Struct Chem (2015) 26:1411–1423
DOI 10.1007/s11224-015-0658-9

ORIGINAL RESEARCH

Accurate modeling of cation–π interactions in enzymes: a case study on the CDPCho:phosphocholine cytidylyltransferase complex

Anikó Lábas[1] · Balázs Krámos[1,2] · Imre Bakó[2] · Julianna Oláh[1]

Received: 30 July 2015 / Accepted: 3 August 2015 / Published online: 19 August 2015
© Springer Science+Business Media New York 2015

Abstract Cation–π interactions are functionally relevant, strong secondary interactions that play versatile roles in a variety of chemical and biological systems. Therefore, it is very important to be able to describe accurately and reliably these interactions. In this study, we propose a methodology for the accurate modeling of cation–π interactions in proteins using QM/MM calculations. We developed a methodology for computing the many-body interaction energy terms and tested the effect of various factors on the accuracy of the binding energy. We found that once well-equilibrated structures were reached in the MD simulations, very similar results can be obtained for the various snapshots taken from the trajectory. The calculated interaction energies were only slightly influenced by electrostatic embedding of the point charges in the QM/MM calculations and by QM/MM geometry optimization. The calculated molecular mechanics interaction energies were off by 50 % for cation–π interactions. Instead, we suggest the calibration of force fields based on fragment-based QM calculations on geometries obtained from MD simulations to yield reliable binding energies at reduced computational cost.

Keywords Cation–π interaction · Energy decomposition · CDPCho:phosphocholine cytidylyltransferase

Dedicated to Professor Magdolna Hargittai on the occasion of her 70th birthday.

Electronic supplementary material The online version of this article (doi:10.1007/s11224-015-0658-9) contains supplementary material, which is available to authorized users.

✉ Julianna Oláh
julianna.olah@mail.bme.hu

Anikó Lábas
labas@ch.bme.hu

[1] Department of Inorganic and Analytical Chemistry, Budapest University of Technology and Economics, Szent Gellért tér 4, Budapest 1111, Hungary

[2] Institute of Organic Chemistry Research Centre for Natural Sciences, Hungarian Academy of Science, Magyar tudósok körútja 2, Budapest, P.O. Box 286, 1519, Hungary

Introduction

Non-covalent interactions, such as hydrogen bonds and salt bridges, play a crucial role in all biological systems [1]. Among others, they contribute to the three-dimensional structure of biomolecules, to ligand binding and to enzymatic catalysis. As such, in silico drug discovery and design methodologies depend to a large extent on proper description of these intermolecular forces. In the course of virtual screening, millions of small compounds are docked into the ligand binding site of receptors or enzymes and scoring functions are used to estimate the strength and likelihood of binding [2].

Cation–π interaction is a relatively newly recognized secondary molecular interaction, where a positively charged molecule or molecular moiety interacts with an aromatic ring [3]. The strength of a cation–π interaction is significant; it is comparable to hydrogen bonds and salt bridges [4]. Several intermolecular forces, such as charge quadrupole, charge dipole, charge transfer, and dispersion forces, contribute to the formation of cation–π interactions, but the most significant energy contribution is derived from electrostatic interactions [5, 6].

Cation–π interactions play a prominent role in a variety of fields, for example in structural biology, materials

science, catalysis and organic synthesis [4, 7]. Their relevance in biological systems where they contribute significantly to the stabilization of protein structure and ligand binding has extensively been demonstrated [8].

The evaluation of 593 high-resolution structures in Protein Data Bank showed that (1) from every 77 residues, one participates in cation–π interaction; (2) arginine is more likely to be involved in cation–π interactions than lysine; and (3) 25 % of tryptophan residues form energetically significant cation–π interaction [9]. Web servers, such as CaPTURE [9] and Protein Explorer [10], use distance-based criteria to find structurally significant cation–π interactions and use force field parameters fitted to ab initio calculations to predict the strength of the interaction. These methods are fast, but very specific, and they can only be applied certain types of interactions.

In contrast to molecular mechanics, quantum mechanics-based methods have the ability to accurately describe inter- and intramolecular interactions in any system, but these methods are too computationally demanding for screening purposes.

Therefore, in the course of drug discovery and design usually scoring functions, approximate mathematical expressions, are used to approximate the binding energy in order to accelerate the computational process. One of the most frequently applied approaches is to use force fields to calculate the interaction energy between the protein and the ligand. This method is fast and easily applied; however, it has serious limitations, for example neglecting electronic polarization, approximate handling of van der Waals interactions, and difficult or no treatment of solvation effects. Although the parameterization of force fields is usually done with great care, extensive validation of the reliability of the force field can only be achieved during extended applications and repeated evaluation. Ideally, one would like to get excellent agreement between calculated and real (measurable) binding energies, or at least obtain the same trend. Achieving such a qualitative agreement is a great advantage and shows the reliability of the method [11]. When experimental data are not available, the computed results can be compared to data calculated at a higher level; for example, the results of MM methods could be compared to ab initio or DFT data.

An important advantage of computed data is that they can reveal the contribution of protein side chains to the overall binding energy of the ligand, which is very difficult to obtain experimentally. Therefore, in this work we (1) design a methodology to obtain accurate binding energies at low computational cost and (2) tackle the factors involved in ligand binding. One of the major issues considered in accurate modeling of protein–ligand interactions is protein flexibility. Therefore, we investigate (1) how thermal motion influences the results by studying the binding energy along the MD trajectory; (2) the generally used approximation of the calculation of the binding energy as a sum of pair potentials by calculating the many-body interaction energy terms; (3) the possible improvement in QM/MM calculations over MM calculations and the effect of electrostatic embedding; and (4) the effect of DFT functional and basis set.

In this work, we use the phosphocholine cytidylyltransferase enzyme (*Pf*CCT) from *Plasmodium Falciparum* causing malaria as a model system. This protein has been recently actively researched as a potentially new drug target against malaria [12, 13]. Furthermore, detailed information exists on the role of the various active site side chains on ligand binding affinity and enzyme activity, and detailed thermodynamic properties of point mutants, which will enable us to compare the computed results to experiments [14].

*Pf*CCT functions as a homodimer (Fig. 1) and plays a key role in the membrane synthesis of the pathogen of malaria. It produces cytidine 5′-diphosphocholine (CDPCho) via the S_N2 reaction of choline phosphate (ChoP) and cytidine 5′-triphosphate (CTP) [12, 15, 16] (Scheme 1). Ligands of *Pf*CCT are bound by several secondary interactions in a highly conserved pocket including a "composite aromatic box" [17]. The cytosine and pyrophosphate groups are stabilized mainly by hydrogen bonds, while the choline group forms cation–π interactions with Trp692 and Tyr714 residues in the aromatic cage and

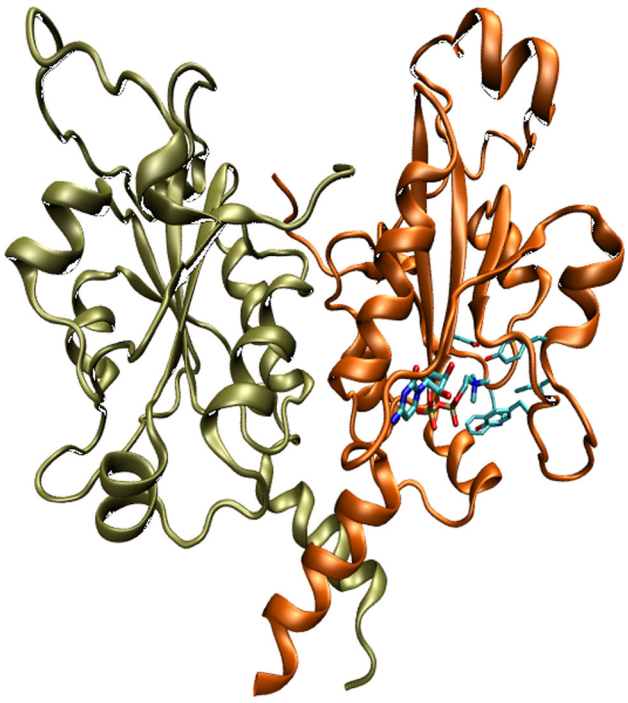
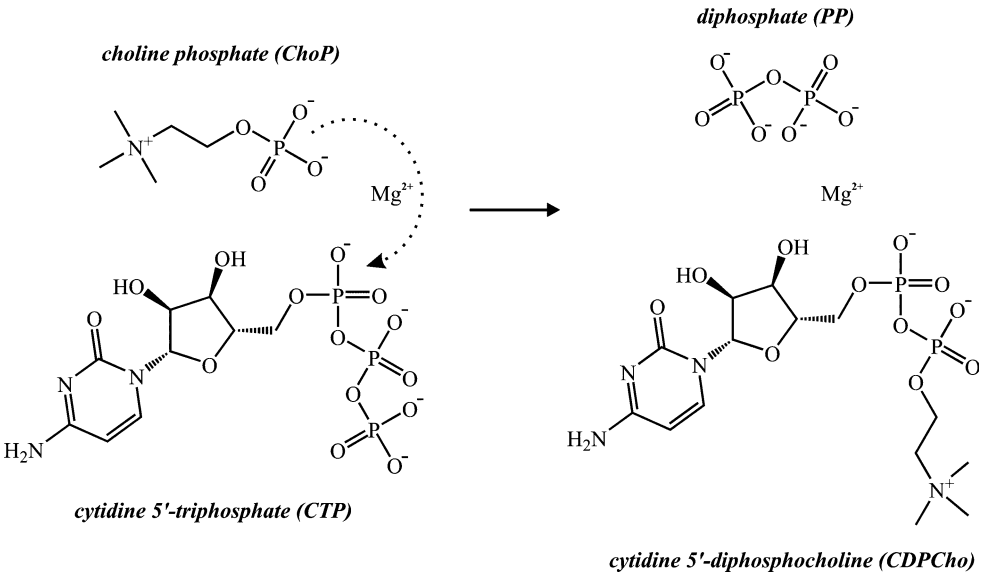


Fig. 1 Three-dimensional representation of the homodimeric structure of the *Pf*CCT enzyme. The cytidine 5′-diphosphocholine (CDPCho) ligand is shown in licorice

Scheme 1 Reaction of choline phosphate (ChoP) and cytidine 5′-triphosphate (CTP) catalyzed by *Pf*CCT producing cytidine 5′-diphosphocholine (CDPCho)



a salt bridge with the side chain of Asp623 (see Fig. 2). As our primary focus was on cation–π interactions, we studied the wild-type (WT) enzyme and two experimentally well-characterized mutant enzymes where the original tryptophan residue at position 692 was mutated to tyrosine (Trp692Tyr) and phenylalanine (Trp692Phe), respectively.

Computational details

Gas-phase calculations

Tetramethylammonium ion (TMA) was used as a model of the choline group and its interaction with the side chains of Trp, Phe, and Tyr amino acids was investigated. As it is known that most common DFT functionals fail to describe non-covalent interactions properly, we used various functionals parametrized for the better description of non-covalent interactions, the ωB97X-D [18] and the M06-2X [19] functionals. We also calculated the interaction energy with the B3LYP [20] and the TPSSh [21] functionals. For B3LYP, the DFT-D3 dispersion correction term [22] was also determined. The structure of the van der Waals complexes of TMA and the amino acid side chains was optimized in vacuo using the ωB97X-D and B3LYP functionals and the 6-31G(d) and cc-pVTZ [23] basis sets using Gaussian09 program [24]. To test the effect of the choice of DFT functional, interaction energies were also

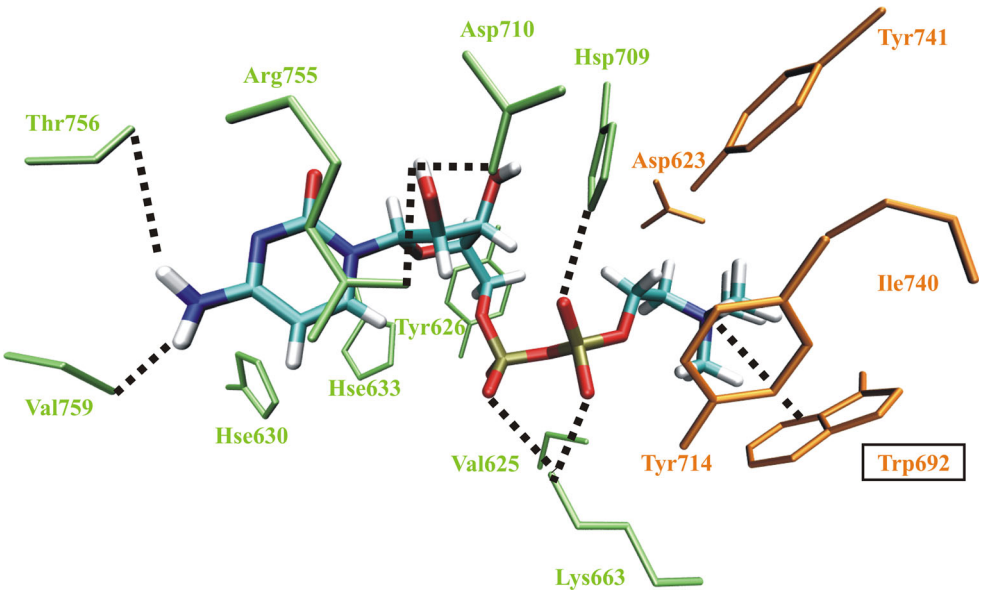


Fig. 2 Most important intermolecular interactions responsible for the binding of cytidine 5′-diphosphocholine in the active site of the *Pf*CCT enzyme

calculated using TPSSH and M06-2X functionals at the ωB97X-D/6-31G(d) geometry. The counterpoise correction [25] was used to avoid basis set superposition error. Natural charges were calculated, and the NEDA analysis was performed with the NBO 5.9 program [26, 27].

MD simulation in enzymes

No crystal structure of the *Pf*CCT enzyme exists; therefore, a previously published homology model (PMDB: PM0078719) [28] of the native enzyme product (CDPCho) complex was used as a starting structure for our MD simulations. In the case of the point-mutated enzymes, the Trp692 residue was mutated in silico to Tyr and Phe residues. The same protocol was followed in the case of the wild-type and mutant enzymes. The protonation state of the ionisable amino acid side chains of the enzymes was predicted by the PROPKA program [29–32]. Based on the estimated pKa values, the Glu638 and the Asp589 residues were protonated and His709 was doubly protonated. All other histidine residues were protonated in the ε position based on careful examination of their hydrogen bond environment. Missing hydrogen atoms were added and minimized according to the standard CHARMM protocol. Buffer region was assigned 21–25 Å away from the alpha-phosphorus atom of the CDPCho ligand, which was set as the center of the system. Charged residues in the buffer region were neutralized using "patch" residues. The structure was solvated within a 60-Å water box containing TIP3P [33] water molecules, and water molecules whose oxygen atom was within 2.8 Å area of other non-hydrogen atoms of the protein or was farther than 25 Å from the alpha-phosphorus of the CDPCho ligand were removed. Added water molecules were equilibrated by stochastic boundary MD at 310 K over 20 ps, followed by a multistep optimization of all atoms within a 25-Å sphere around the center. This was followed by a stochastic boundary MD simulation in which the position of atoms less than 21 Å away from the center of the system was propagated using Newtonian dynamics, and in the buffer region with Langevin dynamics. First, the system was heated to 310 K over 60 ps. At this temperature, MD equilibration was carried out over 100 ps, which was followed by a 20-ns productive MD simulation. Increasing restraints were applied in the buffer region in the course of minimizations and MD simulations in order to restrain the movement of heavy atoms around their initial positions. Atoms farther than 25 Å from the alpha-phosphorus atom were fixed. The CHARMM software package [34] and the CHARMM27 force field [35] were used for these simulations and the VMD program [36] for visualizations. Topology file and parameters for the CDPCho ligand were taken from our previous publication [28]. It should be noted that the parameters for the choline group of the ligand are identical to those from the CHARMM lipid force field [37].

QM/MM calculations

In order to assess the effect of protein flexibility on the results of QM/MM calculations, six snapshots were taken for the three enzyme variants from the corresponding MD trajectories (from the 0th, 16th, 17th, 18th, 19th, and 20th ns) as starting geometries and the geometry of these structures was carefully optimized using molecular mechanics (the CHARMM27 force field) before the QM/MM calculations were started. The CDPCho ligand was represented by a tetramethylammonium (TMA) ion in the quantum mechanically described region (QM region). The QM region also included the closest neighbors of the choline group: the side chains of the Asp623, Tyr714, Ile740, Tyr741 and of the mutated aromatic residue at position 692 (Trp in the wild type and Tyr and Phe in the mutants, respectively). The covalent bond was formally broken between the C_α and C_β atoms of each residue, and hydrogen-type link atoms were used to satisfy the valences of the carbon atoms in the QM/MM calculations. The charge of the group of the MM atoms directly connected to the QM atoms was set to zero to prevent unrealistic polarization effects. Atoms farther than 20 Å from the alpha-phosphorus atom of the CDPCho ligand were fixed, and the charge of atoms farther than 25 Å was set to zero. QM/MM geometry optimizations were carried out at the ωB97X-D/6-31G*/CHARMM27 level using the QoMMMa 8.02 program [38] which couples the output files generated by Gaussian09 and TINKER [39, 40] program packages. QM/MM interaction energies were determined at the optimized geometries using the ωB97X-D, B3LYP, TPSSH and M06-2X functionals and the cc-pVTZ basis set.

Energy decomposition in enzymes

In most force field-based methods, the overall interaction energy is calculated as a sum of the pair interaction energies of the fragments. However, it is also known that in many cases the overall interaction energy cannot be calculated as a simple sum of the pair interaction energies. In these cases, the non-additive term can be approximated by higher-order energy terms. In particular, the three-body term has been shown to have a significant contribution to the overall interaction energy. For example, in the case of water clusters the contribution of the three-body term has been shown to account for almost 20 % of the overall interaction energy at certain geometries, while at other geometries the three-body term was shown to destabilize the structure and oppose binding [41]. Therefore, in the

present study we extended a previously proposed energy decomposition method [42]. Accordingly, the total energy of the system can be written as a sum of one-, two-, three- … n-body terms and the interaction energy (E_{int}) can be calculated according to Eq. (1):

$$\begin{aligned} E_{\text{int}}^{\text{tot}} &= E_{\text{total}} - \sum_{i=1}^{n} E(i) \\ &= \sum_{i=1}^{n-1}\sum_{j>i}^{n} \Delta^2 E(ij) + \sum_{i=1}^{n-2}\sum_{j>i}^{n-1}\sum_{k>j}^{n} \Delta^3 E(ijk) \\ &\quad + \sum_{i=1}^{n-3}\sum_{j>i}^{n-2}\sum_{k>j}^{n-1}\sum_{l>k}^{n} \Delta^4 E(ijkl)\ldots \end{aligned} \tag{1}$$

where n is the number of the fragments in the system and $E(i)$, $E(ij)$, $E(ijk)$, etc. are the energies of the various one-, two-, three-membered clusters in the system, respectively. The pairwise-additive two-body interactions ($\Delta^2 E(ij)$), and the higher three-body ($\Delta^3 E(ijk)$), four-body $\Delta^4 E(ijkl)$, etc. non-additive components are defined as:

$$\Delta^2 E(ij) = E(ij) - \{E(i) + E(k)\}, \tag{2}$$

$$\begin{aligned} \Delta^3 E(ijk) &= E(ijk) - \{E(i) + E(j) + E(k)\} \\ &\quad - \{\Delta^2 E(ij) + \Delta^2 E(ik) + \Delta^2 E(jk)\}, \end{aligned} \tag{3}$$

$$\begin{aligned} \Delta^4 E(ijkl) &= E(ijkl) - \{E(i) + E(j) + E(k) + E(l)\} \\ &\quad - \{\Delta^2 E(ij) + \Delta^2 E(ik) + \Delta^2 E(il) \\ &\quad + \Delta^2 E(jk) + \Delta^2 E(jl) + \Delta^2 E(kl)\} \\ &\quad - \{\Delta^3 E(ijk) + \Delta^3 E(ijl) + \Delta^3 E(ikl) + \Delta^3 E(jkl)\} \end{aligned} \tag{4}$$

This energy decomposition scheme is generally applicable in both quantum mechanics and QM/MM calculations and could be applied to any systems composed of any fragments. However, in the case of QM/MM calculations some extra considerations are needed. One important note is that in this case it is only worth decomposing the interaction energy of the QM region: In the MM region all energies are calculated from pair potentials. Second consideration regards the type of embedding scheme used for the coupling of the QM and MM regions. In the case of QM/MM calculations with mechanical embedding, there is no electrostatic interaction between the QM and MM regions. Therefore, the energy decomposition scheme can be applied similarly to QM only calculations and neglecting the term representing the MM energy. However, in the case of calculations with electrostatic embedding care has to be taken with the treatment of the electrostatic interactions. Therefore, we calculated the energy of each fragment (whether one, two, three, four or five or six body), in the presence of the point charges representing the enzymatic environment, but paid attention to eliminate the self energy of the MM point charges from the calculated energy in order to avoid multiple subtraction.

Our QM region comprised of six fragments (the TMA group of the ligand and the five amino acid side chains), and we determined all 3-, 4-, 5- and 6-body energy terms. These calculations were carried out at the ωB97X-D/cc-pVTZ level and included the full set of basis functions used for the description of the entire QM region, in order to avoid the basis set superposition error.

Results

Gas-phase calculations

We started our investigation of cation–π interactions with determining the geometry of model systems of choline and amino acid residues with aromatic side chains. Tetramethylammonium cation (TMA) was used as a model of the quaternary ammonium group of the choline residue, and the aromatic amino acids were models with their side chains and their β-carbon atom.

The obtained geometries of the complexes are depicted in Fig. 3. The geometries obtained with the B3LYP and ωB97X-D functionals and the 6-31G* and cc-pVTZ basis sets were very similar. However, the interaction energies between the fragments show very strong functional dependence (see Table 1). Functionals parametrized for improved description of non-covalent interactions (M06-2X and ωB97X-D) predicted significantly stronger interaction between TMA and the aromatic fragments than the TPSSh and B3LYP functionals. However, when the D3 dispersion correction is added to the results of the B3LYP calculations, the interaction energy increases significantly and is in the same range as given by M06-2X and ωB97X-D. It is also worth emphasizing that all functionals (whether including or not dispersion correction) predict the same trends for the interaction energies in the studied complexes: Phenylalanine forms the weakest cation–π interaction with TMA, and tryptophan the strongest one. This result is in accordance with the observation that it is primarily tryptophan that is involved in energetically significant cation–π interactions in proteins [43].

The NBO analysis [44] has become one of the most significant tools in the hands of computational chemists to gain insight into the nature of interactions in molecules or complexes. We have performed a NEDA analysis [26] of these complexes. In all cases, delocalization from the bonding orbitals of the aromatic ring to the antibonding orbitals of C–H bonds located nearest in space to the aromatic ring is observed. Interestingly, in the case of the tyrosine–TMA complex, NEDA indicates that the strongest interaction is observed between the oxygen lone pair and

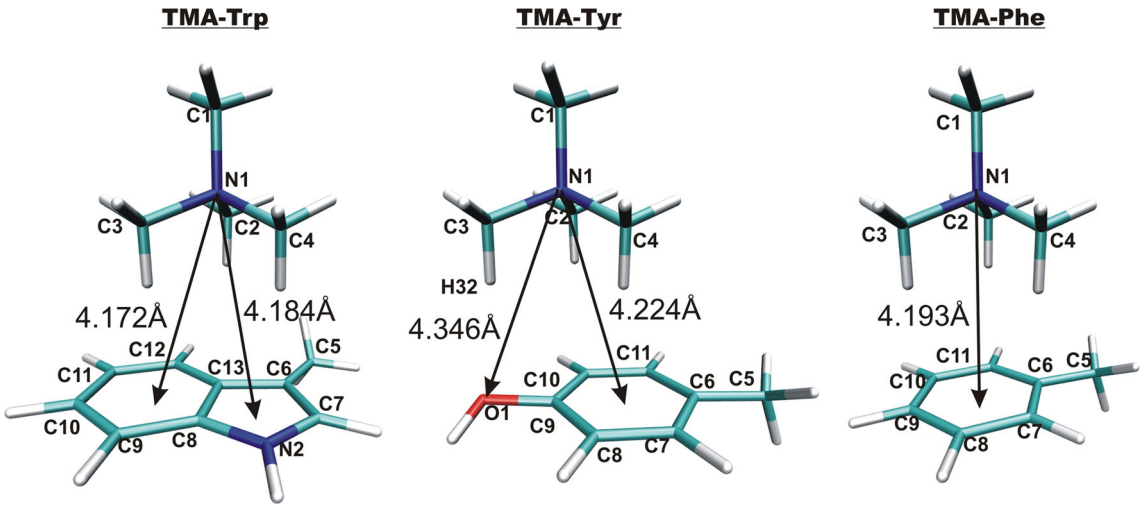


Fig. 3 Optimized structure of investigated models of TMA-Trp/Phe/Tyr complexes at the ωB97X-D/6-31G* level. Selected distances between the quaternary nitrogen atom and the center of the aromatic rings and the tyrosine oxygen are also shown

Table 1 Interaction energies (in kcal/mol) for the gas-phase van der Waals complexes calculated at the ωB97X-D/6-31G* geometries using the cc-pVTZ basis set and the indicated DFT functionals

	ωB97X-D	M06-2X	TPSSh	B3LYP	B3LYP-D3
TMA-Trp	−17.3	−16.0	−9.1	−7.9	−16.5
TMA-Tyr	−13.8	−12.8	−7.7	−7.0	−13.6
TMA-Phe	−11.7	−10.0	−5.9	−5.23	−11.4

antibonding orbital of the methyl group of TMA ($LP_{O1} \rightarrow BD^*_{C3\text{-}H32}$). This interaction is very well reflected by relative positions of the two fragments: TMA is not positioned strictly over the phenyl ring, but moved closer to the hydroxyl group (see Fig. 3).

Molecular dynamics simulations

After having studied the relevant cation–π interactions in gas phase, we turned our attention toward the enzymatic calculations. We did not observe any significant structural distortions of the wild-type and mutant enzymes in the course of 20-ns-long MD simulation, indicating the good quality of the starting structures. We have monitored the distance between the quaternary nitrogen atom of the choline group and the center point of the mutated residue (Trp692, Phe692, Tyr692). As Fig. 4a shows, this distance became slightly shorter during the MD in the wild-type enzyme, while it remained almost constant in the mutants, showing a stronger interaction in the wild-type enzyme. We have also monitored the distance between the nitrogen and the carboxylate oxygens of a nearby aspartate residue (Asp623) (Fig. 4b), as the experimental work

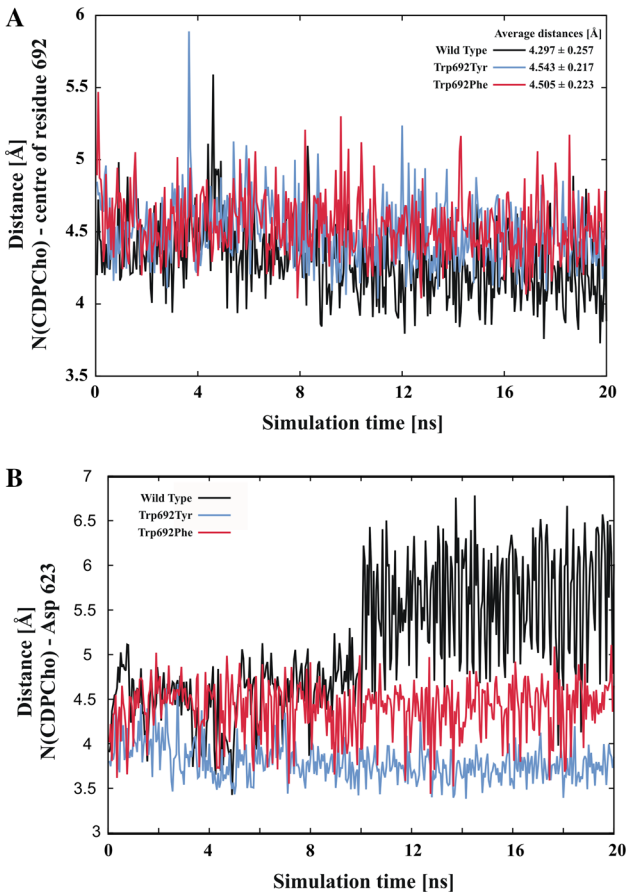


Fig. 4 Distance between the nitrogen atom of CDPCho and the center point of residue 692 (**a**) and the aspartate oxygens (**b**) during the MD simulations in the three enzyme variants

suggested that it contributes to the binding of the ligand via electrostatic interactions [14]. What is quite apparent is that in the mutants the aspartate–choline distance has

decreased, which indicates the strengthening of their interaction.

QM/MM calculations

In order to investigate the effect of the protein flexibility on the results of QM/MM calculations, we have taken six snapshots from the trajectories of the three enzyme variants: at 0th ns (thus the last structure obtained at the heating up period), and at the 16th, 17th, 18th, 19th, and 20th ns of the trajectories. These structures were chosen because we anticipated that the structures will be fully relaxed after 15 ns of MD simulation and we wanted to compare the results obtained for structures before and after the MD simulation. We carried out a full QM/MM optimization of all snapshots at the ωB97X-D/6-31G*/CHARMM27 level and carried out a full energy decomposition of the interaction energy of the system at the ωB97X-D/cc-pVTZ/CHARMM27 level.

In Fig. 5, we have superimposed the QM/MM-optimized geometries of the QM regions of all calculations. It is apparent that compared to the 0-ns structure, the structures underwent some changes, and that there is some variability in the structures between 16 and 20 ns; however, most of the changes are most likely due to thermal motion. These figures also show the abovementioned effect: In the wild-type enzyme, the Asp623 residue went farther away from the ligand than in the mutants.

The experiments on the binding of the CDPCho ligand to the active site of the wild-type and mutant PfCCT enzymes showed that the binding in all cases is enthalpy-driven (see Table 2) [16]; thus, it is a reasonable approximation to compare the calculated interaction energies of the ligand within the active site with the experimentally observed heat of association. Still, it is important to note that the experimental value will include other effects than simply the interaction between the ligands and the protein side chains, as it will necessarily include the heat chain related to the distortion of the ligand/protein upon binding as well as the desolvation of the ligand and of the cavity. According to the experiment, the ligand binds most strongly to the wild-type enzyme, while the mutants show considerably weaker binding, with slight preference for the tyrosine mutant. The gas-phase data (Table 1) correctly reproduce this trend, and on top, the binding energy between the side chain of residue 692 and the ligand at the starting structure (0 ns) and at the last point of the MD (20 ns) shows exactly the same trend. The average of the binding energies of the five snapshots (16- to 20-ns structures) confirms that the wild-type (Trp) enzyme binds the ligand most strongly and indicates similar interaction strength in the case of the Phe and Tyr mutated enzymes, in accordance with the experiments. Therefore, we may conclude that the data show little variation, which suggests that a limited number of snapshots may already yield reliable data. Previous studies on other thermodynamic properties such as the S–O [45] and Fe–O bond enthalpies [46] showed also only very small variation. However, when the overall interaction energy of the ligand with the total QM regions is assessed, the agreement with experiment is not retained, especially not for the 16- to 20-ns structures; here, the wild-type enzyme seems to bind the ligand the most weakly. For this reason we have calculated and extra snapshot taken at the 8th ns of the molecular dynamics trajectory as at this point, the Asp623 residue is still located closer to the ligand in the wild-type structure. As discussed in the following chapter, the Asp residue is responsible for the majority of the binding energy of the ligand in the active site, and based on our results, this term, as it describes a cation–anion interaction, is very sensitive to the exact position of the moieties. However, it is also likely that due to solvation/desolvation effects, this term will be much smaller in reality than predicted by our calculations.

Energy decomposition in enzymes

We carried out a full energy decomposition of the interaction energies at ωB97X-D/cc-pVTZ/MM level. In

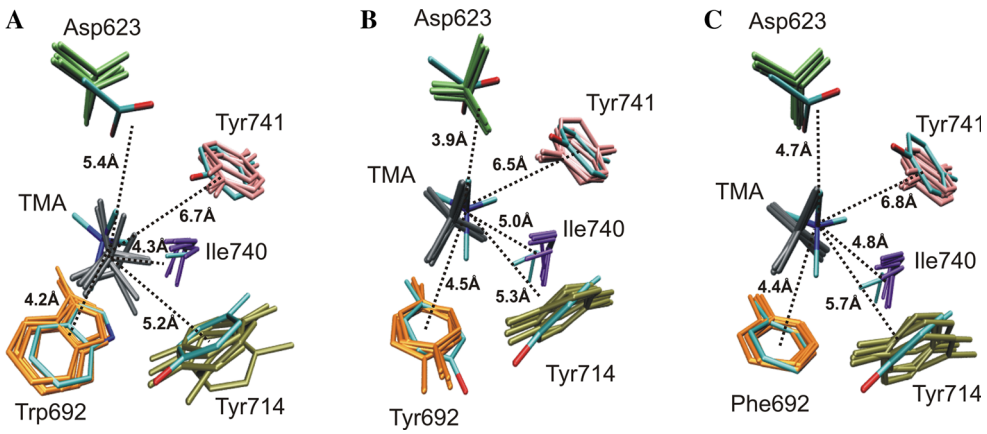


Fig. 5 Superposition of the QM/MM-optimized structures of the QM regions in the **a** wild type **b** Trp692Tyr mutant **c** Trp692Phe mutant enzymes. Hydrogen atoms are omitted for clarity. Coloring code for the atoms of the 0-ns structure: *light blue* carbon, *red* oxygen, *dark blue* nitrogen (Color figure online)

Table 2 Interaction energies (kcal/mol) of the CDPCho ligand with residue 692 in the enzyme (E_{int}(e)) and with the whole QM region in parentheses (in kcal/mol) according to QM/MM calculations, and experimentally measured enthalpy ($\Delta_a H$ in kcal/mol), entropy ($\Delta_a S$ in cal/molK), and Gibbs free energy ($\Delta_a G$ in kcal/mol) of association at 20 °C

	$E_{int}(e)_{0ns}$	$E_{int}(e)_{8ns}$	$E_{int}(e)_{16-20ns}$	$E_{int}(e)_{20ns}$	E_{int}(g)	$\Delta_a H$ [14]	$\Delta_a G$ [14]	$\Delta_a S$ [14]
Wild type	−12.5 (−90.4)	−17.0 (−81.4)	−11.4 ± 0.8 (−71.4 ± 3.2)	−10.8 (−71.5)	−17.3	−22.9 ± 1.0	−5.8 ± 2.5	−58.1 ± 5.0
Trp692Tyr	−8.8 (−93.0)	Not calculated	−6.8 ± 2.1 (−81.5 ± 3.4)	−8.7 (−82.8)	−13.8	−9.8 ± 4.5	−4.9 ± 7.0	−16.8 ± 8.6
Trp692Phe	−6.5 (−86.8)	−7.4 (−78.1)	−7.8 ± 1.0 (−75.9 ± 0.5)	−6.7 (−75.9)	−11.7	−6.8 ± 0.8	−4.3 ± 1.2	−8.9 ± 1.3

For comparison, calculated gas-phase data for the interaction of TMA with Trp, Tyr, and Phe side chains is also indicated (E_{int}(g), kcal/mol)

Table 3, the decomposition of the overall interaction energy to 2-, 3-, 4- and 5-body terms is shown (the contribution of the six-body term is negligible). The calculated overall interaction energy varies considerably between 80 and 92 kcal/mol, but the percentages of the different terms remain almost constant over all snapshots. Interestingly, the two-body term overestimates the binding energy by about 7 kcal/mol, 10 % of the overall interaction energy, which is counteracted by the three-body term, which seems to destabilize the interaction. The sum of the four-body terms accounts for 1.7 % of the overall interaction energy, but taken the large number of four-body terms $\binom{6}{4} = 15$, the contribution of the individual terms is negligible, and those of the five body terms are even smaller. This decomposition shows that the three-body terms contribute indeed significantly to the interaction energy and we observe an overall anticooperative effect as shown by the opposite sign compared to the two-body terms. Interestingly, recently published work on the interplay between various non-covalent interactions found that cation–π interactions acted frequently in an anticooperative manner with other non-covalent interactions, for example with hydrogen bonds [47]. In Table 4, we have only summarized those terms that are related to the binding of the ligand in the active site in contrast to Table 3 where all terms were summarized (i.e., the interaction energy between two amino acid side chains as well). It is remarkable that 90 % of the overall interaction energy is due to ligand binding, and that the ratio of the three-body terms is the same as in the case of the overall interaction energy. The same analyses were carried out for the two mutant structures as well, and the same conclusions were drawn (data not shown).

In Table 5, we show the calculated two-body interaction energy terms and those three-body terms that exceed 1 kcal/mol. It is visible that the standard deviation of the most significant terms is much smaller than the value, but in the case of weak interactions quite large standard deviations are observed. It is also obvious that in all cases the anion–cation interaction between the ligand and Asp623 is responsible for the majority of the interaction energy and that this energy is 6–10 times larger than the interaction between the ligand and the aromatic residue at 692. However, this observation suggests that the interaction with the negative residue is primarily responsible for the strong binding of the ligand in the composite aromatic box, and these numbers should be considered with some caution, as the desolvation effects upon binding of the ligand in the active site are expected to be more significant in the case of anionic residues than of aromatic side chains. It is also clear that the closer the aspartate to the CDPCho

Table 3 Decomposition of the total interaction energy $\left(E_{\text{int}}^{\text{tot}}\right)$ and percentage contribution of the various *n*-body terms to the overall interaction energy for the various QM/MM-optimized structures for the wild-type enzyme

Time (ns)	16	17	18	19	20	Average
E_{tot}	−85.8	−91.9	−89.2	−80.1	−87.2	−86.8 ± 4.4
$\Sigma\Delta^2$	−92.3 (107.6 %)	−99 (107.8 %)	−96.6 (108.3 %)	−85.5 (106.8 %)	−92.9 (106.6 %)	−93.3 ± 5.1
$\Sigma\Delta^3$	7.7 (9.0 %)	8.4 (9.1 %)	8.5 (9.5 %)	6.3 (7.9 %)	6.6 (7.6 %)	7.5 ± 1.0
$\Sigma\Delta^4$	−1.3 (1.5 %)	−1.4 (1.5 %)	−1.3 (1.5 %)	−1 (1.2 %)	−0.9 (1.0 %)	−1.2 ± 0.2
$\Sigma\Delta^5$	0.1 (0.1 %)	0.1 (0.1 %)	0.1 (0.1 %)	0.1 (0.1 %)	0.1 (0.1 %)	0.1 ± 0.0
$\Sigma\Sigma\Delta^{1-5}$	−85.8 (100.0 %)	−91.9 (100.1 %)	−89.2 (100.0 %)	−80.1 (100.0 %)	−87.2 (100.0 %)	−86.8 ± 4.4

Table 4 Decomposition of the interaction energy between the CDPCho ligand and the other fragments $\left(E_{\text{int,CDPCho}}^{\text{tot}}\right)$ and percentage contribution of the various *n*-body terms (compared to the overall interaction energy E_{tot}) for the various QM/MM-optimized structures for the wild-type enzyme

Time (ns)	16	17	18	19	20	Average
$\Sigma\Delta^2$	−77.3 (90.1 %)	−81 (88.2 %)	−81.4 (91.2 %)	−71.7 (89.6 %)	−77.1 (88.4 %)	−77.7 ± 3.9
$\Sigma\Delta^3$	7.6 (8.9 %)	8 (8.7 %)	8.4 (9.4 %)	6.3 (7.9 %)	6.4 (7.3 %)	7.4 ± 0.9
$\Sigma\Delta^4$	−1.2 (1.4 %)	−1.2 (1.3 %)	−1.3 (1.5 %)	−1 (1.2 %)	−0.8 (0.9 %)	−1.1 ± 0.2
$\Sigma\Delta^5$	0.1 (0.1 %)	0.1 (0.1 %)	0.1 (0.1 %)	0.1 (0.1 %)	0.1 (0.1 %)	0.1 ± 0.0
$\Sigma\Sigma\Delta^{1-5}$	−70.8 (82.6 %)	−74.1 (80.7 %)	−74.1 (83.1 %)	−66.3 (82.8 %)	−71.5 (82.0 %)	−71.4 ± 3.2

Table 5 Two-body interaction energy terms (in kcal/mol) and three-body terms exceeding 1 kcal/mol with their standard deviations averaged over the 16- to 20-ns structures for the enzyme variants

Interacting fragments	Wild type	Trp692Tyr	Trp692Phe
12	1.09 ± 0.18	−0.52 ± 0.68	0.36 ± 0.31
13	0.19 ± 0.70	−1.27 ± 0.73	−0.33 ± 0.28
14	−0.83 ± 0.41	−1.63 ± 0.74	−1.36 ± 0.27
15	−7.47 ± 0.72	−28.05 ± 3.04	−10.40 ± 0.20
16	−61.80 ± 3.99	−85.15 ± 2.08	−71.42 ± 1.35
23	−1.83 ± 0.91	−3.43 ± 1.79	−1.42 ± 0.45
24	−0.63 ± 0.32	−0.42 ± 0.14	−0.80 ± 0.17
25	−0.22 ± 0.02	−0.12 ± 0.02	−0.09 ± 0.04
26	−11.37 ± 0.82	−6.80 ± 2.12	−7.84 ± 1.00
34	−2.50 ± 0.23	−3.07 ± 0.77	−2.75 ± 0.55
35	−0.63 ± 0.28	−1.63 ± 0.34	−0.53 ± 0.23
36	−3.61 ± 1.24	−0.59 ± 2.45	−3.84 ± 0.94
45	−2.70 ± 0.24	−2.20 ± 0.59	−2.64 ± 0.15
46	−0.69 ± 0.55	0.05 ± 0.44	0.05 ± 0.33
56	−0.25 ± 0.69	−1.38 ± 1.11	−0.39 ± 0.81
126	2.16 ± 0.18	2.61 ± 0.38	2.01 ± 0.17
136	1.31 ± 0.23	2.02 ± 0.20	1.75 ± 0.15
146	1.01 ± 0.14	1.28 ± 0.37	1.02 ± 0.11
156	2.20 ± 0.24	7.32 ± 0.72	3.07 ± 0.21

Interacting fragments: 1 = Asp623, 2 = Residue 692 (Trp/Tyr/Phe), 3 = Tyr714, 4 = Ile740, 5 = Tyr741, 6 = TMA part of the CDPCho ligand

ligand is, the stronger the interaction is (WT: distance 5.4 Å, E_{int} −61.8 kcal/mol; Trp692Tyr: distance 4.7 Å, E_{int} −71.4 kcal/mol; Trp692Phe: distance 3.9 Å, E_{int} −85.2 kcal/mol). Interestingly, in the case of the Trp692Tyr system the second strongest interaction is found between the Tyr741 and the Asp623 residues. Careful

investigation of the structure revealed that a strong hydrogen bond has been established between Asp623 and Tyr741 in this mutant (average distance 1.62 Å). Furthermore, a very strong interaction network was established between Asp623, Tyr741, and the ligand as shown by the surprisingly large three-body term (7.3 kcal/mol). Furthermore, the data show that only those three-body terms became stronger than 1 kcal/mol which involved the two charged residues (TMA and Asp623). It is also visible that in all cases CDPCho ligand interacts more strongly with residue 692 than with Tyr741.

Assessment of various factors offering possibilities for reduction in computational cost

So far, we have established that our QM/MM calculations correctly predict the effect of the studied point mutations on the ligand binding energy of CDPCho and also found that pair potentials account for 90 % of the interaction energy. It also became evident that there is very little difference between the studied snapshots taken at 1-ns intervals after proper equilibration of the systems. Therefore, we set out to study various other effects on the accuracy of the calculation using the snapshot taken at 20 ns of the MD simulation of wild type.

The calculations presented so far were carried out with a large basis set (cc-pVTZ), so we calculated the pairwise interaction energies with the smaller 6-31G* basis set, and there was practically no effect of changing the basis set (see Table 6). This implies that in these QM/MM calculations no large basis sets are necessary; the transferability of this finding to other systems should be tested. Then, we turned out attention toward the effect of density functionals. Most of the pair potentials show very weak functional dependence, with the exception of the cation–π interaction between the ligand and the aromatic side chains, although this is most apparent for tryptophan residue 692, as it interacts most strongly with the ligand. Here, B3LYP and TPSSH give about 9 kcal/mol weaker interaction energy than M06-2X and ωB97X-D, and this value is in perfect accordance with the results of the gas-phase calculations, showing that from simple gas-phase calculations exact functional dependence of the QM/MM results can be deduced.

One of the most time-consuming part of our calculations was the process of QM/MM geometry optimizations. Therefore, we investigated the effect of QM/MM geometry optimization on the interaction energies. We compared the interaction energy calculated at the geometry of the snapshot taken from the MD trajectory and after QM/MM optimization (Table 6 and Table S2 in the Supporting Information). The observed differences were very small, the interaction energies are slightly more favorable at the QM/MM-optimized geometry. However, this change is so little, that would justify the usage of carefully MM-optimized geometries sampled from MD trajectories for QM/MM interaction energy calculations. The very small interaction energy difference calculated for structures before and after QM/MM optimization points to small differences between these structures. Indeed, when we compared the strain energy for each of side chains included in the QM region, we observed very small effects (Table 6). The effect is largest for the Tyr714 residue and reaches 3 kcal/mol.

QM/MM calculations were developed with the desire and need to be able to take into account the effect of the protein environment in course of calculations. However, with the development of computing resources one is able to treat larger and larger QM region and now there are advocates of both cluster-based QM calculations on enzymes and QM/MM methods as well, both of which are finding their entries into the pharmaceutical industry as well [48]. Therefore, e.g., in the course of modeling enzymatic reactions, if possible, it is worth increasing the size of the QM region to be as large as to have a small effect on the energetic of the reaction, thereby ensuring that all significant interactions are treated at a high level. Too large contribution from the MM region brings up the danger that some essential effects are treated with molecular mechanics; thus, their description might be unreliable. We checked this, and we found that removal of the point charges from the Hamiltonian has a very small effect on the obtained interaction energies, which are slightly stronger in the absence of the point charges.

Based on the above results, it seems likely that reasonable interaction energies could be obtained for the interaction of ligands with active site residues from carefully MM-optimized snapshots taken from a well-equilibrated trajectory, and that these interactions are already well predicted from a single structure, without the need to study an ensemble of structures. However, these calculations still require some quantum mechanics-based calculations. How well are these data reproduced by molecular mechanics? To answer these questions, we have collected all pairwise interaction energies for all QM/MM-optimized structures calculated with the CHARMM27 force field (see Table 6 for the 20-ns WT structure and Table S2 in the Supporting Information for the data of all enzyme variants and all snapshots). It is remarkable that although in the salt bridge interaction between TMA and Asp623 is predicted to be the strongest interaction by both methods in all structures, their strength differs, considerably, by 20 kcal/mol. The second most important interaction is the cation–π interaction between TMA and residue 692; however, the MM method overestimates this interaction by 6–9 kcal/mol, which exceeds by 50 % the QM/MM interaction energy.

Table 6 Effect of various factors (QM/MM geometry optimization, basis set, DFT functional, electrostatic embedding of point charges in the QM region, only MM energy calculation) on the obtained two-body interaction energies between TMA and the residues included in the QM region for the wild-type structure taken at 20 ns from the MD trajectory

Wild type 20 ns	Asp623	Trp692	Tyr714	Ile740	Tyr741
ωB97xD/cc-pVTZ/MM	−59.0	−10.8	−5.3	−1.6	−0.5
ωB97xD/6-31G*/MM	−59.3	−11.2	−5.5	−1.7	−0.7
B3LYP/cc-pVTZ/MM	−58.3	−1.9	−3.0	1.5	0.6
TPSSh/cc-pVTZ/MM	−58.5	−3.1	−3.1	1.2	0.6
M06-2X/cc-pVTZ/MM	−58.8	−9.7	−4.7	−0.8	0.0
ωB97xD/cc-pVTZ/MM[a]	−58.2[a]	−8.3[a]	−5.1[a]	−0.8[a]	−0.2[a]
E^{b}_{strain}	0.8	1.0	3.0	0.5	0.5
ωB97xD/cc-pVTZ[c]	−61.0	−14.1	−7.1	−4.0	−2.5
MM[a]	−17.2[a]	−15.9[a]	−6.8[a,d]	−0.8[a]	1.4[a]

Data are given in kcal/mol

[a] Calculated at the MM-optimized geometry of the snapshot taken from the MD trajectory

[b] E_{srain} was calculated as the energy difference between the geometry of the side chain adopted in the QM/MM-optimized structure and the minimum energy structure in the gas phase

[c] Gas-phase calculation on the QM region

[d] Discrepancy between QM/MM and MM values was much larger for all other structures

The third most important interaction is found between TMA and Tyr714, and both methods predict it to be around 3–5 kcal/mol, which seems to be consistent with the cation–π interaction energy strength calculated in the gas-phase complexes (∼14 kcal/mol), as the distance in the enzymatic pocket (∼5 Å) is much larger between the fragments than in the gas-phase van der Waals complex (∼4 Å). The interaction between Ile740 and Tyr741, and TMA are predicted to be very weak by both methodologies. The discrepancy observed between the QM/MM and MM interaction energies is in line with a recent study on protein–ligand binding energies that found 15 % difference between MM and QM energies even after adjustment for systematic differences [49].

The significant discrepancy between the calculated MM and QM/MM interaction energies suggests that the parameters of the choline group (that were taken from the CHARMM lipid force field [37]) should be treated with great care and may require further validation for enzymatic systems. We also note here that the original parameters were developed for the aim of reliable modeling of membranes, which may be the underlying reason that in the present simulation so large differences were found.

Conclusions

In the present study, we investigated cation–π interactions using a variety of methods focussing on the key enzyme of the membrane synthesis (phosphocholine cytidylyltransferase) of the pathogen of malaria. We have adopted a methodology for the computation of many-body interactions in proteins using QM/MM and showed that in our case the two-body terms overestimate the overall interaction energy of the ligand by about 10 %, while three-body interaction terms, which are dominated by cation–π interactions, counteract the two-body terms and by decreasing the strength of the other interactions. Using accurate QM/MM calculations of the binding energy, we obtained good agreement between the trends of the binding affinities of the wild-type and mutant enzymes. Furthermore, we tested the effect of various factors that could affect the accuracy and reliability of the calculations and sought approximations that could offer reduction in computing time without reducing reliability. We found that once well-equilibrated structures were reached in the MD simulations, very similar results can be obtained for the various snapshots taken from the trajectory. The calculated interaction energies were only slightly influenced by electrostatic embedding of the point charges in the QM/MM calculations, and by QM/MM geometry optimization. However, the calculated molecular mechanics interaction energies were off by more than 50 % for cation–π interactions involving a choline group and even large discrepancies were observed for its interactions with an acidic side chain. Therefore, careful validation of the force fields for the estimation of cation–π interaction energies is highly recommended before using the force field for virtual screening seems. Based on our results, we suggest to calibrate force fields based on fragment-based QM calculations on geometries obtained from MD simulations as this approach is likely to yield reliable binding energies at reduced computational cost.

Acknowledgments The authors thank Dr. Goedele Roos (ULB, Belgium), Gergely N. Nagy, and Dr. Andras T. Rokob (MTA TTK, Hungary) for careful reading of the manuscript and helpful discussions. We are grateful for the support of the New Széchenyi Plan TAMOP-4.2.2/B-10/1-2010-0009 and for the financial support of

OTKA Grant No. 108721. J.O. acknowledges receipt of a Bolyai János Research Fellowship. A.L. acknowledges the financial support of Richter Gedeon Talentum Foundation.

References

1. Cerny J, Hobza P (2007) Non-covalent interactions in biomacromolecules. Phys Chem Chem Phys 9:5291–5303
2. Cheng T, Li Q, Zhou Z, Wang Y, Bryant SH (2012) Structure-based virtual screening for drug discovery: a problem-centric review. AAPS J 14:133–141
3. Zacharias N, Dougherty DA (2002) Cation-pi interactions in ligand recognition and catalysis. Trends Pharmacol Sci 23:281–287
4. Dougherty DA (2013) The cation-π interaction. Acc Chem Res 46:885–893
5. Luhmer M, Bartik K, Dejaegere A, Bovy P, Reisse J (1994) The importance of quadrupolar interactions in molecular recognition processes involving a phenyl group. Bull Soc Chim Fr 131:603–606
6. Williams JH (1993) The molecular electric quadrupole-moment and solid-state architecture. Acc Chem Res 26:593–598
7. McCurdy A, Jimenez L, Stauffer DA, Dougherty DA (1992) Biomimetic catalysis of Sn2 reactions through cation-pi interactions—the role of polarizability in catalysis. J Am Chem Soc 114:10314–10321
8. Ma JC, Dougherty DA (1997) The cation-pi interaction. Chem Rev 97:1303–1324
9. Gallivan JP, Dougherty DA (1999) Cation-pi interactions in structural biology. Proc Natl Acad Sci USA 96:9459–9464
10. Martz E (2002) Protein explorer: easy yet powerful macromolecular visualization. Trends Biochem Sci 27:107–109
11. Scior T, Bender A, Tresadern G, Medina-Franco JL, Martinez-Mayorga K, Langer T, Cuanalo-Contreras K, Agrafiotis DK (2012) Recognizing pitfalls in virtual screening: a critical review. J Chem Inf Model 52:867–881
12. Wein S, Maynadier M, Bordat Y, Perez J, Maheshwari S, Bette-Bobillo P, Ba CTV, Penarete-Vargas D, Fraisse L, Cerdan R, Vial H (2012) Transport and pharmacodynamics of albitiazolium, an antimalarial drug candidate. Br J Pharmacol 166:2263–2276
13. Vial HJ, Wein S, Farenc C, Kocken C, Nicolas O, Ancelin ML, Bressolle F, Thomas A, Calas M (2004) Prodrugs of bisthiazolium salts are orally potent antimalarials. Proc Natl Acad Sci USA 101:15458–15463
14. Nagy GN, Marton L, Contet A, Ozohanics O, Ardelean L-M, Revesz A, Vekey K, Irimie FD, Vial H, Cerdan R, Vertessy BG (2014) Composite aromatic boxes for enzymatic transformations of quaternary ammonium substrates. Angew Chem Int Ed 53:13471–13476
15. Kwak BY, Zhang YM, Yun M, Heath RJ, Rock CO, Jackowski S, Park HW (2002) Structure and mechanism of ctp: phosphocholine cytidylyltransferase (licc) from *Streptococcus pneumoniae*. J Biol Chem 277:4343–4350
16. Veitch DP, Gilham D, Cornell RB (1998) The role of histidine residues in the hxgh site of ctp: phosphocholine cytidylyltransferase in ctp binding and catalysis. Eur J Biochem 255:227–234
17. Lee J, Johnson J, Ding Z, Paetzel M, Cornell RB (2009) Crystal structure of a mammalian ctp: phosphocholine cytidylyltransferase catalytic domain reveals novel active site residues within a highly conserved nucleotidyltransferase fold. J Biol Chem 284:33535–33548
18. Chai JD, Head-Gordon M (2008) Long-range corrected hybrid density functionals with damped atom–atom dispersion corrections. Phys Chem Chem Phys 10:6615–6620
19. Zhao Y, Truhlar DG (2008) The m06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four m06-class functionals and 12 other functionals. Theor Chem Acc 120:215–241
20. Becke AD (1993) Density-functional thermochemistry iii. The role of exact exchange. J Chem Phys 98:5648–5652
21. Tao JM, Perdew JP, Staroverov VN, Scuseria GE (2003) Climbing the density functional ladder: nonempirical meta-generalized gradient approximation designed for molecules and solids. Phys Rev Lett 91:146401
22. Grimme S, Antony J, Ehrlich S, Krieg H (2010) A consistent and accurate ab initio parametrization of density functional dispersion correction (dft-d) for the 94 elements H-Pu. J Chem Phys 132:154104
23. Kendall RA, Dunning TH, Harrison RJ (1992) Electron-affinities of the 1st-row atoms revisited—systematic basis-sets and wavefunctions. J Chem Phys 96:6796–6806
24. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA et al (2009) Gaussian 09. Revision A.1 edn. Gaussian, Inc., Wallingford, CT
25. Boys SF, Bernardi F (2002) The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors (reprinted from molecular physics, vol 19, p 553–566, 1970). Mol Phys 100:65–73
26. Glendening ED, Streitwieser A (1994) Natural energy decomposition analysis—an energy partitioning procedure for molecular-interactions with application to weak hydrogen-bonding, strong ionic, and moderate donor–acceptor interactions. J Chem Phys 100:2900–2909
27. Weinhold F (2012) Natural bond orbital analysis: a critical overview of relationships to alternative bonding perspectives. J Comput Chem 33:2363–2379
28. Nagy GN, Marton L, Krámos B, Oláh J, Révész Á, Vékey K, Delsuc F, Hunyadi-Gulyás É, Medzihradszky KF, Lavigne M, Vial H, Cerdan R, Vértessy BG (2013) Evolutionary and mechanistic insights into substrate and product accommodation of ctp: phosphocholine cytidylyltransferase from *Plasmodium falciparum*. FEBS J 280:3132–3148
29. Li H, Robertson AD, Jensen JH (2005) Very fast empirical prediction and rationalization of protein pk(a) values. Proteins: Struct, Funct, Bioinf 61:704–721
30. Bas DC, Rogers DM, Jensen JH (2008) Very fast prediction and rationalization of pk(a) values for protein-ligand complexes. Proteins Struct Funct Bioinform 73:765–783
31. Olsson MHM, Sondergaard CR, Rostkowski M, Jensen JH (2011) Propka3: consistent treatment of internal and surface residues in empirical pk(a) predictions. J Chem Theory Comput 7:525–537
32. Sondergaard CR, Olsson MHM, Rostkowski M, Jensen JH (2011) Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pk(a) values. J Chem Theory Comput 7:2284–2295
33. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79:926–935
34. Brooks BR, Brooks CL III, Mackerell AD Jr, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S, Caflisch A, Caves L, Cui Q, Dinner AR, Feig M, Fischer S, Gao J, Hodoscek M, Im W, Kuczera K, Lazaridis T, Ma J, Ovchinnikov V, Paci E, Pastor RW, Post CB, Pu JZ, Schaefer M, Tidor B, Venable RM, Woodcock HL, Wu X, Yang W, York DM, Karplus M (2009) Charmm: the biomolecular simulation program. J Comput Chem 30:1545–1614
35. MacKerell AD, Banavali N, Foloppe N (2000) Development and current status of the CHARMM force field for nucleic acids. Biopolymers 56:257–265

36. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. J Mol Graph Model 14:33–38
37. Feller SE, MacKerell AD (2000) An improved empirical potential energy function for molecular simulations of phospholipids. J Phys Chem B 104(31):7510–7515
38. Harvey JN (2004) Spin-forbidden co ligand recombination in myoglobin. Faraday Discuss 127:165–177
39. Tinker—home page. Tinker—software tools for molecular design. http://dasher.Wustl.Edu/tinker/. Accessed 5 Oct 2011
40. Ren P, Wu C, Ponder JW (2011) Polarizable atomic multipole-based molecular mechanics for organic molecules. J Chem Theory Comput 7:3143–3161
41. Hodges MP, Stone AJ, Xantheas SS (1997) Contribution of many-body terms to the energy for small water clusters: a comparison of ab initio calculations and accurate model potentials. J Phys Chem A 101:9163–9168
42. Xantheas SS (2000) Cooperativity and hydrogen bonding network in water clusters. Chem Phys 258:225–231
43. Padgett CL, Hanek AP, Lester HA, Dougherty DA, Lummis SC (2007) Unnatural amino acid mutagenesis of the gaba(a) receptor binding site residues reveals a novel cation-pi interaction between gaba and beta 2tyr97. J Neurosci 27:886–892
44. Glendening ED, Badenhoop JK, Reed AE, Carpenter JE, Bohmann JA, Morales CM, Weinhold F (2009) NBO 5.9. Theoretical Chemistry Institute, University of Wisconsin, Madison, WI
45. Olah J, van Bergen L, De Proft F, Roos G (2015) How does the protein environment optimize the thermodynamics of thiol sulfenylation? Insights from model systems to qm/mm calculations on human 2-cys peroxiredoxin. J Biomol Struct Dyn 33:584–596
46. Lonsdale R, Olah J, Mulholland AJ, Harvey JN (2011) Does compound I vary significantly between isoforms of cytochrome p450? J Am Chem Soc 133:15464–15474
47. Saha S, Sastry GN (2015) Cooperative or anticooperative: how noncovalent interactions influence each other. J Phys Chem B ASAP. doi:10.1021/acs.jpcb.5b03005
48. Xu M, Lill MA (2013) Induced fit docking, and the use of qm/mm methods in docking. Drug Discov Today Technol 10:e411–e418
49. Yilmazer ND, Korth M (2013) Comparison of molecular mechanics, semi-empirical quantum mechanical, and density functional theory methods for scoring protein–ligand interactions. J Phys Chem B 117:8075–8084