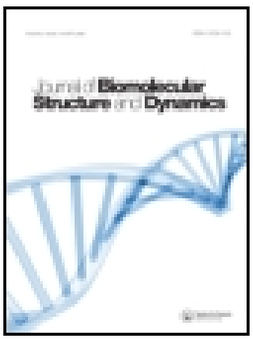


Taylor & Francis
Taylor & Francis Group

Journal of Biomolecular Structure and Dynamics

ISSN: 0739-1102 (Print) 1538-0254 (Online) Journal homepage: https://www.tandfonline.com/loi/tbsd20

Anti-HIV potential of diarylpyrimidine derivatives as non-nucleoside reverse transcriptase inhibitors: Design, synthesis, docking, TOPKAT analysis and molecular dynamics simulations

Vishal K. Singh, Ritika Srivastava, Parth Sarthi Sen Gupta, Farha Naaz, Himani Chaurasia, Richa Mishra, Malay Kumar Rana & Ramendra K. Singh

To cite this article: Vishal K. Singh, Ritika Srivastava, Parth Sarthi Sen Gupta, Farha Naaz, Himani Chaurasia, Richa Mishra, Malay Kumar Rana & Ramendra K. Singh (2020): Anti-HIV potential of diarylpyrimidine derivatives as non-nucleoside reverse transcriptase inhibitors: Design, synthesis, docking, TOPKAT analysis and molecular dynamics simulations, Journal of Biomolecular Structure and Dynamics, DOI: 10.1080/07391102.2020.1748111

To link to this article: https://doi.org/10.1080/07391102.2020.1748111

Anti-HIV potential of diarylpyrimidine derivatives as non-nucleoside reverse transcriptase inhibitors: Design, synthesis, docking, TOPKAT analysis and molecular dynamics simulations

Vishal K. Singh[1], Ritika Srivastava[2], Parth Sarthi Sen Gupta[2], Farha Naaz[1], Himani Chaurasia[1], Richa Mishra[1], Malay Kumar Rana[2] and Ramendra K. Singh[1]*

[1]Bioorganic Research Laboratory, Department of Chemistry, University of Allahabad, Prayagraj- 211002, India

[2]Department of Chemical Sciences, Indian Institute of Science Education and Research Berhampur, Odisha-760010, India

*Corresponding author: Ramendra K. Singh, rksinghsrk@gmail.com

Abstract

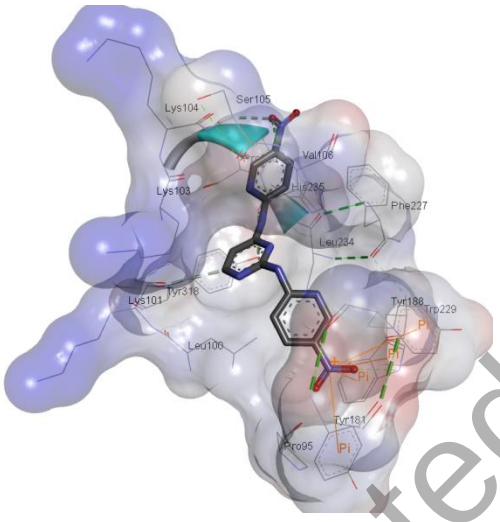
In view of the low toxicity of NNRTIs in comparison to NRTIs, a new series of diarylpyrimidine derivatives has been designed as NNRTIs against HIV-1. *In silico* studies using DS 3.0 software have shown that these compounds behaved as NNRTIs while interacting at the allosteric site of HIV-RT. The designed compounds have shown promising docking results, which revealed that all compounds formed hydrogen bonds with Lys101, Lys103, Tyr181, Tyr318 and π- interactions with Tyr181, Tyr188, Phe227 and Trp229 amino acid residues located in the non-nucleoside inhibitor binding pocket (NNIBP) of HIV-RT protein. The intended molecules have shown high binding affinity with HIV-1 RT, analogous to standard drug molecule - etravirine. TOPKAT results confirmed that the designed compounds were found to be less toxic than the reference drug. Further, employing molecular dynamics simulations, the complexes of the best screened compound 6 and etravirine with the HIV-1 RT protein were analyzed by calculating the RMSD, RMSF, R_g, number of hydrogen bonds, principal components of the coordinates, molecular mechanics-Poisson-Boltzmann surface area-based binding free energy and their decomposition for different interactions. The analysis demonstrated the higher stability of compound 6 than the standard drug etravirine with HIV-1 RT. The interactions like hydrogen-bonding, van-der-Waals, electrostatic and the solvent accessible surface energy have favorable contributions to the complex stability. Thus, the shortlisted designed compound has great promise as a potential inhibitor against HIV-1 RT.

Highlights

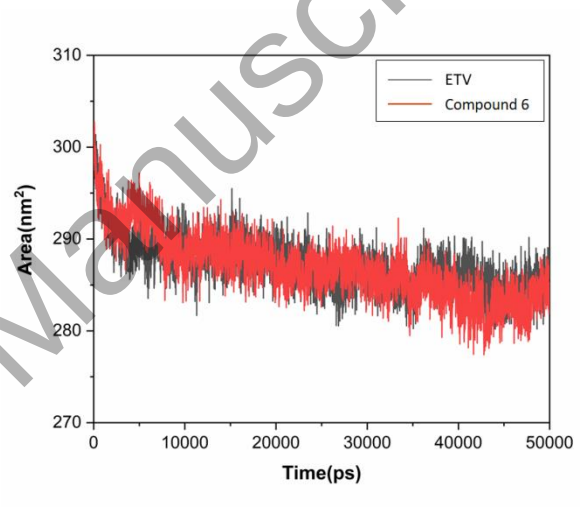
- New diarylpyrimidine derivatives have been designed as potential anti-HIV agents.

- The compounds were docked at the allosteric site of HIV-RT protein (PDB ID: 3MEC) to study the stability of protein-ligand complex.
- Docking studies indicated the stable ligand-protein complexes of all designed compounds.
- The TOPKAT protocol in DS 3.0 software was used to evaluate the toxicity of the designed diarylpyrimidine derivatives.
- Molecular Dynamics studies were performed on best screened compound.

Graphical Abstract



Compound **6**: HIV RT Complex



SASA of Etravirine: HIV RT (black) & Compound **6**: HIVRT (red) Complexes

Keywords: HIV-RT, Docking, QSTR analysis, Molecular dynamics, MM/PBSA, PCA

Abbreviations: ABNR - Adopted Basis Set Newton Raphson, ADME - Absorption Distribution Metabolism Excretion, ART - Antiretroviral therapy, ED - Essential Dynamics, EV - Eigen Vector, DS - Dock Score, DS 3.0 - Discovery Studio 3.0, DTP - Developmental Toxicity Potential, EC_{50} - 50% Effective Concentration, FDA - Food and Drug Administration, GPCR - G-protein coupled receptor, HBA - Hydrogen Bond Acceptor, HBD - Hydrogen Bond Donor, LC_{50} - 50% Lethal Concentration, LD_{50} - 50% Lethal Dose, LIG - Ligand Internal energy, MDS

- Molecular Dynamics Simulation, MM - Molecular Mechanism, MW - Molecular Weight, NNIBP - Non-Nucleoside Inhibitor Binding Pocket, NNRTIs - Non-Nucleoside Reverse Transcriptase Inhibitors, NRTIs - Nucleoside Reverse Transcriptase Inhibitors, PBSA - Poisson-Boltzmann Surface Area, PCA - Principal Component Analysis, PLP - Piecewise Linear Potential, QSTR - Quantitative Structure Toxicity Relationship, R_g - Radius of gyration, RMSD - Root-Mean Square Deviation, RMSF - Root-Mean Square Fluctuation, RT - Reverse Transcriptase, SASA - Solvent-Accessible Surface Area, TPSA - Total Polar Surface Area

1. Introduction

The reverse transcriptase (RT) enzyme of HIV plays a vital role in viral replication and, thus becomes a dedicated target for the discovery of new HIV-RT inhibitors [Jonckheere et al., 2000, Yisma et al., 2014, Bhole et al., 2020]. Till date, several RT inhibitors belonging to the categories either nucleoside reverse transcriptase inhibitors (NRTIs) or non-nucleoside reverse transcriptase inhibitors (NNRTIs) have been discovered [Boone et al., 2006]. Currently, NNRTIs have been used as an important ingredient of drug regimen for HIV-1 treatment in antiretroviral therapy (ART) [Kaufmann et al., 2000, Vella et al., 2000, Wang et al., 2019] due to their high potency and low toxicity. Six NNRTIs, **i-vi** (Fig. 1) have been approved by the US Food and Drug Administration (FDA) and are currently being used in the treatment of HIV/AIDS [Shaung et al., 2019, Gu et al., 2018].

Diarylpyrimidine derivatives, i.e., rilpivirine (**iv**) and etravirine (**v**) represent a new generation of NNRTIs having high potency against wild and mutant type HIV strains [Sarafianos et al., 2009, Chen et al., 2011, Zhan et al., 2016, Zuo et al., 2018]. Therefore, these two NNRTIs have a better chance to reduce the problem of drug resistance, which was observed in earlier NNRTIs. However, resistance mutation in the new generation NNRTIs, i.e., rilpivirine (**iv**) and etravirine (**v**) have also been observed in patients [Anta et al., 2013, Wainberg et al., 2016]. So, the scientific community is still putting serious efforts to discover additional NNRTI drugs, with different heterocyclic scaffolds, belonging to the new generation of molecules having high specificity and low toxicity [Yang et al., 2016, Huang et al., 2015, Chen et al., 2015, Wu et al., 2014].

In continuation of our previous work on first-generation NNRTIs of oxathiadiazole derivatives against HIV-1 [Yadav et al., 2019, Kumari et al., 2016, Singh et al., 2015], we,

hereinreport the design and anti-HIV potential of some new diarylpyrimidine derivatives, **1-8,** as second-generation NNRTIs (Fig. 2) having different substituents on aromatic rings A, B and C.

All these molecules have been designed through docking simulations and essential drug-like characteristics using computational methods. All molecules have been docked in non-nucleoside inhibitor binding pocket (NNIBP), the allosteric site of HIV-RT protein to prove their behaviour as potential NNRTIs.

These substituted diarylpyrimidine derivatives, **1-8**, have been synthesized from 2-4-dichloro-pyrimidine. In compounds **1-8**, modifications were introduced in rings A, B and C of diarylpyrimidine. Introduction of different substituents at 4/5/6 positions of rings B and C of designed compounds established that these groups (hydrophobic/hydrophilic) affect the drug like properties and *in silico* anti-HIV potency of all molecules.

The analysis and comparison of molecular dynamics (MD) simulations and molecular mechanics-Poisson-Boltzmann surface area (MM/PBSA) based binding free energies between the target-ligand and target-etravirine complex and demonstrated a better stability of the former than the latter, which was characterized by a larger contribution of hydrogen bonding, van der Waals, electrostatic interactions and lower RMSD values, resulting in a larger negative binding free energy between the target-ligand complex [Cong et al., 2020, Triki et al., 2019, Sk et al., 2020]

2. Materials and methods

2.1. Design of diarylpyrimidine derivatives as NNRTIs

Pyrimidine moiety plays a significant role in biological system and also works as a building block of nucleic acids. So, considering the importance of pyrimidine nucleobases, 20 pyrimidine derivatives have been designed as NNRTIs against HIV-1 on the basis of Lipinski's rule of five using Moleinspiration and Chemdraw software. After performing the *in silico* structure based assessment, eight (**1**-**8**) promising molecules have been selected for further studies.

Some physicochemical descriptors, like molecular weight (MW), hydrogen bond donor (HBD), hydrogen bond acceptor (HBA), and the logarithmic partition coefficient in octanol/water, and logP (lipophilicity) have been used to investigate or evaluate the drug like properties of the molecules under study [Lipinski et al., 2015].

2.2. Docking analysis

Molecular docking simulation is a computational process used to evaluate the binding efficiency of derived ligands to the target receptor protein by keeping the target receptor protein rigid and the varying conformations and poses of ligands. Docking simulation was performed by Discovery Studio (DS) 3.0 software (Accelrys Ltd, USA). The x-ray crystal structure of wild type HIV-1 RT protein bound to the etravirine was retrieved from the protein data Bank (www.rcsb.org) (PDB ID:3MEC) and used as a docking receptor protein [Yadav et al., 2019].

During the docking process from the receptor target protein 3MEC, ligand was extracted, missing hydrogens added and the position of atoms in the protein was optimized using CHARM force field and the Adopted Basis Set Newton Raphson (ABNR) method available in the protocol of DS 3.0 software using default parameters. The ligand preparation and docking process have been according to the previously published procedure using the DS, 3.0 protocol default parameters [Yadav et al., 2019, Muegge et al., 1999, Lagos et al., 2008, Singh et al., 2016]. Finally, the hydrogen bonding and non-covalent interactions like π-π, π-cation interactions and all scoring functions, such as PLP1, PLP2, LIG, binding energy, EC_{50}, Dock scores, etc., generated during docking simulation, were evaluated that helped predict the exact pose of the designed ligand to the target protein receptor.

2.3. QSTR analysis

All diarylpyrimidine derivatives were also subjected to QSTR (Quantitative Structure Toxicity Relationship) studies to predict their pharmaceutical potential. TOPKAT protocol embedded in the DS 3.0 software was used for QSTR analysis.

2.4. MD simulations

MD simulations were carried out for the target-ligand complex and target-etravirine complex for a period of 50 ns using GROMACS (GROningenMAchine for Chemical Simulations) v5.1 molecular dynamics package [Pronk et al., 2013].The unit cell defined as a cubical box, with a minimal distance of 10 Å from the protein surface to the edges of the box, was solvated using the Simple Point Charge (SPC) water model GROMOS96 53a6 force field [Oostenbrink et al., 2004]. Counter-ions were added to make every system electrically neutral at a salt concentration of 0.15 mol/L. Before the MD run, each system was subjected to energy minimization by employing the steepest descent integrator for 5000 steps with force convergence of <1000 kcal mol^{-1} nm^{-1}.

Thereafter, each protein-ligand complex was equilibrated for 1 ns using canonical (NVT) and isothermal-isobaric (NPT) ensembles. During equilibration, each system was coupled with the Berendsen temperature and Parrinello-Rahman pressure controllers, respectively, to maintain temperature 300 K and pressure 1 bar. The Particle Mesh Ewald (PME) algorithm [Essmann et al., 1995] was employed to deal with the long-range Coulomb interactions with a Fourier grid spacing of 0.12 nm. The short-range van der Waals interactions were given by the Lennard-Jones potential with a cut-off distance of 1 nm. All bond lengths were constrained by the linear constraint solver (LINCS) method [Hess et al., 1997].

Subsequently, 50 ns were considered for production under micro-canonical ensemble by relaxing the couplings with the thermostats. In principle, the same protocol was applied to both systems. A time step of 2 fs was used and the coordinates were saved every 10 ps during the production run. For the structural analyses of every system, the resultant MD trajectories saved at the interval of 100 ps were analyzed using the built-in modules of GROMACS and visual molecular dynamics (VMD 1.9.1) [Schuler et al., 2001]. The 2-D plots depicting the intrinsic dynamical stabilities captured by the root mean square deviation (RMSD), root mean square fluctuation (RMSF), hydrogen bond distribution and radius of gyration (R_g) of the complexes were generated by Grace 5.1.23 program.

2.5. Principal Component Analysis (PCA)

Essential dynamics (ED) or principal component analysis (PCA) is an efficient statistical method applied to reduce the number of dimensions needed to describe protein molecular dynamics through a decomposition process [Humphrey et al., 1996]. The Gromacs module tool called g_covar was used to yield the eigenvalues and eigenvectors by calculating and diagonalizing the covariance matrix, whereas the g_anaeig tool was used to analyze and plot the eigenvectors [Swain et al., 2018].

2.6. Binding Free Energy Calculation

Molecular Mechanics/Poisson-Boltzmann Surface Area (MM/PBSA) widely used for free energy calculation from MD trajectory [Aalten et al., 1995, Wang et., 2018]. The binding free energy (ΔG_{bind}) in a solvent medium was calculated as follows:

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

Where $G_{complex}$is the total free energy of the substrate-protein complex, $G_{protein}$ and G_{ligand} are the total energies of protein and substrate alone in a solvent, respectively.

The free energies for each individual $G_{complex}$, $G_{protein}$ and G_{ligand} were estimated by:

$$\mathbf{G_p = E_{MM} + G_{solv}}$$

Where p can be protein, ligand, or complex. E_{MM} is the average molecular mechanics potential energy in vacuum and G_{solv} is the solvation free energy.

The molecular mechanics potential energy was calculated in the vacuum as follows:

$$\mathbf{E_{MM} = E_{bonded} + E_{non\text{-}bonded} = E_{bonded}(E_{int}) + E_{vdw} + E_{elec}}$$

Where E_{bonded} or (E_{int}) is the total bonded interaction, which includes all bonded interactions like bond, angle, dihedral and improper interactions; $E_{non\text{-}bonded}$ is the total non-bonded interaction consisting of both van der Waals (E_{vdw}) and electrostatic (E_{elec}) interactions. E_{bonded} is always taken as zero.

The solvation free energy (G_{solv}) was estimated as the sum of electrostatic solvation free energy (G_{polar}) and nonpolar solvation free energy ($G_{non\text{-}polar}$) as given below:

$$\mathbf{G_{solv} = G_{polar} + G_{non\text{-}polar}}$$

Where G_{polar}, the polar solvation energy, was determined using the Poisson-Boltzmann (PB) linear equation and the non-polar contribution, $G_{non\text{-}polar}$ was estimated from the solvent-accessible surface area (SASA) as per the following equation:

$$\mathbf{G_{non\text{-}polar} = \gamma SASA + b}$$

Where γ (coefficient related to surface tension of the solvent) = 0.02267 kJ/mol/Å^2 or 0.0054 kcal/mol/Å^{2}and b = 3.849 kJ/mol or 0.916 kcal/mol

The binding free energies for all four complexes were calculated based on 5000 snapshots taken at an equal interval of time from 50 ns MD simulations. The per-residue energy contribution was also computed to understand the contribution of individual amino acids to the total binding energy.

3. Results and discussion

3.1. Physicochemical and drug-like characteristics of diarylpyrimidine derivatives

All diarylpyrimidines (**1**-**8**) have been designed as probable NNRTIs using computational methods. The physicochemical properties of diarylpyrimidine derivatives **1**-**8** and the reference drug etravirine were evaluated using Moleinspiration and Swiss ADME online software [Daina et al., 2017].

All physicochemical data and pharmaceutical properties of compounds **1**-**8** were calculated using Lipinski's rule of five. To evaluate the drug-likeness of any compound,

Lipinski's rule is a thumb rule, which helps to predict the pharmacological and biological activity. All compounds **1**-**8** can ionize at definite pH due to the presence of H-bond donor and H-bond acceptor sites, which helps to increase the solubility of compounds. The compounds having more than one violation are rejected for further studies because of difficulty in solubility and bioavailability.

Lower TPSA values of compounds were expected to exhibit the better cellular internalization in comparison to the reference drug etravirine. Suitable modification in the active moiety was introduced in order to find out the higher drug-likeness scores as explained by the thumb rule of Lipinski. All physicochemical data of diarylpyrimidine derivatives **1**-**8** are summarized in Table 1 and represented as radar graph in Fig.3.

All compounds **1**-**8** were found within the range of Lipinski's rule and expected to behave as probable NNRTIs against wild type HIV-1, like etravirine. The drug-likeness scores of all compounds **1**-**8** and the reference drug etravirine are given in Table 2.

Lower value of lipophilicity indicated that all compounds can easily pass through cell membrane like the reference drug etravirine. Lower TPSA values of all compounds confirmed their better cellular internalization in comparison to reference drug except compounds **6** and **8**.

The drug-likeness scores for proper bioactivity were also evaluated for G-protein coupled receptor (GPCR) ligand, ion channel modulator, kinase inhibitor and nuclear receptor ligand (Table 2). The compounds having drug-likeness bioactivity scores more than zero are found to be pharmaceutically active, whereas the bioactivity scores of compounds between 0.00-0.50, show moderate activity and the scores less than -0.50, show inactivity.

All diarylpyrimidine derivatives **1**-**8** showed better drug-likeness scores than the reference drug etravirine except compounds **2** and **3** in the case of GPCR, compounds **7** and **8** in the case of kinase inhibitor, compounds **4** and **5** in the case of nuclear receptor ligand and compounds **7** and **8** in the case of enzyme inhibitor for bioactivity. These compounds showed moderate *in silico* bioactivity in comparison to other compounds and the reference drug etravirine as shown in Table 2.

3.2. Synthesis of diarylpyrimidine derivative

Substituted diarylpyrimidine derivatives, **1-8**, have been synthesized from 2-4-dichloro-pyrimidine. The further modifications of the replacement of both chloro group of 2-4-dichloro-pyrimidine via ring B and C with different substituents at 4/5/6 positions are clearly outlined in Scheme 1. Purification and characterization data of these compounds have been provided in supporting information.

3.3. Docking analysis of diarylpyrimidine derivatives

To better understand the binding mode of diarylpyrimidine derivatives **1**-**8** and etravirine, molecular docking was performed using DS 3.0 software. The compounds formed H-bonds with Lys101 and Lys103 amino acid residues, π-bonds with Tyr181, Tyr188, Phe227 and Trp229 at the allosteric site in the hydrophobic NNIBP of HIV-RT, with the exception of compounds **7** and **8**.

H-bonds, π-π (hydrophobic), and non-polar π-cation (non covalent) interactions were used to explain the docking results. The existence of these non-covalent interactions implies that the interatomic distances are within the range of 6 Å. The docking results of diarylpyrimidine derivatives **1**-**8** with HIV-RT protein are represented in Table 3.

Docking studies revealed that compound **1** having chloro group at 4^{th} position of aryl rings B and C, and compound **2** having fluoro group at 6^{th} position of aryl rings B and C interacted in similar fashion like the reference drug with HIV-RT and formed a H-bond with Lys103 at the distance of 2.8 Å.

Similarly, compound **3** with bromo group at 6^{th} position and compound **4** with bromo group at 5^{th} position of both aryl rings B and C formed a H-bond (2.9Å) with His235 residue of HIV-RT protein. Additionally, compound **4** also interacted with Tyr188 residue to form a π-bond at a distance of 6Å.

Compound **5** having chloro group at 5^{th} position of both rings B and C formed two H-bonds, i.e., one with Lys101 and another with Lys103 at a distance of 2.6 Å and 2.8 Å, respectively.

Introduction of the nitro group at 5^{th} position of both aryl rings B and C in place of halogen atom stabilized the complex of compound **6** with HIV-RT as the compound **6** interacted with amino acids His235 and Trp229 through the H-bonds. Furthermore, compound **6** also showed four π-cation (non-covalent) interactions with Trp229 (two π-cation bonds), Tyr181 (one

π-cation bond) and Tyr188 (one π-cation bond) amino acid residues at the allosteric site of HIV-RT protein.

Docking interactions of diarylpyrimidine derivatives are shown in Fig. 4. The binding of all these compounds with HIV-RT protein revealed their behaviour as NNRTIs, just like the reference drug.

During docking simulation some scoring functions were generated, which helped in predicting the suitable orientation of ligands within the target protein. The stability of ligand-protein complexes was evaluated on the basis of scoring functions, like PLP1, PLP2, DS, LIG, Ludi values, binding energy and predicted EC_{50}. Scoring functions of all diarylpyrimidine derivatives **1**-**8** are presented in Table 4.

Piecewise Linear Potential (PLP) functions of a ligand are used to correlate the binding affinity of the ligand to a target receptor protein. The compound having higher PLP scores showed stronger binding with receptor protein. All compounds **1**-**8** exhibited considerable binding affinity to the receptor protein (HIV-RT) via PLP1 and PLP2 scores ranging from 96.08 (compound **7**) to 109.74 (compound **6**) and 83.07 (compound **7**) to 95.68 (compound **6**).

The Ligand Internal energy (LIG) consisting of van der Waals and electrostatic interactions was used to evaluate the binding efficiency of ligand to the target protein. All diarylpyrimidine derivatives had much better LIG values than the reference drug etravirine, as shown in Table 4, and thus exhibited higher binding efficiency than the reference drug.

Dock Score (DS), another factor generated from docking simulation again supported the interaction of the designed ligands with the HIV-RT protein in a better way. All compounds **1**-**8** showed a significant value of dock score in comparison to the reference drug etravirine (Table 4).

Ludi2 and Ludi3 are empirical scoring functions derived from the Ludi algorithm and used to select the exact conformations of protein-ligand complexes. From Ludi2 value, the binding energy of the target compound and from Ludi3 value, predicted EC_{50} values of compounds were calculated [Lagos et al., 2008, Kumar et al., 2010]. The binding energy and predicted EC_{50} values of all compounds **1**-**8** were found to be comparable with known drug etravirine as shown in Table 4.

From all scoring values, it was confirmed that compounds **1**-**8** had a similar mode of binding to the RT allosteric site and also had an excellent binding affinity with HIV-RT protein, just like the reference drug used in the docking studies.

3.4. QSTR analysis of diarylpyrimidine derivatives

All compounds **1**-**8** were also subjected to their *in silico* toxicity risk assessment using TOPKAT protocol in DS 3.0 software and data are summarized in Table 5. All the designed molecules and the reference drug were evaluated for carcinogenicity, which helped to recognize the common structural features between the molecules and reference drug. Carcinogenicity of all designed compounds and reference drug etravirine was calculated using two models - NTP Carcinogenicity Call (Female mouse) and NTP Carcinogenicity Call (Male rat) in TOPKAT module. During TOPKAT analysis, neither diarylpyrimidine derivatives; **1**-**8** nor the reference drug showed carcinogenicity become of their negative discriminant scores (Table 5).

TOPKAT mutagenicity predictor (Ames Heteroaromatics Module) was used to predict the mutagenicity of the compounds under study. The compounds with the negative value of discriminant scores were found to be non-mutagenic during TOPKAT mutagenicity prediction process.

TOPKAT analysis revealed that the *in silico* toxicity values, like DTP (Developmental Toxicity Potential), LD_{50} (Lethal Dose causing 50% death of test animals or cells), LC_{50} (50% Lethal Concentration) and EC_{50} (50% Effective Concentration) of all compounds, were found to be less than the reference drug etravirine. It unambiguously indicated that all diarylpyrimidine derivatives **1**-**8** should be less toxic as compared to the reference drug and, in turn, the designed molecules should have a higher safety index.

TOPKAT results also revealed that all compounds exhibited a lower value of LogP (lipophilicity) than the reference etravirine as shown in Table 5, which showed the hydrophobic character of the designed compounds with non-polar atoms, a necessary condition for NNRTIs.

3.5. Molecular dynamics trajectory analysis

The proteins suffer a significant conformational change during the course of interaction with a drug molecule [Amadei et al., 1993, Chou et al., 2006], thus Molecular Dynamics Simulations (MDS) play a significant role to understand the internal motions, conformational changes, stability, etc. of protein-ligand complexes. Using the resulted MDS trajectories, the root-mean square deviation (RMSD), root-mean square fluctuation (RMSF), radius of gyration

(R_g), number of hydrogen bonds, principal component analysis (PCA) and binding free-energy of the complexes were computed to study their structural stabilities, binding modes and binding strengths.

The MDS were performed on HIV-1 RT protein with compound 6 and reference drug etravirine.

3.5.1. Conformational stability

The RMSD plots against simulation time are shown in Figure 5 where the small fluctuations indicate attainment of a stable conformation and vice-versa. While the RMSD plots of the etravirine complex has a relatively larger deviation and strong oscillation (till 28 ns) than that for compound **6**, a stronger binding of the latter with the protein HIV-1 RT than the reference drug is suggested.

3.5.2. Residue flexibility analysis

To evaluate the flexibility of the complexes, fluctuations of each amino acid residue presented by the RMSF plots as a function of residue number are shown in Fig. 6. A large RMSF value indicates the flexibility of the protein structure, loose bonding or the presence of loops; contrarily a small value indicates stability and also the presence of secondary structures such as sheets and helices. While comparing between the two complexes, the mean values of RMSF were found to be in the same, which demonstrates a balance between stability and flexibility, necessary for their activities.

3.5.3. Compactness analysis

The radius of gyration (R_g) accounts for the conformational variation, compactness and tertiary structural volume of the protein-ligand complexes. In general, an expanded structural form of protein shows a higher R_g value than the corresponding globular or compact form. A larger value of R_g describes more elongation, whereas a lower value indicates more compact structure. The plots of R_g versus time (Fig. 7) are quite smooth and almost constant for both complexes, having the mean values at 2.98 nm and 3.34 nm, for compound 6 (red) and etravirine (black), respectively, with HIV-1 RT. A smaller mean R_g value of compound **6** reiterates to the previous finding that it forms a more compact and stable complex with HIV-1 RT, arising from a stronger interaction between them than that for etravirine.

3.5.4. Analysis of hydrogen bonding

Towards the stability of a target-inhibitor complex, the presence of a large number of inter-molecular hydrogen bonds is very crucial. Fig. 8 gives the number of hydrogen bonds formed during the simulations, which is more in case of compound **6** (red) than etravirine (black) with HIV-1 RT. Intactness of hydrogen bonds during dynamics is an indication of gaining more stable conformation and stronger binding between compound **6** and HIV-1 RT. On an average, compound **6** formed approximately three hydrogen bonds, whereas etravirine exhibited only two hydrogen bonds with HIV-1 RT.

3.5.5. Solvent accessible surface area

Solvent Accessible Surface Area (SASA) accounts for the surface area of a protein-ligand complex which directly interacts with the solvent molecules. The increase in SASA denotes relative expansion.

The SASA of the HIV-1 RT complexes of compound **6** and etravirine is ranging between 280 to 303 nm^2 and 285 to 302 nm^2, respectively (Fig. 9).

Within initial few ps, the SASA value drops suddenly and then gradually decreases for both complexes. A similar pattern of fluctuation is noted in both complexes, which is very small. Comparatively, to the end od 45 ns, the compound **6**:HIV-1 RT complex has lower SASA values implying more stability than the etravirine:HIV1 RT complex.

3.6. Essential dynamics

For principal component analysis (PCA), essential dynamics (ED) was employed to represent the directions of principal motions by a set of eigenvectors (EV) called principal or essential modes. PCA is a linear transformation that extracts the most important special data using a covariance matrix using atomic coordinates that describe the accessible degrees of freedom (DOF) of a protein, e.g., the cartesian coordinates that define atomic displacements in each conformation in MD trajectory. This covariance matrix is diagonalized to extract a set of eigenvectors and eigenvalues that reflect the correlated concert motion of the molecule [Khan et al., 2016]. Each MD trajectory was projected onto the phase space to yield a spectrum of EVs, which depicted the vectoral representation of each single component in motion. Each EV holds an eigenvalue that describes the energetic contribution of each component to the motion. The PCA results are represented in Fig.10 and 11.

It was observed that 90-95% of the backbone motion was covered by the first 20 EVs where an exponentially decaying curve of eigenvalues was obtained against the EVs (Fig. 11). The differential scattering of atoms in contour-like plots specifies the occurrence of conformational changes in the complexes in agreement with the other MD analysis.

Taken the first two PCs into consideration, simulation results revealed a higher subspace dimension for the etravirine:HIV-1 RT complex (Fig 10). On the contrary, the compound **6**:HIV-1 RT complex covered the least subspace and showed the least variations. In the 2-D projection plots of trajectories, the etravirine:HIV-1 RT complex showed higher trace values of the covariance matrix than the compound **6**:HIV-1 RT complex.

The overall analysis manifests again that the compound **6**:HIV-1 RT complex had higher stability as reflected from its least conformational changes due to decreasing collective motions than the etravirine:HIV-1 RT complex. Emphatically, the PCA and RMSD analyses elicited the same characteristics about structural fluctuations of the protein-inhibitor complexes.

3.7. Binding free energy and residue interaction energy

MD trajectories of both the complexes were used to calculate binding free energy using MM/PBSA method. Currently, this method has been widely applied in various studies including stability, protein-ligand interactions, drug design, and re-scoring [Genheden et al., 2015, Wang et al., 2018]. Among existing different classical methods, the MM/PBSA approach has been considered as a very efficient and reliable method to study crucial molecular recognition processes [, Kollman et al., 2000, Wang et al., 2019, Gupta et al., 2020].

According to the constituents of the inhibitors, different interactions, e.g., hydrophobic, hydrogen bond, electrostatics and pi-pi interactions prevailed in the corresponding HIV-1 complexes. In this study, 5000 snapshots were extracted at equal intervals from the 50 ns MD trajectories to compute binding free energy for both complexes. The contributions of different interactions were either positive or negative to the overall binding free energy and summarized in Table 6 for both complexes.

The compound **6**:HIV-1RT complex possessed the higher negative binding energy (-112.81 kJ mol^{-1}) than etravirine:HIV-1 RT (-111.4 kJ mol^{-1}). Decomposition into separate energy terms revealed immediately that the polar solvation energy opposes the binding of inhibitors to HIV-1 significantly, and thereby reducing the total binding energy in both complexes due to the positive energy contributions (Table 6).

Among the various interactions, van der Waal energy (ΔE_{vdW}) was the most and solvent-accessible surface area (SASA) energy was the least favourable contributions towards the negative binding free energy of both complexes (Table 6). The greater contribution of electrostatic energy (ΔE_{elec}) and lesser penalty of polar solvation energy in the compound **6**: HIV-1 RT complex gave rise to slightly more negative binding free energy than the etravirine: HIV-1 RT complex, making the former relatively more stable than the latter.

In addition to the information presented in Table 6, the contributions of individual amino acids to the binding free energy ($\Delta G_{bind)}$ were also computed using the MM/PBSA method and presented in Figure 12. For clarity, only actively contributing residues towards the positive and negative binding energies are highlighted.

The per-residue interactional energy profiles revealed that Pro97, Lys103, Glu169, Asn175, Asp177, Lys233, Met230 and Glu308 actively participated in interaction to give rise to a stronger binding and stability in both complexes. Except Thr165, Leu234 and Asn343, most of the amino acid residues showed a negative binding energy contribution to the stability of the complexes.

Interestingly, amino acid residues contributing favorably or unfavorably in both complexes were the same; the only difference was that a common amino acid residue Lys103 contributed more favorably to the compound 6: HIV-1 RT complex than to the reference drug etravirine: HIV-1 RT complex. The most favourably contributing residue for both the complexes was Lys223 having a negative binding energy of -13.23 kJ/mol and -12.98 kJ/mol, for the compound 6: HIV-1RT and etravirine: HIV-1RT complexes, respectively.

In order to know the conformation of the compound 6 after MD simulations, we have captured snapshots of the protein-ligand complex every 10 ns, starting from 0 ns to 50 ns. From the Fig. 13, it was observed that the compound 6 (pink color) remained stable in the binding site of the target protein throughout 50 ns MD simulation process.

HIV RT consists of two domains; polymerase and Rnase H, in which the catalytic site is located on the former site. While the NNRTIs bind in the binding site (polymerase domain), etravirine complexes are mostly seen in the Rnase H domain, Fig. 13 and 14, respectively, Any flexibility indicated in the RMSF for the Rnase H domain does not affect the stability of the compound 6-HIV RT complex, marked by lower RMSD values, as the binding site is located far from the

RNase H domain, which starts from residue 425 and is connected with a loop. (Kohlstaedt et al., 1992)

4. Conclusion

A new series of diarylpyrimidine derivatives **1**-**8** has been developed as potent and selective inhibitors of HIV-1 RT by using *in silico* structure-based approach and is synthesized in the laboratory. Docking results revealed that the designed compounds showed high potency and some of the compounds were predicted be more potent than the known inhibitor etravirine. A good correlation had been observed between the predicted EC_{50}, various energy terms and descriptors for all newly reported compounds against HIV-1 RT. Investigation of these relationships shall definitely be useful for the prediction of the activity of novel derivatives. Based on structural fluctuations, dominating interactions and binding free energies of the complexes of HIV-1 RT, the molecular dynamics simulation results demonstrated higher stability of the shortlisted compound **6** and a better inhibitor capability in comparison to the reference drug. The present study, thus, provided a valuable advance in the search for novel second generation NNRTIs and can be considered as a starting point for lead optimization.

Acknowledgement

Financial assistance to Vishal K Singh in the form of Junior ResearchFellowship (Ref No: 349/CSIR-UGC NET DEC. 2017) by University Grants Commission is sincerely acknowledged. Ritika Srivastava and Parth Sarthi Sen Gupta sincerely thank IISER Berhampur, Odisha for postdoctoral fellowship as well as computational support.

Conflict of interest

The authors report no conflicts of interest.

References

A. Amadei, A. B. M. Linssen, H. J. C. Berendsen, 1993. Essential dynamics of proteins. *Pro.: Str., Fun., and Gen.*, 17, 412-425. doi:org/10.1002/prot.340170408

A. Daina, O. Michielin, V. Zoete, 2017. Swiss ADME: a free web tool to evaluate pharmacokinetics, drug likeness and medicinal chemistry friendliness of small molecules. *Sci. Rep*., 7, 42717. doi: org/10.1038/srep42717

A. Singh, D. Yadav, M. Yadav, A. Dhamanage, S. Kulkarni, Ramendra K. Singh, 2015. Molecular Modeling, Synthesis and Biological Evaluation of N-Heteroaryl Compounds as Reverse Transcriptase Inhibitors Against HIV-1. *Chem. Biol. Drug Des*., 85, 336-347. doi: org/10.1111/cbdd.12397

A. Singh, M. Yadav, R. Srivastava, 2016. Design and anti-HIV activity of aryl sulphonamides as non-nucleoside reverse transcriptase inhibitors. *Med. Chem. Res.*, 25, 2842-2859. doi: org/10.1007/s00044-016-1707-7

B. Huang, X. Liang, C. Li, W. Chen, T. Liu, X. Li, Y. Sun, L. Fu, H. Liu, E. D. Clercq, C. Pannecouque, P. Zhan, X. Liu, 2015. Fused heterocycles bearing bridgehead nitrogen as potent HIV-1 NNRTIs. Part 4: design, synthesis and biological evaluation of novel imidazo[1,2-a] pyrazines. *Eur. J. Med. Chem*., 93, 330-337. doi: org/10.1016/j.ejmech.2015.02.022

C. A. Lipinski, F. Lambardo, B.W. Dominy, P.J. Fenney, 2012. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev*., 64, 4-17. doi: org/10.1016/j.addr.2012.09.019

C.F. Lagos, J. Caballero, F.D. Gonzalez-Nilo, D.P.C. Mahana,T.Perez-Acle, 2008 Docking and quantative structure-activity relationship studies for the bisphenyl benzimidazole family of non-nucleoside inhibitors of HIV-1 reverse transcriptase. *Chem. Biol. Drug Des*., 72, 360-369. doi: org/10.1111/j.1747-0285.2008.00716.x

C. Oostenbrink, A. Villa, A. E. Mark, W. F. Van Gunsteren, 2004. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *J. of Com. Che.*, 25, 1656-1676. doi: org/10.1002/jcc.20090

C. Wang, D. Greene, L. Xiao, R. Qi, R. Luo, 2018. Recent Developments and Applications of the MMPBSA Method. *Fro. in Mol. Bio*., 4, 1-18. doi:org/10.3389/fmolb.2017.00087

C. Wang, F. Schmich, S. Srivatsa, J. Weidner, N. Beerenwinkel, A. Spang, 2018 .Context-dependent deposition and regulation of mRNAs in P-bodies. *Cel. Bio. Chr. and Gen. Exp*., doi: 10.7554/eLife.29815

D. M. F. va.Aalten, J. B. C. Findlay, A. Amadei, H. J. C. Berendsen, 1995. Essential dynamics of the cellular retinol-binding protein evidence for ligand-induced conformational changes. *Pro. Eng. Des. and Sel*., 8, 1129-1135. doi: org/10.1093/protein/8.11.1129

D. Triki, M. Kermarrec, B. Visseaux, D. Descamps, D. Flatters, A.-C. Camproux, & L. Regad, 2019. Exploration of the effects of sequence variations between HIV-1 and HIV-2 proteases on their three-dimensional structures. *J. of Bio. Str. and Dyn.*, 1–20. doi:10.1080/07391102.2019.1704877

E.Yisma, B. Dessalegn, A. Astatkie,N. Fesseha, G. Benagiano, C. Bastianelli, S. Brusamento, H. Elmoniry, M. Velthoven, U. van, Pape, 2014. Global report: UNAIDS report on the global AIDS epidemic 2013, *Rep. Hea*., 10, 23. doi: org/10.1186/1742-4755-10-23

G. Kumai, Ramendra K. Singh, 2015. Molecular Modeling, Synthesis, and Anti-HIV Activity of Novel Isoindolinedione Analogues as Potent Non-nucleoside Reverse Transcriptase Inhibitors. *Chem. Biol. Drug Des*., 87, 200-212. doi: org/10.1111/cbdd.12620

G.R. Kaufmann, D.A. Cooper, 2000. Antiretroviral therapy of HIV-1 infection: established treatment strategies and new therapeutic options. *Curr. Opin. Microbiol.*, 3, 508 -514. doi: org/10.1016/s1369-5274(00)00131-4

G.-X. Shuang, T. Xiao, Y.-Y. Zhu, G.-Y. Liu, F.-E. Chen, 2019. Recent progress in HIV-1 inhibitors targeting the entrance channel of HIV-1 non-nucleoside reverse transcriptase inhibitor binding pocket. *Eur. J. Med. Cem.*, 174, 277-291. doi: org/10.1016/j.ejmech.2019.04.054

H. Jonckheere, J. Ann, E.De Clercq, 2000. The HIV-1 Reverse Transcription (RT) Process as Target for RT Inhibitors. *Med. Res. Rev.*, 20 (2), 129-154. doi: org/10.1002/(sici)1098-1128(200003)20:2%3C129::aid-med2%3E3.0.co;2-a

I. Muegge, Y.C. Martin, 1999. A general and fast scoring function for protein-ligand interactions:a simplified potential approach. *J. Med. Chem*., 42 (5), 791-804. doi: org/10.1021/jm980536j

J.Yang, W. Chen, D. Kang, X. Lu, X. Li, Z. Liu, B. Huang, D. Daelemans, C. Pannecouque, E. D. Clercq, P. Zhan, X. Liu, 2016. Design, synthesis and anti-HIV evaluation of novel diarylpyridine derivatives targeting the entrance channel of NNRTI binding pocket. *Euro. J. Med. Chem*., 109, 294-304. doi: org/10.1016/j.ejmech.2015.11.039

L. A. Kohlstaedt, J. Wang, J. M. Friedman, P. A. Rice and T. A. Steitz, 1992. *Science*, 256, 1783-1790. doi: org/10.1126/science.1377403

L. Anta, J.M. Llibre, E. Poveda, J.L. Blanco, M. Alvarez, M.J. Perez-Elias, A. Aguilera, E. Caballero, V. Soriano, C. deMendoza, 2013. Rilpivirine resistance mutations in HIV patients failing non-nucleoside reverse transcriptase inhibitor-based therapies, *AIDS*. 27 (1), 81-85. doi: org/10.1097/QAD.0b013e3283584500

L. D. Schuler, X. Daura, W. F. Van Gunsteren, 2001. An Improved GROMOS96 Force Field for Aliphatic Hydrocarbons in the Condensed Phase. *J. of Com. Che.*, 22, 1205-1218. doi: org/10.1002/jcc.1078

L.R.Boone, 2006. Next-generation HIV-1 non-nucleoside reverse transcriptase inhibitors, *Curr. Opin. Investig. Drugs* (London, Engl. 2000). 7 (2), 128-135.

M.A.Wainberg, 2013. Combination therapies, effectiveness, and adherence in patients with HIV infection: clinical utility of a single tablet of emtricitabine, rilpivirine, and tenofovir . *HIV/AIDS*. 5, 41-49. doi.org/10.2147/HIV.S32377

M. Kumar, R.Vijayakrishnan, Gita S.Rao, 2010. In silico structure-based design of a novel class of potent and selective small peptide inhibitor of Mycobacterium tuberculosis Dihydrofolate reductase, a potential targetfor anti-TB drug discovery. *Mol. Divers*. 14, 595-604. doi: org/10.1007/s11030-009-9172-6

M. F. Sk, R. Roy & P. Kar, (2020). Exploring the potency of currently used drugs against HIV-1 protease of subtype D variant by using multiscale simulations. *J. of Bio. Str. and Dyn.*, 1–30. doi:10.1080/07391102.2020.1724196

M. Yadav, R. Srivastava, F. Naaz, A. Singh, R. Verma, Ramendra K.Singh, 2019. In silico studies on new oxathiadiazoles as potential anti-HIV agents. *Gen. Rep.* 14, 87-93. doi: org/10.1016/j.genrep.2018.12.004

P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, T. E. Cheatham, 2000. Calculating Structures and Free Energies of Complex Molecules: Combining

Molecular Mechanics and Continuum Models. *Acc. of Che. Res.*, 33, 889-897. doi: org/10.1021/ar000033j

P. S. S. Gupta, R. N. U. Islam, S. Banerjee, A. Nayek, M. K. Rana, A. K. Bandyopadhyay, 2020. Screening and Molecular Characterization of Lethal Mutations of Human Homogentisate 1, 2 Dioxigenase. J. *of Bio. Str. and Dyn.,* 1–16. doi:10.1080/07391102.2020.1736158

P. Zhan, C. Pannecouque, E. D. Clercq, X. Liu, 2016. Anti-HIV drug discovery and development: Current innovations. *J. Med. Chem.,* 59 (7), 2849-2878. doi: org/10.1021/acs.jmedchem.5b00497

R. P. Bhole, C. G. Bonde, S. G. Bonde, R. V. Chikhale, & R. D. Wavhale, 2020. Pharmacophore Model and Atom-Based 3D Quantitative Structure activity Relationship (QSAR) of Human Immunodeficiency Virus-1 (HIV-1) Capsid Assembly Inhibitors. *J. of Bio. Str. and Dyn.*, 1–14. doi:10.1080/07391102.2020.1715258

S. Genheden, U. Ryde, 2015. The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Exp. Opi. on Dru. Dis.,* 10 (5), 449-461. doi: org/10.1517%2F17460441.2015.1032936

S.G. Sarafianos, B. Marchand, K. Das, D. M. Himmel, M. A. Parniak, S. H. Hughes, E. Arnold, 2009. Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition. *J. Mol. Biol.* 385 (3), 693-713. doi: org/10.1016/j.jmb.2008.10.071

S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, E. Lindahl, 2013. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*, 29 (7), 845-854. doi: org/10.1093/bioinformatics/btt055

S. S. Swain, S. K. Paidesetty, B. Dehury, J. Sahoo, S. C. Vedithi, N. Mahapatra, R. N. Padhy, 2018. Molecular docking and simulation study for synthesis of alternative dapsone derivative as a newer antileprosy drug in multidrug therapy. *J. of Cel. Bio.*,119 (12), 9838-9852. *doi: org/10.1002/jcb.27304*

S.Vella, L.Palmisanom, 2000. Antiretroviral therapy: state of the HAART. *Antiviral Res.*, 45 (1), 1-7. doi: org/10.1016/s0166-3542(99)00068-6

S.-X. Gu, H.-H. Lu, G.-Y. Liu, X.-L. Ju, Y.-Y. Zhu, 2018. Advances in diarylpyrimidines and related analogues as HIV-1 nonnucleoside reverse transcriptase inhibitors. *Eur. J. Med. Cem.*, 158, 371-392. doi: org/10.1016/j.ejmech.2018.09.013

T.-C. Chou, 2006. Theoretical Basis, Experimental Design, and Computerized Simulation of Synergism and Antagonism in Drug Combination Studies. *Pha. Rev.*, 58 (3), 621-681. doi: org/10.1124/pr.58.3.10

U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, 1995. A smooth particle mesh Ewald method. *The J. of Che. Phy.*, 103, 8577-8593. doi: org/10.1063/1.470117

U. Hess, S. Blairy, R. E. Kleck, 1997. The Intensity of Emotional Facial Expressions and Decoding Accuracy. *J. of Nver. Beh.*, 21, 241-257. doi: org/10.1023/A:1024952730333

W. Humphrey, A. Dalke, K. Schulten, 1996. VMD: Visual molecular dynamics. *J. of Mol. Gra.*, 14, 33-38. doi: org/10.1016/0263-7855(96)00018-5

X. Chen, P. Zhan, D. Li, E. D. Clercq, X. Liu, 2011. Recent Advances in DAPYs and Related Analogues as HIV-1 NNRTIs. *Curr. Med. Chem.*, 18 (3), 359-376. doi: org/10.2174/092986711794839142

X. Chen, Q. Meng, L. Qiu, P. Zhan, H. Liu, E. D. Clercq, C. Pannecouque, X. Liu, 2015. Design, synthesis, and anti-HIV evaluation of novel triazine derivatives targeting the entrance channel of the NNRTI binding pocket, *Chem. Biol. Drug Des*., 86, 122-128. doi: org/10.1111/cbdd.12471

X. Zuo, Z. Huo, D. Kang, G. Wu, Z. Zhou, X. Liu, P. Zhun, 2018. Current insights into anti-HIV drug discovery and development: A review of recent patent literature (2014-2017). *Exp. Opi. on Thera. Pat.* 28 (4), 299-316. doi.org/10.1080/13543776.2018.1438410Y. Cong, L. Duan, K. Huang, J. Bao & J. Z. H. Zhang, 2020. Alanine Scanning Combined with Interaction Entropy Studying the Differences of Binding Mechanism on HIV-1 and HIV-2 Proteases with Inhibitor. *J. of Bio. Str. and Dyn.*, 1–30. doi:10.1080/07391102.2020.1734488

Y. Khan, M. Garg, Q. Gui, M. Schadt, A. Gaikwad, D. Han,A. C. Arias, 2016. Flexible Hybrid Electronics: Direct Interfacing of Soft and Hard Electronics for Wearable Health Monitoring. *Adv. Fun. Mat.*, 26, 8764-8775. doi: 10.1002/adfm.201603763

Y. Wang, X. Wang, Y. Xiong, A. C. Kaushik, J. Muhammad, A. Khan, H. Dai, D. Q. Wei, 2019. New strategy for identifying potential natural HIV-1 non-nucleoside reverse transcriptase inhibitors against drug-resistance: an *in silico* study, *J. of Bio. Str. and Dyn. doi: 10.1080/07391102.2019.1656673.*

Z.-Y. Wu, N. Liu, B. Qin, L. Huang, F. Yu, K. Qian, S. L. Morris- Natschke, S.Jiang, C. H. Chen, K.-H. Lee, L. Xie, 2014. Optimization of the Antiviral Potency and Lipophilicity of Halogenated 2,6-Diarylpyridinamines as a Novel Class of HIV-1 NNRTIS. *Chem. Med. Chem.* 9 (7), 1546-1555. doi: org/10.1002/cmdc.201400075

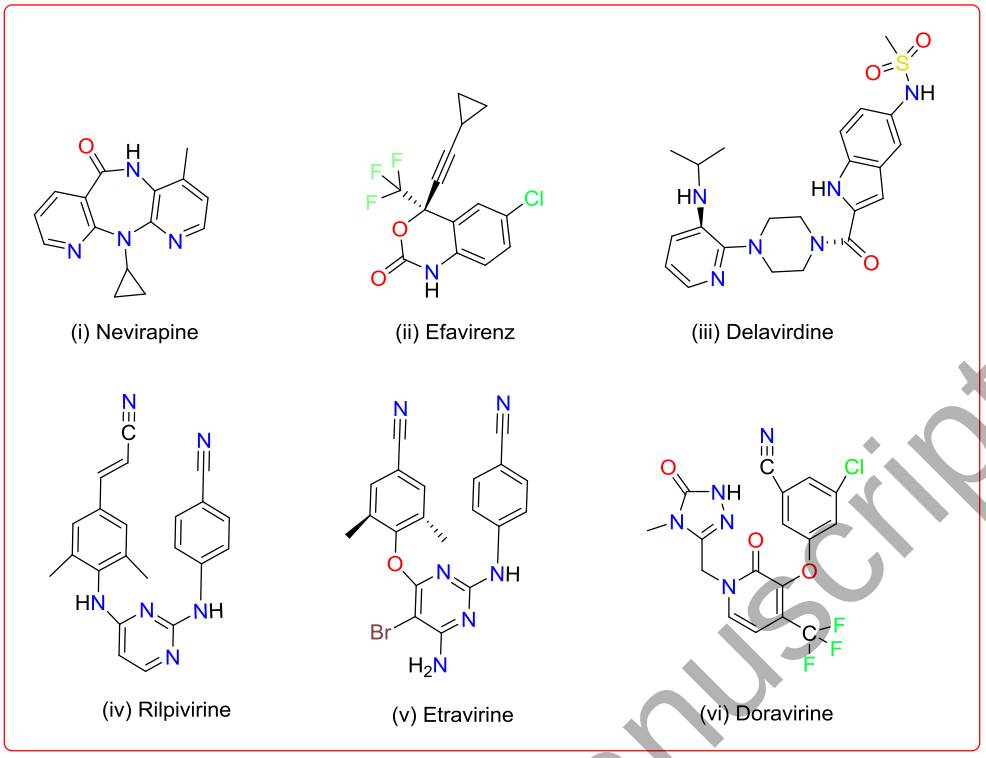


Fig. 1. FDA approved NNRTI drugs

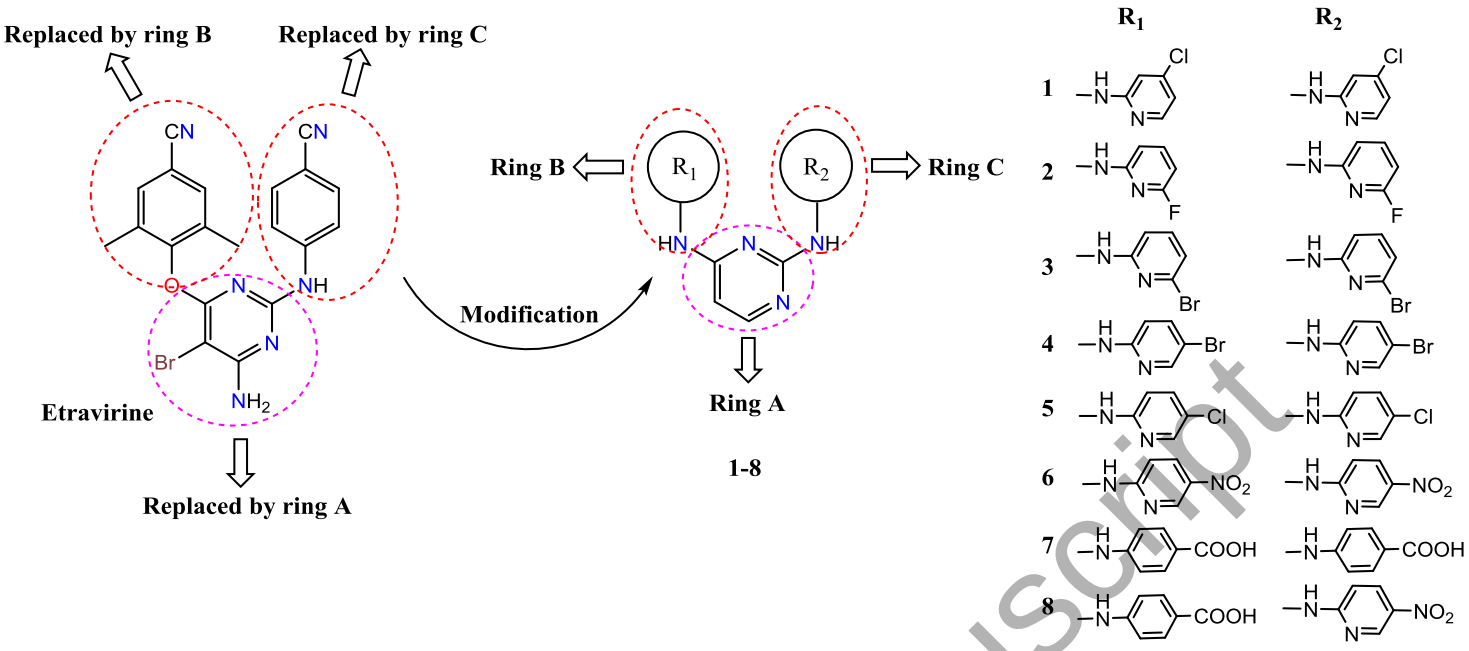
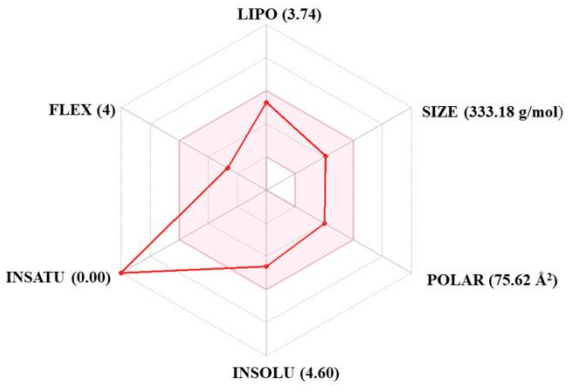
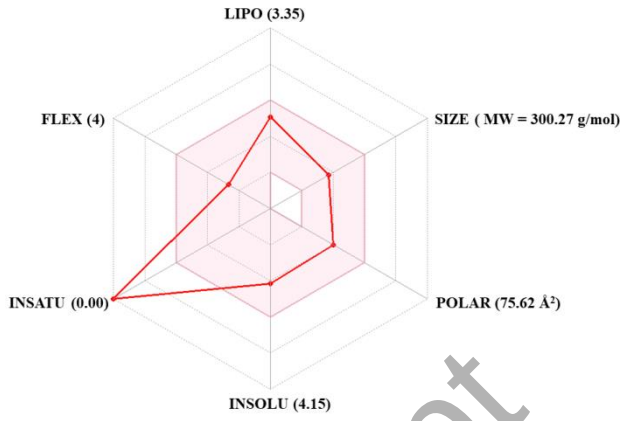


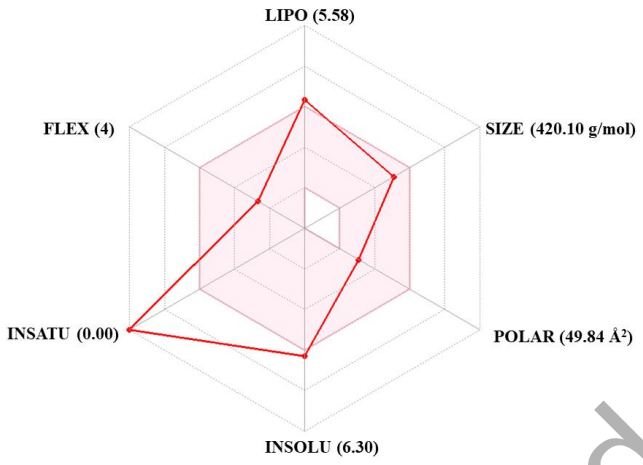
Fig. 2. Design consideration of substituted diarylpyrimidine derivatives, **1-8,** as NNRTIs



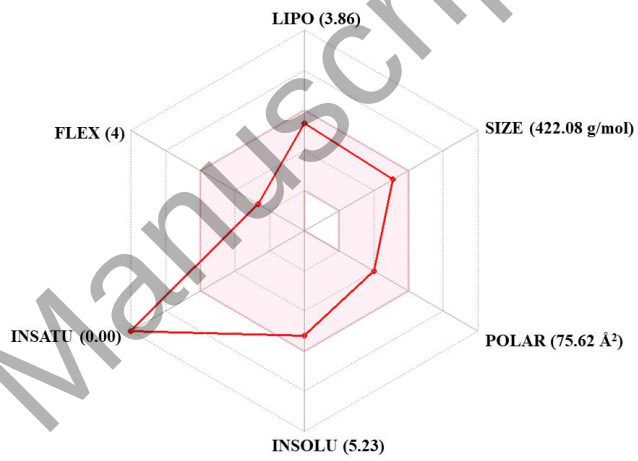
1



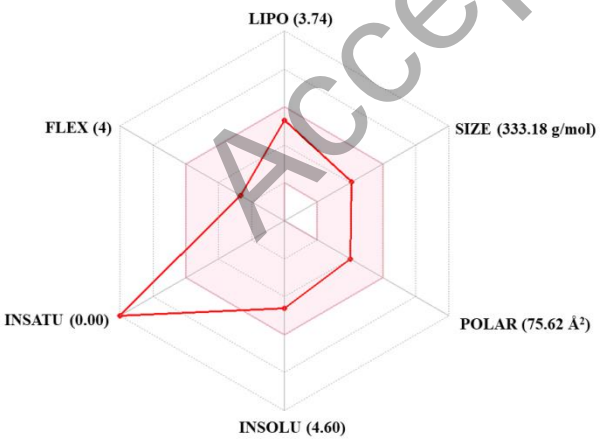
2



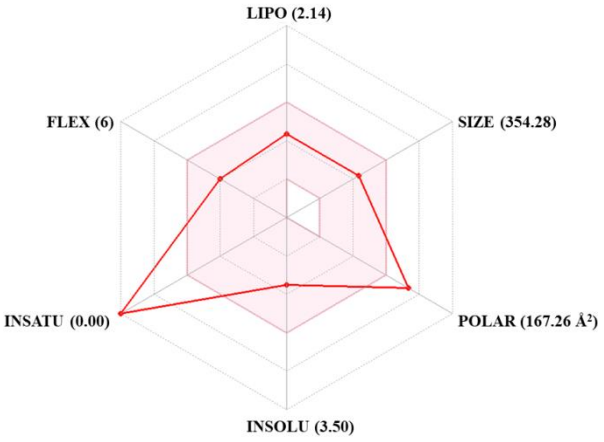
3



4



5



6

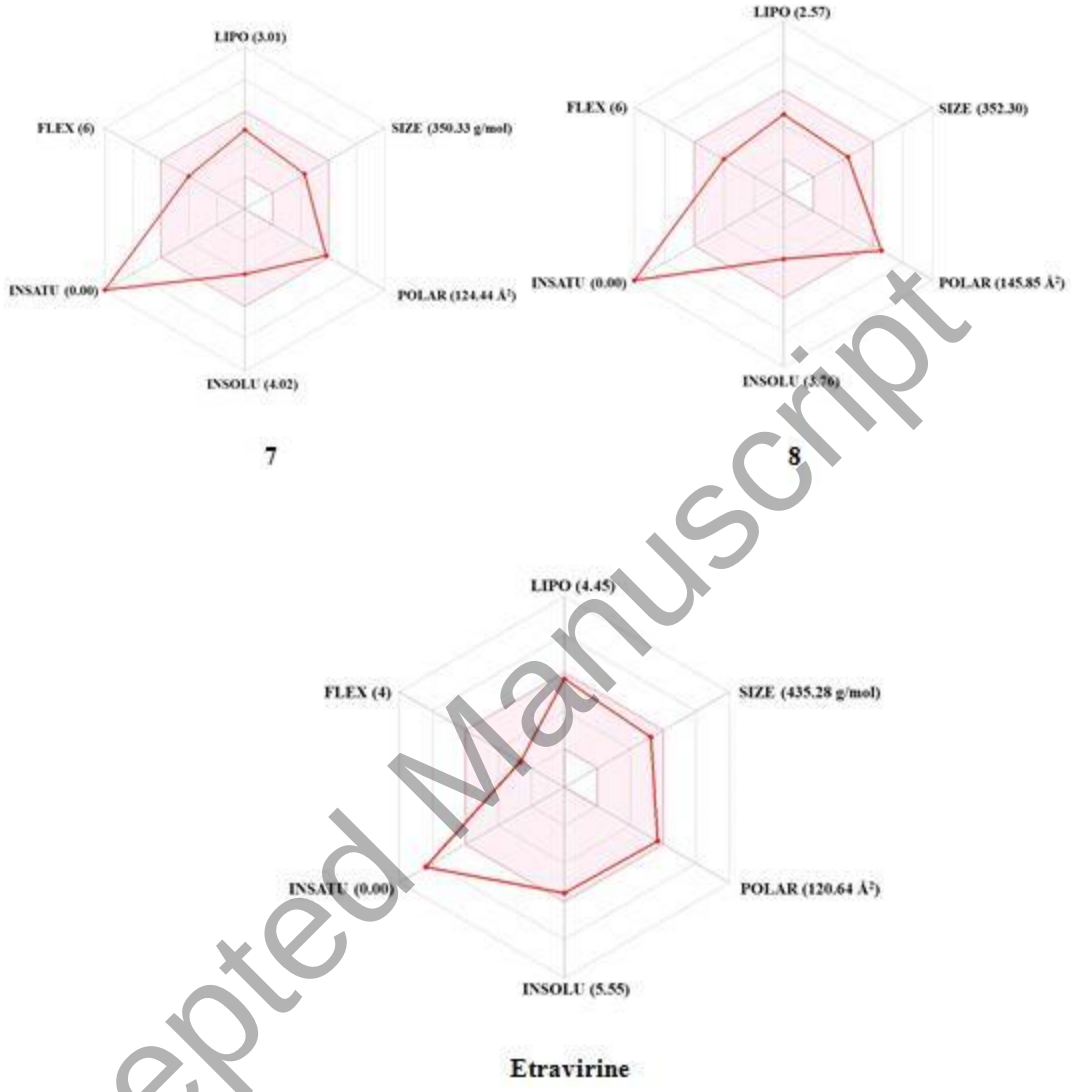
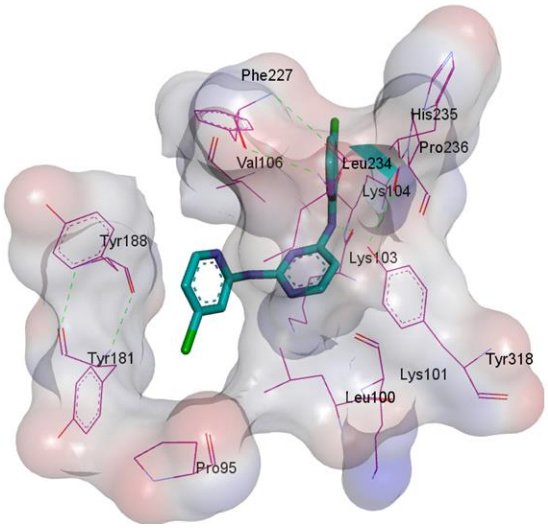


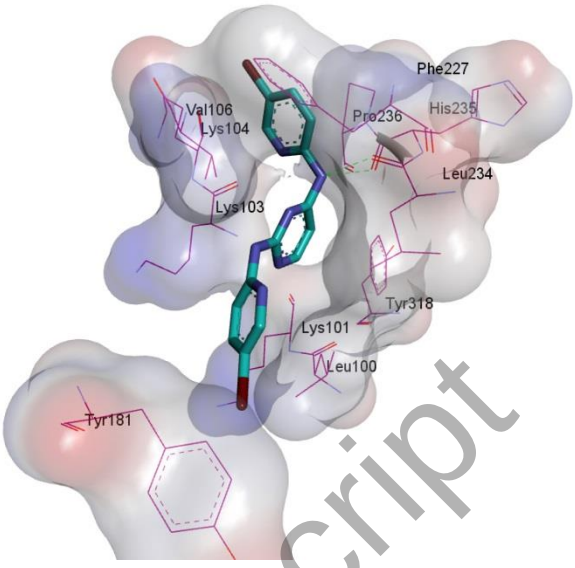
Fig. 3. Radar graph of diarylpyrimidine derivatives; **1**, **2,** **3**, **4**, **5**, **6**, **7**, **8** and reference drug showing drug likeness characteristics (pink area focuses the permissible values of drug likeness characteristics of molecule), (LIPO – Lipophilicity (≤**5**), SIZE - Molecular weight (≤500) , FLEX - Rotatable bonds (≤9), *INSOLU – Insolubility (≤**6**) , **INSATU – Insaturation (≤**1**))

*INSOLU: Insolubility of a compound is one of the physicochemical descriptors used to define the physicochemical properties of a drug to evaluate their drug likeness.

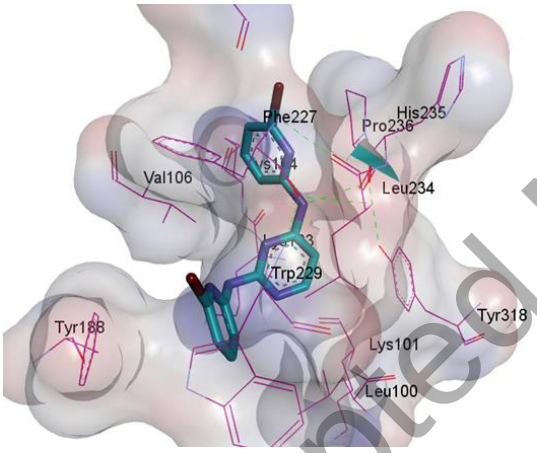
**INSATU: Insaturation in the designed molecule can be defined as "the fraction of carbons in the sp3 hybridization is increased from 0.25 to a higher value". This parameter is also considered as a physiochemical descriptor in drug likeness evaluation.



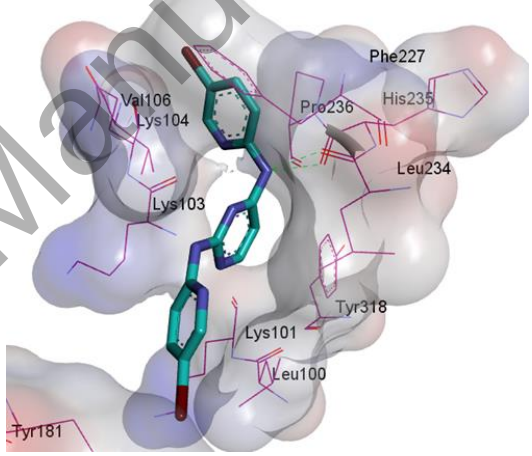
Compound 1



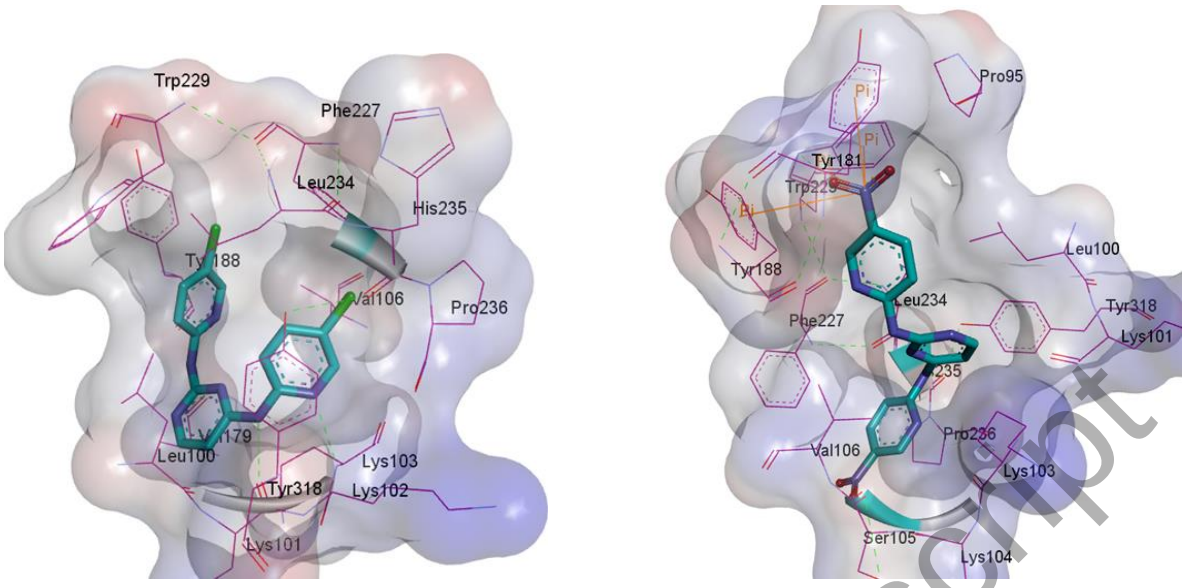
Compound 2



Compound 3



Compound 4



Compound 5 **Compound 6**

Fig. 4. Docking interactions of diarylpyrimidine derivatives; **1**-**6** within NNIBP of RT allosteric site of wild type HIV-1.

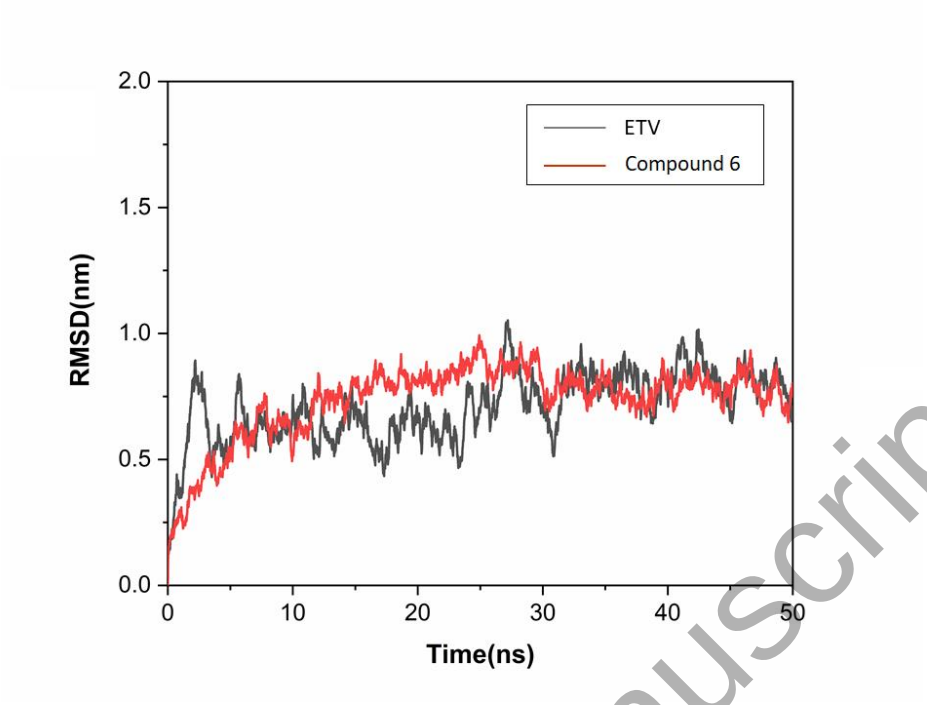


Fig. 5. RMSD plots of the HIV-RT complexes of Etravirine (ETV) and compound **6**. (shown in black and red, respectively)

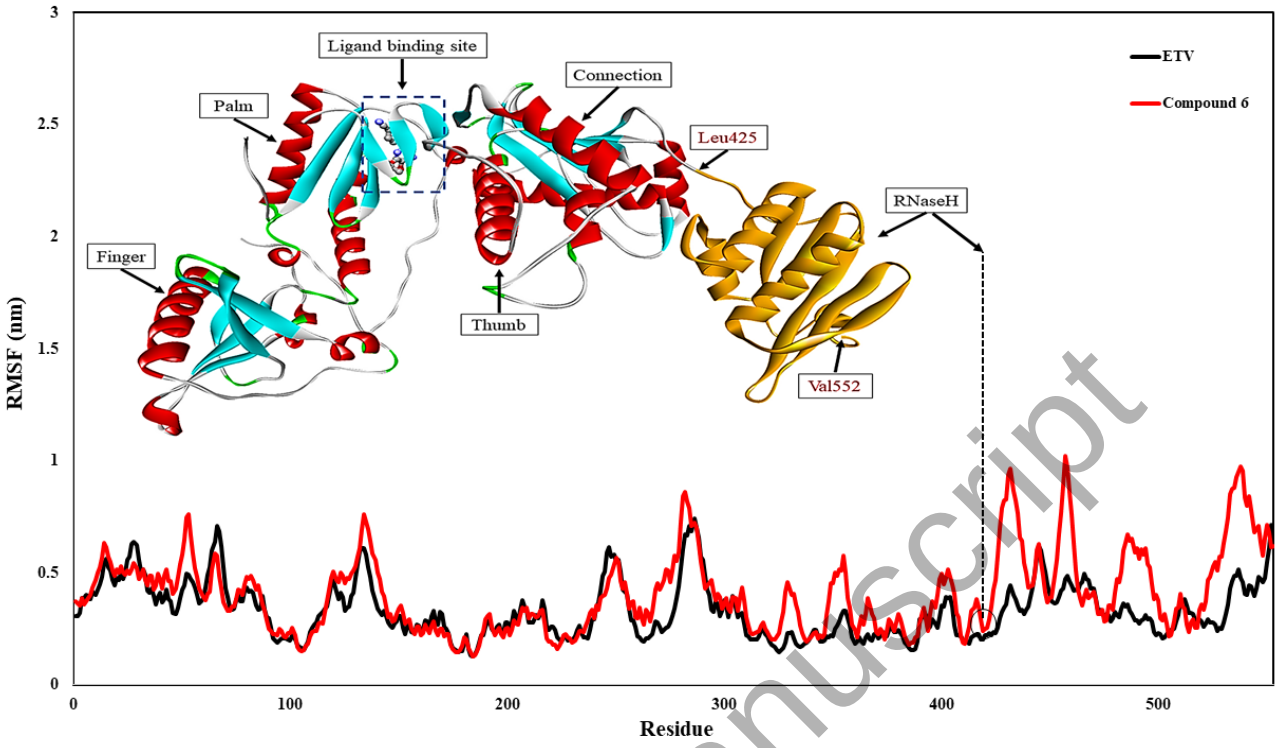


Fig. 6. RMSF plots of the protein backbone of HIV-RT:Etravirine (ETV) (black) and HIV-RT:Compound **6** (red) complexes. Ribbon representation of HIV RT (PDB id: 3MEC) with Rnase H domain shown in golden color

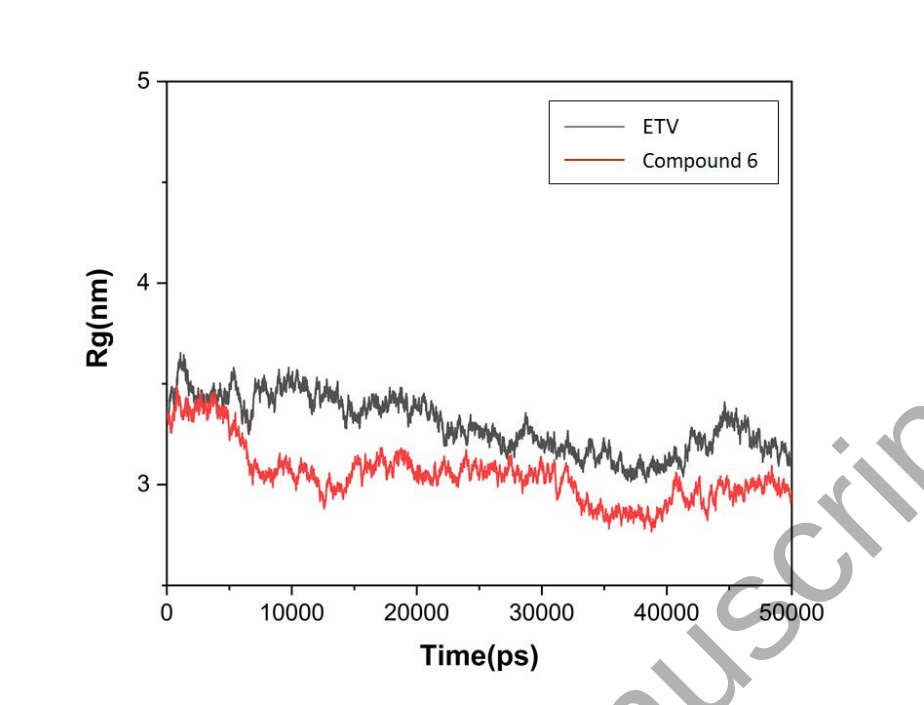


Fig. 7. R_g (radius of gyration) plots of HIV-RT:Etravirine (ETV) (black) and HIV-RT:Compound **6** (red) complexes

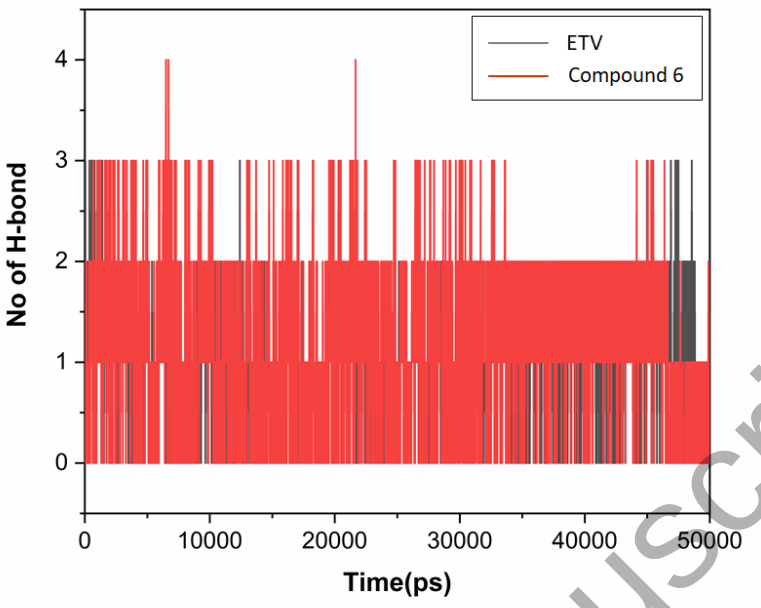


Fig. 8. Graphical representation of no. of H-bond formed in HIV-RT:Etravirine (ETV) (black) and HIV-RT:Compound **6** (red) complexes

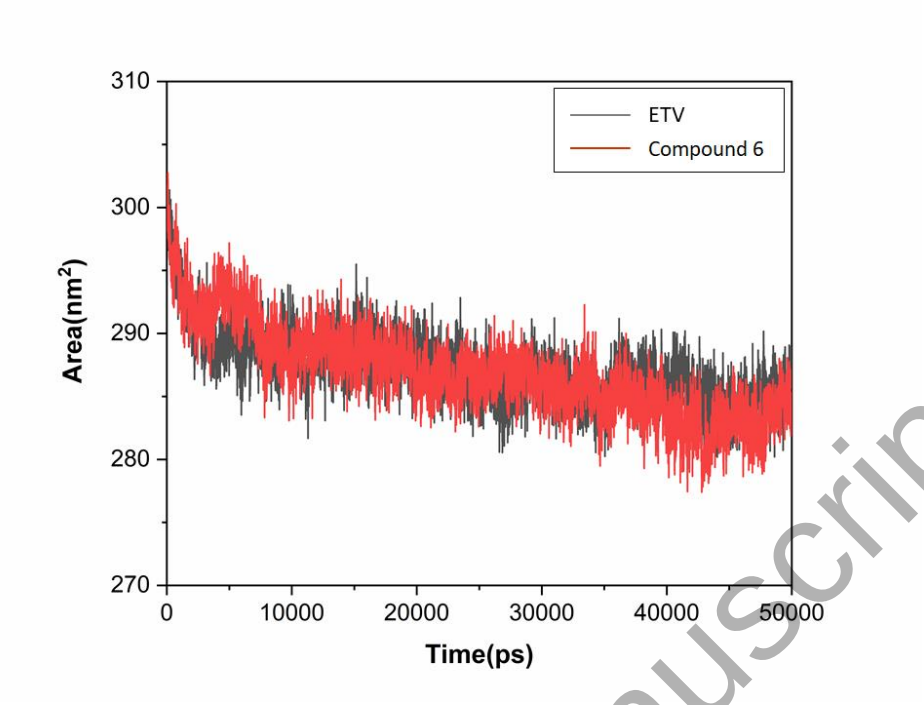


Fig. 9. Graphical representation of solvent accessible surface area of HIV-RT:Etravirine (ETV) (black) and HIV-RT:Compound **6** (red) complexes

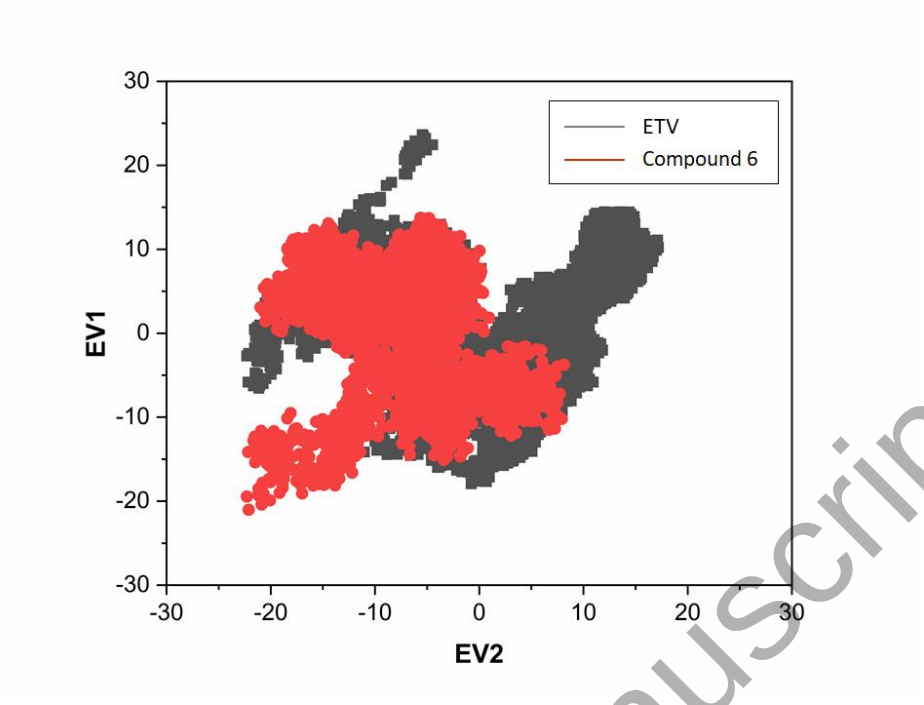


Fig. 10. 2-D scatter plots of protein-ligand complexes, HIV-RT:Etravirine (ETV) (black) and HIV-RT:Compound **6** (red), projecting the motion of the protein in phase space for the first two principle components. (EV1 and EV2 are eigenvector 1 and 2, respectively)

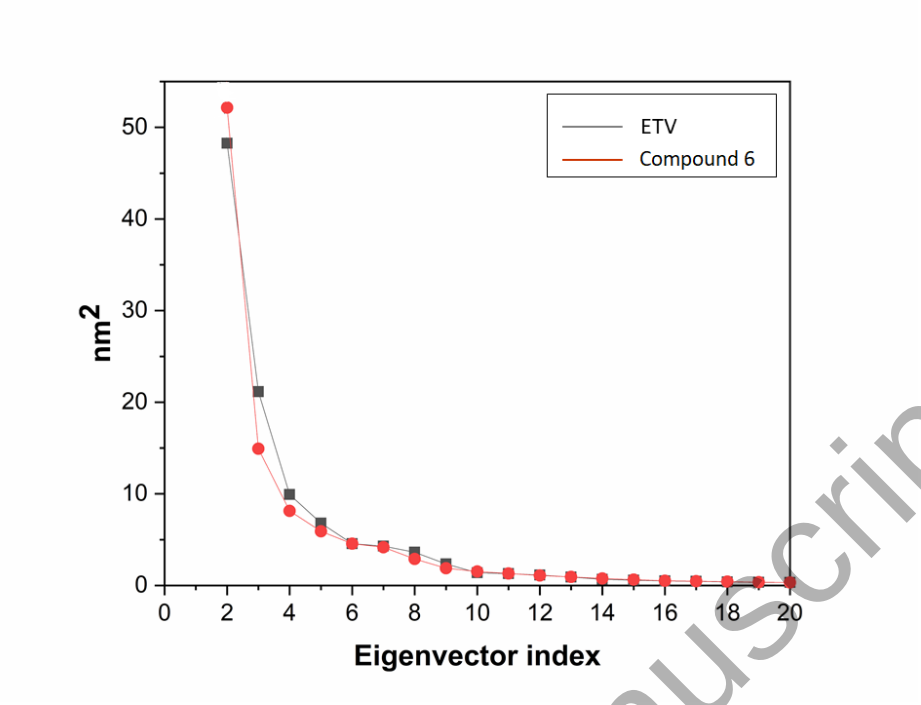


Fig. 11. Plot of eigenvalues of the fluctuations of protein backbone Vs the respective eigenvector indices of HIV-RT:Etravirine (ETV) (black) and HIV-RT:Compound **6** (red) complexes

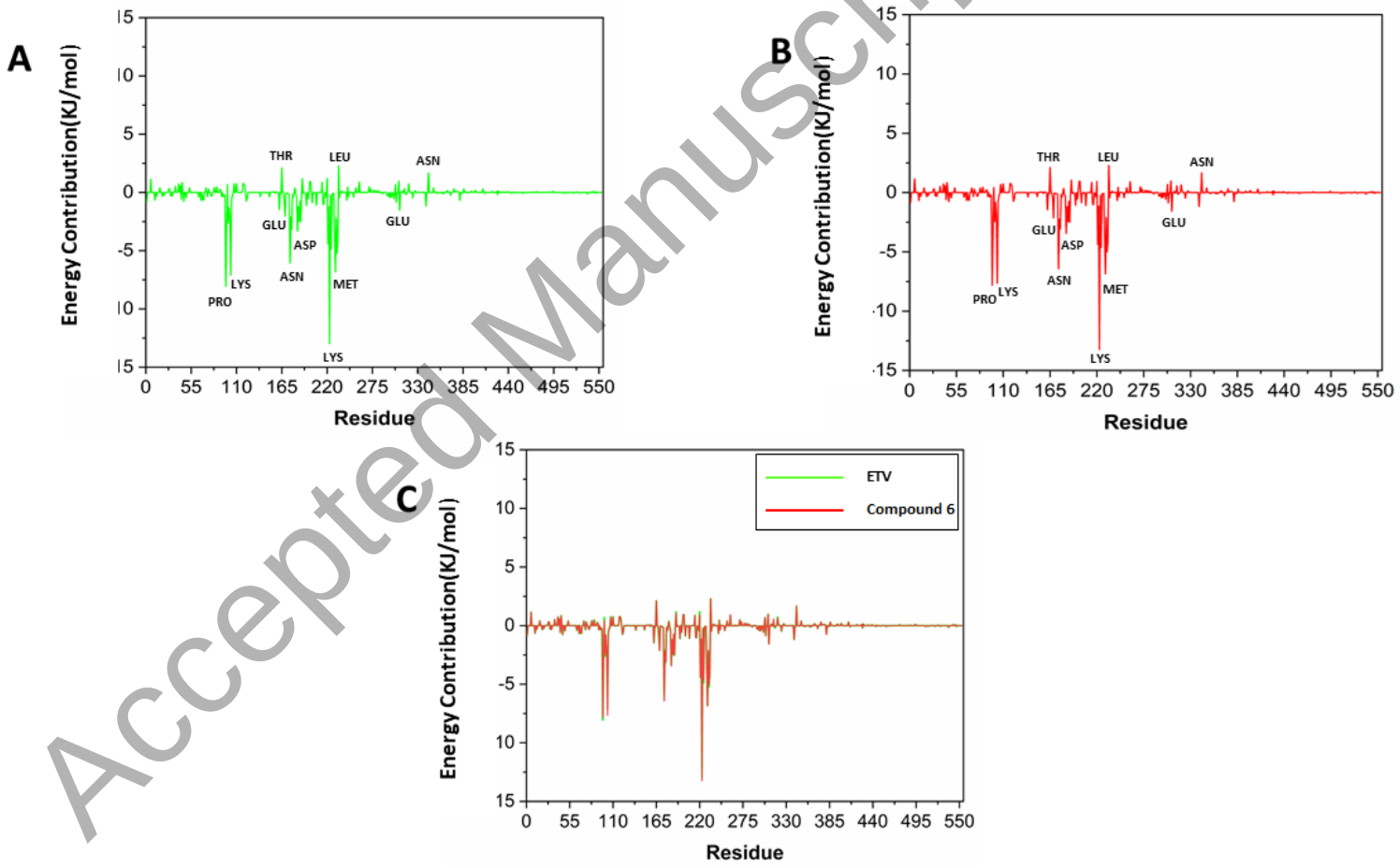


Fig. 12. Energy contributions of individual amino acid residues to the binding free energy ($\Delta G_{bind)}$ shown for **(A)** (HIV-RT:Compound **6**, **(B)** HIV-RT:Etravirine (ETV) and **(C)** both complexes combinedly. Residues only actively contributing either the positive or negative binding energies to ΔG_{bind} are highlighted

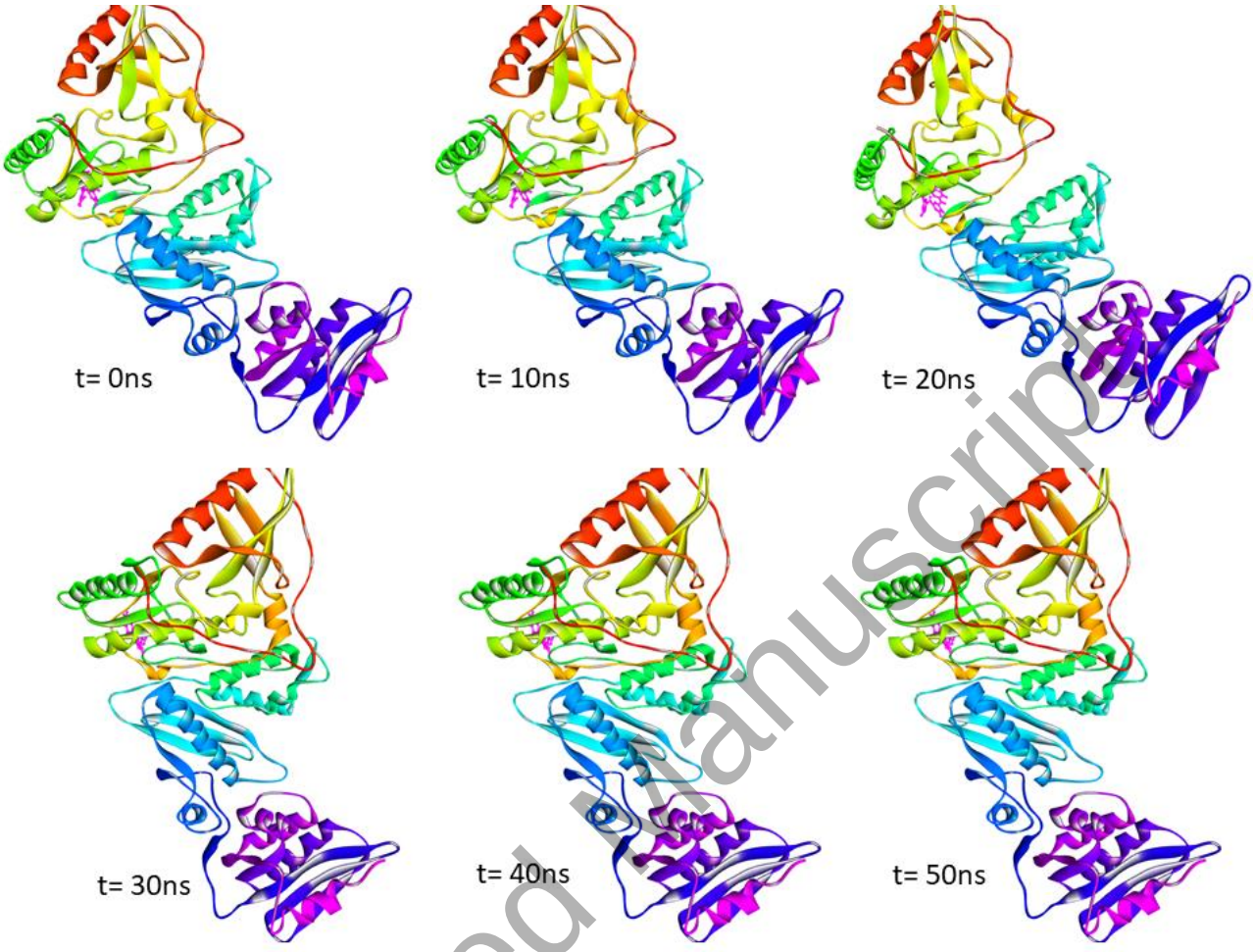


Fig. 13. Snapshots of the compound 6-HIV-RT complex from 0 to 50ns molecular dynamics simulation

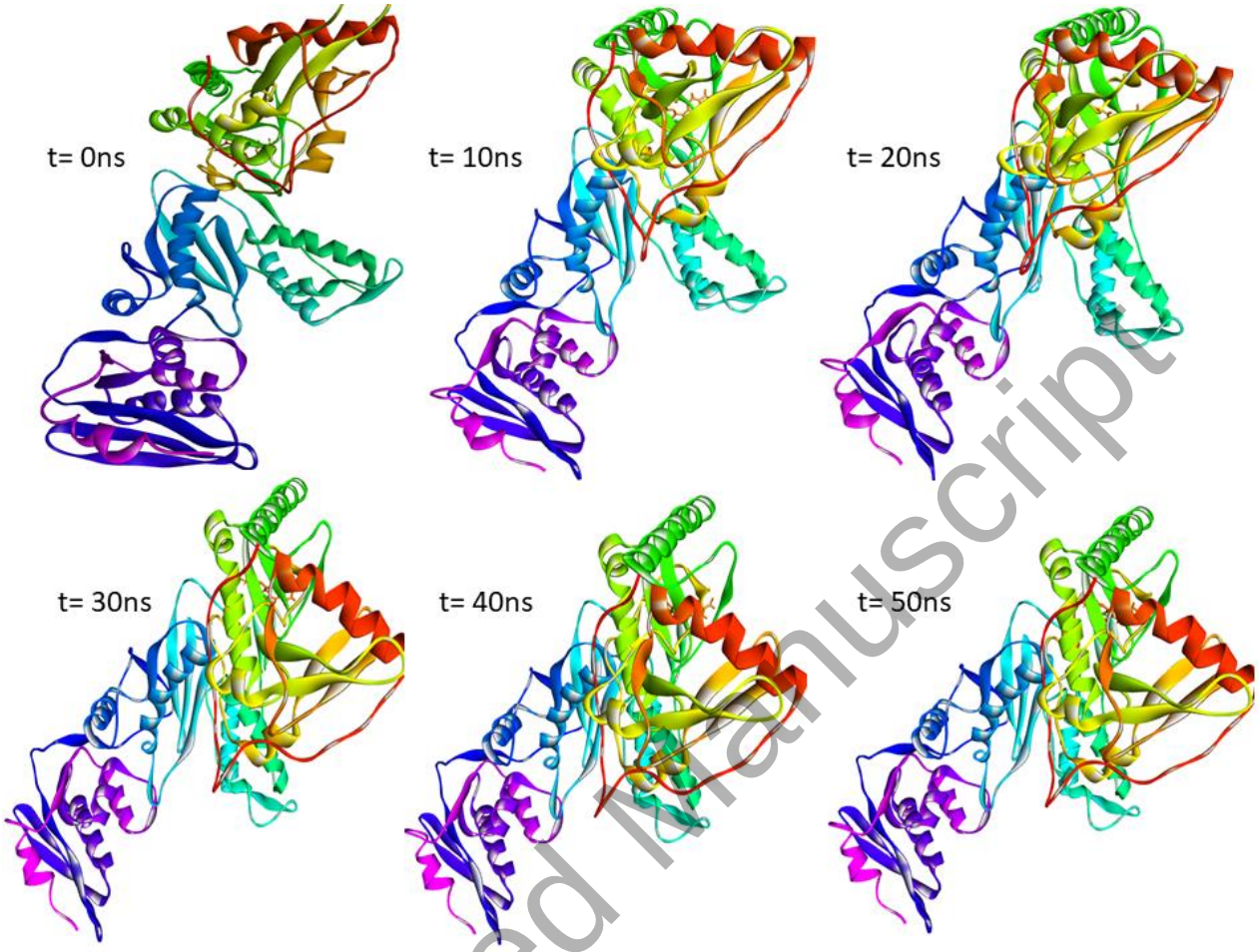


Fig. 14. Snapshots of the Etravirine-HIV-RT complex from 0 to 50 ns molecular dynamics simulation

Table 1. Physicochemical data of diarylpyrimidine derivatives **1-8**

Compound	MW <500	Log P ≤5	TPSA ≤140	HBA ≤10	HBD ≤5	Violation ≤1
1	333.18	3.90	75.62	6	2	0
2	300.27	3.65	75.62	6	2	0
3	422.08	4.94	75.62	6	2	0
4	422.08	4.60	75.62	6	2	0
5	333.18	4.34	75.62	6	2	0
6	354.29	2.90	167.27	12	2	1
7	350.33	4.26	124.44	8	4	0
8	352.31	3.58	145.85	10	3	0
Etravirine	435.28	5.02	120.65	7	3	1

TPSA = Total polar surface area, HBA = no. of H-bond acceptors, HBD = no. of H-bond donor, Violation = violence of Lipinski's rule of five

Table 2. Drug-likeness properties of diarylpyrimidine derivatives **1**-**8**

Compound	GPCR ligand	Ion channel modulator	Kinase Inhibitor	Nuclear receptor ligand	Protease inhibitor	Enzyme inhibitor
1	0.24	0.12	0.84	-0.32	-0.14	0.24
2	0.63	0.46	0.46	0.00	0.21	0.20
3	0.51	0.47	0.48	-0.14	-0.01	0.18
4	0.17	0.08	0.46	-0.48	-0.15	0.24
5	0.28	0.18	0.48	-0.44	-0.04	0.23
6	0.12	0.10	0.44	-0.35	-0.11	0.21
7	0.14	0.04	0.57	-0.16	-0.04	0.27
8	0.10	0.05	0.66	-0.18	-0.05	0.25
Etravirine	-0.01	-0.08	0.49	-0.31	0.00	0.22

GPCR = G-protein coupled receptors

Table 3. Docking interaction of diarylpyrimidine derivatives, **1-8,** with HIV-RT in ligand - receptor docked complex

Compound	No. of H-B	Amino acid in H-B	H-B Type	D(Å)	D-A	A-A	No. of π- π / π-cation B	Amino acid in π- π / π-cation B	'π-π/ π+' monitor			
									Bond	D(Å)	End1	End2
1	1	Lys103	**1**: N9 - Lys103: O	2.8	N9	O	--	--	--	--	--	--
2	1	Lys103	**2**: N9 - Lys103: O	2.8	N9	O	--	--	--	--	--	--
3	1	His235	**3**: N8 - His235: O	2.9	N8	O	--	--	--	--	--	--
4	1	His235	**4**: N8 - His235: O	2.9	N8	O	1(π-π)	Tyr188	Tyr188 - **4**	6.0	Tyr188	**4**
5	1	Lys101	**5**: N8 - Lys101: O **5**: N9 - Lys103: O	2.6 2.8	N8 N9	O O	--	--	--	--	--	--
6	2	His235 Trp229	**6**: N8 - His235: O Trp229: NE1- **6**: O23	2.8 3.1	N8 NE1	O O23	4 (π+-B)	Trp229 Tyr188 Tyr181	Trp229 - **6** Trp229 - **6** Tyr188 - **6** Tyr181 - **6**	5.8 5.9 5.7 6.0	Trp229 Trp229 Tyr188 Tyr181	**6** **6** **6** **6**
7	--	--	--	--	--	--	--	--	--	--	--	--
8	--	--	--	--	--	--	--	--	--	--	--	--
Ref.	1	Lys101	Ref: N19 - Lys101:O	3.0	N19	O	2 (π-π)	Tyr318 Tyr181	Tyr318 - **Ref.** Tyr181 - **Ref.**	5.0 6.0	Ref. Ref.	Tyr318 Tyr181

H-B = Hydrogen bond, **D** = Distance (Å), **D-A** = Donor Atom, **A-A** = Acceptor Atom, **π - π B** = Pi-Pi bond, **π - cation B** = Pi - cation bond, Ref. = Etravirine

Table 4. Docking scores of diarylpyrimidine derivatives, **1-8,** with HIV-RT in ligand - receptor docked complex

Compound	PLP1	PLP2	DS	LIG	Ludi2	Ludi3	ΔG	EC_{50} (Predicted)
1	99.89	92.33	59.45	-4.12	496	729	-7.03	5.1×10^{-8}
2	98.02	85.48	54.93	-4.187	450	728	-6.38	5.2×10^{-8}
3	102.41	93.92	59.42	-5.03	527	741	-7.47	3.9×10^{-8}
4	97.08	89.40	57.34	-4.94	491	710	-6.96	7.9×10^{-8}
5	97.24	88.51	57.42	-5.05	491	702	-6.96	9.5×10^{-8}
6	109.74	95.68	58.29	-3.89	383	608	-5.43	8.3×10^{-7}
7	96.08	83.07	53.89	-3.55	369	598	-5.24	1.0×10^{-6}
8	101.48	93.20	57.36	-4.56	471	579	-6.68	1.6×10^{-6}
Etravirine	111.55	103.86	55.05	1.97	570	732	-7.19	4.8×10^{-8}

PLP = Piecewise Linear Potential, DS = Dock Score, LIG = Ligand Internal Energy, Ludi2 and Ludi3 = Empirical scoring function, ΔG = Binding Energy (Kcal/mol), EC_{50} predicted = 50% effective concentration of compound

Table 5. TOPKAT analysis of diarylpyrimidine derivatives **1-8**

Ligand	Molecular formula	DTP	LD_{50} Log		LC_{50} Log		ALogP	EC_{50} Log		Rotatable bonds (<10)	Carcinogenicity/Mutagenicity (Discriminant score)
			Predicted Value	Confidence limit	Predicted Value	Confidence limit		Predicted Value	Confidence limit		
1	$C_{14}H_{10}Cl_2N_6$	0.021	172.4 mg/kg	981mg/kg	2.05	10	3.04	8.216	13.0	4	-15.50
2	$C_{14}H_{10}F_2N_6$	0.000	1.3g/kg	7.1g/kg	2.15	5.4	3.52	8.216	11.7	4	-15.55
3	$C_{14}H_{10}Br_2N_6$	0.000	179.3 mg/kg	765.2 mg/kg	2.19	7.8	3.31	8.216	13.6	4	-15.58
4	$C_{14}H_{10}Br_2N_6$	0.028	122.1 mg/kg	746.3 mg/kg	2.13	5.4	3.21	8.218	16.5	4	-10.42
5	$C_{14}H_{10}Cl_2N_6$	0.028	153.3 mg/kg	873.8 mg/kg	2.06	10	3.68	8.216	13.0	4	-15.56
6	$C_{14}H_{10}N_8O_4$	0.013	342.4 mg/kg	2.0g/kg	2.04	7.1	1.50	7.189	17.7	6	-15.77
7	$C_{18}H_{14}N_4O_4$	0.558	286.5 mg/kg	1.8g/kg	2.53	6.5	3.33	8.513	65.1	4	-14.42
8	$C_{16}H_{12}N_6O_4$	0.682	147.4 mg/kg	893.3 mg/kg	2.65	10	4.35	8.565	16.5	6	-15.56
Etravirine	$C_{20}H_{15}BrN_6O$	1.000	182.2 mg/kg	795.8 mg/kg	2.22	8.2	5.50	9.590	15.1	4	-16.22

DTP = Developmental Toxicity Potential (< 0.70 score shows less toxicity)

LD_{50} = 50% Lethal Dose of a chemical that kills 50% of a sample population

LC_{50} = 50% Lethal Concentration (1/mol/h)

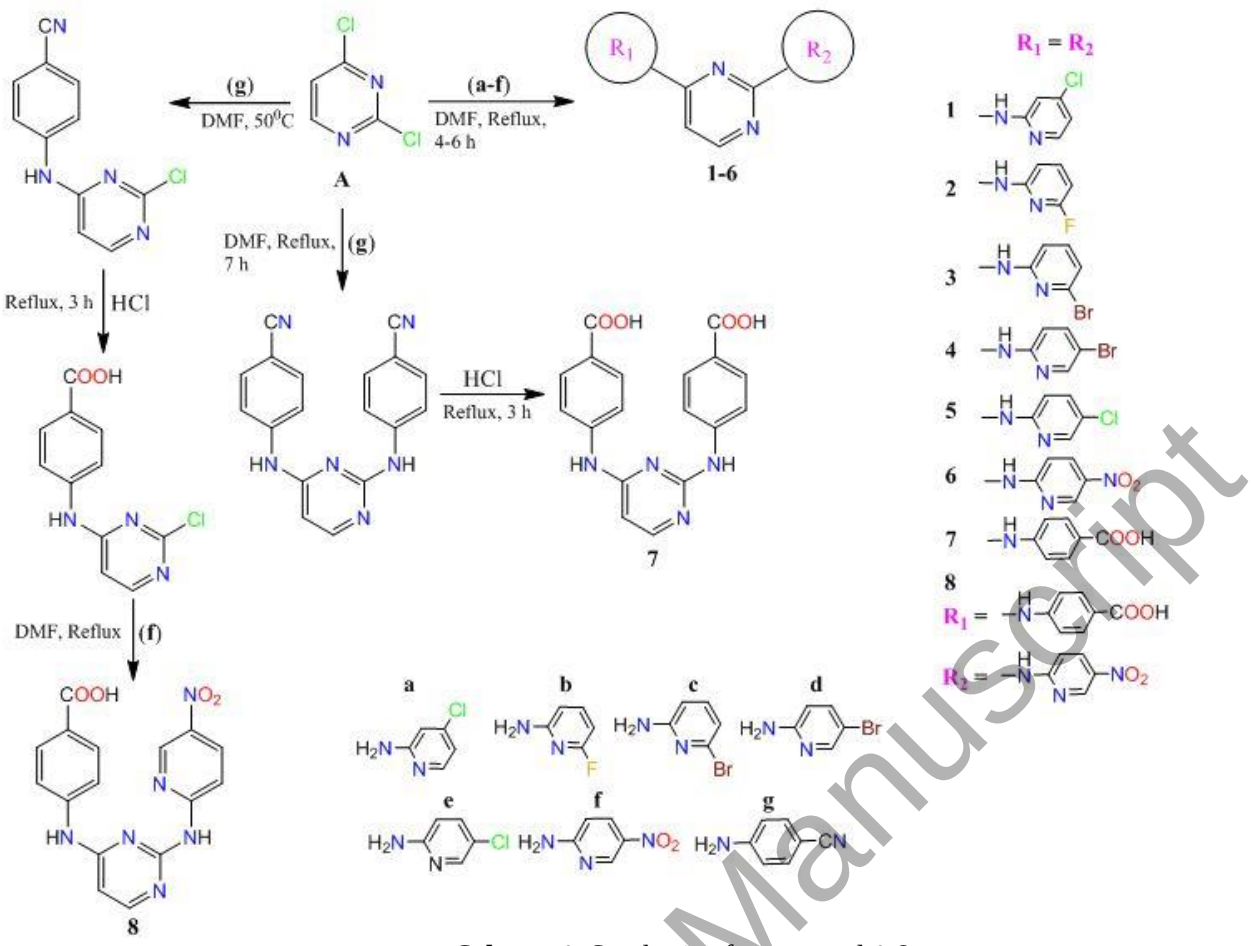
EC_{50} = 50% effective Concentration (1/mol)

ALogP = Lipophilicity (<5 value shows good lipophilicity)

Carcinogenicity = -ve discriminant score shows no carcinogenicity

Table 6. Molecular interaction energy data and the total binding free energies of selected protein-ligand complexes after molecular dynamics simulations

Energy components (kJ/mol)	**Compound 6: HIV-RT complex**	**Etravirine (ETV): HIV-RT complex**
van der Waal energy (ΔE_{vdW})	-242.063	-243.628
Electrostatic energy (ΔE_{elec})	-67.274	-64.596
Polar solvation energy	213.993	214.262
SASA energy	-17.465	-17.477
Binding energy (ΔG_{bind})	-112.809	-111.439



Scheme 1. Synthesis of compound **1-8**.