Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry 13 (2005) 5492–5501
Bioorganic & Medicinal Chemistry

Combinatorial design of nonsymmetrical cyclic urea inhibitors of aspartic protease of HIV-1

Vladimír Frecer,[a,b] Enrico Burello[c] and Stanislav Miertus[a,*]

[a]International Centre for Science and High Technology, UNIDO, AREA Science Park, Padriciano 99, I-34012 Trieste, Italy
[b]Cancer Research Institute, Slovak Academy of Sciences, Vlárska 7, SK-83391 Bratislava, Slovakia
[c]Department of Biochemistry, Biophysics and Macromolecular Chemistry, University of Trieste, Via L. Giorgieri 1, I-34127 Trieste, Italy

Available online 27 July 2005

Abstract—The aspartic protease (PR) of the human immunodeficiency virus type 1 (HIV-1) is an important target for the design of specific antiviral agents dedicated to treatment of HIV-1 infection. We have employed computer-assisted combinatorial chemistry methods to design a small focused virtual library of nonsymmetrically substituted cyclic urea inhibitors of the PR. Nonsymmetrical compounds with decreased peptidic character were namely found to inhibit the PR with comparable inhibition potencies as their C_2-pseudosymmetric counterparts and to possess superior pharmacokinetic properties. To generate the virtual library of fully nonsymmetrical cyclic urea analogs, diverse reagents were selected from databases of available chemicals with characteristics similar to those of the building blocks of known potent PR inhibitors. The X-ray structure of the protease–inhibitor complex PR–XV-638 was used as the receptor model in the structure-based focusing and in silico screening of the virtual library. A target-specific LUDI-type scoring function, parameterized for a QSAR training set of known cyclic urea inhibitors and validated on a set of compounds not included into the training set, was used to predict the inhibition constants (K_i) of the generated analogs toward the HIV-1 PR. The fragments most frequently occurring in the analogs with the highest predicted inhibition potencies ($K_i^* < 10$ pM) were then selected to constitute a highly focused library subset containing novel nonsymmetrical cyclic ureas with predicted K_i^*s 1 order of magnitude lower than the most potent known cyclic urea inhibitors. ADME properties calculated for the most promising analogs suggested that the cyclic ureas are endowed with a wide range of favorable pharmacokinetic properties, which may favor the discovery of a potent orally administrable antiviral drug.

1. Introduction

The aspartic protease (PR) of the human immunodeficiency virus type 1 (HIV-1) cleaves the viral *gag-pol* fusion precursor polyprotein into active viral structural proteins and replicative enzymes such as reverse transcriptase, endonuclease, and integrase, thus playing an essential role in the maturation of HIV-1 particles and virus replication.[1] Therefore, PR is an important target for the design of specific antiviral agents dedicated to treatment of HIV-1 infection and acquired immunodeficiency syndrome (AIDS).[2] Unfortunately, constant viral RNA mutations and the ability of HIV-1 to rapidly generate drug-resistant PR mutants imply a continuing need for the discovery of new more potent protease inhibitors with improved pharmacokinetic properties active against a wider spectrum of PR mutant forms.[3,4]

Recently, cyclic ureas have been reported to constitute an entirely new class of potent and perspective non-peptidic inhibitors of PR, Fig. 1.[5–13] A fundamental feature of the cyclic urea inhibitors is the carbonyl oxygen that mimics the hydrogen-bonding features of the key structural water molecule present in the active site of the PR.[8] The presence of this structural water distinguishes the retroviral PR from human aspartic proteases pepsin and rennin. The success of the design in both displacing and mimicking the structural water molecule that bridges the substrate to the flaps of the protein by hydrogen bonds was confirmed by X-ray crystallographic studies.[9] Cyclic ureas proved to be highly selective and preorganized structures with decreased conformational flexibility, low molecular

Keywords: Inhibitor design; Aspartic protease of HIV-1; Library of nonsymmetrical cyclic ureas; Structure-based focusing; In silico screening; ADME properties.
* Corresponding author. Tel.: +39 040 922 8114; fax: +39 040 922 8115; e-mail: miertus@ics.trieste.it

0968-0896/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2005.06.026

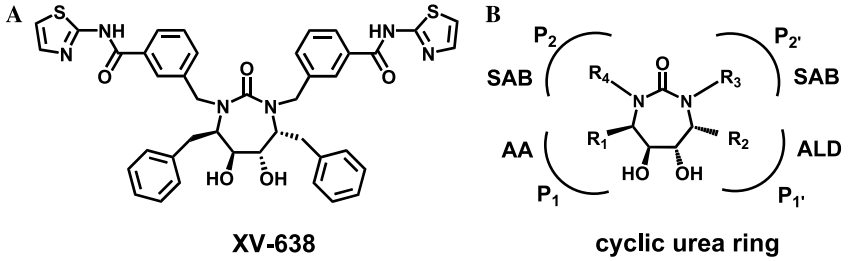


Figure 1. (A) Cyclic urea inhibitor XV-638 of DuPont Merck[23] with inhibition constant K_i of 30 pM against the wild type PR of HIV-1. (B) R-groups of cyclic ureas and their equivalence with the building blocks: amino acids (AA), aldehydes (ALD), and substituted arylbromides (SAB). The corresponding peptidic substrate residues notation scheme of Schechter and Berger[24] is also shown.

weight, improved water solubility, and oral bioavailability.[6]

Combinatorial chemistry techniques have been developed to supply large number of molecules prepared via parallel synthesis of compound libraries, which are screened in high throughput bioassays to increase the rate of generating new lead compounds with the desired potency and specificity.[14,15] Computational methods are being increasingly used to assist the combinatorial library design, focusing, and virtual screening by introducing selection criteria such as molecular diversity, drug likeness, predicted receptor binding, and ADME properties of analogs. Selection and focusing methods using these descriptors are employed to reduce the size of the combinatorial libraries to be prepared and screened.[16–18] Computational approaches can thus significantly reduce the cost, time, and labor required to synthesize and screen large libraries, as well as enhance the success rate in lead compound generation.

In this article, we have employed computer-assisted combinatorial chemistry methods to design, focus, and screen in silico a virtual library of nonsymmetrically $P_1/P_{1'}$ and $P_2/P_{2'}$ substituted cyclic urea inhibitors. Nonsymmetrical cyclic urea inhibitors were namely found to inhibit PR of HIV-1 with comparable potencies as their C_2-pseudosymmetric counterparts.[7,10] In addition, they possessed superior solubility and oral bioavailability, reduced molecular mass, and increased potential to translate strong enzyme inhibition activity into high antiviral potency.[7] We have derived a small highly focused library subset, which contains analogs with substituted aryl side chains that are predicted to contain cyclic ureas with PR inhibition constants (K_i) in the low picomolar range, a wide range of lipophilicity, and water solubility with a potential to be developed into an antiviral agent.

2. Results and discussion

2.1. Library design and fragment-based focusing

The central core of the cyclic urea inhibitor (seven-membered ring, Fig. 1) is synthesized from amino acid (AA) and aldehyde (ALD) building blocks (R_1 and R_2 substitution points), while alkylation of the urea nitrogens by substituted arylbromides (SAB) at positions R_3 and R_4 yields the complete inhibitor.[11] The number of cyclic urea analogs that can potentially be synthesized in a combinatorial experiment by using all the amino acids, aldehydes, and substituted arylbromides available in the commercial databases of suppliers of chemicals[19] ($\sim 10^{12}$ compounds) can easily exceed the capacity of any synthesis and screening equipment. Therefore, we have introduced a set of 'structural' filters and penalties to reduce the number of fragments employed in the library generation. The strategy of the fragment-based focusing relied on the predetermined optimum ranges of structural properties of the inhibitor building blocks, which were obtained by analyzing the crystal structures of potent peptidic and nonpeptidic inhibitors bound to the native PR (Table 1). To describe the properties and structural requirements for a successful fragment, 11 different descriptors related to shape, size, polarity, hydrogen bonding, lipophilicity, solubility, and conformational flexibility (Table 1) relevant to the match between the fragments and the characteristics of the S_3 to $S_{3'}$ specificity pockets of the PR-binding site, were computed using the Cerius^2 program.[20] Fragments whose descriptor values lied outside the optimal ranges and similar fragments were filtered out due to high penalty scores and diversity criteria. Finally, 6 amino acids, 3 aldehydes, and 18 substituted arylbromides were selected as suitable and diverse cyclic urea building blocks (Fig. 2). A higher number of SABs were chosen in line with the higher variability in the P_2 and $P_{2'}$ fragments of known potent peptidic and nonpeptidic PR inhibitors.[2] The fragment-based library focusing thus reduced the size of the designed cyclic urea library to:

$$6(R_1 - AA) \times 3(R_2 - ALD) \times 18(R_3 - SAB) \times 18(R_4 - SAB) = 5832 \text{ cyclic ureas.}$$

Generation of the virtual library was carried out by the attachment of the relevant R groups (side chains) of the selected sets of fragments to the central core (Fig. 1).[20]

2.2. Analog-based library focusing

The size of the generated fragment-focused library of cyclic ureas was further reduced by applying

Table 1. Analysis of physicochemical properties of known potent pseudopeptidic and nonpeptidic inhibitors of the HIV-1 PR and their building blocks in terms of ranges of selected descriptors[20]

Descriptor	Descriptor range[a]						
	P_3[b]	P_2	P_1	$P_{1'}$	$P_{2'}$	$P_{3'}$	Inhibitor[c]
Dipole moment (Debye)	1.3–5.3	0.9–6.9	0.1–4.4	0.6–6.3	0.7–6.4	1.2–6.0	1.2–24.2
Surface area (Å^2)	112–219	132–207	152–198	122–185	115–213	115–252	580–953
Principal moment of inertia (Da Å^2)	51–231	32–437	88–198	57–173	25–239	51–421	2221–9800
Molecular volume (Å^3)	81–162	89–157	94–146	94–150	78–162	83–191	460–754
Molecular weight (Da)	93–170	73–226	92–145	92–167	73 – 164	93–213	505–825
No. of rotatable bonds	0–3	0–3	0–2	0–2	0–3	0–4	12–23
No. of hydrogen bond acceptors	1–3	0–5	0–3	0–2	0–4	1–6	3–9
No. of hydrogen bond donors	0–3	0–4	0–2	0–2	0–3	0–3	2–7
Lipophilicity (AlogP98)	−0.5 to 2.8	−1.3 to 2.4	0.4–2.8	1.8–2.5	−0.3 to 2.4	−0.5 to 2.8	2.8–8.2
Solvation in water (kcal/mol)	−20.7 to −3.5	−20.2 to 1.2	−10.2 to −0.7	−10.2 to −0.9	−20.7 to −0.9	−20.6 to −3.5	−66.3 to −31.0
Solvation in octanol (kcal/mol)	−19.5 to 4.9	−2.7 to 0.6	−13.7 to 1.4	−13.7 to −4.5	−20.7 to −2.7	−3.7 to 1.2	−66.4 to −36.3

[a] Descriptor ranges have been derived by analysing the properties of building blocks (residues) of selected potent pseudopeptidic and nonpeptidic HIV-1 PR inhibitors, for which crystal structures of PR–I complex have been published in the Protein Data Bank.[37]

[b] Fragments P_3–$P_{3'}$ represent amino acid residues or the corresponding building blocks of cyclic ureas that occupy the specificity pockets S_3-$S_{3'}$ of the active site of the HIV-1 PR[24] (Fig. 1).

[c] Set of HIV PR inhibitors considered (PDB entry):[37] ritonavir {ABT-538} (1HXW), amprenavir {VX-478} (1HPV), indinavir {L-735,524} (2BPX), saquinavir {Ro 31-8959} (1HXB), nelfinavir {AG-1343} (1OHR), lopinavir {ABT-378} (1MUI), tipranavir {PNU-140690} (1D4Y), A77003 (1HVI), KNI-272 (1HPX), KNI-764 {AG-1776} (1MSM), SB-203386 (1SBG), GR-137615 (1HTG), L-700417 (4PHV), DMP-323 (1QBS), DMP-450 (1DMP), XV-638 (1QBR), SD-146 (1QBT), XK-263 (1HVR).

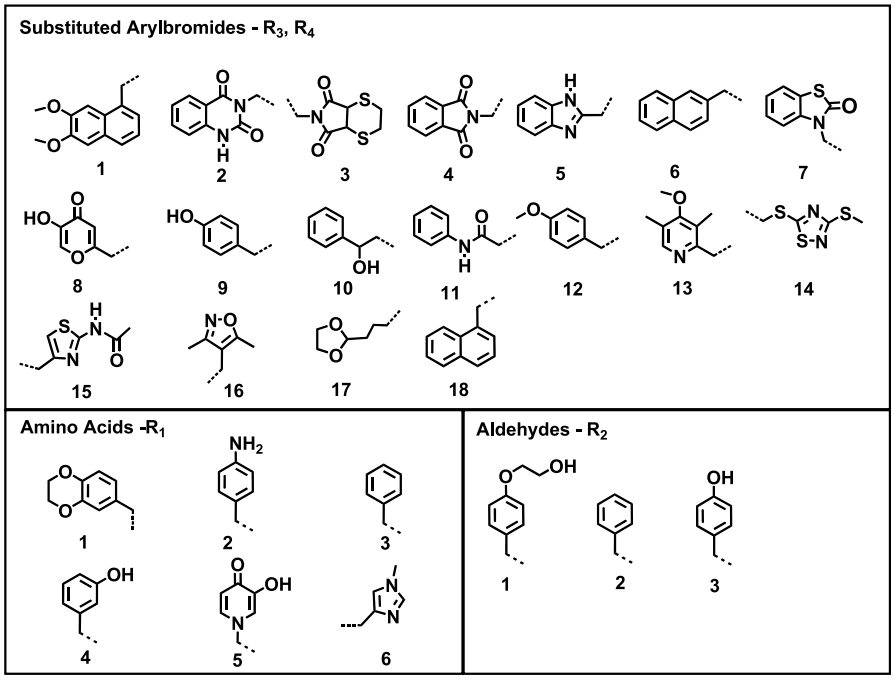


Figure 2. The R-groups (fragments, building blocks)—'side chains' amino acids (AA), aldehydes (ALD), and substituted arylbromides (SAB) selected by fragment-based library focusing.

analog-based focusing methods. Analog filtering procedure was based on the molecular physicochemical descriptors and optimum molecular property ranges calculated for the set of bound conformations of the known PR inhibitors (Table 1), which permitted to select the 1000 most diverse cyclic urea analogs with suitable molecular characteristics for the subsequent structure-based focusing.

2.3. Structure-based library focusing

The size of the analog-focused library of cyclic ureas was further reduced by applying structure-based focusing and in silico screening methods with the goal to select a small restricted combinatorial subset containing analogs with the best binding affinities to HIV-1 PR as possible lead compounds for antiviral drug development. All generated cyclic urea analogs were docked to the model of the PR receptor derived from the crystal structure of PR–XV-638 complex[23] by employing the Monte Carlo ligand fit algorithm of Cerius[2].[20,36] The docking procedure yielded 10 best binding conformers per analog that were clustered into five conformational families based on their mutual rms deviations. In each cluster, the conformer that displayed the highest docking score was selected for virtual screening.

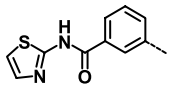
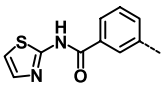
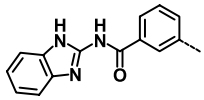
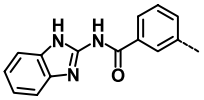
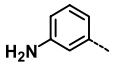
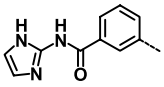
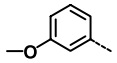
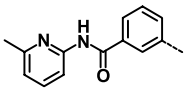
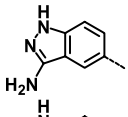
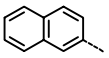
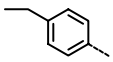
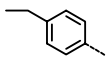
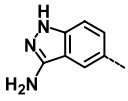
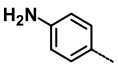
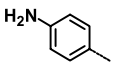
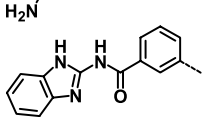
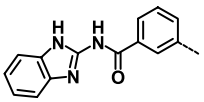
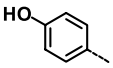
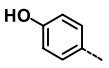
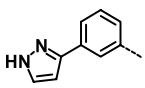
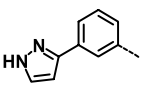
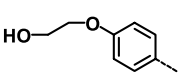
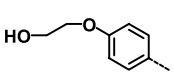
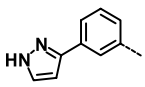
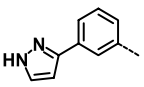
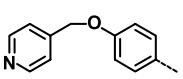
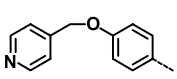
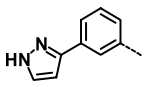
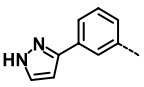
2.4. QSAR analysis of known cyclic urea inhibitors and scoring function parameterization

To obtain a LUDI-type scoring function[26] capable of predicting in vitro inhibition potencies of cyclic ureas against the wildtype PR of HIV-1, we have correlated experimentally determined K_is of a training set of 12 known cyclic urea PR inhibitors (Table 2) with the scoring function variables computed for the known inhibitors. The two quantities HB_{score} and VDW_{score} characterize the strength of hydrogen bonding and lipophilic interactions between the ligand and the receptor.[26] The known inhibitors were submitted to the same Monte Carlo docking procedure to the HIV-1 PR receptor model that was applied to the library of the designed nonsymmetrical cyclic ureas. The resulting QSAR equation was obtained by linear regression (Fig. 3):

$$pK_i = 0.0016 \cdot HB_{score} - 0.0038 \cdot VDW_{score} + 1.8915 \quad (1)$$

[number of samples $n = 12$, correlation coefficient $R^2 = 0.90$, leave-one-out cross validated correlation coefficient $R^2_{xv} = 0.82$, standard error = 0.11, Fisher F-test = 40.12, confidence level $\alpha > 95\%$, K_i (pM)]. Relatively high value of the crossvalidated R^2_{xv} indicates that major portion of the variance of the training set data were successfully explained by our QSAR model. The quality of the derived scoring function was validated on a set of five cyclic ureas with known inhibition constants (not included into the training set), which confirmed its predictive power (ratio $pK_i^{calcd}/pK_i^{exp} \sim 1$; Table 2). The target-specific scoring function was then applied in the in silico screening step to predict the K_i^* constants toward the HIV-1 PR for the designed nonsymmetrical cyclic urea analogs. The successful QSAR model confirmed the validity of the computational procedures used and the correct choice of adjustable parameters in the ligand docking algorithm.[20,36]

Table 2. Symmetrical and nonsymmetrical cyclic ureas used as the training and validation sets in the QSAR model of HIV PR inhibition and in parameterization of the scoring function used for in silico screening of the designed library

Cyclic urea	R_1[a]	R_2	R_3	R_4	K_i^{exp}[b] (pM)
Training Set					
CU-1 XV-638					30
CU-2 SD-146					24
CU-3					12
CU-4					53
CU-5					23
CU-6					72
CU-7					16
CU-8					16
CU-9					59
CU-10					71
CU-11					130
CU-12					61
Validation set					pK_i^{calcd}/pK_i^{exp}[c]
CU-13 DMP-450					1.21
CU-14					0.98
CU-15					1.25
CU-16					1.12
CU-17					0.89

[a] For position of R_1–R_4 substituents see Figure 1, dashed bonds indicate the attachment points.

[b] Experimental inhibition constants of the training and validation sets of cyclic ureas toward the wildtype HIV PR were taken from Refs. 7,10,12,28–30.

[c] Ratio of predicted and experimental pK_is of the validation set; pK_i^{calcd} was computed from Eq. 1.

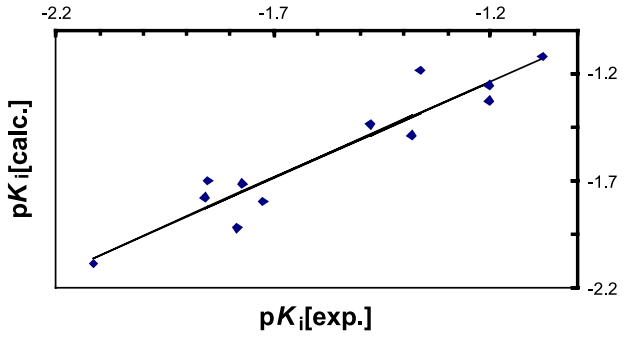


Figure 3. Linear regression analysis of a QSAR training set of known cyclic urea inhibitors of HIV-1 PR. Eq. 1 relates the experimental inhibition constants K_i (Table 2), to the LUDI-type scoring function variables[26] to give a target-specific scoring function.

2.5. In silico screening

The best conformers of the cyclic urea analogs were screened and ranked according to the pK_i^* values predicted by the trained scoring function (Eq. 1, Fig. 3). This in silico screening for potential lead compounds resulted in the identification of nine most promising fully nonsymmetrical cyclic urea analogs (Fig. 4) with predicted K_i^* constants in the low picomolar and subpicomolar range (Table 3). The predicted K_i^*s of the leads are up to 2 orders of magnitude lower than the K_is of the cyclic ureas of the training set (Table 2). This is not surprising since the replacement of phenyl rings of residues R_3 and R_4 (Fig. 1) by bulkier bisbenzamides, aniline amides, and aminoindazoles led to a significant decrease in the PR inhibition constants of both symmetrical and nonsymmetrically substituted cyclic ureas.[7,10] For example, substitution of phenyl rings of the R_3 and R_4 groups by hydrogen bond acceptors in the positions 3 and 4 of phenyl ring resulted in up to a 200-fold increase in the PR inhibitory potencies of azacyclic ureas due to favorable interaction with the main chain amide groups of the PR residues Asp:A29, Asp:B29 and Asp:A30, and Asp:B30.[31] The 3D structure of the most promising designed cyclic urea (analog 5-3-16-10; Fig. 4) bound at the active site of the wildtype PR of HIV-1 is shown in Fig. 5. The R groups of the analog form mainly hydrophobic interactions with the residues lining the S_2, S_1, $S_{1'}$, and $S_{2'}$ specificity pockets of the PR-binding site.[24]

Predicted lipophilicity ($\log P_{o/w}$)[32], aqueous solubility ($\log S_w$)[38,39] Caco-2 cell-membrane permeability (BIP_{caco}[39]), and number of likely metabolic reactions

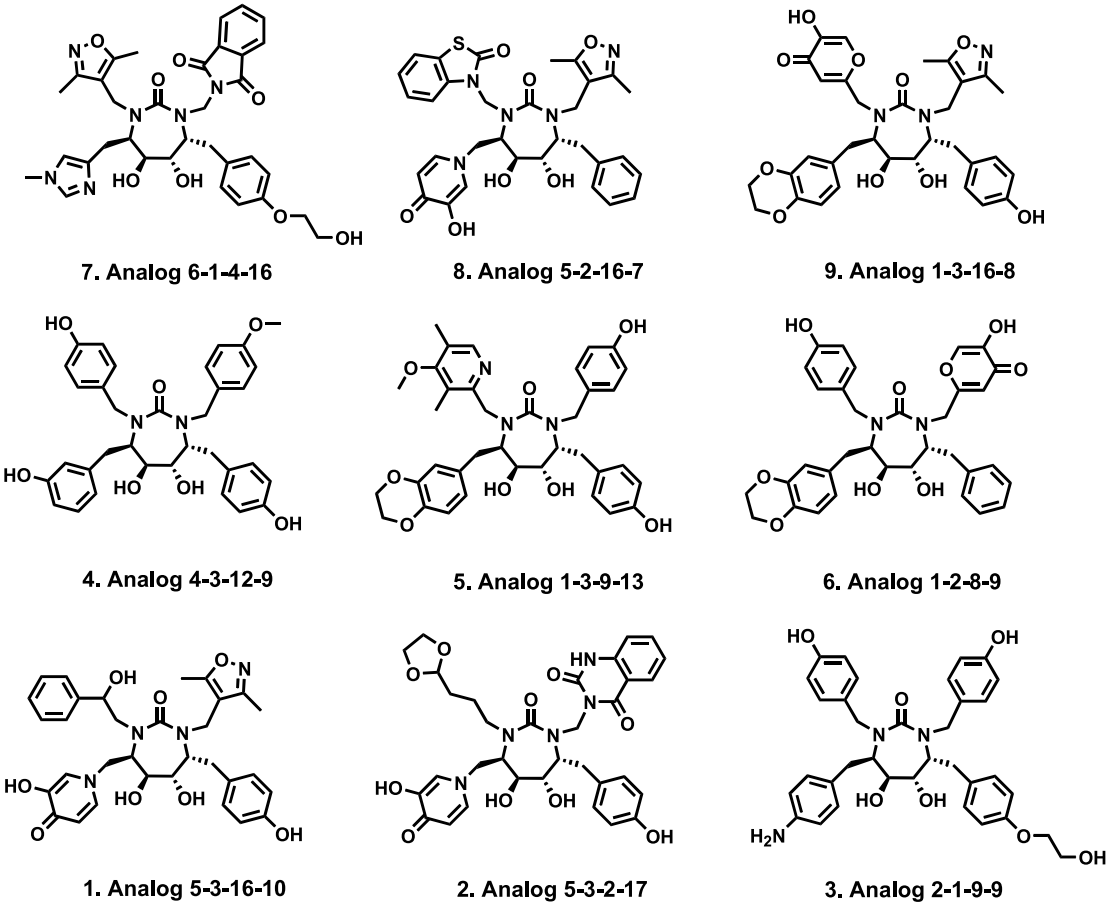


Figure 4. Chemical structures of nine best cyclic urea lead compounds obtained by structure-based library focusing with predicted K_i^*s against the wild type PR of HIV-1 in the low picomolar and subpicomolar range.

Table 3. List of the nine most potent cyclic urea inhibitors against the PR of HIV-1 (Fig. 4) generated via structure-based library focusing and identified through in silico screening

Cyclic urea analog	Docking score[a]	LUDI score[b]	HB_{score}	VDW_{score}	Predicted K_i^*[c] (pM)	M_r[d] (Da)	log $P_{o/w}$[e]	log S_w[f]	BIP_{caco}[g]	#Metab[h]
Analog 5-3-16-10	147.2	801	470	781	0.5	605	2.5	−4.5	2.7	12
Analog 5-3-2-17	−52.3	823	490	834	0.6	664	1.7	−3.7	4.3	9
Analog 2-1-9-9	−222.4	852	479	849	0.8	614	3.5	−5.0	2.0	11
Analog 4-3-12-9	146.8	951	491	885	0.9	584	4.4	−5.4	17.8	10
Analog 1-3-9-13	186.5	1021	508	938	1.1	656	4.9	−5.6	108.4	12
Analog 1-2-8-9	203.1	824	435	814	1.2	614	3.4	−4.7	19.9	8
Analog 6-1-4-16	−442.1	633	372	737	1.7	658	1.4	−2.1	42.0	9
Analog 5-2-16-7	−454.4	634	346	713	2.1	632	2.1	−3.6	9.0	10
Analog 1-3-16-8	−8.3	732	379	778	2.2	634	2.3	−4.2	5.3	10

[a] Docking score was derived by Monte Carlo assisted fitting of flexible ligand to rigid PR receptor-relaxed crystal structure of PR-XV-638[23], Protein Data Bank[37] entry 1QBR.

[b] Score of original LUDI ligand scoring function[26] that predicts the binding affinity of analogs to PR based on the computed ligand–receptor hydrogen bond energies (HB_{score}) and lipophilic interaction (VDW_{score}).

[c] Predicted HIV PR inhibition constant of designed cyclic urea inhibitors computed by LUDI scoring function Eq. 1, parameterized via QSAR analysis of a training set of potent PR inhibitors (Table 2, Fig. 3).

[d] Relative molecular mass (range for 95% of known drugs: from 130 to 725 Da).[39]

[e] log $P_{o/w}$ calculated partitioning coefficient in the system *n*-octanol/water (range for 95% of known drugs: from −2.0 to 6.0).[39]

[f] Predicted aqueous solubility, S_w in mol/dm^3 (range for 95% of known drugs: from −6.0 to 0.5).[38,39]

[g] Predicted Caco-2 cell membrane permeability using the Boehringer–Ingelheim scale, BIP_{caco} in nm/s (< 5 nm/s means low permeability, and more than 100 nm/s predicts high membrane permeability).[39]

[h] Predicted number of likely metabolic reactions.[39]

(No. Met. [39]) of the nine leads (Table 3) cover a relatively wide range of these quantities, thus opening the possibility to select and develop orally bioavailable antiviral agents with improved pharmacokinetic properties. The molecular mass of the leads is slightly over 600 Da and thus exceeds somewhat the threshold recommended by Lipinski et al.[32] However, it remains within the limits of potent experimental cyclic urea PR inhibitors,[33] and for four analogs, it also meets the molecular mass threshold for good oral bioavailability of cyclic ureas: $M_r < 620$ Da.[7]

2.6. Combinatorial subset selection

We have analyzed the composition of 90 best cyclic urea analogs with the highest predicted PR inhibition potencies ($K_i^* < 10$ pM) in terms of the frequency of occurrence of individual R groups in the positions R_1–R_4 within this set (Fig. 6). The fragments that displayed their frequency of occurrence above the average were selected to constitute a highly focused combinatorial subset with high predicted inhibition activities toward the PR of HIV-1 (Fig. 7). The size of this restricted library subset was reduced up to:

$$3(AA_1) \times 2(ALD_2) \times 4(SAB_3) \times 3(SAB_4) = 72 \text{ cyclic ureas.}$$

If we consider nonsymmetrical as well as symmetrical cyclic ureas, we may extend the combinatorial subset by combining the R_1, R_2 and R_3, and R_4 substituents to obtain slightly larger subset of the size:

$$5(AA_1) \times 5(ALD_2) \times 6(SAB_3) \times 6(SAB_4) = 900 \text{ cyclic ureas,}$$

keeping in mind that amino acids and aldehydes with identical 'side chains' were present in the initial pool of 850 AA_1 and 735 ALD_2.[19]

3. Conclusions

The presented study yielded a small highly focused virtual combinatorial subset of fully nonsymmetrical cyclic ureas, which contains potential lead compounds with high predicted inhibitory potencies against the wildtype form of PR of HIV-1. The predicted inhibition constants of the new leads are up to 2 orders of magnitude lower than the K_is of the training set. The subset contains potent cyclic urea PR inhibitors endowed with a wide range of aqueous solubility, lipophilicity, number of hydrogen-bonding groups, and cell membrane permeability rate that may allow discovery of a potent orally administrable antiviral drug with favorable pharmacokinetic properties.

The study can help to draw the attention of synthetic chemists working on the preparation of a next generation of nonsymmetrical cyclic ureas to a particular subset of the chemical space that is predicted to contain compounds with high HIV-1 PR inhibition potencies as well as favorable pharmacokinetic properties.

4. Methods

4.1. Virtual library enumeration

The library of cyclic urea analogs was generated by attaching the side chains of selected reagents available in the commercial databases of suppliers of chemicals[19] to the central core by means of Combi-Chem module of

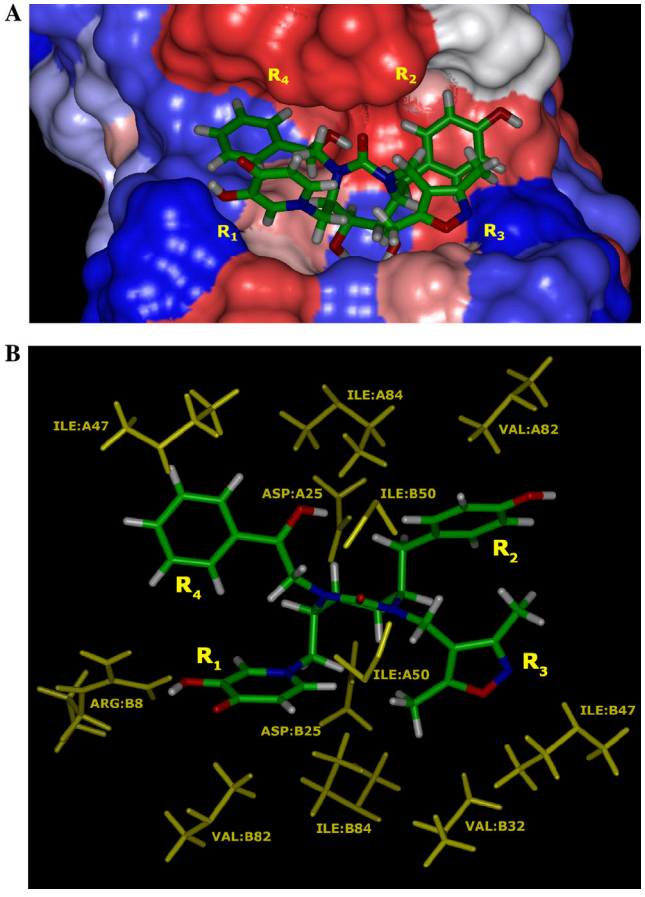


Figure 5. (A) 3D-structure of the most promising cyclic urea inhibitor (Analog 5-3-16-10) in stick representation positioned at the catalytic site of wild type PR modelled with Insight-II.[34] Connolly surface: an axial crossection of the binding pocket and catalytic site is coloured according to the Engleman–Steitz hydrophobicity scale[35] for amino acids: blue—hydrophilic, white-intermediate, red—hydrophobic resides. (B) The residues of the catalytic site of the HIV-1 PR that strongly interact with the bound Analog 5-3-16-10 are shown in stick representation (the –OH groups interacting with the catalytic Asp residues A25 and B25 are hidden underneath the N–C and C–C bonds of the cyclic urea ring in the 'top' view of the analog).

the Cerius^2.[20] The assembled analogs were energy minimized using molecular mechanics and cff91 force field[21], Rappé and Goddard equilibrated charges,[22] and smart minimizer with high convergence criteria (energy difference of 10^{-4} kcal mol^{-1}, rms displacement of 10^{-5} Å).

4.2. Fragment- and analog-based focusing

Descriptors that characterize molecular shape, size, polarity, hydrogen bonding, lipophilicity, solubility, and conformational flexibility of known HIV PR inhibitors and their building blocks were computed using the Cerius^2 program.[20] The optimum ranges of the properties were defined in terms of upper and lower bounds and average values by computing and analyzing the descriptors for 18 known potent HIV PR inhibitors and their building blocks (Table 1) for their bound conformations.[27] Penalty scores were assigned to fragments whose descriptor values lie outside the optimal ranges to filter out those with different properties. The diversity between the fragments was evaluated via the distance-based MaxMin function and the topological indices of Balaban, Hosoya, Wiener, and Zagreb.[25]

4.3. Structure-based focusing

X-ray structure of PR from the complex PR–XV-638 obtained from the Protein Data Bank[37] was used as the receptor model for docking of the generated cyclic urea analogs. This complex contains one of the largest cyclic urea inhibitors, the XV-638 of DuPont Merck,[23] a tightly binding inhibitor that occupies S_2, S_1, $S_{1'}$, and $S_{2'}$ specificity pockets of the PR-binding site.[24] Shape and size of the receptor binding site was defined with the help of the bound ligand (XV-638) mapped onto an energy grid with a resolution of 0.25 Å, which was enlarged by 3 Å. Then the docking of flexible analogs to the binding site was carried out by generating conformers of each cyclic urea analog by randomizing the dihedral angles (10^4 Monte Carlo steps) and fitting them to the site model by comparing principal moments of inertia of the site and the cyclic urea analog.[36] The docking score (ligand–receptor nonbonding interaction energy) was computed via molecular mechanics and cff91 force field[21] for each conformer using the grid representation of the rigid protein receptor, 15 Å cutoff distance on the nonbonding interactions and a dielectric constant of 4. The 10 best-fitting conformers were then energy minimized at the PR-binding site and clustered into five conformational families according to their mutual rms deviations using the Jarvis–Patrick method.[25]

4.4. In silico screening

The best member of each cluster which displayed the highest docking score was then selected for virtual screening using a LUDI-type scoring function[26] that can predict a binding (inhibition) constant of a ligand–receptor complex by combining hydrogen bonding (HB_{score}) and lipophilic (VDW_{score}) interactions between the ligands and receptor as: $pK_i = a \cdot HB_{score} + b \cdot VDW_{score} + c$. The parameters of this target-specific scoring function a, b, and c were obtained beforehand by linear regression analysis of PR binding of a QSAR training set of 12 known symmetrical and $P_2/P_{2'}$ non-symmetrical cyclic urea inhibitors (Table 2) with experimentally determined inhibition constants toward the PR of HIV-1. The training set of inhibitors underwent the same docking procedure to the binding site of the receptor model with subsequent calculation of HB_{score} and VDW_{score} descriptors. The predictive ability of the derived scoring function was validated on a set of five symmetrical cyclic ureas with known inhibition constants, which were not included into the training set.

4.5. ADME properties prediction

Partitioning coefficient of the cyclic urea analogs in the system *n*-octanol/water (log $P_{o/w}$) as a simple parameter describing molecular lipophilicity was calculated by the method of additive fragment constants.[39] Aqueous solubility (log S_w), Caco-2 cell membrane permeability in the Boehringer–Ingelheim scale (BIP_{caco}), and number of

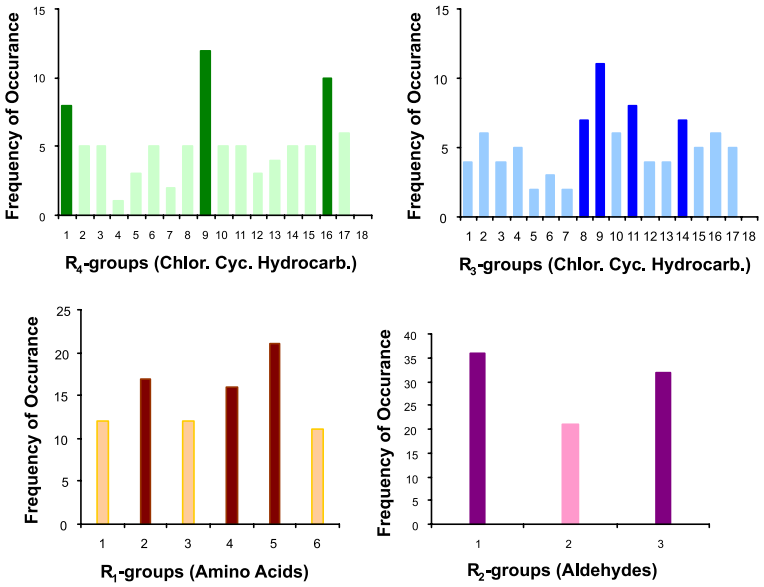


Figure 6. Histograms showing frequency of occurrence of individual R-groups (Fig. 1) in 90 cyclic ureas with the highest predicted PR inhibitory potencies ($pK_i^* > -1.0$). For R-group numbers see Figure 2.

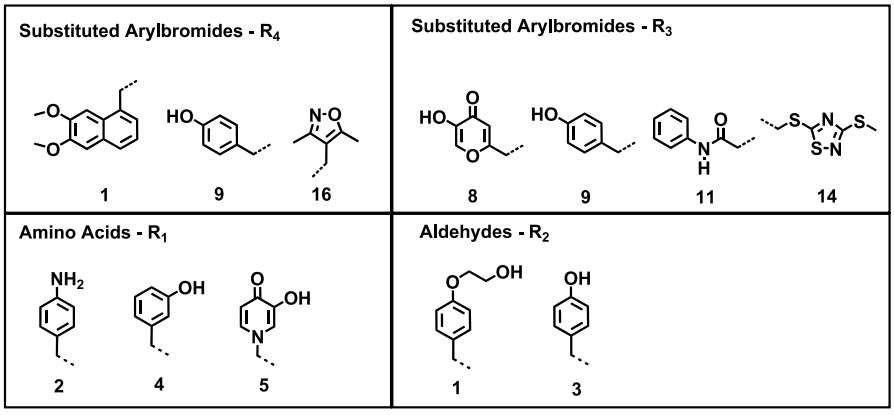


Figure 7. Subset of amino acids (AA), aldehydes (ALD), and substituted arylbromides (SAB) occurring with the highest frequency among the R-groups of 90 cyclic ureas with the highest predicted PR inhibitory potencies.

likely metabolic reactions (No. Met.) were calculated using the QikProp program of Schrödinger.[38,39] These simple parameters can serve as useful guidelines in prioritization of virtual hits for the synthesis.

Acknowledgment

AIDS research grant of the Istituto Superiore di Sanità, Rome, Italy, number 40 D 59 is gratefully acknowledged.

References and notes

1. Katz, R. A.; Skalka, A. M. *Annu. Rev. Biochem.* **1994**, *63*, 133–173.
2. De Clercq, E. *Biochem. Biophys. Acta* **2002**, *1587*, 258–275.
3. Boden, D.; Markowitz, M. *Antimicrob. Agents Chemother.* **1998**, *42*, 2775–2783.
4. Kantor, R.; Katzenstein, D. *AIDS Rev.* **2003**, *5*, 25–35.
5. Lam, P. Y.; Jadhav, P. K.; Eyermann, C. J.; Hodge, C. N.; Ru, Y.; Bacheler, L. T.; Meek, J. L.; Otto, M. J.; Rayner, M. M.; Wong, Y. N.; Chang, C.-H.; Weber, P. C.; Jackson, D. A.; Sharpe, T. R.; Erickson-Viitanen, S. *Science* **1994**, *263*, 380–384.
6. Hodge, C. N.; Aldrich, P. E.; Bacheler, L. T.; Chang, C. H.; Eyermann, C. J.; Garber, S.; Grubb, M.; Jackson, D. A.; Jadhav, P. K.; Korant, B.; Lam, P. Y.; Maurin, M. B.; Meek, J. L.; Otto, M. J.; Rayner, M. M.; Reid, C.; Sharpe, T. R.; Shum, L.; Winslow, D. L.; Erickson-Viitanen, S. *Chem. Biol.* **1996**, *3*, 301–314.
7. Rodgers, J. D.; Lam, P. Y.; Johnson, B. L.; Wang, H.; Li, R.; Ru, Y.; Ko, S. S.; Seitz, S. P.; Trainor, G. L.; Anderson, P. S.; Klabe, R. M.; Bacheler, L. T.; Cordova, B.; Garber, S.; Reid, C.; Wright, M. R.; Chang, C. H.; Erickson-Viitanen, S. *Chem. Biol.* **1998**, *5*, 597–608.
8. De Lucca, G. V.; Jadhav, P. K.; Waltermire, R. E.; Aungst, B. J.; Erickson-Viitanen, S.; Lam, P. Y. *Pharm. Biotechnol.* **1998**, *11*, 257–284.
9. Lam, P. Y.; Ru, Y.; Jadhav, P. K.; Aldrich, P. E.; DeLucca, G. V.; Eyermann, C. J.; Chang, C. H.; Emmett,

G.; Holler, E. R.; Daneker, W. F.; Li, L.; Confalone, P. N.; McHugh, R. J.; Han, Q.; Li, R.; Markwalder, J. A.; Seitz, S. P.; Sharpe, T. R.; Bacheler, L. T.; Rayner, M. M.; Klabe, R. M.; Shum, L.; Winslow, D. L.; Kornhauser, D. M.; Hodge, C. N. *J. Med. Chem.* **1996**, *39*, 3514–3525.
10. Wilkerson, W. W.; Dax, S.; Cheatham, W. W. *J. Med. Chem.* **1997**, *40*, 4079–4088.
11. Benedetti, F.; Miertus, S.; Norbedo, S.; Tossi, A.; Zlatoidzky, P. *J. Org. Chem.* **1997**, *62*, 9348–9353.
12. Patel, M.; Kaltenbach, R. F.; Nugiel, D. A.; McHugh, R. J.; Jadhav, P. K.; Bacheler, L. T.; Cordova, B. C.; Klabe, R. M.; Erickson-Viitanen, S.; Garber, S.; Reid, C.; Seitz, S. P. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 1077–1082.
13. Kotamarthi, B.; Bonin, I.; Benedetti, F.; Miertus, S. *Biochem. Biophys. Res. Commun.* **2000**, *268*, 384–389.
14. Schneider, G. *Curr. Med. Chem.* **2002**, *23*, 2095–2101.
15. Hobbs, D. W.; Guo, T. *J. Recept. Signal Transduct. Res.* **2001**, *21*, 311–356.
16. Martin, Y. C. *J. Comb. Chem.* **2001**, *3*, 231–250.
17. Böhm, H. J.; Stahl, M. *Curr. Opin. Chem. Biol.* **2000**, *4*, 283–286.
18. Beavers, M. P.; Chen, Z. *J. Mol. Graph. Model.* **2002**, *20*, 463–468.
19. Available Chemicals Directory version 95.1, MDL Information Systems, San Leandro, CA. http://cds3.dl.ac.uk/cds/cds.html.
20. Cerius^2 Life Sciences, version 4.5, Accelrys, San Diego, CA, 2000.
21. Maple, J. R.; Hwang, M. J.; Stockfish, T. P.; Dinur, U.; Waldman, M.; Ewing, C. S.; Hagler, A. T. *J. Comput. Chem.* **1994**, *15*, 162–182.
22. Rappé, A. K.; Goddard, W. A., III *J. Phys. Chem.* **1991**, *95*, 3358–3363.
23. Jadhav, P. K.; Ala, P.; Woerner, F. J.; Chang, C. H.; Garber, S. S.; Anton, E. D.; Bacheler, L. T. *J. Med. Chem.* **1997**, *40*, 181–191, Protein Data Bank entry 1QBR.
24. Schechter, I.; Berger, A. *Biochem. Biophys. Res. Commun.* **1967**, *27*, 157–162.
25. Willett, P. Similarity-Searching and Clustering Algorithms for Processing Databases of Two-dimensional and Three-dimensional Chemical Structures. In *Molecular Similarity in Drug Design*; Dean, P. M., Ed.; Chapman and Hall: Glasgow, 1994, pp 110–137.
26. Böhm, H. J. *J. Comput. Aided Mol. Des.* **1994**, *8*, 243–256.
27. Frecer, V.; Burello, E.; Miertus, S. Design of Protease Inhibitors by Computer-Assisted Combinatorial Chemistry. In *Combinatorial Chemistry and Technology*; Miertus, S., Fassina, G., Eds., 2nd ed.; M. Dekker: New York, 2004, pp 55–74, Chapter 4.
28. Ala, P. J.; DeLoskey, R. J.; Huston, E. E.; Jadhav, P. K.; Lam, P. Y.; Eyermann, C. J.; Hodge, C. N.; Schadt, M. C.; Lewandowski, F. A.; Weber, P. C.; McCabe, D. D.; Duke, J. L.; Chang, C. H. *J. Biol. Chem.* **1998**, *273*, 12325–12331.
29. Patel, M.; Bacheler, L. T.; Rayner, M. M.; Cordova, B. C.; Klabe, R. M.; Erickson-Viitanen, S.; Seitz, S. P. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 823–828.
30. Kaltenbach, R. F., III; Patel, M.; Waltermire, R. E.; Harris, G. D.; Stone, B. R. P.; Klabe, R. M.; Garber, S.; Bacheler, R. T.; Cordova, B. C.; Logue, K.; Wright, M. R.; Erickson-Viitanen, S.; Trainor, G. L. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 605–608.
31. Sham, H. L.; Zhao, C.; Stewart, K. D.; Betebenner, D. A.; Lin, S.; Park, C. H.; Kong, X. P.; Rosenbrook, W., Jr.; Herrin, T.; Madigan, D.; Vasavanonda, S.; Lyons, N.; Molla, A.; Saldivar, A.; Marsh, K. C.; McDonald, E.; Wideburg, N. E.; Denissen, J. F.; Robins, T.; Kempf, D. J.; Plattner, J. J.; Norbeck, D. W. *J. Med. Chem.* **1996**, *39*, 392–397.
32. Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Deliv. Rev.* **2001**, *46*, 3–26.
33. Wilkerson, W. W.; Akamike, E.; Cheatham, W. W.; Hollis, A. Y.; Collins, R. D.; DeLucca, I.; Lam, P. Y.; Ru, Y. *J. Med. Chem.* **1996**, *39*, 4299–4312.
34. Insight-II version 2000, Accelrys, San Diego, CA, 2000.
35. Engelman, D. M.; Steitz, T. A. *Cell* **1981**, *23*, 411–422.
36. Peters, K. P.; Fauck, J.; Frommel, C. *J. Mol. Biol.* **1996**, *256*, 201–213.
37. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucl. Acids Res.* **2000**, *28*, 235–242.
38. Jorgensen, W. L.; Duffy, E. M. *Adv. Drug Deliv. Rev.* **2002**, *54*, 355–366.
39. QikProp, version 2.0, Schrödinger, New York, NY, 2002.