PCCP

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

PCCP

PAPER

Theoretical studies of energetics and binding isotope effects of binding a triazole-based inhibitor to HIV-1 reverse transcriptase

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

A. Krzemińska,[a] K. P. Świderek[a,b] and P. Paneth[a*]

Understanding of protein-ligands interactions is crucial for rational drug design. Binding isotope effects, BIEs, can provide intimate details of specific interactions between individual atoms of an inhibitor and the binding pocket. We have applied multi-scale QM/MM simulations to evaluate binding energetics of novel triazole-based non-nucleoside inhibitor of HIV-1 reverse transcriptase and to calculate associated BIEs. The binding sites can be distinguished based on ^{18}O-BIE.

Introduction

Human immunodeficiency virus (HIV-1) is a cytopathic human retrovirus, that primarily infects CD+ helper T lymphocytes leading to the loss of their number and function. The progressive failure of the immune system associated with loss of predominantly T lymphocyte and other cells critical to this system e.g. macrophages or dendritic cells, leads to acquired immune deficiency syndrome (AIDS). Progressive disappearance of immunity weakens the body, allows opportunistic infections to evolve, leads to cancers and ultimately to premature death.[1] Nowadays, treatment of HIV-1 infected patients relies on cocktails of different drugs. Due to undesired side-effects, especially occurring during long-term treatments, there has been considerable interest in developing new inhibitors targeting three HIV-1 enzymes: integrase (IN), protease (PR) and reverse transcriptase (RT).[2] Herein we focus on an inhibitor of this latter enzyme.

HIV-1 RT is an asymmetric heterodimer, composed of p66 and p51 subunits, which catalyses transcription of single-strand viral RNA into viral DNA double helix.[3] The enzyme contains one DNA polymerase active site and one ribonuclease H (RNase H) active site, both of which reside in the p66 subunit at spatially distinct regions while the smaller p51 chain is responsible for structural scaffolding. The hydrophilic polymerase active site is located between palm (light blue), fingers (blue) and thumb (cyan) subdomains (see Fig.1) and is made of Asp110, Asp185 and Asp186 amino acids, that can coordinate two Mg^{2+} or two Mn^{2+} cations.[4] The allosteric cavity, composed by Lys103, Tyr181, Tyr188, Phe227, Trp229, Leu234 and Tyr318 residues[5] (highlighted in Fig.1) lies approximately 10 Å from the polymerase active site. The RNase H active site comprising Asp443, Asp498 and Asp549 amino acids and two bound Mg^{2+} cations[6] is located 60 Å from the allosteric pocket. The two active sites and the allosteric cavity constitute a key target of antiviral research. Nucleoside reverse transcriptase inhibitors (NRTIs) inhibit polymerase active site and need to be phosphorylated by cellular kinases to active triphosphate form, which competes with natural dNTPs and finally becomes incorporated into the growing primer. Nucleotide reverse transcriptase inhibitors (NtRTIs), in contrast to NRTIs, are structures already equipped with phosphonate group and after the phosphorylation step are able to terminate DNA synthesis also by the competitive mechanism.[7] Non-nucleoside RT inhibitors (NNRTIs), on the other hand, follow non-competitive mechanism of inhibition. They are considered to be highly specific compounds that bind tightly in the allosteric pocket, acting on the enzyme subunits as a molecular wedge.[8,9] Simultaneous inhibition by combining nucleoside and non-nucleoside inhibitors is also currently explored.[10,11] Up to date eight NRTIs, two NtRTIs and five NNRTIs have been approved by FDA.[12] In presented studies we focus on a ligand that can bind either in allosteric cavity or RNase H active site of HIV-1 RT and compare it to ligands designed specifically for one of these sites (structural formulas of all ligands are presented in Supplementary Information Fig.S1).

Accurate description of binding affinity of an inhibitor to an enzyme is still a challenging task.[13,14] Nevertheless, using Monte Carlo simulation, Jorgensen and co-workers, estimated binding affinities of several NNRTIs.[15-18] Binding energies of NNRTIs have also been studied by MM-PBSA[19] and quantum calculations[20] including the ONIOM method.[21] Warshel and co-workers obtained the absolute binding free energies of NNRTIs by PDLD/S-LRA/b method.[22] Our own experience, first validated on the FDA approved drugs, indicate that binding free energies of HIV-1 RT complexes with NNRTIs and NRTIs can be efficiently obtained using a thermodynamic cycle from alchemical free energy perturbation (FEP) calculations.[23,24]

Recently, Schramm and co-workers showed that binding isotope effects, BIEs,[25] exhibit different values upon binding of inhibitor, para-aminobenzoic acid, to an enzyme (dihydropteroate synthase) from different sources (*Staphylococcus aureus and Plasmodium*

[a] *Institute of Applied Radiation Chemistry, Lodz University of Technology, 90-924 Lodz, Poland. E-mail: paneth@p.lodz.pl*
[b] *Department de Quimica Fisica i Analitica, Universitat Jaume I, 12071 Castello, Spain.*

† Electronic Supplementary Information (ESI) available: structural formulas for the NRTIs and NNRTIs, computed values of BIEs, the average of key distances for L-1 bounded in allosteric cavity and in RNase H active site, evolution of key distances during dynamic simulation for L-1 bounded in allosteric cavity and in RNase H active site. See DOI: 10.1039/x0xx00000x

falciparum). In this communication we extend these findings to identification of the binding site within a single enzyme using BIEs on the example of a novel potent HIV-1 RT triazole-based inhibitor.

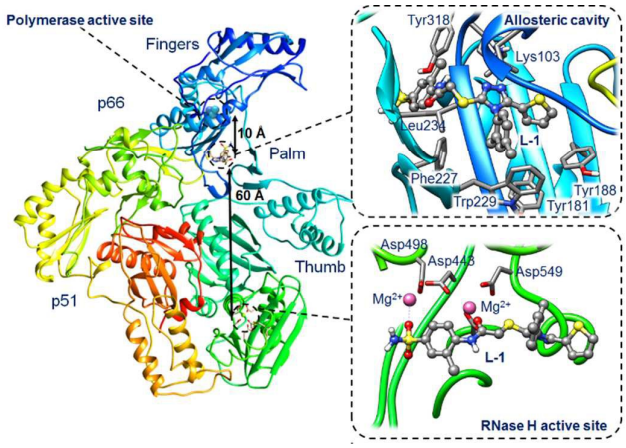

Fig. 1 Representation of HIV-1 RT with two binding sites, allosteric cavity (upper right box) and RNase H active site (lower right box) with bounded L-1 inhibitor.

Computational modelling and methods

System setup

Novel triazole derivative N-(2-chloro-4-sulphamoylphenyl)-2-((4-(2,4-dimethyl-phenyl)-5-(thiophen-2-yl)-4H-1,2,4-triazol-3yl)sulphanyl)-acetamide, **L-1**,[26] that was found to be a potent NNRTI has been docked to the allosteric cavity and RNase H active site of HIV-1 RT based on the crystallographic structure from Protein Data Bank with PDB ID 2RKI[27] using Glide package as implemented in Schrödinger program.[28] Both complexes are depicted in Fig.1. The 2RKI structure was crystalized with 4-benzyl-3-[(2-chlorobenzyl)sulfanyl]-5-thiophen-2-yl-4H-1,2,4-triazole mole-molecule bounded in allosteric cavity. Native ligand was removed from 2RKI and L-1 was docked to the allosteric cavity and to the active site using Glide XP.[29] The pK_a for titratable amino acids was calculated using PROPKA3.0.[30-32] Assuming physiological pH value, hydrogen atoms were added using tLEAP module of AMBER program.[33] Neutralization of L-1-allosteric cavity and L-1-RNase H active site complexes was completed by added 15 and 9 Cl^- counterions, respectively. Finally, both complexes were placed in orthorhombic boxes of TIP3P[34] water molecules of 144x160x144 Å^3. Parameters for L-1 were obtained from GAFF[35] as implemented in AMBER Tools. Afterwards, several optimization and dynamics simulations were performed using AMBER ff03[36] implemented in NAMD[37] program with time step 1 fs and periodic boundary conditions using the particle mesh Ewald method.[38] Cut-off for nonbonding interactions were applied using a smooth switching function with a range radius from 14.5 to 16 Å. First energy optimizations were carried out by means of conjugated gradient algorithm. The systems were heated by increasing temperature from 0 to 300 K with 0.001 K increment. The equilibration of the systems was achieved during 300 ps of Langevin-Verlet dynamics at 300 K. Then 2 ns of NVT MM MD were carried out. Subsequently, 200 ps for L-1-cavity and 1 ns for L-1-RNase H complexes were simulated using QM/MM MD simulations at AM1[39]/AMBER:TIP3P theory level as implemented in fDynamo.[40] All atoms beyond 20 Å from L-1 were kept frozen. In order to describe binding energy of L-1 to HIV-1 RT, a reference system of L-1 in an aqueous solution of the equivalent size (144x160x144 Å^3) was simulated during 200 ps of AM1/TIP3P dynamics.

Alchemical Free Energy Perturbation method (FEP)

The enzyme-ligand binding affinity is a main goal of drug design. Among different methods developed for this purpose, the alchemical free energy perturbation (FEP) combined with QM/MM molecular dynamics, while not providing detailed information about the full binding process, delivers reliable values of binding free energy, $\Delta\Delta G_{bind}$, at a reasonable computational cost. The FEP method by means of thermodynamic cycle was first introduced by Warshel and co-workers.[41] Herein, a series of QM/MM MD have been carried out at 300 K in the NVT ensemble with two parameters λ and γ, which corresponds to electrostatic and Van der Waals interactions respectively:

$$\boldsymbol{E_{QM/MM}(\lambda,\gamma)} = \langle \boldsymbol{\Psi}|\widehat{\boldsymbol{H_0}}|\boldsymbol{\Psi}\rangle + \boldsymbol{\lambda}\left(\sum\left\langle \boldsymbol{\Psi}\left|\frac{\boldsymbol{q_{MM}}}{\boldsymbol{r_{e,MM}}}\right|\boldsymbol{\Psi}\right\rangle + \sum\sum\frac{\boldsymbol{Z_{QM}q_{MM}}}{\boldsymbol{r_{QM,MM}}}\right) + \boldsymbol{\gamma E^{vdW}_{QM/MM}} + \boldsymbol{E_{MM}} \quad (1)$$

Both, λ and γ were smoothly changed from 0 to 1 with 0.02 increments. According to our own experience with HIV-1 RT FEP calculations in each window 5 ps relaxation followed by 100 ps QM/MM MD need to be carried out and a simplified thermodynamic cycle, depicted in Fig.2, can be used to obtain $\Delta\Delta G_{bind}$ based on equation 2.

$$\Delta\Delta G_{bind} = \Delta G_W - \Delta G_E = \Delta\left(G^{elec}_{QM/MM} + G^{vdW}_{QM/MM}\right)_W - \Delta\left(G^{elec}_{QM/MM} + G^{vdW}_{QM/MM}\right)_E \quad (2)$$

The potential energy

In order to describe individual interactions between allosteric cavity amino acids and the triazole molecule, the analysis of averaged potential energy have been performed through QM/MM MD based on equation 3

$$\boldsymbol{E^{INT}_{QM/MM}} = \boldsymbol{E^{elect}_{QM/MM}} + \boldsymbol{E^{vdW}_{QM/MM}} = \sum_{MM}\left[\left\langle \boldsymbol{\Psi}\left|\frac{\boldsymbol{q_{MM}}}{\boldsymbol{r_{e,MM}}}\right|\boldsymbol{\Psi}\right\rangle + \sum_{QM}\frac{\boldsymbol{Z_{QM}q_{MM}}}{\boldsymbol{r_{QM,MM}}}\right] + \sum\sum\boldsymbol{4\varepsilon_{QM,MM}}\left[\left(\frac{\boldsymbol{\sigma_{QM,MM}}}{\boldsymbol{r_{QM,MM}}}\right)^{\boldsymbol{12}} - \left(\frac{\boldsymbol{\sigma_{QM,MM}}}{\boldsymbol{r_{QM,MM}}}\right)^{\boldsymbol{6}}\right] \quad (3)$$

where $\boldsymbol{E^{INT}_{QM/MM}}$ is composed of electrostatic, $\boldsymbol{E^{elect}_{QM/MM}}$, and van der Waals, $\boldsymbol{E^{vdW}_{QM/MM}}$, interaction terms as a function of interaction energy between QM and MM parts. The polarized QM subsystem include the coulombic interaction of the QM nuclei (Z_{QM}) and the electrostatic interaction of the polarized electronic wave-function (Ψ) with enzyme amino acids charges (q_{MM}), while the MM part brings non-polarizable potential contribution.

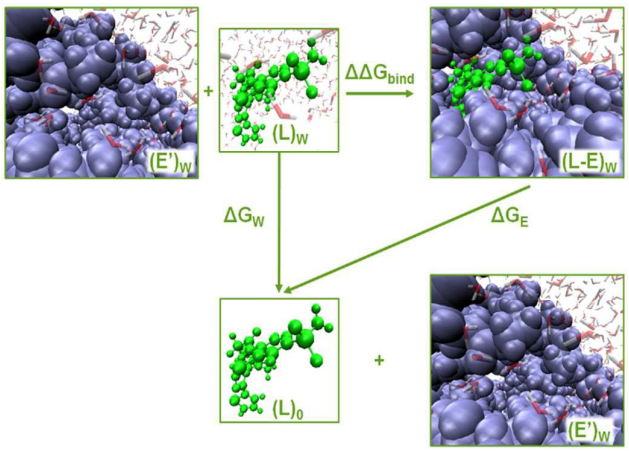

Fig. 2 Thermodynamic cycle to compute enzyme-ligand binding free energies from the FEP method, were enzyme (E) with ligand (L) in its binding site, E' is apo form of enzyme, $(L)_0$ is the ligand in the gas phase.

Molecular Electrostatic Potential (MEP)

The Molecular Electrostatic Potential, MEP, describe interaction energy between the charge distribution of a molecule and a unit positive charge. This method allows predicting nuclear and electronic charge distribution of a given molecule. Despite the fact that polarization is not included, charge distribution remains unperturbed by the external test charge, the MEP is still effective for interpretation and prediction of chemical reactivity.[42] Three-dimensional MEP surfaces for the L-1 extracted from allosteric cavity and RNase H active site were obtained at the B3LYP/6-31+G(d,p) theory level with Gaussian09.[43] Negative electronic potential corresponds to the most nucleophilic regions (red), while positive electronic potential indicate the most electrophilic regions (blue).

Binding Isotope Effects (BIEs)

Binding isotope effects can provide information on conformation changes and ligand-enzyme interactions.[44] In combination with the experimental values they can also be used to determine the actual binding site. We have already successfully used BIEs to analyse interactions between ligands and HIV-1 RT.[24] BIEs can be expressed by the binding free energy of the light, L, and heavy, H, isotopologs considering the relationship between binding free energy and the equilibrium constant:

$$BIE = \frac{\left(\frac{Q_e}{Q_w}\right)_L}{\left(\frac{Q_e}{Q_w}\right)_H} e^{-(\Delta ZPE_L - \Delta ZPE_H)/RT} \quad (4)$$

where the total partition function Q_e, for the ligand in the enzyme-ligand complex, and Q_w, for the ligand in the ligand-water complex, are computed as products of the translational, rotational and vibrational partition functions. The different contributions were separately computed using the Born-Oppenheimer, rigid rotor and harmonic oscillator approximations. Due to the fact that in each case the ligand is in a condense phase, rotational and translational motions are replaced by six librational motions of the ligand within its external environment. However, we decided to subject the full 3N × 3N Hessians for the ligand to a projection procedure to eliminate translational and rotational components, which give rise to small non-zero frequencies, as previously described.[44,45] Thus, it has been assumed that the 3N - 6 vibrational degrees of freedom are separable from the 6 translational and rotational degrees of freedom of the substrate (contrary to ref.[46]). Although the environment (protein or water) surrounding the ligand affects the values of each Hessian, there is no coupling between the ligand and its environment. Eleven snapshots from last 20 ps of QM/MM MD calculations were extracted for L-1-water, L-1-allosteric cavity and L-1-RNase H active site complexes. Combination of 121 BIEs individual values obtained with eleven enzyme-ligand and eleven ligand-water (11 x 11) structures allowed for calculations of BIEs with their uncertainties. All formalisms used for BIEs calculations yielded practically same values (see Tables S1 and S2 in the Supplementary Information).

Results and discussion

The obtained total free energies of binding inhibitors to HIV-1 RT binding sites are collected in Table 1. More extensive comparison is presented in Fig.3. There are some difficulties in validation of the $\Delta\Delta G_{bind}$ using experimental data. Our experience shows that IC_{50} measurements are not suitable for this purpose[23,24] in contrast to the reciprocal of the dissociation constant (K_I) of the protein-ligand complex.[47] Unfortunately, no systematic study of K_I values of different NNRTIs measured at the same experimental conditions can be found in literature. The evaluation of our FEP calculations was made by using NRTIs[23] and NNRTIs[24] drugs approved by FDA, where structures of NRTIs correspond to 2,7-dihydroxy-4-1(methylethyl)-2,4,6-cycloheptatrien-1-one (**L_{NA}-1**), 3-cyclopentyl-1,4-dihydroxy-1,8-naphthyridin-2(1H)-one (**L_{NA}-2**), ethyl 1,4-di-hydroxy-2-oxo-1,2-dihydro-1,8-naphthyridine-3-carboxylate (**L_{NA}-3**), and 3-[4-(diethylamino)phenoxy]-6-(ethoxycarbonyl)-5,8-dihydroxy-7-oxo-7,8-dihydro-1,8-naphthyridin-1-ium (**L_{NA}-4**), while structures of NNRTIs correspond to nevirapine (**L_{NN}-1**), efavirenz (**L_{NN}-2**), rilpivirine (**L_{NN}-3**), etravirine (**L_{NN}-4**), and delavirine (**L_{NN}-5**), for more details see Fig.S1 in Supplementary Information. In a hydrophobic pocket van de Waals interactions are the most favourable, as observed for NNRTIs. Van der Waals interactions are responsible for strong (-29.2 kcal·mol^{-1}) binding of L-1 in the allosteric cavity. It is worth noticing that electrostatic interactions between L-1 and both water molecules and the allosteric site have the same ΔG_{elect} value what indicates that electrostatic term does not influence protein-ligand interactions. In this case the magnitude of the total energy of binding of L-1 to allosteric cavity is thus solely due to the van der Waals term (see the green bars in Fig.3), which is preferable in the hydrophobic pocket. Usually, van der Waals term is much smaller than electrostatic term, hence relative $\Delta\Delta G_{vdW}$ value obtained for L-1 is the highest among NNRTIs. This makes L-1 one of the most potent NNRTIs. Binding of L-1 to RNase H active site is due to preferable electrostatic forces with $\Delta\Delta G_{bind}$ equal to -36.5 kcalmol^{-1}.

PCCP

PAPER

Table 1 The total free energy of binding (kcal·mol^{-1}), $\Delta\Delta G_{bind}$, inhibitors to HIV-1 RT binding sites obtained from FEP method. ΔG_{elect} and ΔG_{vdW} components correspond to electrostatic and van der Waals terms, respectively, aq indicate ligand in aqueous solution while HIV in enzymatic environment.

inhibitor	cavity	$\Delta G_{elect\text{-}aq}$	$\Delta G_{elect\text{-}HIV}$	$\Delta\Delta G_{elect}$	$\Delta G_{vdW\text{-}aq}$	$\Delta G_{vdW\text{-}HIV}$	$\Delta\Delta G_{vdW}$	$\Delta\Delta G_{bind}$
L-1	RNase H	-33.6	-56.2	-22.5	-37.2	-51.2	-14.0	-36.5
L-1	Allosteric	-33.6	-33.6	0.0	-37.2	-66.4	-29.2	-29.2
L_{NA}-3	RNase H	-8.2	-37.7	-29.5	-23.1	-23.8	-0.7	-30.2
L_{NN}-3	Allosteric	-7.1	-14.9	-7.8	-29.5	-50.9	-21.4	-29.2

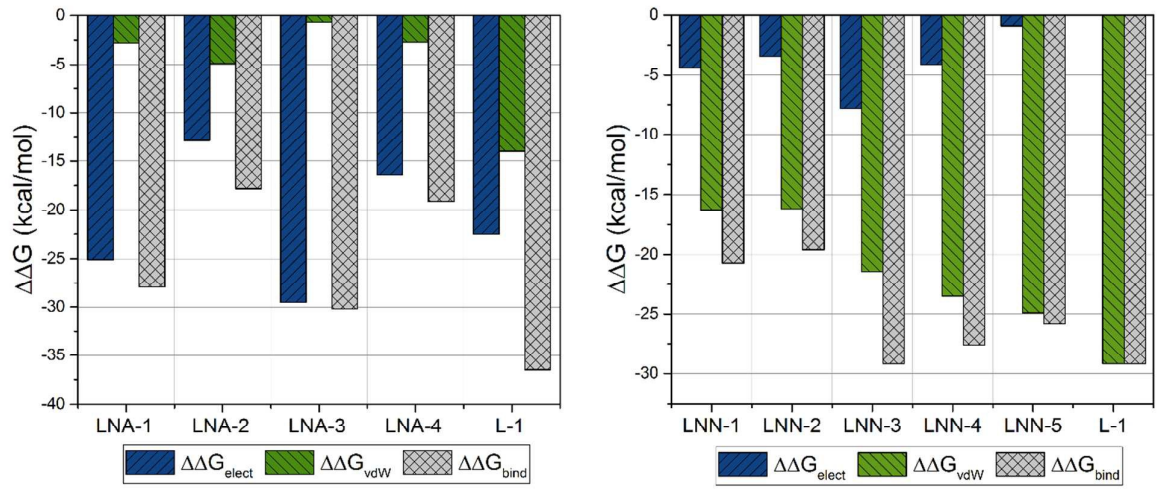

Fig. 3 Contributions of electrostatic and van der Waals terms to total binding free energy obtained from FEP calculations.

There is a meaningful difference between L-1 interactions in aqueous solution and enzyme environment in both electrostatic and Van der Waals terms. NRTIs show the same behaviour in case of electrostatic interactions, but opposite in van der Waals interactions. Consequently, binding L-1 to RNase H site is accompanied by both strong electrostatic and van der Waals terms, and indicates that L-1 can successfully bind into the RNase H active site. The potential energy of electrostatic and van der Waals interactions between L-1 and allosteric cavity amino acids has been obtained along the QM/MM dynamics; averaged results are displayed in Fig.4. Negative values correspond to favourable interactions and stabilize enzyme-ligand complex, while positive values represent unfavourable interactions. The strength of interactions depends on the nature of the ligand (size, stereochemistry, number and type of functional groups) and differs among NNRTIs. The crystallographic structures[5,48] indicate that Leu100, Lys101, Lys103, Val106, Thr107, Val108, Val179, Tyr181, Tyr188, Val189, Gly190, Phe227, Trp229, Leu234, Tyr318 all from chain A and Glu138 from chain B are usually responsible for inhibitor-enzyme interactions. The L-1-allosteric cavity complex has meaningful interactions with Leu100, Lys103, Lys104, Val106 and Pro236 as shown in Fig.4. L-1 interacts also with Phe227, His235 and Tyr318. Residues Lys101, Thr107, Val108, Val189, Gly190, Trp229, Leu234, and Tyr318 are too far, while Val179, Tyr181, Tyr188, and Glu138 from chain B exercise negligible bending affinity toward L-1. The main contribution of favourable electrostatic interactions, shown in Fig.4 and Fig.5, are observed between 1,2-triazole nitrogen atoms and Lys103. However, these interactions do not enhance binding affinities, because interactions of L-1 with hydrogen atoms of solvent molecules fully compensate them. Significantly weaker are favourable electrostatic interactions with Lys104 and Tyr318, wherein carbonyl oxygen from Lys104 interaction with the hydrogen atom from the aromatic ring is further stabilized by the π-stacking with Tyr318. These positive electrostatic interactions are compensated by electrostatic repulsion of Asp192. The determinant factor of L-1-RT attraction comes from the van der Waals interactions, not present in water, especially with Val106, Leu100, Pro236, and His235. Residues

Tyr181 and Tyr188 are stabilizing one aromatic ring of L-1 with two methyl groups and Phe227 interacts by π-stacking with another aromatic ring neighbouring the sulphate group. The electrostatic positive and negative terms are equal, therefore the binding process is a result of van der Waals interactions between L-1 and the allosteric cavity; the potential energy of van der Waals interactions is sufficient for incorporating and maintaining L-1 in this hydrophobic pocket.

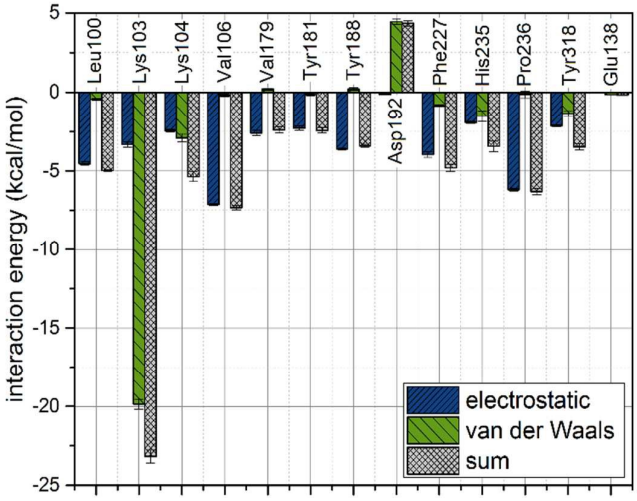

Fig. 4 Analysis of potential energy of interactions between L-1 and allosteric cavity amino acids.

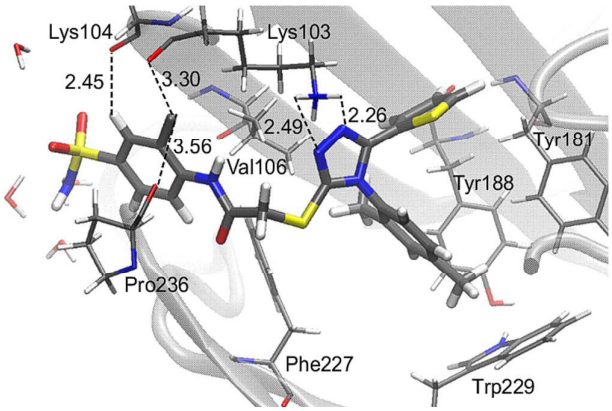

Fig. 5 Arrangement of L-1 at the allosteric cavity with average of key distances in Å.

In contrast to L-1-allosteric cavity complex, the electrostatic interactions between L-1 and RNase H active site are crucial to hold ligand in the active site centre. RNase H active site is made of Asp443, Glu478, Asp498 and Asp549 aminoacids, which coordinate two Mg^{2+}. The magnesium cation tends to create six coordination bonds but in the active site usually only five are occupied. Hence, two cations can interact with ligands by two relatively strong coordination bonds.

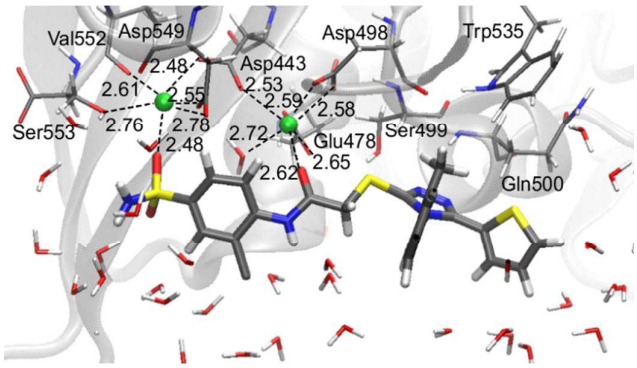

Fig. 6 Arrangement of L-1 at the RNase H active site with average key distances in Å. Magnesium cations highlighted as a green balls.

Small ligands, like L_{NA}-1,[23] interact with both Mg^{2+} without reorganizing active site. Opposite situation is observed for bigger molecules such as L_{NA}-4 and L-1. Both alter the RNase H active site in the same way, causing increase of the distance between magnesium cations and forcing coordination of Val552 and Ser553 by one Mg^{2+}. The open and flexible character of the RNase H region supports reorganization of the active site. This dynamic effect, observed along MM MD, supports binding of a bigger ligand such as L-1.

Recently, effective medical treatment of HIV infection is focused on composite approach combining different targets (enzymes) of the virus as well as different sites within a single enzyme. It is assumed that inhibitors of different types bind to different sites of the enzyme. However, recent reports on finding several alternative binding sites within HIV-1 RT shed new light on the actual binding sites of particular inhibitors.[49] In order to analyse specificity of the triazole-based inhibitors we have performed calculations of the binding energetics of L-1 in two binding spots; the allosteric centre and the RNase H active site. After QM/MM MD simulations we have extracted the last structure of L-1 from L-1-cavity and L-1-RNaseH complexes, and calculated three-dimensional MEP surfaces. Obtained MEP surfaces are shown in Fig.7, where the most nucleophilic regions are in red (-0.1 a.u.) and the most electrophilic regions are in blue (+0.1 a.u.). Both structures have similar distribution of nuclear and electronic charges. As expected, the triazole ring and the sulphate group exhibit negative electronic potential. Sulphate group is dangling almost outside of the hydrophobic pocket, thus it can interact with water molecules. In L-1-RNase H complex, the active site is open to aqueous environment and also in this case sulphate group can interact with solvent molecules. The L-1 nitrogen atoms are deep inside the allosteric cavity and in the presence of Lys103 they act as a nucleophile, while in the RNase H active site they interact electrostatically with Asp498, Gln500 and Tyr501. Consequently, there is no big difference between charge distribution of L-1 extracted from an allosteric cavity and from RNase H active site. These results indicate that L-1 can bind to both binding sites.

Another analysis of interactions of L-1 in complexes with allosteric cavity and RNase H active site has been based on heavy atoms (^{13}C, ^{15}N, ^{18}O, ^{37}Cl and ^{34}S) BIEs. Calculated values are collected in Fig.7. BIEs can provide information on differences in interactions between L-1 dissolved in water and bounded in RT binding sites. Despite the dissimilarity in molecular character of hydrophobic cavity and hydrophilic RNase H active site, in both RT binding sites lack of BIEs for most of atoms has been observed. The largest isotope effects were obtained for the ^{18}O of the sulphate group and apparently this isotope effect can be experimentally used to distinguish the place in which L-1 is bounded; normal ^{18}O-BIE (1.016 ± 0.002) is expected when L-1 binds in the allosteric cavity, where oxygen atom interacts (see Fig.4) with the hydrogen atom from Val106 (the distance is 4.23 Å) and is also surrounded by three water molecules. On the other hand inverse BIE (0.988 ± 0.002) for sulphate group ^{18}O atoms is expected for the complex of L-1 with RNase H, originating in strong interactions with one of the active site Mg^{2+} cations (the distance is 2.41 Å) and additionally from stabilization by hydrogen atom (the distance is also 2.41 Å) from the methyl group of Ala445 and one water molecule. Isotope effect differentiation of the other oxygen of the sulphonyl group is not possible as it is only slightly inverse (0.997 ± 0.001) when L-1 is bounded in the RNase H active site in contrast to lack of isotope effect of this oxygen atom in the allosteric cavity. This slightly inverse value of BIE originates in the interaction with the second Mg^{2+} cation (the distance is 2.56 Å). Although electrostatic interaction between nitrogen atom of the triazole ring and the hydrogen atom from Lys103 (the distance is 2.52 Å) is unfavourable in allosteric cavity, while in the RNase H active site electrostatic interactions of this atom with Gln500 and Tyr501 hydrogen atoms tend to stabilize the complex (the distances are 2.55 Å and 2.10 Å respectively) there is no reflection of these differences in the isotopic binding patterns with both ^{15}N-BIEs being slightly inverse (0.997 ± 0.001) and equal. Finally, negligible normal BIE (1.002 ± 0.002) was calculated for ^{37}Cl-BIE when L-1 binds in the RNase H active site, and similar value was obtained when L-1 is complexed in the allosteric cavity. Although it should be kept in mind that equilibrium isotope effects, such as BIEs, are generally small and hard to measure experimentally, the above analysis illustrates that it is possible to distinguish binding sites based on their values in favourable cases.

Conclusions

The novel inhibitor of HIV-1 RT, L-1, has been proposed and has been examined based on theoretical simulations. The L-1 molecule was docked to both RT binding sites: allosteric cavity and RNase H active site. Final complexes of L-1-binding site have been obtained after long QM/MM dynamics. Once the stability of both systems was confirmed, the interactions between L-1 and RT binding sites have been quantified by free energy of binding computed with FEP method. The analysis of this affinity has been also enriched by potential energies calculations, three-dimensional MEP isosurfaces, structural analysis and heavy atoms (^{13}C, ^{15}N, ^{18}O, ^{37}Cl and ^{34}S) BIEs. Obtained results confirmed differences between allosteric cavity and RNase H binding sites, as well as dissimilarity in the ligand binding pathways. In allosteric cavity the inhibitor binds by van der Waals affinities, which are preferable in this hydrophobic pocket, while the electrostatic interactions are unfavourable. The opposite situation was observed for binding ligand to RNase H active site, where binding occurs by electrostatic forces.

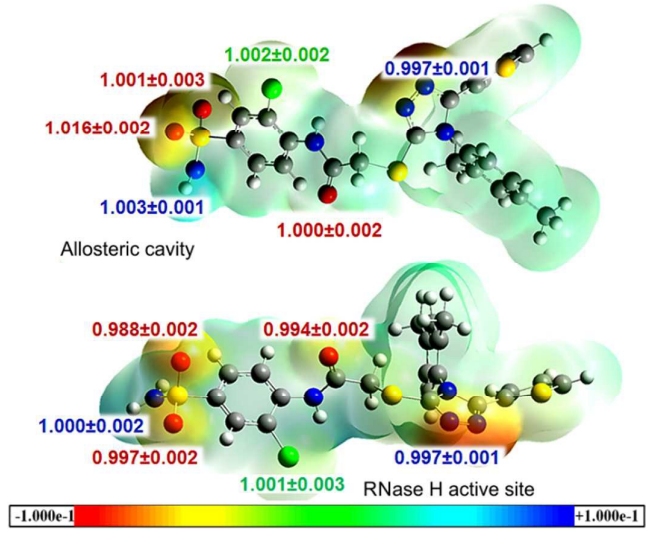

Fig. 7 Three-dimensional MEP surfaces for the L-1 extracted from allosteric cavity and RNase H active site with heavy atoms BIEs. Negative electronic potential corresponds to the most nucleophilic regions (red), while positive electronic potential indicate the most electrophilic regions (blue).

Inhibition is accompanied with release of standard free energy, $\Delta\Delta G_{bind}$. The values of $\Delta\Delta G_{bind}$ were obtained by alchemical FEP method. Comparison of different NNRTIs, NRTIs and L-1 bounded in allosteric cavity and RNase H active site shows that L-1 has the largest negative standard free energy of binding, indicating that L-1 is able to successfully inhibit both allosteric cavity and RNase H sites. Also three-dimensional MEP isosurfaces display show that there is no large difference between L-1 charge distribution in allosteric cavity and RNase H active site.

BIEs values are almost identical for both L-1-enzyme complexes, with one exception. In contrast to L-1-allosteric cavity complex, where a normal (larger than unity) ^{18}O-BIE is expected, in the L-1-RNase H active site an inverse BIE for two oxygens from sulphate group is predicted. This occurs due to the strong interactions with two Mg^{2+} originally belonging to RNase H active site. Thus BIEs allow distinguishing to which site L-1 binds. Experimental determination of this ^{18}O-BIE is being currently carried out in our laboratory.

Acknowledgements

This work has been supported by the grants 2011/02/A ST4/00246 (Maestro) from the Polish National Research Center (NCN) and

0478/IP3/2015/73 (Iuventus Plus) from the Polish Ministry of Science and Higher Education. The authors also acknowledge the Servei d´Informatica, Universitat Jaume I for generous allotment of the computer time.

References

1 R. A. Weiss, *Science*, 1993, **260**, 1273-1279.
2 J. A. Moss, *Radiol. Technol.*, 2013, **84**, 247-267.
3 J. M. Coffin, S. H. Hughes, H. E. Varmus, *Retroviruses,* Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press, 1997.
4 B. A. Larder, D. J. Purifoy, K. L. Powell, G. Darby, *Nature*, 1987, **327**, 716-717.
5 J. Ding, K. Das, Y. Hsiou, S. G. Sarafianos, A. D. Jr Clark, A. Jacobo-Molina, C. Tantillo, S. H. Hughes, E. Arnold, *J. Mol. Biol.*, 1998, **284**, 1095-1111.
6 E. Rosta, M. Nowotny, W. Yang, G. Hummer, *J. Am. Chem. Soc.*, 2011, **23**, 8934-8941.
7 Y. Mehellou, E. D. Clercq, *J. Med. Chem*., 2010, **2**, 521-538.
8 F. Esposito, A. Corona, E. Tramontano, *Mol. Biol. Int.*, 2012, ID **586401**, 23 pages.
9 S.G. Sarafianos, B. Marchand, K. Das, D. M. Himmel, M. A. Parniak, S. H. Hughes, E. Arnold, *J. Mol. Biol.,* 2009, **385**, 693–713.
10 K. Das, S.E. Martinez, J.D. Bauman, E. Arnold, *Nature Struct. Mol. Biol.*, 2012, **19**, 253-259.
11 R. Meleddu, S. Distinto, A. Corona, G. Bianco, V. Cannas, F. Esposito, A. Artese, S. Alcaro, P. Matyus, D. Bogdan, F. Cottiglia, E. Tramontano, E. Maccioni, *Eur. J. Med. Chem.*, 2015, **93**, 452-460.
12 www.fda.gov
13 A. R. Leach, B. K. Shoichet, C. E. Peishoff, *J. Med. Chem.*, 2006, **49**, 5851-5855.
14 C. J. Woods, M. Malaisree, S. Hannongbua, A. J. Mulholland, *J. Chem. Phys.,* 2011, **134**, 054114.
15 R. H. Jr. Smith, W. L. Jorgensen, J. Tirado-Rives, M. L. Lamb, P. A. Janssen, C. J. Michejda, M. B. Kroeger Smith, *J. Med. Chem.*, **41**, 1998, 5272-5286.
16 R. C. Rizzo, J. Tirado-Rives, W. L. Jorgensen, *J. Med. Chem.*, 2001, **44**, 145-154.
17 W. L. Jorgensen, M. Bollini, V. V. Thakur, R. A. Domaoal, K. A. Spasov, K. A. Anderson, *J. Am. Chem. Soc.*, 2011, **133**, 15686-15696.
18 D. J. Cole, J. Tirado-Rives, W. L. Jorgensen, *Biochim. Biophys. Acta.*, 2015, **1850**, 966-971.
19 J. Wang, P. Morin, W. Wang, P. A. Kollman, *J. Am. Chem. Soc.,* 2001, **123**, 5221-5230.
20 X. He, Y. Mei, Y. Xiang, Da W. Zhang, J. Z. H. Zhang., *PROTEINS: Struct, Func. And Bioinf.*, 2005, **61**, 423-432.
21 S. Saen-oon, M. Kuno, S. Hannongbua, *PROTEINS, Struct. Func. And Bioinf.*, 2005, **61**, 859-869.
22 N. Singh, A. Warshel, *PROTEINS, Struct. Func. And Bioinf.*, 2010, **78**, 1705-1723.
23 K. Świderek, A. Martí, V. Moliner, *Phys. Chem. Chem. Phys.*, 2012, **14**, 12614-12624.
24 A. Krzemińska, P. Paneth, V. Moliner, K. Świderek, *J. Phys. Chem. B*, 2015, **119**, 917-927.
25 C. F. Stratton, H. A. Namanja-Magliano, S. A. Cameron, V. L. Schramm, *ACS Chemical Biology,* **ASAP**, DOI: 10.1021/acschembio.5b00490.
26 T. Fraczek, A. Paneth, R. Kaminski, A. Krakowiak, P. Paneth, *J. Enzym. Inhib. Med. Chem.*, 2015, Epub ahead of print. DOI: 10.3109/14756366.2015.1039531.
27 T. A. Kirschberg, M. Balakrishnan, W. Huang, R. Hluhanich, N. Kutty, A. C. Liclican, D. J. McColl, N. H. Squires, E. B. Lansdon, *Bioorg. Med. Chem. Lett.*, 2008, **18**, 1131-1134.
28 R. A. Friesner, R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren, P. C. Sanschagrin, D. T. Mainz, *J. Med. Chem.,* 2006, **49**, 6177-6196.
29 T. Fraczek, A. Siwek, P. Paneth, *J. Chem. Inf. Model*., 2013, **53**, 3326-3342.
30 H. Li, A. D. Robertson, J. H. Jensen, *PROTEINS: Struct, Func. And Bioinf.,* 2005, **61**, 704-721.
31 D. C. Bas, D. M. Rogers, J. H. Jensen, *PROTEINS: Struct, Func. And Bioinf.,* 2008, **73**, 765-783.
32 M. H. M. Olsson, C. R. Sndergard, M. Rostkowski, J. H. Jensen, *J. Chem. Theory Comput.*, 2011, **7**, 525-537.
33 D. A. Case, T. A. Darden, T. E. Cheatham III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A. W. Goetz, I. Kolossva´ry, K. F. Wong, F. Paesani, J. Vanicek, R. M. Wolf, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J. Wang, M.-J. Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko and P. A. Kollman, AMBER 12, University of California, San Francisco, 2012.
34 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys*., 1983, **79**, 926-935.
35 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem*., 2004, **25**, 1157-1174.
36 Y. Duan, C. Wu, S. Chowdhury, M. C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang, P. Kollman, *J. Comput. Chem.,* 2003, **24**, 1999-2012.
37 J. C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R. D. Skeel, L. Kale, K. Schulten*, J. Comput. Chem.,* 2005, **26**, 1781-1802.
38 T. Darden, D. York, L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089-10092.
39 M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, J. J. P. Stewart, J. *Am. Chem. Soc.,* 1985, **107**, 3902-3909.
40 M. J. Field, M. Albe, C. Bret, F. Proust-De Martin, A. Thomas, *J. Comput. Chem*., 2000, **21**, 1088-1100.
41 A. Warshel, F. Sussman, G. King, *Biochem.*, 1986, **25**, 8368-8372.
42 C. K. Bagdassarian, V. L. Schramm, S. D. Schwartz, *J. Am. Chem. Soc*., 1996, **118**, 8825-8836.
43 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O¨. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, D. J. Fox, Gaussian 09, Revision D.01, Gaussian, Inc., Wallingford, CT, 2009.
44 K. Świderek, P. Paneth, *Chem.Rev.*, 2013, **113**, 7851-7879.
45 G. D. Ruggiero, S. J. Guy, S. Martí, V. Moliner, I. H. Williams, J. Phys. Org. Chem., 2004, 17, 592-601.
46 J. J. Ruiz-Pernía, I. H. Williams, *Chem. Eur. J*. 2012, **18**, 9405-9414.
47 Y. Cheng, W. H. Prusoff, *Biochem. Pharmacol*. 1973, **22**, 3099-3108.

48 L. A. Kohlstaedt, J. Wang, J. M. Friedman, P. A. Rice, T. A. Steitz, *Science*, 1992, **256**, 1783–1790.
49 J. D. Bauman, D. Patel, C. Dharia, M. W. Fromer, S. Ahmed, Y. Frenkel, R. S. K. Vijayan, J. T. Eck, W. C. Ho, K. Das, A. J. Shatkin, E. Arnold, *J. Med. Chem*., 2013, **56**, 2738-2746.