International Journal of PharmTech Research
CODEN (USA): IJPRIF, ISSN: 0974-4304, ISSN(Online): 2455-9563
Vol.14, No.01, pp 39-50, 2021

Discovery of HIV-I protease inhibitors: Pharmacophore mapping, Virtual Screening, Docking, and *In Silico* Pharmacokinetic and Toxicities Prediction

Hardik G. Bhatt[1], Paresh K. Patel*[1, 2], Dilip G. Maheshwari[2]

[1]Institute of Pharmacy, Nirma University, S.G. Highway, Ahmedabad 382 481. India.
[2]Department of Medicinal Chemistry & Quality Assurance, L.J. Institute of Pharmacy, S.G. Road 382 210. India.

Abstract : HIV-I protease is one of three critical enzymes for survival of virus which is responsible for Acquired Immunodeficiency Syndrome. The currently available protease inhibitors have demerits viz. drug resistance, severe side effects and poor pharmacokinetic profile. In course of our current research to discover novel HIV-I protease inhibitors, combination of ligand and structure based drug design approaches were used. A ligand based pharmacophore hypotheses were generated from clinically used FDA approved using DISCOtech and was refined using GASP. A structure based pharmacophore hypothesis was generated using phase. The features obtained from four different pharmacophore were hydrogen bond donor, hydrogen bond acceptor and hydrophobic region with distance geometry. These features were used to obtain substructures from databases like ZINC and NCI. A total 148 substructures having Q_{fit} value more than 99 were studied to design 25 potent triazine derivatives which were docked to find 10 hits. *In silico* pharmacokinetic and toxicities properties were calculated usingSWISSADME and pkCSM. The outcome obtained was satisfactory and all compounds had better pharmacokinetic profiles and they were free from toxicities. The present study will help to discover novel and bioactive leads for protease inhibitory activity.
Key words : HIV protease; Pharmacophore hypothesis; Molecular docking; Molecular dynamics; Virtual screening.

Introduction

Human Immunodeficiency virus (HIV) accounts for acquired immunodeficiency syndrome (AIDS) which spreads and transmits through blood and sex. According to WHO, around 41,758,381 people have already suffered from AIDS since 1986 with African and Asian countries have maximum number of HIV patients[1].A total number of 566,113 people had died due to AIDS and related illness. In 2019, more than 24

DOI= http://dx.doi.org/10.20902/IJPTR.2021.140104

million people were accessing antiretroviral therapy. It is estimated that US$ 26.2 billion will be required for the AIDS response in 2020[2]. It is estimated that 1 in every 3 patients living with AIDS are died because of tuberculosis which is most commonly opportunistic infection found in AIDS[3]. HIV is a retrovirus containing reverse transcriptase enzyme which is responsible for conversion of single stranded RNA to double stranded DNA. The life cycle of virus also depends on integrase enzyme, responsible for integration of virus into human genome. HIV protease enzyme forms functional and small polypeptides from large and non-functional polypeptide. The aspartic protease of human immunodeficiency virus (HIV PR) is in control for the cleavage of the viral Gag and Gag-Pol polyproteins precursors into advanced, purposeful viral enzymes and structural proteins. This process, called viral maturation, which clues to the final morphological rearrangements, is necessary for production of infectious viral particles[4]. If HIV-PR is repressed, the budding virions cannot affect new cells and the scattering of HIV is therefore terminated. The combination therapy referred as Highly Active Antiretroviral Therapy (HAART), comprise of drugs which inhibit these 3 essential enzymes. Though there is availability of 13 FDA approved protease inhibitors, but demerits are also associated with them which include drug resistance, poor bioavailability and severe adverse drug reactions. Due to these limitations there is a need to discover safe and potent protease inhibitors[5].

Due to synthesis and evaluation of large number of compounds, drug discovery was costly and time consuming which have been overcome by Computer Assisted Drug Design (CADD)[6]. Various approaches including ligand based and structure based have proven successfully in identification and optimization of lead molecules in last two decades[7]. Ligand based drug design is a methodology used in the absenteeism of the receptor 3D structure and it depend on knowledge of ligand molecules which may bind to target. Ligand based pharmacophore modeling and 3D quantitative structure activity relationships (3D QSAR) are the most important and widely used techniques in ligand based drug design and can be generated using chemical compounds which have proved their pharmacological activity[8]. In case of structure based drug designing approach, there must be availability of 3D structure of protein or receptor on which molecular docking and *in silico* screening can be performed. However, homology modeling can also be used to build structure of protein so that molecular docking along with pharmacophore can be performed. Virtual screening technique is an integral part of both ligand and structure based drug design methodologies used to filter drug like compounds by applying filters like Lipinski's rule of five[9]. The impression of molecular dynamics (MD) simulations in chemoinformatics and drug design has prolonged intensely in last few decade. These simulations are used to capture the behavior of target protein along with ligand at very fine resolution. The molecular dynamic simulation ensures and provides confidence for structure based drug design[10]. *In silico* pharmacokinetic and toxicity prediction plays a significant part in drug discovery and development because ADMET (absorption, distribution, metabolism and excretion and toxicities) are major reasons for failure of drug candidates during various phases of clinical trials. *In silico* properties prediction of proposed structure molecules before synthesis can reduce time and cost involved in drug development[11].

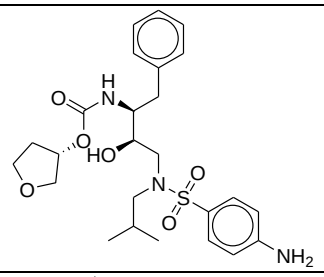
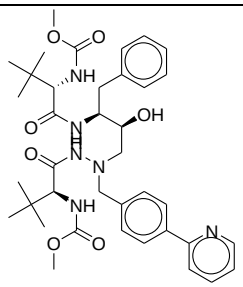
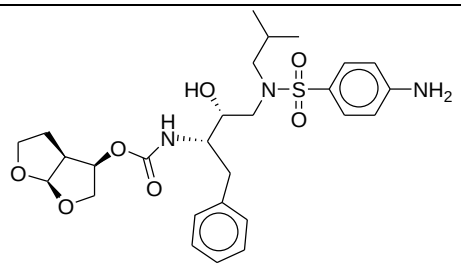
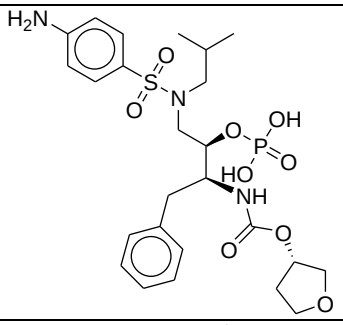
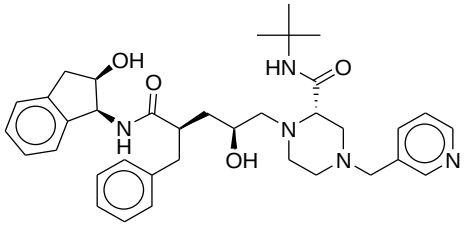
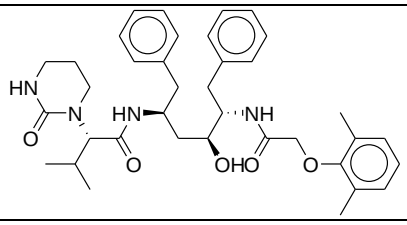
Material and Methods

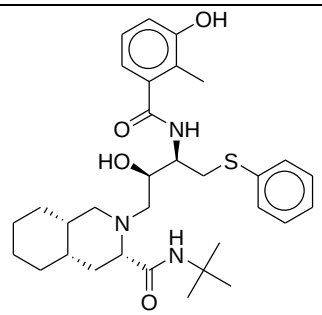
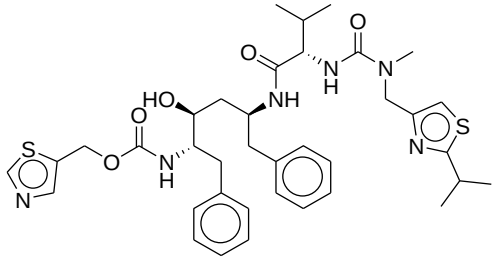
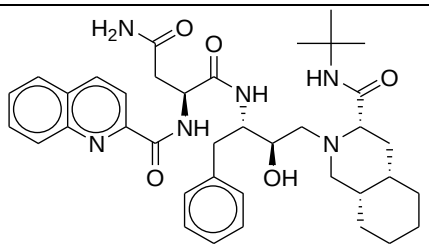
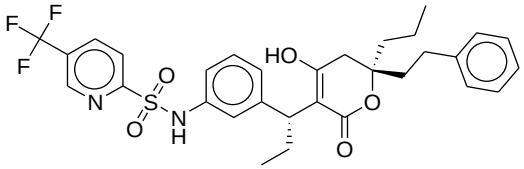
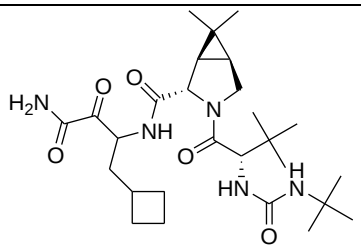
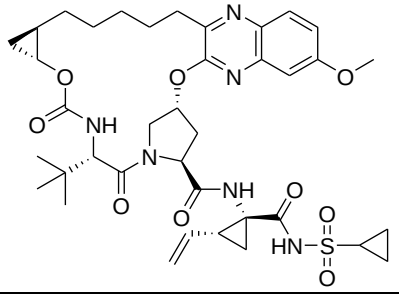
Pharmacophore Hypothesis-1(Ligand Based Pharmacophore Mapping)

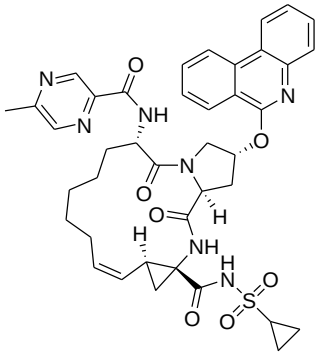
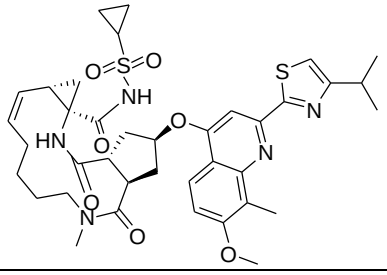
Ligand based pharmacophore mapping was generated using Sybyl X software from Tripos Inc., St. Louis, MO, USA[12].All the compounds were drawn by SKETCH function of SybylX. Partial atomic charges were calculated by the Gasteiger Huckel method was used to calculate charges. A Tripos force field was used with a distance-dependent dielectric and the Powell conjugate gradient algorithm conjunction criterion of 0.01 kcal/mol Å during the process of energy minimization. A total number of 10 models were generated for each hypothesis using DISCOtech. The best model generated for each hypothesis by DISCOtech was again selected for refinement by Genetic algorithm similarity program (GASP) which was based on genetic algorithm to characterise common features of different molecules from which hypothesis generated. The objective of this step was to produce 3D query through proper alignment, to flatten out rings so that spatial query can be created to use in substructure searching. By carrying out a GASP alignment, it was guaranteed that conformations which were used as input would look at least once in optimization. During this optimization, only stable and least energy conformer would be retained and propagated to the next iteration of the genetic algorithm. All parameters were kept as default other than population size (125), mutation weight (96), and fitness increment (0.02) and number of alignment (10)[13]. Clinically approved drugs have attracted researchers to design new

drugs as they have already proved efficacy and safety. They also provided another advantage that they contained chemical groups or features which were free from toxicities and had better pharmacokinetic properties.There are 14 FDA approved HIV protease inhibitors used in clinical practice[14, 15]. Hypothesis-1 was generated using these marketed drugs and Table 1 shows name and structure of these drugs.

Table 1 Structures of compounds used for pharmacophore hypothesis-1 generation

Structure of compounds used for generation of Pharmacophore hypothesis-1			
Sr. no	**Chemical derivatives**	**Structure**	**IC50 (nM)**
1	Amprenavir		0.027
2	Atazanavir		0.003
3	Darunavir		5
4	Fosamprenavir		0.41
5	Indinavir		0.05
6	Lopinavir		0.0013

7	Nelfinavir		0.002
8	Ritonavir		0.14
9	Saquinavir		0.0016
10	Tipranavir		0.002
11	Boceprevir		0.014
12	Grazoprevir		0.007

13	Paritaprevir		0.006
14	Simeprevir		0.0013

Pharmacophore Hypothesis-2(StructureBased Pharmacophore Mapping)

Structure based pharmacophore hypothesis was developed using E-pharmacophore tool of Schrödinger.A total number of three protein was used to generate hypothesis (PDBs: 1HXB, 3EKT, 3SPK). All these protein contained HIV-I protease enzyme bound with ligand molecule. All these proteins were prepared and minimized using protein preparation wizard tool of Schrödinger. Enzyme binding interactions were defined using amino acids residue involved with inhibitory molecules in complex. A Van der Waals scaling factor of 0.5 was used for excluded volumes in hypothesis to describe the shape of the protein binding site[16].

Validation of Pharmacophore Hypotheses

Statistical Validation

Sensitivity and specificity are statistical measures of the performance of a binary classification test, also known in statistics as classification function. Sensitivity measures the proportion of actual positives which are correctly identified as such. Specificity measures the proportion of negatives which are correctly identified.

- Sensitivity relates to the test's ability to identify positive results.

$$\textbf{Sensitivity} = \frac{\text{Number of true positives}}{\text{Number of true positives} + \text{Number of false negatives}}$$

$$\textbf{Specificity} = \frac{\text{Number of true negatives}}{\text{Number of true negatives} + \text{Number of false positives}}$$

Güner-Henry (GH) Scoring Method

The GH score is effectively useful to quantify the generated best model. The GH score should be from 0 to 1, where GH score near to 1 indicate perfect model and it should be above 0.7. It is one of the most appropriate method as it takes into account of both percentage yield of actives in database which is known as %Y recall and percent ratio of actives in hit list which is known as %A precision. The GH scoring method was applied to the previously mentioned 1395active HIV protease inhibitors with their 27900 conformers and the

decoy dataset molecules taken from http://dude.docking.org/to validate the pharmacophore models. The structures of this data set are provided in supplementary information. The GH analysis was carried out by considering following parameters:(a) the percent active (%A), which represented the coverage of activity space by the model; (b) the percent yield (%Y), which was a measure of the selectivity of the model; (c) the enrichment factor (E) and (d) the GH score. These variables were determined using information derived from the total number of compounds in the drug database (D), the number of actives in the database (A), the number of actives retrieved by the model (Ha), and the total number of hits retrieved by the model (Ht). The above mentioned parameters were calculated using following equation[17, 18].

$$\%A = \left(\frac{\mathrm{Ha}}{\mathrm{A}}\right) \times 100$$

$$\%Y = \left(\frac{\mathrm{Ha}}{\mathrm{Ht}}\right) \times 100$$

$$E = \left(\frac{\mathrm{Ha}}{\mathrm{Ht}}\right)/\left(\frac{\mathrm{A}}{\mathrm{D}}\right)$$

$$GH = \{\mathrm{Ha}(3\mathrm{A} + \mathrm{Ht})/4\ \mathrm{HtA}\} \times [1 - (\mathrm{Ht\text{-}\ Ha})/(\mathrm{D\text{-}A})]$$

Virtual Screening and Molecular Docking

Both ligand based and structure based pharmacophore hypothesis were followed by substructure searching from NCI and ZINC databases. The substructures were again screened for drug like properties by applying Lipinski's rule of five. The objective behind this filter was to obtain drug like candidate,so designed molecules may not be failed in any stage of drug development. The compounds thus obtained from databases after Lipinski's rule of five were followed by Q_{fit} value. Q_{fit} is parameter to check matching of pharmacophore with substructure means higher value of $Q_{fit,}$ better is matching and accuracy with structure. A total number of 256 substructures having Q_{fit} value greater than 99% were selected to design potent HIV-I protease inhibitors after considering knowledge based structure activity relationship i.e. Bioisosterism replacement. The major ring system obtained and used to design compounds were tetrazole, triazine, quinoline, pyrimidine, pteridine, benzthiazole etc. From these ring system, a total number of 50 compounds having triazine ring were selected for further studies[19].

Surflex docking tool of Sybyl X was used to dock all 25 designed molecules on HIV protease-I enzyme (PDB: 1HXB). The protein structure was downloaded from Protein data Bank (PDB). All designed molecules were minimized using Tripos force field after adding Gasteiger–Hückle charges. The energy minimization of enzyme was also an essential as ligand bound with enzyme when it's in stable and lower energy state. In order to produce stable conformer of HIV-I protease, a ligand bound with an enzyme and water molecules were removed during minimization of protein. Amino acid residues were fixed and hydrogen atoms were added. A protomol containing binding pocket was created where designed molecules may bind and may inhibit enzyme[20].

***In Silico* Pharmacokinetic, Toxicity and Bioactivity Prediction**

This methodology is more dominant during lead optimization as it can offer guidance regarding modification of structure which will improve physicochemical properties. ADMET (absorption, distribution, metabolism, excretion and toxicity) properties comprise the evaluation of various factors which are relevant not only for the estimation of the drug pharmacologically active concentration at the therapeutic targets but also for adverse effect and evaluation of toxicity. The designed compounds were tested for different pharmacokinetic properties such as blood-brain barrier (BBB), human intestinal absorption (HIA), caco-2 permeability (Caco-2), P-glycoprotein inhibitor (PGI), CYP450 (CYP) substrate/inhibitor, Ames toxicity and carcinogenicity[21].

Table 2 Result of top ranked pharmacophore mapping generated by DISCOtech (hypothesis-1)

Hypothesis	size	Hits	Score	Tolerance	D_{mean}
1	4	14	0.4848	0.2500	4.3538
2	4	14	0.4629	0.2500	4.3538
3	3	14	0.5091	0.2500	4.1245

Table 3 Results top ranked Pharmacophore hypotheses refined by GASP (hypothesis-1)

Hypothesis	Fitness	Size	Hits	D_{mean}
1	1625.8200	4	14	2.8562
2	1462.8800	4	14	3.3245
3	2232.1400	3	14	3.2547

Results and Discussion

Pharmacophore Hypothesis-1 (Ligand Based Pharmacophore Mapping)

The results of pharmacophore hypothesis-2 generated from 14 FDA approved HIV-I protease inhibitors obtained by DISCOtech are shown in Table2. The outcomes of GASP, after refinement, are shown in Table3. The pharmacophore hypothesis-1 contained1 donor site, 1 donor atom, 1 acceptor atom and 1 hydrophobic region which is displayed in Fig.1c as a three dimensional arrangement. A two dimensional query of hypothesis-2 is exhibited in Fig.1d.

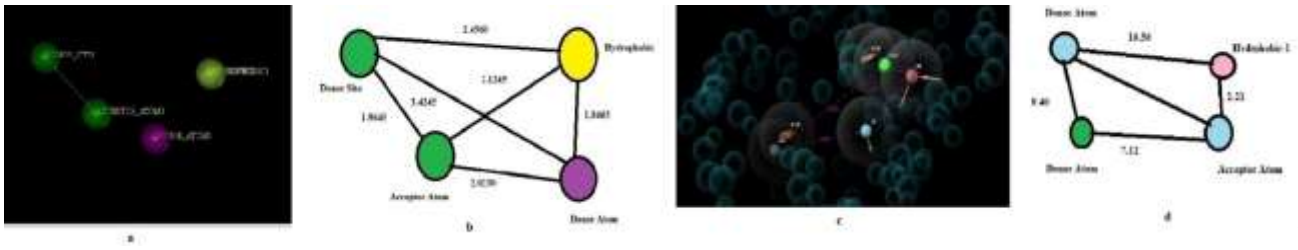


Fig. 1 Pharmacophore Hypothesis-1: (a) 3D representation; (b) 2D representation. Pharmacophore Hypothesis-2: (c) 3D representation; (d) 2D representation. Pharmacophore

Pharmacophore Hypothesis-2 (Structure Based Pharmacophore Mapping)

The results obtained from structure based pharmacophore from E-Pharmacophore is shown in Table4. The top ranked hypothesis which comprised of all 3 protein structures contained 1 hydrogen bond acceptor, 2 hydrogen bond donor and 1 hydrophobic feature. A Fig.1g and Fig.1hshowed three dimensional and two dimensional arrangement of pharmacophore features with distance geometry, respectively.

Table 4 Result of pharmacophore hypothesis-4 generated by E-Pharmacophore

Model	Features	Provenance
Model 1	ADDH	Common pharmacophore
Model 2	ADDHR	Common pharmacophore
Model 3	ADHR	Common pharmacophore
Model 4	AADHR	E-pharmacophore derived from 1HXB
Model 5	ADDHR	E-pharmacophore derived from 3EKT
Model 6	AADDRR	E-pharmacophore derived from 3SPK
A-Hydrogen bond acceptor, **D**-Hydrogen bond donor, **H**-Hydrophobic, **R**-Aromatic group		

Validation of Pharmacophore Hypotheses

Statistical Validation by Sensitivity and Specificity

Sensitivity and specificity of best model was evaluated by using in house database consisting of 41 compounds. Results of sensitivity and selectivity for 2 pharmacophore hypothesis including various parameters

for its calculations are given in Table5.1 which satisfied required criteria.The sensitivity was found as79, 75 for hypothesis-1 and 2 respectively, while specificity was found as 91 and 95 in same order.

Güner-Henry (GH) studies

All three pharmacophore hypotheses were further validated using Güner-Henry (GH) scoring method. The GH analysis was carried out by computing all parameters which are mentioned in material and methods section. The values of all parameters for all 3 hypotheses are shown in Table5.2. All parameters are statistically significant. The GH values of hypothesis-1, 2 3 and 4 are 0.83, 0.87, 0.85 and 0.89, respectively and are higher than 0.80.

Table 5.1 Result of validation of ligand based pharmacophore hypothesis by sensitivity and selectivity (hypothesis 1 and2)

Parameters	Hypothesis-1	Hypothesis-2
Number of true positives	15	3
Number of false negatives	4	1
Number of true negatives	20	38
Number of false positives	02	2
Sensitivity	79	75
Specificity	91	95

Table 5.2Pharmacophore model evaluation based on the Güner-Henry (GH) scoring method

Parameters	Statistical values (Hypothesis-1)	Statistical values (Hypothesis-2)
%A	89.9	92.71
%Y	92.82	91.79
E	3.99	3.07
GH	0.87	0.89
D	1436	1436
A	446	398
H_t	432	402
H_a	401	369

Ht- number of hits retrieved; Ha- number of actives in hit list; D - number of compounds in a database; %A- ratio of actives retrieved in hit list; %Y is a fractions of hits relative to size of database (hit rate or selectivity), E-enrichment of active bin by model relative to random screening ;GH - Güner-Henry score.

Virtual Screening and Molecular Docking

All pharmacophore hypotheses, after validation, were used for substructure searching in NCI and ZINC databases. All parameters were retained unmoved, but priority was kept high. A total number of 10247 and 1602 substructures were obtained for hypothesis-1 and 2 respectively. These substructures were again filtered to get drug like compounds by applying Lipinski rule of five and finally 8541and 1335 compounds were obtained from hypothesis-1, and 2 respectively. A total number of 9876 compounds were obtained through virtual screening followed by selection of compounds based on Q_{fit} value. The best 50 compounds from each hypothesis having Q_{fit} value greater than 99 were selected to design potent and bioactive HIV-I protease inhibitors. The detail and stepwise procedure for virtual screening of hypothesis-1 and 2 are given in Table 6. The common structure of designed molecules is displayed in Fig.2.

Fig. 2General structure of designed triazine derivatives.

Table 6. Result of virtual screening of Hypothesis-1 and 2

	Hypothesis-1			Hypothesis-2		
Steps	**In**	**out**	**Omission**	**In**	**out**	**Omission**
Remove counter ions	20127	20085	42	2048	2048	0
Remove duplicate structures	20085	15324	4761	2048	2048	0
Remove structures with bad fragments	15324	10247	5077	2048	1602	446
Apply Lipinski filtering	10247	8541	1706	1602	1335	267
Validate and convert structure formats	8541	8541	0	1335	1335	0

Table 7 Docking Results of Top 10 Molecules (Triazine derivatives on PDB Id: 1HXB)

Compound no.	Total score	Crash	Polar	Similarity
PP-HB-9	7.4200	-0.3800	0.0100	0.0800
PP-HB-21	7.3200	-0.1400	0.0200	0.0300
PP-HB-6	7.0900	-0.6400	0.0100	0.0600
PP-HB-13	6.8600	-0.2400	1.4300	0.0100
Saquinavir	6.6300	-0.3600	0.6800	0.0400
PP-HB-17	6.2200	-0.6300	1.1900	0.0040
PP-HB-14	6.1400	-0.8800	2.5700	0.0060
PP-HB-19	5.9200	-0.6500	0.0200	0.0100
PP-HB-12	5.8600	-0.6100	0.0100	0.0100
PP-HB-15	5.7000	-0.8900	0.0200	0.0100
PP-HB-2	5.6800	-0.3200	0.0100	0.0100
Amprenavir	5.7900	-1.6900	0.4600	0.0030

The docking results of top 10 compounds from designed molecules are given in Table7which showed better docking score along with proper binding affinity and hydrogen bonding interactions. A compound PP-HB-9 has the best docking score and binding interactionsof compound is shown inFig.3a. A total number of three hydrogen bonds are made by a compound with HIV-I protease. An Ile-50 made 2 hydrogen bonds, one with oxygen atom of ether linkage at distance of 2.17 Å and other hydrogen bond is made with oxygen atom of ketone group present on triazine ring. The third hydrogen bond is formed between oxygen atom of ketone and Gly-51. These non-covalent interactions of ligand with enzyme further confirmed binding and effective inhibition of protease enzyme. The effective binding of PP-HB-9 in binding pocket of HIV protease is displayed in Fig.3b.

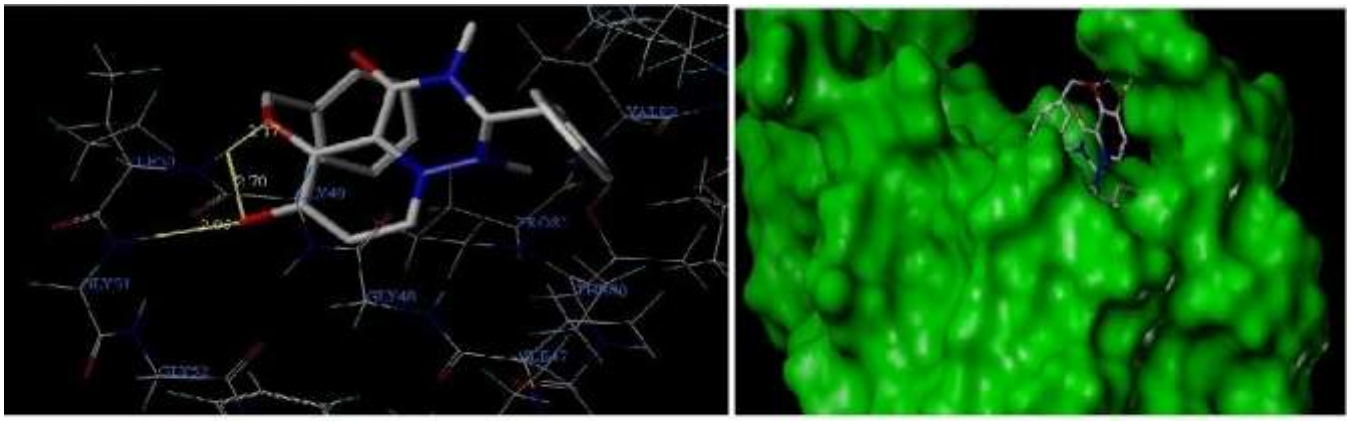


Fig. 3docking of compound PP-HB-9 on 1HXB: (a) binding interaction; (b) binding pose in cavity.

Table 8 ADMET prediction of hit compounds by SWISSADME and pkCSM

PROPERTIES	**PH-9**	**PH-21**	**PH-6**	**PH-13**	**PH-17**	**PH-14**	**PH-19**	**PH-12**	**PH-15**	**PH-2**
GI Absorptions	Low	Low	Low	High	Low	Low	High	Low	Low	Low
BBB Permeability	No	No	No	No	Yes	No	No	No	No	No
P- Glycoprotein Substrate	Yes	No	Yes	Yes	No	Yes	No	Yes	Yes	No
CYP1A2 inhibitor	No	No	No	No	Yes	No	No	No	Yes	No
CYP2C19 inhibitor	Yes	No	No	No	No	No	No	Yes	No	No
CYP2C9 inhibitor	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
CYP2D6 inhibitor	No	No	Yes	No	No	No	No	No	Yes	No
CYP3A4 inhibitor	Yes	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes
Log Kp (skin permeation)	-6.58 cm/s	-5.42 cm/s	-7.12 cm/s	-7.94 cm/s	-8.40 cm/s	-4.14 cm/s	-5.21 cm/s	-8.12 cm/s	-7.41 cm/s	-6.74 cm/s
Total clearance (log ml/min/kg	0.28	0.21	0.47	0.54	0.30	0.29	0.31	0.40	0.39	0.38
Renal OCT2 substrate	No	No	Yes	No	No	No	Yes	No	No	No
AMES toxicity	No	No	No	No	No	No	No	No	No	No
hERG I inhibitor	No	No	Yes	No	No	No	No	No	No	No
hERG I inhibitor	No	No	No	No	No	No	No	No	No	Yes
Hepatotoxicity	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes
Skin sensitization	No	No	Yes	No	No	No	No	Yes	No	No

*PH: PP-HB

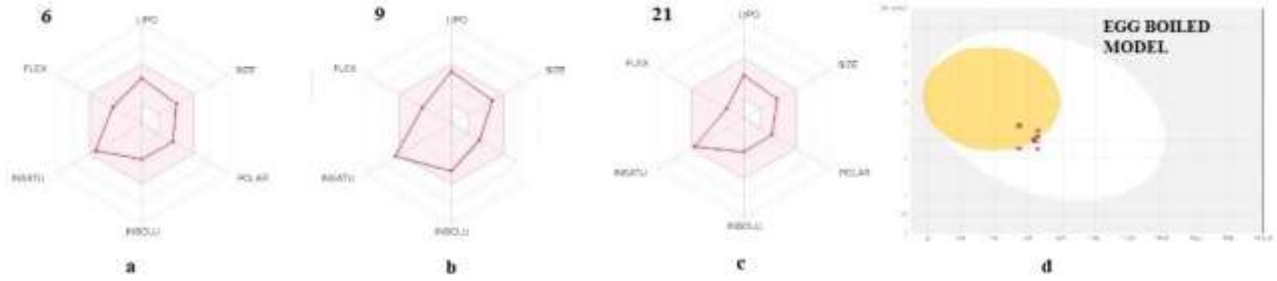


Fig. 4 ADMET prediction plot: (a) PP-HB-6; (b) PP-HB-9; (c) PP-HB-21; (d) egg boiled model for all 10 hits.

In silico Pharmacokinetic and Toxicity Prediction

ADMET prediction of hit compounds by SWISSADME and pkCSM are shown in Table8 which indicated that proposed molecules are safe and have better pharmacokinetic profiles. Most of compounds have less GI absorption while none of compounds have ability to cross BBB. P- Glycoprotein substrate affinity accounts for absorption and most of compounds are able to bind with it, which indicated that they may have good absorption. Most of the compounds did not inhibit any cytochrome enzymes. As the cytochrome p450 monooxygenase (CYP) enzymes super-family shows a key role in drug metabolism in the liver, especially 2D6, 2C9 and 3A4 which are the most important systems in human. According to the results, most of the drug candidates were seen to be negative to CYP2D6. All compounds showed good permeability towards biological membrane. AMES toxicity test was used to assess the compounds ability to damage DNA. None of compounds produced any damage or mutation in DNA which further proved safety of designed molecules. Most of compounds do not produce any sensitization on skin. The total body clearance is vital for estimation of dose and its frequencies which was measured by Log (CLtot) value with pkCSM online server to calculate the permutation of hepatic clearance (metabolism in the liver and biliary clearance) and renal clearance (excretion via the kidneys). Radar plots for PP-HB-6, PP-HB-9 and PP-HB-21 are shown in Fig.4a, 4b and 4c, respectively,while egg boiled model is given in Fig.4d. The radar plots are used to predict bioavailability. It can be observed that all three compounds have parameters in pink region which indicated that they have good oral bioavailability. Brain Or IntestinaL Estimate D permeation method (BOILED-Egg) is proposed as a precise predictive model which works on lipophilicity and hydrophilicity of drug like small molecules. Most of designed hits are observed near to white coloured space in model which indicated that compounds are well observed in GIT. The results obtained from Swiss ADME are promising and indicated that designed protease inhibitors have better bioavailability, hydrophilicity and lipophilicity properties.

Conclusion

The present study focus on combination of ligand based and structure based drug design approaches. Ligand based pharmacophore mapping was done using three different types of molecules including structurally diversified molecules, clinically used protease inhibitors and phytoconstituents which have proven protease inhibitory activities. Structure based pharmacophore mapping was done by taking 3 PDBs of HIV-I protease enzyme using E-Pharmacophore tool of Schrödinger. The chemical features generated from all pharmacophore were used to search for substructure followed by virtual screening to filter drug like compounds. A triazine derivatives were designed using these features and were docked to find initial hits. Molecular dynamics simulation study was again performed to access binding affinity of hit compound towards enzyme. The top ranked 10 compounds were also studied for *in silico* ADMET study so that compounds may not be failed in any stage of drug development. These 10 hits will be synthesized and evaluated for their anti-HIV activity to find lead compounds. The present study has discovered novel, potent and bioactive HIV-I protease inhibitor having better pharmacokinetic profile.

Acknowledgement

The authors PP and HB are thankful to Nirma University, Ahmedabad, India for providing necessary facilities to carry out the work, which is a part of Doctor of Philosophy (Ph.D.) research work of Mr. Paresh Patel, to be submitted to Nirma University, Ahmedabad, India. The authors PP and DM are also thankful to L. J. Institute of Pharmacy, Ahmedabad, India.

References

1. https://aidsinfo.unaids.org/ (Accessed on 10th April 2020)
2. Fabian L., Taverna Porro. M., Gómez N., Salvatori M., Turk G., Estrin D., Design, synthesis and biological evaluation of quinoxaline compounds as anti-HIV agents targeting reverse transcriptase enzyme, Eur J Med Chem., 2020, 188, 111987.
3. https://www.worldometers.info/aids/ (Accessed on 12th April 2020)
4. Lu D., Sham Y.Y., Vince R.,Design, asymmetric synthesis, and evaluation of pseudosymmetric sulfoximine inhibitors against HIV-1 protease, Bioorg Med Chem., 2010, 18, 2037–2048.

5. Dou Y., Zhu M., Dong B., Wang J.X., Zhang G.N., Zhang F., Design, synthesis and biological evaluation of HIV-1 protease inhibitors with morpholine derivatives as P2 ligands in combination with cyclopropyl as P1′ ligand, Bioorg Med Chem Lett., 2020, 30, 127019.
6. Agrafiotis D.K., Bandyopadhyay D., Wegner J.K., van Vlijmen. H.,Recent Advances in Chemoinformatics, J Chem Inf Model.,2007, 47, 1279–1293.
7. Patel S., Patel B., Bhatt H.,3D-QSAR studies on 5-hydroxy-6-oxo-1,6-dihydropyrimidine-4-carboxamide derivatives as HIV-1 integrase inhibitors, J. Taiwan Insti. Chem. Eng.,2016, 59, 61-68.
8. Dudek A.Z., Galvez T.A., Computational Methods in Developing Quantitative Structure-Activity Relationships (QSAR): A Review, Comb Chem High Throughput Screen., 2006, 9, 213–228.
9. Novikov F.N., Chilov G.G.,Molecular docking: theoretical background, practical applications and perspectives, Mendeleev Commun., 2009, 19, 237–242.
10. Deschamps N., Simões-Pires C.A., Carrupt P.A., Nurisso A.,How the flexibility of human histone deacetylases influences ligand binding: an overview. Drug Discov Today., 2015, 20, 736–742.
11. Ferreira L.L.G., Andricopulo A.D.,ADMET modeling approaches in drug discovery, Drug Discov Today., 2019, 24, 1157–1165.
12. Sybyl X (2011) Molecular modelling software. Tripos Certara,V.1.2, St. Louis
13. Bhatt H.G., Patel P.K.,Pharmacophore modeling, virtual screening and 3D-QSAR studies of 5-tetrahydroquinolinylidine aminoguanidine derivatives as sodium hydrogen exchanger inhibitors, Bioorg Med Chem Lett.,2012, 22, 3758–3765.
14. Wensing A.M.J., van Maarseveen N.M., Nijhuis M.,Fifteen years of HIV Protease Inhibitors: raising the barrier to resistance, Antiviral Res., 2010, 85, 59–74.
15. Voshavar C.,Protease Inhibitors for the Treatment of HIV/AIDS: Recent Advances and Future Challenges, Curr Top Med Chem., 2019, 19, 1571-1598.
16. Schrödinger Release 2020-1: Schrödinger, LLC, New York, NY, 2020.
17. Chaube U.J., Vyas V.K., Bhatt H.G., Design and synthesis of potent N-phenylpyrimidine derivatives for the treatment of skin cancer, RSC Adv., 2016, 6, 10285.
18. Bhatt H., Patel P., Pannecouque C., Discovery of HIV-1 Integrase Inhibitors: Pharmacophore Mapping, Virtual Screening, Molecular Docking, Synthesis, and Biological Evaluation, Chem Biol Drug Des., 2014, 83, 154–166.
19. Ozgencil F., Eren G., Ozkan Y., Guntekin-Ergun S., Cetin-Atalay R., Identification of small-molecule urea derivatives as novel NAMPT inhibitors via pharmacophore-based virtual screening, Bioorg Med Chem., 2020, 28, 115217.
20. Mohammadi A.A., Taheri S., Amouzegar A., Ahdenov R., Halvagar M.R., Sadr A.S., Diastereoselective synthesis and molecular docking studies of novel fused tetrahydropyridine derivatives as new inhibitors of HIV protease, J Mol Struct., 2017, 139, 166–174.
21. Hu B., Joseph J., Geng X., Wu Y., Suleiman M.R., Liu X., Refined pharmacophore features for virtual screening of human thromboxane A2 receptor antagonists, Comput Biol Chem., 2020, 86, 107249.

*****