Journal of Molecular Graphics and Modelling 27 (2008) 376–387

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Design of peptidomimetic inhibitors of aspartic protease of HIV-1 containing –PheΨPro– core and displaying favourable ADME-related properties

Vladimir Frecer [a,*], Federico Berti [b], Fabio Benedetti [b], Stanislav Miertus [c]

[a] Cancer Research Institute, Slovak Academy of Sciences, Vlarska 7, Bratislava SK-83391, Slovakia
[b] Department of Chemical Sciences, University of Trieste, Trieste I-34127, Italy
[c] International Centre for Science and High Technology, UNIDO, AREA Science Park, Trieste I-34012, Italy

ARTICLE INFO

Article history:
Received 13 March 2008
Received in revised form 13 June 2008
Accepted 17 June 2008
Available online 26 June 2008

Keywords:
Aspartic protease of HIV-1
Peptidomimetic inhibitors
Transition-state-isostere core
Structure-based molecular design
Molecular modelling
QSAR
ADME-related properties prediction

ABSTRACT

Aspartic protease (PR) of HIV-1 virus represents a valid therapeutic target for the design of antiviral agents suitable for treatment of AIDS. We have designed peptidomimetic PR inhibitors containing a novel dihydroxyethylenediamine –Phe-Ψ[CHOH–CHOH]-Pro– core using molecular modelling approach that predicts the inhibitory potencies (IC_{50}^{pre}) in terms of computed relative enzyme–inhibitor complexation Gibbs free energies ($\Delta\Delta G_{comp}$). The modelling approach considers not only the enzyme–inhibitor interactions, but includes also the solvent and entropic effects affecting the enzyme inhibition. The objectives of this study were to optimize the number and type of flanking residues that occupy the S_3, S_2 and $S_{2'}$, $S_{3'}$ positions in the PR binding pocket and to select potent lead candidates, which display also favourable ADME-related properties. The structure-based design was combined with a synthetic strategy used to prepare a training set of 10 analogues sharing the –PheΨPro– core. This strategy couples stereochemical control with full flexibility in the choice of the flanking residues and *in vitro* activity assays. A QSAR model correlating calculated $\Delta\Delta G_{comp}$ with the measured IC_{50}^{exp} values for the training set was prepared and confirmed that our computational approach can serve for reliable prediction of PR inhibitory potencies of peptidomimetics. The appropriate choice of the flanking residues allowed us to design virtual lead compounds, such as **FP14**, **FP23** and **FP76**, with reduced molecular weight, predicted inhibitory potencies in the picomolar range, promising ADME profiles and a potential to escape drug resistance due to favourable interactions with the PR backbone.

1. Introduction

The aspartic protease (PR) of HIV-1, responsible for cleaving the viral polyprotein precursor into structural proteins and enzymes, plays an essential role in the viral replication and maturation. Inhibition of the PR blocks the infectivity of the HIV virus. The PR has thus become the target of numerous efforts to design antiviral therapeutic agents suitable for the treatment of AIDS [1,2]. The HIV-1 PR is a homodimer consisting of two monomers of 99 residues each related by a pseudo-two-fold symmetry. The catalytic site is formed by two triads Asp:25, Thr:26, Gly:27 with the catalytic residues Asp:A25 and Asp:B25 positioned at the centre of the substrate binding cleft, which is covered by a pair of flexible protease flaps that assist the substrate entry and binding [3]. The nomenclature of Schechter and Berger for proteases [4] designates the specificity pockets of the PR's active site as $-S_3-S_2-S_1\ldots S_{1'}-S_{2'}-S_{3'}-$, etc. with the catalytic centre positioned between the pockets S_1 and $S_{1'}$ and the C-terminal portion of the substrate occupying the prime site. The HIV-1 PR can recognize Phe-Pro and Tyr-Pro sequences as the specific cleavage sites [3].

A number of substrate or transition-state-mimetic PR inhibitors have reached the stage of clinical trials and currently more than ten inhibitors have been approved by the FDA for anti-HIV therapy (http://www.fda.gov/oashi/aids/virals.html). These specific inhibitors cause significant initial lowering of the viral load in tissues and a general improvement in the clinical state of AIDS patients. However, PR-targeted therapy leads to rapid selection of drug-resistant mutants and development of viral resistance. Thus a continuing need exists for the discovery of new more potent orally bioavailable inhibitors with simpler dosing regimens active against a wider spectrum of PR mutant forms [5]. The mutated residues in the drug-resistant PR forms isolated from patients treated with PR inhibitors appear at positions close to the active site as well as in more distant parts of the protein [6]. The active site mutations decrease the substrate/inhibitor binding affinity while the non-active site mutations increase the catalytic

* Corresponding author. Tel.: +421 2 59327 255; fax: +421 2 59327 250.
E-mail address: vladimir.frecer@savba.sk (V. Frecer).

1093-3263/$ – see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmgm.2008.06.006

efficiency and partially compensate for the decreased substrate binding [7].

Our laboratories have developed an integrated method for the design and synthesis of peptidomimetic C_2-symmetric inhibitors based on a modular assembly of appropriate flanking residues onto non-hydrolysable mono and dihydroxyethylenediamine cores [8–13]. The synthetic route used by our group is based on available starting materials and allows full flexibility and stereoselective synthesis of the central hydroxylated ethylenediamine core with the possibility to vary independently the P_1 and $P_{1'}$ residues in the central unit as well as the number and type of P_n, $P_{n'}$ residues flanking the central core on either side [7,14–16]. Using this methodology, we have prepared a series of peptidomimetic inhibitors containing 4, 5 or 6 residues and ranging from C_2-symmetrical to non-symmetrical ones composed of building blocks of natural and non-natural amino acids and carboxylic acids [8–13].

The choice of the central core and of the flanking residues is guided by structure-based design approach that uses crystal structures of PR:inhibitor complexes. This approach employs molecular mechanics, Monte Carlo method, Poisson–Boltzmann equation and normal mode analysis to calculate complexation Gibbs free energies of binding of the candidate inhibitor to the PR and utilizes a QSAR model to predict the inhibitory potencies.

General strategy to combat the viral resistance problem is to design inhibitors that interact with the backbone atoms of the active site rather than the residue side chains only. Alternatively, more flexible molecules can readily adapt to structural and conformational changes occurring in the emerging resistant PR forms. This strategy was successfully applied by Mitsuya and co-workers during the development of the PR inhibitor TMC114 (analogue of amprenavir), which is active against multi-PR-inhibitor-resistant HIV-1 variants [17]. In addition, more hydrophobic inhibitors driven to form the complexes with receptor predominantly by the effect of the solvent should be also less sensitive to the receptor mutations conferring drug resistance.

The objective of this study was to rationally design a series of PR inhibitors containing novel –Phe-Ψ[CHOH–CHOH]-Pro– transition-state-isostere core and optimize the number and type of the flanking residues to reduce the molecular weight of the inhibitors while preserving the PR inhibitory potency and improving the ADME-related properties. Previous studies on inhibitors containing allophenylnorstatin core and other similar transition-state mimetics (Fig. 1) namely indicated that this core can generate potent pseudopeptidic HIV PR inhibitors [18–20]. Another objective was to introduce more flexible flanking residues and residues capable of interacting with the PR backbone in an attempt to design inhibitors capable of preserving high binding affinity also to the emerging mutant resistant PR forms.

The design started by a QSAR analysis of inhibitory potencies (IC_{50}^{exp}) of 10 synthesized and tested compounds sharing the same –PheΨPro– core (Fig. 2), continued with structure-based optimization of building blocks attached to either side of the core and concluded with selection of lead candidates. Calculated ADME-related descriptors as predictors of pharmacokinetic properties, were employed to prioritize the selection of potent and bioavailable leads suitable for further development.

2. Methods

2.1. Calculation of binding affinity

The IC_{50} (concentration of a competitive tight binding inhibitor that causes 50% reduction of the rate of catalytic substrate

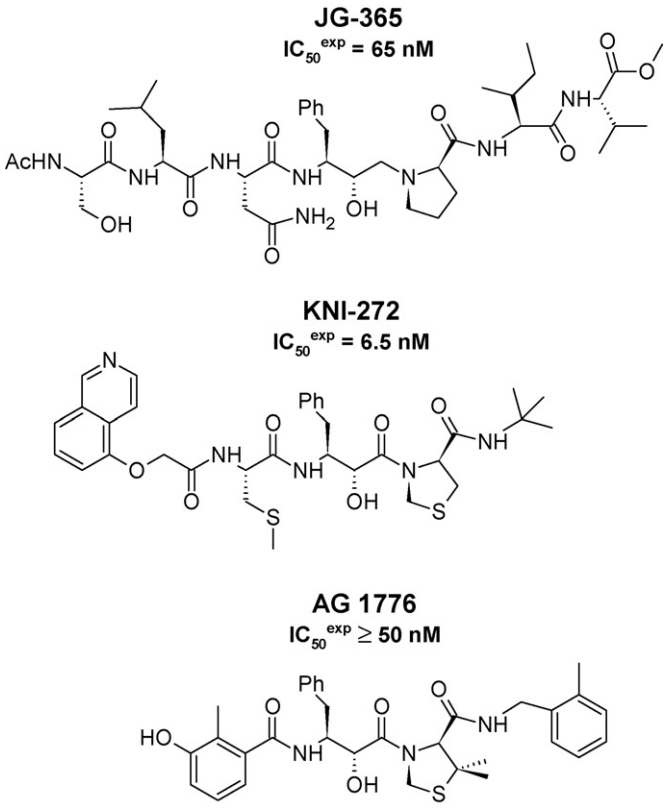

Fig. 1. Chemical structures of HIV-1 PR inhibitors JG-365 [18], KNI-272 [19] and AG 1776 [20] containing variants of the studied –PheΨPro– transition-state-isostere cores.

conversion) of a reversible inhibitor depends on the enzyme inhibition constant (K_i) as: $IC_{50}^{exp} = K_i + [S]\cdot K_i/K_M + [E]/2$, where $[S]$ and $[E]$ are the substrate and active enzyme concentrations and K_M the Michaelis constant [21]. The IC_{50} value can thus be predicted from the standard Gibbs free energy change of the enzyme:inhibitor complex formation, $\Delta G_{comp} = -RT \ln K_i$, assuming the following equilibrium in solution:

$$\{PR\}_{aq} + \{I\}_{aq} \Leftrightarrow \{PR:I\}_{aq} \quad (1)$$

where $\{PR\}_{aq}$, $\{I\}_{aq}$ and $\{PR:I\}_{aq}$ indicate solvated enzyme (PR), inhibitor and their complex. The standard Gibbs free energy change of the reaction (1) can be derived by molecular simulations of the complex and the free reactants:

$$\Delta G_{comp} = G\{PR:I\} - \{PR\}\{I\} \quad (2)$$

In this work we approximate the exact values of standard Gibbs free energies for larger systems such as enzyme:inhibitor complexes by the expression [9,11,13]:

$$G\{PR:I\} \cong [E_{MM}\{PR:I\} + RT - TS_{trv}\{PR:I\}] + G_{sol}\{PR:I\} \quad (3)$$

where $E_{MM}\{PR:I\}$ stands for the molecular mechanics total energy of the complex (including bonding and non-bonding contributions), $G_{sol}\{PR:I\}$ is the solvation Gibbs free energy and $TS_{trv}\{PR:I\}$ the entropic term:

$$TS_{trv}\{PR:I\} = TS_{trans}\{PR:I\} + TS_{rot}\{PR:I\} + TS_{vib}\{PR:I\} \quad (4)$$

i.e. a sum of contributions arising from translational, rotational and vibrational motions. When assuming that the translational and rotational terms for the free enzyme and for the enzyme:inhibitor

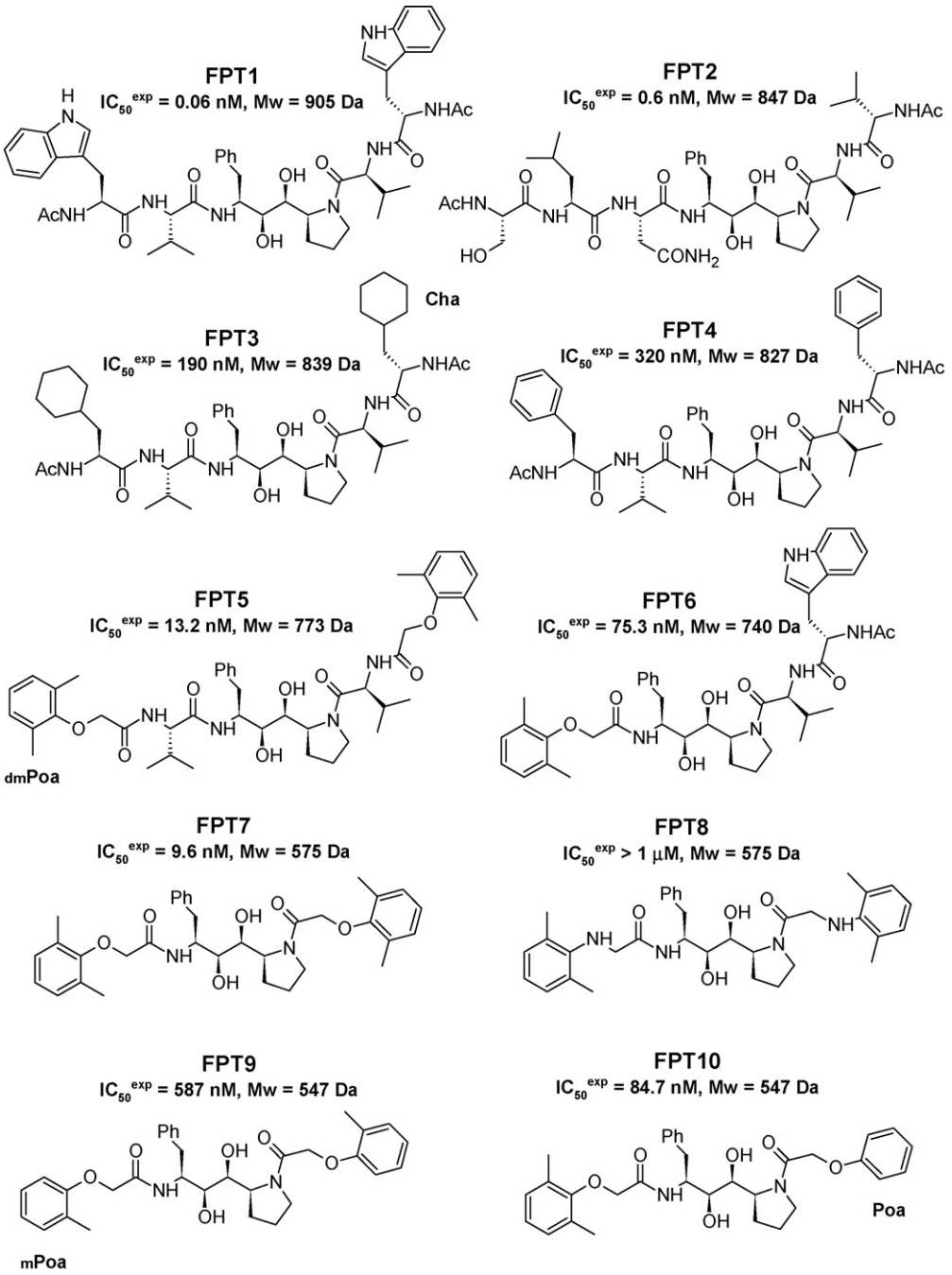

Fig. 2. Chemical structures of the synthesized training set of HIV-1 PR inhibitors **FPT1–FPT10** sharing the common –Phe-Ψ[CHOH–CHOH]-Pro– (*s*,*s*,*s*,*s*) stereoisomer core and their experimental inhibitory potencies IC_{50}^{exp} and molecular weight.

complex are approximately equal ($TS_{trans}\{PR{:}I\} \approx TS_{trans}\{PR\}$ and $TS_{rot}\{PR{:}I\} \approx TS_{rot}\{PR\}$), we obtain

$$\begin{aligned}\Delta G_{comp} \cong\ & [E_{MM}\{PR:I\} - E_{MM}\{PR\} - E_{MM}\{I\}] + [G_{sol}\{PR:I\} \\ & - G_{sol}\{PR\} - G_{sol}\{I\}] + TS_{trans}\{I\} + TS_{rot}\{I\} \\ & - [TS_{vib}\{PR:I\} - TS_{vib}\{PR\} - TS_{vib}\{I\}] \\ =\ & \Delta H_{MM} + TS_{trans}\{I\} + TS_{rot}\{I\} - T\Delta S_{vib} + \Delta G_{sol} \end{aligned} \quad (5)$$

where $TS_{trans}\{I\}$ and $TS_{rot}\{I\}$ describe the translational and rotational entropy terms of the free inhibitor and $T\Delta S_{vib}$ represents the vibrational entropy change upon the complex formation.

Comparison between different inhibitors was done via relative changes in the complexation Gibbs free energy with respect to a reference inhibitor, I_{ref}, assuming ideal gas behaviour for the rotational and translational motions of the inhibitors:

$$\begin{aligned}\Delta\Delta G_{comp} &= \Delta G_{comp}(I) - \Delta G_{comp}(I_{ref}) \\ &= \Delta\Delta H_{MM} - T\Delta\Delta S_{vib} + \Delta\Delta G_{sol} \end{aligned} \quad (6)$$

The evaluation of relative changes is preferable as it is expected to lead to partial cancellation of errors caused by the approximate

nature of the molecular mechanics method as well as solvent and entropic effects description.

2.2. Model building

Molecular models of the PR:inhibitor complexes, free PR and inhibitors were prepared from a high-resolution crystal structure of the reference complex containing substrate-based hydroxyethylamine inhibitor JG-365 [22] (Fig. 1) (Protein Data Bank [23] entry code 7HVP) using *Insight-II* molecular modelling program [24]. The wild type structures of the PR and PR:inhibitor complexes were considered to be at pH of 7 with neutral N- and C-terminal residues, all protonizable and ionisable residues were charged, except for the active site Asp:A25 and Asp:B25, which shared one proton [25]. A crystallographic water molecule located in the proximity of the non-scissile bond, which forms hydrogen bonds with the enzyme flaps, was also included into the model. The inhibitors were built into the crystal structure of the PR:inhibitor complex by replacing the appropriate flanking residues. An exhaustive conformational search over all rotatable bonds of the replacing residues coupled with energy minimization of inhibitor and the PR active site was employed to identify the low-energy bound conformations of the modified inhibitor. The resulting low-energy structure of the PR:inhibitor complex was then carefully refined by minimization of the whole complex.

2.3. Molecular mechanics

Refinement of the models was carried out by molecular mechanics (MM) using all-atom representation and the class II consistent force field CFF91 [26–28]. A dielectric constant of 4 was used for all MM calculations in order to take into account dielectric shielding effects in proteins. Minimization of the PR:inhibitor complexes and of the free enzyme was carried out by relaxing the structures gradually, starting with the residue side chains and concluding with the relaxation of all atoms including the protein backbone. In the geometry optimization, a sufficient number of conjugate gradient iterative cycles, with a convergence criterion set to the average gradient of 0.01 kcal mol^{-1} $Å^{-1}$, has been used.

2.4. Conformational search

Free inhibitor conformations were derived from their bound conformations in the PR:inhibitor complexes by gradual relaxation to the nearest local energy minimum. Then a Monte Carlo search (with an upper limit of 50,000 iterations) for low-energy conformations over all rotatable bonds except those in the rings, was carried out using *Cerius*^{2} molecular modelling package [29]. Two hundred unique conformations were generated per each inhibitor by randomly varying torsion angles of the last accepted conformer by $\pm 15^\circ$ at 5000 K followed by subsequent energy minimization. During the minimization a dielectric constant $\varepsilon = 80$ was used to account approximately for the dielectric screening effect of hydration upon the generated conformers. The conformer with the lowest total energy was selected and re-minimized at $\varepsilon = 4$.

2.5. Solvation Gibbs free energies

The electrostatic component of solvation Gibbs free energy that incorporates also the effects of ionic strength through the solution of nonlinear Poisson–Boltzmann equation [30,31] was computed by the *DelPhi* software package [24]. The program treats the solvent as a continuous medium of high dielectric constant ($\varepsilon_o = 80$) and the solute as a cavity with low dielectric ($\varepsilon_i = 4$) with boundaries linked to the solute's molecular surface, which encloses the solute's atomic charges. The program uses a finite difference method to solve for the molecular electrostatic potential and reaction field around the solute. *DelPhi* calculations were carried out on a ($235 \times 235 \times 235$) cubic lattice grid for the complexes and free PR and ($75 \times 75 \times 75$) grid for the free inhibitors with full Coulombic boundary conditions. Two subsequent focusing steps led in both cases to a similar final resolution of about 0.3 Å per grid unit at 70% filling of the grid by the solute. Physiological ionic strength of 0.145 mol dm^{-3}, atomic partial charges and radii defined in the CFF91 parameter set [26–28] and a probe sphere radius of 1.4 Å were used. The electrostatic component of the solvation Gibbs free energy was calculated as the reaction field energy [30–34].

2.6. Entropic term

The vibrational entropy change during the inhibitor binding to the PR was calculated by normal mode analysis of the inhibitor vibrations using a simplified method of Fischer et al. [35,36]. In this approach vibrational analyses of the inhibitor bound at the active site of a 'frozen' PR and of the low-energy conformer of the free inhibitor are computed for fully minimized structures using *Discover* package [24] and $T\Delta S_{vib} = TS_{vib}\{I\}_{PR} - TS_{vib}\{I\}$. It has been shown previously that for small and relatively stiff ligands this method gives a good approximation of the vibrational entropy change of the fully flexible system, i.e. including the degrees of freedom of the protein receptor [35,36]. The $TS_{vib}\{I\}$ term accounts for vibrational motions of the free inhibitor and represents an indicator of conformational flexibility of the molecule. Namely, low frequency vibrations, which correspond to collective motions of a number of atoms with larger amplitudes, i.e. conformational changes, contribute most to this term. Relative values of $T\Delta\Delta S_{vib}$ with respect to the reference inhibitor were used to compensate partially for the restricted flexibility of the PR.

2.7. ADME-related properties

Prediction of descriptors related to adsorption, distribution, metabolism and excretion (ADME properties) of inhibitors was carried out by the *QikProp* program [37] based on the method of Jorgensen [38–40]. The program computes pharmaceutically relevant properties such as octanol/water partitioning coefficient, aqueous solubility, brain/blood partition coefficient, Caco-2 cell permeability, serum protein binding, number of likely metabolic reactions, and others. Compliance with the ADME descriptors of clinically used PR inhibitors was assessed with the help of a penalty function. This function combines the tendencies of individual descriptors to fall into the optimum value ranges ($\bar{u}_i \pm u_{SDi}$, $\bar{u}_i$ is the mean value of the descriptor and u_{SDi} denotes its standard deviation) derived from a set of seven clinically used orally bioavailable HIV PR inhibitors (Fig. 3). The penalty function: $P = 1/N_p \sum_i^{N_p}(w_i P_i)$ is composed of individual penalties $P_i = |u_i - \bar{u}_i|/(nu_{SDi})$ of $N_p = 16$ ADME-related descriptors u_i and their statistical weight factors w_i as defined in the *Cerius*^{2} program [29]. Individual penalty maximum was reached at $n = 1$ standard deviation. In the overall penalty function the weight factors were set to $w_i = 1$, however, the Lipinski's rule of five descriptors [41,42] and the predicted activity (IC_{50}^{pre}) had their weight factor increased to 5 and 20, respectively, to bias the penalty function towards highly active orally bioavailable drug-like molecules.

2.8. Synthesis and activity determination

The synthesis of the training set compounds from commercially available reagents has been described elsewhere [8,14–16].

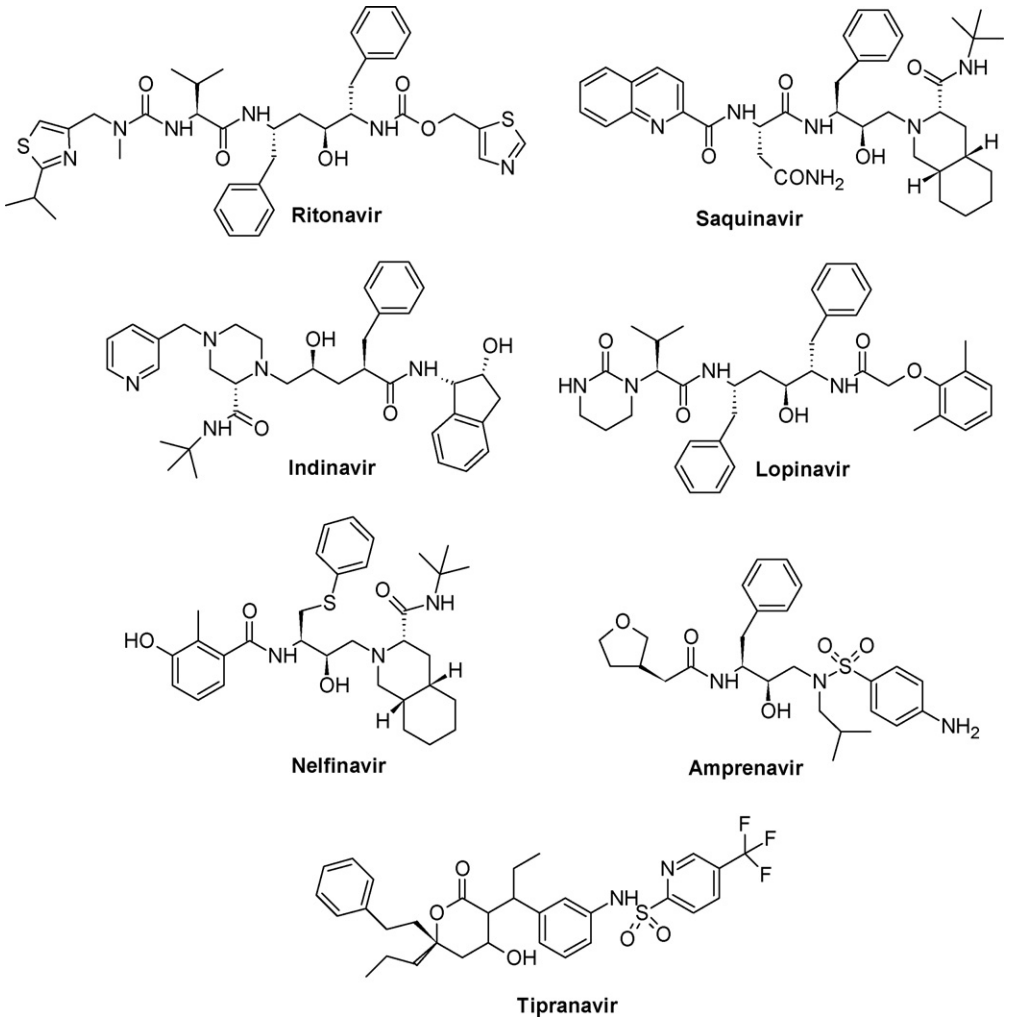

Fig. 3. Chemical structures of selected HIV-1 PR inhibitors approved by FDA (Food and Drug Administration) used for the determination of optimum value ranges of ADME-related descriptors.

Table 1
Training set of synthesized and assayed C_2-symmetric peptidomimetic inhibitors of HIV-1 PR containing –Phe-Ψ[CHOH–CHOH]-Pro– core **FPT1–FPT10**

Inhibitors	Chemical structure, training set of HIV-1 PR inhibitors[a]	M_w[b] (g mol^{-1})	$\Delta\Delta H_{MM}$[c] (kcal mol^{-1})[d]	$\Delta\Delta G_{sol}$[e] (kcal mol^{-1})	$-\Delta\Delta TS_{vib}$[f] (kcal mol^{-1})	$\Delta\Delta G_{comp}$[g] (kcal mol^{-1})	IC_{50}^{exp}[h] (nM)
FPT1	*Ac*-Trp-Val-**PheΨPro**-Val-Trp-*Ac*	905	0.0	0.0	0.0	0.0	0.06
FPT2	*Ac*-Ser-Ile-Asn-**PheΨPro**-Val-Val-*Ac*	847	−2.9	4.6	−1.3	0.4	0.6
FPT3	*Ac*-Cha-Val-**PheΨPro**-Val-Cha-*Ac*	839	3.3	9.4	−1.7	11.0	190
FPT4	*Ac*-Phe-Val-**PheΨPro**-Val-Phe-*Ac*	827	2.0	12.2	−1.4	12.8	320
FPT5	dmPoa-Val-**PheΨPro**-Val-dmPoa	773	4.1	3.3	−1.8	5.6	13.2
FPT6	dmPoa-**PheΨPro**-Val-Trp-*Ac*	740	4.6	4.1	0.5	9.2	75.3
FPT7	dmPoa-**PheΨPro**-dmPoa	575	5.4	−4.2	5.5	6.7	9.6
FPT8	+dmPaa-**PheΨPro**-dmPaa+	575	30.2	−9.4	6.0	26.8	8400
FPT9	mPoa-**PheΨPro**-mPoa	547	23.3	−8.6	7.1	21.8	587
FPT10	dmPoa-**PheΨPro**-Poa	547	18.6	−9.5	7.4	16.5	84.7

The table lists computed contributions to the complexation Gibbs free energy and inhibitory concentrations (IC_{50}^{exp}) determined in the HIV-1 PR inhibition assay.

[a] For the chemical structures of the training set of inhibitors see Fig. 2.

[b] M_w is the molecular mass of the inhibitors.

[c] $\Delta\Delta H_{MM}$ is the relative enthalpic contribution to the Gibbs free energy change related to the intermolecular interactions in the enzyme–inhibitor complex derived by molecular mechanics: $\Delta\Delta H_{MM} = [E_{MM}\{PR{:}I\} - E_{MM}\{I\}] - [E_{MM}\{PR{:}\mathbf{FTP1}\} - E_{MM}\{\mathbf{FPT1}\}]$, **FPT1** is the reference inhibitor.

[d] 1 kcal mol^{-1} = 4.1840 kJ mol^{-1}.

[e] $\Delta\Delta G_{sol}$ is the relative solvation Gibbs free energy contribution to the Gibbs free energy of the enzyme–inhibitor complex formation: $\Delta\Delta G_{sol} = [G_{sol}\{PR{:}I\} - G_{sol}\{I\}] - [G_{sol}\{PR{:}\mathbf{FPT1}\} - G_{sol}\{\mathbf{FPT1}\}]$.

[f] $-\Delta\Delta TS_{vib}$ is the relative entropic contribution of the inhibitor to the Gibbs free energy related to the enzyme–inhibitor complex formation: $\Delta\Delta TS_{vib} = [TS_{vib}\{I\}_{PR} - TS_{vib}\{I\}] - [TS_{vib}\{\mathbf{FPT1}\}_{PR} - TS_{vib}\{\mathbf{FPT1}\}]$.

[g] $\Delta\Delta G_{comp}$ is the relative Gibbs free energy change related to the enzyme–inhibitor complex formation: $\Delta\Delta G_{comp} = \Delta\Delta H_{MM} - \Delta\Delta TS_{vib} + \Delta\Delta G_{sol}$.

[h] IC_{50}^{exp} is the inhibitor concentration that causes 50% decrease of the rate of substrate conversion by the HIV-1 PR measured in the enzyme inhibition assay $IC_{50}^{exp} = K_i\,[S]\cdot(K_i/K_M) + [PR]/2$ (where [*S*] and [PR] are the substrate and active enzyme concentrations and K_M the Michaelis constant) [21] determined in HIV-1 PR inhibition assay [43].

The *in vitro* inhibitory potencies of the compounds (IC_{50}^{exp}) were evaluated by HIV-1 PR inhibition assay reported previously [43].

3. Results and discussion

3.1. Training set of inhibitors

QSAR analysis of the training set of PR inhibitors (Fig. 2) was based on the availability of high-resolution crystal structure of the HIV-1 PR complexed with inhibitor JG-365 [22] (Fig. 1). The JG-365 contains a non-scissile hydroxyethylenediamine amide bond transition-state-isostere similar to the studied C_2-symmetric core formed by the dihydroxyethylenediamine –Phe-Ψ[CHOH–CHOH]-Pro– (*s*,*s*,*s*,*s*) stereoisomer [15]. The PR:inhibitor complexes of the ten training set compounds **FPT1**–**FPT10** were modelled from the experimental structure of the PR:JG-365 complex as described in Section 2. The *in vitro* PR inhibition activities (IC_{50}^{exp}) of ten synthesized training set compounds with the –PheΨPro– core flanked by a variable number and type of residues (Fig. 2) were evaluated in the PR inhibition assay. The determined IC_{50}^{exp} displayed a sufficient range of measured potencies to establish a QSAR model. In this model, the experimental activities were correlated with the calculated relative enzyme–inhibitor binding affinities expressed as the relative complexation Gibbs free energies, $\Delta\Delta G_{comp}$, Eq. (6). These were computed for the complexes obtained by modifying *in situ* the structure of the PR:JG-365 reference complex to that of the training set compounds followed by an extensive refinement. The $\Delta\Delta G_{comp}$ quantity reflects the enzyme's affinity for a potential inhibitor structure, however, due to the approximate character of its derivation a regression analysis was used to establish a linear correlation between the experimental and computed quantities (Table 1). The following statistically significant correlation equation was obtained (Fig. 4):

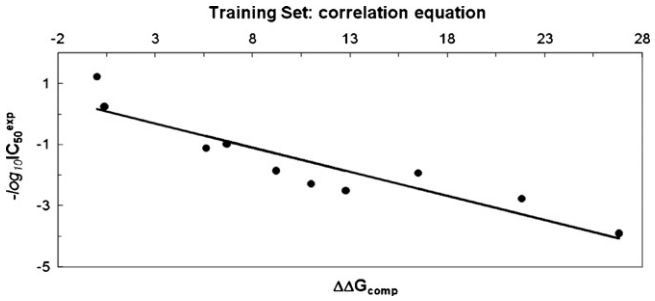

Fig. 4. Plot of the correlation equation between pIC_{50}^{exp} and relative complexation Gibbs free energies of the training set of PR inhibitors $\Delta\Delta G_{comp}$.

$$pIC_{50} = -\log_{10} IC_{50}^{exp} = -0.1577 \cdot \Delta\Delta G_{comp} + 0.1536 \quad (7)$$

(number of compounds $n = 10$, squared correlation coefficient of regression $R^2 = 0.86$, leave-one-out cross-validated squared correlation coefficient $R_{xv}^2 = 0.81$, standard error of the regression $\sigma = 0.59$, statistical significance of the regression (Fischer *F*-test) $F = 44.6$, level of the statistical significance $\alpha > 95\%$, range of experimental IC_{50}^{exp} values: 0.06–8400 nM).

It is interesting to notice that from the individuals contributions to the $\Delta\Delta G_{comp}$ the enzyme–inhibitor interaction displayed the highest correlation with the experimental inhibition potencies (correlation of pIC_{50} vs. $\Delta\Delta H_{MM}$ showed R_{xv}^2 of 0.56 only).

Using the regression Eq. (7), the $\Delta\Delta G_{comp}$ term can be used to predict PR inhibitory potencies (IC_{50}^{pre}) of peptidomimetics similar to the training set **FPT1**–**FPT10** provided that they preserve the same mode of PR binding. Computational prediction can considerably reduce the number of molecules, which need to be synthesized in a rational drug development project to obtain new

Table 2
Optimization of designed hexameric, pentameric and tetrameric C_2-symmetric peptidomimetic inhibitors of HIV-1 PR containing –Phe-Ψ[CHOH–CHOH]-Pro– core and range of flanking residues

Inhibitors	Chemical structure, designed HIV-1 PR Inhibitors[a]	M_w[b] (g mol^{-1})	$\Delta\Delta H_{MM}$[c] (kcal mol^{-1})	$\Delta\Delta G_{sol}$[d] (kcal mol^{-1})	$-\Delta\Delta TS_{vib}$[e] (kcal mol^{-1})	$\Delta\Delta G_{comp}$[f] (kcal mol^{-1})	IC_{50}^{pre}[g] (nM)
FP1	*Ac*-Trp-Tcl-**PheΨPro**-Tcl-Trp-*Ac*	961	0.6	−3.7	6.3	3.2	2.2
FP2	*Ac*-Phe-Val-**PheΨPro**-Val-dmPhe-*Ac*[h]	855	−13.2	0.8	6.4	−6.0	0.08
FP3	*Ac*-Btz-Val-**PheΨPro**-Val-Btz-*Ac*	941	−15.0	2.3	8.1	−4.6	0.1
FP4	*Ac*-Thi-Val-**PheΨPro**-Val-Thi-*Ac*	841	−12.8	2.9	7.1	−2.8	0.3
FP5	*Ac*-Tha-Val-**PheΨPro**-Val-Tha-*Ac*	839	−21.1	4.8	10.2	−6.1	0.08
FP6	*Ac*-Tha-Val-**PheΨPro**-Val-Thd-*Ac*	849	−19.2	6.0	4.3	−8.9	0.02
FP7	*Ac*-Tha-Val-**PheΨPro**-Val-Thm-*Ac*	835	−15.7	2.3	1.6	−11.8	0.01
FP8	*Ac*-Tcl-Val-**PheΨPro**-Val-Tcl-*Ac*	787	−15.0	−0.3	−1.1	−16.4	0.002
FP9	*Ac*-Val-Val-**PheΨPro**-Val-Val-*Ac*	731	−8.0	−0.9	0.4	−8.5	0.03
FP10	*Ac*-Ala-Val-**PheΨPro**-Val-Ala-*Ac*	675	−6.9	−1.3	4.4	−3.8	0.2
FP11	*Ac*-Gly-Val-**PheΨPro**-Val-Gly-*Ac*	647	9.8	−18.6	5.2	−3.6	0.2
FP12	Poa-Val-**PheΨPro**-Val-Poa	717	−3.1	−5.7	9.0	0.2	0.8
FP13	*Ac*-Tcl-Val-**PheΨPro**-Val-*Ac*	660	−7.5	0.6	1.3	−5.6	0.09
FP14	*Ac*-Tcl-Val-**PheΨPro**-Tcl-*Ac*	688	−4.9	−2.7	−1.8	−9.4	0.02
FP15	*Ac*-Tcl-**PheΨPro**-Val-Tcl-*Ac*	688	−4.9	0.8	0.0	−4.1	0.2
FP16	*Ac*-Val-**PheΨPro**-Val-Tcl-*Ac*	660	−1.0	−7.1	4.7	−3.4	0.2
FP17	*Ac*-Tcl-Val-**PheΨPro**-Mim-*Ac*	741	−12.0	−2.3	−0.1	−14.4	0.004
FP18	*Ac*-Mim-**PheΨPro**-Val-Tcl-*Ac*	741	−16.0	3.5	6.1	−6.4	0.07
FP19	*Ac*-Tcl-Val-**PheΨPro**-mPya	668	0.0	−4.6	2.6	−2.0	0.3
FP20	mPya-**PheΨPro**-Val-Tcl-*Ac*	668	4.9	−3.9	5.5	6.5	7.4
FP21	*Ac*-Tcl-Val-**PheΨPro**-mPyf	669	−4.7	−5.2	0.5	−9.4	0.02
FP22	mPyf-**PheΨPro**-Val-Tcl-*Ac*	669	6.1	−5.2	5.2	6.1	6.4
FP23	*Ac*-Tcl-Val-**PheΨPro**-dmPoa	681	−9.0	−4.2	1.6	−11.6	0.01
FP24	dmPoa-**PheΨPro**-Val-Tcl-*Ac*	681	−11.8	7.7	0.0	−4.1	0.2
FP25	dmPoa-**PheΨPro**-Val-dmPoa	674	7.7	−16.5	8.4	−0.4	0.6
FP26	*Ac*-Val-**PheΨPro**-Val-*Ac*	533	15.3	−11.2	5.9	10.0	26.5
FP27	*Ac*-Dtl-**PheΨPro**-Dtl-*Ac*	561	16.5	−8.1	1.9	10.3	29.5
FP28	*Ac*-Ile-**PheΨPro**-Ile-*Ac*	561	18.5	−9.5	4.0	13.0	79
FP29	*Ac*-Tcl-**PheΨPro**-Tcl-*Ac*	589	4.6	−2.8	4.0	5.8	5.8

Table 2 (*Continued*)

Inhibitors	Chemical structure, designed HIV-1 PR Inhibitors[a]	M_w[b] (g mol^{-1})	$\Delta\Delta H_{MM}$[c] (kcal mol^{-1})	$\Delta\Delta G_{sol}$[d] (kcal mol^{-1})	$-\Delta\Delta TS_{vib}$[e] (kcal mol^{-1})	$\Delta\Delta G_{comp}$[f] (kcal mol^{-1})	IC_{50}^{pre}[g] (nM)
FP30	*Ac*-Phe-**PheΨPro**-Phe-*Ac*	629	10.9	−10.2	9.0	9.7	23.8
FP31	*Ac*-Cha-**PheΨPro**-Cha-*Ac*	641	11.1	−9.1	2.2	4.2	3.2
FP32	*Ac*-Mim-**PheΨPro**-Mim-*Ac*	695	1.5	−7.9	16.0	9.6	22.9
FP33	CH_3-Mim-**PheΨPro**-Mim-CH_3	639	26.4	−6.7	8.8	28.5	>1 μM
FP34	*Ac*-Mib-**PheΨPro**-Mib-*Ac*	663	−4.0	−5.6	12.3	2.7	1.8
FP35	CH_3-Mib-**PheΨPro**-Mib-CH_3	607	36.5	−7.0	9.3	38.8	≫1 μM
FP36	*Ac*-Mic-**PheΨPro**-Mic-*Ac*	695	−2.4	−4.5	17.6	10.7	34.2
FP37	*Ac*-Mid-**PheΨPro**-Mid-*Ac*	691	0.8	−7.8	13.2	6.2	6.7
FP38	*Ac*-Mie-**PheΨPro**-Mie-*Ac*	719	0.1	−4.6	11.9	7.4	10.3
FP39	*Ac*-Min-**PheΨPro**-Min-*Ac*	695	1.4	−4.2	16.4	13.6	98
FP40	*Ac*-Mio-**PheΨPro**-Mio-*Ac*	697	2.7	−4.6	12.4	10.5	31.8
FP41	*Ac*-Mim-**PheΨPro**-mPya	622	6.9	−10.9	12.5	8.5	15.4
FP42	mPya-**PheΨPro**-Mim-*Ac*	622	13.3	−11.9	12.7	14.1	117
FP43	*Ac*-Mim-**PheΨPro**-mPyf	623	6.8	−10.5	12.1	8.4	14.8
FP44	mPyf-**PheΨPro**-Mim-*Ac*	623	12.7	−8.9	12.5	16.3	261
FP45	*Ac*-Mim-**PheΨPro**-dmPoa	635	−1.4	−6.9	12.1	3.8	2.8
FP46	dmPoa-**PheΨPro**-Mim-*Ac*	635	5.9	−9.8	14.6	10.7	34.2
FP47	Qna-**PheΨPro**-Kni	553	−10.1	15.4	13.2	18.5	581
FP48	Qna-**PheΨPro**-Kno	554	3.4	3.1	14.2	20.7	>1 μM
FP49	Qna-**PheΨPro**-Kna	570	2.1	1.4	14.9	18.4	560
FP50	dmPoa-**PheΨPro**-Kni	560	13.1	−14.3	13.0	11.8	51
FP51	dmPoa-**PheΨPro**-Kna	577	13.8	−14.8	11.6	10.6	32.9
FP52	Hfa-**PheΨPro**-Kni	511	18.3	−7.4	8.4	19.3	776
FP53	Tho-**PheΨPro**-Kni	539	22.8	−15.5	12.2	19.5	834
FP54	dmPha-**PheΨPro**-dmPha	571	−2.1	5.9	6.3	10.1	27.5
FP55	mePha-**PheΨPro**-mePha	603	6.3	−8.5	4.7	2.5	1.7
FP56	ePha-**PheΨPro**-ePha	571	14.1	−10.2	5.6	9.5	22.1
FP57	mPya-**PheΨPro**-mPya	549	18.5	−22.0	13.9	10.4	30.6
FP58	mPyb-**PheΨPro**-mPyb	549	19.5	−13.8	13.1	18.8	647
FP59	dmPyc-**PheΨPro**-dmPyc	577	6.4	−12.8	15.8	9.4	21.3
FP60	Qnn-**PheΨPro**-Knn	554	−1.8	2.6	14.1	14.9	157
FP61	Qnb-**PheΨPro**-Knn	504	1.2	2.3	11.7	15.2	175
FP62	Qnb-**PheΨPro**-Kni	503	−1.0	3.7	13.5	16.2	252
FP63	Qnb-**PheΨPro**-Knd	518	1.8	4.0	10.9	16.7	302
FP64	Qnc-**PheΨPro**-Knd	518	1.6	4.4	11.6	17.6	419
FP65	Qnc-**PheΨPro**-Kne	518	5.9	6.0	11.9	23.8	>1 μM
FP66	Pta-**PheΨPro**-Pta	555	21.2	−19.6	15.9	17.5	404
FP67	Ptd-**PheΨPro**-Ptd	553	15.2	−16.4	14.0	12.8	73
FP68	mPye-**PheΨPro**-mPye	551	22.3	−17.9	16.3	20.7	>1 μM
FP69	mPyf-**PheΨPro**-mPyf	551	22.6	−22.7	15.2	15.1	169
FP70	Tha-**PheΨPro**-Tha	641	7.6	−10.1	15.6	13.1	82
FP71	mhPoa-**PheΨPro**-mhPoa	607	10.9	−12.7	11.8	10.0	26.5
FP72	mgPoa-**PheΨPro**-mPoa	563	4.6	−12.2	11.7	4.1	3.1
FP73	mPya-**PheΨPro**-mPyf	550	25.2	−22.6	10.6	13.2	85
FP74	mPyf-**PheΨPro**-mPya	550	23.4	−22.4	12.5	13.5	94
FP75	mPya-**PheΨPro**-dmPoa	562	19.8	−19.3	11.1	11.6	47.4
FP76	dmPoa-**PheΨPro**-mPya	562	3.4	−19.8	12.0	−4.4	0.1
FP77	mPyf-**PheΨPro**-dmPoa	563	17.6	−20.6	15.5	12.5	66
FP78	dmPoa-**PheΨPro**-mPyf	563	10.1	−19.0	14.4	5.5	5.2
FP79	admPoa-**PheΨPro**-mPoa	576	10.0	−12.6	8.5	5.9	6.0

Computed contributions to the complexation Gibbs free energy $\Delta\Delta G_{comp}$ and predicted inhibitory concentrations IC_{50}^{pre}, are shown.

[a] For the chemical structures of the training set of inhibitors see Fig. 2.

[b] M_w is the molecular mass of the inhibitors candidates.

[c] $\Delta\Delta H_{MM}$ is the relative enthalpic contribution to the Gibbs free energy change related to the enzyme–inhibitor complex formation derived by molecular mechanics (MM): $\Delta\Delta H_{MM} \cong [E_{MM}\{PR{:}I\} - E_{MM}\{I\}] - [E_{MM}\{PR{:}\mathbf{FPT1}\} - E_{MM}\{\mathbf{FPT1}\}]$, **FPT1** is the reference inhibitor.

[d] $\Delta\Delta G_{solv}$ is the relative salvation Gibbs free energy contribution to the Gibbs free energy change related to the enzyme–inhibitor complex formation: $\Delta\Delta G_{solv} = [G_{solv}\{PR{:}I\} - G_{solv}\{I\}] - [G_{sol}\{PR{:}\mathbf{FPT1}\} - G_{sol}\{\mathbf{FPT1}\}]$.

[e] $-\Delta\Delta TS_{vib}$ is the relative entropic contribution of the inhibitor to the Gibbs free energy related to the enzyme–inhibitor complex formation: $\Delta\Delta TS_{vib} = [TS_{vib}\{I\}_{PR} - TS_{vib}\{I\}] - [TS_{vib}\{\mathbf{FPT1}\}_{PR} - S_{vib}\{\mathbf{FPT1}\}]$.

[f] $\Delta\Delta G_{comp}$ is the relative Gibbs free energy change related to the enzyme–inhibitor complex formation: $\Delta\Delta G_{comp} \cong \Delta\Delta H_{MM} + \Delta\Delta G_{solv} - \Delta\Delta TS_{vib}$.

[g] IC_{50}^{pre} is the predicted IC_{50} of the designed inhibitor candidates, which was calculated from the regression equation: $IC_{50}^{pre} = 10^{(0.1577\Delta\Delta G_{comp} - 0.1536)}$, which was obtained by the QSAR model (Fig. 4).

[h] Reagents used: Btz, β-(benzothiazol-2-yl)-L-alanine; Cha, β-cyclohexyl-L-alanine; Dtl, L-α-*tert*-butylglycine; Hfa, carbonic acid mono-(tetrahydro-furan-3-yl) ester; Kna, carbonic acid mono-(2-methoxy-benzyl) ester; Knd, (3,5-dimethyl-pyridin-4-ylmethyl)-carbamic acid; Kne, (2,4-dimethyl-pyridin-3-ylmethyl)-carbamic acid; Kni, (2-methyl-benzyl)-carbamic acid; Knn, (3-methyl-pyridin-2-ylmethyl)-carbamic acid; Kno, carbonic acid mono-(2-methyl-benzyl) ester; Mib, deoxymimosine; Mic, 2-hydroxy-mimosine; Mid, 2-methyl-deoxymimosine; Mie, 2,6-dimethyl-deoxymimosine; Mim, mimosine; Min, β-(5-hydroxy-4-oxo-1,4-dihydro-pyridin-2-yl)-L-alanine; Mio, β-(5-hydroxy-6-oxo-1,6-dihydro-pyrimidin-2-yl)-L-alanine; Tha, β-(2-thienyl)-L-alanine; dmPaa+, protonated (2,6-dimethyl-phenylamino)-acetic acid; dmPha, 3-(2,6-dimethyl-phenyl)-propionic acid; ePha, 3-(2-ethyl-phenyl)-propionic acid; mePha, 3-(2-methyl-6-ethyl-phenyl)-propionic acid; dmPhe, 2,6 dimethyl phenylalanine; admPoa, (4-amino-2,6-dimethyl-phenoxy)-acetic acid; dmPoa, 2,6-dimethyl-phenoxyacetic acid; mPoa, 2-methyl-phenoxyacetic acid; mgPoa, 6-hydroxy-2-methyl-phenoxyacetic acid; mhPoa, 3-hydroxy-2-methyl-phenoxyacetic acid; Poa, phenoxyacetic acid; Pta, (pyrimidin-2-ylsulfanyl)-acetic acid; Ptd, (pyridin-4-ylsulfanyl)-acetic acid; mPya, (3-methyl-pyridin-2-yloxy)-acetic acid; mPyb, (4-methyl-pyridin-3-yloxy)-acetic acid; dmPyc, (3,5-dimethyl-pyridin-4-yloxy)-acetic acid; mPyc, (3-methyl-pyridin-4-yloxy)-acetic acid; mPye, (3-methyl-pyrazin-2-yloxy)-acetic acid; mPyf, (5-methyl-pyrimidin-4-yloxy)-acetic acid; Qna, quinoline-2-carboxylic acid; Qnb, pyridine-2-carboxylic acid; Qnc, nicotinic acid; Qnn, quinoline-7-carboxylic acid; Tcl, *tert*-L-leucine; Tha, β-(2-thienyl)-L-alanine; Thd, β-(5,5-dimethyl-cyclopenta-1,3-dienyl)-L-alanine; Thm, β-(5-methyl-cyclopenta-1,3-dienyl)-L-alanine; Thi, β-(2-thiazolyl)-L-alanine; Tho, carbonic acid mono-thiazol-5-ylmethyl ester.

Table 3
Mean values of contributions to the PR binding affinity averaged over the groups of hexameric, pentameric and tetrameric PR inhibitors

Inhibitor size	Inhibitors	Count	Mean $M_w \pm$ S.D.[a] (g mol^{-1})	Mean $\Delta\Delta H_{MM} \pm$ S.D. (kcal mol^{-1})	Mean $\Delta\Delta G_{sol} \pm$ S.D. (kcal mol^{-1})	Mean $-\Delta\Delta TS_{vib} \pm$ S.D. (kcal mol^{-1})
Hexameric	**FP1–FP12**	12	807 ± 98	−9.9 ± 8.9	−0.9 ± 6.5	5.2 ± 3.4
Pentameric	**FP13–FP25**	13	684 ± 27	−4.1 ± 8.3	−3.0 ± 6.2	2.6 ± 3.4
Tetrameric	**FP26–FP79**	54	587 ± 55	9.8 ± 9.5	−8.9 ± 8.5	11.6 ± 3.7

[a] Mean value of the molecular weight and individual contributions to $\Delta\Delta G_{comp}$ for the groups of inhibitors with the same number of residues (Table 1) and their standard deviations.

lead compounds with high binding affinities and improved characteristics.

3.2. Design of new inhibitors

The structure-based design of new PR inhibitors departed from the model of the complex of the most active training set inhibitor-reference compound **FPT1** (Fig. 2), which was modified to obtain the PR:inhibitor complexes of the new derivatives. During the design, structural complementarity of the flanking residues of the new candidates with the binding pockets of the enzyme's active site was taken into consideration. In our previous studies using this approach, it was possible to successfully reduce the peptidic nature and molecular mass of some larger PR inhibitors, without decreasing their potency, by replacing bulky acetylated aromatic amino acid residues in the P_3 or $P_{3'}$ positions with kynurenic acid and in P_2 and $P_{2'}$ with phenoxyacetic acid (Poa) [9–13]. Especially Poa, which can form favourable hydrophobic contacts with some

Table 4
Ten lead compounds identified out of 79 designed inhibitor candidates by means of a multiparameter penalty function, which combines predicted IC_{50}^{pre} and predicted ADME-related descriptors computed by QikProp program [37]

Inhibitor[a]	# Stars[b]	M_w[c] (g mol^{-1})	Dipole[d] (Deb)	S_{mol}[e] (Å^2)	$S_{mol,hfob}$[f] (Å^2)	V_{mol}[g] (Å^3)	# R.B.[h]	HB_{don}[i]	HB_{acc}[j]	$\log_{10} P_{o/w}$[k]	$\log_{10} S_{wat}$[l] (g dm^{-3})	$\log_{10} K_{HSA}$[m]	$\log_{10} B/B$[n]	BIP_{caco}[o] (nm s^{-1})	# Metab[p]	$IC_{50}^{exp}/IC_{50}^{pre}$[q] (nM)	Penalty[r]
Ritonavir	7	721	1.3	1051	450	2122	18	3	11	6.2	−6.3	0.8	−2.6	80	9	3	–
Lopinavir	2	629	2.4	963	418	1943	16	4	10	5.4	−5.1	0.6	−2.1	266	8	0.5	–
Saquinavir	0	671	6.8	929	438	1924	14	5	14	1.9	−1.0	−0.5	−1.8	11	6	2	–
Indinavir	2	614	5.3	999	412	1932	14	4	14	2.5	−2.0	−0.1	−1.0	4	11	1	–
Nelfinavir	0	568	6.2	885	487	1759	12	4	10	4.4	−4.3	0.7	−1.0	52	6	2	–
Amprenavir	0	506	9.5	795	361	1526	13	4	11	3.2	−4.6	0.1	−1.9	123	4	1	–
Tipranavir	1	603	9.1	914	253	1713	11	2	11	5.2	−7.7	0.9	−2.4	12	8	33	–
Min[s]	0	450	1.0	750	250	1500	5	2	8	1.5	−7.0	−0.5	−2.6	50	0	0	
Max	8	750	10.0	1100	500	2200	18	5	15	6.5	−1.0	1.0	−1.0	300	12	1[t]	–
Weight fact.	1	5	1	1	1	1	5	5	5	5	1	1	1	1	1	20	
FP76	0	562	9.2	798	290	1659	14	3	11	3.4	−2.3	−0.2	−1.6	513	11	0.1	0.02
FP24	2	681	4.3	938	562	2021	18	4	13	3.0	−0.9	−0.6	−2.0	601	9	0.2	0.04
FP19	1	668	9.6	883	556	1949	18	4	13	2.5	0.0	−0.8	−2.0	527	9	0.3	0.05
FP25	1	674	6.1	818	446	1897	17	3	13	3.6	−0.5	−0.3	−1.5	1695	12	0.6	0.07
FP16	3	660	5.3	1007	697	2092	18	4	14	2.1	−0.6	−1.0	−2.5	165	6	0.2	0.10
FP23	3	681	2.9	984	626	2084	18	4	13	3.5	−1.7	−0.5	−1.9	761	9	0.01	0.10
FP13	2	660	4.4	969	704	2049	18	4	14	1.9	0.0	−1.1	−2.4	225	6	0.09	0.12
FP14	3	688	10.4	997	738	2134	19	4	14	2.6	−0.3	−0.9	−2.2	504	6	0.02	0.14
FP15	4	688	10.3	1001	732	2132	19	4	14	2.7	−0.4	−0.9	−2.0	939	6	0.2	0.19
FP21	3	669	13.4	804	522	1886	18	4	15	1.5	1.7	−1.9	−1.9	605	10	6.4	0.21

[a] Descriptors were computed for minimized inhibitor structures of the low-energy conformers of free designed and clinically used inhibitors.
[b] Drug likeness—the number of property descriptors (from the full list of 36 descriptors of QikProp, ver. 2.0.005 [37]) that fall outside the range of values for 95% of known drugs.
[c] Molecular weight in Da (range for 95% of drugs: 130–725 Da).
[d] Dipole moment of the molecule in Debye (1 Debye = 3.335 64 × 10^{-30} C m) (range for 95% of drugs: 1.0–12.5 Debye).
[e] Total solvent-accessible molecular surface, in Å^2 (probe radius 1.4 Å) (range for 95% of drugs: 300–1000 Å^2).
[f] Hydrophobic portion of the solvent-accessible molecular surface, in Å^2 (probe radius 1.4 Å) (range for 95% of drugs: 0–750 Å^2).
[g] Total volume of molecule enclosed by solvent-accessible molecular surface, in Å^3 (probe radius 1.4 Å) (range for 95% of drugs: 500–2000 Å^3).
[h] Number of rotatable bonds (range for 95% of drugs: 0–15).
[i] Number of hydrogen bonds donated by the molecule (range for 95% of drugs: 0–6).
[j] Number of hydrogen bonds accepted by the molecule (range for 95% of drugs: 2–20).
[k] Logarithm of partitioning coefficient between n-octanol and water phases (range for 95% of drugs: −2 to 6).
[l] Logarithm of aqueous solubility (range for 95% of drugs: −6.0 to 0.5).
[m] Logarithm of predicted binding constant to human serum albumin (range for 95% of drugs: −1.5 to 1.2).
[n] Logarithm of predicted blood/brain barrier partition coefficient (range for 95% of drugs: −3.0 to 1.0).
[o] Predicted apparent Caco-2 cell membrane permeability in Boehringer–Ingelheim scale, in (nm/s) (range for 95% of drugs: <5 low, >100 high).
[p] Number of likely metabolic reactions (range for 95% of drugs: 0–15).
[q] Experimental IC_{50}^{exp} values of clinically used and designed HIV PR inhibitors. IC_{50}^{exp} of Ritonavir, Indinavir and Nelfinavir were taken from Ref. [43], Lopinavir [46], Saquinavir [47], Amprenavir [48], and Tipranavir [49]. IC_{50}^{pre} is the predicted potency of the designed PR inhibitors, Table 2.
[r] Priority of the lead selection was based on the multiparameter activity-ADME penalty function $P = 1/N_p \sum_i^{N_p} (w_i P_i)$ derived from 16 individual penalties $P_i = |u_i - \bar{u}_i|/(nu_{SD})$ of descriptor u ranges ($\bar{u}_i$ is the mean value), standard deviations (u_{SD}) and statistical weight factors w_i as defined in Cerius^2 [29].
[s] Property value ranges (minimum value, maximum value, weight factor) of ADME property-related descriptors of clinically used HIV PR inhibitors computed via *QickProp* [37]. The ranges were rounded off or extended to include desired favourable descriptor values (e.g. lower M_w, lower S_{mol}, etc.).
[t] Maximum value of the IC_{50}^{pre} value range was decreased to 1 nM and standard deviation to 0.5 nM in order to bias the selection towards the designed analogues with the highest predicted inhibitory activity.

residues lining the lipophilic S_2 and $S_{2'}$ pockets (Ala28, Val32, Ile47, Ile 50 and Ile84) can be further improved by using 2,6-dimethyl-Poa (dmPoa) and other Poa derivatives and analogues, e.g. such as (pyrimidin-2-ylsulfanyl)-acetic acid (Pta), (3-methyl-pyridin-2-yloxy)-acetic acid (mPya) or (5-methyl-pyrimidin-4-yloxy)-acetic acid (mPyf).

While keeping the –PheΨPro– core conserved we have designed a series of over 150 inhibitor candidates out of which we report here 79 symmetrical and nonsymmetrical inhibitors with up to two flanking residues occupying the S_3, S_2, and $S_{2'}$, $S_{3'}$ sites (Table 2). Amino acid and carboxylic acid residues were selected on the basis of their potential for forming specific (hydrogen bonding) or hydrophobic interactions with the binding pockets of the PR substrate binding cleft as well as on the availability of the relevant reagents from the suppliers of chemicals. Although smaller tetrameric inhibitors were preferred in an attempt to reduce molecular weight and improve the pharmacokinetic profile of the inhibitors, the highest predicted inhibitory potencies were displayed by hexameric inhibitors (**FP1**–**FP12**) and pentameric inhibitors (**FP13**–**FP25**), (Table 2). The predicted potencies for the most potent analogues: **FP8**, **FP17**, **FP7**, **FP6**, **FP14**, **FP21** and **FP23** which were obtained from the QSAR model by converting computed binding affinities ($\Delta\Delta G_{comp}$) into IC_{50}^{pre} using the regression Eq. (7), reached the low picomolar range.

In the most promising tetrameric inhibitors **FP29**, **FP31**, **FP34**, **FP37**, **FP45**, **FP55**, **FP72**, **FP76**, **FP78** and **FP79** with the predicted IC_{50}^{pre} reaching low nanomolar range (<7 nM) the central –PheΨPro– core was flanked by acetylated β-cyclohexyl-alanine,

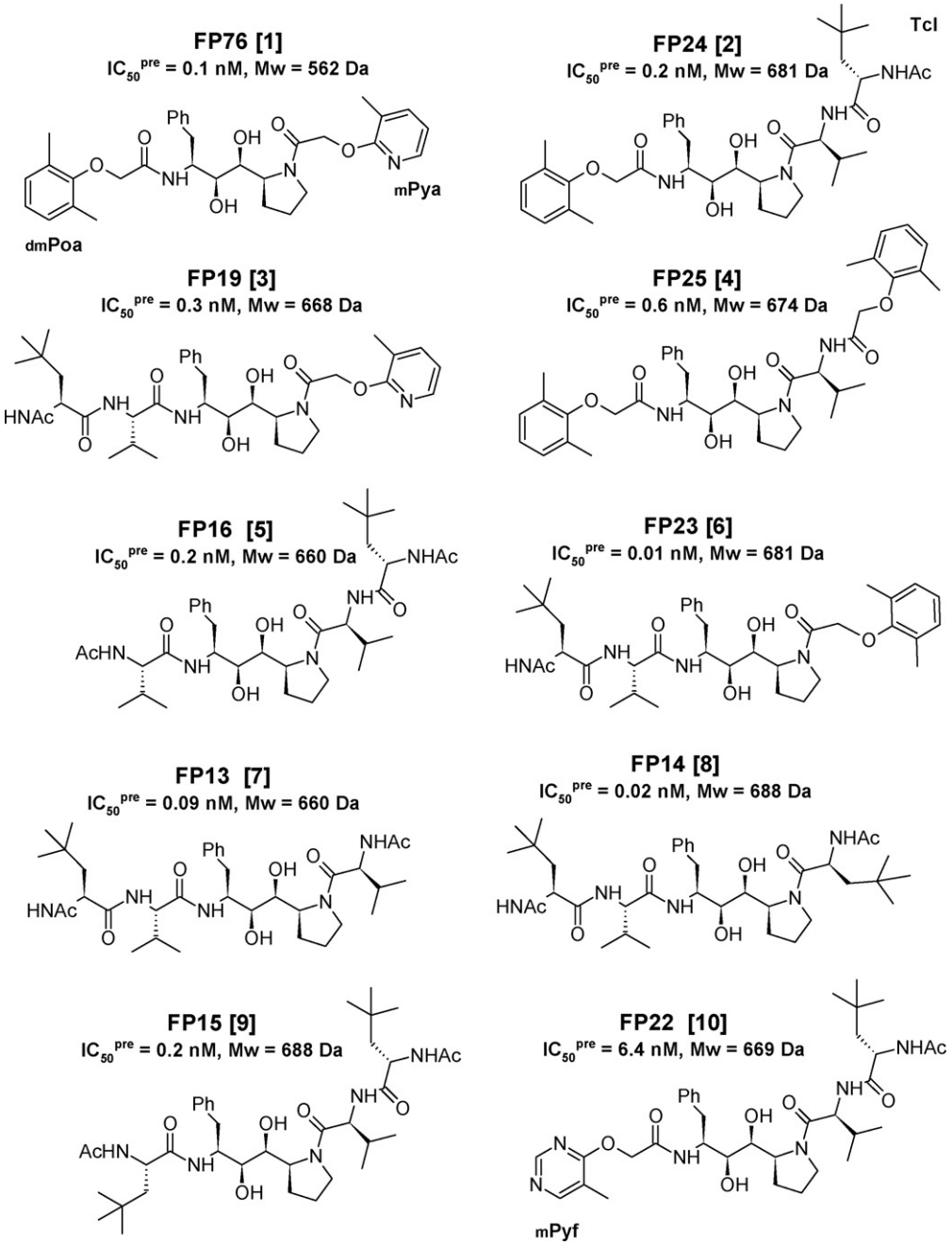

Fig. 5. Chemical structures of ten lead candidates endowed with the best combination of high-predicted inhibitory potencies (IC_{50}^{pre}) towards the HIV-1 PR and favourable ADME-related properties.

Tcl or derivatives of mimosine: Mim, Mib and Mid and derivatives of phenoxyacetic acid: dmPoa, mPya, mPyf and admPoa (for details see the footnote of Table 2). These residues represent suitable hydrophobic capping groups that seal both ends of the PR substrate binding cleft.

To determine the main driving force contribution to the binding of the hexameric, pentameric and tetrameric inhibitor candidates to the PR receptor we have averaged the individual contributions to the binding affinity over the group members (Table 3). In case of the larger hexameric peptidomimetics the main contribution driving the receptor binding comes form the interaction with the receptor ($\Delta\Delta H_{MM}$). For the pentameric analogues the receptor binding and solvent effect contribute about equally to the driving force, while for the smaller tetrameric analogues the solvent effect contributes most to the receptor binding. The entropic effect opposes the receptor binding for analogues of all sizes. The tetrameric inhibitor **FP76** driven to form the complex with the PR chiefly by the effect of the solvent may be expected to display a reduced propensity for diminished activity towards drug resistant PR forms.

3.3. ADME-related properties prediction

Numerous drugs at a late stage of pharmaceutical development and many more lead compounds fail due to adverse pharmacokinetic properties [44,45]. It is therefore important to incorporate ADME properties prediction into lead compound selection.

We have selected a number of relevant molecular properties that may help to predict the pharmacokinetic behaviour of PR inhibitors and prioritize selection of lead compounds for further development. A set of 15 ADME-related descriptors calculated by the *QikProp* program [37] is given in Table 4. To establish the optimum values of these descriptors for orally bioavailable peptidomimetics, we have determined the ranges of the descriptor

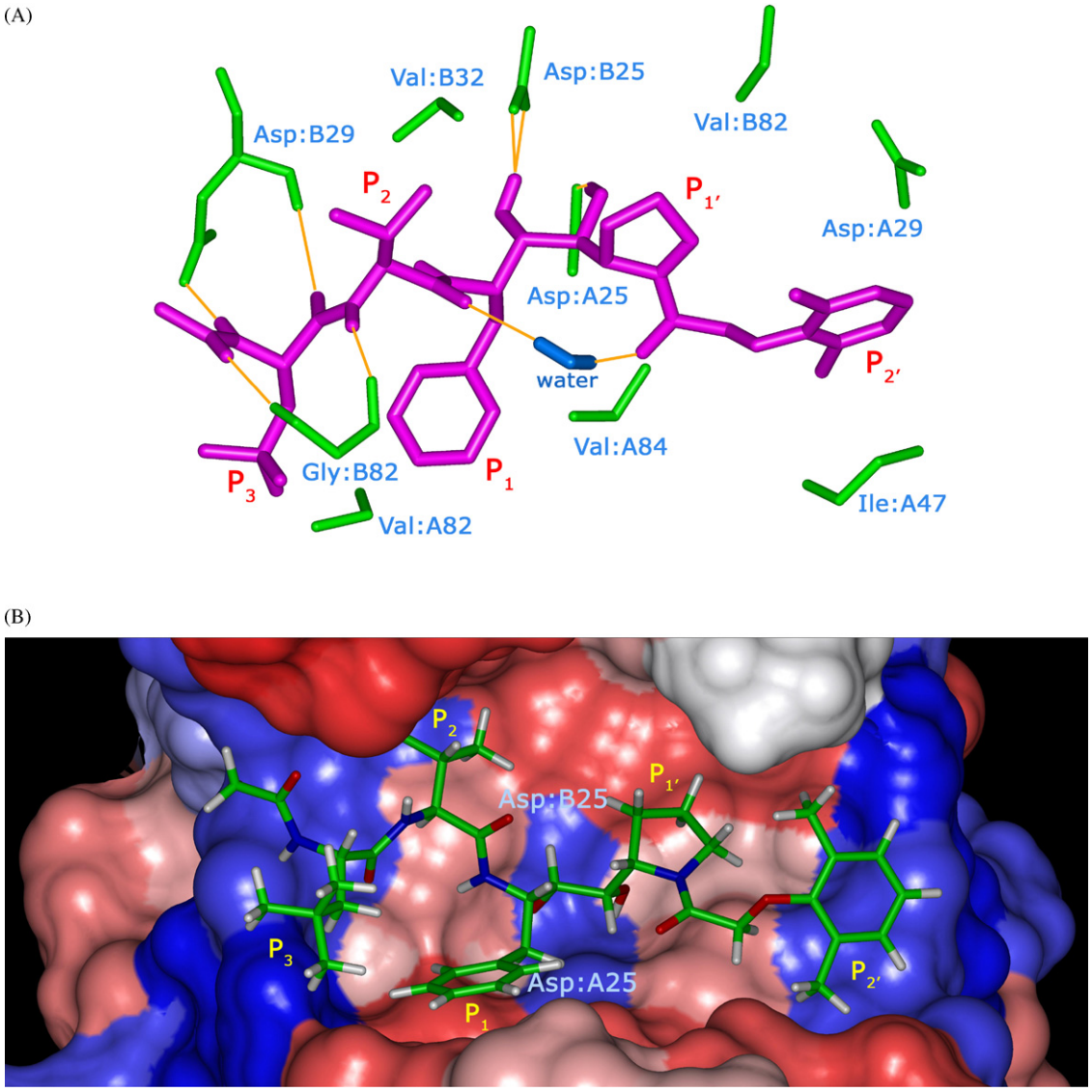

Fig. 6. (A) Selected lead compound **FP23** (*Ac*-Tcl–Val–Phe-Ψ[CHOH–CHOH]-Pro–dmPoa) in stick representation, purple colour) at the active site of HIV-1 PR. Hydrogen bonds between the lead, structural water molecule and catalytic residues Asp:A25 and Asp:B25 of the HIV-1 PR are shown in orange colour. Hydrogen bonds between the residues *Ac*-Tcl–Val– in positions P_3 and P_2 of the **FP23** and the backbone atoms of residues of Asp:B29 and Gly:B48 are also shown. (B) Cross-section of the binding pocket of HIV-1 PR with complexed inhibitor **FP23** (in stick representation). Connolly surface of the active site is coloured by the PR residue hydrophobicity (hydrophilic residues: blue, intermediate residues: white, hydrophobic residues: red colour). Atom type colouring scheme for the inhibitor: carbon: green, oxygen: red, nitrogen: blue, hydrogen: white, sulphur: yellow. Schechter and Berger notation [4] of residues of protease substrates/inhibitors: $P_3P_2P_1–P_1'P_2'$ is shown for the **FP23**. Front part of the binding pocket has been removed.

values for seven clinically used HIV-1 PR inhibitors (Fig. 3) (Table 4). Thus, the selection the inhibitor candidates for further development was based not only on the predicted inhibitory activities but also on their compliance with optimum value ranges of the ADME-related descriptors. The overall rank order of the analogues was calculated with help of a penalty function, which reflects the total score of individual ADME descriptor penalties and the predicted activities (IC_{50}^{pre}).

The analogues with best predicted PR inhibitory potencies and favourable ADME-related properties, which scored at first ten positions of the penalty function (shown in Fig. 5) can be selected as the lead candidates. Eight of the selected candidates are pentameric peptidomimetics with picomolar inhibitory potencies and molecular weight within the range 660–688 Da. One tetrameric candidate **FP76** similar to the training set compounds **FPT7**, **FPT9** and **FPT10**, which bears some similarity to the inhibitor AG 1776 [20] (Fig. 1), displays molecular weight of 562 Da, favourable ADME-related properties and predicted activity in the picomolar range.

The combination of *Ac*-*tert*-leucine–valine– (*Ac*-Tcl–Val–) flanking residues occurring in analogues with the highest predicted activities seems to be very effective on both prime and non-prime sites of the –PheΨPro– core, probably due to the elevated conformational flexibility and space filling properties of the hydrophobic side chains occupying the S_3, S_2 or $S_{2'}$, $S_{3'}$ pockets. In addition, in the modelled complexes the peptide bond atoms of the *Ac*-Tcl and Val residues form hydrogen bonds to the backbone atoms of Gly:B48 and Asp:B29 while the acetyl cap forms a hydrogen bond to the amino group of Gly:B48 (**FP23**, Fig. 6). As mentioned above, specific interactions of inhibitors with the backbone atoms of the receptor may reduce the potential loss of the inhibitory potency of a drug towards the emerging drug-resistant mutant forms of its receptor [17]. In **FP14** the binding to the PR is driven in part also by the entropic term (Table 2), making this lead candidate somewhat less sensitive to PR mutations conferring the resistance.

The analogues **FP14**, **FP23** and **FP76** represent thus potent virtual lead candidates, which might be suitable for further development.

4. Conclusions

In conclusion, we presented here an integrated method for designing, synthesizing and pharmacokinetic profile predicting of potential inhibitors of HIV-1 aspartic protease sharing the –Phe-Ψ[CHOH–CHOH]-Pro– transition-state-isostere core. The emphasize was put on the structure-based design method and optimization of the flanking residues, which relies on the calculated complexation Gibbs free energies of binding to the PR receptor. A statistically significant QSAR model was established between the calculated binding affinities, $\Delta\Delta G_{comp}$, and the measured activities IC_{50}^{exp} of the training set of inhibitors, which confirmed that our computational approach can serve for prediction of PR inhibitory potencies of peptidomimetics. The appropriate choice of the flanking residues allowed us to identify lead candidates **FP14**, **FP23** and **FP76** with reduced peptidic nature, predicted inhibitory potencies in the picomolar range and favourable ADME profiles. The candidates **FP14** and **FP23** display also increased potential to partially escape drug resistance due to hydrogen bonding interactions of the *Ac*-Tcl–Val– residues to atoms of the PR backbone. The binding of the lead **FP76** to the PR predicted to be driven mainly by the effect of the solvent may be less affected by the mutations of the receptor.

The presented study can thus help to direct the attention of pharmaceutical companies working on the preparation of a next generation of anti-HIV drugs towards the sampled portion of the chemical space, which is predicted to contain compounds with picomolar HIV-1 PR inhibitory potencies and favourable ADME-related properties.

References

[1] E. De Clercq, New approaches toward anti-HIV chemotherapy, J. Med. Chem. 48 (2005) 1297–1313.
[2] H. Mitsuya, K. Maeda, D. Das, A.K. Ghosh, Development of protease inhibitors and the fight with drug-resistant HIV-1 variants, Adv. Pharmacol. 56 (2008) 169–197.
[3] A. Wlodawer, A. Miller, M. Jaskólski, B.K. Sathyanarayana, E. Baldwin, I. Weber, L.M. Selk, L. Clawson, J. Schneider, S.B.H. Kent, Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease, Science 245 (1989) 616–623.
[4] I. Schechter, A. Berger, On the size of the active site in proteases. I. Papain, Biochem. Biophys. Res. Commun. 27 (1967) 157–162.
[5] R. Kantor, D. Katzenstein, Polymorphism in HIV-1 non-subtype B protease and reverse transcriptase and its potential impact on drug susceptibility and drug resistance evolution, AIDS Rev. 5 (2003) 25–35.
[6] S.Y. Rhee, W.J. Fessel, A.R. Zolopa, L. Hurley, T. Liu, J. Taylor, D.P. Nguyen, S. Slome, D. Klein, M. Horberg, J. Flamm, S. Follansbee, J.M. Schapiro, R.W. Shafer, S.Y. Rhee, HIV-1 Protease and reverse-transcriptase mutations: correlations with antiretroviral therapy in subtype B isolates and implications for drug-resistance surveillance, J. Infect. Dis. 192 (2005) 456–465 (HIV drug resistance database, http://hivdb.stanford.edu).
[7] H.B. Schock, V.M. Garsky, L.C. Kuo, Mutational anatomy of an HIV-1 protease variant conferring cross-resistance to protease inhibitors in clinical trials. Compensatory modulations of binding and activity, J. Biol. Chem. 271 (1996) 31957–31963.
[8] F. Benedetti, S. Miertus, S. Norbedo, A. Tossi, P. Zlatoidzky, Versatile and stereoselective synthesis of diamino diol dipeptide isosteres, core units of pseudopeptide HIV protease inhibitors, J. Org. Chem. 62 (1997) 9348–9353.
[9] V. Frecer, S. Miertus, A. Tossi, D. Romeo, Rational design of inhibitors for drug-resistant HIV-1 aspartic protease mutants, Drug Des. Disc. 15 (1998) 211–231.
[10] A. Tossi, I. Bonin, N. Antcheva, S. Norbedo, F. Benedetti, S. Miertus, A.C. Nair, T. Maliar, F. Dal Bello, G. Palù, D. Romeo, Aspartic protease inhibitors. An integrated approach for the design andsynthesis of diaminodiol-based peptidomimetics, Eur. J. Biochem. 267 (2000) 1715–1722.
[11] V. Frecer, S. Miertus, Interactions of ligands with macromolecules: Rational design of specific inhibitors of aspartic protease of HIV-1, Macromol. Chem. Phys. 203 (2002) 1650–1657.
[12] V. Frecer, E. Burello, S. Miertus, Combinatorial design of nonsymmetrical cyclic urea inhibitors of aspartic protease of HIV-1, Bioorg. Med. Chem. 13 (2005) 5492–5501.
[13] V. Frecer, A. Jedinak, A. Tossi, F. Berti, F. Benedetti, D. Romeo, S. Miertus, Structure based design of inhibitors of aspartic protease of HIV-1, Lett. Drug Des. Disc. 2 (2005) 638–646.
[14] F. Benedetti, F. Berti, S. Norbedo, Epoxyalcohol route to hydroxyethylene dipeptide isosteres, Stereodivergent synthesis of the diamino alcohol core of ritonavir and its C-2 epimer, J. Org. Chem. 67 (2002) 8635–8643.
[15] F. Benedetti, F. Berti, F. Dinon, G. Nardin, S. Norbedo, Synthesis of a Val-Pro diaminodiol dipeptide isostere by epoxyamine cyclization, Org. Lett. 6 (2004) 1017–1019.
[16] F. Dinon, P. Campaner, F. Berti, A. Tossi, F. Benedetti, Italian Patent PCT/EP2005/052770.
[17] Y. Koh, H. Nakata, K. Maeda, H. Ogata, G. Bilcer, T. Devasamudram, J.F. Kincaid, P. Boross, Y.F. Wang, Y. Tie, P. Volarath, L. Gaddis, R.W. Harrison, I.T. Weber, A.K. Ghosh, H. Mitsuya, Novel bis-tetrahydrofuranylurethane-containing nonpeptidic protease inhibitor (PI) UIC-94017 (TMC114) with potent activity against multi-PI-resistant human immunodeficiency virus in vitro, Antimicrob. Agents Chemother. 47 (2003) 3123–3129.
[18] D.H. Rich, C.-Q. Sun, J.V.N. Vara Prasad, A. Pathiasseril, M.V. Toth, G.R. Marshall, M. Clare, R.A. Mueller, K. Houseman, Effect of hydroxyl group configuration in hydroxyethylamine dipeptide isosteres on HIV protease inhibition. Evidence for multiple binding modes, J. Med. Chem. 34 (1991) 1222–1225.
[19] T. Mimoto, J. Imai, S. Kisanuki, H. Enomoto, N. Hattori, K. Akaji, Y. Kiso, Kynostatin (KNI)-227 and -272, highly potent anti-HIV agents: conformationally constrained tripeptide inhibitors of HIV protease containing allophenylnorstatine, Chem. Pharm. Bull. (Tokyo) 40 (1992) 2251–2253.
[20] Y. Ohno, Y. Kiso, Y. Kobayashi, Solution conformations of KNI-272, a tripeptide HIV protease inhibitor designed on the basis of substrate transition state: determined by NMR spectroscopy and simulated annealing calculations, Bioorg. Med. Chem. 4 (1996) 1565–1572.
[21] S. Cha, Tight-binding inhibitors. I. Kinetic behavior, Biochem. Pharmacol. 24 (1975) 2177–2185.
[22] A.L. Swain, M.M. Miller, J. Green, D.H. Rich, J. Schneider, S.B.H. Kent, A. Wlodawer, X-ray crystallographic structure of a complex between a synthetic protease of human immunodeficiency virus 1 and a substrate-based hydroxyethylamine inhibitor, Proc. Natl. Acad. Sci. U.S.A. 87 (1990) 8805–8809.
[23] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne, The protein data bank, Nucl. Acids Res. 28 (2000) 235–242.
[24] Insight-II version 2000 molecular modelling package, Discover version 2.98 simulation package, and DelPhi version 3.0 solvation program, 2000, Accelrys, Inc., San Diego, Calif.

[25] Y.X. Wang, D.I. Freedberg, T. Yarnazaki, P.T. Wingfield, S.J. Stahl, J.D. Kaufman, Y. Kiso, D.A. Torchia, Solution NMR evidence that the HIV-1 protease catalytic aspartyl groups have different ionization states in the complex formed with the asymmetric drug KNI-272, Biochemistry 35 (1996) 9945–9950.
[26] J.R. Maple, U. Dinur, A.T. Hagler, Derivation of force fields for molecular mechanics and dynamics from ab initio energy surfaces, Proc. Natl. Acad. Sci. U.S.A. 85 (1988) 5350–5354.
[27] U. Dinur, A.T. Hagler, New approaches to empirical force fields, in: K.B. Lipkowitz, D.B. Boyd (Eds.), Reviews in computational chemistry, vol. 2, VCH Publishers, New York, 1991, pp. 99–164.
[28] J.R. Maple, M.J. Hwang, T.P. Stockfish, U. Dinur, M. Waldman, C.S. Ewing, A.T. Hagler, Derivation of class II force fields. I. Methodology and quantum force field for the alkyl functional group and alkane molecules, J. Comput. Chem. 15 (1994) 162–182.
[29] Cerius^2 Life Sciences molecular simulation software, version 4.5, 2000, Accelrys Inc., San Diego, Calif.
[30] M.K. Gilson, B. Honig, The inclusion of electrostatic hydration energies in molecular mechanics calculations, J. Comput. Aided Mol. Des. 5 (1991) 5–20.
[31] W. Rocchia, S. Sridharan, A. Nicholls, E. Alexov, A. Chiabrera, B. Honig, Rapid grid-based construction of the molecular surface and the use of induced surface charge to calculate reaction field energies: applications to the molecular systems and geometric objects, J. Comput. Chem. 23 (2002) 128–137.
[32] C.J.F. Böttcher, Theory of Electric Polarization, Elsevier Press, Amsterdam, 1973.
[33] S. Miertus, E. Scrocco, J. Tomasi, Electrostatic interaction of a solute with a continuum. A direct utilization of ab initio molecular potentials for the prevision of solvent effects, Chem. Phys. 55 (1981) 117–129.
[34] V. Frecer, S. Miertus, Polarizable continuum model of salvation for biopolymers, Int. J. Quant. Chem. 42 (1992) 1449–1468.
[35] S. Fischer, J.C. Smith, C. Verma, Dissecting the vibrational entropy change on protein/ligand binding: burial of a water molecule in bovine pancreatic trypsin inhibitor, J. Phys. Chem. B 105 (2001) 8050–8055.
[36] S.M. Schwarzl, T.B. Tschopp, J.C. Smith, S. Fischer, Can the calculation of ligand binding free energies be improved with continuum solvent electrostatics and an ideal-gas entropy correction? J. Comput. Chem. 23 (2002) 1143–1149.
[37] QikProp, version 2.0, 2002, Schrödinger, Inc., New York, NY.
[38] E.M. Duffy, W.L. Jorgensen, Prediction of properties from simulations: free energies of solvation in hexadecane, octanol, and water, J. Am. Chem. Soc. 122 (2000) 2878–2888.
[39] W.L. Jorgensen, E.M. Duffy, Prediction of drug solubility from Monte Carlo simulations, Bioorg. Med. Chem. Lett. 10 (2000) 1155–1158.
[40] W.L. Jorgensen, E.M. Duffy, Prediction of drug solubility from structure, Adv. Drug Deliv. Rev. 54 (2002) 355–366.
[41] C.A. Lipinski, Drug-like properties and the causes of poor solubility and poor permeability, J. Pharmacol. Toxicol. 44 (2000) 235–249.
[42] C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, Adv. Drug Deliv. Rev. 46 (2001) 3–26.
[43] A. Tossi, F. Benedetti, S. Norbedo, D. Skrbec, F. Berti, D. Romeo, Small hydroxyethylene-based peptidomimetics inhibiting both HIV-1 and C. albicans aspartic proteases, Bioorg. Med. Chem. 11 (2003) 4719–4727.
[44] D.J. Kempf, H.L. Sham, HIV Protease Inhibitors, Curr. Pharm. Design 2 (1996) 225–246.
[45] F. Darvas, G. Keseru, A. Papp, G. Dorman, L. Urge, P. Krajcsi, In silico and ex silico ADME approaches for drug discovery, Curr. Top. Med. Chem. 2 (2002) 1287–1304.
[46] H.L. Sham, C. Zhao, L. Li, D.A. Betebenner, A. Saldivar, S. Vasavanonda, D.J. Kempf, J.J. Plattner, D.W. Norbeck, Novel lopinavir analogues incorporating non-aromatic P-1 side chains-synthesis and structure-activity relationships, Bioorg. Med. Chem. Lett. 12 (2002) 3101–3103.
[47] C. Parolin, B. Gatto, C. Del Vecchio, T. Pecere, E. Tramontano, V. Cecchetti, A. Fravolini, S. Masiero, M. Palombo, G. Palu, New anti-human immunodeficiency virus type 1 6-aminoquinolones: mechanism of action, Antimicrob. Agents. Chemother. 47 (2003) 889–896.
[48] L. Rocheblave, F. Bihel, C. De Michelis, G. Priem, J. Courcambeck, B. Bonnet, J.C. Chermann, J.L. Kraus, Synthesis and antiviral activity of new anti-HIV amprenavir bioisosteres, J. Med. Chem. 45 (2002) 3321–3324.
[49] J.Y. Lee, S.N. Kim, C.K. Lee, H. Park, Y.S. Lee, 1,4-Dioxane-fused 4-anilinoquinazoline as inhibitors of epidermal growth factor receptor kinase, Arch. Pharm. (Weinheim) 334 (2001) 255–257.