THE JOURNAL OF PHYSICAL CHEMISTRY B

ARTICLE

pubs.acs.org/JPCB

Semiempirical Quantum Mechanical Method PM6-DH2X Describes the Geometry and Energetics of CK2-Inhibitor Complexes Involving Halogen Bonds Well, While the Empirical Potential Fails

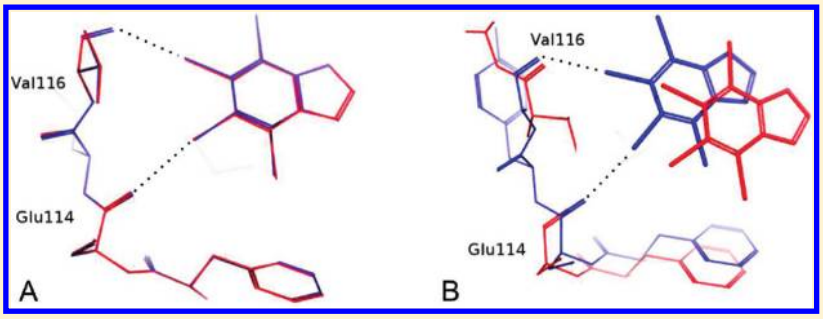
Petr Dobeš,[†,§] Jan Řezáč,[†] Jindřich Fanfrlík,[†] Michal Otyepka,[*,‡] and Pavel Hobza[*,†,‡]

[†]Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic and Center for Biomolecules and Complex Molecular Systems, 166 10 Prague, Czech Republic

[‡]Regional Centre of Advanced Technologies and Materials, Department of Physical Chemistry, Faculty of Science, Palacky University, 771 46 Olomouc, Czech Republic

[§]Center of Molecular Biology and Gene Therapy, Department of Internal Medicine — Hematooncology, University Hospital Brno, 625 00 Brno, Czech Republic

ABSTRACT:



In the present study, we have investigated complexes of CK2 protein kinase with halogenated inhibitors by means of the advanced semiempirical quantum mechanical (SQM) PM6 method (called PM6-DH2X), which describes various types of noncovalent interactions including halogen bonding well. The PM6-DH2X method provides reliable geometries of those CK2 protein kinase-inhibitor complexes involving halogen bonds that agree well with the X-ray crystal structures. When the Amber empirical potential is applied, this agreement becomes considerably worse. Similarly, the binding free energies determined by the PM6-DH2X SQM method are much closer to the experimental inhibition constants than those based on the Amber empirical potential.

INTRODUCTION

Protein kinase CK2 (formerly known also as casein kinase-2) is a pleiotropic Ser/Thr protein kinase with hundreds of regulation targets, which have a variety of cellular functions. CK2 plays an important role in gene-expression regulation, the synthesis and degradation of proteins, as well as the signaling and suppression of apoptosis.[1,2] CK2 is abnormally highly active in many tumor cells,[3] thus the attenuation of CK2 activity can be a strategy for the treatment of different neoplastic diseases.[4,5] CK2 is composed of two catalytic subunits (alpha) and two regulatory subunits (beta).[6] So far, a number of CK2 inhibitors have been described in micromolar and submicromolar ranges.[7−9] Halogenated benzoimidazoles belong to a widely and successfully tested family of CK2 inhibitors competing with ATP. There are a number of crystal structures available.[9−11] Of these, eight crystal structures contain aromatic tetrabromo derivatives and one a tetraiodobenzimidazole inhibitor (cf. Table 1A).[11] The inhibitors are bound to a CK2 α-catalytic subunit, specifically to a small hydrophobic cavity of the ATP binding site (Figure 1). The binding mode could explain the fairly narrow selectivity of those CK2 inhibitors. Further analysis highlights the major role of the hydrophobic effect for increasing potency within this inhibitor class.[7,12,13] It also shows that polar interactions are responsible more so for the orientation in the active site and, further, that the respective binding arises from the existence of two halogen bonds of the O···Br type.[12] Recently, it has been shown that novel tetraiodinated benzimidazoles[8] are more powerful inhibitors of CK2 than their tetrabrominated analogues.[11] Evidently, this class of CK2 inhibitors represents an ideal target for the study of halogen bonding in protein−ligand complexes.

A recently published survey of protein and nucleic acid structures has revealed the halogen bond as a stabilizing intermolecular interaction.[14] Lately, some papers have even been aimed at using halogen bonding in rational drug design,[15−21] in crystal engineering,[22,23] and directing macromolecular conformation.[24] The strength of the halogen bonding depends on its

Received: March 7, 2011
Revised: May 13, 2011
Published: June 07, 2011

Table 1. Summary of the Experimental Data and Calculated Terms for (A) The First Set of CK2 Inhibitors (cf. Figure 2A) and (B) The Second Set of CK2 Inhibitors (cf. Figure 2B)[a]

(A)				PM6-DH2X					AMBER			
system	K_i (μM)	ln K_i (M)	Res.(Å)	ΔH_w	$\Delta\Delta G_w$(I)	ΔE_{def}(I)	$T\Delta S_w$	$\Delta G'_w$	ΔH_w	ΔE_{def}(I)	$T\Delta S_w$	$\Delta G'_w$
1J91[10]	0.400[12]	−14.73	2.22	−15.13	−4.70	0.67	−15.67	−3.49	−24.81	0.27	−15.67	−8.87
1ZOE[12]	0.045[12]	−16.92	1.77	−38.09	−7.16	0.87	−19.53	−24.85	−31.76	1.53	−19.53	−10.70
1ZOG[12]	0.070[12]	−16.47	2.30	−28.81	−6.80	1.04	−18.83	−15.74	−23.15	1.31	−18.83	−3.02
1ZOH[12]	0.100[12]	−16.12	1.81	−24.70	−7.13	2.58	−27.95	−1.30	−42.83	0.77	−27.95	−14.11
2OXD[56]	0.150[56]	−15.71	2.30	−22.29	−5.25	0.83	−10.62	−16.09	−8.24	0.68	−10.62	3.06
2OXX[56]	0.200[56]	−15.42	2.30	−16.68	−10.43	0.93	−3.09	−23.09	−18.99	0.32	−3.09	−15.58
2OXY[56]	0.300[56]	−15.02	1.81	−20.87	−5.87	0.17	1.31	−27.88	−14.15	0.14	1.31	−15.32
3KXG[11]	0.048[50]	−16.85	1.70	−30.50	−4.48	0.47	−1.87	−32.64	−11.33	0.31	−1.87	−9.15
3KXN[11]	0.023[8]	−17.59	2.00	−34.97	−4.90	1.02	−6.53	−32.32	−14.27	0.51	−6.53	−7.22

(B)			PM6-DH2X					AMBER			
system	K_i (μM)	ln K_i (M)	ΔH_w	$\Delta\Delta G_w$(I)	ΔE_{def}(I)	$T\Delta S_w$	$\Delta G'_w$	ΔH_w	ΔE_{def}(I)	$T\Delta S_w$	$\Delta G'_w$
10a	0.120	−15.94	−17.56	−4.87	0.42	−14.24	−7.77	−35.67	0.32	−14.24	−21.12
12a	0.120	−15.94	−21.99	−5.70	1.21	−9.14	−17.34	−27.12	0.17	−9.14	−17.81
13a	0.090	−16.22	−22.76	−5.99	0.25	−11.52	−16.97	−29.35	0.28	−11.52	−17.55
15	0.050	−16.81	−18.88	−9.21	0.89	−13.28	−13.92	−28.39	0.20	−13.28	−14.91
17	0.050	−16.81	−24.85	−4.85	0.27	−9.71	−19.70	−24.56	1.17	−9.71	−13.68
7a	0.024	−17.55	−24.40	−5.61	0.22	−19.42	−10.38	−33.87	0.30	−19.42	−14.15
K10	0.370	−14.81	−17.21	−4.23	0.11	−13.59	−7.75	−14.47	0.37	−13.59	−0.50
K20	0.370	−14.81	−16.59	−5.49	0.14	−13.43	−8.51	−25.85	0.17	−13.43	−12.24
K21	0.250	−15.20	−17.74	−5.11	1.82	−10.54	−10.48	−26.22	0.10	−10.54	−15.57
K22	0.200	−15.42	−16.50	−10.64	0.46	−15.98	−10.70	−39.35	0.14	−15.98	−23.23
K32	0.180	−15.53	−20.57	−5.10	0.16	−11.38	−14.13	−25.18	0.91	−11.38	−12.89
1a	0.023	−17.59	−24.73	−4.82	0.25	−11.81	−17.49	−28.45	0.12	−11.81	−16.53
4a	0.460	−14.59	−21.21	−2.57	0.11	−10.28	−13.39	−24.91	0.60	−10.28	−14.04
8a	0.330	−14.92	−18.93	−2.55	0.12	−6.58	−14.79	−20.94	0.77	−6.58	−13.59

[a] K_i is the inhibition constant (in μM); Res. is the experimental mean resolution of the respective X-ray structure (in Å); ΔH_w is the interaction enthalpy; $\Delta\Delta G_w$(I) is the correction for the ligand desolvation free energy; ΔE_{def}(I) is the inhibitor deformation energy; $T\Delta S_w$ is the entropic contribution; $\Delta G'_w$ is the total score (all of the energies are in kcal/mol).

substitution[25] and solvent effects.[26] It has been shown that the origin of halogen bonding is in an electrostatic attraction between the positive σ-hole on the tip of the halogen atom (opposite to the covalent bond between the halogen and carbon atoms) and the negative Lewis base[27−29] and that it has a directional property similar to the hydrogen bond.[20,27] Moreover, the positive σ-hole is largest for the iodine atom and is not present for a fluorine atom; i.e., fluorinated analogues cannot make halogen bonds.[23] It should be noted that the interaction energy of a halogen bond can be comparable with the interaction energy of a hydrogen bond;[30,31] therefore, halogen bonds are very appealing in the process of rational drug design.[18] Recently, a halogen bond has been successfully exploited in the design of a selective CDK9 inhibitor.[32] The opening of the σ-hole is clearly a quantum effect, and consequently, halogen bonding cannot be described by current classical force fields using atom-centered partial charges. In addition, the description of the halogen bond might also be challenging for scoring functions, which are based on principles similar to empirical force fields. Without any doubt, to investigate the interaction of the inhibitor with a protein, where halogen bonds matter, it is necessary to employ quantum mechanical (QM) methods. Recently, we have developed a semiempirical QM (SQM) method (PM6-DH2),[33,34] which performs well in the description of various types of noncovalent complexes (see Methods for details). The PM6-DH2 method has also been successfully applied for scoring the inhibitors of the HIV protease[35] and CDK2 kinase.[36]

The aim of the present study is to investigate complexes of CK2 protein kinase with different inhibitors, each having several halogens engaged in the halogen bonding. We will demonstrate that force-field methods completely fail to describe the structure as well as binding free energy of these complexes, while the SQM method provides geometries in good agreement with the X-ray structures. Similarly, the scoring function based on the SQM method yields a good correlation between the binding energy and inhibition activity.

METHODS

The formation of the protein (P)−inhibitor (I) complex from free (hydrated) subsystems is represented by the following equation

$$P_W + I_W \rightarrow P{-}I_W \quad (1)$$

The binding of a competitive inhibitor is expressed by an inhibition constant (K_i), which is related to the change of the

dx.doi.org/10.1021/jp202149z |J. Phys. Chem. B 2011, 115, 8581–8589

Figure 1. ATP competitive inhibitor (here K25 in sticks) binds to a deep cleft in the CK2α subunit structure (PDB ID code: 1ZOE), which is shown in a cartoon model (the red curled strips represent α-helices and the yellow stripes β-strands).

free energy of binding (binding free energy) as follows

$$\Delta G_{\mathrm{W}} = RT \ln(K_{\mathrm{i}}) \quad (2)$$

The binding free energy was approximated by the total score ($\Delta G'_{\mathrm{w}}$), expressed by the following equation (for details, see our previous paper).[35]

$$\Delta G'_{\mathrm{W}} = \Delta H_{\mathrm{W}} - T\Delta S_{\mathrm{W}} + \Delta E_{\mathrm{def}}(\mathrm{I}) + \Delta\Delta G_{\mathrm{W}}(\mathrm{I}) + \Delta E_{\mathrm{def}}(\mathrm{P}) + \Delta\Delta G_{\mathrm{W}}(\mathrm{P}) \quad (3)$$

The individual terms in $\Delta G'_{\mathrm{w}}$ (eq 3) represent the interaction enthalpy (ΔH_{w}), the interaction entropy (as $-T\Delta S_{\mathrm{w}}$), the inhibitor, and the protein deformation energy $\Delta E_{\mathrm{def}} = E(\mathrm{I})_{\mathrm{w}}^{\mathrm{PI}} - E(\mathrm{I})_{\mathrm{w}}^{\mathrm{I}}$ (analogously for the protein P; the upper indexes PI and I standing for the geometry of the inhibitor (protein) in a protein/inhibitor complex and the relaxed inhibitor (protein) geometry in water, respectively) as well as the corrections for the inhibitor (protein) hydration free energy $\Delta\Delta G_{\mathrm{w}}(\mathrm{I}) = \Delta G_{\mathrm{w}}^{\mathrm{MOPAC}}(\mathrm{I})^{\mathrm{PI}} - \Delta G_{\mathrm{w}}^{\mathrm{G09}}(\mathrm{I})^{\mathrm{PI}}$ (analogously for protein P; the lower index w stands for a water environment). The last term reflects the fact that the structure of the PI complex (in a water environment) is determined using the hydration model considering the electrostatic term only (COSMO in the MOPAC code), while an accurate treatment requires the use of a hydration model considering the nonelectrostatic term as well. The majority of the scoring functions uses empirical potentials, also called molecular mechanics (MM) methods. The major drawback of all of the classical empirical potentials is their limited accuracy. The most important of them are the polarization effects, proton and electron transfer, and halogen bonding. All of these quantum effects are satisfactorily described by QM methods. The use of nonempirical QM methods is, however, limited to systems having no more than several hundred atoms. The P−I complexes are mostly much larger, with the smallest ones possessing several thousand atoms. The first possibility of how to solve the problem is to use methods like the ab initio fragment MO (FMO) method.[37] Two limitations of the FMO method should be taken into account, i.e., the applicability to charged protein−charged inhibitor complexes and the inability to perform geometry optimizations. The second possibility is the use of a hybrid QM/MM model.[38] When, however, the QM part is limited to a few hundred atoms, some artifacts arising from the communication between the layers can occur. The third possibility is the application of SQM methods.[39] The use of standard SQM methods is not straightforward owing to their poor treatment of dispersion energy and hydrogen bonding.[40] This problem has recently been solved in our laboratory by using Stewart's PM6[41] SQM method corrected for dispersion and hydrogen bonding.[33,34] The developed PM6-DH2 method provides accurate results for various types of noncovalent complexes including hydrogen-bonded and dispersion-controlled complexes, which arises from the inclusion of corrections to the dispersion and hydrogen-bonding energies.[33,34] The PM6-DH2 method overestimates, however, the strength of the halogen bond and provides distances between halogen and electronegative atoms that are too short. This is because of an insufficient repulsion in the PM6. We have addressed this problem in a recent letter[42] by an additional repulsion term fitted specifically to halogen bonding. The corrected method, abbreviated as PM6-DH2X, is able to describe the halogen bond with accuracy close to the high-level QM methods. For more details, see ref 42.

Nonelectrostatic terms associated with the formation of a cavity to accommodate the solute, and the van der Waals interaction between solute and solvent, are neglected in the former approach. This approximation might work well for inhibitors of a similar shape and the same charge, but in the case of entirely different inhibitors it can cause a serious error. For this reason, we have adopted a more accurate calculation of the complete solvation free energy of the ligand with a COSMO[43] model based on B3LYP calculations (Gaussian 09 code[44]).

The formation of a PI complex restricted the motions of the ligand as well as the torsional motion of the protein. Furthermore, the vibrational entropy of the ligand was also restricted by the formation of a PI complex. The entropic contributions were determined using the rigid rotor/harmonic oscillator approximation based on a Cornell et al.[45] AMBER force field. The structures of all of the systems were reoptimized using the empirical potential considering the continuous water (generalized Born model), and the same method was used for the calculation of the second derivatives of the total energy.

Structure Preparation. Two sets of structural data were considered with appropriate inhibition constants. The first set of complexes (Figure 2A) involves nine crystal structures (eight tetrabromo derivatives and one tetraiodobenzimidazole) of a maize CK2α subunit with an inhibitor for which both structural data as well as inhibition constants are available (Table 1A). The structures of all of the complexes were taken from the RSCB Protein Data Bank. When two CK2α units were present, only chain A was further considered. The hydrogen atoms were added to all of the structures, and their positions were minimized using AMBER[46] − ff03[45] and gaff[47] force fields in Chimera software.[48] All of the crystal waters were removed except for those in the 1ZOE complex, where the Cl^- ion is located in an active site near an inhibitor. Consequently, nine waters located up to 5 Å from this inhibitor were considered; these waters are important for the shielding of the Cl^- ion. Generally, two different conformations of an inhibitor exist; the only exception is the 1J91. Those conformations differ in rotation by about 60° around the axis perpendicular to the plane of the inhibitors.

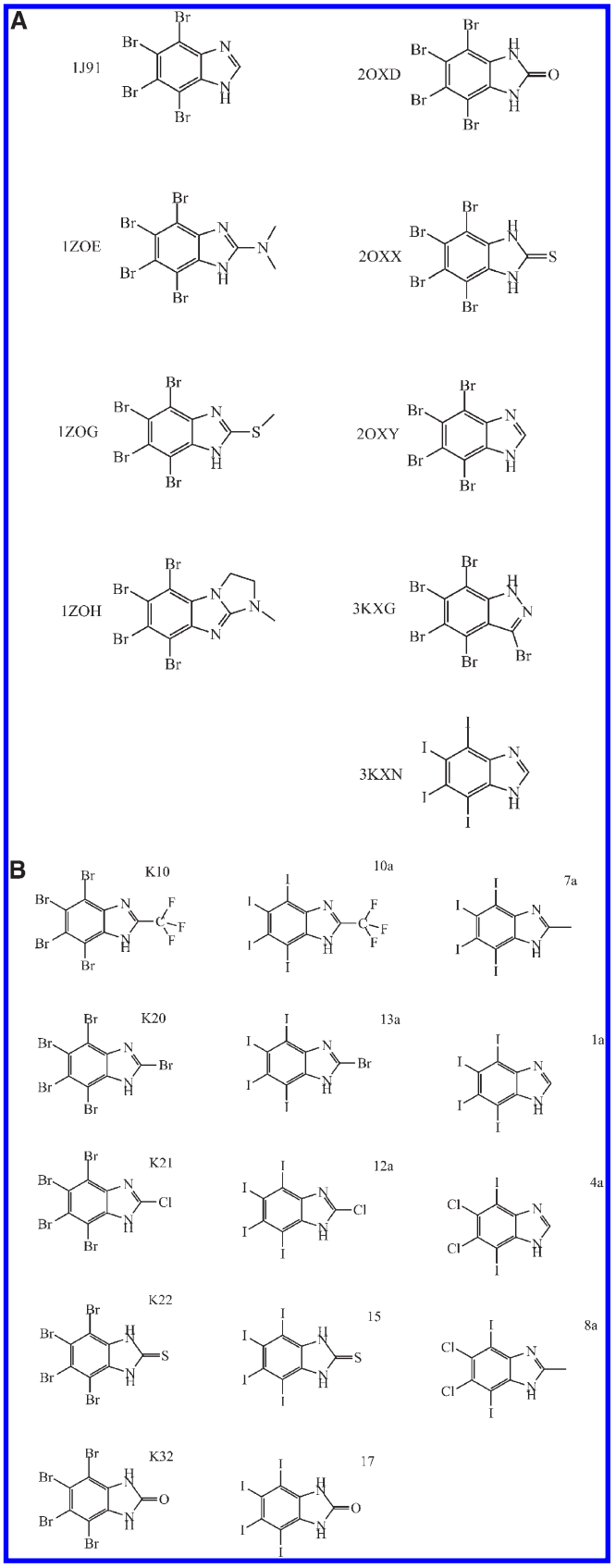


Figure 2. (A) The structures of the nine inhibitors considered in the first set. (B) The structures of the fourteen inhibitors considered in second set.

In the first conformation (1ZOG, 1ZOH, 3KXG, 3KXN), Br5 and Br6 interact with an oxygen of Glu114 and Val116 through halogen bonds, and in the second conformation (1ZOE, 2OXD, 2OXX, 2OXY) Br4 and Br5 interact with an oxygen of Glu114 and Val116 through halogen bonds. More specifically, in the case of 1ZOH and 3KXN complexes, the inhibitor is found in two positions. In the present study, we have systematically considered only one position, which belongs to the first conformation. In the cases of 1ZOG and 3KXN inhibitors, also the second conformation was taken into account. All of the inhibitors investigated were assumed to be electroneutral. The only exception represents the 1ZOE inhibitor where besides neutral form also protonated form was considered. The positions of the hydrogen atoms in the 1ZOE complex having nine waters connected to the Cl^- ion were optimized by means of PM6 (in the MOPAC code[49]) with the aim of obtaining a H-bonded network of waters, ion, CK2, and the inhibitor. The resulting structure as well as all other structures were further fully reoptimized using the PM6-D2X method (see later).

The second set was created by using a structural alignment, where one human CK2α subunit was taken from the 1JWH (a CK2complex with ANP, which is an ATP analogue) and the structure of the inhibitor was taken from the maize CK2α subunit (1ZOE). In the model structures, the inhibitors were modified according to their chemical structure to prepare complexes (Figure 2B) with known inhibition constants (Table 1B).[8] Like in the previous case, hydrogen atoms were added to all of the structures considered, and their positions were minimized using AMBER[46] – ff03[50] and gaff[47] force fields in Chimera software.[48] All of the inhibitors in this group were considered to be electroneutral.

Strategy of Calculations. The structures of all of the protein–inhibitor complexes were systematically optimized by the PM6-D2X method in a continuum COSMO solvent model as implemented in the MOPAC code[49] with the following optimization criteria (ΔE = 0.0300 kcal/mol, maxGrad = 6.0 kcal/mol/Å, and rmsGrad = 3 kcal/mol/Å). Note that the hydrogen bonding correction is not used for the optimization because our model calculations suggest that the most accurate interaction energies are obtained using the H-bond-corrected PM6-DH2X energies on geometries obtained without the correction (PM6-D2X). The respective interaction enthalpies ($\Delta H_w = \Delta H_w(\mathrm{PI}) - (\Delta H_w(\mathrm{P}) + \Delta H_w(\mathrm{I}))$) were determined using these optimized structures. In the second step, all of the complexes were reoptimized with AMBER ff99[51] and gaff force fields.[47] The RESP charges[52] were computed at the HF/6-31G* level; in the case of iodine, the HF/SDD[53] level with the respective pseudopotentials as implemented in Gaussian 09 was considered. In all of the cases, a generalized Born solvent model in the Nucleic Acid Builder (NAB, from the AMBER[46] package) was utilized. The entropy term (T = 298 K, p = 1 atm) was determined using the ideal gas, rigid-rotor harmonic-oscillator approximation using the empirical potential; all of the geometries were reoptimized with the Newton–Raphson method and the L-BFGS TNCG algorithm to a gradient of 10^{-12} kcal/mol/Å. The interaction enthalpies ($\Delta H_w^{\mathrm{AMBER}}$) were also determined with the same empirical force field (T = 298 K, p = 1 atm) and the generalized Born solvent model.

The inhibitor deformation energy ($\Delta E_{\mathrm{def}}(\mathrm{I})$) was calculated as the difference between the energy of the inhibitor in the geometry taken from the protein/inhibitor complex ($E(\mathrm{I})^{\mathrm{PI}}$) and the energy of the fully optimized inhibitor in water ($E(\mathrm{I})^{\mathrm{w}}$). It should be mentioned that the inhibitors investigated in the

Table 2. Comparison of the Structural X-ray Data with the PM6-D2X- and AMBER-Optimized Geometries for CK2−Inhibitor Complexes ((A), the First Set, cf. Figure 2A[a] and (B) the Second Set, cf. Figure 2B)[b]

(A)

	rmsd		halogen bonds (PM6-D2X and AMBER relative changes to X-ray)					
system	PM6-D2X	AMBER	X-ray		PM6-D2X		AMBER	
			Arg47 N	Val44 O	Arg47 N	Val44 O	Arg47 N	Val47 O
1J91	0.10	1.30	2.99	3.99	0.13	0.21	1.42	−0.41
			Glu114 O	Val116 O	Glu114 O	Val116 O	Glu114 O	Val116 O
1ZOE	0.11	1.12	3.43	3.24	−0.02	−0.02	0.90	0.94
1ZOG	0.14	1.30	3.23	2.98	−0.02	0.08	1.54	1.06
1ZOH	0.14	1.33	3.38	3.23	0.02	−0.18	0.23	0.40
2OXD	0.31	0.98	3.27	2.80	0.25	0.18	1.46	1.93
2OXX	0.11	0.87	3.73	2.72	0.11	0.14	0.52	1.53
2OXY	0.10	1.24	3.16	2.95	−0.04	0.09	1.75	1.96
3KXG	0.13	1.03	3.57	2.85	−0.33	0.11	0.69	1.41
3KXN	0.12	1.13	3.34	3.17	−0.06	−0.11	1.39	1.56

(B)

	halogen bonds			
system	PM6-D2X		AMBER	
	Glu114 O	Val116 O	Glu114 O	Val116 O
10a	3.06	2.89	4.47	4.06
12a	3.02	2.94	5.41	5.47
13a	3.01	2.96	5.4	5.19
15	3.03	2.94	5.25	5.25
17	3.01	2.97	4.94	3.76
7a	3.05	3.00	5.52	5.34
K10	3.06	3.07	5.75	5.8
K20	3.06	3.05	5.44	5.23
K21	3.10	3.03	5.54	5.44
K22	3.05	3.08	5.56	5.59
K32	3.04	3.07	4.82	3.68
1a	3.04	3.00	5.51	5.58
4a	3.04	3.04	4.67	3.29
8a	3.04	3.04	4.63	3.29

[a] The rmsd of the Cα atoms of the PM6-D2X- and AMBER-optimized structures from the X-ray structure; the fourth and the fifth columns show the distances between the oxygens of the CK2 amino acids and the halogen atoms of the inhibitor as found in the respective crystal structure, whereas the sixth and the seventh columns and the eighth and ninth columns show the same distances after the PM6-D2X and AMBER optimization (all of the distances are in Å). [b] The second and third columns and the fourth and fifth columns, show the distances between the oxygens of the CK2 amino acids and the halogen atoms of the inhibitor after PM6-D2X and AMBER optimization (in Å).

present paper are relatively rigid when compared to the floppy inhibitors considered in our previous studies.[35,36] For a further discussion of the inclusion of the deformation and desolvation energies of the ligand, see also another of our recent papers.[54] The PM6-D2X methods with the COSMO solvation model were used for these calculations. The correction for the inhibitor deformation energy ($\Delta E_{def}^{AMBER}(I)$) was also calculated at the empirical level using AMBER force fields with a generalized Born solvent model.

The correction of the inhibitor desolvation free energy $\Delta\Delta G_w$-(I) (T = 298 K, p = 1 atm) was determined for an inhibitor structure taken from the optimized CK2−inhibitor complex. It was evaluated as the difference between the solvation free energies calculated by the PM6/COSMO model ($\Delta G_w^{MOPAC}(I)^{PI}$), where only the electrostatic terms were taken into consideration, and the COSMO model (C-PCM from B3LYP/6-31G* wave function; in the case of iodine, the B3LYP/SDD level with the respective pseudopotentials as implemented in Gaussian 09 was taken into account), where also all of the nonelectrostatic terms were considered.

The deformation energy $\Delta E_{def}(P)$ and correction for the desolvation free energy $\Delta\Delta G_w(P)$ of the protein were not considered since it is expected that these terms are changed only marginally for the structurally very similar inhibitors investigated in the present study.

RESULTS AND DISCUSSION

Structural Comparisons. Analyzing the structures of all of the complexes, we have found two basic positions of inhibitors with

dx.doi.org/10.1021/jp202149z |J. Phys. Chem. B 2011, 115, 8581–8589

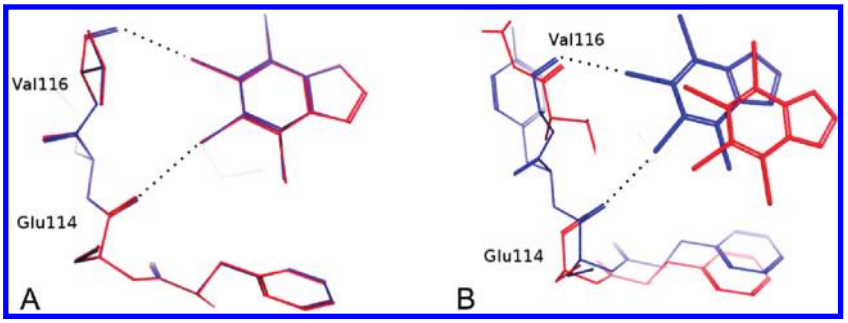


Figure 3. (A) PM6-D2X optimized structure (red lines) of the inhibitor bound to the CK2α active site agrees well with the 2OXY X-ray structure (blue lines), (B) whereas the AMBER optimized structure (red lines) shows substantial differences from the X-ray geometry (blue lines). The failure of the AMBER empirical potential is caused by its inability to describe halogen bonds (shown by black dotted lines).

respect to the CK2 active site: (i) as in 1ZOG, 1ZOH, 3KXG, 3KXN and (ii) as in 1ZOE, 2OXD, 2OXX, 2OXY with the exception of 1J91. Tables 2A and 2B show a comparison of the geometrical parameters of the experimental (or model) structures with the optimized ones for all of the complexes of the first and second group when the PM6-D2X and AMBER methods were utilized. The PM6-D2X method provides much better rmsd's of the Cα atoms from the X-ray than the AMBER force field. Specifically, the averaged rmsd for the nine complexes optimized by the PM6-D2X method of the first group amounts to 0.14 Å, and when the AMBER method is used the rmsd value increases considerably to 1.14 Å. The halogen···oxygen distances in both of the halogen bonds between the inhibitor and the CK2 active site determined by the PM6-D2X agree well with the experimental data (Table 2A). On the other hand, when the AMBER force field is used, the optimized distances are 1−2 Å larger (Table 2A). The most apparent difference, found for the 2OXY complex, is shown in Figure 3. Again, the PM6-D2X optimized structure agrees well with the experimental X-ray structures, while the AMBER-optimized structure differs considerably. It is evident that the inhibitor is shifted out of the cavity region in the AMBER optimized structure. The same applies for all of the inhibitors involved in the study.

All of these results clearly demonstrate that the SQM method reproduces the structure and geometry of CK2−inhibitor complexes well including the geometry of halogen bonds, while the MM method fails. MM fails because of its inability to describe the halogen bond between the inhibitor and CK2 kinase. The reason is that the AMBER force field is not capable of describing the σ-hole on the halogen atom because the halogen atom is represented as a homogeneous Lennard-Jonnes sphere bearing a usually negative partial charge. Consequently, the interaction between the halogen atom and halogen-bond acceptor (also bearing a negative partial charge) is repulsive in the MM calculation. This fact explains why the inhibitors are shifted out of the cavity region in the MM optimized structures. All optimized structures are available upon request.

Binding Free Energy. In the present study, two sets of CK2 inhibitors with known inhibition constants (K_i) were considered. The X-ray structures of the maize CK2 were available for the first set, while the model structures were derived for the second set. The primary sequence similarity between maize and human (or rat) CK2 is about 70%. The ATP binding site is almost identical, and the only deference nearby CK2 inhibitor is the Val66Ile mutation. The available inhibition constants from rat CK2 (Table 1B) were used for the maize CK2 structures in the

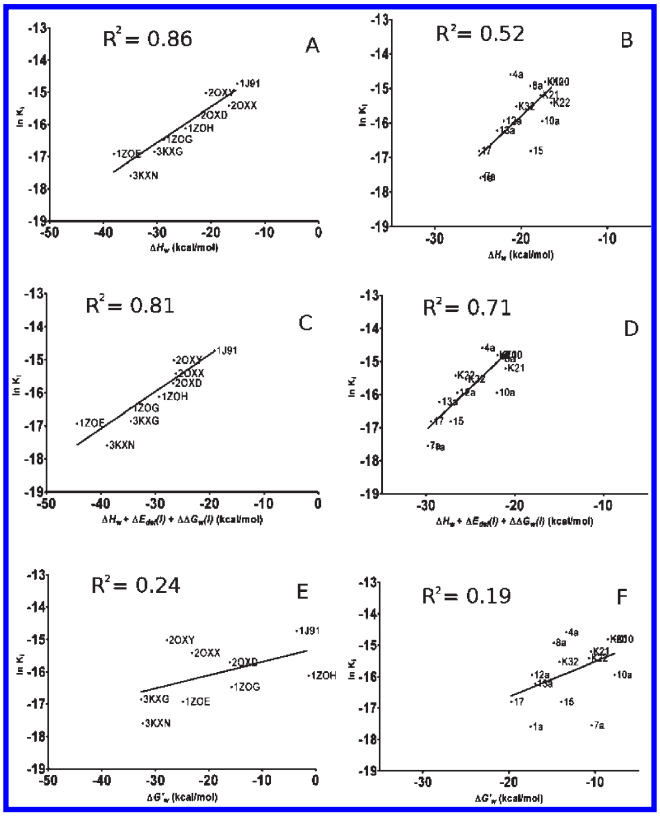


Figure 4. Correlations between the logarithm of the inhibition constants (ln K_i, in μM) and the energy terms (all of the energies are in kcal/mol) calculated by the PM6-DH2X method, i.e., with (A) the interaction enthalpy (ΔH_w) for the first set (R^2 = 0.86), (B) the interaction enthalpy (ΔH_w) for the second set (R^2 = 0.52), (C) the sum of the interaction enthalpy, the inhibitor deformation, and the corrections for the desolvation energies ($\Delta H_w + \Delta E_{def}(I) + \Delta\Delta G_w(I)$) for the first set ($R^2$ = 0.81), (D) the same as in C but for the second set (R^2 = 0.71), (E) the total score ($\Delta G'_w$) calculated as the sum of the interaction enthalpy, the interaction entropy, and the corrections for the inhibitor deformation and the desolvation energies ($\Delta G'_w = \Delta H_w - T\Delta S_w + \Delta E_{def}(I) + \Delta\Delta G_w(I)$) for the first set ($R^2$ = 0.24), (F) the same as in E but for the second set (R^2 = 0.19).

case of the first group based on crystal data.[55] On the contrary, the second group created from a model based on human CK2 structures used inhibition constants measured on rat CK2.[8] We thus investigated the correlation between the experimental and

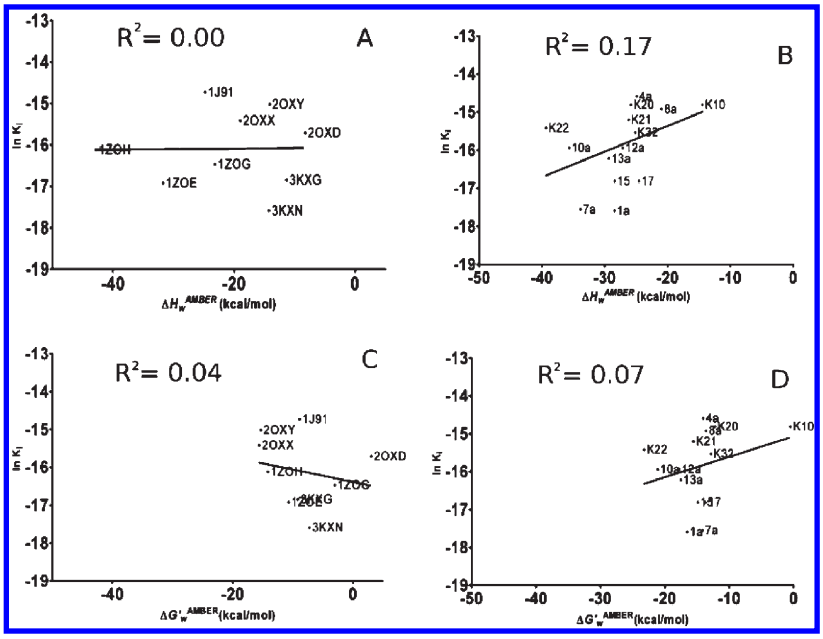


Figure 5. Correlations between the inhibition constants (ln K_i) and the AMBER characteristic calculated with (A) the interaction enthalpy ΔH_w^{AMBER} for the first set ($R^2 = 0.00$), (B) the interaction enthalpy ΔH_w^{AMBER} for the second set ($R^2 = 0.17$), (C) the total score $\Delta G'_w{}^{AMBER}$ for the first set ($R^2 = 0.04$), and (D) the total score $\Delta G'_w{}^{AMBER}$ for the second set ($R^2 = 0.07$).

computed binding free energy (inhibition constants) for the nine CK2−inhibitor complexes from the first group (cf. Figure 2A). In addition, the correlation between the experimental and computed binding free energy for the fourteen CK2−inhibitor complexes from the second group (cf. Figure 2B) was also evaluated. In both cases, the energy terms were calculated on fully optimized structures. This is important because it makes the method independent of an exact knowledge of the experimental geometry.

Tables 1A and 1B summarize the single terms from the total score for both of the inhibitor groups calculated by the SMQ method. Figure 4A shows a tight correlation ($R^2 = 0.86$) between the PM6-DH2X interaction enthalpy (ΔH_w) and the experimental inhibition constants (ln K_i) for the first group of complexes (with known crystal structures). It should be mentioned that the protonated form of the 1ZOE inhibitor provided less favorable interaction enthalpies than the neutral form and was not considered. Further, the second orientation of the 3KXN inhibitor provided considerably less favorable interaction enthalpy than the first one, and thus, only the first orientation was considered. The results of the analogous calculations performed for the second group of complexes (where model structures were generated) are shown in Figure 4B, but the correlation is considerably worse (R^2 = 0.52). Despite the fact that the interaction enthalpies were calculated in a water environment, several important contributions to the total score (which estimates the free energy of binding) are still missing. Adding inhibitor deformation (ΔE_{def}(I)) and correction for desolvation free energies ($\Delta\Delta G_w$(I)) to the interaction enthalpy does not change the correlation of the first group (Figure 4C) significantly ($R^2 = 0.81$), while it considerably improves the correlation of the second group ($R^2 = 0.71$, Figure 4D). When, however, the total score including also the entropy term is considered, the correlations for both of the groups of complexes (cf. Figures 4E and 4F) strongly deteriorate ($R^2 = 0.24$ and 0.19, respectively). It is worth noting that the entropy term is evaluated by the empirical potential (on structures minimized by the empirical potential) because the calculation of the second derivatives by the SQM method employed is still impractical for such large systems. The weak correlation is explained by the usage of the empirical potential for the evaluation of the entropy term, which fails in its description of halogen bonding, essential for the binding of the considered inhibitors to CK2. In other words, the empirical force field is not able to describe halogen bonding, as a result of which the optimized structures of inhibitor−CK2 complexes involving halogen bonds are incorrect, as shown in the previous paragraph. The only possibility for overcoming the problems with entropy is then to evaluate this term either by using the proper QM or SQM methods or by using the modified empirical potential. In our laboratory, we are working in both equally important directions.

For the sake of completeness, we have also analyzed the correlation of the inhibition constants (ln K_i) with the energy terms calculated consistently by the AMBER force field on the AMBER optimized structures. For the first (Figure 5C, Table 1A) as well as the second group (Figure 5D, Table 1B), no statistically significant correlation between the inhibition constant and the total score was found (R^2 = 0.04 and 0.08, respectively). The same applies for the correlation of the interaction enthalpy with the inhibition constant (Figures 5A and 5B). The empirical potential fails in ranking the CK2 inhibitors involving halogen bond in binding. This finding is alarming because the majority of the empirical scoring functions used in docking is based on the same principles as the AMBER empirical potential. The failure of the AMBER empirical potential in the geometry and energy of the protein−inhibitor complexes involving a halogen bond warns that the docking and scoring of the inhibitors bearing potential halogen-bond donors (Cl, Br, and I atoms typically on the aromatic rings or neighboring the strong electron-acceptor groups) cannot be blind and the results should be carefully analyzed while keeping in mind the Achilles′ heel of the empirical potential in the description of the halogen bond.

CONCLUSIONS

(i) The PM6-D2X optimized structures of the eight tetrabromobenzimidazoles and one tetraiodobenzimidazole in a CK2α active site agreed well with the X-ray data, with the rmsd of the Cα atoms being equal to 0.14 Å. When the AMBER empirical potential was applied, the agreement became considerably worse with an average rmsd of 1.14 Å. The experimental distances between the halogen and oxygen in the halogen bonds were very well reproduced by the PM6-D2X SQM method (with the average difference being below 0.2 Å), while the AMBER force field showed significant differences (with the average difference being greater than 1.0 Å).

(ii) The correlation between the PM6-DH2X interaction enthalpies and the inhibition constants for the CK2−inhibitor complexes for which experimental structural data exist was high (R^2 = 0.86), and it was lower (R^2 = 0.52) for the modeled CK2−inhibitor complexes. When the interaction enthalpy was augmented by the inhibitor deformation energy and corrections for the inhibitor desolvation free energy, the correlation was practically the same in the first set of inhibitors (from r^2 = 0.86 to 0.81) and was significantly improved in the second set (from R^2 = 0.52 to 0.71). Adding the entropy term, which was evaluated at the empirical level, deteriorated the correlation strongly (to R^2 = 0.24 and 0.19, for the first and the second set, respectively). This shows that the entropy term determined for the AMBER-optimized structures cannot be used in the construction of a total scoring function. The systematic use of the AMBER force field instead of the PM6-D2X SQM method does not lead to any improvement.

(iii) The AMBER force field fails to describe halogen bonding and, consequently, also the structures of the CK2−inhibitor complexes. The opposite is true for the PM6-D2X SQM method. The AMBER force field also fails to predict the interaction enthalpies or the total score for the CK2−inhibitor complexes, and again, the PM6-DH2X method provides quite reliable values.

AUTHOR INFORMATION

Corresponding Author

*E-mail: pavel.hobza@uochb.cas.cz. Tel.: +420 220 410 311. Fax: +420 220 410 320.

ACKNOWLEDGMENT

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 funded by Grant Nos. LC512 and MSM6198959216 from the Ministry of Education, Youth and Sports of the Czech Republic. This work was also supported by grant GACR P208/11/0295 and by the Operational Program Research and Development for Innovations − European Social Fund (CZ.1.05/2.1.00/03.0058). The support of Praemium Academiae, Academy of Sciences of the Czech Republic, awarded to P.H. in 2007 is also acknowledged. We would like to thank both referees for their excellent comments which improved quality of the present paper.

REFERENCES

(1) Pinna, L. A.; Allende, J. E. *Cell. Mol. Life Sci.* **2009**, *66*, 1795–1799.
(2) Filhol, O.; Cochet, C. *Cell. Mol. Life Sci.* **2009**, *66*, 1830–1839.
(3) Ruzzene, M.; Pinna, L. A. *Biochim. Biophys. Acta* **2010**, *1804*, 499–504.
(4) Sarno, S.; Pinna, L. A. *Mol. BioSyst.* **2008**, *4*, 889–894.
(5) Trembley, J. H.; Wang, G.; Unger, G.; Slaton, J.; Ahmed, K. *Cell. Mol. Life Sci.* **2009**, *66*, 1858–1867.
(6) Niefind, K.; Raaf, J.; Issinger, O. G. *Cell. Mol. Life Sci.* **2009**, *66*, 1800–1816.
(7) Pagano, M. A.; Bain, J.; Kazimierczuk, Z.; Sarno, S.; Ruzzene, M.; Di Maira, G.; Elliott, M.; Orzeszko, A.; Cozza, G.; Meggio, F.; Pinna, L. A. *Biochem. J.* **2008**, *415*, 353–365.
(8) Gianoncelli, A.; Cozza, G.; Orzeszko, A.; Meggio, F.; Kazimierczuk, Z.; Pinna, L. A. *Bioorg. Med. Chem.* **2009**, *17*, 7281–7289.
(9) Cozza, G.; Bortolato, A.; Moro, S. *Med. Res. Rev.* **2010**, *30*, 419–462.
(10) De Moliner, E.; Brown, N. R.; Johnson, L. N. *Eur. J. Biochem.* **2003**, *270*, 3174–3181.
(11) Sarno, S.; Papinutto, E.; Franchin, C.; Bain, J.; Elliott, M.; Meggio, F.; Kazimierczuk, Z.; Orzeszko, A.; Zanotti, G.; Battistutta, R.; Pinna, L. A. *Curr. Topics Med. Chem.* **2011**, *11*, 1340–1351.
(12) Battistutta, R.; Mazzorana, M.; Sarno, S.; Kazimierczuk, Z.; Zanotti, G.; Pinna, L. A. *Chem. Biol.* **2005**, *12*, 1211–1219.
(13) Wasik, R.; Lebska, M.; Felczak, K.; Poznanski, J.; Shugar, D. *J. Phys. Chem. B* **2010**, *114*, 10601–10611.
(14) Auffinger, P.; Hays, F. A.; Westhof, E.; Ho, P. S. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 16789–16794.
(15) Hernandes, M. Z.; Cavalcanti, S. M.; Moreira, D. R.; de Azevedo Junior, W. F.; Leite, A. C. *Curr. Drug Targets* **2010**, *11*, 303–314.
(16) Lu, Y.; Shi, T.; Wang, Y.; Yang, H.; Yan, X.; Luo, X.; Jiang, H.; Zhu, W. *J. Med. Chem.* **2009**, *52*, 2854–2862.
(17) Voth, A. R.; Ho, P. S. *Curr. Top. Med. Chem.* **2007**, *7*, 1336–1348.
(18) Hardegger, L. A.; Kuhn, B.; Spinnler, B.; Anselm, L.; Ecabert, R.; Stihle, M.; Gsell, B.; Thoma, R.; Diez, J.; Benz, J.; Plancher, J. M.; Hartmann, G.; Banner, D. W.; Haap, W.; Diederich, F. *Angew. Chem., Int. Ed. Engl.* **2010**, *50*, 314–318.
(19) Lu, Y.; Wang, Y.; Zhu, W. *Phys. Chem. Chem. Phys.* **2010**, *12*, 4543–4551.
(20) Metrangolo, P.; Neukirch, H.; Pilati, T.; Resnati, G. *Acc. Chem. Res.* **2005**, *38*, 386–395.
(21) Kortagere, S.; Ekins, S.; Welsh, W. J. *J. Mol. Graphics Modell.* **2008**, *27*, 170–177.
(22) Metrangolo, P.; Meyer, F.; Pilati, T.; Resnati, G.; Terraneo, G. *Angew. Chem., Int. Ed. Engl.* **2008**, *47*, 6114–6127.
(23) Politzer, P.; Murray, J. S.; Concha, M. C. *J. Mol. Model.* **2007**, *13*, 643–650.
(24) Voth, A. R.; Hays, F. A.; Ho, P. S. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 6188–6193.
(25) Torii, H.; Yoshida, M. *J. Comput. Chem.* **2010**, *31*, 107–116.
(26) Sarwar, M. G.; Dragisic, B.; Salsberg, L. J.; Gouliaras, C.; Taylor, M. S. *J. Am. Chem. Soc.* **2010**, *132*, 1646–1653.
(27) Politzer, P.; Murray, J. S.; Clark, T. *Phys. Chem. Chem. Phys.* **2010**, *12*, 7748–7757.
(28) Politzer, P.; Lane, P.; Concha, M. C.; Ma, Y.; Murray, J. S. *J. Mol. Model.* **2007**, *13*, 305–311.
(29) Legon, A. C. *Phys. Chem. Chem. Phys.* **2010**, *12*, 7736–7747.
(30) Zierkiewicz, W.; Wieczorek, R.; Hobza, P.; Michalska, D. *Phys. Chem. Chem. Phys.* **2011**, *13*, 5105–5113.
(31) Li, Q. Z.; Jing, B.; Li, R.; Liu, Z. B.; Li, W. Z.; Luan, F.; Cheng, J. B.; Gong, B. A.; Sun, J. Z. *Phys. Chem. Chem. Phys.* **2011**, *13*, 2266–2271.
(32) Baumli, S.; Endicott, J. A.; Johnson, L. N. *Chem. Biol.* **2010**, *17*, 931–936.
(33) Korth, M.; Pitonak, M.; Rezac, J.; Hobza, P. *J. Chem. Theory Comput.* **2010**, *6*, 344–352.
(34) Rezac, J.; Fanfrlik, J.; Salahub, D.; Hobza, P. *J. Chem. Theory Comput.* **2009**, *5*, 1749–1760.

(35) Fanfrlik, J.; Bronowska, A. K.; Rezac, J.; Prenosil, O.; Konvalinka, J.; Hobza, P. *J. Phys. Chem. B* **2010**, *114*, 12666–12678.
(36) Dobes, P.; Fanfrlik, J.; Rezac, J.; Otyepka, M.; Hobza, P. *J. Comput.-Aided Mol. Des.* **2011**, *25*, 223–235.
(37) Mazanetz, M. P.; Ichihara, O.; Law, R. J.; Whittaker, M. *J. Cheminform.* **2011**, *3*, 2.
(38) Gleeson, M. P.; Gleeson, D. *J. Chem. Inf. Model.* **2009**, *49*, 670–677.
(39) Raha, K.; Peters, M. B.; Wang, B.; Yu, N.; Wollacott, A. M.; Westerhoff, L. M.; Merz, K. M., Jr. *Drug. Discovery Today* **2007**, *12*, 725–731.
(40) Dobes, P.; Otyepka, M.; Strnad, M.; Hobza, P. *Chem.—Eur. J.* **2006**, *12*, 4297–4304.
(41) Stewart, J. J. *J. Mol. Model.* **2009**, *15*, 765–805.
(42) Rezac, J.; Hobza, P. *Chem. Phys. Lett.* **2011**, *506*, 286–289.
(43) Klamt, A.; Eckert, F.; Hornig, M. *J. Comput.-Aided Mol. Des.* **2001**, *15*, 355–365.
(44) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A.; , Jr., Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; , . Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, Fox, D. J. *Gaussian 09*, Revision A.1; Gaussian, Inc.: Wallingford CT, 2009.
(45) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. *J. Comput. Chem.* **2003**, *24*, 1999–2012.
(46) Case, D. A.; Cheatham, T. E., 3rd; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M., Jr.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. *J. Comput. Chem.* **2005**, *26*, 1668–1688.
(47) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. *J. Comput. Chem.* **2004**, *25*, 1157–1174.
(48) 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*, 1605–1612.
(49) Stewart, J. J. *J. Comput.-Aided Mol. Des.* **1990**, *4*, 1–105.
(50) Duncan, J. S.; Litchfield, D. W. *Biochim. Biophys. Acta* **2008**, *1784*, 33–47.
(51) Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049–1074.
(52) Bayly, C. I.; Cieplak, P.; Cornell, W.; Kollman, P. A. *J. Phys. Chem.* **1993**, *97*, 10269–10280.
(53) Dunning, T. H., Jr.; Hay, P. J. In *Modern Theoretical Chemistry*; Schaefer, H. F., III, Ed.; Plenum: New York, 1976; Vol. 3, pp 1−28.
(54) Kolar, M.; Fanfrlik, J.; Hobza, P. *J. Phys. Chem. B* **2011**, *115*, 4718–4724.
(55) Pagano, M. A.; Meggio, F.; Ruzzene, M.; Andrzejewska, M.; Kazimierczuk, Z.; Pinna, L. A. *Biochem. Biophys. Res. Commun.* **2004**, *321*, 1040–1044.
(56) Battistutta, R.; Mazzorana, M.; Cendron, L.; Bortolato, A.; Sarno, S.; Kazimierczuk, Z.; Zanotti, G.; Moro, S.; Pinna, L. A. *ChemBioChem* **2007**, *8*, 1804–1809.