Full Length Research Paper

Creating a high-throughput screening database to propose ligands capable of modulating the HIV-1 protease receptor

C. Micallef*, C. Shoemake and L. M. Azzopardi

Department of Pharmacy, Faculty of Medicine and Surgery, University of Malta, Msida-Malta.

This is a report of a *de novo* design study that aimed to identify novel structures capable of inhibiting the human immune deficiency virus (HIV-1) protease ligand binding pocket (HIV-1 PR_LBP). Baseline information regarding ligand binding modality and affinity was obtained through analysis of the pdb crystallographic depositions describing the HIV-1 PR enzyme complexed with small molecule inhibitors currently available on the Protein Data Bank (PDB). Molecular visualisation and modelling was carried out using SYBYL® 1.1, and *in silico* predicted ligand binding affinity (LBA) was quantified using XSCORE_V1.3. The *de novo* design phase of the study was based on the utilisation of the bound co-ordinates of Lopinavir. This particular HIV-1 protease (HIV-1 PR) inhibitor was selected as a template owing to its superior *in vivo* activity and unique binding modality. Based on literature derived data, the cyclic urea moiety of Lopinavir was retained as a seed fragment, overlaid onto its counterpart moiety and planted into the HIV-1 PR_LBP with growth being allowed according to defined parameters utilising the genetic algorithm embedded in the GROW module of LIGBUILDER®V1.2. The result was the identification of 200 *de novo* designed structures with a predicted *in silico* ligand binding affinity (LBA) (pK_d) ranging between 9.63 and 10.00. A smaller cohort (n = 35) was also Lipinski rule of 5 complaint. The implication of this study consequently is that this series of novel structures may be compiled into a library that may be of utility in high throughput screening (HTS) processes and future iterative optimisation.

Key words: High-throughput screening (HTS), protein data bank (PDB) depositions, ligand binding pocket (LBP), ligand binding affinity (LBA), human immune deficiency virus (HIV)-1 protease (HIV-1 PR), acquired immune deficiency syndrome (AIDS).

INTRODUCTION

Contemporary rational drug design is often attempted through a *de novo* approach. Ligand binding pocket (LBP) mapping and the identification of molecular moieties identified through X-ray crystallography as critical for molecular stabilisation are retained in the context of the creation of seed on which growing sites may be assigned such that novel affinity attachment may be sustained. Available software typically allows this process to be highly user driven such that the *de novo* designed structures lie well within Lipinski's rules for predicted bioavailability and low toxicity, and are also synthetically feasible (Lipinski et al., 2001). The value of the *de novo*

*Corresponding author. E-mail: chantelle.micallef89@hotmail.com.
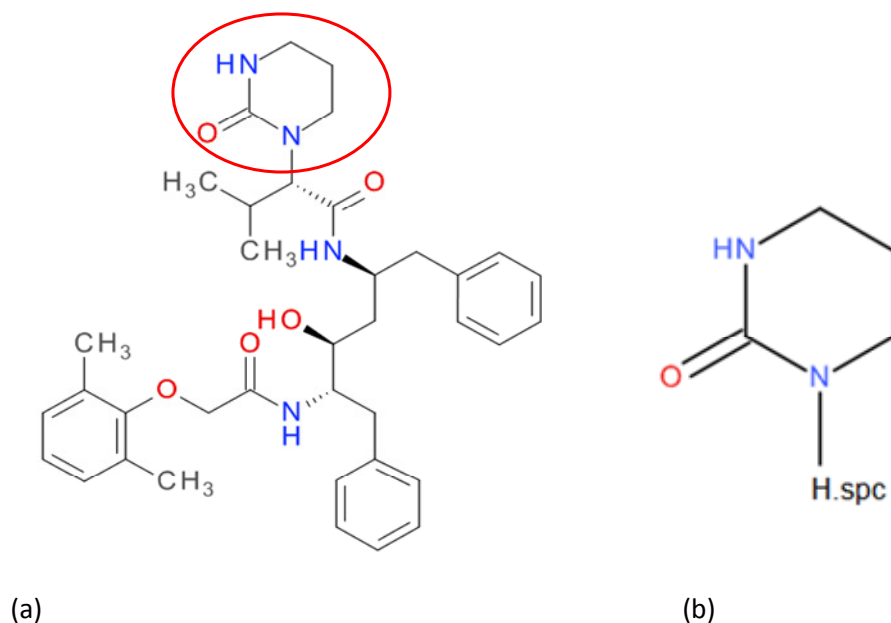
Figure 1. (a) The 2-D structure of the Lopinavir molecule with the cyclic urea moiety encircled in red whose chemical change is shown in Figure 1b. (b) The cyclic urea moiety whose adjacent Sp3 carbon atom was changed to H.spc.

approach is speed and low cost, together with the ability to yield novel structures with proven affinity for their target and which consequently represent valid candidates for further optimization (Audie et al., 2013; Mandal et al., 2009). They are also valuable candidates for inclusion into molecular databases for high-throughput screening (HTS).

MATERIALS AND METHODS

The X-ray crystallographic depositions describing the bound co-ordinates of the HIV-1 PR enzyme bound to small molecule inhibitors available on the Protein Data Bank (PDB) between 1999 to 2004 were identified. These 12 depositions specifically 1EBW (Andersson et al., 2003), 1EBY (Andersson et al., 2003), 1EBZ (Andersson et al., 2003), 1EC0 (Lindberg et al., 2004), 1EC1 (Andersson et al., 2003), 1EC2 (Andersson et al., 2003), 1EC3 (Andersson et al., 2003), 1D4I (Andersson et al., 2003), 1D4H (Andersson et al., 2003), 1W5V (Lindberg et al., 2004), 1W5X (Lindberg et al., 2004) and 1W5Y (Lindberg et al., 2004) were read into the molecular visualisation and editing programme SYBYL® 1.1 (SYBYL 7.3 Tripos International., 1699). Each individual deposition was edited in SYBYL® 1.1 such that at the end of this process a file saved in *mol2* format containing the bound co-ordinates of each small molecule and another saved in PDB format containing the *apo* receptor devoid of all water molecules lying at the distance ≤ 5Å of the LBP were generated. These files were exported to XSCORE_V1.3 (Wang et al., 1998). The score algorithm was then utilised in order to calculate the *in silico* predicted affinity (pK_d) between each HIV-1 PR conformation and its respective cognate small molecule which outcomes are shown in Figure 8. This

process effectively produced baseline affinity data which was set as benchmark for comparision in the *de novo* design phase of the study.

The *de novo* design phase of the study was based on the bound co-ordinates of the HIV-1 PR inhibiting small molecule Lopinavir complexed with the HIV-1 PR enzyme as described in the PDB deposition 1MUI. This particular deposition was selected based on the fact that Lopinavir is recognised to be a high affinity ligand with a unique binding modality for the HIV-1 PR_LBP (Stoll et al., 2002). The bound crystallographic co-ordinates of Lopinavir were consequently read into SYBYL® 1.1 such that seed fragments could be created for use in the *de novo* drug design phase of the study. During the seed creation phase, the cyclic urea moiety was retained. Retention of this cyclic urea was based on the fact that crystallographic evidence was suggestive of the fact that this moiety was capable of forging a novel hydrogen bonding arrangement within the HIV-1 PR_ LBP, specifically with Asp[29] that had not been observed with other HIV-1 PR inhibitors. Furthermore, hydrogen bonding at Asp[29] has been correlated with significant potency gains for this class of drug.

Based on this evidence, all the chemical moieties extraneous to the cyclic urea of Lopinavir were edited in SYBYL® 1.1. Subsequently, an atom type change from Sp3 carbon atom to H.spc hydrogen was effected as shown in Figure 1b. This was important in order to direct molecular growth during the *de novo* design process.

De novo design was carried out using LIGBUILDER®V1.2 (Wang et al., 2000). During the first stage of the process, the bound co-ordinates of Lopinavir were used as probes in order to delineate the 3-dimensional (3D) volume and the chemical nature of the perimeter of HIV-1 PR enzyme as described in PDB ID 1MUI and in Figures 9-11. This was carried out using the POCKET algorithm of LIGBUILDER®V1.2.
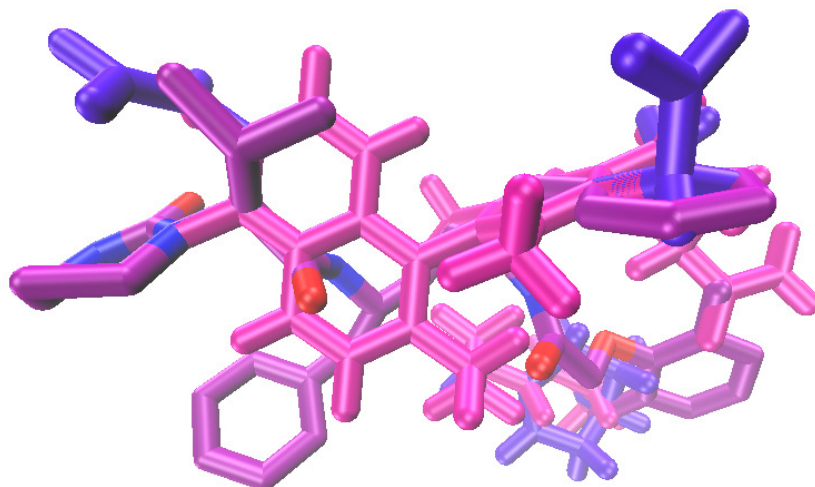
Figure 2. Superimposition of the highest affinity *de novo* molecules for family 1 (shown in violet and magenta) onto Lopinavir (shown according to molecular type).

Figure 3. Superimposition of the highest affinity *de novo* molecules for family 2 (shown in violet and magenta) onto Lopinavir (shown according to molecular type).
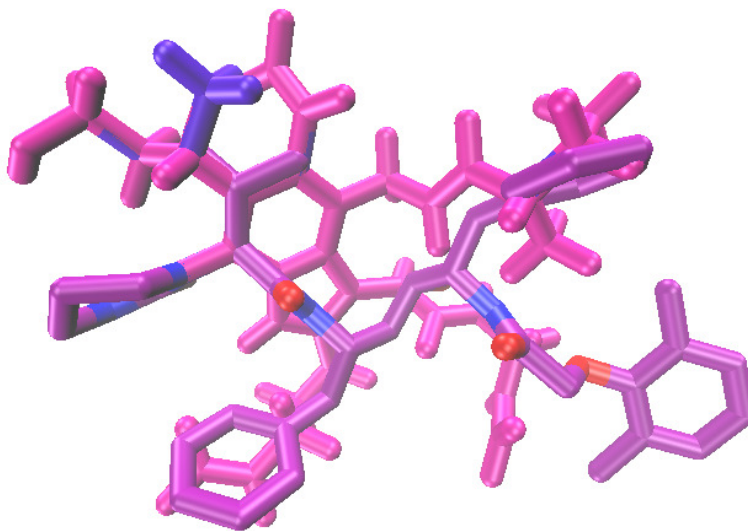
The seed fragment created in SYBYL® 1.1 with 3D co-ordinates identical to its counterpart moiety in Lopinavir and having a pre-designated growing site was then directed into the HIV-1 PR_LBP with molecular growth being allowed to occur according to the parameters of the generic algorithm embedded in the GROW module of LIGBUILDER®V1.2. This process resulted in the elaboration of a number of analog series (n = 6) containing varying amounts of molecular structures whose predicted *in silico* affinity (pK$_d$), molecular weight (Daltons/Da) and logP were quantified as shown in Figure 12-14 respectively. The binding poses of the highest affinity members of each family are superimposed onto the bound co-ordinate of Lopinavir as shown in Figures 2 to 7.

RESULTS AND DISCUSSION

All of the small molecule HIV-1 PR inhibitors which were utilised during the baseline data establishment stage of the study were assessed for Lipinski rule of 5 (predictors of *in vivo* bioavailability) rule compliance. It was interesting to note that all of these molecules which are either currently being utilised clinically namely Lopinavir or which are currently under experimental evaluation did not comply with Lipinski's rules.
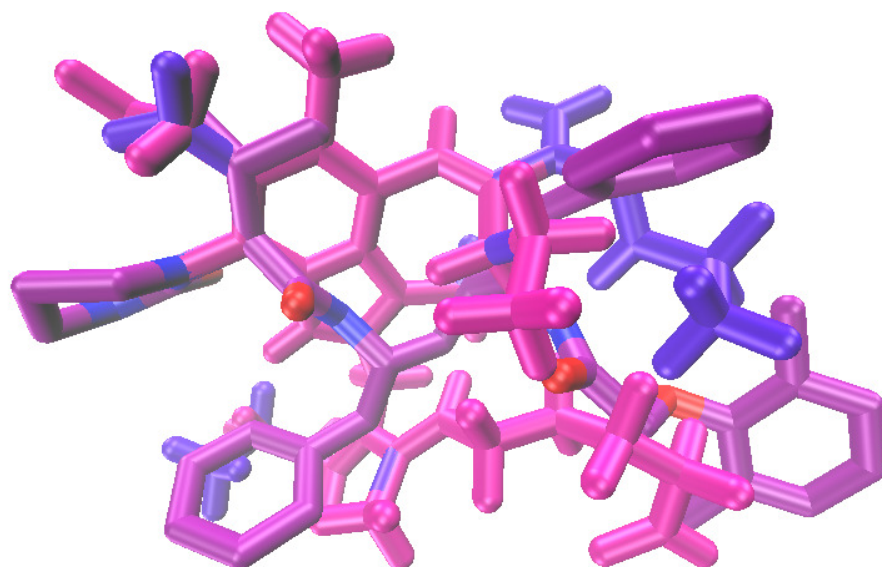
Figure 4. Superimposition of the highest affinity *de novo* molecules for family 3 (shown in violet and magenta) onto Lopinavir (shown according to molecular type).

Figure 5. Superimposition of the highest affinity *de novo* molecules for Family 4 (shown in violet and magenta) onto Lopinavir (shown according to molecular type).
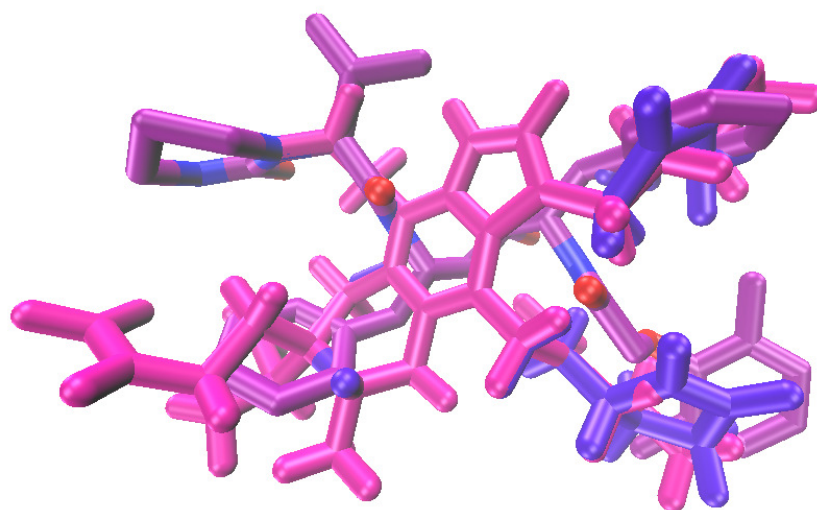
In all cases, non-compliance was attributed to a molecular weight that exceeded the cut-off value of 500 recommended by Lipinski et al. (2001) for acceptable *in vivo* bioavailability. Their molecular weight ranged from a minimum of 610 for 1D4H to a maximum molecular weight of 778 for 1EC1. All other molecules except 1D4H and 1D4I further violated Lipinski's rules with respect to the number of hydrogen bond donating and accepting

moieties, and 1EC1 also had a logP value of 6. All respective outcomes are shown in Table 1.

This data further reinforces the notion that there is scope for the design of improved novel members of this drug class. Given that Lipinski et al. (2001) recommend compliance with their rules in order to ensure acceptable bioavailability, and given also that evidence points to the fact the greater the number and extent to which these
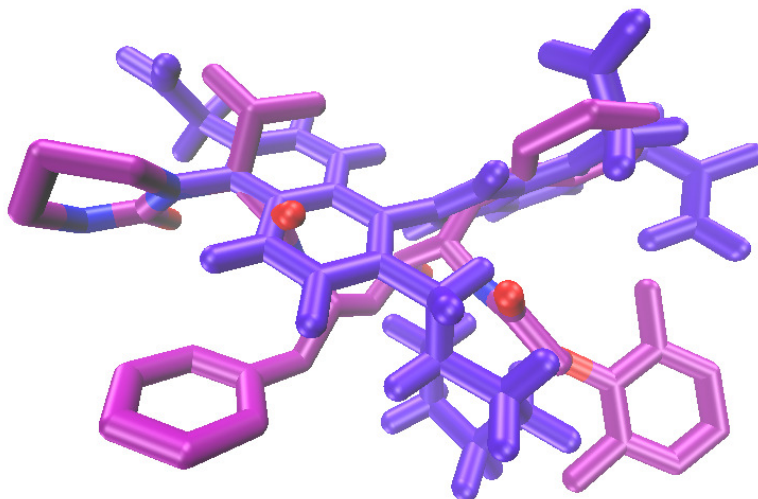
Figure 6. Superimposition of the only *de novo* molecule for family 5 (shown in violet) onto Lopinavir (shown according to molecular type).
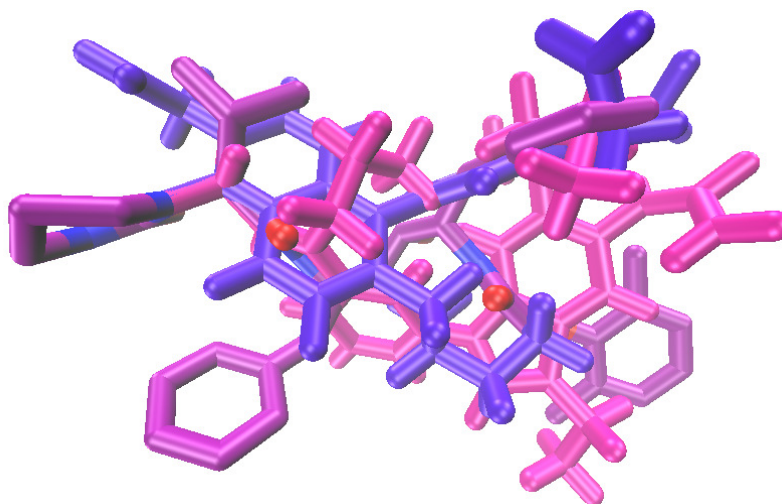

Figure 7. Superimposition of the highest affinity *de novo* molecules for family 4 (shown in violet and magenta) onto Lopinavir (shown according to molecular type).

rules are violated the lower is the predicted bioavailability, then the implication is that fine tuning of these HIV-1 PR inhibitors which are already being used successfully could result in novel structures with enhanced *in vivo* performance which could, in turn, imply decreased frequency of administration, better patient compliance and better overall disease management (David et al., 2013).. Furthermore, in a scenario in which drug resistance is an issue, increased patient compliance may also reduce the emergence of resistant strains.
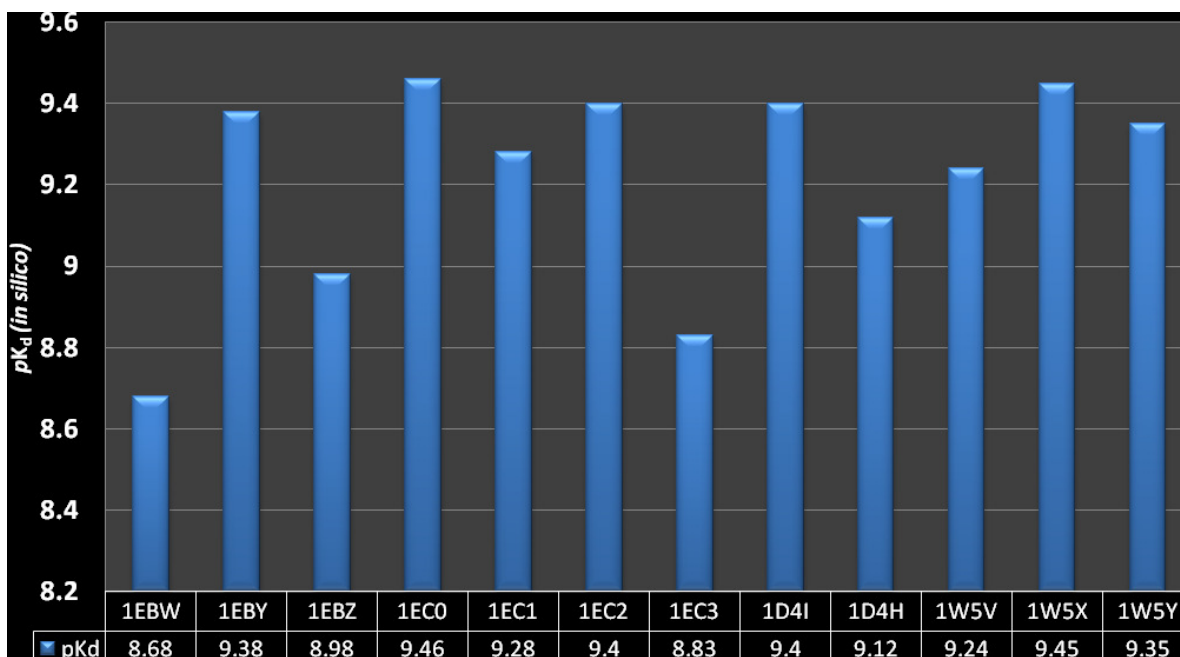
The *de novo* approach adopted in this study consequently attempted to identify novel high affinity HIV-1 PR inhibitors that were also Lipinski rule compliant. Lopinavir (PDB ID 1MUI) was selected as a template from which fragment seed structures were created owing to its quantified *in silico* ($pK_d = 8.44$) and *in vitro* (Ki = 1.3 pM) high affinity for the HIV-1 PR. Furthermore, its unique binding modality, coupled with the fact that its cyclic urea moiety was observed cystallographically to be capable of forging a hydrogen bond arrangement with

Table 1. The Lipinski Rule of 5 parameters for the selected PDB depositions and for Lopinavir.

PDB ID (Ligand ID)	Property	Pass/Fail the Lipinski rule of 5
IMU1 (AB1_100)	Molecular Weight = 628	Fail
	LogP = 4.8	Pass
	Number of Hydrogen donors = 4	Pass
	Number of Hydrogen acceptors = 9	Pass
1D4H (BEH)	Molecular Weight = 610	Fail
	LogP = 4.5	Pass
	Number of Hydrogen bond donors = 5	Pass
	Number of Hydrogen bond acceptors = 9	Pass
1D4I (BEG)	Molecular Weight = 636	Fail
	LogP = 3.7	Pass
	Number of Hydrogen bond donors = 5	Pass
	Number of Hydrogen bond acceptors = 9	Pass
1EBW (BEI)	Molecular Weight = 642	Fail
	LogP = 3.9	Pass
	Number of Hydrogen donors = 6	Fail
	Number of Hydrogen acceptors = 12	Fail
1EBY (BEB)	Molecular Weight = 652	Fail
	LogP = 3.4	Pass
	Number of Hydrogen bond donors = 6	Fail
	Number of Hydrogen bond acceptors = 10	Pass
1EBZ (BEC)	Molecular Weight = 633	Fail
	LogP = 3.2	Pass
	Number of Hydrogen bond donors = 6	Fail
	Number of Hydrogen bond acceptors = 11	Fail
1EC0 (BED)	Molecular Weight = 688	Fail
	LogP = 3.7	Pass
	Number of Hydrogen bond donors = 6	Fail
	Number of Hydrogen bond acceptors = 10	Pass
1EC1 (BEE)	Molecular Weight = 778	Fail
	LogP =6.0	Fail
	Number of Hydrogen bond donors = 6	Fail
	Number of Hydrogen bond acceptors =12	Fail
1EC2 (BEJ)	Molecular Weight = 768	Fail
	LogP = 4.3	Pass
	Number of Hydrogen bond donors = 6	Fail
	Number of Hydrogen bond acceptors = 14	Fail
1EC3 (MS3)	Molecular Weight = 768	Fail
	LogP = 3.8	Pass
	Number of Hydrogen bond donors = 6	Fail
	Number of Hydrogen bond acceptors = 14	Fail

Table 1. Contd.

1W5V (BE3_1100)	Molecular Weight = 688	Fail
	LogP = 3.7	Pass
	Number of Hydrogen bond donors = 6	Fail
	Number of Hydrogen bond acceptors = 10	Pass
1W5X (BE5_501)	Molecular Weight = 724	Fail
	LogP = 4	Pass
	Number of Hydrogen bond donors = 6	Fail
	Number of Hydrogen bond acceptors = 10	Pass
1W5Y (BE6_1100)	Molecular Weight = 724	Fail
	LogP = 4	Pass
	Number of Hydrogen bond donors = 6	Fail
	Number of Hydrogen bond acceptors = 10	Pass

	1EBW	1EBY	1EBZ	1EC0	1EC1	1EC2	1EC3	1D4I	1D4H	1W5V	1W5X	1W5Y
pKd	8.68	9.38	8.98	9.46	9.28	9.4	8.83	9.4	9.12	9.24	9.45	9.35

Figure 8. A graph showing the pK_d (*in silico*) for the twelve selected PDB depositions.

Asp[29] which has not, to date, been observed with other approved HIV-1 PR inhibitors, and established this drug as particularly interesting from a drug design point of view (Stoll et al., 2002).

Also, as evidenced in Table 1, Lopinavir was found to be Lipinski rule non-compliant exclusively from a molecular weight perspective, the implication being that it was a molecule that did not require radical intervention in order to ensure Lipinski rule compliance. The *de novo* approach adopted in this study consequently attempted to preserve the HIV-1 PR inhibitory action of Lopinavir and to design lower molecular weight molecules with a binding modality similar to that of Lopinavir. Maintenance of a binding modality to the HIV-1 PR enzyme similar to that of Lopinavir was approached through the creation of a seed structure that incorporated the moieties considered essential to the unique interaction of Lopinavir to the HIV-1 PR_LBP. Specifically, this included the
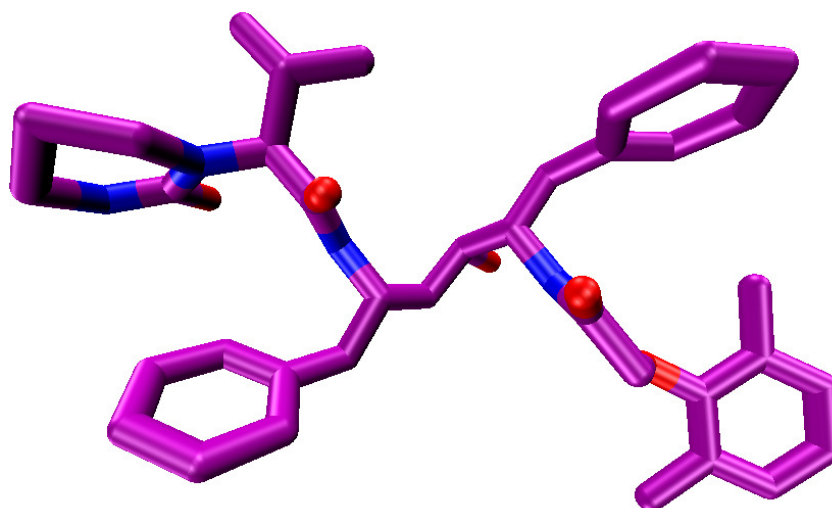
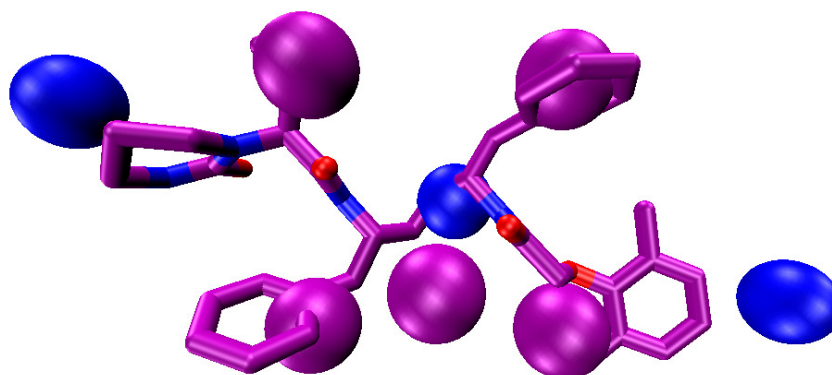**Figure 9.** The Lopinavir molecule with the cyclic urea encircled rendered in VMD (Humphrey et al., 1996).

Figure 10. The Lopinavir molecule onto its pharmacophore shown in beads rendered in VMD.
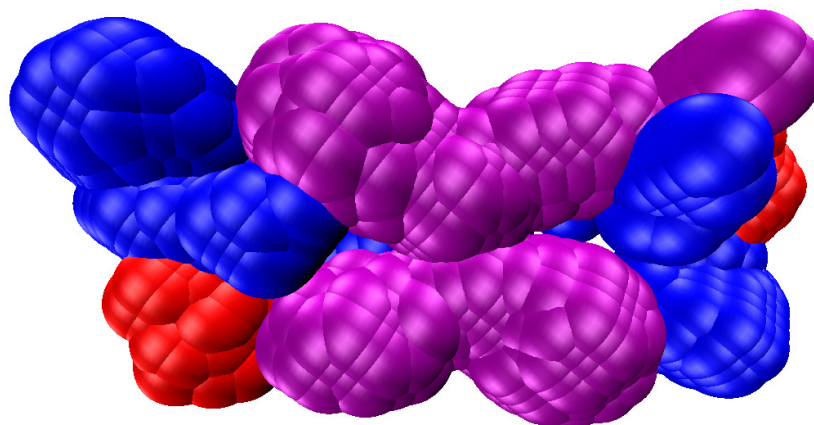

Figure 11. Key interactions sites of Lopinavir (hydrogen donor sites in blue, hydrogen acceptor sites in red and hydrophobic sites in purple) rendered in VMD.
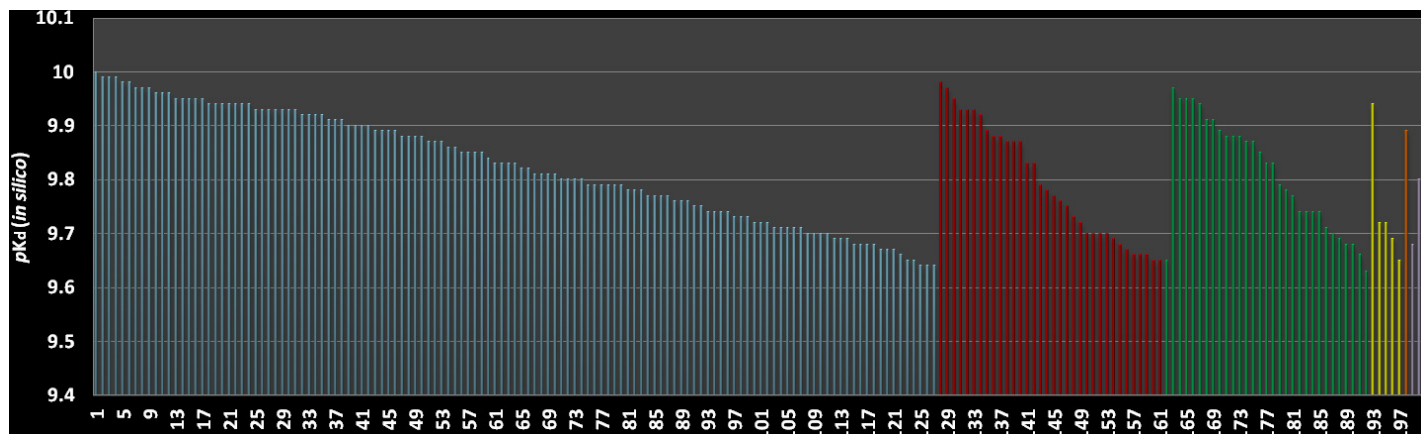
Figure 12. A graph showing the pK_d (*in silico*) for the 200 *de novo* designed ligands; the 6 different colours imply the 6 analog series.
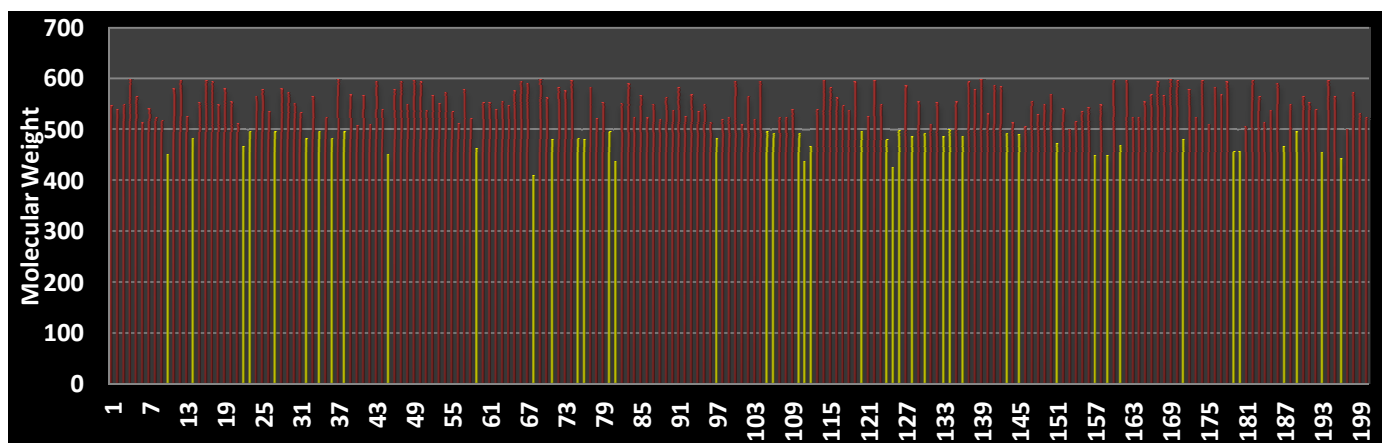

Figure 13. A graph showing the molecular weight for the 200 *de novo* ligands. The yellow colour indicates a molecular weight of less than 500 and thus compliance of such molecules with Lipinski rule of 5 with respect to molecular weight only.
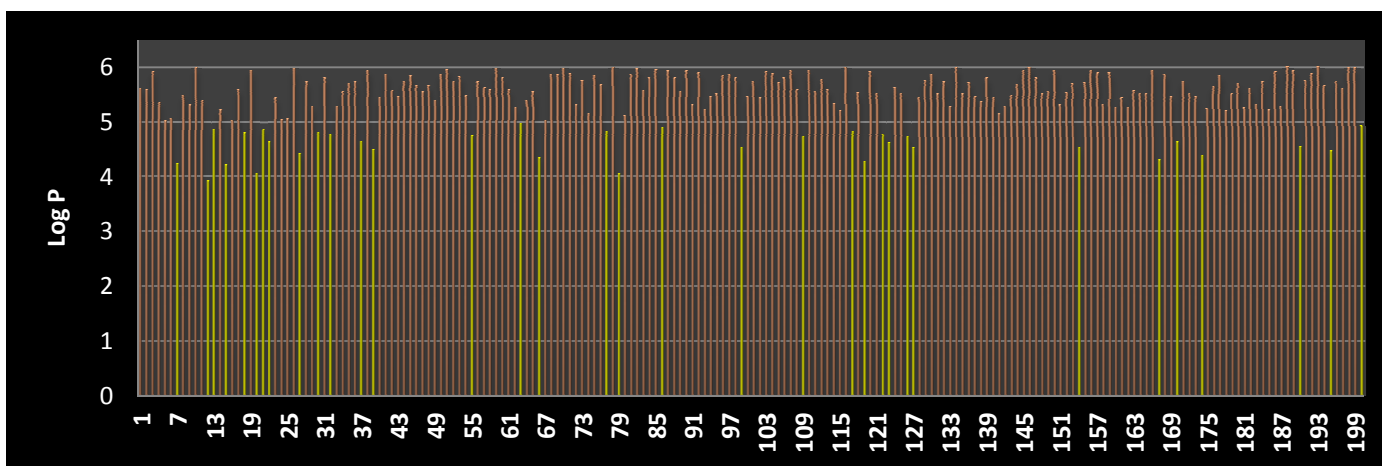

Figure 14. A graph showing the value of LogP for the 200 *de novo* ligands. The yellow colour indicates a value of less than 5 and thus compliance of such molecules with the Lipinski Rule of 5 with respect to LogP only.

Table 2. Molecular formula, binding affinity pK_d (*in silico*) and binding energy (Kcal mol^{-1}). These values are indicated in red for Lopinavir as a reference.

Ligand ID	Molecular formula	pK_d	Binding energy
Lopinvir	C37H48N4O5	8.84	20.643
10	C28H26N4O2	9.97	160.72
14	C28H26N4O4	9.95	88.228
22	C28H26N4O3	9.94	59.372
23	C29H28N4O4	9.94	48.969
27	C28H25N5O4	9.93	164.259
32	C28H27N5O3	9.93	60.724
34	C29H28N4O4	9.92	88.826
36	C28H26N4O4	9.92	79.936
38	C30H30N4O3	9.91	100.353
45	C28H26N4O2	9.89	148.948
59	C29H24N3O3	9.85	39.432
68	C26H23N3O3	9.81	142.23
71	C30H31N4O2	9.81	168.857
80	C32H29N5O3	9.79	51.441
81	C36H39N5O3	9.79	160.137
97	C29H39N5O3	9.74	66.713
105	C29H27N3O4	9.71	65.09
106	C29H29N5O3	9.71	187.298
110	C30H26N3O4	9.7	58.065
112	C28H23N3O4	9.7	44.389
120	C29H28N4O4	9.67	70.668
124	C29H28N4O3	9.65	111.497
125	C26H23N3O3	9.65	40.342
126	C28H27N5O4	9.64	64.22
128	C29H34N5O2	9.98	117.954
130	C28H40N6O2	9.95	150.674
133	C29H34N5O2	9.93	152.434
136	C29H36N5O2	9.88	145.68
143	C28H38N5O3	9.79	256.98
145	C28H36N5O3	9.77	320.506
157	C28H36N5O3	9.66	142.921
159	C26H34N5O2	9.66	171.695
179	C26H34N5O2	9.79	340.53
180	C28H32N4O3	9.78	264.248
187	C29H30N4O2	9.7	298.461

design process still lead molecules, implying that further optimisation must be carried out prior to *in vitro* evaluation. Ideally at this stage of the study, the *de novo* designed leads would have been rule of 3 compliant (Ress et al., 2004). The implications consequently are that a tight rope is still being walked in the quest to design clinically useful rule of 5 compliant high affinity HIV-1 PR inhibitors. However, these 35 molecules whose LBAs for the HIV-1 PR enzyme range from pK_d *in silico* of 9.63 to 10.00, warrant further investigation and evalua-tion. It must also be highlighted that this rational drug design exercise employed a static model that did not take into account the mutations that the HIV is known to undergo. However, the identification of novel structures potentially capable of inhibiting the HIV-1 PR enzyme is also valid in this context in which the availability of different molecular structures which stabilise and antagonise the HIV-1 PR_LBP could have a retarding effect on the development of mutations.

CONCLUSION

Finally, it must be stressed that in the context of the design of a chemical library of known affinity for the HIV-1 PR enzyme, the remaining 165 molecules should also be included. There are 2 very important reasons which justify such a decision - the first is that, as already pointed out, all the HIV-1 PR inhibitors in successful current use are Lipinski rule violators. The second is the very real threat of the emergence of resistant strains of HIV which necessitates the existence of a broad based drug armamentarium at the disposal of clinicians. Furthermore, these molecules could be further optimised, through for example, bioisosteric replacement, such that their molecular weight could be reduced to conform to Lipinski et al. (2001) requirements.

REFERENCES

Audie J, Swanson J (2013). Advances in the Prediction of Protein–Peptide Binding Affinities: Implications for Peptide-Based Drug Discovery. Chem. Biol. Drug Des. 81:50–60.
David J, Craik DJ, Fairlie DP, Liras S, Price D (2013). The Future of Peptide-based Drugs. Chem. Biol. Drug Des. 81: 136–147.
Humphrey W, Dalke A, Schulten K (1996). VMD – Visual Molecular Dynamics. J. Molec. Graphics. 14(1):33-38..
Lindberg J, Pyring D, Löwgren S, Rosenquist A, Zuccarello G, Kvarnström I, Zhang H, Vrang L, Classon B, Hallberg A, Samuelsson B, Unge T (2004). Symmetric fluoro-substituted diol-based HIV protease inhibitors. Ortho-fluorinated and meta-fluorinated P1/P1'-benzyloxy side groups significantly improve the antiviral activity and preserve binding efficacy. Eur. J. Biochem. 271(22):4594-602.
Lipinski CA (2000). Drug-like properties and the causes of poor solubility and poor permeability. J. Pharmacol. Toxicol. Methods. 44:235–249.
Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (2001). Experimental and computational approaches to estimate solubility and permeability

retention of cyclic urea with a change of the attached Sp3 carbon atom to an H.spc hydrogen as shown in Figures 1a, b and 9.

The genetic algorithm embedded in XSCORE_V1.3 generated 200 novel structures, 35 of which were Lipinski rule of 5 compliant (Table 2). This result must be dis-cussed critically with a view to further iterative optimisation in the drug design process. The 35 Lipinski rule of 5 compliant molecules are, at this stage of the

in drug discovery and development settings. Adv. Drug Deliv. Rev. 46:3-26.

Mandal S, Moudgil M, Mandal SK (2009). Rational drug design. Eur. J. Pharmaco. 625:90–100.

Stoll V, Qin W, Stewart KD, Jakob C, Park C, Walter K, Simmer RL, Helfrich R, Bussiere D, Kao J, Kempf D, Sham HL, Norbeck DW (2002). X-ray crystallographic structure of ABT-378 (lopinavir) bound to HIV-1 protease. Bioorg. Med. Chem. 10(8): 2803-2807.

SYBYL 7.3, Tripos International (1699). South Hanley Rd., St. Louis, Missouri, 63144, USA.

Wang R, Gao Y, Lai L (2000). LigBuilder: A Multi-Purpose Program for Structure-Based Drug Design. J. Mol. Model. 6: 498-516.

Wang R, Liu L, Lai L, Tang Y (1998). SCORE: A New Empirical Method for Estimating the Binding Affinity of a Protein-Ligand Complex. J. Mol .Model. 4:379-390.