J Comput Aided Mol Des (2011) 25:729–742
DOI 10.1007/s10822-011-9448-7

A QM/MM study of the binding of RAPTA ligands to cathepsin B

Antonella Ciancetta · Samuel Genheden · Ulf Ryde

Received: 2 March 2011 / Accepted: 13 June 2011 / Published online: 24 June 2011

Abstract We have carried out quantum mechanical (QM) and QM/MM (combined QM and molecular mechanics) calculations, as well as molecular dynamics (MD) simulations to study the binding of a series of six RAPTA (Ru(II)-arene-1,3,5-triaza-7-phosphatricyclo-[3.3.1.1] decane) complexes with different arene substituents to cathepsin B. The recently developed QM/MM-PBSA approach (QM/MM combined with Poisson–Boltzmann solvent-accessible surface area solvation) has been used to estimate binding affinities. The QM calculations reproduce the antitumour activities of the complexes with a correlation coefficient (r^2) of 0.35–0.86 after a conformational search. The QM/MM-PBSA method gave a better correlation ($r^2 = 0.59$) when the protein was fixed to the crystal structure, but more reasonable ligand structures and absolute binding energies were obtained if the protein was allowed to relax, indicating that the ligands are strained when the protein is kept fixed. In addition, the best correlation ($r^2 = 0.80$) was obtained when only the QM energies were used, which suggests that the MM and continuum solvation energies are not accurate enough to predict the binding of a charged metal complex to a charged protein. Taking into account the protein flexibility by means of MD simulations slightly improves the correlation ($r^2 = 0.91$), but the absolute energies are still too large and the results are sensitive to the details in the calculations, illustrating that it is hard to obtain stable predictions when full flexible protein is included in the calculations.

Keywords QM/MM · Ligand-binding affinities · Ruthenium · Anticancer drugs · Cathepsin B · Continuum solvation · QM/MM-PBSA

Introduction

During the latest 30 years, the field of metal-based anticancer drugs has broaden the search for new active compounds to an increasing range of transition metals, leading to the design and the evaluation of the antitumour properties of a wide selection of organometallic compounds [1–3]. As a consequence, several new classes of potential anticancer drugs have emerged and among them, ruthenium-based complexes have shown to represent a promising alternative to the platinum-based anticancer therapy [4]. In particular, the use of Ru(III) complexes such as *trans*-Ru(imidazole)$_2$Cl$_4$ [5, 6] and *trans*-Ru(dimethylsulphoxide)(imidazole)Cl$_4$ [7, 8] as antitumour and antimetastatic agents, respectively, has drawn attention to this group of compounds.

The evidence that both these compounds readily interact with biological reductants [9–11] has suggested that the reduction of the Ru(III) pro-drug to Ru(II) is a required step to accomplish their biological activity (the so-called activation-by-reduction hypothesis) [12, 13]. According to this suggestion, the anticancer potential of Ru(II) compounds has been explored. In particular, a growing number of studies on Ru(II)-arene phosphine derivatives has

Electronic supplementary material The online version of this article (doi:10.1007/s10822-011-9448-7) contains supplementary material, which is available to authorized users.

A. Ciancetta
Dipartimento di Scienze del Farmaco, Università degli Studi "G. D'Annunzio" Chieti-Pescara, Via dei Vestini 31, 66100 Chieti, Italy

S. Genheden · U. Ryde (✉)
Department of Theoretical Chemistry, Lund University, Chemical Centre, P. O. Box 124, 221 00 Lund, Sweden
e-mail: ulf.ryde@teokem.lu.se

emerged [13, 14] and the in vitro and in vivo activity of ruthenium(II)-arene compounds incorporating 1,3,5-triaza-7-phosphatricyclo-[3.3.1.1] decane (PTA) as a co-ligand [15, 16] have been studied in detail [17–19].

Despite the large number of studies, the mechanism of action of these new complexes is still largely unknown. The poor correlation between the binding to DNA and their cytotoxicity has suggested that Ru(II)-arene PTA (RAPTA) compounds act through a mechanism different from that established for classical platinum anticancer drugs. Hence, it is believed that the interaction of the RAPTA complexes with intracellular or extracellular proteins [13, 20, 21] may represent their primary mechanism of action.

Among many potential anticancer targets that have been identified, cathepsin B (catB), an enzyme involved in cellular metabolism, has been shown to be implicated in tumour progression and metastasis processes [22, 23], and to represent a suitable target for the development of anti-metastatic drugs. Recently, the inhibitory properties of a series of RAPTA compounds against catB have been tested [24]. The study revealed that these complexes can inhibit this enzyme, suggesting their potential use against tumour growth and metastases. The inhibition is believed to arise from the coordination of the Ru(II) centre to a protein residue in the active site and docking studies have proposed that the catalytic cysteine (Cys) is involved in the binding.

This suggestion has been also supported by a quantum mechanical (QM) study of a series of organometallic compounds of general formula (arene)M(PTA)Cl_2, with arene = η^6-$C_{10}H_{14}$ or η^5-C_5Me_5) and M = Ru(II), Os(II), Rh(III), or Ir(III) [25]. Adducts with a model of the Cys residue in the catB active site revealed that the metal–sulphur binding is thermodynamically favoured for the active compounds containing Ru(II), Ru(III), and Os(II) centres, whereas the inactive Ir(III) and Rh(III) complexes give rise to weak M–S bonds.

However, to obtain an accurate evaluation of the binding free energy of structurally closely-related complexes to catB, not only the formation of the covalent M–S bond has to be considered, but also the non-covalent interactions between the metal-bound ligands and the neighbouring residues of the active site. Indeed, it is likely that the latter interactions fine-tune the inhibition properties of these complexes, as suggested by their narrow range of activities. To suitably take into account both types of contributions, a QM/MM (QM combined with molecular mechanics) treatment of the system is required [26]. Such a method is also indispensable to describe systems containing metallic centres for which no accurate force-field parameters are available, and, in the specific case of RAPTA-type complexes, to achieve accurate coordination geometries of the ligands to the active site.

QM/MM methods have been used to study enzyme mechanisms for a long time [26, 27]. More recently, they have also started to be used to study docking and ligand binding [28–32]. To that aim, the QM/MM energies are often supplemented by molecular dynamics (MD) simulations and additional energy terms, e.g. describing bulk solvation and entropy. Several such approaches have been suggested, e.g. the QM/MM-PBSA approach (QM/MM combined with Poisson–Boltzmann and surface-area solvation) [33–37] and the QM/MM linear response method [38, 39]. They have been used to gain structural insights into the active site, to predict and rank ligands, and to explain the variance in observed binding affinities.

In this paper, we use QM, MD, QM/MM, and QM/MM-PBSA methods to study the binding of a series of six RAPTA ligands (Fig. 1) to catB. This seems to be the first time such methods are used for the binding of a ligand that is composed of a metal complex. Considering the growing number of new classes of compounds based on various transition metals that have recently emerged, there is a strong need for studies aimed at evaluating the reliability and applicability of such methods also to those systems.

Methods

Quantum mechanical calculations

Quantum mechanical (QM) calculations were carried out with density functional theory (DFT), using the τ-dependent gradient-corrected functional of Tao et al. [40] (TPSS), as implemented in the Turbomole package [41]. The def2-SV(P) basis set by Schäfer et al. [42, 43] was used for all the atoms, implying a 28-electron relativistic effective core potential for Ru. The calculations were sped up by expansion of the Coulomb interactions in auxiliary basis sets, the resolution-of-identity approximation [44, 45]. Structures were optimised until the change in energy between two iterations was below 2.6 J/mol (10^{-6} a.u.) and the maximum norm of the internal gradients was below 10^{-3} a.u.

Continuum solvation effects were estimated by single-point calculations using the continuum conductor-like screening model (COSMO) [46, 47]. These calculations were performed at the same level of theory as the geometry optimisation and with default values for all parameters (implying a water-like probe molecule) and a dielectric constant of 4 (to simulate a protein-like environment) and 80 for the water. For the generation of the cavity, we used the optimised COSMO radii in Turbomole (1.30, 2.00, 1.83, 1.72, 2.05, and 2.16 Å for H, C, N, O, Cl, and S, respectively [48]) and 2.17 and 2.00 Å for phosphorus and ruthenium, respectively.

Zero-point energies and thermal correction to the Gibbs free energy (at 298 K and 1 atm pressure) were calculated

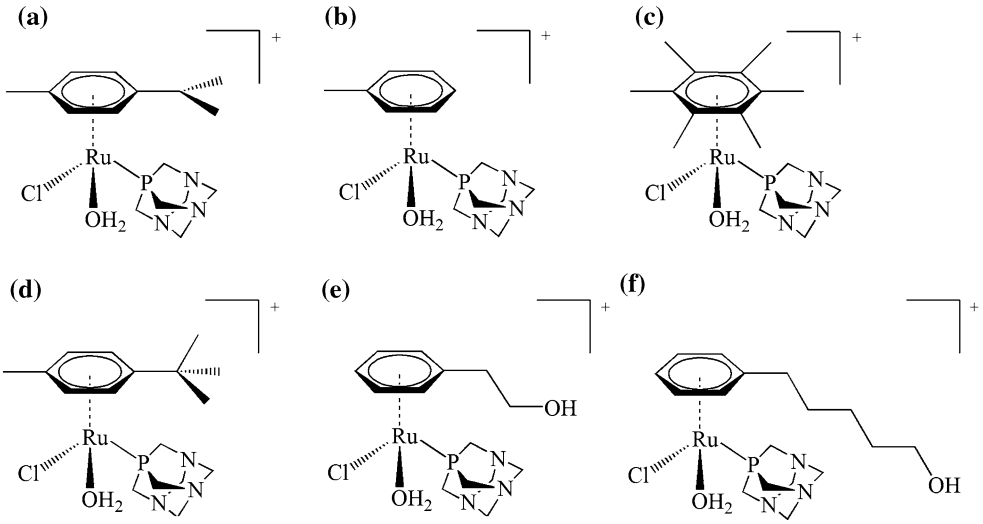


Fig. 1 Structures of the six RAPTA analogues studied: **a** RAPTA-C, **b** RAPTA-T, **c** RAPTA-H, **d** RAPTA-TBMe, **e** RAPTA-OH, and **f** RAPTA-PentaOH

from frequency calculations, performed at the same level of theory as the geometry optimisation, using a normal-mode, harmonic-oscillator, and ideal-gas approximation [49]. To obtain more accurate energies, single-point calculations were performed with the three-parameter hybrid functional B3LYP method [50, 51] using the larger def2-TZVP basis set [44].

The protein

The calculations in this investigation are based on the X-ray structure of human cathepsin B (catB) in complex with 2-pyridinethiol, PDB access code 2IPP (2.15 Å resolution, *R* factor 0.160) [52]. As shown schematically in Fig. 2, catB consists of a heavy and a light subunit, linked by six disulphide bridges. The catalytic domain resides in the light chain and consists of a deprotonated cysteine residue (Cys-29). Both subunits and all the crystal water molecules were used in the calculations.

Hydrogen atoms were added with the leap module of the AMBER software [53], assuming that all Asp and Glu residues are negatively charged and all Lys and Arg residues are positively charged according to their normal protonation state at pH 7.0. The protonation state of the His residues was determined by inspection of the hydrogen bond pattern, the surroundings, and the solvent exposure of each residue: His-45 and 97 were assumed to protonated on the $N^{\delta 1}$ atom, whereas the remaining histidine residues (His-110, 111, 190, 199, 239) were assumed to be doubly protonated on both the $N^{\delta 1}$ and $N^{\varepsilon 1}$ atom (and therefore positively charged). The protonation states of all residues were checked with the PROPKA software [54]. This charge-assignment gave a total charge of −6 for the free protein (−5 with the RAPTA ligands bound).

The protein was solvated in a sphere of explicit TIP3P water molecules [55] with a radius of 36 Å.

Fig. 2 Cathepsin B complexed with 2-pyridinethiol (*orange*) from the 2IPP file [52]. The sulphur atom of the catalytic cysteine is shown in *yellow*. The light chain is *blue* and the heavy chain is *cyan*. The *dark blue* spheres are crystal water molecules

Approximatively 4,286 water molecules were added, yielding a total of ∼17,040 atoms in the simulations. The added water molecules were kept inside the sphere by a force constant of 6.3 kJ/mol/Å^2. The position of all the hydrogen atoms and added water molecules were first minimised and then equilibrated for 300 ps using a simulated annealing protocol (i.e. heating the system to 370 K and cooling it slowly to 0 K), keeping the heavy atoms in their crystallographic positions.

In this investigation, we study the six RAPTA ligands (Fig. 1) that were docked into catB by Casini et al. [24]. Starting from the equilibrated structures of the enzyme, we followed two different protocols. In the first procedure, we removed the added water molecules (but retained the

crystal water molecules) and then placed the ligands into the active site according to the docked structures [24]. After graphical inspection, we removed overlapping crystal water molecules for each ligand and the resulting ligand–protein complex was solvated in a sphere of explicit TIP3P water molecules with a radius of 36 Å and was subjected to minimisation and equilibration as described above, yielding six different ligand–protein complexes.

In the second procedure, we superimposed the docked ligands into the active site of the fully solvated protein and, after graphical inspection, removed the water molecules overlapping with any of the ligands, yielding exactly the same protein and water surroundings for all six complexes. The resulting ligand–protein complexes were then subjected to QM/MM optimisation with no further minimisation or equilibration.

QM/MM calculations

The QM/MM calculations were carried out with the ComQum program [56, 57]. In the current version, it combines Turbomole 5.9 [41] for the QM part with AMBER 10 [53] for the MM part. In this approach, the protein and solvent are split into three subsystem: The QM region (system 1) contains the Ru(II) ion with all its first-sphere molecules (shown in Fig. 1), including either a water molecule for the free ligand, or a CH_3S model of Cys-29, when bound to the protein. This region was relaxed by QM methods. System 2 consists of all residues within 8 Å of any atom in system 1 and was relaxed by a full MM minimisation in each step of the optimisation of the QM system (protein-free calculations) or kept fixed (protein-fixed calculations). Finally, system 3 contains the remaining part of the protein and surrounding solvent molecules and was kept fixed at the original structure (the crystal structure for the protein).

Geometry optimisations were carried out in two steps. First, systems 2 and 3 were kept fixed at the original coordinates and only the quantum system was optimised (geometries obtained at the end of this optimisation step will be referred to as protein fixed). Second, systems 1 and 2 were allowed to relax, whereas system 3 was still kept fixed (protein free). In each step of the QM geometry optimisation of system 1, system 2 was fully relaxed by MM. In this optimisation of system 2, the charges on the quantum atoms were updated after each QM energy calculation [57] using the Merz-Kollman approach [58]. In the "Results" section, the protein-fixed structures are based on the second protocol of the protein setup, whereas the protein-free structures are based on the first protocol.

In the QM calculations, system 1 is represented by a wavefunction, whereas all of the other atoms are represented by partial point charges, one for each atom, taken from MM libraries. Thereby, the polarisation of the QM system by the surroundings is included in a self-consistent manner (electrostatic embedding).

In Cys-29, there is a bond between systems 1 and 2 (a junction), for which the hydrogen link-atom approach is used [59]: The QM system was capped with hydrogen atoms (H junction atoms), the position of which are linearly related to the corresponding carbon atoms (C junction atoms) in the full system [56]. In this investigation, charges on all atoms in the MM system around the junction atoms (besides the C junction atoms themselves) were included as point charges in the QM calculations [60].

The total QM/MM energy is calculated as [56, 57]:

$$E_{\mathrm{QM/MM}} = E_{\mathrm{QM1+ptch23}} - E_{\mathrm{MM1,no\_q}} + E_{\mathrm{MM123,no\_q1}} \quad (1)$$

where $E_{\mathrm{QM1+ptch23}}$ is the QM energy of the QM system truncated by a hydrogen atom, including the interaction of the quantum system with the surrounding point charges (MM part), but excluding the self energy of the point charges. $E_{\mathrm{MM1,no\_q}}$ is the MM energy of the QM system, still truncated by a hydrogen atom, but without any electrostatic interactions. Finally, $E_{\mathrm{M123,no\_q1}}$ is the classical energy of all atoms with a carbon link atom in Cys-29 and the charges of the QM system set to zero (to avoid double-counting of the electrostatic interactions). By this approach, which is similar to the one used in the Oniom method [61], errors caused by the truncation of the QM system should cancel.

The geometry optimisations were run until the energy change between two iterations was less than 2.6 kJ/mol (10^{-6} a.u.) and the maximum norm of the Cartesian gradients was below 10^{-3} a.u.

MM and MD calculations

All MM calculations were run with the sander module of the AMBER software [53], using the AMBER 1999SB [62–64] force field and the general AMBER force field (GAFF) [65] for the protein and the ligands, respectively. The Ru ion was described by only a non-bonded potential (because it was either fixed or treated by QM), using $r^* = 1.33$ Å and $\varepsilon = 1.56$ kJ/mol [66].

The QM system was represented by charges fitted to the electrostatic potential (ESP), calculated in the following way: The ligands (with a CH_3S model of the Cys-29 group) were optimised at DFT level with the TPSS functional and def2-SV(P) basis set. Then, the ESP was calculated with a single-point B3LYP calculation, employing the 6-31G* basis set for all atoms, except ruthenium for which LANL2DZ [67] basis set was used. The B3LYP method was used here instead of Hartree–Fock, which was used in the construction of the Amber 1999SB force field, because the latter method typically gives poor electronic structures

for transition metal complexes [68]. The ESP points were selected according to the Merz-Kollman scheme [58], but using a higher-than-default density of points ($\sim$2,500 points/atom). These calculations were performed with the Gaussian03 package [69]. Atomic charges were than fitted to the ESPs using the restrained electrostatic potential (RESP) procedure [70], as implemented in AMBER 10 [53]. The charge on the C^{β} carbon-link atom in Cys-29 was adapted so that the total charge of the protein (including both QM and MM atoms) was an integer (−5) [56]. Thereby, we allow charge transfer within the QM system (Cys-29 and the other Ru ligands have non-integer total charges). Moreover, the charge on the C junction atom is changed from what is typical for a hydrogen atom to what is more typical for carbon atoms.

In the MD simulations, bond lengths involving hydrogen atoms were constrained using the SHAKE algorithm [71]. The water solvent was described explicitly with the TIP3P model [55]. The electrostatics were treated with the particle-mesh Ewald method [72, 73] with a grid size of 80^3, a fourth-order B-spline interpolation, a tolerance of 10^{-5}, and a real-space cut-off of 8 Å. The temperature was kept constant at 300 K and the pressure was kept constant at 1 atm using the Berendsen weak-coupling algorithm [74] with a time constant of 1 ps. The MD time step was 2 fs and the non-bonded pair list was updated every 50 fs.

QM/MM-PBSA

The QM/MM-PBSA method [35] can be seen as a post-processing of QM/MM calculations to obtain more stable energies, including a proper solvation energy by using continuum solvation methods. It is an adaptation of the widely used MM/PBSA approach [75] for QM/MM calculations. In this approach, an approximation to the total free energy for each state is obtained from:

$$G = <E'_{\mathrm{QM/MM}}> + <G_{\mathrm{solv}}> + <G_{np}> - T<S_{\mathrm{QM/MM}}> \quad (2)$$

where $E'_{\mathrm{QM/MM}}$ is a QM/MM energy, slightly different from that in Eq. 1, as is discussed below, and G_{solv} is the polar solvation energy, estimated by a continuum approach, in our case obtained either by solving the Poisson–Boltzmann (PB) equation by the PB solver in AMBER [53] or by the generalised Born method GB^{OBC2} (with α, β, and γ set to 1.0, 0.8 and 4.85, respectively) [76], as implemented in AMBER [53]. G_{np} is the non-polar solvation energy, estimated from the solvent-accessible surface area (SASA) [77], through the linear relation $G_{np} = 0.0227$ SASA (in Å^2) + 3.85 kJ/mol [75, 78], where SASA was determined with the Molsurf program [53, 77], using a probe radius of 1.4 Å. Finally, T is the temperature, and $S_{\mathrm{QM/MM}}$ is the entropy. In this paper, we assume that only the QM system gives significant contributions to the entropy. This entropy was calculated for the harmonic frequencies of the isolated QM systems, as described above.

Among several available variants of QM/MM-PBSA [35], we use version 2 in [35], in which the electrostatic interaction energy between the QM and the MM system is calculated at the MM level, to make this contribution pairwise separable and to avoid problems arising from the use of link-atoms [60]. However, two different ways to calculate the QM/MM energy were tested. In both cases, the QM/MM energy is calculated from the sum of three terms:

$$E'_{\mathrm{QM/MM}} = E_{\mathrm{QM1}} - E_{\mathrm{MM1}} + E_{\mathrm{MM123'}} \quad (3)$$

where E_{QM1} is the QM energy of the QM system with hydrogen link atom (without any point-charge model of the surroundings), $E_{\mathrm{MM123'}}$ is the MM energy of systems 1, 2, and 3, with carbon link atoms and full charges, but with all water molecules stripped off, and E_{MM1} is the MM energy of the QM system, still with carbon link atoms and full charges. The two variants differ in how the QM energy and the charges on the QM atoms are calculated. In the first approach, the QM energy and the charges are calculated from a vacuum wavefunction. In the second approach, the QM wavefunction is optimised with a point-charge model of system 2 and 3. However, the final energy and the QM charges were calculated without the point charges, without re-optimising the wavefunction. Thereby, the QM system is polarised by the MM surroundings and the cost of the polarisation is included in the calculations. The latter approach was used in previous studies [35]. Finally, the calculations could be either based on a set of snapshots from the MD simulations (the $<>$ brackets in Eq. 2 indicate averages over these snapshots) or on a single QM/MM optimised structure. This gives rise to a number of different combinations of approaches. However, in the "Results" section, only results with PB solvation and with the polarised wavefunction are given, because the other approaches (as is described in the supplementary material) gave similar or worse results.

The MD simulations for the QM/MM-PBSA calculations were set up in the following way: First, each ligand–protein complex was optimised by QM/MM, using a spherical system, as described above. Second, the system was further solvated and moved into a periodic truncated octahedral box (with at least 9 Å from the already solvated QM/MM system to the periodic boundary) and a 100-step minimisation was performed, keeping the QM system fixed and restraining all heavy atoms, besides water molecules, towards their position in the QM/MM structure by a force constant of 418 kJ/mol/Å^2. Then, two 20 ps MD simulations were run with the same heavy atoms restrained. The

Table 1 Experimental activities of the RAPTA ligands against bovine catB [24]

Name	IC_{50}^{a}	ΔG_{exp}^{b}
RAPTA-C	2.5	−32.2
RAPTA-T	1.5	−33.4
RAPTA-H	>200	<−21.2
RAPTA-TBMe	82	−23.5
RAPTA-OH	1.6	−33.3
RAPTA-PentaOH	7	−29.6

[a] IC_{50} values in μM from Casini et al. [24]

[b] ΔG_{exp} was estimated as $RT\ln(IC_{50})$ and is presented in kJ/mol

first simulation was run with a constant volume, the second simulation with a constant pressure. Finally, the box size was equilibrated by a 50 ps MD simulation with a constant pressure, with only the heavy atoms in the QM system restrained. Since it is not possible to keep the QM atoms fixed in the constant-pressure simulation, the QM atoms were moved back to the QM/MM structure before an equilibration of 200 ps and a production of 400 ps were run with a constant volume and with the QM system fixed. During the production run, snapshots were collected every 2 ps.

In the QM/MM and QM/MM-PBSA energy calculations, no periodic boundary conditions were used. Instead, the MM system was centred on the QM system and an infinite cut-off was employed. The QM/MM-PBSA calculations were automatised and performed by Linux shell scripts, which are available from the authors on request. Further details of the calculations are available in http://www.teokem.lu.se/~ulf/Methods/qmmmpbsa.html.

Result and discussion

In this investigation, we have studied the binding of the six structurally related RAPTA ligands in Fig. 1 to the protein cathepsin B (catB), with known experimental affinities (IC_{50}), cf. Table 1 [24]. Different methods to predict the structure of the protein–ligand complexes and to estimate the binding free energies have been explored.

Vacuum calculations

First, we studied the intrinsic affinity of the six RAPTA ligands to the cysteine ligand using QM calculations in vacuum and in a continuum solvent. All the considered complexes are expected to undergo rapid hydrolysis in water to give water-derivatives, which easily undergo further substitution by endogenous nucleophiles. Experimental studies have shown that the most abundant hydrolysis product of RAPTA-C is the mono-aquo species $[Ru(\eta^6\text{-arene})(PTA)Cl(H_2O)]^+$, shown in Fig. 1a [79]. Therefore, the same species has been assumed for all the six ligands. Thus, it was assumed that inhibition of catB involves the substitution of the water ligand on the metal centre of the respective mono-aquo species by the sulphur atom of the catalytic Cys-29 residue site, which is believed to be the anchoring site for the ruthenium centre. The Cys-29 residue is expected to be in its anionic form at physiological pH and was therefore modelled by CH_3S^-. Consequently, we studied the model reaction in Scheme 1.

The Ru–ligand bond lengths are collected in Table 2. They are very similar among the six RAPTA complexes: For the CH_3S^- complexes, the Ru–Cl bond lengths are 2.42–2.44 Å, the Ru–P bonds are 2.31–2.33 Å, the Ru–S bonds are 2.44–2.45 Å, and the average Ru–C bonds are 2.24–2.26 Å (2.18–2.32 Å for the individual Ru–C bonds). Of course, the Ru–O bonds in the water complexes are shorter than the Ru–S bonds, 2.19–2.22 Å. Consequently, the Ru–P bonds are slightly longer, 2.35–2.37 Å, but the Ru–Cl and average Ru–C bonds are actually slightly shorter, 2.41–2.43 and 2.23–2.24 Å, respectively.

Next, we calculated the energy of the ligand-exchange reaction in Scheme 1. As described in the "Methods" section, we modelled the reaction in vacuum, as well in continuum solvents with a dielectric constant of 4 or 80. We also calculated reaction energies with or without thermal corrections to the Gibbs free energy from the vibrational frequencies (ΔE and ΔG). The results of these calculations are collected in Table 3.

The results show that the ligand-exchange energies strongly depend on the dielectric constant of the continuum solvation model: For example, ΔE for RAPTA-C decreases from −524 kJ/mol in vacuum to −112 kJ/mol in a water-like solvent. The reason for this is of course that the

Scheme 1 The reaction studied for the calculation of the binding free energies at QM level

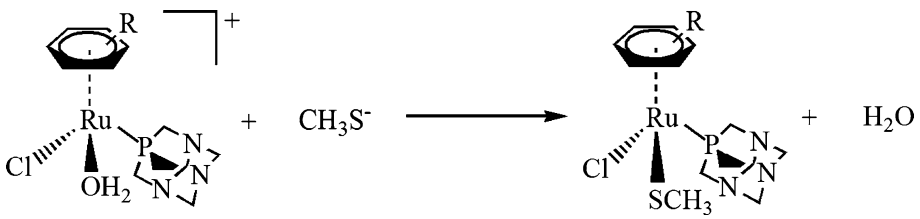


Table 2 Main geometrical parameters (bond lengths in Å) in the optimised vacuum (with either a H_2O or a CH_3S^- ligand) and QM/MM structures (with either the surrounding protein fixed, Fix, or allowed to relax, Free)

	Vacuum		QM/MM	
	H_2O	CH_3S^-	Fix	Free
	Ru–O		Ru–S	
RAPTA-C	2.20	2.44	2.49	2.49
RAPTA-T	2.19	2.44	2.54	2.43
RAPTA-H	2.22	2.44	**4.14**	2.44
RAPTA-TBMe	2.20	2.45	2.47	2.44
RAPTA-OH	2.20	2.44	2.30	2.43
RAPTA-PentaOH	2.20	2.44	2.61	2.42
	Ru–P			
RAPTA-C	2.36	2.32	2.56	2.41
RAPTA-T	2.35	2.33	2.43	2.39
RAPTA-H	2.37	2.31	2.33	2.38
RAPTA-TBMe	2.36	2.32	2.59	2.38
RAPTA-OH	2.36	2.32	**4.89**	2.35
RAPTA-PentaOH	2.35	2.32	2.45	2.39
	Ru–Cl			
RAPTA-C	2.41	2.43	2.29	2.39
RAPTA-T	2.42	2.42	2.29	2.39
RAPTA-H	2.43	2.43	2.30	2.37
RAPTA-TBMe	2.42	2.44	2.35	2.38
RAPTA-OH	2.42	2.43	2.33	2.38
RAPTA-PentaOH	2.42	2.43	2.35	2.37
	$Ru–C_{av}$			
RAPTA-C	2.23	2.24	2.32	2.27
RAPTA-T	2.23	2.24	2.33	2.24
RAPTA-H	2.24	2.26	2.23	2.28
RAPTA-TBMe	2.24	2.25	2.31	2.25
RAPTA-OH	2.22	2.24	2.23	2.23
RAPTA-PentaOH	2.23	2.24	2.27	2.24
	Cl–HN			
RAPTA-C			2.59	**4.77**
RAPTA-T			2.51	**3.40**
RAPTA-H			2.59	2.84
RAPTA-TBMe			3.05	2.95
RAPTA-OH			3.16	2.98
RAPTA-PentaOH			2.80	2.97

$Ru–C_{av}$ is the average of the six Ru–C bonds to the arene. Cl–HN is the hydrogen bond between the Cl ligand and the backbone NH group of Cys-29. Deviating distances are marked in bold face

binding involves two charged reactants but two neutral products. Therefore, solvation will strongly favour the reactants.

The energies presented in Table 3 are obtained by single-point B3LYP/def2-TZVP energy calculations on the TPSS/def2-SV(P) structures. This improvement of the functional and basis set has a rather small effect, which is almost the same for all ligands, making the predicted affinities 24–29 kJ/mol more positive (cf. Table S1 in the supplementary material). The effect of zero-point energies and entropies is even smaller, but it is more varying. It is positive and 5–13 kJ/mol for all ligands, except for RAPTA-pentaOH, with its long tail, for which it is −6 kJ/mol.

The energy of this exchange reaction can be used as an estimate of the intrinsic affinity of the six RAPTA ligands for the catalytic Cys-29 residue in catB. In absolute terms, the predicted affinities are too negative, even in water, indicating that some terms still are missing. However, in relative terms, the affinities are reasonable, giving a range of values that are similar to the experimental range, e.g. 13–23 kJ/mol for the ΔE results, compared to 12 kJ/mol for the experimental estimates. The range decreases with the dielectric constant.

This small range of the experimental results makes it a formidable task to reproduce the experimental affinities with theoretical methods. Still, as can be seen in Table 3,

Table 3 Binding energies obtained from the QM calculations in a continuum solvent with three different values of the dielectric constant (ε)

	ΔE			ΔG			Exp.
	$\varepsilon = 1$	$\varepsilon = 4$	$\varepsilon = 80$	$\varepsilon = 1$	$\varepsilon = 4$	$\varepsilon = 80$	
RAPTA-C	−524.3	−245.6	−112.3	−519.7	−241.0	−107.7	−32.2
RAPTA-T	−512.6	−234.0	−101.6	−503.4	−224.8	−92.4	−33.4
RAPTA-H	−505.2	−230.5	−99.6	−497.7	−222.9	−92.0	<−21.2
RAPTA-TBMe	−511.2	−237.0	−106.2	−498.5	−224.3	−93.5	−23.5
RAPTA-OH	−528.0	−246.5	−111.5	−523.2	−241.7	−106.7	−33.3
RAPTA-PentaOH	−513.9	−238.7	−106.8	−520.0	−244.9	−113.0	−29.6
r^2	0.57	0.41	0.30	0.52	0.35	0.26	
Range	23	16	13	26	22	21	12

All energies (in kJ/mol) are obtained by single-point B3LYP/def2-TZVP energy calculations on the TPSS/def2-SV(P) structures. The ΔG values include zero-point and thermal corrections to Gibbs free energy, calculated from frequencies obtained at the TPSS/def2-SV(P) level

the calculated affinities show some correlation to the experimental data, even for these small-model QM calculations, with correlation coefficients (r^2) of 0.26–0.57. The correlation is best in vacuum and without thermal corrections. The B3LYP/def2-TZVP results are slightly better than the raw TPSS/def2-SV(P) results. The results are plotted in Fig. 3. It can be seen that RAPTA-T gives the largest deviation, having a too low affinity.

The optimised vacuum structures of the Ru(η^6-arene)(PTA)Cl(CH_3S) and [Ru(η^6-arene)(PTA)Cl(H_2O)]$^+$ complexes of RAPTA-C are shown in Fig. 4. It can be seen that the structure is mainly determined by possible weak interactions between the Cl^-, water, and CH_3S^- ligands and the CH groups of the arene and PTA ligands. Thus, the complexes are rather flexible. Therefore, it is conceivable that docked conformations used in this investigation are not optimal for all ligands. In order to test that possibility, we performed a conformational search for all six ligands, except the symmetric RATA-H, by systematically rotating the arene ligand for both the CH_3S and H_2O complexes in steps of 30°, calculating TPSS/def2-SV(P) energies in both vacuum and in water.

This procedure gave improved structures for all complexes, except two (RAPTA-T and RAPTA-OH with a water ligand), with the energies lowered by up to 18 kJ/mol. The searches in vacuum and water gave identical results in all cases, except for two water complexes (RAPTA-TBMe and RAPTA-T), but the difference in energy was less than 1 kJ/mol. The new structures have nearly identical bond lengths to Ru, with differences less than 0.02 Å (Table S2).

Using these structures, energies for the ligand-exchange reaction in Scheme 1 were calculated in the same way as before. The results in Tables 4 and S3, as well as in Fig. 5 show that the results are significantly improved by the conformation search. The correlation coefficients are now 0.35–0.86, with similar trends as before the conformational

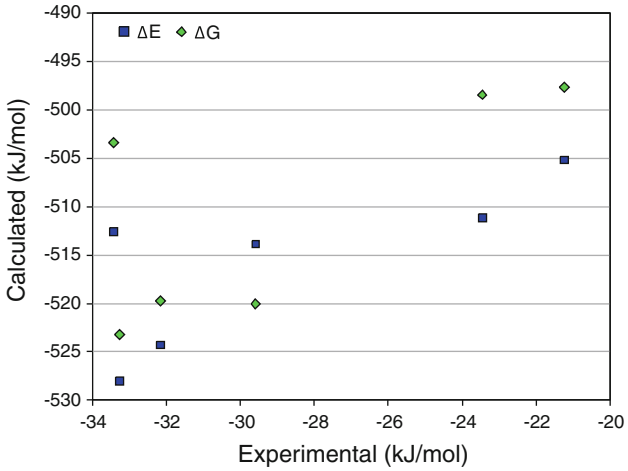


Fig. 3 Correlation between the experimental affinities and those calculated from the optimised QM structures in vacuum. Results are shown for ΔE and ΔG without and with thermal corrections

search, with the exception that the best ΔE results are obtained in a protein-like environment with a dielectric constant of 4. In fact, as can be seen from Fig. 4, the calculations without thermal corrections now rank all the ligands in the correct order, whereas the ΔG calculations have problem with RAPTA-C.

These results suggest that we can quite successfully use the intrinsic QM energy of the substitution of the water ligand with the sidechain of Cys-29 to explain the different activities of the RAPTA complexes, provided that a proper conformational search is performed.

QM/MM geometries

Next, we investigated whether the predictions can be improved by take the surrounding protein into consideration by the QM/MM approach. In the QM/MM calculations, we still assume that the RAPTA ligands bind to the

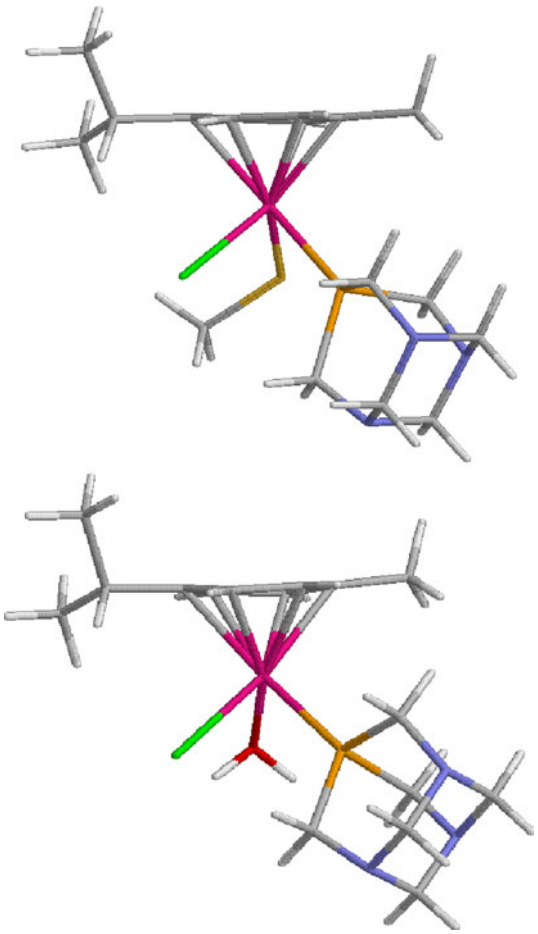


Fig. 4 Optimised QM structures of RAPTA-C with a CH_3S^- (*top*) or H_2O (*bottom*) ligand

sidechain of Cys-29 in catB, replacing a water ligand, i.e. we studied the ligand-exchange reaction in Scheme 2. The Cys-29 ligand was modelled by CH_3S^- in the QM system with a link atom replacing C^{α}. It is well-known that QM/MM energies are sensitive to differences in the conformation of the surrounding protein. Therefore, we employed two sets of structures. In the first (called protein fixed), the surrounding protein was not allowed to move, only the RAPTA ligand and the side chain of Cys-29 (i.e. the QM system). Thereby, we ensure that all protein structures are identical, so that the energies are directly comparable. On the other hand, this does not allow the surroundings to adapt to the ligand. In the second set of structures (called protein free), the surrounding protein was also optimised.

In the two last columns of Table 2, the main geometrical parameters are collected for the complexes after the QM/MM optimisations: the Ru–S, Ru–P, Ru–Cl, and the average Ru–C bond lengths, as well as the distance between the chloride ligand and the backbone HN group of Cys-29. The latter interaction was found to play an important role in a previous docking study [24].

It can be seen that the Ru–Cl bond length is similar in all optimised structures, 2.29–2.39 Å. This is expected, because Cl is a small ligand that is not significantly affected by the surroundings. The bond length is slightly shorter than in the vacuum (2.41–2.43 Å). It is also shorter in the protein-fixed structures than in the protein-free structures (by 0.03–0.10 Å).

The Ru–P bond is also similar in most complexes, 2.33–2.59 Å. However, for the RAPTA-OH ligand, the PTA ligand dissociates from Ru if the protein is not allowed to relax and moves to a hydrophilic pocket. However, if the protein is also relaxed, the PTA ligand remains bound to Ru. The PTA ligand also remains bound to Ru when the other solvation protocol is used without any relaxation of the protein, although it is elongated (2.79 Å; Table S2). The Ru–P distances in the relaxed structures are more similar than in the fixed structures, 2.35–2.41 Å. Still the QM/MM Ru–P bond lengths are slightly longer than in the vacuum structures (2.31–2.33 Å).

The Ru–S bond shows a similar behaviour. If the protein is allowed to relax, it is 2.42–2.49 Å, i.e. similar to what is found in the vacuum structures (2.44–2.45 Å). However, if the protein is kept fixed, it is typically longer (2.47–2.61 Å), except in the RAPTA-OH structure, for which the PTA ligand dissociated (2.30 Å). Moreover, for RAPTA-H, this bond is broken (Ru–S = 4.14 Å; the same applies for the other solvation protocol). Altogether, these results indicate that a conformational relaxation of the protein is needed before the RAPTA ligands can bind properly, especially for the large RAPTA-H complex.

The change in the structures of RAPTA-H with a fixed or relaxed protein is illustrated in Fig. 6. It can be seen that Ru ion and the PTA group hardly move ($\sim$0.3 Å), whereas the arene rotates by $\sim$20°. On the other hand, as the Cys ligand dissociates, the Cl ion nearly takes its former position.

Finally, the Ru–C bond lengths are similar to those in the vacuum optimisations when the protein is allowed to relax (2.23–2.28 Å, compared to 2.22–2.24 Å). However, for most of the ligands, the Ru–C_{av} distance is longer if the protein is kept fixed (2.27–2.33 Å), except for the two complexes with a dissociated ligand (2.23 Å), again indicating that the protein-fixed complexes are somewhat strained.

In most of the complexes, the binding of the RAPTA ligands is enhanced by a hydrogen bond between the chloride anion and the backbone NH group of Cys-29. With a fixed protein, the hydrogen bond length is 2.51–2.80 Å, but for the RAPTA-TBMe and RAPTA-OH complexes, this bond is almost broken (3.05–3.16 Å). Interestingly, the Cl–HN hydrogen bond is broken (RAPTA-C and RAPTA-T) or weakened when the protein is allowed to relax. This may be an effect of the shortened Ru–S bond.

Table 4 Binding energies obtained from the QM calculations after the conformational search. Results are presented from calculations in a continuum solvent with three different values of the dielectric constant (ε)

	ΔE			ΔG			Exp.
	$\varepsilon = 1$	$\varepsilon = 4$	$\varepsilon = 80$	$\varepsilon = 1$	$\varepsilon = 4$	$\varepsilon = 80$	
RAPTA-C	−517.8	−239.1	−106.1	−510.1	−231.4	−98.4	−32.2
RAPTA-T	−528.5	−246.3	−111.5	−524.3	−242.1	−107.3	−33.4
RAPTA-H	−505.2	−230.5	−99.6	−497.7	−222.9	−92.0	<−21.2
RAPTA-TBMe	−512.2	−235.6	−103.9	−511.3	−234.7	−103.0	−23.5
RAPTA-OH	−525.6	−243.7	−108.6	−523.4	−241.5	−106.3	−33.3
RAPTA-PentaOH	−515.1	−239.9	−108.1	−518.3	−243.1	−111.3	−29.6
r^2	0.83	0.86	0.81	0.68	0.51	0.35	
Range	23	16	12	27	20	19	12

All energies (in kJ/mol) are obtained by single-point B3LYP/def2-TZVP energy calculations on the TPSS/def2-SV(P) structures. The ΔG values include zero-point and thermal corrections to Gibbs free energy, calculated from frequencies obtained at the TPSS/def2-SV(P) level

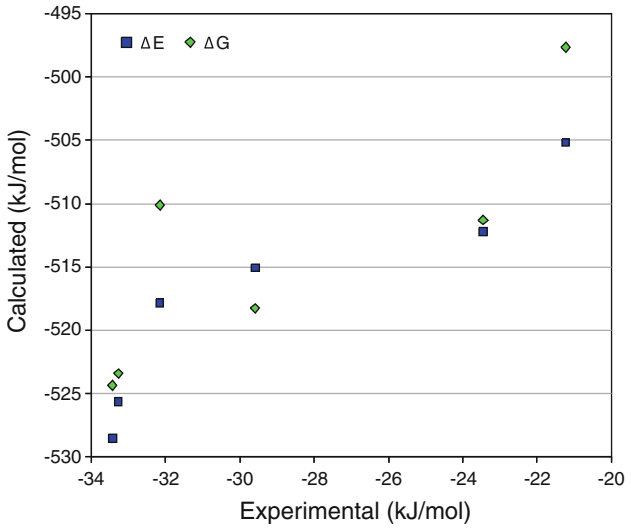


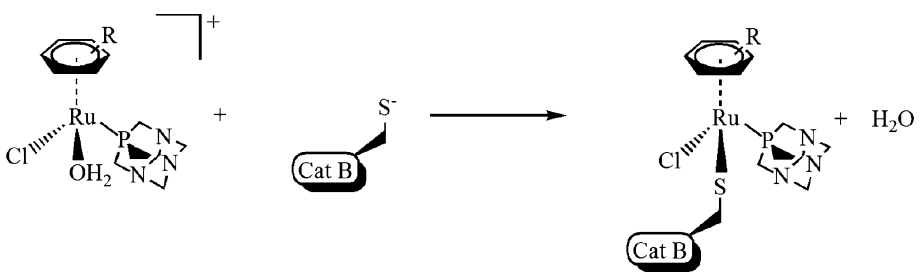
Fig. 5 Correlation between the experimental affinities and those calculated from the optimised QM structures in vacuum after the conformational search. Results are shown for ΔE without and ΔG with thermal corrections

QM/MM energies

Next, we calculated energies for the binding of the six RAPTA complexes, according to the reaction in Scheme 2. These energies are based on the QM/MM structures described in the previous section (i.e. both with a fixed and with a relaxed protein). To obtain more stable energies, we employed the QM/MM-PBSA procedure [35], according to Eq. 2, i.e. the QM/MM energies were supplemented by continuum solvation energies calculated by solving the PB equation and non-polar energies calculated from the change in the SASA during ligand binding. For the free ligand and the released water molecule, we used vacuum optimised structures, including solvation effects, calculated by the COSMO model. The free protein was modelled by a QM/MM-optimised structure with a deprotonated Cys-29 group. The charges of the QM system were obtained by a wavefunction polarised by the point charges of the surroundings. For comparison, QM energies for the isolated QM systems were also calculated for all geometries. Several other variants of the QM/MM-PBSA method were tested, as is detailed in the supplementary material (Tables S4–S6), e.g. using the generalised Born solvation energy or by using a vacuum wavefunction. However, they gave similar results.

The calculated energies are collected in the six first columns of Table 5. It can be seen that the calculations with a fixed protein give the best correlation with the experimental results, $r^2 = 0.59$. On the other hand, the calculations with a relaxed protein give an appreciably worse correlation, $r^2 = 0.00$. This indicates that the local-minima problem with the relaxed calculations may be serious. However, it is worth noting that the predicted

Scheme 2 The reaction studied for the calculation of the binding free energies at QM/MM level



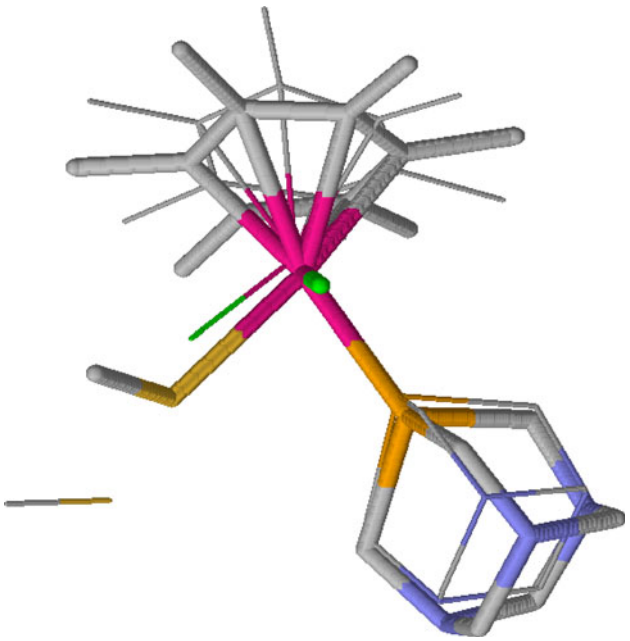


Fig. 6 Overlay of the QM/MM structures of RAPTA-H, obtained when the protein was fixed (*thin line*) or relaxed (*thick lines*). Only the QM system is shown

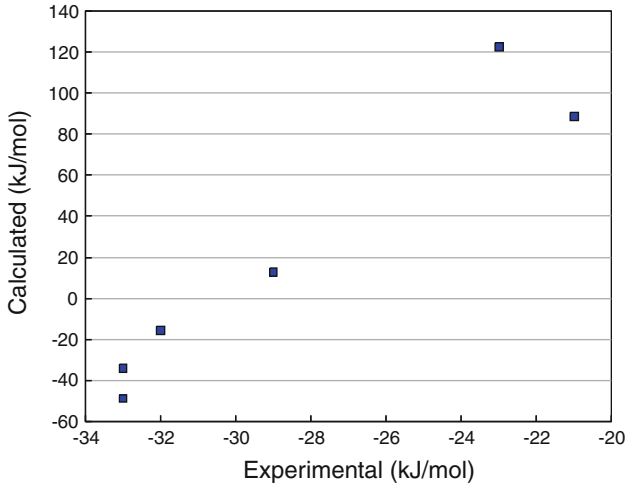


Fig. 7 Correlation between the experimental affinities and those calculated from the optimised QM/MM structures, followed by a MD simulation, from which QM/MM-PBSA energies were calculated for 200 snapshots

binding energies become much more favourable when the protein is allowed to relax (−239 to 500 kJ/mol, compared to 100–457 kJ/mol with the fixed protein). This shows that the ligands are severely strained when the protein is not allowed to relax. The range of the predicted binding energies among the six ligands is also somewhat reduced, from 357 to 96 kJ/mol with PB solvation, although it is still larger than the experimental range.

It can also be seen that for this system, the raw QM/MM energies give similar correlations ($r^2 = 0.58$ and 0.08), although the net binding energies change by about 530 and 730 kJ/mol for the protein-fixed and free calculations, respectively.

Interestingly, the correlations are always improved if only the QM energies are used (QM column in Table 5), to $r^2 = 0.80$ with the relaxed protein and 0.19 with the fixed protein. The range of the binding energies is also reduced, to 195 and 47 kJ/mol. This confirms that it is a formidable task for the continuum models to predict the binding of a charged metal complex to a charged protein.

Finally, we made an attempt to further study the effect of the surrounding protein and to reduce the local-minima problem by running a MD simulation for each catB–RAPTA complex with a fixed ligand structure (the QM/MM structure) but allowing the surrounding protein and solvent to fully relax. Snapshots were saved every 2 ps and QM/MM-PBSA energies were calculated for each snapshot. The results of these QM/MD-PBSA calculations are also collected in Table 5 (columns MD/Fix and MD/Free).

Table 5 QM/MM and QM/MM-PBSA binding energies (in kJ/mol) for the QM/MM structures, obtained with either the protein fixed (Fix) or free to relax (Free)

	Fix			Free			MD/Fix		MD/Free		Exp.
	QM/MM	PB	QM	QM/MM	PB	QM	PB	QM	PB	QM	
RAPTA-C	−336	189	−261	−795	−151	−451	−16 ± 2	−260 ± 1	−98 ± 2	−444 ± 1	−32
RAPTA-T	−376	133	−316	−897	−173	−497	−53 ± 4	−307 ± 1	−153 ± 1	−482 ± 1	−33
RAPTA-H	−270	254	−177	−936	−183	−450	88 ± 1	−182 ± 1	−142 ± 2	−435 ± 1	<−21
RAPTA-TBMe	−72	457	−145	−918	−180	−477	123 ± 2	−153 ± 1	−162 ± 1	−462 ± 1	−23
RAPTA-OH	−434	100	−340	−930	−195	−480	−34 ± 5	−343 ± 1	−200 ± 1	−478 ± 1	−33
RAPTA-PentaOH	−362	235	−211	−1,009	−246	−463	13 ± 2	−204 ± 1	−223 ± 2	−449 ± 1	−29
r^2	0.58	0.59	0.80	0.08	0.00	0.19	0.91	0.76	0.01	0.35	
Range	362	357	195	214	96	47	176	190	125	47	12

Results for 200 MD snapshots are also included (MD). For each structure, results are given both with PB solvation energies included or excluded (QM)

It can be seen that the binding energies for all complexes with a fixed protein are improved (lowered), by 134–343 kJ/mol, again illustrating that protein relaxation is important for the binding of the RAPTA ligands. For the protein-free calculations, the change is varying. The standard error of the energies are 1–5 kJ/mol, showing that the calculations are well converged and the precision is good. The MD simulations also suppress some of the variation among the ligands, but the range is still much larger than for the experimental affinities, 176 and 125 kJ/mol for the protein-fixed and protein-free calculations, respectively. Still, the calculations with a fixed protein give the best correlations to the experimental results, and with PB solvation, it is as high as $r^2 = 0.91$, as is illustrated in Fig. 7.

Conclusions

In this investigation, we have explored several QM-based methods to predict the structure of protein–ligand complexes and to estimate the binding free energies of a series of six Ru(II)-arene PTA (RAPTA) ligands to catB. QM calculations in vacuum and in a continuum solvent have been used to estimate the intrinsic binding affinity of the ligands, using CH_3S^- as model for the catalytic Cys-29 residue in the active site, and studying the energy of the exchange reaction of a water ligand with the sidechain model. Although the predicted absolute values are too negative, the calculated relative affinities are reasonable. The correlation coefficient (r^2) between the calculated and experimental affinities improved from 0.26–0.57 to 0.35–0.86 after a conformational search of the flexible arene group, suggesting that the intrinsic QM energy is quite successful in explaining the activities of the various RAPTA complexes, once the flexibility of the ligand is properly accounted for.

Next, we investigated whether the predictions can be improved by taking the surrounding protein into account by obtaining structures by QM/MM and binding energies by the QM/MM-PBSA procedure. We used two different approaches: In the first set, the surrounding protein was not allowed to move, in order to minimise the risk of ending up in different local minima for the surrounding protein before and after ligand binding. In the second set, the surrounding protein was also optimised, in order to allow structural relaxation upon ligand binding. The results showed that if the protein was not allowed to relax, some of the metal ligands dissociated during ligand binding and all structures had elongated bond lengths. Moreover, the binding becomes stronger with all methods if the protein also was optimised, thus indicating that protein relaxation is important during ligand binding.

Unfortunately, the best correlation between experimental and calculated affinities was obtained with the fixed protein ($r^2 = 0.58$–0.80), whereas the calculations with a relaxed surrounding gave poor correlations ($r^2 = 0.00$–0.19), although the range of the calculated affinities is smaller. This indicates severe problems with local-minima in the MM region. Moreover, for both sets, the correlations were improved if only the QM energies were used, suggesting that the MM energies and the continuum solvation model are not able to accurately predict the binding of a charged metal complex to a charged protein.

In an attempt to reduce the local-minima problem, we performed MD simulations for each catB–RAPTA complex with the ligand fixed at the QM/MM structure, but allowing the protein to fully relax. These QM/MD-PBSA results showed that the binding energies for all QM/MM calculations with a fixed protein are improved, whereas for the protein-free calculations, the change is smaller and more varying. These results further emphasize that the protein relaxation is important for the binding of the RAPTA ligands. The standard error of the energies is 1–5 kJ/mol, showing that the calculations are well converged and the precision is good. In most cases, the correlation to the experimental results is slightly improved (to $r^2 = 0.01$–0.91), but the results depend quite strongly on the details of the calculations.

Altogether, these results suggest that both the flexibility of the ligands and the surrounding protein play a role in the binding of the six RAPTA complexes to catB. Unfortunately, available methods seem to have problems to describe in a balanced way both solvation effects and the interaction of the ligand with the protein. Therefore, the best results are not obtained with the seemingly best methods (QM/MM-PBSA energies with MD sampling for the relaxed structures). Instead, the protein-fixed calculations and the raw QM energies often give better results. In fact, the QM-only calculations give nearly equally good results as the QM/MM calculations.

Acknowledgments We thank Dr. Alessandro Marrone who provided us the initial docked structures. This investigation has been supported by grants from the Swedish research council (project 2010-5025) and from the Research school in pharmaceutical science. It has also been supported by computer resources of Lunarc at Lund University and HPC2N at Umeå University.

References

1. Hartinger CG, Dyson PJ (2009) Chem Soc Rev 38:391–401
2. Nobili S, Mini E, Landini I, Gabbiani C, Casini A, Messori L (2010) Med Res Rev 30:550–580
3. Gianferrara T, Bratsos I, Alessio E (2009) Dalton Trans 37: 7588–7598
4. Clarke MJ, Zhu F, Frasca DR (1999) Chem Rev 99:2511–2534
5. Hartinger CG, Zorbas–Seifried S, Jakupec MA, Kynast B, Zorbas H, Keppler BK (2006) J Inorg Biochem 100:891–904

6. Groessl M, Hartinger CG, Polec-Pawlak K, Jarosz M, Dyson PJ, Keppler BK (2008) Chem Biodivers 5:1609–1614
7. Sava G, Capozzi I, Bergamo A, Gagliardi R, Cocchietto M, Masiero L, Onisto M, Alessio E, Mestroni G, Garbisa S (1996) Intern J Cancer 68:60–66
8. Rademaker-Lakhai JM, Van den Bongard D, Pluim D, Beijnen JH, Schellens JHM (2004) Clin Cancer Res 10:3717–3727
9. Sava G, Bergamo A, Zorzet S, Gava B, Casarsa C, Cocchietto M, Furlani A, Scarcia V, Serli B, Iengo E, Alessio E, Mestroni G (2002) Eur J Cancer 38:427–435
10. Schluga P, Hartinger CG, Egger A, Reisner E, Galanski M, Jakupec MA, Keppler BK (2006) Dalton Trans 14:1796–1802
11. Brindell M, Piotrowska D, Shoukry AA, Stochel G, Eldik R (2007) J Biol Inorg Chem 12:809–818
12. Clarke MJ, Zhu F, Frasca DR (1999) Chem Rev 99(9):2511–2534
13. Dyson PJ, Sava G (2006) Dalton Trans 16:1929–1933
14. Yan YK, Melchart M, Habtemariam A, Sadler PJ (2005) Chem Commun 38:4764–4776
15. Phillips AD, Gonsalvi L, Romerosa A, Vizza F, Peruzzini M (2004) Coord Chem Rev 248:955–993
16. Bravo J, Bolaño S, Gonsalvi L, Peruzzini M (2010) Coord Chem Rev 254:555–607
17. Allardyce CS, Dyson PJ, Ellis DJ, Heath SL (2001) Chem Commun 15:1396–1397
18. Scolaro C, Bergamo A, Brescacin L, Delno R, Cocchietto M, Laurenczy G, Geldbach TJ, Sava G, Dyson PJ (2005) J Med Chem 48:4161–4171
19. Bergamo A, Masi A, Dyson PJ, Sava G (2008) Intern J Oncol 33:1281–1289
20. Casini A, Gabbiani C, Michelucci E, Pieraccini G, Moneti G, Dyson PJ, Messori L (2009) J Biol Inorg Chem 14:761–770
21. Casini A, Karotki A, Gabbiani C, Rugi F, Vasak M, Messori L, Dyson PJ (2009) Metallomics 1:434–441
22. Mohanam S, Jasti SL, Kondraganti SR, Chandrasekar N, Lakka SS, Kin Y, Fuller GN, Yung AWK, Kyritsis AP, Dinh DH, Olivero WC, Gujrati M, Ali–Osman F, Rao JS (2001) Oncogene 20:3665–3673
23. Koblinski JE, Ahram M, Sloane BF (2000) Clin Chim Acta 291:113–135
24. Casini A, Gabbiani C, Sorrentino F, Rigobello MP, Bindoli A, Geldbach TJ, Marrone A, Re N, Hartinger CG, Dyson PJ, Messori L (2008) J Med Chem 51:6773–6781
25. Casini A, Edafe F, Erlandsson M, Gonsalvi L, Ciancetta A, Re N, Ienco A, Messori L, Peruzzini M, Dyson PJ (2010) Dalton Trans 39:5556–5563
26. Senn H, Thiel W (2009) Angew Chem Int Ed 48:1198–1229
27. Warshel A, Levitt M (1976) J Mol Biol 103:227–249
28. Alex A, Finn P (1997) J Mol Struct Theochem 398:551–554
29. Beierlein F, Lanig H, Schürer G, Horn AHC, Clark T (2003) Mol Phys 15:2469–2480
30. Raha K, Peters MB, Wang B, Yu N, Wollacott AM, Westerhoff LM, Merz KM (2007) Drug Discov Today 12:725–731
31. Menikarachchi LC, Gascón JA (2010) Curr Top Med Chem 10:46–54
32. Söderhjelm P, Kongsted J, Genheden S, Ryde U (2010) Interdiscip Sci Comput Life Sci 2:21–37
33. Gräter F, Schwarzl SM, Dejaegere A, Fischer S, Smith JC (2005) J Phys Chem B 109:10474–10483
34. Wang M, Wong CF (2007) J Chem Phys 126:026101 3 pages
35. Kaukonen M, Söderhjelm P, Heimdal J, Ryde U (2008) J Phys Chem B 112:12537–12548
36. Retegan M, Milet A, Jamet H (2009) J Chem Inf Model 49:963–971
37. Shi J, Lu Z, Zhang Q, Wang M, Wong CF, Liu J (2010) J Theor Comput Chem 9:543–559
38. Khandelwal A, Lukacova V, Comez D, Kroll DM, Raha S, Balaz S (2005) J Med Chem 48:5437–5447
39. Khandelwal A, Balaz S (2007) Proteins 69:326–339
40. Tao JM, Perdew JP, Staroverov VN, Scuseria GE (2003) Phys Rev Lett 91:146401 4 pages
41. Treutler O, Ahlrichs R (1995) J Chem Phys 102:346–354
42. Schäfer A, Horn H, Ahlrichs R (1992) J Chem Phys 97:2571–2577
43. Weigend F, Ahlrichs R (2005) Phys Chem Chem Phys 7: 3297–3305
44. Eichkorn K, Treutler O, Öhm H, Häser M, Ahlrichs R (1995) Chem Phys Lett 242:652–660
45. Eichkorn K, Weigend F, Treutler O, Ahlrichs R (1997) Theor Chem Acc 97:119–124
46. Klamt A, Schüürmann J (1993) J Chem Soc Perkin Trans 2:799–805
47. Schäfer A, Klamt A, Sattel D, Lohrenz JCW, Eckert F (2000) Phys Chem Chem Phys 2:2187–2193
48. Klamt A, Jonas V, Burger T, Lohrenz JCW (1998) J Phys Chem A 102:5074–5085
49. Jensen F (1999) Introduction to computational chemistry. Wiley, Chichester
50. Becke AD (1993) J Chem Phys 98:1372–1377
51. Hertwig RH, Koch W (1997) Chem Phys Lett 268:345–351
52. Huber CP, Campbell RL, Hasnain S, Hirama T to be published PDB file 2IPP
53. Case DA, Darden TA, Cheatham TE III, Simmerling CL, Wang J, Duke RE, Luo R, Walker RC, Zhang W, Merz KM, Roberts BP, Wang B, Hayik S, Roitberg A, Seabra G, Kolossváry I, Wong KF, Paesani F, Vanicek J, Wu X, Brozell SR, Steinbrecher T, Gohlke H, Yang L, Tan C, Mongan J, Hornak V, Cui G, Mathews DH, Seetin MG, Sagui C, Babin V, Kollman PA (2008) AMBER 10. University of California, San Francisco
54. Li H, Robertson AD, Jensen JH (2005) Proteins 61:704–721
55. Jorgensen WL, Chandrasekhar J, Madura J, Impey RW, Klein ML (1983) J Chem Phys 79:926–935
56. Ryde U (1996) J Comput Aided Mol Des 10:153–164
57. Ryde U, Olsson MHM (2001) Intern J Quant Chem 81:335–347
58. Besler BH, Merz KM Jr, Kollman PA (1990) J Comp Chem 11:431–439
59. Reuter NI, Dejaegere A, Maigret B, Karplus M (2000) J Phys Chem A 104:1720–1735
60. Hu L, Söderhjelm P, Ryde U (2011) J Chem Theory Comput 7:761–777
61. Svensson M, Humbel S, Froese RDJ, Matsubara T, Sieber S, Morokuma K (1996) J Phys Chem 100:19357–19363
62. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM Jr, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995) J Am Chem Soc 117:5179–5197
63. Wang J, Cieplak P, Kollman PA (2000) J Comput Chem 21:1049–1074
64. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C (2006) Proteins: Struct Funct Bioinform 65:712–725
65. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) J Comput Chem 25:1157–1174
66. Bartolotti LJ, Pedersen LG, Charifson PS (1991) J Comput Chem 12:1125–1128
67. Hay PJ, Wadt WR (1985) J Chem Phys 82:270–283
68. Sigfridsson E, Ryde U (1998) J Comp Chem 19:377–395
69. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA Jr, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Bakken V, Adamo C, Jaramillo J,

Gomperts R, Stratmann RE, Yazyev O, Austin A J, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA (2004) Gaussian 03, Revision C.02, Gaussian, Wallingford

70. Bayly CI, Cieplak P, Cornell WD, Kollman PA (1993) J Phys Chem 97:10269–10280
71. Ryckaert JP, Ciccotti G, Berendsen HJC (1977) J Comput Phys 23:327–341
72. Darden T, York D, Pedersen L (1993) J Chem Phys 98: 10089–10092
73. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) J Chem Phys 103:8577–8593
74. Berendsen HJC, Postma JPM, van Gunsteren WF, Di Nola A, Haak JR (1984) J Chem Phys 81:3684–3690
75. Kollman PA, Massova I, Reyes C, Kuhn B, Huo S, Chong L, Lee M, Lee T, Duan Y, Wang W, Donini O, Cieplak P, Srinivasan J, Case DA, Cheatman TE (2000) Acc Chem Res 33:889–897
76. Onufriev A, Bashford D, Case DA (2004) Proteins 55:383–394
77. Connolly ML (1983) J Appl Cryst 16:548–558
78. Kuhn B, Kollman PA (2000) J Med Chem 43:3786–3791
79. Scolaro C, Hartinger CG, Allerdyce CS, Keppler BK, Dyson PJ (2008) J Inorg Biochem 102:1743–1748