IDENTIFICATION OF HIV-1 PROTEASE INHIBITORS, USING OLEANOLIC ACID AS REFERENCE LIGAND

ADITI GANGOPADHYAY[a], *, SAYAK GANGULI[a], ABHIJIT DATTA[c]

[a]*DBT Centre for Bioinformatics, Presidency University, Kolkata*
[c]*Department of Botany, Jhargram Raj College, Jhargram*

The HIV-1 Protease is a homodimeric enzyme which is required in the HIV life cycle for cleavage of virus proteins prior to maturation. Although quite a number of HIV-1 Protease inhibitor drugs have been established, the major setback faced by most Anti Retroviral Therapies (ARTs) is the fact that the HIV is very susceptible to getting mutated, and some of these mutations are able to confer drug resistance to the HIV. Therefore it is imperative to identify new molecules which would be active against the mutant strains of HIV. In this study, we have employed a hybrid structure-based and ligand-based computational approach to identify novel lead compounds against the wild type and mutant strains of HIV-1 Protease, using Oleanolic acid as the reference ligand. Comparison of the docking results of our screened library with established drugs indicated that some plant steroidal alkaloids, including Condurangamine B and Rostratamine, showed better docking results than the established drugs, and therefore could be further explored as lead compounds.

The HIV-1 Protease is a homodimer, falling within the class of Aspartyl Proteases, which have a conserved catalytic triad of Asp-Thr-Gly (Mager 2001). In case of the HIV-1 Protease, this comprises of Asp25-Thr26-Gly27 of both monomers. The Protease comes to use in the virus pathogenesis cycle following virus release, where it functions to cleave the virus proteins and promote virus maturation Kohl *et al* (1988). Therefore, any molecule that works to inhibit this Protease would, in effect, help to keep the virus in its immature state, and thus reduce virus infectivity. Established HIV-1 Protease inhibitor drugs include Saquinavir, Ritonavir, Indinavir and Neflinavir among others (Flexner, 1998).

In this study, we have employed a combined structure-based and ligand-based *in silico* method to identify lead compounds against the HIV-1 Protease. The ligand-based approach was pursued according to the molecular similarity principle, which states that the activity of a compound is a function of its structure; therefore compounds with similarity in their structures are presumed to function similarly. In the ligand-based approach we used Oleanolic acid as the reference ligand, since it has been shown to have anti-HIV-1 Protease activity *in vitro* Mengoni *et al* (2002). Using Oleanolic acid as the reference, we generated a screened compound library comprising of a total of 445 compounds from PubChem [Bolton *et al* (2008)] and our in-house library. In the structure-based aspect of the work, we subjected the structure of the wild type Protease to in silico mutagenesis to generate the structures of the mutant Proteases which have been associated in patients showing resistance to ARTs [Kandathil *et al* (2008)]. The mutations considered in the study were based on clinical findings reported in a South Indian population. Molecular docking simulations of

***Corresponding author.** *aig.bioinfo@gmail.com*

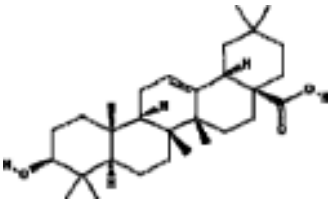
the screened library to the wild type and mutant Protease structures were performed using an hierarchical docking algorithm, and were scored using an empirical scoring function. The top scoring hits were identified and their interactions were studied to identify the lead compounds. Lead identification was also guided by the values of the ligand efficiency of the compounds.

The results indicated that some steroidal alkaloids from our in-house compound library showed strong binding affinities, in many cases these were well above those of the established drugs, with ligand efficiency values near to that of oral drugs, and therefore can be further studied and optimised as lead compounds.

MATERIALS AND METHODS

Generation of screened compound library: Using Oleanolic acid as the reference ligand, we identified 4264 compounds from PubChem [Bolton *et al* (2008)], which were screened on the basis of Lipinski's rule of five for drug-likeliness and ADMET risks, to generate 147 compounds. Since the structure of Oleanolic acid consists of a classical steroidal nucleus (Fig.1) we proceeded to generate an in-house library of plant steroidal alkaloids. The structures of the alkaloids were drawn manually using Marvin Sketch (2014), using information from literature Southon *et al* (2014). The 3D structures were generated using Ligprep (2009), and the compound pool was subsequently similarly screened on the basis of drug-likeliness and ADMET risks to generate 298 screened plant steroidal alkaloids.

Fig. 1: Oleanolic acid, a naturally occurring triterpene was used as the reference ligand for identification of similar compounds. Figure source: PubChem



Generation of mutant proteases using in silico mutagenesis, structural analyses:

The structure of the wild type HIV-1 Protease was retrieved from the PDB (PDB ID: 4DJR) Berman *et al* (2000), and the ligands were removed and the structure was prepared Fig.2. In silico mutagenesis of the wild type protease was performed using SPDBViewer (version 4.10) [Guex *et al* (1997)], to generate 3 mutants, which were refined and prepared using the protein preparation wizard of Schrodinger [Wizard (2007)]. The mutations introduced in the three mutants are shown in Table 1. The druggable pockets of the wild type and mutant proteases were identified with DogSiteScorer [Volkamar *et al* (2012)], which were used for defining the receptor grids [Table 2].

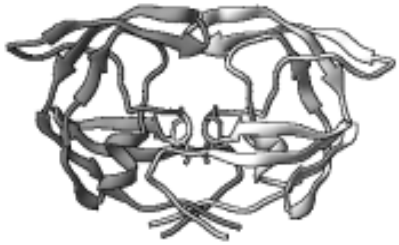


Fig. 2: The wild type HIV-1 Protease. PDB ID: 4DJR; Chain A: blue, Chain B: grey; the conserved catalytic triad of chain A and B has been highlighted in yellow and orange respectively

Table 1: Mutations introduced to generate the 3 mutant HIV-1 Proteases

Protease Mutant	Mutations
1	D30N,M46I, N88D, L10M, K20T, E35K, M36I, I93L
2	M46I, I47V, I54V, L10R, K20I, E35D, M36I, A71V
3	V82I, L90M, E35D, M36I, A71V, I93V

Table 2: Druggable pockets in the wild type and mutant HIV-1 Proteases

HIV-1 Protease	Pocket volume	Druggability score
Wild type	967.17	0.80
Mutant 1	899.65	0.81
Mutant 2	1107.33	0.81
Mutant 3	979.01	0.81

Docking, interaction analyses, comparison with protease drugs, lead identification:

The screened compound library was docked to the wild type and mutant Proteases, using the hierarchical docking algorithm of Glide [Friesner *et al*, 2004]. For identification of the top scoring hits, a two step approach was used: the compounds were first docked using the Glide Standard Precision (SP) mode, and the top scoring compounds from the SP run were re-docked using the Extra Precision (XP) mode, and the top scoring hits were identified. The interactions of the compounds with the proteases were studied with Maestro. The lead compounds were identified on the basis of their interactions with the residues of the catalytic triad (Asp25-Thr26-Gly27), XP scores, ligand efficiency values, which were compared to those of four established protease drugs: Saquinavir, Ritonavir, Neflinavir and Indinavir. The interactions were analysed with Maestro [Maestro Version 9.8].

Comparison of ADMET risks, drug-likeliness of leads with established drugs:

The ADMET (Absorption, Distribution, Metabolism, Excretion and Toxicity) risks and the drug-likeliness of the identified lead compounds were compared to those of four established HIV-1 Protease drugs - Ritonavir, Saquinavir, Indinavir and Neflinavir. For evaluation of the ADMET risks, the Qikprop suite of Schrodinger [QikProp, Version 3.2 (2012)] was used, while drug-likeliness was evaluated using Lipinski's rule of five [Hilbig *et al* (2013)].

RESULTS AND DISCUSSION

The structure of the refined HIV-1 Protease retrieved from the PDB (PDB ID: 4DJR) is represented in Fig. 2. Results of the docking studies of the screened compounds with the wild type protease indicate that Condurangamine B, a steroidal alkaloid from the plant Marsdenia condurango, interacted with the residues of the catalytic triad [Fig. 3a], and also had stronger binding affinity (XP score -8.891), than the established drugs Indinavir and Neflinavir (XP scores -8.429 and -7.857 respectively) [Table 3]. The ligand efficiency value of Condurangamine B was 0.262, which is very close to that of oral drugs (0.29).

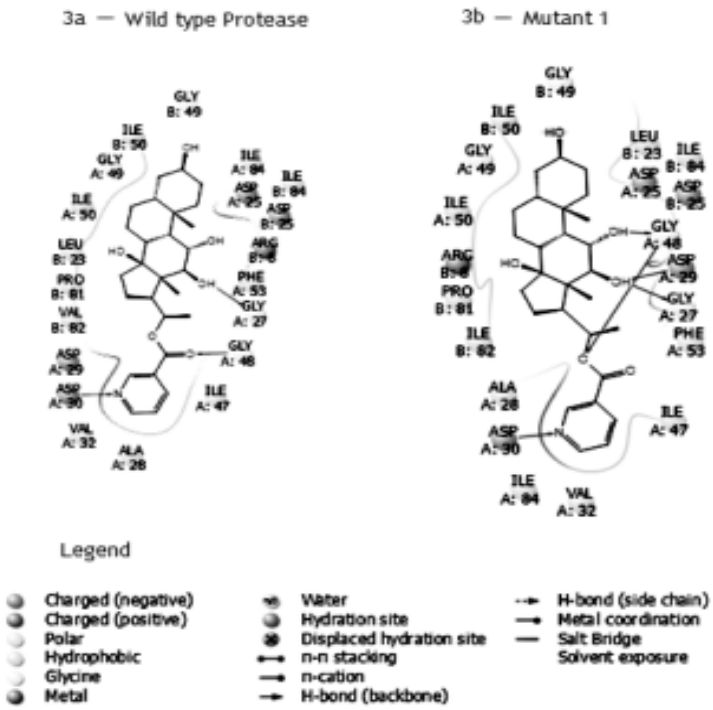


Fig. 3a, 3b: Analysis of interaction of Condurangamine B with the wild type and Mutant 1 HIV-1 Protease indicates several points of interaction with the residues of the catalytic triad

The docking scores of Condurangamine B with protease mutant 1 although were poorer than the established drugs in terms of binding affinity, the value of ligand efficiency (0.191) was better than that of all the established drugs, with the exception of Indinavir (0.216) [Table 3]. Analyses of the interaction of Condurangamine B with mutant 1 reveals that it interacts with the protease via 5 hydrogen bonds to atoms of the protein backbone [Fig. 3b], therefore indicating that the nature of the interaction is strong. In comparison, the established drugs Indinavir, Ritonavir, Saquinavir and Neflinavir interacts via 4, 4, 3 and 2 hydrogen bonds respectively with the protease mutant 1.

Results of the dockings with mutant 2 revealed that Rostratamine, another plant steroidal alkaloid showed the strongest binding affinity in terms of XP score, which was higher than that of the four established protease drugs - Neflinavir, Saquinavir, Indinavir and Ritonavir [Table 3]. The ligand efficiency value of Rostratamine was also very high (-0.252), very close to the range for oral drugs (Schultes *et al* 2010) [Table 3]. Analyses of the interactions showed that Rostratamine interacted with the protease via 2 hydrogen bonds to Arg25 - one of the residues of the catalytic triad [Fig. 4].

Table 3: XP Docking scores, ligand efficiency values of the top scoring compounds and established drugs with the wild type and mutant proteases

HIV-1 Protease	Compound	XP Score (kCal/mol)	Ligand efficiency (kCal/mol)
Wild type	Saquinavir	-10.568	-0.216
	Ritonavir	-9.758	-0.195
	Condurangamine_B	-8.891	-0.262
	Indinavir	-8.429	-0.187
	Nelfinavir	-7.857	-0.196
Mutant 1	Indinavir	-9.718	-0.216
	Ritonavir	-8.634	-0.173
	Saquinavir	-8.194	-0.167
	Nelfinavir	-7.481	-0.187
	Condurangamine_B	-6.479	-0.191
	Pingpeimine_C	-6.092	-0.179
	Rostratamine	-8.834	-0.252
Mutant 2	Indinavir	-8.368	-0.186
	Zygadenilic_Acid_Delta-Lactone	-6.360	-0.182
	Saquinavir	-5.327	-0.109
	Ritonavir	-5.228	-0.105
	Neflinavir	-4.664	-0.117
Mutant 3	Nelfinavir	-8.480	-0.212
	Indinavir	-8.400	-0.187
	Ritonavir	-7.851	-0.157
	Pingpeimine_C	-7.121	-0.209
	Saquinavir	-4.551	-0.093

For mutant 3, the binding affinity of Pingpeimine C, a steroidal alkaloid from the bulbs of Fritillaria ussuriensis, was stronger than that of the established drug Saquinavir, as indicated in the XP docking score [Table 3]. The compound performed better than Ritonavir and Saquinavir in terms of its ligand efficiency value.

Comparison of the ADMET risks and drug-likeliness of the identified leads with four established drugs (Saquinavir, Ritonavir, Neflinavir and Indinavir) indicate that all these four protease drugs have a number of ADMET risks, and violate some of the parameters of Lipinski's drug-likeliness thumb rule [Fig. 5]. On the other hand, the natural compound leads identified in the study show no violations of the drug-likeliness parameters, and have no ADMET risks, reflecting on the fact that the identified natural compounds are safer than the established drugs.

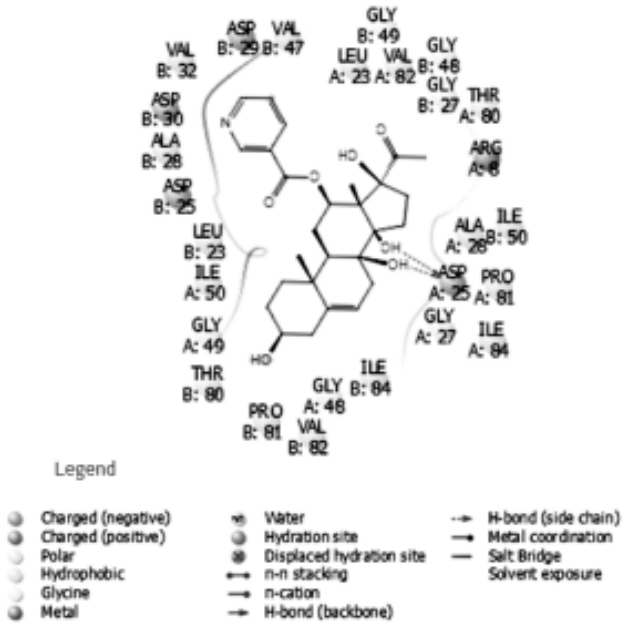


Fig. 4 Interaction of Rostratamine with Protease mutant 2 via hydrogen bonds with a residue of the conserved catalytic residue - Arg 25

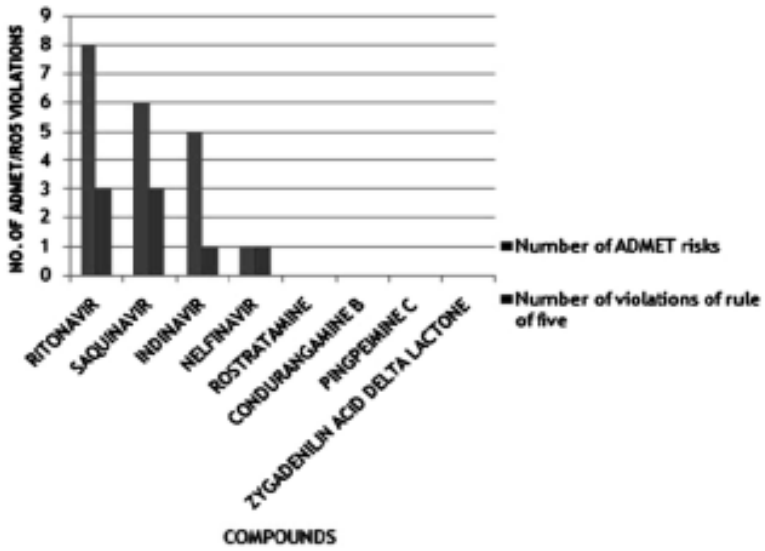


Fig. 5 Comparison of ADMET risks and drug-likeliness of the identified hits to established drugs

CONCLUSION

The docking results indicate that Condurangamine B, a steroidal alkaloid of plant origin, could be a better inhibitor of the wild type HIV-1 protease than the established drugs Indinavir and Neflinavir. Similar activity of Condurangamine B might also be proposed in case of protease mutant 1, where, although its binding affinity is poorer than the established drugs, it scores better in terms of its ligand efficiency value. The plant steroidal alkaloid Rostratamine, can be further explored for its inhibitory activity against the mutant HIV-1 protease carrying M46I, I47V, I54V, L10R, K20I, E35D, M36I and A71V mutations.

ACKNOWLEDGEMENTS

The authors remain grateful to Council for Scientific and Industrial Research (CSIR), India for providing funds, and to DBT, BIF-BTBI scheme, GoI for providing infrastructure. The authors would also like to thank the Team at the DBT Centre for Bioinformatics at Presidency University, Kolkata for being a constant source of support, help and advice.

REFERENCES

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., ... & Bourne, P. E. (2000): The protein data bank. Nucleic acids research, **28** (1), 235-242.

Bolton, E. E., Wang, Y., Thiessen, P. A., & Bryant, S. H. (2008): PubChem: integrated platform of small molecules and biological activities. Annual reports in computational chemistry, **4**, 217-241.

Flexner, C. (1998). HIV-protease inhibitors. N Engl J Med, **338**(18), 1281-92.

Friesner, R. A., Banks, J. L., Murphy, R. B., Halgren, T. A., Klicic, J. J., Mainz, D. T., ... & Shenkin, P. S. (2004): Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. Journal of medicinal chemistry, **47**(7), 1739-1749.

Guex, N., & Peitsch, M. C. (1997): SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modeling. electrophoresis, 18(15), 2714-2723.

Hilbig, M., Urbaczek, S., Groth, I., Heuser, S., & Rarey, M. (2013). MONA-Interactive manipulation of molecule collections. J. Cheminformatics, **5**, 38.

Kandathil, A. J., Kannangai, R., Abraham, O. C., Sudarsanam, T. D., Pulimood, S. A., & Sridharan, G. (2008): Genotypic resistance profile of HIV-1 protease gene: a preliminary report from Vellore, south India. Indian journal of medical microbiology, **26**(2), 151.

Kohl, N. E., Emini, E. A., Schleif, W. A., Davis, L. J., Heimbach, J. C., Dixon, R. A., & Sigal, I. S. (1988): Active human immunodeficiency virus protease is required for viral infectivity. Proceedings of the National Academy of Sciences, **85**(13), 4686-4690.

LigPrep 2.3 (2009): Schrödinger. LLC, New York

Mager, P. P. (2001): The active site of HIV-1 protease. Medicinal research reviews, **21**(4), 348-353.

Maestro, version **9.8**, Schrödinger, LLC, New York

Marvin Sketch version **6.2**, (2014): ChemAxon Ltd.

Mengoni, F., Lichtner, M., Battinelli, L., Marzi, M., & Mastroianni, C. M. (2002): In vitro anti-HIV activity of oleanolic acid on infected human mononuclear cells. pharmacological research, **3,** 8.

QikProp, version **3.2**, Schrödinger, LLC, New York, NY, 2012.

Schultes, S., de Graaf, C., Haaksma, E. E., de Esch, I. J., Leurs, R., & Krämer, O. (2010): Ligand efficiency as a guide in fragment hit selection and optimization. Drug Discovery Today: Technologies, **7 (3)**, e157-e162.

Southon, I. W., & Buckingham, J. (Eds.). (2010): Dictionary of Alkaloids, with CD-ROM (Vol. 2). CRC Press.

Volkamer, A., Kuhn, D., Rippmann, F., & Rarey, M. (2012): DoGSiteScorer: a web server for automatic binding site prediction, analysis and druggability assessment. Bioinformatics, **28** (15), 2074-2075.

Wizard, P. P. (2007): Schrödinger, LLC. New York.