ON THE RELIABILITY OF THE CORRECTED SEMIEMPIRICAL QUANTUM CHEMICAL METHOD (PM6-DH2) FOR ASSIGNING THE PROTONATION STATES IN HIV-1 PROTEASE/INHIBITOR COMPLEXES

Adam PECINA[1,#], Ondřej PŘENOSIL[2,#], Jindřich FANFRLÍK[3], Jan ŘEZÁČ[4], Jaroslav GRANATIER[5], Pavel HOBZA[6,*] and Martin LEPŠÍK[7,*]

Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, v.v.i., and Center for Biomolecules and Complex Molecular Systems, 166 10 Prague 6, Czech Republic; e-mail: [1] pecina@uochb.cas.cz, [2] ondrej.prenosil@uochb.cas.cz, [3] jindrich.fanfrlik@uochb.cas.cz, [4] rezac@uochb.cas.cz, [5] j.granatier@gmail.com, [6] hobza@uochb.cas.cz, [7] lepsik@uochb.cas.cz

Received February 3, 2011
Accepted March 11, 2011
Published online April 20, 2011

Dedicated to Dr. Zdeněk Havlas on the occasion of his 60th birthday.

A novel computational protocol for determining the most probable protonation states in protein/ligand complexes is presented. The method consists in treating large parts of the enzyme using the corrected semiempirical quantum chemical (QM) method – PM6-D2 for optimization and PM6-DH2 for single-point energies – while the rest is calculated using molecular mechanics (MM) within a hybrid QM/MM fashion. The surrounding solvent is approximated by an implicit model. This approach is applied to two model systems, two different carboxylate pairs in one general and one unique HIV-1 protease/inhibitor complex. The effect of the size of the movable QM part is investigated in a series of several sizes, 3-, 6-, 8- and 10-Å regions surrounding the inhibitor. For the smallest region (< 450 atoms) the computationally more costly DFT QM/MM optimizations are performed as a check of the correctness. Proton transfer (PT) phenomena occur at both the PM6-D2 and DFT levels, which underlines the requirement for a QM approach. The barriers of PT are checked in model carboxylic acid pairs using the highly accurate MP2 and CCSD(T) values. An important result of this study is the fine-tuning of the protocol which can be used in further applications; its limitations are also shown, pointing to future developments. The calculations reveal which protonation variants of the active site are the most stable. In conclusion, the presented protocol can also be utilized for defining probable isomers in biomolecular systems. It can also serve as a preparatory step for further interaction-energy and binding-score calculations.

Keywords: HIV-1 protease inhibition; Protonation; QM/MM calculations; Semiempirical quantum chemical method; Proton transfer; Drug design; Inhibitors; X-ray crystallography.

[#] Both authors contributed equally to the work.

Collect. Czech. Chem. Commun. 2011, Vol. 76, No. 5, pp. 457–479

doi:10.1135/cccc2011035

Protonation and thus also the charge of amino-acid residues in proteins are defined by their pK_a values and pH [1]. The determination of the pK_a values of titratable residues in enzyme active sites is thus a key prerequisite for a molecular understanding of the reaction mechanisms and inhibition. Indeed, it has been shown that the calculated protein/ligand interaction energy is sensitive to the protonation state of the active site[2]. However, owing to the interactions between protonation sites, the protein titration curves may deviate from the standard Henderson–Haselbalch curves[3,4]. A number of experimental and computational approaches have been devised to determine the pK_a values of amino-acid side chains in proteins. A few of them related to this work are mentioned here. A classical experimental approach is based on measuring the pH dependence of a reaction[5]. Atomic resolution (R < 1.1 Å) X-ray crystal structures have also been used to infer the protonation states of titratable residues in enzyme active sites[6,7]. From the theoretical side, electrostatic pK_a calculations[4,8], quantum-mechanics (QM)-based calculations[9–11] or force-field molecular dynamics simulations[12,13] have been utilized to determine the protonation states.

HIV protease (PR) is one of the most intensively studied pharmaceutical targets. Its C_2-symmetrical dimeric structure features two catalytic aspartates (Asp25/Asp25′) in its active site (Fig. 1). These two carboxyl moieties are coplanar and so close to each other that one Asp has a shifted pK_a of ~6 while the other stays at ~3.5 [14]. As a result of this, one proton connects this Asp dyad of an unliganded PR via a double-well low-barrier hydrogen bond[15]. Bidirectional proton hopping between the two aspartates in this system has been simulated in an ab initio molecular dynamics study using a six-residue fragment of the active site[16]. In complexes with inhibitors, the catalytic Asp dyad of PR is monoprotonated in most cases (inhibitors featuring hydroxyl isostere; ref.[17] and references therein) and less frequently diprotonated (statine-based inhibitors)[6,18].

Proton transfer (PT) is one of the most important quantum effects, which are, by definition, not covered by empirical force fields. Quantum mechanical calculations, in contrast, inherently describe not only PT but also other quantum effects like charge redistribution, electron transfer or halogen bonding. The QM methods are thus the proper tool to use in order to determine the protonation states in the active sites of proteins where these phenomena might be important. However, because of the high computational costs, usually only a few residues in the active site could be treated[11,19]. Recent progress in the development of linear-scaling semiempirical quantum chemical (SQM) methods has offered the possibility to treat the whole biomolecular system containing several thousand atoms. However, the ac-

curacy of such methods was quite low by QM standards. Therefore, corrected versions have recently been introduced.

In our laboratory, we have chosen the novel SQM method PM6 with a parametrization for 70 elements[20], which is very well suited for the modeling of protein/ligand complexes, thanks to among others its linear-scaling algorithm MOZYME[21]. To increase further its accuracy for noncovalent binding, we have corrected this method with dispersion and hydrogen-bonding corrections (PM6-DH2) to reproduce closely benchmark CCSD(T) data[22,23]. As the solvent effects influence the biomolecular structures and energies, we use an implicit solvent of the COSMO[24] or generalized Born type[25] around the proteins while some important explicit water molecules from the crystal structure may be added. Hybrid QM/MM calculations using PM6 in a large QM part are also possible.

For this pilot study, we have chosen two model systems, one general and one unique HIV-1 PR/inhibitor complex. The inhibitors are: (i) the clinically successful nonpeptidic inhibitor darunavir (DRV, TMC-114, UIC-94017; K_i value of 5.3 pM)[26] and (ii) a phenylnorstatine-based peptidomimetic in-

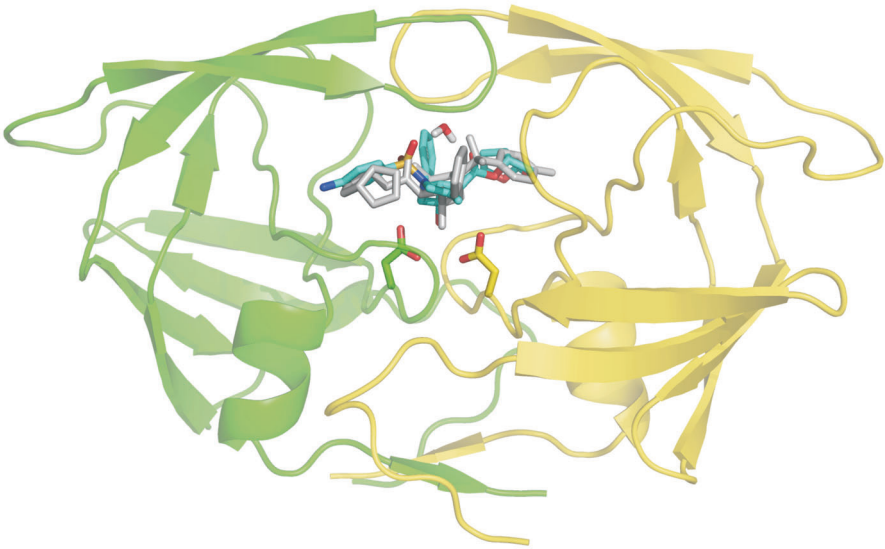
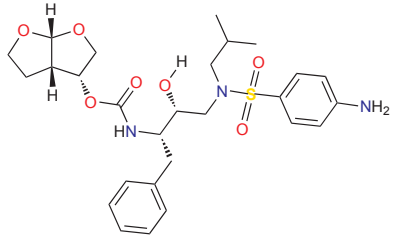


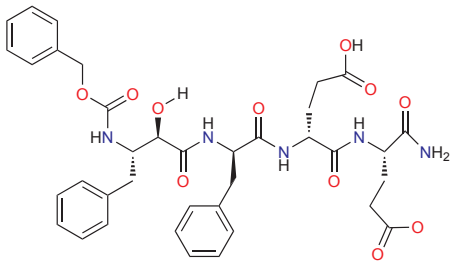
FIG. 1
The ribbon structure of the dimeric wild-type HIV-1 protease (PR) in complex with the darunavir (DRV) inhibitor (PDB code 3QOZ). One PR monomer is shown in green, the other in yellow. The two catalytic aspartates are depicted as sticks. Both of the crystallographic orientations of DRV are shown in cyan and grey. The structural flap water is shown with the hydrogens added. The oxygen atoms are in red, nitrogen atoms in blue, hydrogens in white

hibitor KI2 (K_i value of 180 pM)[27] (Fig. 2). The complex of wild-type (wt) PR with DRV crystallized in the hexagonal $P6_1$ space group yielded two orientations of the inhibitor[28] in the pseudo-C_2-symmetrical enzyme (the conformations of the A and B chain PR residues differ slightly owing to the binding of the asymmetrical inhibitor; Figs 1, 3A). Moreover, there are four possible variants of the monoprotonated catalytic Asp dyad for both inhibitor orientations (see the Methods).

The KI2 inhibitor formed a unique complex with the PR in which two molecules of KI2 bound to the enzyme; one was localized in the active site[6] and the other at the outer part of the PR, which allowed an atomic resolution of the crystal structure (1.03 Å)[29]. With such a high quality of the X-ray structure, the protonation state of the active site could be inferred from measuring the highly accurate CG-OD1/OD2 distances (Fig. 3B)[6]. Furthermore, the electron-density maps allowed a resolution of the P2 benzyloxycarbonyl group of KI2 to conformations (depicted in cyan and grey in Fig. 3B) with an alternative possibility of hydrogen bonding (H22 bonding inter-molecularly to the OD2 of the Asp25′ or intra-molecularly to the O01 of the KI2; Fig. 3B). In another part of the active site, another carboxyl–



DRV



KI2

FIG. 2
The structures of two potent protease inhibitors studied in this work

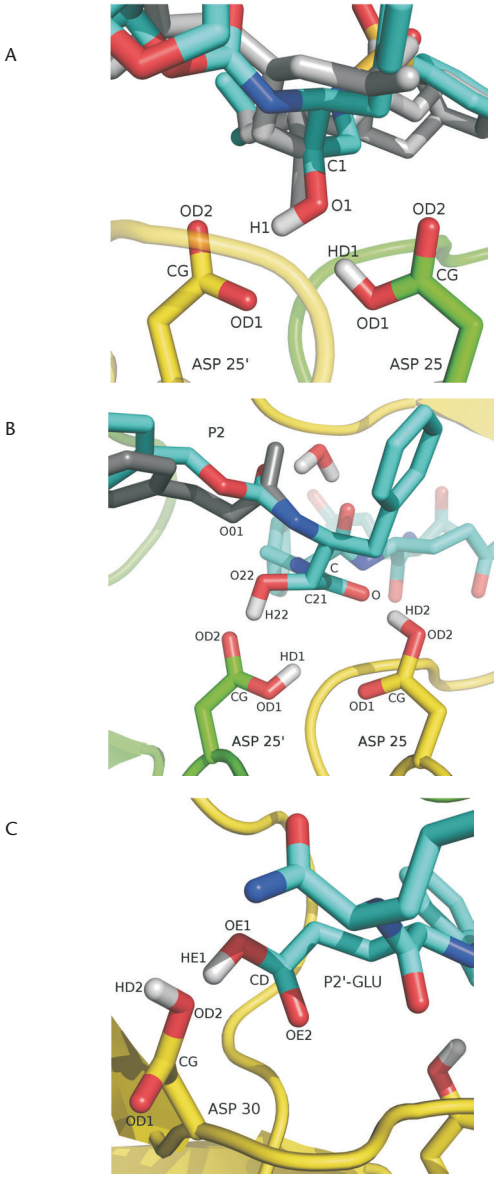


FIG. 3
Details of the active sites of protease/inhibitor complexes. The color coding is the same as in Fig. 1. The hydroxyl (O1–H1) of darunavir binding to a monoprotonated Asp25/25′ dyad of protease (A). The norstatine hydroxyl (O22–H22) and carbonyl (C–O) of KI2 binding to a diprotonated Asp25/25′ dyad of protease (B). The protonated Glu-P2′ of KI2 binding to Asp30 of protease (C)

carboxyl interaction was observed, namely between the side chain of the Asp30 of the PR and the P2′ glutamate moiety of the KI2 (Fig. 3C). The highly accurate CG-OD1/OD2 and CD-OE1/OE2 distances only reveal a preference for a protonation at the OD2 of the Asp30 and the OE1 of the Glu-P2′. This could be explained by either the fact that both of these oxygen atoms are protonated or the fact that a single proton connects them via

TABLE I
The protonation, orientation and conformation variants of the calculated structures

Variant	Inhibitor orientation	Active-site proton location
A. The PR/DRV complex		
D1	A	Asp25:OD1
D2	A	Asp25′:OD1
D3	A	Asp25:OD2
D4	A	Asp25′:OD2
D5	B	Asp25:OD1
D6	B	Asp25′:OD1
D7	B	Asp25:OD2
D8	B	Asp25′:OD2

Variant	Inhibitor chain/conformation	H22 hydrogen bond acceptor[a]	HE1/HD2 location in Asp30/Glu-P2′ pair[b]
B. The PR/KI2 complex			
K1	I/A	OD2	OD2
K2	I/A	OD2	OE1
K3	I/A	OD2	OD2, OE1
K4	I/B	OD2	OD2
K5	I/B	O01	OD2
K6	I/B	OD2	OE1
K7	I/B	O01	OE1
K8	I/B	OD2	OD2, OE1

[a] cf. Fig. 3B; [b] cf. Fig. 3C.

a hydrogen bond which can be either localized or mobile (low-barrier hydrogen bond). In order to shed light on the probability of these variants, we have constructed the respective molecular models and explored them computationally.

In this paper, several variants of the proton locations on the carboxylic moieties in question in both the wtPR/DRV and KI2 complexes have been investigated. Using QM/MM optimizations, these structures have been sorted by the relative energies of their QM parts; the lowest energy variants correspond to the most stable ones. The corrected PM6 method (PM6-D2 for optimization and PM6-DH2 for single-point energies) is used for the QM regions extending up to 10 Å from the inhibitors, while the results are checked using the DFT QM/MM optimizations on the 3-Å surroundings. The observed proton-transfer phenomena are further checked on small monoprotonated carboxylate pair models using high-level MP2 and CCSD(T) methods. In summary, we present a novel computational protocol which not only can be used for determining the protonation in the active site of the HIV PR but which also represents a general computational procedure enabling an objective decision on which potential isomers in a biomolecule/ligand complex will be populated.

METHODS

Systems Studied

The complexes of wild-type HIV-1 protease with two potent inhibitors, darunavir (DRV, TMC-114, UIC-94017)[26] and KI2 [27], were studied (Figs 1, 2). The current crystallographic structure of the wtPR/DRV complex (PDB code 3QOZ) represents a general case of the most common hydroxyl isostere inhibitors binding to the PR. The complex crystallized in the hexagonal $P6_1$ space group, which resulted in a superposition of two orientations of the inhibitor in the pseudo-C_2-symmetrical PR. It should be noted that there exist two other wtPR/DRV structures (PDB codes 2IEN [30] and 1T3R [31]) that crystallized in a less common orthorhombic $P2_12_12$ space group with a single inhibitor orientation. These structures superpose with our structure with a root-mean-square deviation (RMSD) of Cα atoms of the protein of 0.3 and 0.4 Å, respectively, showing their high similarity. The one proton present in the catalytic aspartate dyad can be placed on either the OD1 or OD2 atoms of either the Asp25 or Asp25′ (Fig. 3A). These four variants were studied for both of the inhibitor orientations (Table IA). The flap water was included in the model; the others were discarded.

The wtPR/KI2 crystal structure was determined at an atomic resolution of 1.03 Å (PDB code 1NH0), which allowed the deduction of the protonation state of the catalytic aspartates (Fig. 3B)[6]. Further in the binding cavity, another carboxyl–carboxyl interaction was observed, namely between the Asp30 of the PR and the P2′ glutamate moiety of the KI2 (Fig. 3C). The highly accurate CG-OD1/OD2 and CD-OE1/OE2 distances only revealed a preference for a protonation at the OD2 of the Asp30 or the OE1 of the Glu-P2′ or both (cf. the K1, K2 and K3 variants in Table IB). Furthermore, two conformations, A and B, of the inhibitor benzyloxycarbonyl group with relative occupancies of 54 and 46%, respectively, were fitted to the electron density maps (EDM). Several PR residues were also refined to alternate conformations – the active-site examples are the Asp 30′ (54:47), Val 32′ (55:45), Val 82/Val 82′ (65:35) or Ile 84/Ile 84′ (61:39). The pairs of the major (K1-K3 variants, Table IB) and minor (K4-K8) conformations of all of the PR residues with alternate conformations of the KI2 inhibitor were constructed. In the B conformation of the inhibitor, the acceptor of the H22 atom of the KI2 could be either the OD2 atom of the Asp25′ (K4, K6 and K8 variants) or the O01 atom of the KI2 (K5, K7). Taken together, a set of eight variants was prepared for the PR/KI2 complex (Table IB). The flap water was included in the model; the others were discarded.

System Setup

The structures with the protonation variants shown in Table I were prepared using a special protocol developed with the aim of enabling a comparison of the stabilities of differentially protonated structures. Thus the steps of hydrogen-atom addition and relaxation (see below) were performed only for the D1, K1 and K4 variants. Only then were the protonation states exchanged to include all of the variants and the position of the added proton was optimized.

Hydrogen atoms were added to D1, K1 and K4 variant structures using the UCSF Chimera program[32] for the ligand and the Reduce[33] and LEaP programs available in the AMBER 10 simulation package[34] for the protein. To mimic the pH of 5.0 and 5.6 used for the crystallization experiments of the wtPR/DRV [28] and wtPR/KI2 [6] complexes, respectively, the arginine, lysine and histidine residues as well as the N-termini were modeled as positively charged, whereas the aspartic and glutamic acid side-chains (with the exceptions of the Asp25/25′, Asp30 and the P2′-Glu of the KI2) as well as the C-termini were in their anionic forms. The DRV inhibitor was neutral and the KI2 was in a mono- or dianionic state depending on the protonation

variant. The positions of the added hydrogens were relaxed using the SANDER module of AMBER by up to 10,000 steps of the steepest descent, and conjugate gradient optimizations until the root-mean-square gradient fell below 0.001 kcal/mol/Å. The position of the flap water was then optimized using the same criteria as above, because in the case of the DRV, the crystallographic position of the flap water was an average of two slightly differing positions for each inhibitor orientation.

The parameters for these AMBER calculations were as follows: the ff03 force field[35] was used for the protein and the General AMBER force field (GAFF)[36] for the ligands. The charges for the ligand were obtained using a restrained fit to the electrostatic potential (RESP) calculated at the HF/6-31G* level[37].

QM/MM Setup

The hybrid QM/MM calculations were set up as follows: for large QM regions (6-, 8- and 10-Å surroundings of the inhibitor), the QM part was treated using the semiempirical quantum chemical PM6 method[20] corrected for dispersion (PM6-D2; optimization) and hydrogen bonding (PM6-DH2; single-point energies)[22,23]. The hydrogen-bond corrections could not be used for optimizations because of their inability to tackle PT. All of the PM6 calculations were performed using the MOZYME linear-scaling algorithm available in the MOPAC code[21]. The benchmark QM/MM calculations were performed on the smallest region of the 3.0- (DRV) or 2.5-Å (KI2) surroundings of the inhibitor using the calculations on the DFT and compared with the PM6 level. An RI-DFT-D approach (the accelerated resolution-of-the-identity variant[38] enhanced with empirical dispersion[39] was used with the TPSS/TZVP//B-LYP/SVP functional/basis set combination for single-point and optimization, respectively) using the Turbomole program, version 6.2 [40]. To accelerate the SCF convergence, a levelshift of 0.25 a.u. was applied. The QM part was calculated in vacuum. The MM part was treated using AMBER and the parameters listed above.

The coupling between the QM and MM parts was done with an in-house program using a subtractive scheme of an ONIOM-type[41]. The protein/inhibitor complex was surrounded by the generalized Born (GB)[25] implicit solvent model. To speed up the convergence of optimizations, the outer part of the protein was kept frozen.

RESULTS AND DISCUSSION

DFT and PM6 QM/MM Optimizations

The RI-DFT-D QM/MM calculations (the B-LYP/SVP optimizations followed by the TPSS/TZVP single-point energies) were used as a reference for both the PR/DRV and PR/KI2 complexes (eight variants for each, see Table I). Thus, if the structures and energies of the DFT and PM6 QM/MM calculations differed qualitatively, conclusions were drawn from the former ones. For the feasibility of the DFT calculations, the size of the QM parts was selected not to exceed 450 atoms. Thus, the 3-Å surroundings of the Asp25/Ap25′ pair for the PR/DRV case and the 2.5-Å surroundings of the Asp30/Glu-P2′ pair including the Asp25/Asp25′ dyad for the PR/KI2 case were chosen. The single-point energies of the QM part in vacuum were sorted with respect to the most stable one (Table II).

The RI-DFT-D calculations on the PR/DRV system have revealed that either orientation of the DRV in the complex (A and B) yields two stable structures (D1, D2, and D5, D6 variants; see Table IIA) within 3 kcal/mol (for a discussion of the energy cutoff, see ref.[42]). The other variants are 4.3–13.0 kcal/mol less stable. It is important to bear in mind that because of the C_2-pseudosymmetrical structure of the PR/DRV complex, the D1–D6, D2–D5, D3–D8 and D4–D7 pairs are symmetry equivalents as regards the protonation state relative to the inhibitor orientation.

In contrast to the DFT, in the PM6-D2 QM/MM optimizations four of the eight structures resulted in a proton transfer in the active site (which transformed D1 and D3 to D2 and D6 and D8 to D5; denoted with an asterisk, Table IIA). To check the validity of such an observation, we investigated the heights of the PT barriers given by these two methods on the model systems derived from the PR/DRV and PR/KI2 complexes and compared their values to the benchmark values at the RI-MP2 and CCSD(T) levels. The preliminary data confirmed the well-known tendency of the DFT generalized gradient approximation (GGA) functionals to underestimate the reaction barrier heights[43] but also showed an even greater underestimation using the PM6-D2 [44]. Taking the PT into account, two energetically best structures at the PM6-DH2//PM6-D2 level corresponded to the D5 and D2 variants, the symmetrically equivalent pair. These two structures were the most stable ones at the DFT level as well. The minor consequences of the inability of the PM6-DH2//PM6-D2 approach to localize also the D1–D6 pair as equally stable are discussed below. It is of interest that even after the PT occurred, the D1, D3, D5 and D6 variants remained by 5.4–7.1 kcal/mol less

stable than the D8* structure. A visual inspection of the optimized structures showed that these variants became trapped in the local minima, differing in the geometry of the active site from the D2 and D8* structures. In the next section, we have investigated whether allowing more relaxation in the more distant surroundings of the active site could help bring these structures to the global minimum.

In the second system, the PR/KI2 complex, we started by studying the mono- and diprotonated variants of the Asp30/Glu-P2′ pair. For the former case (all of the variants except for the K3 and K8), the DFT QM/MM optimizations resulted in a PT from the Glu-P2′ of the KI2 to the Asp30 PR residue (Table IIB). Nevertheless, the PM6-D2 QM/MM optimizations re-

TABLE II
The relative energies (kcal/mol) of the QM parts (3-Å surroundings for DRV, 2.5 Å for KI2) of the protonation variants on the QM/MM optimized geometries

Variant	DFT	PM6
A. The wtPR/DRV complex		
D1	2.7	6.7[a]
D2	2.0	2.6
D3	6.0	7.1[a]
D4	13.0	20.2
D5	0.0	5.4
D6	1.1	5.5[a]
D7	6.7	16.7
D8	4.3	0.0[a]
B. The wtPR/KI2		
K1	0.0	0.3
K2	0.1[b]	0.0
K4	10.5	8.5
K5	20.2	9.8
K6	10.1[b]	9.0
K7	19.3[b]	7.9

[a] Denotes a proton transfer which transforms the D1 and D3 structures to D2 and the D6 and D8 structures to D5. [b] Denotes a proton transfer which transforms the K2 structure to K1, K6 to K4 and K7 to K5

sulted in an intermediate structure in which the HE1 proton is localized between the two oxygen atoms of the Glu-P2′ and Asp30 with typical O–H distances of 1.2–1.3 Å. This again can be explained by the shape of the PT curve for the PM6-D2 method, which for this model system has an energy minimum at the intermediate positions, unlike the DFT. However, the preliminary data show that even the DFT curve differs from the MP2 and CCSD(T) ones in the details of the shape and energetics[44]. In the diprotonated case (K3 and K8 variants), a PT occurred in neither the DFT nor PM6-D2 QM/MM optimizations. A comparison of the mono- and diprotonated variants to the crystal structure revealed large deviations of the Asp30/Glu-P2′ pair for the diprotonated variants as opposed to the energetically most stable monoprotonated variants (RMSD of 0.89 vs 0.14 and 0.94 vs 0.19 Å, respectively). We thus conclude that the monoprotonated variant is going to be more probable than the diprotonated one.

Comparing the relative stabilities of the A and B conformations of the KI2 and the PR residues in the PR/KI2 complex, the PM6-D2 and DFT QM/MM optimizations consistently show that the former is more stable by roughly 10 or 10–20 kcal/mol, respectively (Table IIB). Although the energy difference is too high (several possible reasons are discussed below) to allow the population of the B conformations, this result qualitatively agrees with the higher occupation of 54–65% for the A conformations observed in the crystal structure[6].

Another structural detail investigated in the PR/KI2 complex was the hydrogen bond formed by the hydroxyl of the KI2 (O22–H22 atoms). In the intermolecular case, i.e. binding to the OD2 atom of the Asp25′ (K4, K6 and K8 variants), the atoms involved in this hydrogen bond did not undergo any sizeable movements. However, in the intra-molecular case, i.e. binding to the O01 oxygen of the KI2 (K5, K7 variants), the O01 oxygen moved away (in the direction of its position in the A conformation) to increase its distance to the O22 atom from 2.28 Å in the B-conformation of the X-ray structure to 3.01 Å in the DFT QM/MM optimized structure. This shift of the O01 oxygen suggests that this intramolecular hydrogen bond would not be stable.

Size of the QM Moving Region in the QM/MM Calculations

To investigate the effect of the size of the moving QM part on the determination of the most stable protonation states, a region comprising the 6-Å surroundings of the inhibitor in the PR/DRV complex was set up and extended to include the 8- and 10-Å surroundings. The respective sizes of

these QM regions were 967, 1325 and 1696 atoms for the PR/DRV system, and 1050, 1374 and 1733 atoms for the PR/KI2. It should be noted that such an extension in a molecular system is not a smooth one as charged groups can be included in the QM part upon its extension and affect substantially its electrostatics.

In the PR/DRV case, the results were qualitatively similar to those found on the small QM part of 3 Å (cf. Table IIA); the same four variants underwent PT (D1 and D3 to D2 and D6 and D8 to D5; denoted with an asterisk, Table IIIA). Again, considering the PT, the D2–D5 equivalent pair had the lowest energy, whereas the other variants were less stable by 15.9–39.0 kcal/mol (Table IIIA) as compared to 16.7–20.2 kcal/mol and 1.1–13.0 kcal/mol for the smallest 3-Å region in the PM6-DH2//PM6-D2 and DFT QM/MM calculations, respectively (cf. Table IIA). This comparison shows that even allowing large parts of the protein move does not alleviate the problem of trapping the unstable D4 and D7 variants in the local minima of higher energy.

A similar set of calculations has been conducted for eight variants (K1–K8) of the PR/KI2 complex. For the monoprotonated Glu-P2′/Asp30 pair, the relative energies of the QM part are shown in Table IIIB. The sizes of the 6- and 8-Å surroundings of the inhibitor are energetically consistent with the DFT and corrected PM6 QM/MM optimizations on a small 2.5-Å region in that the QM parts of the A conformations (K1, K2 variants) are by 13.9–20.8 kcal/mol more stable than the B conformations (K4–K7 variants). In the larger region of 10 Å, however, another variant (K5 variant) approached the stability of the most stable variant, K1. Due to the large size of the QM region (1733 atoms), we wanted to verify whether the energy differences stemmed from the differences in the active site because of the different protonation variants (that is the goal of our investigation) or whether some structurally unrelated changes occurred at more distant parts of the PR (these would be unwanted effects that we would wish to avoid). We therefore reoptimized the 10-Å region optimized geometries using a smaller 8-Å region and compared the relative single-point energies using the 8-Å region. The last column of Table IIIB shows that the K5 variant again became less stable, which suggests that unrelated structural changes in the farther (in this case 8–10 Å) region can significantly influence the energetics of the QM part. A visual inspection of the optimized geometries revealed, like in the PM6-DH2//PM6-D2 QM/MM calculations in the smallest 2.5-Å region, that in all the monoprotonated structures the HE1 proton ended in an intermediate position between the OE1 of the Glu-P2′ and the OD2 of the Asp30. This corresponds to the shift of the position of the mini-

mum on the PM6-D2 hypersurface (HE1–OE1 and HE1–OD2 distances of 1.2 and 1.3 Å, respectively) relative to DFT (1.6 and 1.1 Å)[44].

The diprotonated variant of the A and B models of the PR/KI2 complex (K3 and K8 variants, respectively) showed a large structural deviation from the crystal structure, during which the Asp30 side chain turned outwards

TABLE III
The relative energies (kcal/mol) of the QM parts for the protonation variants on the PM6-D2 QM/MM optimized geometries with a varying size of the moving QM part

Variant	6 Å	8 Å	10 Å
A. The wtPR/DRV complex			
D1	17.6[a]	3.4[a]	20.5[a]
D2	7.5	7.5	1.3
D3	0.0[a]	0.0[a]	21.8[a]
D4	28.0	23.0	16.4
D5	18.3	3.8	26.8
D6	19.8[a]	4.1[a]	0.0[a]
D7	35.3	33.4	39.0
D8	25.3[a]	2.9[a]	8.2[a]

Variant	6 Å	8 Å	10 Å	8 Å//10 Å[b]
B. The PR/KI2 complex				
K1	0.0	0.0	0.0	0.0
K2	3.7	2.9	5.7	6.7
K4	15.9	17.3	8.8	11.3
K5	13.9	18.9	0.9	10.5
K6	17.5	18.0	19.8	21.0
K7	20.8	20.7	14.8	20.3

[a] Denotes a proton transfer which transforms the D1 and D3 structures to D2 and the D6 and D8 structures to D5. [b] The 8 Å//10 Å column means an 8-Å region reoptimization and single-point energies on the 10-Å region optimized geometries.

from its original position. Table IV shows the RMSDs with respect to the X-ray structure of the non-hydrogen atoms of the mono- and diprotonated Asp30/Glu-P2′ dyad obtained using corrected PM6 QM/MM optimizations using different sizes of the QM part. The small values for the monoprotonated variants found consistently for both the A and B conformations (cf. Table IVA and IVB) suggest that either the models of the diprotonated variants are more sensitive to the used approximations such as the lack of explicit water molecules or that the Asp30/Glu-P2′ dyad is only singly protonated.

Methodological Issues

In this study, we present hybrid QM/MM calculations on a biomolecular system in which protons play a pivotal role. For two molecular complexes, a QM region of approximately 400 atoms has been chosen, including not only the protonated carboxylate pairs but also their close (~3-Å) surroundings. The QM part was treated with the DFT method, which has been used

TABLE IV
The root-mean-square deviations (Å) with respect to the crystal structure of the non-hydrogen atoms of the mono- and diprotonated Asp30/Glu-P2′ dyad obtained using QM/MM optimizations with different sizes of the QM parts

QM region, Å	K1:OD2 or OE1, monoprotonated	K3:OD2, OE1, diprotonated
A. The K1 and K3 variants		
2.5	0.14	1.45
6	0.24	0.89
8	0.21	1.05
10	0.24	0.93
QM region, Å	K4–K7:OD2 or OE1, monoprotonated	K8:OD2, OE1, diprotonated
B. The most stable of K4-K7 and K8 variants		
2.5	0.16	1.10
6	0.22	0.52
8	0.21	0.52
10	0.24	0.63

frequently in biomolecular QM/MM calculations[45–47]. Yet, even for the accelerated RI variant[38] of the DFT, this size currently represents the upper limit. However, it is important to test the effect of increasing the QM part further.

To be able to treat the biomolecular systems of several thousand atoms, we have turned to semiempirical methods and because of its superior performance we have chosen the PM6 method[20] with a linear-scaling algorithm[21]. However, its description of the noncovalent interactions had to be enhanced by introducing empirical corrections for dispersion and hydrogen-bonding[22,23]. The newly developed method has been successfully applied for two biomolecular systems[48,49].

The present study is the first one to study the protonation phenomena in biomolecules using the corrected PM6-DH2 method. It should be stressed that a QM approach is the only one to be used (in contrast to the MM methods) to describe correctly a molecular system in which PT phenomena can occur[11]. For the two molecular complexes studied here, a PT occurred in one of them (PR/KI2 case) on the DFT level, while in the other system (PR/DRV; D1–D8 variants) no PT did take place. The results from the corrected PM6-DH2 method differed qualitatively; in one system (PR/KI2), the proton ended between the two oxygen atoms, whereas in the second system (PR/DRV) a PT was observed. We have therefore conducted a preliminary study of the PT barrier heights on a model of monoprotonated carboxylate pairs using high-level computational chemistry methods. The results have not only confirmed a well-known tendency of DFT GGA functionals to underestimate the reaction barriers[43] but also showed an even greater underestimation on the PM6-D2 level and shifting of the minimum toward the intermediate positions of the proton between the two oxygens[44]. This finding thus points to the need for better corrections or even new reparametrizations of the PM6 method which would also describe PT. Moreover, our results show that the frequently used DFT calculations must be taken with caution and preferably checked against higher-level QM calculations.

Owing to recent developments in linear-scaling semiempirical quantum chemical methods, we have been able to increase the size of the QM part stepwise up to approximately 1733 atoms (10-Å region surrounding the ligand) and optimize it at the PM6-D2 level. Although there were quantitative differences, the most stable variants (D2, D5 and K1, taking PT into account; cf. Table III) were found consistently in the 2.5–3-, 6-, 8- and 10-Å regions. However, in the 10-Å region (and significantly more in 12-Å regions, not shown) optimizations, unrelated structural changes occurred far

from the active site that affected the relative stabilities. Poised between the Scylla of allowing sufficient relaxation and the Charybdis of avoiding structural changes far from the active site, we recommend an optimal size of the QM region for HIV protease studies of ~8-Å surrounding the ligand. We add that this size may differ for other protein/ligand systems, depending on the flexibility of the complex, the hydration of the active site, etc.

A small note regarding the preparation of the structures should be made here. We have endeavored to develop a protocol which would be useful not only for a comparison of various protonation variants in the HIV protease but more generally of possible constitutional isomers (tautomers, conformers) in complex biomolecular systems. We therefore urge that the variant structures be prepared carefully and consistently (cf. the hydrogen-atom addition and relaxation performed solely for representative structures, only then setting up the protonation variants) to eliminate unwanted geometrical and energy differences.

As mentioned, the current protocol can be utilized in a QM-based scoring of the HIV protease/ligand complexes to select the most probable protonation variants for further scoring calculations. Including the PT, the PM6-DH2//PM6-D2 method has correctly found the D2–D5 pair as the most stable (Table IIA). However, the D1–D6 pair, which was the second most stable on the DFT level, transformed due to PT into the former one. This was caused by the underestimation of the PT barriers on the PM6-D2 level, which is even greater than that of the DFT level as shown by comparison with high-level MP2 and CCSD(T) calculations. The less stable D3–D8 pair again transformed in the PM6-D2 QM/MM optimizations to the most stable D2–D5 pair owing to PT, indicating its low stability. Finally, the least stable D4–D7 pair on the DFT level was also the least stable on the PM6-DH2//PM6-D2 level. Taken together, a semi-quantitative agreement of the corrected PM6 energies with the DFT ones can be obtained in cases where an incorrect PT does not occur. This points to a need of further adjustment of the corrected PM6 protocol for biomolecular systems by either restraining the O–H bonds in question or introducing another reparametrization.

Biomolecular Findings

In order to draw conclusions for the two HIV protease/inhibitor complexes studied in this work, we must be aware of the strengths and weaknesses of the crystallographic structures and computational methods/protocols employed. As already mentioned, the wtPR/DRV complex crystallized in the

common hexagonal $P6_1$ space group, in which the electron-density maps (EDM) for the two orientations of the inhibitor in the pseudo-C_2-symmetrical enzyme overlapped. Some inaccuracies of the starting structure may hence stem from fitting the two inhibitor orientations into these EDMs. In contrast, the highly accurate X-ray structure of the PR/KI2 complex allowed an inference of the proton locations of the catalytic Asp25/25′ while providing hints for the Asp30/Glu-P2′ pair. The P2 benzyloxycarbonyl of the KI2 inhibitor, on the other hand, had a poor omit EDM, which was explained by its higher mobility and an alternative conformation[6].

As regards the accuracy of the PM6-DH2 method, it has been established on several datasets of noncovalently interacting model complexes that it performs equally well as the DFT-D within a chemical accuracy[22,23]. However, we should bear in mind that these values hold for equilibrium geometries obtained on accurate MP2 and CCSD(T) geometries[50]. For less accurate geometries (as for example the PM6-D2 level used here), the error will increase.

For the PR/DRV complex, which represents a general case of the HIV PR/inhibitor complexes in which two orientations of the inhibitor were refined, we have found that the symmetry-related pairs of the protonation variants are also energy-related. The structural similarities of the two inhibitor orientations in PR were acknowledged in analyses of PR/DRV X-ray structures[30]. However, in the calculations of HIV protease/inhibitor complexes, it has been a common practice to use only the first orientation of the inhibitor (see e.g. ref.[17] and the references therein or ref.[51]). To the best of our knowledge, we present in this paper the very first study to confirm that, if using a QM-relaxed region, only one inhibitor orientation is sufficient to correctly describe the energetics in the active site of HIV PR/inhibitor complexes. In the DFT QM/MM calculations, the D1–D6 and D2–D5 symmetry-related pairs proved to be the most stable, separated from the less stable pairs by 4.3–13.0 kcal/mol. In an atomic-resolution (1.1 Å) crystal structure of DRV in complex with the PR Val82Ala mutant, a streak of positive electron density in the omit map appeared, suggesting the location of a proton[30]. This finding presents an experimental verification of our approach, as this corresponds to our stable D2 variant. In a molecular mechanics-based study of the DRV and a related inhibitor amprenavir (APV) binding to the PR, several protonation variants were tried[51]; although an equivalent of the D2 variant (beware of the fact that the PR chain notation is reversed with respect to our study) had the most favorable interaction energy with the APV (see the Supporting Information to ref.[51]), another variant (an equivalent of D3) was chosen because of its structural

similarity to the crystal structure after a molecular dynamics run. However, force-field-based methods may be unreliable for a structural description of the active site of enzymes, in which quantum effects such as PT or charge redistribution can occur.

The X-ray structure of the PR/KI2 complex is unique in that the atomic resolution of 1.03 Å enabled the deduction of the protonation states of the catalytic aspartates[6]. However, despite the high quality of the crystal structure, three molecular features remained questionable: (i) the protonation state of the Asp30/Glu-P2′ carboxylate pair, (ii) the relative stabilities of the A and B conformations of the P2 group of KI2 and several PR residues, and (iii) the acceptor of the hydrogen bond from the KI2 hydroxyl. The DFT QM/MM calculations revealed that of the three possible variants of the Asp30/Glu-P2′ protonation, the diprotonated variants (K3 and K8) could be excluded based on geometrical criteria, whereas the inhibitor OE1 oxygen-protonated variants (K2, K6 and K7) transformed during optimizations into the respective Asp30:OD2 protonated variants of K1, K4 and K5. The higher stability of the K1, K4 and K5 variants was also corroborated by the preliminary high-level QM calculations on a model system derived from the Asp30/Glu-P2′ pair of this crystal structure. We have thus determined using our computations that the OD2 atom of Asp30, and not the OE1 of Glu-P2′of KI2, will be protonated in the PR/KI2 complex. This is an interesting and farther reaching conclusion since the Glu residue is present as the P2′ moiety not only in several inhibitors[27,52,53] but also in the substrate derived from the CA-p2 cleavage site[54,55].

The second molecular feature of the PR/KI2 complex which deserved attention was the stability of the major (A) and minor (B) conformations of the P2 moiety of the KI2 and surrounding PR residues. The DFT QM/MM calculations have identified the A conformation as more stable than B, which is in qualitative agreement with the higher occupancy of the former over the latter observed in the crystal structure. However, the energy difference of 10.1–20.2 kcal/mol (Table IIB) is too high to interpret the crystallographic occupancy ratio of 54:46. Several limitations of the presented computational approach as well as crystallographic issues can be responsible. We could envisage that allowing structural relaxation of the active site surroundings would bring the two alternative conformations closer in energy, but this possibility was disproved by the corrected PM6 QM/MM calculations in the larger regions (Table IIIB). The lack of the explicit description of the vibrational energy and the dynamics may be another reason for the high energy difference. Regarding the X-ray structure, the experimental electron density maps (EDM) reveal that the P2 moiety is quite

flexible (among others reflected by the presence of two alternative conformations) as compared to the rest of the inhibitor. Upon a closer inspection of the EDM of the PR/KI2 complex, we observed that whereas the A conformation of the P2 was fitted into a well-defined EDM, the alternative B conformation could be fitted in several ways. It may be that some of these possible alternative B conformations would be lower in energy than the one present in the crystal structure.

The third molecular feature was the identity of the acceptor oxygen for the hydrogen bond of the KI2 hydroxyl in the B conformation. It was consistently found, for both the inter- and intra-molecular variants, that the O01 oxygen deviated from its crystallographic position toward a position found in the A conformation. This can either be a proof that (i) the hydrogen-bond acceptor of the KI2 hydroxyl would rather be OD2 of Asp25′ than O01 of the KI2 or (ii) the B conformation of the P2 moiety present in the crystal may be less stable than a potential other alternative conformation which could be fitted into EDM (see above). In summary, even very high quality crystal structures, such as that of the PR/KI2 complex, pose some unknowns for computational chemists. However, these may be elucidated by means of calculations.

CONCLUSIONS

Based on the results of this pilot computational study, several methodological and biomolecular conclusions have been drawn.

1) We have presented a novel computational protocol for determining the probable protonation states based on the quantum mechanical energy. This approach is general and can be utilized for assessing the stabilities of various conformers/tautomers in biomolecular systems.

2) A comparison with the benchmark MP2 and CCSD(T) data on protonated carboxylate pair model systems revealed that the DFT using a GGA functional and even more the PM6-D2 underestimate the PT barriers.

3) The corrected PM6 QM/MM calculations using a QM region extending up to 3 Å from the inhibitor found the same stable protonation states in the two HIV protease complexes as DFT. The extension of the QM region from 3 to 8 Å gave the same qualitative picture on the corrected PM6 level.

4) Allowing relaxation of overly large regions in the QM part (>10 Å) increases the risk of distant unrelated structural changes occurring, which can affect the energetics of the active site.

5) The symmetry-related pairs of the HIV PR/inhibitor complexes with two orientations of the inhibitors are also energy-related. It has been shown here on the PR/DRV complex.

6) The Asp30/Glu-P2′ carboxylate pair is monoprotonated on the Asp30 as shown for the PR/KI2 complex by (i) the geometrical instability of the diprotonated variants and (ii) the PT in the QM/MM calculations, corroborated by PT transfer barriers in model systems obtained with high-level QM calculations. This finding has consequences for other HIV PR inhibitors and substrates containing a Glu moiety at P2′.

7) The acceptor of the hydrogen bond from the hydroxyl group of the KI2 is most probably the OD2 oxygen of the Asp25′. There is, however, a possibility that an intramolecular hydrogen bond could form transiently with a structure from a dynamic equilibrium of alternative P2 conformations.

8) The major A conformation of the KI2 and the surrounding PR residues is more stable than the B conformation. This agrees with the experimental crystallographic finding of its higher occupation.

We thank Dr. J. Brynda for the electron-density maps of wtPR/KI2, Dr. P. Řezáčová and Dr. K. Grantz-Šašková for releasing the unpublished wtPR/DRV structure prior to the publication. Additionally, we are indebted to the Reviewer 1 for the valuable comments and suggestions. This work was a part of Research Project No. Z40550506 of the Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic and was supported by Grants No. LC512 and No. 1M0508 from the Ministry of Education, Youth and Sports of the Czech Republic and the Grant No. P208/11/0295 from the Grant Agendy of the Czech Republic. The support of the Praemium Academiae, Academy of Sciences of the Czech Republic, awarded to P.H. in 2007, is also acknowledged.

REFERENCES

1. Bombarda E., Ullmann G. M.: *J. Phys. Chem. B* **2010**, *114*, 1994.
2. Tawa G. J., Topol I. A., Burt S. K., Erickson J. W.: *J. Am. Chem. Soc.* **1998**, *120*, 8856.
3. Bashford D., Karplus M.: *J. Phys. Chem.* **1991**, *95*, 9556.
4. Ullmann G. M.: *J. Phys. Chem. B* **2003**, *107*, 1263.
5. Berg J. M., Tymoczko J. L., Stryer L.: *Biochemistry*, 5th ed. W. H. Freeman, New York 2002.
6. Brynda J., Řezáčová P., Fábry M., Hořejší M., Stouračová R., Sedláček J., Souček M., Hradilek M., Lepšík M., Konvalinka J.: *J. Med. Chem.* **2004**, *47*, 2030.
7. Podjarny A., Cachau R. E., Schneider T., Van Zandt M., Joachimiak A.: *Cell. Mol. Life Sci.* **2004**, *61*, 763.
8. Trylska J., Antosiewicz J., Geller M., Hodge C. N., Klabe R. M., Head M. S., Gilson M. K.: *Protein Sci.* **1999**, *8*, 180.
9. Rajamani R., Reynolds C. H.: *J. Med. Chem.* **2004**, *47*, 5159.

10. Yu N., Hayik S. A., Wang B., Liao N., Reynolds C. H., Merz K. M.: *J. Chem. Theory Comput.* **2006**, *2*, 1057.
11. Piana S., Sebastiani D., Carloni P., Parrinello M.: *J. Am. Chem. Soc.* **2001**, *123*, 8730.
12. Harte W. E., Beveridge D. L.: *J. Am. Chem. Soc.* **1993**, *115*, 3883.
13. Park H., Lee S.: *J. Am. Chem. Soc.* **2003**, *125*, 16416.
14. Hyland L. J., Tomaszek T. A. J., Meek T. D.: *Biochemistry* **1991**, *30*, 8454.
15. Northrop D. B.: *Acc. Chem. Res.* **2001**, *34*, 790.
16. Piana S., Carloni P.: *Proteins Struct. Funct. Genet.* **2000**, *39*, 26.
17. Lepšík M., Kříž Z., Havlas Z.: *Proteins Struct. Funct. Bioinf.* **2004**, *57*, 279.
18. Velazquez-Campoy A., Luque I., Todd M. J., Milutinovich M., Kiso Y., Freire E.: *Protein Sci.* **2000**, *9*, 1801.
19. Sirois S., Proynov E. I., Truchon J. F., Tsoukas C. M., Salahub D. R.: *J. Comput. Chem.* **2003**, *24*, 1110.
20. Stewart J. J. P.: *J. Mol. Mod.* **2007**, *13*, 1173.
21. Stewart J. J. P.: *J. Mol. Mod.* **2009**, *15*, 765.
22. Řezáč J., Fanfrlík J., Salahub D., Hobza P.: *J. Chem. Theory Comput.* **2009**, *5*, 1749.
23. Korth M., Pitoňák M., Řezáč J., Hobza P.: *J. Chem. Theory Comput.* **2010**, *6*, 344.
24. Klamt A., Schuurmann G.: *J. Chem. Soc., Perkin Trans. 2* **1993**, 799.
25. Tsui V., Case D. A.: *Biopolymers* **2001**, *56*, 275.
26. Koh Y., Nakata H., Maeda K., Ogata H., Bilcer G., Devasamudram T., Kincaid J. F., Boross P., Wang Y. F., Ties Y. F., Volarath P., Gaddis L., Harrison R. W., Weber I. T., Ghosh A. K., Mitsuya H.: *Antimicrob. Agents Chemother.* **2003**, *47*, 3123.
27. Rinnová M., Hradilek M., Bařinka C., Weber J., Souček M., Vondrášek J., Klimkait T., Konvalinka J.: *Arch. Biochem. Biophys.* **2000**, *382*, 22.
28. Kožíšek M., Šašková K. G., Lepšík M., Brynda J., Konvalinka J., Řezáčová P.: Manuscript in preparation.
29. Brynda J., Řezáčová P., Fábry M., Hořejší M., Stouračová R., Souček M., Hradilek M., Konvalinka J., Sedláček J.: *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2004**, *60*, 1943.
30. Tie Y. F., Boross P. I., Wang Y. F., Gaddis L., Hussain A. K., Leshchenko S., Ghosh A. K., Louis J. M., Harrison R. W., Weber I. T.: *J. Mol. Biol.* **2004**, *338*, 341.
31. Surleraux D. L. N. G., Tahri A., Verschueren W. G., Pille G. M. E., de Kock H. A., Jonckers T. H. M, Peeters A., De Meyer S., Azijn H., Pauwels R., de Bethune M. P., King N. M., Prabu-Jeyabalan M., Schiffer C. A., Wigerinck P. B. T. P.: *J. Med. Chem.* **2005**, *48*, 1813.
32. Pettersen E. F., Goddard T. D., Huang C. C., Couch G. S., Greenblatt D. M., Meng E. C., Ferrin T. E.: *J. Comput. Chem.* **2004**, *25*, 1605.
33. Word J. M., Lovell S. C., Richardson J. S., Richardson D. C.: *J. Mol. Biol.* **1999**, *285*, 1735.
34. Case D. A., Darden T. A., Cheatham III T. E., Simmerling C. L., Wang J., Duke R. E., Luo R., Crowley M., Walker R. C., Zhang W., Merz K. M., Wang B., Hayik S., Roitberg A., Seabra G., Kolossváry I., Wong K. F., Paesani F., Vanicek J., Wu X., Brozell S. R., Steinbrecher T., Gohlke H., Yang L., Tan C., Mongan J., Hornak V., Cui G., Mathews D. H., Seetin M. G., Sagui C., Babin V., Kollman P. A.: *AMBER 10*. University of California, San Francisco 2008.
35. Duan Y., Wu C., Chowdhury S., Lee M. C., Xiong G. M., Zhang W., Yang R., Cieplak P., Luo R., Lee T., Caldwell J., Wang J. M., Kollman P.: *J. Comput. Chem.* **2003**, *24*, 1999.
36. Wang J. M., Cieplak P., Kollman P. A.: *J. Comput. Chem.* **2000**, *21*, 1049.
37. Bayly C. I., Cieplak P., Cornell W. D., Kollman P. A.: *J. Phys. Chem.* **1993**, *97*, 10269.

38. Feyereisen M., Fitzgerald G., Komornicki A.: *Chem. Phys. Lett.* **1993**, *208*, 359.
39. Jurečka P., Černý J., Hobza P., Salahub D. R.: *J. Comput. Chem.* **2007**, *28*, 555.
40. Ahlrichs R., Bär M., Häser M., Horn H., Kölmel C.: *Chem. Phys. Lett.* **1989**, *162*, 165.
41. a) Svensson M., Humbel S., Froese R. D. J., Matsubara T., Sieber S., Morokuma K.: *J. Phys. Chem.* **1996**, *100*, 19357; b) Dapprich S., Komaromi I., Byun K. S., Morokuma K., Frisch M. J.: *J. Mol. Struct. – THEOCHEM* **1999**, *461–462*, 1.
42. Fanfrlík J., Brynda J., Řezáč J., Hobza P., Lepšík M.: *J. Phys. Chem. B* **2008**, *112*, 15094.
43. Koch W., Holthausen M. C.: *A Chemist's Guide to Density Functional Theory*, 2nd ed. Wiley-VCH, Weinheim 2000.
44. Granatier J., Fanfrlík J., Hobza P., Lepšík M.: Unpublished results.
45. Banáš P., Jurečka P, Walter N. G., Šponer J., Otyepka M: *Methods* **2009**, *49*, 202.
46. Cui Q., Karplus M.: *J. Phys. Chem. B* **2002**, *106*, 1768.
47. Senn H. M., Thiel W.: *Angew. Chem. Int. Ed.* **2009**, *48*, 1198.
48. Fanfrlík J., Bronowska A. K., Řezáč J., Přenosil O., Konvalinka J., Hobza P.: *J. Phys. Chem. B* **2010**, *114*, 12666.
49. Dobeš P., Fanfrlík J., Řezáč J., Otyepka M., Hobza P.: *J. Comput.-Aided Mol. Des.* **2011**, *25*, 223.
50. Jurečka P., Šponer J., Černý J., Hobza P.: *Phys. Chem. Chem. Phys.* **2006,** *8*, 1985.
51. Hou T. J., Yu R.: *J. Med. Chem.* **2007**, *50*, 1177.
52. Urban J., Konvalinka J., Stehlíková J., Gregorova E., Majer P., Souček M., Andreansky M., Fábry M., Štrop P.: *FEBS Lett.* **1992**, *298*, 9.
53. Weber I. T., Wu J., Adomat J., Harrison R. W., Kimmel A. R., Wondrak E. M., Louis J. M.: *Eur. J. Biochem.* **1997**, *249*, 523.
54. Richards A. D., Phylip L. H., Farmerie W. G., Scarborough P. E., Alvarez A., Dunn B. M., Hirel P. H., Konvalinka J., Štrop P., Pavlíčková L., Kostka V., Kay J.: *J. Biol. Chem.* **1990**, *265*, 7733.
55. Prabu-Jeyabalan M., Nalivaika E., Schiffer C. A.: *J. Mol. Biol.* **2000**, *301*, 1207.