

Protein preparation

The three-dimensional structure of target HIV reverse transcriptase (PDBID: 1REV) was retrieved from protein data bank (http://www.pdb.org) at 2.60 Å RMSD resolution. 1REV has a molecular weight of 116.34 kDa and is active at a pH of (6.5–8.1).[27]

The protein molecule 1REV was prepared using Swiss-PdbViewer v4.1 and autodock 4.2. The protein was in complex with a ligand, water molecules and heteroatoms. Water molecules, inhibitor, and other heteroatoms from the protein were removed using notepad + + and used for docking. The preparation process involved: deleting water molecules and co-crystallized DNA primer complex, adding hydrogen atoms at a pH range of 6.5–8.1 for effective ligand binding using Swiss-PdbViewer v4.1[28] and autodock 4.2.[29] The atom constraints were applied to the protein backbone and the Magnesium Ion were fixed to avoid any modifications in the experimental structure, and then saved in PDB format for energy minimization.

Active site prediction

The active site of reverse transcriptase (PDB ID: 1REV) was predicted using MetaPocket 2.0 online server (https://projects.biotec.tu-dresden.de/metapocket/).[30,31] The processed protein data file without heteroatoms was uploaded and the top result from the three best (based on the z-score) potential ligand-binding site was chosen for docking. Active site prediction using MetaPocket 2.0 has revealed presence of 35 amino acid residues in the ligand binding pocket of 1REV. The amino acid residues predicted were then compared with the amino acids in the active site of the ligand-1REV complex. This was done by using LigPlot + v 2.2[32] and discovery studio softwares.[18] Molecular docking steps were performed after this step.

Molecular docking experiments

Docking experiment was performed with potential active site on HIV reverse transcriptase enzyme using AutoDock 4.2.6 (http://autodock.scripps.edu/).[33] During docking at first the explicit hydrogens, charges, flexible torsions, were assigned by the AD program for both the protein and ligands.

Polar H-atoms were added to the target protein for correct ionization and tautomeric states of amino acid residues. Kollman united charge and Gasteiger charge were added to protein and ligands, respectively. Rigid roots were also assigned to the ligand and five bonds were made "active" or rotatable.[34] The modified 3-dimensional structure of HIV-1RT and ligands accounting for the flexibility of its bonds were converted to PDBQT format as required in AutoDock calculations.

The Lamarckian genetic algorithm (LGA) was utilized to search for the conformations using the following docking parameters: a population size of 150 dockings, a maximum number of generations of 27,000, a maximum energy evaluation of 25 million, 50 docking runs, and random initial positions and conformations.[33] Other parameters such as crossover rate and mutation rate, were used in the default mode. The autogrid program was used to generate grid size for specifying the search space and grid box was set with center x = -36.956; y = 31.989; z = -19.75 and size x = 40; y = 60; z = 50 centered on the predicted cavities with a default grip maps space 0.375 Å spacing. Pre-calculated grid maps, which store grids of energy based on the interaction of the ligand atom probes with receptor targets, were obtained using AutoGrid4.2.

The least binding energy conformation was considered as the most favorable docking pose. The images and output of AutoDock and all modeling studies were analyzed using PyMOL.[35] The interaction between ligand and receptor and hydrogen bond lengths were analyzed using LigPlot + and protein ligand interaction profiler server https://plip-tool.biotec.tu-dresden.de/plip-web/plip/index.[32]

Validation of the docking method

For the validation of docking tool, a decoy set of ligands was used along with the active ligands. For decoy dataset generation, the co-crystal ligand (PDB ID: 1REV) was used. The physicochemical properties of co-crystal that is, molecular weight, number of hydrogen bond acceptors, number of hydrogen bond donors and LogP were used for decoy dataset generation in ChEMBL database. After generation of decoy dataset, all the ligands were prepared by MGL tool for docking. Autodock was used to dock all the ligands, that is, actives and decoys to the specific site of protein.[36]

BeiGene Exhibit 1138.002
BeiGene v. Pharmacyclics
PGR2024-00009

Molecular dynamic simulation

Ligand-protein complexes that displayed better docking poses were subjected to molecular dynamic (MD) simulation. Molecular dynamics simulation was performed for 100 nanoseconds using Desmond-Maestro module 2020, a Package of Schrödinger LLC.[37] The initial stage of protein and ligand complexes for molecular dynamics simulation were obtained from the docking studies. The protein–ligand complex was preprocessed using Protein Preparation Wizard or Maestro, which also included optimization and minimization of complexes. All systems were prepared by the System Builder tool.

The docked complex was first submerged in Transferable Intermolecular Interaction Potential 3 Points (TIP3 P) water model in an orthorhombic shape.[38] The optimized potentials for the liquid simulations (OPLS) - 2005 force field was used in the simulation and for energy calculation. The models were made neutral by adding counter ions where needed. To mimic the physiological conditions, 0.15 M salt (NaCl) was added. The MD simulation was performed under thermodynamically stable conditions.

The NPT ensemble with 300 K temperature and 1 atm pressure was select for complete simulation. The models were relaxed before the simulation. The trajectories of each complex were saved after every 100 ps (0.1 ns) for analysis, and the stability of simulations was evaluated by calculating the root mean square deviation (RMSD) of the protein and ligand over time. In addition, each complex was subjected to specific parameters such as root mean square fluctuation (RMSF), radius of gyration, conformational modification of ligands and intermolecular interactions to analyze the level of structural changes.

Free energy binding calculation

To estimate and compare the binding affinity, the ligand binding energy of the phytochemicals was calculated using the molecular mechanics-generalized born surface area (MMGBSA) module in the Schrodinger Suite 2014[39,40] (Archontis, 2012; Ylilauri & Pentikäinen, 2013). Binding free energy was averaged over 1000 snapshots extracted from the 100 ns trajectory. The binding free energy $\Delta Gbind$ was calculated with the MMGBSA methodology was applied based on stable MD trajectory.

$$\Delta Gbind = Gcomplex - Gprotein - Gligand \quad (1)$$

$Gcomplex$ is the free energy of the complex; $Greceptor$ is the free energy of the receptor; $Gligand$ is the free energy of the ligand.

Results and Discussions

Validation of the docking method

The docking was validated by the area under curve (AUC) graph and the early enrichment factor (E.F) at 1% and 20%. The AUC value for the docking was 0.7040 while the E.F value at 1% and 20% was 2.38 and 2.48 respectively. The AUC values indicate that the docking tool picks the active compounds and rank them better than decoy compounds. Which means that the tool gives 70% true positive results. There were a total 15 active ligands and the E.F at 1% shows that there were 8 ligands in top 1% results. From the results of this step, it can be concluded that docking tools is validated.[36] The AUC graph is shown in Figure S1 in Supplementary.

Molecular docking analysis

Our computational docking study revealed that 10 of the docked compounds had a greater binding efficiency ranging from -6.9 to -7.48 kcal/mol which is much better as compared to the positive control drugs. The binding energies of FDA-approved drugs ranged from -5.63 to -6.85 kcal/mol. Table 1 summarizes the docking results for the phytochemicals and FDA-approved antiretroviral drugs.

Aleuritolic acid (L12), had the best binding conformation with the reverse transcriptase enzyme with a binding energy of -8.48 kcal/mol followed by Crotoxide A (L135), Crothalimene A (L292) and Crotodichogamoin B (L216) with -7.73, -7.48, -7.42 kcal/mol respectively (Table 1). The binding efficiency of the top 10 phytochemicals was greater than the binding efficiency of the FDA approved drugs delaviridine (-6.85 kcal/mol), zidovudine (-5.68 kcal/mol), nevirapine (-5.65 kcal/mol) and abacavir (-5.63 kcal/mol) which confirms that these phytochemicals might have potential reverse transcriptase inhibitory activity.

Molecular dynamic simulation

Binding free energy. As molecular docking only measures the geometric fit of ligands at the active site of a protein in static conditions, molecular dynamics simulations were run for 100 ns to allow for the ligands to become "comfortable" within an enzyme's binding site, to assess the binding free energy of the system. The more negative the values, the better the binding free energy between the enzyme (HIV-1 RT 1REV) and the ligands.

The binding free energy ($\Delta G bind$) of FDA approved drugs and 13 phytochemical compounds isolated from *C. dichogamus* were determined using the MMGBSA method. As shown in Table 2, the free binding energy of the phytochemicals and the FDA-approved drugs is in agreement with the molecular docking results. Delavirdine (DLV) showed the highest binding energy (-50.85 kcal/mol) than the other FDA approved reverse transcriptase inhibitors and phytochemicals compounds. Among the phytochemical compounds aleuritolic acid displayed the highest free binding energy of -173.52 kcal/mol, followed by furocrotinsulolide A (-40.53 kcal/mol), crotoxide A (-38.07 kcal/mol), crotohaumanoxide (-35.78 kcal/mol) and Crothalimene (-32.73 kcal/mol). These five phytochemical

BeiGene Exhibit 1138.003
BeiGene v. Pharmacyclics
PGR2024-00009

Table 1. Molecular docking analysis of phytochemicals isolated from *Croton dichogamus against HIV-1 reverse transcriptase (PDB: 1REV)*.

SER. NO.	CODE	NAME OF PHYTOCHEMICALS	BINDING ENERGY (KCAL/MOL)	INHIBITION CONSTANT, KI IN UM
FDA approved drugs				
1.	DLV	Delaviridine	−6.85	9.54
2.	NVP	Nevirapine	−5.65	72.21
3.	AZT	Zidovudine	−5.68	68.71
4.	ABC	Abacavir	−5.63	74.74
Phytochemicals isolated from *C. dichogamus*				
5.	L12	Aleuritolic acid	−8.48	0.61
6.	L135	Crotoxide A	−7.73	2.12
7.	L292	Crothalimene A	−7.48	3.3
8.	L216	Crotodichogamoin B	−7.42	3.62
9.	L104	Crotonolide E	−7.31	4.42
10.	L136	Crotoxide B	−7.15	5.72
11.	L105	Furocrotinsulolide A	−6.92	8.43
12.	L293	Crothalimene B	−6.92	11.05
13.	L214	Depressin	−6.92	8.53
14.	L215	Crotodichogamoin A	−6.90	11.23
15.	L436	Cadalene	−6.77	10.9
16.	L140	Crotohaumanoxide	−6.76	8.41
17.	L441	4-patchoulene (cyperene)	−6.75	8.77
18.	L440	4-patchoulen-3-one (cyperotundone)	−6.64	13.53
19.	L443	15,16-Epoxy-5,13(16),14-ent-halimatriene-3-ol	−6.57	15.28
20.	L444	15,16-epoxy4(18),13(16),14-ent-clerodatrien-3a-ol (gbaninol)	−6.30	23.96
21.	L439	3β,4β,15,16-diepoxy-13(16),14-ent-clerodadiene-17,(12S)-olide	−6.27	25.5
22.	L442	15,16-epoxy13(16),14-ent-clerodadien-3-one (trans-cascarillone)	−6.01	39.62
23.	L447	1(6),7,9-cadinatriene-4α,5β-diol (4α,5β-corocalanediol)	−6.00	39.79
24.	L320	1,3,5-cadinatriene-(7R,10S)-diol	−5.87	50.13
25.	L437	6α-methoxy-patchoulan-4-ene (6a-methoxycyperene)	−5.6	78.12
26.	L438	3β,4β,15,16-diepoxy-13(16),14-clerodadiene	−5.49	95.32
27.	L446	15,16-epoxy-3β-hydroxy-5(10),13(16),14-ent-halimatriene-17,(12S)-olide	−5.12	176.46
28.	L449	15,16-epoxy-3α-hydroxy-4(18),13(16),14-ent-clerodatrien-17,(12S)-olide (crotonolide F)	−4.98	224.48

Abbreviations: ABC, abacavir; AZT, zidovudine; DLV, delaviridine; FDA, Food and Drug Administration; NVP, nevirapine.

compounds displayed higher free binding energy as compared to the ARV drugs.

From our MMGBSA calculations it is possible to note that electrostatic ($\Delta Gbind\ coulomb$) and van der Waals ($\Delta G\ bind\ vdW$) energies play the main role in the binding free energy of the ligands. Overall, the results provide a set of guidelines for design novel and more potent reverse transcriptase inhibitors.

It was also interesting to note that although compounds crotonolide E, crotodichogamoin B and crotoxide B demonstrated

BeiGene Exhibit 1138.004
BeiGene v. Pharmacyclics
PGR2024-00009

Table 2. Binding free energy for phytochemical compounds from *C. dichogamus* and FDA approved drugs to HIV-RT 1REV.

CODE	NAME OF COMPLEX	ENERGY COMPONENT (KCAL/MOL)				
		ΔGBIND	ΔGBIND COULOMB	ΔGBIND COVALENT	ΔGBIND SOLVGB	ΔGBIND VDW
FDA approved drugs						
DLV	Delaviridine	−50.85 ± 0.25	−19.24 ± 0.21	−0.11 ± 0.06	−52.49 ± 0.26	−32.09 ± 0.11
ABC	Abacavir	−29.01 ± 0.22	−3.38 ± 0.09	−0.22 ± 0.04	−43.09 ± 0.26	−20.58 ± 0.08
NVP	Nevirapine	−28.06 ± 0.17	−6.37 ± 0.11	−0.11 ± 0.09	−28.19 ± 0.19	−21.08 ± 0.09
AZT	Zidovudine	−27.39 ± 0.24	−3.58 ± 0.10	−0.59 ± 0.03	−40.57 ± 0.28	−23.04 ± 0.11
Phytochemical compounds						
L12	Aleuritolic acid	−173.52 ± 1.28	−35.35 ± 0.64	−0.41 ± 0.06	−114.78 ± 1.17	−27.79 ± 0.15
L105	Furocrotinsulolide A	−40.53 ± 0.22	−6.38 ± 0.13	−0.41 ± 0.08	−30.58 ± 0.23	−31.14 ± 0.18
L135	Crotoxide A	−38.07 ± 0.16	−8.9 ± 0.09	−0.67 ± 0.04	−25.47 ± 0.16	−28.4 ± 0.10
L140	Crotohaumanoxide	−35.78 ± 0.19	−8.76 ± 0.09	−0.99 ± 0.07	−37.55 ± 0.21	−24.93 ± 0.13
L292	Crothalimene A	−32.73 ± 0.17	−5.22 ± 0.07	−0.11 ± 0.07	−31.65 ± 0.18	−24.99 ± 0.11
L216	Crotodichogamoin B	−31.98 ± 0.21	−7.75 ± 0.11	−1.58 ± 0.09	−29.76 ± 0.18	−24.18 ± 0.13
L104	Crotonolide E	−30.95 ± 0.21	−5.73 ± 0.13	−0.11 ± 0.06	−26.85 ± 0.19	−29.04 ± 0.15
L215	Crotodichogamoin A	−28.53 ± 0.16	−4.19 ± 0.06	−0.02 ± 0.07	−29.60 ± 0.17	−22.29 ± 0.09
L293	Crothalimene B	−26.07 ± 0.19	−2.96 ± 0.12	−0.44 ± 0.07	−26.11 ± 0.18	−20.46 ± 0.11
L136	Crotoxide B	−24.05 ± 0.23	−3.21 ± 0.09	−0.69 ± 0.14	−43.69 ± 0.24	−18.98 ± 0.15
L214	Depressin	−24.04 ± 0.17	−1.56 ± 0.07	−0.21 ± 0.06	−21.88 ± 0.14	−21.49 ± 0.14
L436	Cadalene	−22.83 ± 0.15	−1.99 ± 0.04	−0.24 ± 0.03	−15.23 ± 0.11	−19.12 ± 0.12
L440	4-patchoulen-3-one (cyperotundone)	−19.28 ± 0.23	−2.14 ± 0.09	−0.42 ± 0.07	−15.51 ± 0.19	−15.83 ± 0.12

Abbreviations: ABC, abacavir; AZT, zidovudine; DLV, delaviridine; FDA, Food and Drug Administration; NVP, nevirapine.

relatively high docking scores, binding free energy calculations for these systems indicated dissimilar results. This validates the need for molecular dynamic simulations, which may allow for a compound to become "comfortable" within an enzyme's binding site.

Structural analysis of the most optimal phytochemical-HIV RT complexes. To further establish the mechanistic inhibitory characteristics of the top ten selected phytochemical compounds with antiviral activity against HIV1-RT and to better understand the complex stability and backbone fluctuation, Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF), Radius of Gyration (RoG) and ligand interaction plots were assessed as previously described by previous studies.[38,41,42]

The Root Mean Square Deviation (RMSD) of the protein backbone was calculated to check the stability of structure during the simulation period. Figure 1 depicts the RMSD plot for the four phytochemical compounds and the FDA approved drugs. In this study, RMSD values for the C-alpha atoms of the structures were determined. The RMSD of each complex was compared to protein RMSD and the RMSD values of the four FDA approved drugs. Deviation in a range of 1-2 Å between the RMSD of the protein alone and RMSD of the complex is considered as acceptable and stable. RMSD measures protein stability as the simulation progresses.

The first 30 ns of the simulation of Crotohalimene A (L292), Crotodichogamoin B (L216) and Aleuritolic acid (L12) showed instability of the enzyme, but from 30 to 100 ns of simulation the enzyme was stable. The RMSD plots of Aleuritolic acid (L12) and Crotodichogamoin B (L216) with average values of 6.75 ± 0.03 Å and 6.72 ± 0.03 Å respectively are similar to the RMSD of nevirapine (6.71 ± 0.03 Å). This indicates the same enzyme stability between the phytochemical compounds and

BeiGene Exhibit 1138.005
BeiGene v. Pharmacyclics
PGR2024-00009

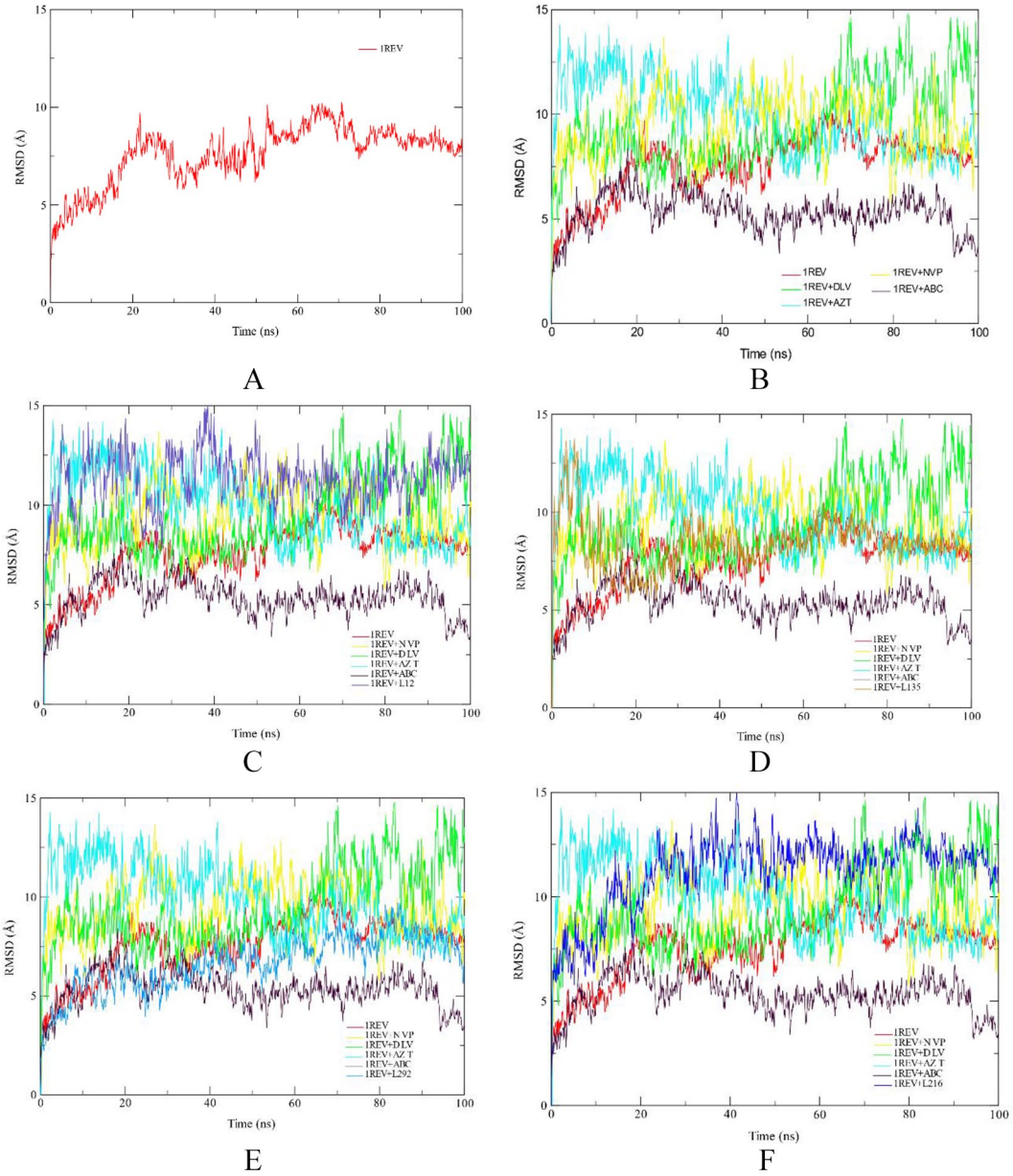


Figure 1. RMSD profile of protein backbone atoms of HIV1-RT 1REV (A), FDA approved antiretreoviral drugs delaviridine-DLV, Nevirapine NVP, Zidovudine AZT and Abacavir ABC drugs (B), Aleuritolic acid, L12 (C); Crotoxide A, L135 (D); Crothalimene A, L292 (E); Crotodichogamoin B, L216 (F), calculated over the course of 100 ns molecular dynamic simulation.
ABC indicates abacavir; AZT, zidovudine; DLV, delaviridine; FDA, Food and Drug Administration; NVP, nevirapine; RMSD, root mean square deviation.

the FDA-approved drugs. Similarly, the RMSD plots of Crotohalimene A (L292) and Depressin (L214) with average values of 8.85 ± 0.06 Å and 7.09 ± 0.04 Å, respectively are related to the RMSD of the apo-enzyme 1REV (7.52 ± 0.05 Å) (reverse transcriptase enzyme without ligand), which indicates the same stability between the phytochemical compounds and the apo-enzyme.

The calculated RMSD between the phytochemical compounds aleuritolic acid (L12), crotoxide A (L135), crothalimene (L292), crotohaumanoxide (L140) and crotodichogamoin B (L216) is within 0.1 nm (1 Å) as compared with the protein HIV-RT (1REV), and the control drugs nevirapine, etravirine and delaviridine indicating only a very small change in the ligands position during the simulation period. This will tell us that these ligands were well stabilized in the protein binding site during the period of simulation, which will infer a more likely similar stability in biological system.

For a drug to bring about an agonistic or antagonistic effect to a protein it should first be stable within the protein. Achieving a given response by a drug requires that the drug stabilize specific conformational states of the receptor and thus specific conformational states in the protein binding pocket. So measuring RMSD of a drug candidate and comparing it with reference drugs is important to determine the stability of the drug within the protein. Therefore, the lower RMSD is, the better the model will be in comparison to the target structure.

BeiGene Exhibit 1138.006
BeiGene v. Pharmacyclics
PGR2024-00009

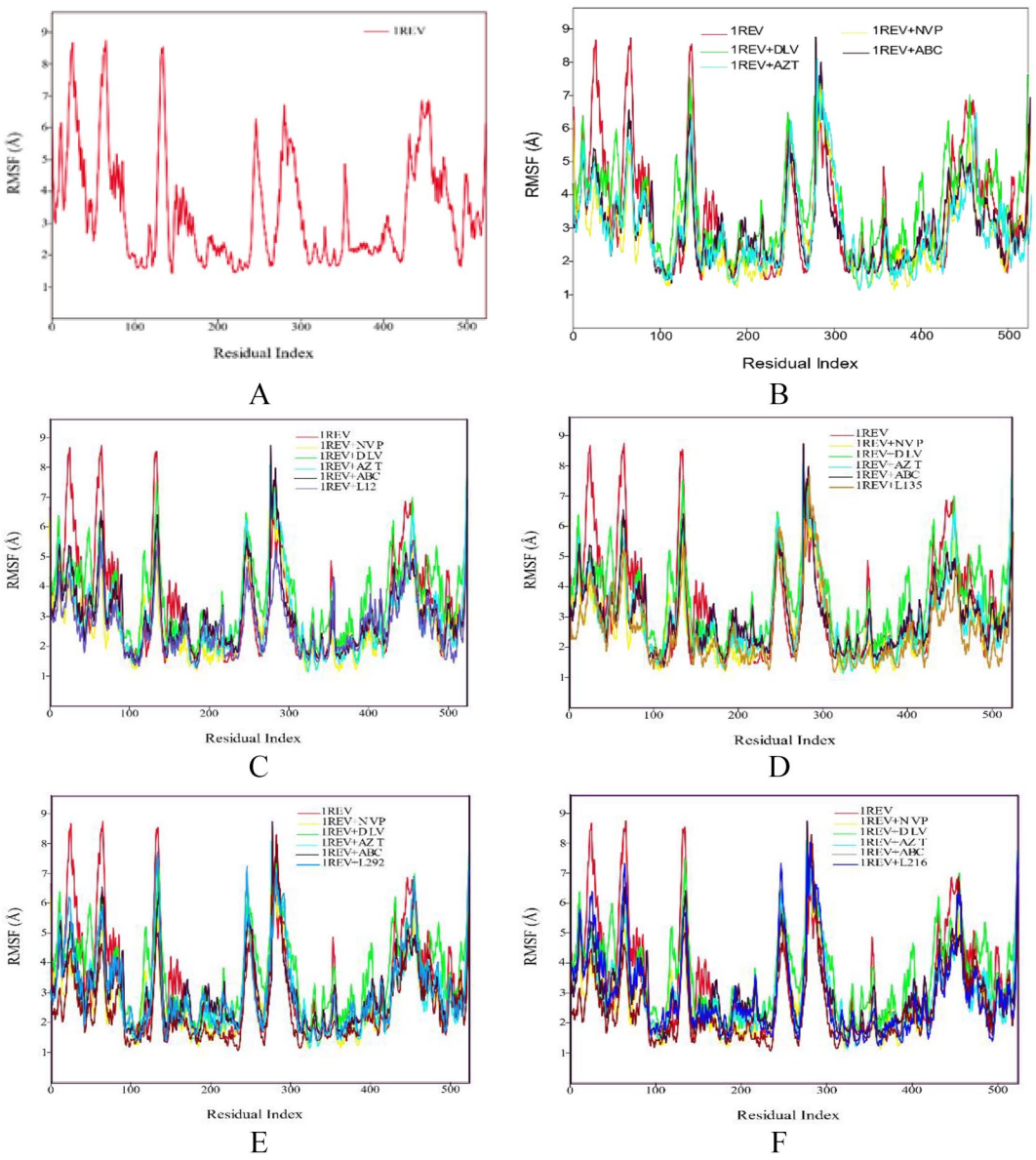


Figure 2. RMSF profile of protein backbone atoms of HIV1-RT 1REV (A), FDA approved antiretreoviral drugs delaviridine-DLV, Nevirapine NVP, Zidovudine AZT and Abacavir ABC drugs (B), Aleuritolic acid, L12 (C); Crotoxide A, L135 (D); Crothalimene A, L292 (E); Crotodichogamoin B, L216 (F), calculated over the course of 100 ns molecular dynamic simulation.
ABC indicates abacavir; AZT, zidovudine; DLV, delaviridine; FDA, Food and Drug Administration; NVP, nevirapine; RMSF, root mean square fluctuation.

The Root Mean Square Fluctuation (RMSF) was calculated to check the flexibility of amino acids residues during simulation. The RMSF values monitor the fluctuation of each amino acid residue as they interact with the ligand throughout a trajectory. The RMSF values of the phytochemicals were compared to the RMSF of the four FDA approved drugs.

Figure 2A Shows the RMSF plot of reference (1REV). The amino acid residues that showed low RMSF values remained rigid during simulation and the residues with higher RMSF values showed fluctuations during simulation. RMSF plot shows a minor fluctuation in the region of ~ 11 to 26 amino acids. There are major fluctuations seen in the regions of ~ 30 to 50, 70 to 80, 120 to 140, 230 to 270, 280 to 300. It showed that these amino acids residues were present on the cytoplasmic side of C terminal helix. The plot showed fluctuations at some positions ignoring the C-terminal helix while remaining protein showed less fluctuations.

Figure 2B shows the RMSF plots of controls, that is, DLV, AZT, NVP, and ABC. This plot shows that the amino acid residues of protein behaved like reference 1REV during simulation. All the fluctuations are same as reference. A minor difference in the loop region (11 to 26 and 70 to 80) was observed. The overall plots showed similar trends. As shown in Figure 2C, RMSF plot of 1REV + L12 complex shows a similar trend like NVP control.

As depicted in Figure 2, the RMSF plot of 1REV + L135 showed that the amino acid residues have same trend in fluctuations as the control ABC, while 1REV + L292 and 1REV + L216 complexes showed related RMSF plot to NVP. RMSF is useful for characterizing local changes along the pro-

BeiGene Exhibit 1138.007
BeiGene v. Pharmacyclics
PGR2024-00009

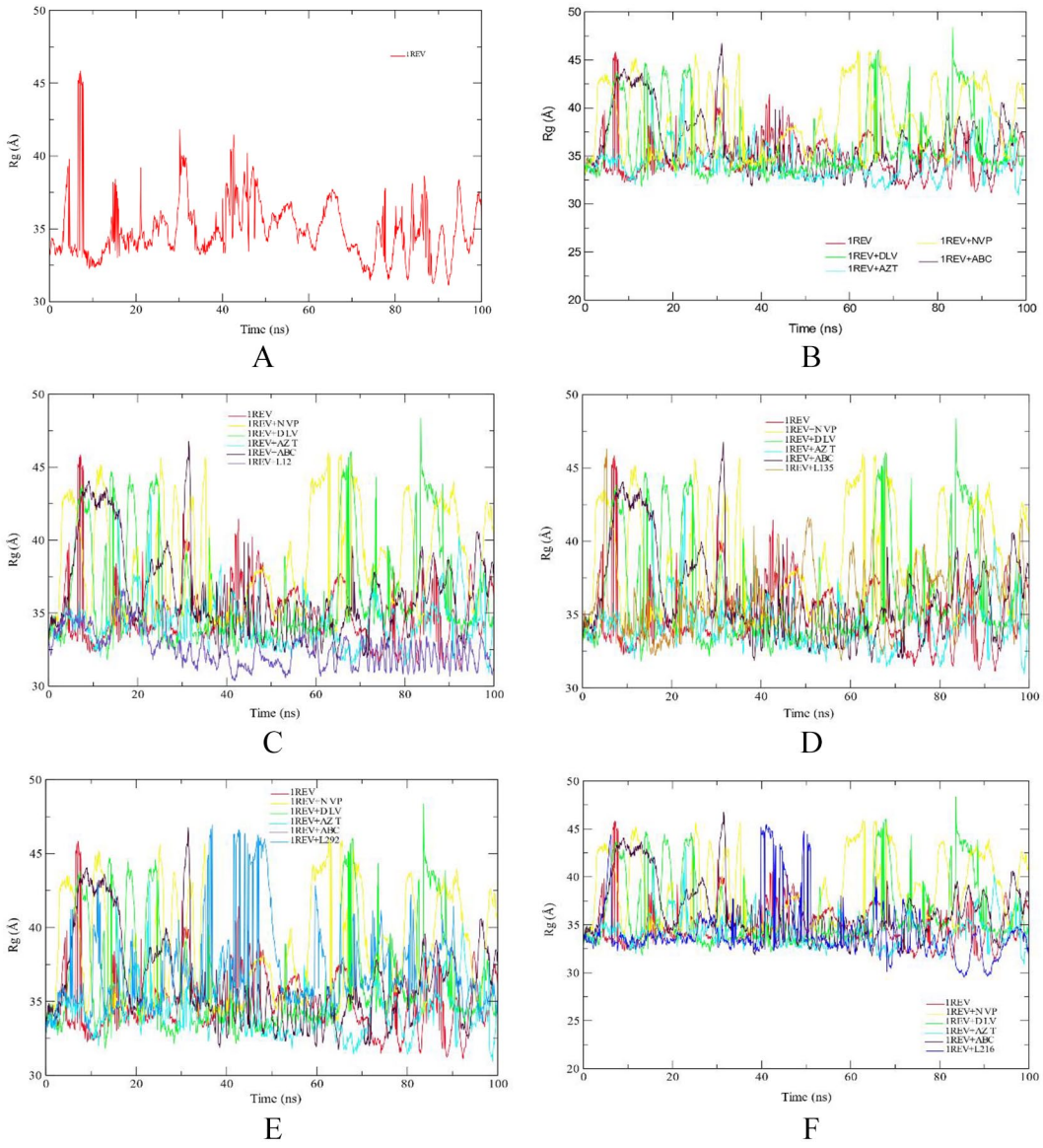


Figure 3. RoG profile of protein backbone atoms of HIV1-RT 1REV (A), FDA approved antiretreoviral drugs delaviridine-DLV, Nevirapine NVP, Zidovudine AZT and Abacavir ABC drugs (B), Aleuritolic acid, L12 (C); Crotoxide A, L135 (D); Crothalimene A, L292 (E); Crotodichogamoin B, L216 (F), calculated over the course of 100 ns molecular dynamic simulation.
ABC indicates abacavir; AZT, zidovudine; DLV, delaviridine; FDA, Food and Drug Administration; NVP, nevirapine; RoG, radius of gyration.

tein chain. On RMSF graphs, peaks indicate areas of the protein that fluctuate the most during the simulation.

The radius of gyration was calculated to examine the compactness of system. The high value of RoG shows the unfolding events during simulation. Figure 3A shows the RoG plot of reference (1REV). From the plot, it can be observed that the system remained compact till ~ 5 ns and then it showed distortion till 10 ns and then it gained stability. Protein showed some unfolding events during ~ 30 to 35 ns, 41 to 44 ns, but it was compact during remaining time. From the plot, it can be observed that, besides for some time, the protein remained compacted during 100 ns long simulation.

The RoG values of crotodichogamoin B, L216 (34.19 Å), furocrotinsulolide A, L105 (34.12 Å) show similarity with average RoG values with zidovudine (34.29 Å) and the reverse transcriptase enzyme (34.97 Å). The RoG values of crothalimene B, L293 (39.57 Å), crothalimene A, L292 (37.03 Å), and crotohaumanoxide, L140 (37.83 Å) show similarity with the RoG of Nevirapine, NVP (38.45 Å).

Similarly, the RoG values of Crotonolide E, L104 (33.83 Å) and Cadalene, L436 (33.65 Å) show related RoG values with zidovudine (34.29 Å) while the RoG values of Crotoxide B, L136 (36.52 Å) and Crotoxide A, L135 (35.76 Å) is related with the RoG value of Delaviridine (36.42 Å) (Figure 3).

If a protein is stably folded, it will likely maintain a relatively steady value of RoG. If a protein unfolds, its RoG will change over time. The RoG is used to assess the overall dimensions and stabilities of the enzyme-ligand complex and is a function of the mass-weighted RMS distances of atoms from the center of mass.

BeiGene Exhibit 1138.008
BeiGene v. Pharmacyclics
PGR2024-00009

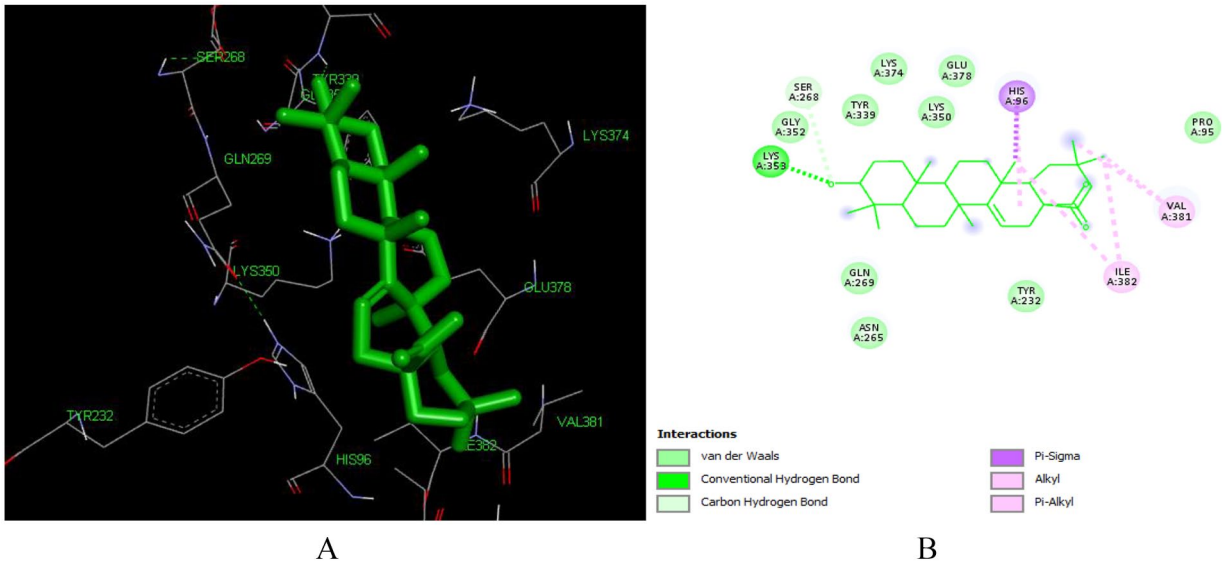


Figure 4. Docked poses of Aleuritolic acid (L12) with the active site region of reverse transcriptase (PDB ID: 1REV) enzyme (binding energy -8.48 kcal/mol) (A) 3D Aleuritolic acid with surrounding amino acids of 1REV; (B) 2D view of interaction type of Aleuritolic acid with surrounding amino acids of 1REV.

Ligand-RT interaction with different amino acids

Active site prediction using MetaPocket 2.0 revealed presence of 35 amino acid residues in the ligand binding pocket of 1REV. To gain insight on the binding modes of studied phytochemicals as per the molecular docking results, visual poses inspection analysis was performed for the top 10 phytochemicals which exhibited highest binding energy in the active site pocket of the HIV-RT (PDBID: **1REV**).

The interactions of phytochemicals with key residues of HIV-1 reverse transcriptase demonstrated that the ligands interact with most of residues of the hydrophobic pocket as shown in Table S1 in Supplementary. Most of the residues involved on the hydrophobic interaction include ASN 265, GLU 378, GLY 352, HIS 96, ILE 382, SER 268, TRP 266.

In silico binding studies suggest that inhibitors that undergo hydrogen bonding with the main chain backbone of Lys101, LYS 350, LYS 353 and pi-pi interaction with the aromatic side chain of Trp229 improves the inhibitor selectivity for RT and thus helps in further drug design attempts to obtain potent phytochemical compounds or their derivatives.[43]

Our results also demonstrate that most of the ligands formed hydrogen bonds with at least one key residue of the enzyme, the most represented are ARG 355, ARG 356, ARG 358, GLN 269, ILE 94, LEU 92, LYS 350, LYS 353, LYS 374, TYR 232 and the distances of hydrogen bonds vary between 2 Å and 4 Å.

Interaction of aleuritolic acid (L12). Aleuritolic acid (L12) is a pentacyclic triterpenoid isolated from the stem bark of *C. megalocarpus*,[44] *C. dichogamus, C. psudopulchellus* and *C. oligandrus*,[10] *C. urucurana*[45] and *Jatropha isabellei*.[46] Previous studies have proven anti-inflammatory,[47] antifilarial,[48] anti-HIV,[49] antinociceptive[50] and antioxidant[46] activity of aleuritolic acid its derivative acetyl aleuritolic acid.

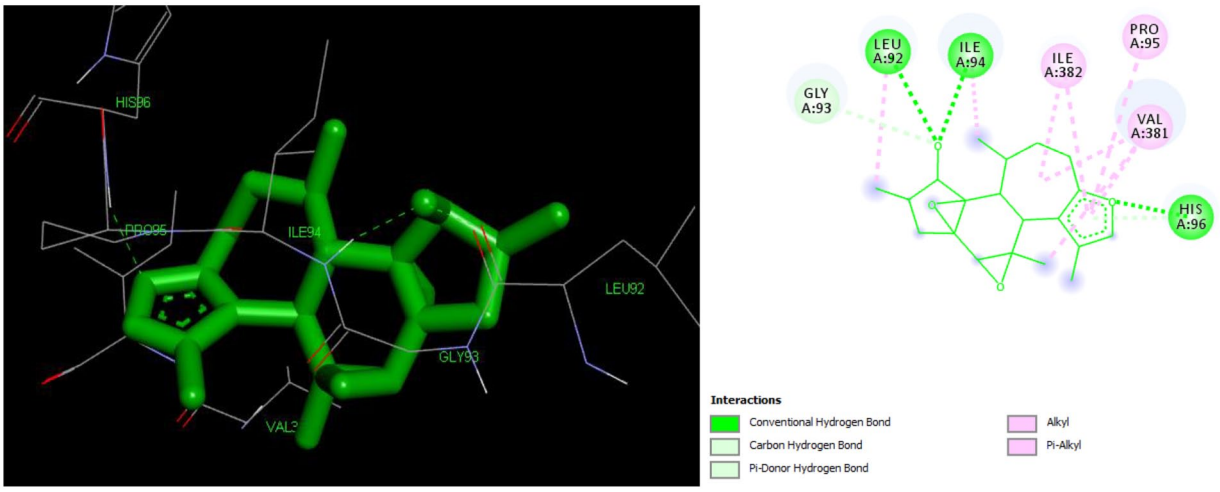
Aleuritolic acid forms strong binding to the receptor with an estimated free binding energy of -8.48 kcal/mol and Ki of 0.61 uM making it superior in binding efficiency as compared to all docked compounds including the FDA-approved drugs.

This high binding energy of aleuritolic acid in the binding site is attributed to the Pi-alkyl interactions with ILE 382 and VAL 382, Pi-Sigma interactions with HIS 96 which involves charge transfer and helps in intercalating the compound in the binding site of the receptor (1REV) as shown in Figure 4. Furthermore, it displayed hydrophobic interactions with residues in the active binding site including – ASN 265, GLN 269, GLU 378, GLY 352, HIS 96, ILE 94, LYS 350, LYS 353, LYS 374, PRO 95, SER 268, TYR 232, TYR 339. Similar interacting amino acids were reported previously by Singh et al,[51] Seal et al.[52]

Aleuritolic acid forms two hydrogen bonds with LYS353 with an interatomic distance of 2.77 Å and 3.62 Å. The hydroxyl group at position 1 and the carboxyl acid moiety at the 11th position have a role in forming these hydrogen bonds. The in-silico predicted inhibition constant (Ki) value of aleuritolic acid was 0.61 uM (Table 1). Inhibition constant value is the half-maximum inhibition of an enzyme by a chemical compound and is used to estimate the potential of substrate/inhibitor in enhancing/inhibiting the biological and function of enzymes.[53] Compounds with an inhibition constant less than 100 mM are considered to be potential inhibitors whereas inhibition constant greater than 100 mM are non-potent inhibitors.[54] Based on this, aleuritolic acid could be a potential potent inhibitor of reverse transcriptase enzyme.

Interaction of crotoxide A (L135) and crotoxide B (L136). Crotoxide A (L135) is a crotofolane-type diterpenoid isolated from leaves of *C. dichogamus*.[55] The compound has a very good docking pose with the reverse transcriptase enzyme (PDB: 1REV) with a binding energy of -7.73 kcal/mol and inhibitory constant

BeiGene Exhibit 1138.009
BeiGene v. Pharmacyclics
PGR2024-00009



A B

Figure 5. Docked poses of Crotoxide A (L135) with the active site region of reverse transcriptase (PDB ID: 1REV) enzyme (binding energy -7.73 kcal/mol) (A) 3D Crotoxide A (L135) with surrounding amino acids of 1REV; (B) 2D view of interaction type of Crotoxide A (L135) with surrounding amino acids of 1REV.

of 2.11 µM making it the second most active compound to inhibit the enzyme. As depicted in Figure 5, corotoxide interacts with the following amino acid residues of the enzyme LEU 92, HIS 96, VAL 381, ILE 382, GLY 93, PRO 95, ILE 94, ILE 94. It forms two hydrogen bonding with ILE 94 (2.01 Å and 3.01 Å) and His 96 (3.05 Å). It also forms Pi-alkyl bond with VAL 381, PRO 95 and ILE 382, these interactions are playing a crucial role in the recognition of ligand by protein.[56]

The furan ring is responsible for formation of the hydrogen bonding with HIS 96 and for the formation of Pi-Alkyl bond with ILE 382, PRO 95 and VAL382. The hydroxyl group at position 12 is responsible for the formation of hydrogen bonding with LEU 92 and ILE 94 (Table S1 in Supplementary).

From structure activity relationship of Crotoxide A and B, as shown in Figure 6, substitution of C-12 hydroxyl group by acetyl group like in the case of Crotoxide B will reduce the binding efficiency (binding energy -7.15 kcal/mol) of the ligand to the receptor. Hence Crotoxide A was ranked the second while Crotoxide B (L136) was ranked as the sixth most active compound to inhibit reverse transcriptase enzyme (PDB ID: 1REV) as shown in Table 1.

Crotoxide B (L136) forms 7 hydrogen bonds with LYS 353, LYS 374 (3.43 Å), ALA 355, TYR 339 (2.79 Å), ASN 265 (2.87 Å), LYS 374 (3.43 Å) (Figure 7). The acetyl moiety at C-12, the hydroxyl group at C-7, and the furan ring has played crucial role in the hydrogen bond formation. This implies that Crotoxide B can be a potential hit compound that can inhibit reverse transcriptase enzyme.

Interaction of crothalimene A (L292) and crothalimene B (L293). Crothalimene A (L292) is a halimene type diterpenoid isolated from *C. dichogamus*.[10] It binds to the reverse transcriptase enzyme (PDB ID: 1REV) satisfactorily with a binding energy of -7.48 kcal/mol and Ki of 3.3 µM making it the third most active drug with the ability to inhibit the enzyme. As depicted in Figure 8, crothalimene A (L292) forms two hydrogen bonds with Ile 382 (2.31 Å) and Tyr 232 (2.69 Å). The hydrogen acceptor groups at position number 2 and 17 are responsible for formation of these hydrogen bonds. It also forms Pi-alkyl interactions with HIS 96, ILE 94, VAL 381, and PRO95 which is important for conformational stability of the compound and also for recognition of ligand by protein.[56] It also displayed a number of hydrophobic contacts with 10 amino acid residues HIS 96 (3.30 Å), VAL 381 (3.56 Å), ILE 382 (3.60 Å), ILE 94 (2.94 Å), ILE 382 (2.31 Å), HIS 96 (5.29 Å), TYR 232 (2.69 Å), PRO 95, LEU 92, GLY 93. The naphthalene moiety could be responsible for these hydrophobic contacts.

Crotohalmane A (L292) and crothalimene B (L293) have related chemical structures as depicted in Figure 9. Their structural difference lies in the presence of additional tetrahydro 4H pyran-2-one ring in the structure of Crothalimene. The tetrahydro 4H pyran-2-one ring has a role in forming strong binding in the receptor pocket. The structural difference has impacted the binding energy, inhibitory constant and the different bonds created between the compound and the amino acids within the active enzyme site. The binding energy and inhibition constant for crothalimene B was -6.92 kcal/mol and 11.92 µM, respectively.

Crothalimene B forms hydrophobic contacts with the following amino acids ILE 94, LYS 350, GLU 378, HIS 96, ILE 94, PRO 95, GLN 269, and TYR 232 in the active site of the reverse transcriptase (PDB ID: 1 REV). It also forms Pi-alkyl bonding with ILE 94, ILE 382 and VAL 381. The pyran and naphthalene rings are responsible for this binding and contribute for the stability of the conformation.

As depicted in Figures 8 and 10, C-2 of both compounds is involved in hydrogen bonding with ILE 382 and LYS 350

BeiGene Exhibit 1138.010
BeiGene v. Pharmacyclics
PGR2024-00009

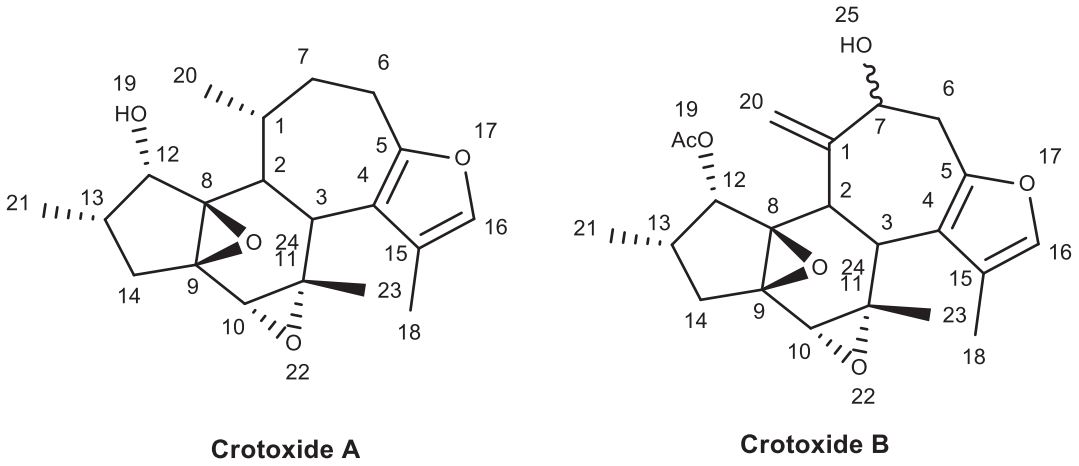


Figure 6. Chemical structure of Crotoxide A (L135) and Crotoxide B (L136).

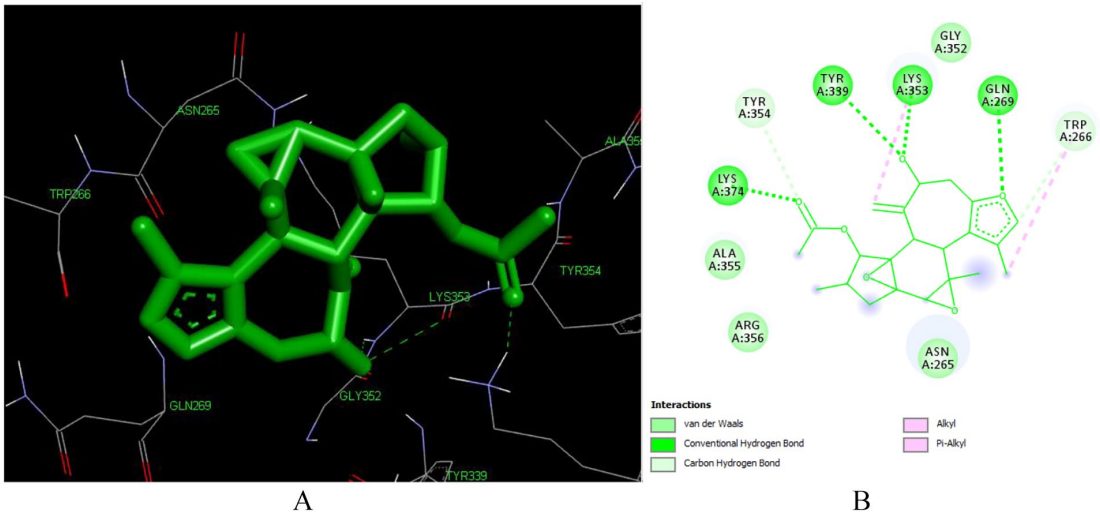


Figure 7. Docked poses of Crotoxide B (L136) with the active site region of reverse transcriptase (PDB ID: 1REV) enzyme (binding energy -7.15 kcal/mol) (A) 3D Crotoxide B (L136) with surrounding amino acids of 1REV; (B) 2D view of interaction type of Crotoxide A (L136) with surrounding amino acids of 1REV.

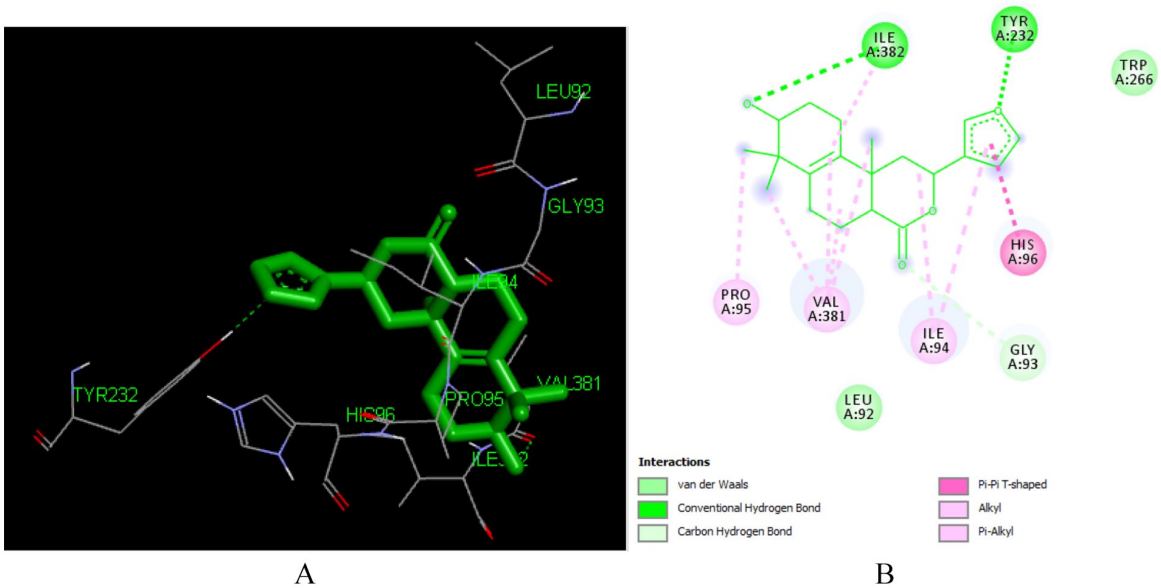


Figure 8. Docked poses of Crothalimene A (L292) with the active site region of reverse transcriptase (PDB ID: 1REV) enzyme (binding energy -7.48 kcal/mol) (A) 3D Crothalimene A (L292) with surrounding amino acids of 1REV; (B) 2D view of interaction type of Crothalimene A (L292) with surrounding amino acids of 1REV.

BeiGene Exhibit 1138.011
BeiGene v. Pharmacyclics
PGR2024-00009

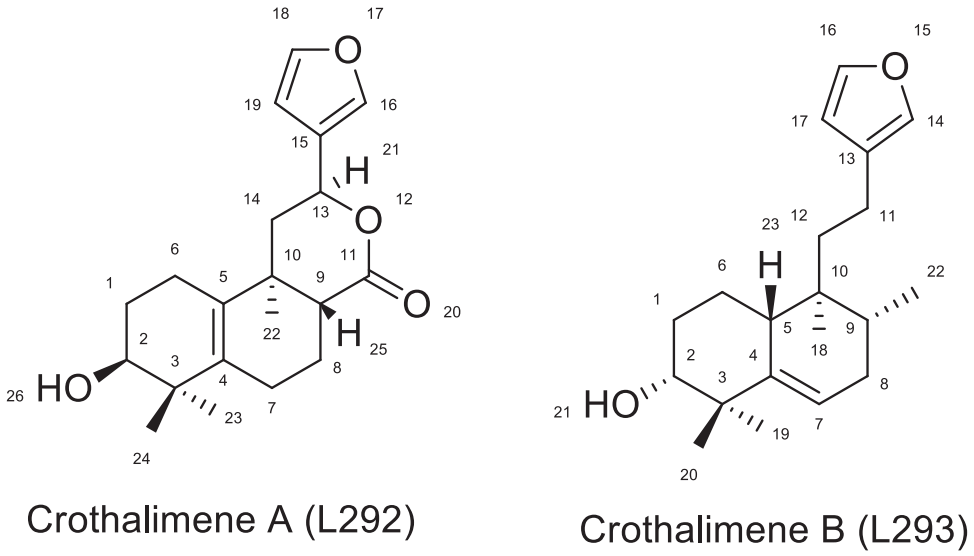


Figure 9. Structure of Crothalimene A and Crothalimene B.

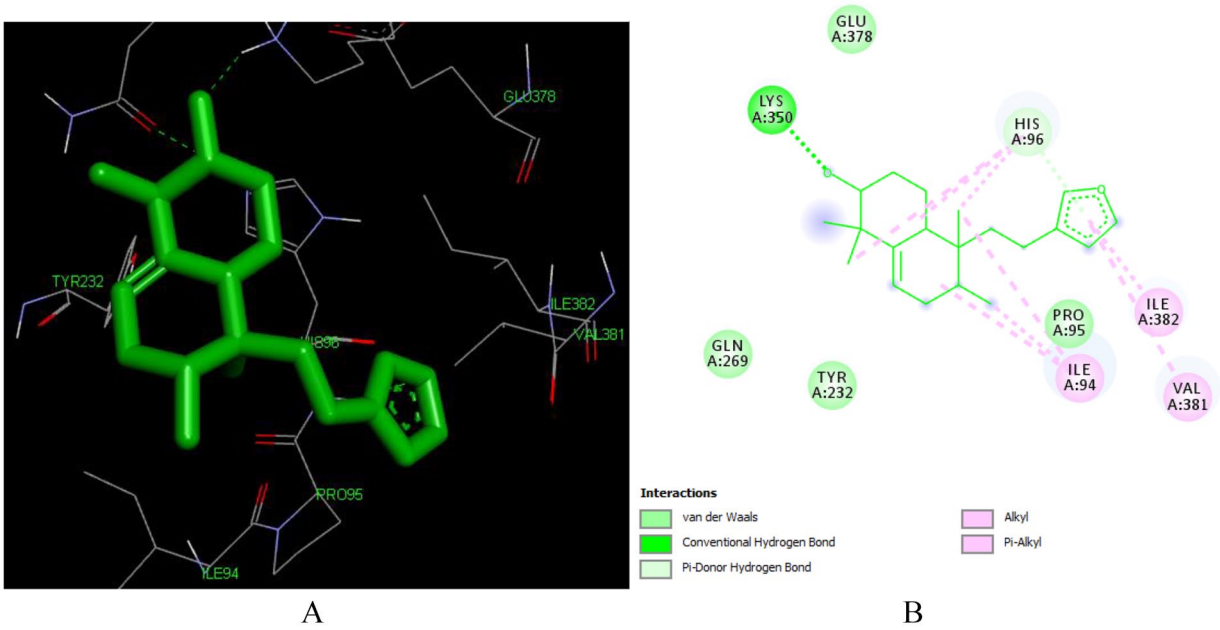


Figure 10. Docked poses of Crothalimene B (L293) with the active site region of reverse transcriptase (PDB ID: 1REV) enzyme (binding energy -7.48 kcal/mol) (A) 3D Crothalimene B (L293) with surrounding amino acids of 1REV; (B) 2D view of interaction type Crothalimene B (L293) with surrounding amino acids of 1REV.

respectively for crothalimene A and B. The furan moiety in crothalimene B has no role in formation of the hydrogen bond but is engaged in the Pi-alkyl bonding with ILE 382 and VAL 381 which contributes for receptor fitting and stabilization.

Interaction of crotodichogamin A (L215) and crotodichogamoin B (L216). **Crotodichogamoin B (L216)** is a crotofolane type diterpenoids isolated from the roots of *C. dichogamus*.[10] The compound binds well with 1REV ligand binding site with a free binding energy of -7.42 kcal/mol and inhibition constant of 3.62 µM making it the fourth most active molecule to inhibit reverse transcriptase (1REV). As illustrated in Figures 9 and 11, crotodichogamin B forms hydrophobic contacts with 5 amino acids TYR 232, MET 230, GLN 269, LYS 350, PRO 95 and four hydrogen bonding with LYS 350 (2.91 Å), TYR 232 (2.27 Å), GLU 378 (3.72 Å), GLU 378 (3.13 Å), and HIS 96 (4.53 Å). The stability of the complex in the binding pocket can be linked to the Pi-stacked interaction including Pi-cation interaction with HIS 96, Pi-alkyl interactions with ILE 94, ILE 382, VAL 381 and MET 230. The phenyl group is responsible for the formation of hydrogen bonding with LYS 350 and Pi-cation interaction with HIS 96.

Crotodichogamin A (L215) has a lower binding efficiency as compared to crotodichogamin B. Its binding energy and

BeiGene Exhibit 1138.012
BeiGene v. Pharmacyclics
PGR2024-00009

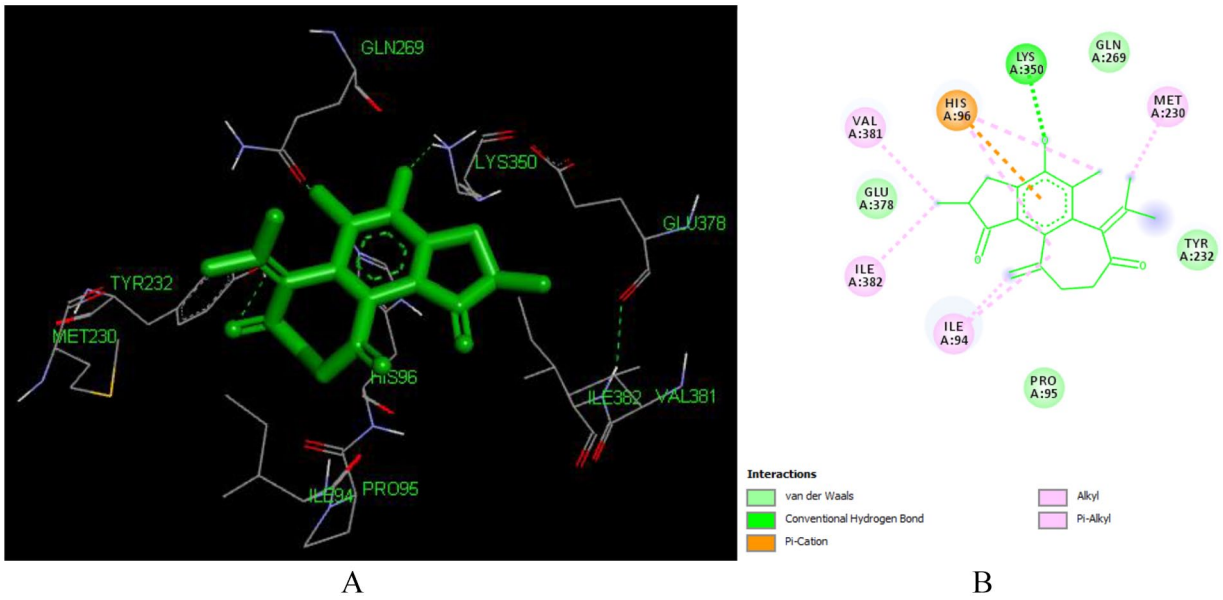


Figure 11. Docked poses of Crotodichogamoin B (L216) with the active site region of reverse transcriptase (PDB ID: 1REV) enzyme (binding energy -7.42 kcal/mol) (A) 3D Crotodichogamoin B (L216) with surrounding amino acids of 1REV; (B) 2D view of interaction type Crotodichogamoin B (L216) with surrounding amino acids of 1REV.

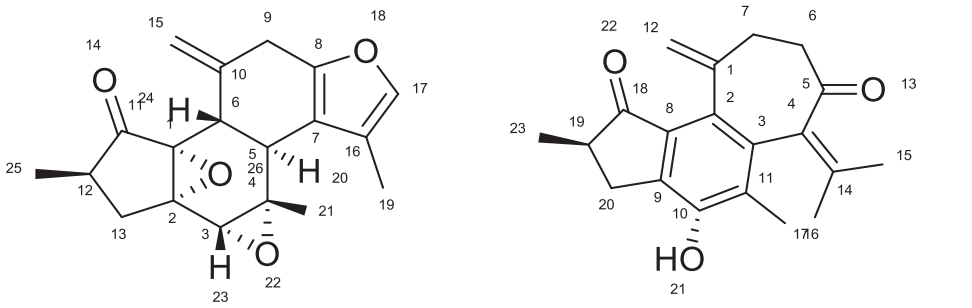


Figure 12. Chemical structures of Crotodichogamin A and B.

inhibitor constant is -6.9 kcal/mol and 11.23 uM respectively ranking as the 10th active compound to inhibit the reverse transcriptase (PDB ID: 1REV) enzyme. As depicted in Figure 11, it forms 8 hydrophobic contacts with ILE 94, MET 230, TRP 266, GLN 269, GLU 378, LYS 350, HIS 96, and TYR 232 amino acids in the active site of the enzyme. It also forms hydrogen bonding with TYR 232 (1.78 Å), HIS 96 (4.34 Å) and LYS 350. Crotodichogamin A forms Pi-cation interaction with HIS 96 and Pi-alkyl interactions ILE 94, ILE 382, TRP 266.

The two epoxide moieties play a role in forming hydrogen bond with Tyr 232, while the furan ring is engaged in hydrogen bonding with LYS 350, and formation of HIS 96. As shown in Figure 12, the difference in binding energy among crotodichogamin A and B could be attributed to the presence of cycloheptane ring in crotodichogamin B which increases its hydrophobicity and receptor affinity which is displayed by the pi-alkyl bond formed with ILE 94 and MET230

Interaction of crotonolide E (L104) with reverse transcriptase enzyme (PDB ID: 1REV). Crotonolide E (L104) is a clerodane type diterpenoid isolated from the roots of *C. dichogamus* and *C. megalocarpus*.[49,57] As visualized in Figure 13, it forms six hydrophobic contacts ILE 94 (3.29 Å), ILE 94 (3.52 Å), HIS 96 (2.94 Å), TRP 266 (3.82 Å), PRO 95, GLU 378, and one hydrogen bond with TYR 232 (2.79 Å). The compound also forms Pi-alkyl bond with TRP 266, ILE 382, PRO 95. Crotonolide E has free binding energy of -7.31 kcal/mol and inhibition constant of 4.42 µM with reverse transcriptase (PDB ID: 1REV) enzyme.

Interaction of furocrotinsulolide A (L105) with reverse transcriptase enzyme (PDB ID: 1REV). Furocrotinsulolide A (L 105) is a clerodane type diterpenoid isolated from the roots of *C. dichogamus*, *C. megalocarpus* and *C. insularis*.[57,58] It forms hydrogen bonding with LYS 353 (1.94 Å), LYS 353 (2.65 Å), LYS 374 (1.95 Å) in the active site of the enzyme. The hydroxyl groups in C-2 and C-3 are responsible for the formation of the

BeiGene Exhibit 1138.013
BeiGene v. Pharmacyclics
PGR2024-00009

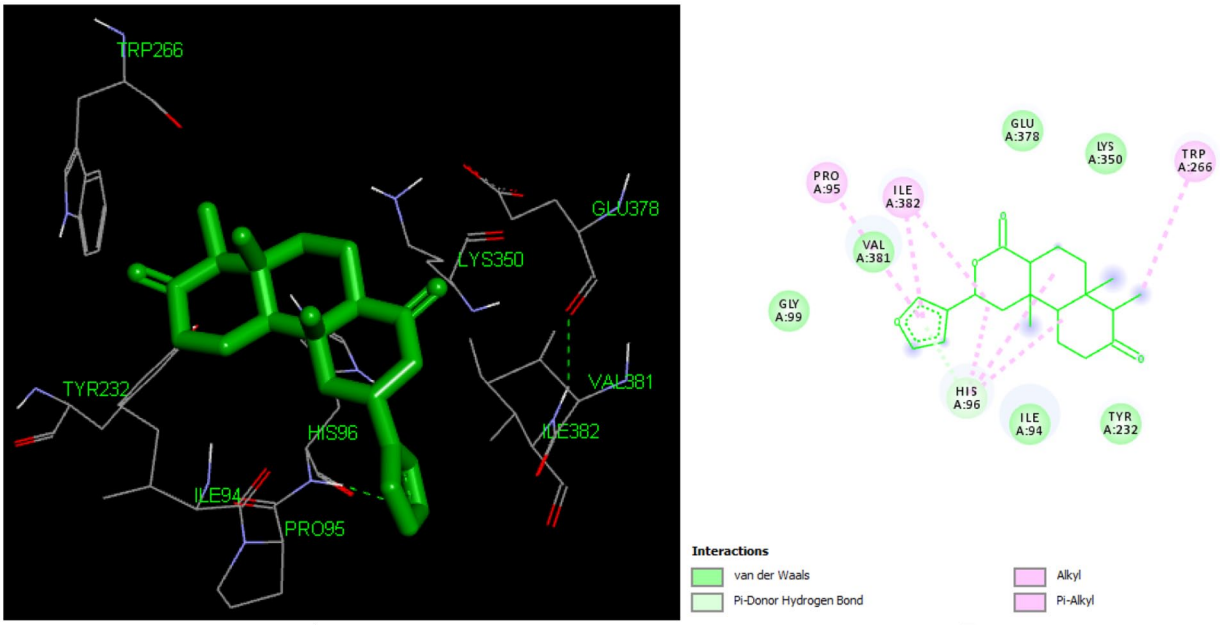


Figure 13. Docked poses of Crotonolide E (L104) with the active site region of reverse transcriptase (PDB ID: 1REV) enzyme (binding energy –7.31 kcal/mol) (A) 3D Crotonolide E (L104) with surrounding amino acids of 1REV; (B) 2D view of interaction type Crotonolide E (L104) with surrounding amino acids of 1REV.

hydrogen bond. Our computations studies indicated that the binding energy of Furocrotinsulolide A with the receptor is -6.92 kcal/mol (Figure 14).

Interaction of depressin (L214) with reverse transcriptase enzyme (PDB ID: 1REV). Depressin (L214) is Caspian diterpenoid isolated from roots of *C. dichogamus*.[10,59] It forms good docking pose with the reverse transcriptase (PDB ID: 1REV) at free binding energy of -6.92 kcal/ml and inhibition constant of 8.53 μM. This interaction was supported by hydrophobic interactions with ILE 94 (3.26 Å), ILE 94 (3.10 Å), HIS 96 (3.22 Å, TYR 232 (3.43 Å), TRP 266, GLN 269, MET 230 found in the active site of the enzyme. It also forms hydrogen bonding with LYS 350 (2.58 Å), HIS 96 (3.15 Å). The carbonyl group at C-13 (Figure 15) is responsible for the formation of hydrogen bond with Lys 350.

In silico pharmacokinetic ADMET prediction. The phytochemical compounds investigated fulfill the Lipinski rule of drug likeness and have acceptable molecular weight and solubility profile except aleuritolic acid (L12) and Depressin (L214) which had logP value of 6.06 and 7.31, respectively, Table S2 in Supplementary. The polar surface area of the phytochemical compounds was predicted to be less than 100 indicating that these compounds had good oral absorption or membrane permeability.[60] Among the FDA-approved drugs RPV and NVP were predicted to have good absorption while the other drugs had displayed strong polarity. Generally, drugs with smaller PSA are more easily absorbed.[61,62]

The phytochemical compounds and the FDA-approved drugs were predicted as having ideal lipophilicity (AlogP98 [WLOGP] ⩽ 5) except aleuritolic acid (L12) and depressin (L214) that display poor lipophilicity, AlogP98 (WLOGP) > 5.[63] With regard to intestinal absorption (human), absorbance of less than 30% is considered to be poorly absorbed. The phytochemical compounds and the FDA-approved drugs were predicted to have good absorption. Gastrointestinal (GIT) absorption is significant for the maintenance of optimal drug levels in the systemic circulation. For drugs or potential compounds to reach their target, they must be absorbed from the GIT and enter the systemic circulation in enough amount or quantities.[64] Highly absorbed drugs from the GIT will easily attain optimal concentration and exert a pharmacological effect at its target site.

With regard to skin permeability, the log Kp > -2.5, the compound is considered to be relatively low skin permeability.[63] Crothalimene B (L293) and depressin (L215) had predicted low skin permeability, while the other phytochemical compounds and ARV drugs were predicted to have high skin permeability. Skin permeability is a significant consideration for many consumer products efficacy, and of interest for the development of transdermal drug delivery.[65]

As shown in Table 3, among the FDA approved antiretroviral drugs, DLV, RPV, ETV and ABC are predicted to be P-glycoprotein substrates. Similarly, the 10 phytochemical compounds isolated from *C. dichogamus* are predicted to be P-glycoprotein substrates, hence they may be actively exuded from cells by P-glycoprotein and while compounds L292, L216, L293, L2215 and L12 are predicted to be P-glycoprotein inhibitors.[65] Among the phytochemical compounds L12 and L214 were predicted to be pumped out of the cell by P-glycoprotein efflux pump. While L293, L216, L104, L292, L135, L215 and L105 were predicted to cross the BBB.

BeiGene Exhibit 1138.014
BeiGene v. Pharmacyclics
PGR2024-00009

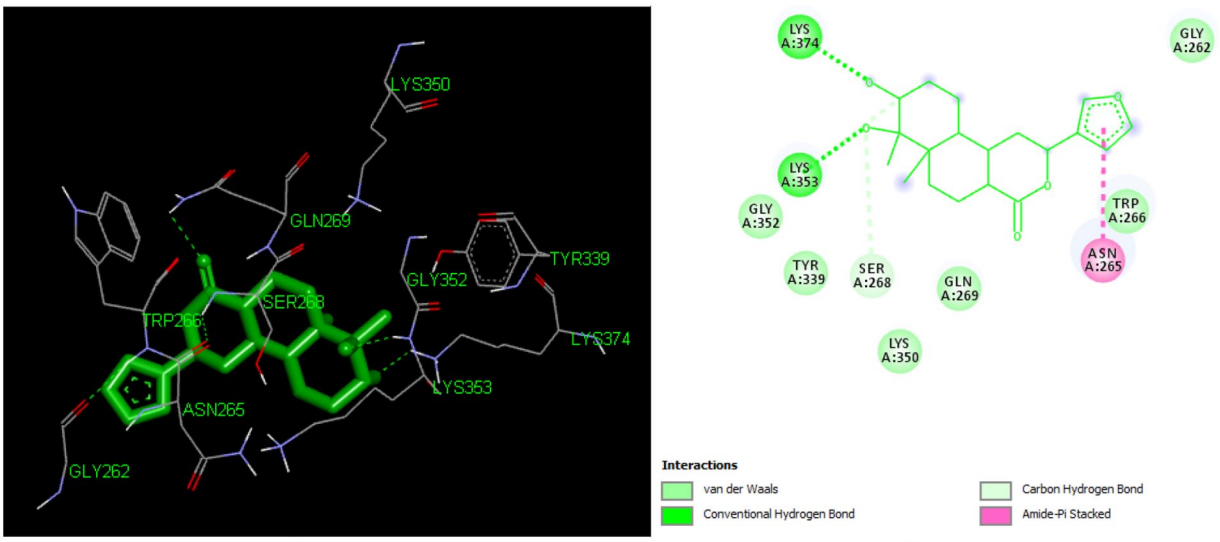


Figure 14. Docked poses of Furocrotinsulolide A (L105) with the active site region of reverse transcriptase (PDB ID: 1REV) enzyme (binding energy −6.92 kcal/mol) (A) 3D Furocrotinsulolide A (L105) with surrounding amino acids of 1REV; (B) 2D view of interaction type Furocrotinsulolide A (L105) with surrounding amino acids of 1REV.

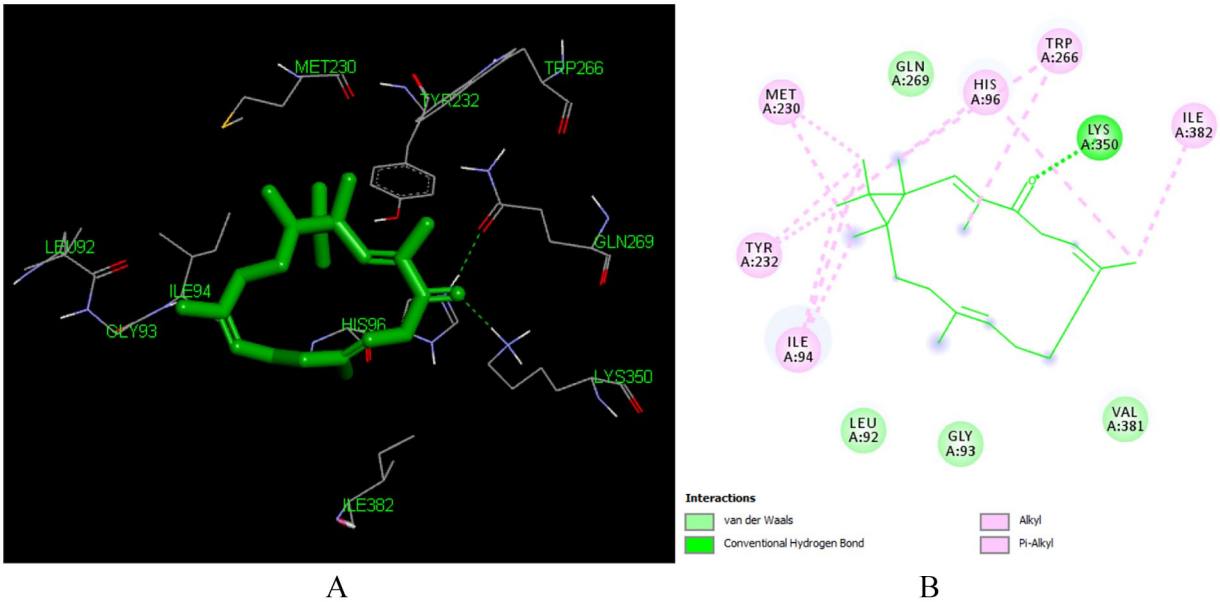


Figure 15. Docked poses of Depressin (L214) with the active site region of reverse transcriptase (PDB ID: 1REV) enzyme (binding energy −6.92 kcal/mol) (A) 3D Depressin (L214) with surrounding amino acids of 1REV; (B) 2D view of interaction type Depressin (L214) with surrounding amino acids of 1REV.

The "Brain Or Intestinal Estimated permeation, (BOILED-Egg)" method was utilized as it computes the lipophilicity and polarity of small molecules.[66] As depicted in Figure 16, in the BOILED-Egg model, the white region represents the passive absorption of the GI tract, while the BBB penetration is represented by the yellow region (yolk) represents. The blue color indicator represents a molecule which is actively effluxed by P-glycoprotein (PGP+), whereas the red color indicator shows the nonsubstrate P-gp (PGP-).[67]

The results showed that the DLV and RPV have high predicted distribution volume (VDss), as compared to the other antiretroviral drugs. Among the phytochemical compounds L135, L216, L293, L214 and L215 were predicted to have high VDss. It was also noted that compounds with higher AlogP98/WLOGP values had high predicted VDss, a good example being L293 and RPV.

In addition, aleuritolic acid (L12), furocrotinsulolide A (L105), crotoxide B (L136), crotohaumanoxide (L140), depressin (L214) and cadalene (L4360) were predicted not to cross the BBB. Among the phytochemical compounds L293 and L214 were predicted to cross the BBB as determined by the logBB value > 0.3, while the other compounds were not predicted to cross the BBB.

BeiGene Exhibit 1138.015
BeiGene v. Pharmacyclics
PGR2024-00009

Table 3. Predicted ADMET properties of phytochemical compounds isolated from *C. dichogamus* and FDA approved drugs.

PROPERTIES	L12	L135	L292	L216	L104	L136	L105	L293	L214	L215	DLV	RPV	NVP	ETR	AZT	ABC	TDF
PSA	57.53	58.43	59.67	54.37	56.51	84.73	79.9	33.37	17.07	55.27	118.45	97.42	63.57	120.64	134.07	101.88	146.19
WLOGP	7.23	2.95	3.84	4.25	3.95	2.41	2.5	4.98	6.41	2.69	2.65	4.88	2.42	4.73	−0.52	0.85	−0.04
Absorption																	
Water solubility (logmol/L)	−4.757	−3.897	−4.543	−4.956	−4.8	−4.044	−3.977	−5.177	−6.505	−3.385	−3.159	−4.777	−3.527	−4.303	−3.204	−1.976	−1.859
Caco-2 permeability (log Papp in 10^{-6} cm/s)	1.369	1.287	1.391	1.366	1.447	1.452	0.962	1.543	1..563	1.271	0.329	0.664	1.285	0.079	−0.11	−0.118	−0.135
Intestinal absorption (human) (% absorbed)	97.287	95.906	94.821	96.218	96.989	97.511	95.376	93.67	95.501	98.295	76.683	97.766	97.084	87.101	69.077	77.633	70.097
Skin permeability (log Kp)	−2.73	−3.216	−3.202	−3.36	−3.129	−3.337	−3.918	−2.327	−2.183	−3.127	−2.735	−2.806	−2.819	−2.741	−3.506	−2.735	−2.735
P-Glycoprotein substrate	No	No	No	No	No	No	No	No	No	No	Yes	Yes	No	Yes	No	Yes	No
P-Glycoprotein I inhibitor	No	No	No	Yes	Yes	No	No	Yes	No	Yes	Yes	Yes	No	Yes	No	No	No
P-Glycoprotein II Inhibitor	Yes	No	No	No	No	No	No	No	No	No	No	Yes	No	Yes	No	No	No
Distribution																	
VDss (human) (log L/kg)	−0.784	0.584	0.228	0.479	0.232	0.324	−0.279	0.719	0.613	0.55	1.098	0.503	−0.123	−0.126	−0.423	0.158	0.051
Fraction unbound (human) (Fu)	0	0.289	0.227	0.092	0.111	0.265	0.207	0.085	0.118	0.16	0.264	0.072	0.388	0.094	0.738	0.668	0.695
BBB permeability (logBB)	−0.25	0.264	0.071	−0.099	−0.034	−0.321	−0.17	0.306	0.527	0.169	−1.363	−0.191	0.084	−0.574	−1.21	−1.088	−1.726
CNS permeability (log PS)	−1.011	−2.836	−2.765	−1.533	−1.626	−2.883	−3.296	−2.597	−2.613	−1.993	−2.799	−1.804	−2.865	−1.941	−3.632	−3.572	−3.474
Subcellular localization	Mi	Mi	Mi	Mi	Mi	Mi	MI	Mi	Mi	Mi	Mi	Mi	Mi	Mi	Mi	Ly	Nu

(Continued)

BeiGene Exhibit 1138.016
BeiGene v. Pharmacyclics
PGR2024-00009

Table 3. (Continued)

PROPERTIES	L12	L135	L292	L216	L104	L136	L105	L293	L214	L215	DLV	RPV	NVP	ETR	AZT	ABC	TDF
Metabolism																	
CYP2D6 substrate	No	No	No	No	No	No	No	No	No	No	Yes	No	No	No	No	No	No
CYP3A4 substrate	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	No	No	No
CYP1A2 inhibitor	No	No	No	Yes	No	No	No	Yes	No	No	No	Yes	Yes	Yes	No	No	No
CYP2C19 inhibitor	No	No	No	No	Yes	No	No	Yes	Yes	Yes	Yes	Yes	No	Yes	No	No	No
CYP2C9 inhibitor	No	No	No	No	No	No	No	No	No	No	Yes	Yes	No	Yes	No	No	No
CYP2D6 inhibitor	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
CYP3A4 inhibitor	No	No	No	No	Yes	No	No	No	No	No	Yes	Yes	No	Yes	No	No	No
Excretion																	
Total clearance (log ml/min/kg)	−0.116	0.482	0.749	0.87	0.8778	0.44	0.924	0.977	0.892	0.749	0.441	0.19	0.002	−0.434	0.048	1.161	0.709
Renal OCT2 substrate	No	No	No	No	No	No	No	Yes	Yes	No	No	No	No	No	No	No	No
Toxicity																	
AMES toxicity	No	No	No	No	No	No	No	No	No	No	No	No	No	No	Yes	No	No
hERG I inhibitor	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No
hERG II inhibitor	No	No	No	Yes	No	No	No	No	No	No	Yes	Yes	No	Yes	No	No	No
Hepatotoxicity	No	No	Yes	No	No	No	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Skin sensitization	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No	No

Abbreviation: FDA, Food and Drug Administration.
Key for color coding: Red for highly positive, yes; Green for negative, no.

BeiGene Exhibit 1138.017
BeiGene v. Pharmacyclics
PGR2024-00009

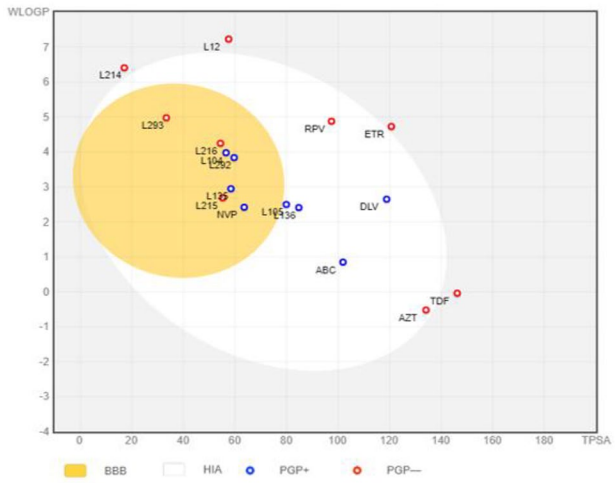


Figure 16. BOILED-Egg. Plot of 10 phytochemical compounds isolated from *C. dichogamus* and FDA approved antiretroviral drugs. FDA indicates Food and Drug Administration.

Among the FDA approved drugs, delaviridine was predicted to be the only substrate for CYP2D6 enzyme, while all the other drugs were predicted to be substrates for CYP2D6 and CYP3A4 (except AZT, TDF and ABC), which is in agreement with previous reports by Gong et al.[68] Similarly, all the phytochemical compounds were predicted to be substrates for CYP3A4 (except depressin, L214), suggesting that these compounds may be metabolized in the liver. Crotonolide E (L104), crotohaimene B (L293), depressin (L214) and crotodichogamoin A (L215) were predicted to be CYP2C19 inhibitors.

Comparing the in silico PK results on CYP and P-gp, it was found out that crothalimene B (L293) and crotodichogamoin A (L215) are predicted to inhibit CYP2C19 and P-glycoprotein. The inhibition of CYP3A4 and P-gp by the phytochemical compounds could decrease the elimination and pumping of other antiretreoviral drugs from the systemic circulation and the cells respectively.

Drug clearance prediction shows that the total clearance of L105, L293 and ABC is the highest followed by L299, L216, L194, L214, L215 and TDF. In addition, L293 and L214 have shown to be renal OCT2 substrates. Renal organic cation transporter 2 is a renal uptake transporter that plays an important role in disposition and renal clearance of drugs and endogenous compounds. OCT2 substrates also have the potential for adverse interactions with co-administered OCT2 inhibitors. Assessing candidate's potential to be transported by OCT2 provides useful information regarding not only its clearance but potential contraindication.

Drug clearance occurs primarily as a combination of hepatic clearance (metabolism in the liver and biliary clearance) and renal clearance (excretion via the kidneys). It is related to molecular weight, hydrophilicity and bioavailability of compounds, and is important for determining dosing rates to achieve steady state concentrations.

One of the main step in drug discovery is ensuring new drug candidates are safe to humans, animals, plants or the environment. Toxicity studies help in determining the harmful effect of drugs. Toxicity studies can be done in vivo, in vitro or in silico. Toxicity studies involving animals have a number of challenges with regard to time, ethical consideration and financial burden. Even though there are efforts to perform in vitro toxicity studies like cytotoxicity studies, these approaches are still costly and time consuming. In comparison to experimental approaches, computational methods of toxicity prediction are considered as fast, cheap and useful methods to analyze, simulate, visualize to predict the toxicity of chemicals.[69] Currently, many software and web servers can predict chemical toxicity before synthesis

AMES test formulated by Bruce Ames is a recognized *in vitro* assay that uses bacteria, *Salmonella typhimurium*, to test whether a given compound is mutagenic and therefore may act as a carcinogen.[70] A positive test indicates that the compound is mutagenic and therefore may act as a carcinogen. It predicts whether a given compound is likely to be Ames positive and hence mutagenic. In our study it is only zidovudine that was found to be Ames positive, which is in agreement with,[71] who reported "template-switch mutagensis" by zidovudine, through its action as a chain terminator during DNA replication. In silico prediction tools for AMES mutagenicity (*Salmonella typhimurium* reverse mutation assay) represent a cost-effective high throughput approach for the prioritization of compounds before experimental testing.[72,73]

Crotodichogamoin B (L216) and three FDA approved drugs (DLV, RPV and ETR) were predicted to inhibit hERG. Inhibition of the potassium channels encoded by the hERG (human ether-a-go-go gene) are the principal causes for the development of acquired long QT syndrome leading to fatal ventricular arrhythmia. Inhibition of hERG channels has resulted in withdrawal of many substances from the pharmaceutical market.

All the FDA approved drugs studied have predicted hepatotoxicity, which is in agreement with previous reports and clinical findings as liver toxicity is one of the most relevant adverse effects of antiretroviral therapy.[74,75] Among the phytochemicals crothalimene A (L292) have predicted hepatotoxicity.

All the phytochemical compounds and FDA approved drugs have not shown skin sensitization and predicted toxicity. Skin sensitization is a potential adverse effect for dermal applied products. The evaluation of whether a product, which may have encountered the skin can induce allergic contact dermatitis is an important safety concern.

Physicochemical properties associated with chemical compounds that have good oral bioavailability, low toxicity and optimum values of physicochemical properties are key parameters for the anti-HIV drug discovery.

Conclusions

In the present study binding mechanism of a phytochemical compounds isolated from *C. dichogamus* in the active site of HIV-1 RT have been elucidated using molecular docking and molecular dynamics simulation studies. Based on our results,

BeiGene Exhibit 1138.018
BeiGene v. Pharmacyclics
PGR2024-00009

we obtained five potential inhibitors of HIV-1 RT, including aleuritolic acid, crotoxide A, Crothalimene A, crotodichogamoin B, and crotonolide E with respective binding energy of -8.48 (Ki = 0.61 μM), -7.73 (Ki = 2.11 μM), -7.48 (Ki = 3.3 μM), -7.42 (Ki = 3.62 μM), -7.31 (Ki = 4.42 μM). These compounds have shown high binding energy as compared to standard FDA approved antiretreoviral drugs. Our computational studies have also shown that these phytochemicals form hydrophobic interactions with ASN 265, GLU 378, GLY 352, HIS 96, ILE 382, SER 268, TRP 266, hydrogen bonding with ARG 355, ARG 356, ARG 358, GLN 269, ILE 94, LEU 92, LYS 350, LYS 353, LYS 374, TYR 232 amino acids in the active site of the enzyme. Thus, we report these top 5 phytochemicals as potentially potent, selective, orally bioavailable, and nontoxic leads based on the ADMET screening and effective binding analysis in the active site of the reverse transcriptase (PDBID: 1REV) for further consideration. The experimental validation of the results should be warranted in future studies.

Acknowledgements

We sincerely thank the United States International University-Africa, University of Nairobi and Kenya Medical Research institute for their support toward the successful completion of the research work.

Author Contributions

Authors' contributed in the conceptualization, methodology, investigation, data analysis and write-up of the manuscript.

Supplemental Material

Supplemental material for this article is available online.

REFERENCES

1. UNAIDS. Global HIV & AIDS statistics – 2020 fact sheet. UNAIDS. https://www.unaids.org/en/resources/fact-sheet. Published 2020. Accessed April 17, 2021.
2. Kitamura K, Honda M, Yoshizaki H, et al. Baicalin, an inhibitor of HIV-1 production in vitro. *Antiviral Res*. 1998;37:131-140.
3. Zhou P, Takaishi Y, Duan H, et al. Coumarins and bicoumarin from Ferula sumbul: anti-HIV activity and inhibition of cytokine release. *Phytochemistry*. 2000;53:689-697.
4. Cichewicz RH, Kouzi SA. Chemistry, biological activity, and chemotherapeutic potential of betulinic acid for the prevention and treatment of cancer and HIV infection. *Med Res Rev*. 2004;24:90-114.
5. Holz-Smith SL, Sun IC, Jin L, Matthews TJ, Lee KH, Chen CH. Role of human immunodeficiency virus (HIV) type 1 envelope in the anti-HIV activity of the betulinic acid derivative IC9564. *Antimicrob Agents Chemother*. 2001;45:60-66.
6. Loya S, Rudi A, Kashman Y, et al. Polycitone A, a novel and potent general inhibitor of retroviral reverse transcriptases and cellular DNA polymerases. *Biochem J*. 1999;344:85-92.
7. Dey B, Lerner DL, Lusso P, Boyd MR, Elder JH, Berger EA. Multiple antiviral activities of cyanovirin-N: blocking of human immunodeficiency virus type 1 gp120 interaction with CD4 and coreceptor and inhibition of diverse enveloped viruses. *J Virol*. 2000;74:4562-4569.
8. Uckun FM, Rajamohan F, Pendergrass S, Ozer Z, Waurzyniak B, Mao C. Structure-based design and engineering of a nontoxic recombinant pokeweed antiviral protein with potent anti-human immunodeficiency virus activity. *Antimicrob Agents Chemother*. 2003;47:1052-1061.
9. Asres K, Seyoum A, Veeresham C, Bucar F, Gibbons S. Naturally derived anti-HIV agents. *Phytother Res*. 2005;19:557-581.
10. Aldhaher A, Langat M, Ndunda B, et al. Diterpenoids from the roots of Croton dichogamus Pax. *Phytochemistry*. 2017;144:1-8.
11. Kokwaro JO. Medicinal plants of East Africa. https://agris.fao.org/agris-search/search.do?recordID=KE2005100575. Published 1976. Accessed April 18, 2021.
12. Jeruto P, Lukhoba C, Ouma G, Otieno D, Mutai C. An ethnobotanical study of medicinal plants used by the Nandi people in Kenya. *J Ethnopharmacol*. 2008;116:370-376.
13. Fratkin E. Traditional medicine and concepts of healing among Samburu pastoralists of Kenya. *J Ethnobiol*. 1996;16:63-97.
14. Matara DN, Nguta JM, Musila FM, Mapenay I. Phytochemical analysis and investigation of the antimicrobial and cytotoxic activities of Croton dichogamus Pax crude root extracts. *Evid Based Complement Alternat Med*. 2021;2021:2699269.
15. Schneidman-Duhovny D, Inbar Y, Polak V, et al. Taking geometry to its edge: fast unbound rigid (and hinge-bent) docking. *Proteins Struct Funct Genet*. 2003;52:107-112.
16. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. PatchDock and SymmDock: servers for rigid and symmetric docking. *Nucleic Acids Res*. 2005;33:W363-W367. doi:10.1093/nar/gki481.
17. Duhovny D, Nussinov R, Wolfson HJ. Efficient unbound docking of rigid molecules. In: Guigó R, Gusfield D, eds. *Lecture Notes in Computer Science* (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics). Berlin, Germany: Springer Verlag; 2002:185-200.
18. Dassault Systèmes BIOVIA. Discovery Studio Visualization.Version 19. San Diego, CA: Dassault Systèmes. https://www.3dsbiovia.com/products/collaborative-science/biovia-discovery-studio/visualization-download.php. Published 2018. Accessed June 17, 2020.
19. Wang Y, Xiao J, Suzek TO, Zhang J, Wang J, Bryant SH. PubChem: a public information system for analyzing bioactivities of small molecules. *Nucleic Acids Res*. 2009;37:W623-W633.
20. PubChem. https://pubchem.ncbi.nlm.nih.gov/. Accessed June 17, 2020.
21. Sander T, Freyss J, Von Korff M, Rufener C. DataWarrior: an open-source program for chemistry aware data visualization and analysis. *J Chem Inf Model*. 2015;55:460-473.
22. Benet LZ, Hosey CM, Ursu O, Oprea TI. BDDCS, the Rule of 5 and drugability. *Adv Drug Deliv Rev*. 2016;101:89-98.
23. Lipinski CA. Lead- and drug-like compounds: the rule-of-five revolution. *Drug Discov Today Technol*. 2004;1:337-341.
24. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv Drug Deliv Rev*. 2001;46:3-26.
25. Daina A, Michielin O, Zoete V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. *Sci Rep*. 2017;7:42717.
26. Cheng F, Li W, Zhou Y, et al. AdmetSAR: a comprehensive source and free tool for assessment of chemical ADMET properties. *J Chem Inf Model*. 2012;52:3099-3105.
27. Ren J, Esnouf R, Hopkins A, et al. The structure of HIV-1 reverse transcriptase complexed with 9-chloro-TIBO: lessons for inhibitor design. *Structure*. 1995;3:915-926.
28. Guex N, Peitsch MC. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*. 1997;18:2714-2723.
29. Meenambiga SS, Rajagopal K, Durga R. In silico docking studies on the components of inonotus sp., a medicinal mushroom against cyclooxygenase-2 enzyme. *Asian J Pharm Clin Res*. 2015;8:142-145.
30. Zhang Z, Li Y, Lin B, Schroeder M, Huang B. Identification of cavities on protein surface using multiple computational approaches for drug binding site prediction. *Bioinformatics*. 2011;27:2083-2088.
31. Huang B. Metapocket: a meta approach to improve protein ligand binding site prediction. *OMICS*. 2009;13:325-330.
32. Wallace AC, Laskowski RA, Thornton JM. Ligplot: a program to generate schematic diagrams of protein-ligand interactions. *Protein Eng*. 1995;8:127-134.
33. Morris GM, Ruth H, Lindstrom W, et al. Software news and updates AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J Comput Chem*. 2009;30:2785-2791.
34. Seniya C, Yadav A, Khan GJ, Sah NK. In-silico studies show potent inhibition of HIV-1 reverse transcriptase activity by a herbal drug. *IEEE/ACM Trans Comput Biol Bioinform*. 2015;12:1355-1364.
35. DeLano WL. The PyMOL molecular graphics system. Scientific Research Publishing. https://www.scirp.org/(S(i43dyn45teexjx455qlt3d2q))/reference/ReferencesPapers.aspx?ReferenceID=732727. Published 2002. Accessed September 9, 2020.
36. Granchi C, Capecchi A, Del Frate G, et al. Development and validation of a docking-based virtual screening platform for the identification of new lactate dehydrogenase inhibitors. *Molecules*. 2015;20:8772-8790.
37. Jorgensen WL, Maxwell DS, Tirado-Rives J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *J Am Chem Soc*. 1996;118:11225-11236.

BeiGene Exhibit 1138.019
BeiGene v. Pharmacyclics
PGR2024-00009

38. Garg S, Anand A, Lamba Y, Roy A. Molecular docking analysis of selected phytochemicals against SARS-CoV-2 M^{pro} receptor. *Vegetos*. 2020;33:766-781.
39. Hayes JM, Archontis G. MM-GB(PB)SA calculations of protein-ligand binding free energies. In: Wang L, ed. *Molecular Dynamics—Studies of Synthetic and Biological Macromolecules*. London: IntechOpen; 2012. doi:10.5772/37107.
40. Ylilauri M, Pentikäinen OT. MMGBSA as a tool to understand the binding affinities of filamin-peptide interactions. *J Chem Inf Model*. 2013;53:2626-2633.
41. Esposito F, Corona A, Tramontano E. HIV-1 reverse transcriptase still remains a new drug target: structure, function, classical inhibitors, and new inhibitors with innovative mechanisms of actions. *Mol Biol Int*. 2012;2012:586401.
42. Kakarala KK, Jamil K. Screening of phytochemicals against protease activated receptor 1 (PAR1), a promising target for cancer. *J Recept Signal Transduct Res*. 2015;35:26-45.
43. Kumar Maurya A, Sarkar N. In-silico inhibition of galectin-1 during HIV-1 pathogensis: a pharmacophore based virtual screening, molecular docking & QSAR studies. http://ethesis.nitrkl.ac.in/5499/1/E-THESIS_1.pdf. Published 2014.
44. Addae-Mensah I, Achenbach H, Thoithi GN, Waibel R, Mwangi JW. Epoxychiromodine and other constituents of Croton megalocarpus. *Phytochemistry*. 1992;31:2055-2058.
45. Peres MT, Delle Monache F, Cruz AB, Pizzolatti MG, Yunes RA. Chemical composition and antimicrobial activity of Croton urucurana Baillon (Euphorbiaceae). *J Ethnopharmacol*. 1997;56:223-226.
46. Fröhlich JK, Froeder AL, Janovik V, et al. Antioxidant capacity, antimicrobial activity and triterpenes isolated from Jatropha isabellei Müll Arg. *Nat Prod Res*. 2013;27:1049-1059.
47. Khiev P, Oh SR, Chae HS, et al. Anti-inflammatory diterpene from Thyrsanthera suborbicularis. *Chem Pharm Bull (Tokyo)*. 2011;59:382-384.
48. Nyasse B, Ngantchou I, Nono JJ, Schneider B. Antifilarial activity in vitro of polycarpol and 3-O-acetyl aleuritolic acid from cameroonian medicinal plants against Onchocerca gutturosa. *Nat Prod Res*. 2006;20:391-397.
49. Liu CP, Xu JB, Zhao JX, et al. Diterpenoids from croton laui and their cytotoxic and antimicrobial activities. *J Nat Prod*. 2014;77:1013-1020.
50. Bezerra AJN, Silva FCO, da Silva AW, et al. Antinociceptive effect of triterpene acetyl aleuritolic acid isolated from Croton zehntneri in adult zebrafish (Danio rerio). *Biochem Biophys Res Commun*. 2021;534:478-484.
51. Singh R, Nath A, Sharma B. Docking studies of HIV-1 reverse transcriptase and HIV-1 integrase with phytocompounds of Carissa Carandas L. *J Clin Res HIV AIDS Prev*. 2019;3:10-19.
52. Seal A, Aykkal R, Ghosh MG. Docking study of HIV-1 reverse transcriptase with phytochemicals. *Bioinformation*. 2011;5:430-439.
53. Zhao BQ, Peng S, He WJ, Liu Y-H, Wang J-F, Zhou X-J. Antitubercular and cytotoxic tigliane-type diterpenoids from Croton tiglium. *Bioorg Med Chem Lett*. 2016;26:4996-4999. doi:10.1016/j.bmcl.2016.09.002.
54. Zheng X, Polli J. Identification of inhibitor concentrations to efficiently screen and measure inhibition Ki values against solute carrier transporters. *Eur J Pharm Sci*. 2010;41:43-52.
55. Jogia MK, Andersen RJ, Párkányi L, Clardy J, Dublin HT, Sinclair ARE. Crotofolane diterpenoids from the African shrub Croton dichogamus Pax. *J Org Chem*. 1989;54:1654-1657.
56. Patel R, Chudasama R, Solanki R, Patel P, Parmar K, Munshi NS. Structure prediction and molecular docking studies of aromatic hydrocarbon sensing proteins TbuT, HbpR and PhnR to detect priority pollutants. *J Environ Sci Health A Tox Hazard Subst Environ Eng*. 2020;55:126-141.
57. Alqahtani A. *Phytochemical investigation of members of the Asparagaceae and Euphorbiaceae families* [PhD dissertation]. Guildford, England: University of Surrey; 2015.
58. Graikou K, Aligiannis N, Chinou I, Skaltsounis A-L, Tillequin F, Litaudon M. Chemical constituents from Croton insularis. *Helv Chim Acta*. 2005;88:2654-2660.
59. Li Y, Carbone M, Vitale RM, et al. Rare casbane diterpenoids from the hainan soft coral sinularia depressa. *J Nat Prod*. 2010;73:133-138.
60. Qidwai T. QSAR modeling, docking and ADMET studies for exploration of potential anti-malarial compounds against Plasmodium falciparum. *Silico Pharmacol*. 2017;5:6.
61. Yang L, Wu ZN, Zhang YB, et al. Three new diterpenoids from Croton laui Merr. et Metc. *Nat Prod Res*. 2017;31:1028-1033.
62. Ma L, Fenalti G, Li J, et al. Structure of the CCR5 chemokine receptor—HIV entry inhibitor Maraviroc complex. *Science*. 2013;341:1387-1390.
63. Han Y, Zhang J, Hu CQ, Zhang X, Ma B, Zhang P. In silico ADME and toxicity prediction of ceftazidime and its impurities. *Front Pharmacol*. 2019;10:434.
64. Kremers P. In vitro tests for predicting drug-drug interactions: the need for validated procedures. *Pharmacol Toxicol*. 2002;91:209-217.
65. Pires DEV, Blundell TL, Ascher DB. pkCSM: predicting small-molecule pharmacokinetic and toxicity properties using graph-based signatures. *J Med Chem*. 2015;58:4066-4072. doi:10.1021/acs.jmedchem.5b00104.
66. Seca AML, Pinto DCGA. Biological potential and medical use of secondary metabolites. *Medicines (Basel)*. 2019;6:66. doi:10.3390/medicines6020066.
67. Zafar F, Gupta A, Thangavel K, et al. Physicochemical and pharmacokinetic analysis of anacardic acid derivatives. *ACS Omega*. 2020;5:6021-6030.
68. Gong Y, Haque S, Chowdhury P, et al. Pharmacokinetics and pharmacodynamics of cytochrome P450 inhibitors for HIV treatment. *Expert Opin Drug Metab Toxicol*. 2019;15:417-427.
69. Yang H, Sun L, Li W, Liu G, Tang Y. In silico prediction of chemical toxicity for drug design using machine learning methods and structural alerts. *Front Chem*. 2018;6:30.
70. Kier LD. Use of the Ames test in toxicology. *Regul Toxicol Pharmacol*. 1985;5:59-64.
71. Seier T, Zilberberg G, Zeiger DM, Lovett ST. Azidothymidine and other chain terminators are mutagenic for template-switch-generated genetic mutations. *Proc Natl Acad Sci*. 2012;109:6171-6174.
72. Hansen K, Mika S, Schroeter T, et al. Benchmark data set for in silico prediction of Ames mutagenicity. *J Chem Inf Model*. 2009;49:2077-2081.
73. Yang H, Li J, Wu Z, Li W, Liu G, Tang Y. Evaluation of different methods for identification of structural alerts using chemical Ames mutagenicity data set as a benchmark. *Chem Res Toxicol*. 2017;30:1355-1364.
74. Rivero A, Mira JA, Pineda JA. Liver toxicity induced by non-nucleoside reverse transcriptase inhibitors. *J Antimicrob Chemother*. 2007;59:342-346.
75. Montessori V, Harris M, Montaner JSG. Hepatotoxicity of nucleoside reverse transcriptase inhibitors. *Semin Liver Dis*. 2003;23:167-172.

BeiGene Exhibit 1138.020
BeiGene v. Pharmacyclics
PGR2024-00009