Complexes of a Zn-Metalloenzyme Binding Site with Hydroxamate-Containing Ligands. A Case for Detailed Benchmarkings of Polarizable Molecular Mechanics/Dynamics Potentials When the Experimental Binding Structure is Unknown

Nohad Gresh,*[a,b] David Perahia,[c] Benoit de Courcy,[a,b] Johanna Foret,[d] Céline Roux,[d] Lea El-Khoury,[a,e] Jean-Philip Piquemal,[a,f] and Laurent Salmon[d]

Zn-metalloproteins are a major class of targets for drug design. They constitute a demanding testing ground for polarizable molecular mechanics/dynamics aimed at extending the realm of quantum chemistry (QC) to very long-duration molecular dynamics (MD). The reliability of such procedures needs to be demonstrated upon comparing the relative stabilities of competing candidate complexes of inhibitors with the recognition site stabilized in the course of MD. This could be necessary when no information is available regarding the experimental structure of the inhibitor–protein complex. Thus, this study bears on the phosphomannose isomerase (PMI) enzyme, considered as a potential therapeutic target for the treatment of several bacterial and parasitic diseases. We consider its complexes with 5-phospho-D-arabinonohydroxamate and three analog ligands differing by the number and location of their hydroxyl groups. We evaluate the energy accuracy expectable from a polarizable molecular mechanics procedure, SIBFA. This is done by comparisons with *ab initio* quantum-chemistry (QC) calculations in the following cases: (a) the complexes of the four ligands in three distinct structures extracted from the entire PMI-ligand energy-minimized structures, and totaling up to 264 atoms; (b) the solvation energies of several energy-minimized complexes of each ligand with a shell of 64 water molecules; (c) the conformational energy differences of each ligand in different conformations characterized in the course of energy-minimizations; and (d) the continuum solvation energies of the ligands in different conformations. The agreements with the QC results appear convincing. On these bases, we discuss the prospects of applying the procedure to ligand-macromolecule recognition problems. © 2016 Wiley Periodicals, Inc.

DOI: 10.1002/jcc.24503

Introduction

Polarizable molecular mechanics/dynamics potentials (PMM) could constitute an extremely promising extension of quantum chemistry (QC) if their accuracy were convincingly demonstrated in a diversity of test situations. This is a particularly strong requirement with a major class of targets for drug design, namely Zn-metalloproteins,[1] on account of the importance of many-body polarization and charge-transfer effects which take place in, and around, the metal-binding site. One of our Laboratories has recently achieved very promising enrichments of a variant of the "Tinker" software,[2] denoted as Tinker-HP,[3] which parallelizes massively on several thousands of processors, and handles the polarization contribution stably and efficiently.[4] Tinker-HP enables MD simulations on complexes of several hundreds of thousands of atoms, paving the way to very long duration MD on drug–metalloprotein complexes. This makes it all the more critical to evaluate beforehand the accuracy expectable from PMM when confronted to QC.

Accuracy issues have been addressed in previous publications resorting to the SIBFA (Sum of Interactions Between Fragment Ab initio computed) procedure,[5] notably concerning Zn(II) oligoligated complexes[6–8] and Zn-metalloproteins.[9–12] As a

[a] *N. Gresh, B. de Courcy, L. El-Khoury, J.-P. Piquemal*
Laboratoire de Chimie Théorique, Sorbonne Universités, UPMC, UMR 7616 CNRS, Paris, France
Email: gresh@lct.jussieu.fr
[b] *N. Gresh, B. de Courcy*
Chemistry and Biology, Nucleo(s)tides and Immunology for Therapy (CBNIT), UMR 8601 CNRS, UFR Biomédicale, Paris, France
[c] *D. Perahia*
Laboratoire de Biologie et Pharmacologie Appliquées (LBPA), UMR 8113, Ecole Normale Supérieure, Cachan, France
[d] *J. Foret, C. Roux, L. Salmon*
Laboratoire de Chimie Bioorganique et Bioinorganique, Institut de Chimie Moléculaire et des Matériaux d'Orsay (ICMMO), Univ Paris-Saclay, Univ Paris-Sud, UMR 8182 CNRS, rue du Doyen Georges Poitou, Orsay F-91405, France
[e] *L. El-Khoury*
Centre d'Analyses et de Recherche, UR EGFEM, LSIM, Faculté de Sciences, Saint Joseph University of Beirut, BP 11-514, Riad El Solh, Beirut 1116-2050, Lebanon
[f] *J.-P. Piquemal*
Department of Biomolecular Engineering, The University of Texas at Austin, Texas 78712

Contract grant sponsor: National de Calcul Intensif (GENCI); Contract grant sponsor: Institut du Développement et des Ressources en Informatique Scientifique (IDRIS); Contract grant sponsor: Centre Informatique de l'Enseignement Supérieur (CINES, France, project No. x2009-075009); Contract grant sponsor: the Centre de Ressources Informatiques de Haute Normandie (CRIHAN, Rouen, France, project 1998053)

continuation of previous work,[10,11] we consider ligand binding to the type I Zn-dependent phosphomannose isomerase (PMI) enzyme, which catalyzes the reversible interconversion of D-mannose 6-phosphate into D-fructose 6-phosphate.[13] This reaction is the first step in the metabolism of mannose resulting in the formation of GDP-D-mannose, an important precursor to many mannosylated structures, such as glycoproteins, nucleotides, glycolipids, bacterial exopolysaccharides, or essential components for the integrity of the cell wall of fungi.[14] PMI has proven essential to the survival or pathogenicity of microorganisms such as bacteria (*Mycobacterium smegmatis*,[15] *Pseudomonas aeruginosa*[16]) protozoan parasites (*Leishmania mexicana*[17]) and yeast (*Saccharomyces cerevisiae*,[18] *Cryptococcus neoformans*,[19] *Aspergillus nidulans*,[20] and *Candida albicans*[21]). It is thus a target in several microbial and parasitic diseases, but in spite of these early demonstrations, there are yet no PMI inhibitors of clinical interest.[15–17] Toward the design of such inhibitors, 5-phospho-D-arabinonohydroxamate (5PAH) was designed and synthesized.[22,23] This molecule embodies a Zn-ligating hydroxamate moiety at one end, and a dianionic phosphate at the other, separated by three —-CHOH groups. 5PAH has a submicromolar inhibition potency of PMI.[24]

Hydroxyl groups constitute important modulators of binding affinities, with their dual H-bond donor and acceptor properties. They are also important tools in drug design studies, enabling to evaluate the impact on the binding affinities of hydroxylation/dehydroxylation at targeted sites of the ligand. Along these lines, three dehydroxylated 5PAH analogs have been synthesized (detailed in Supplementary material S1) and their inhibitory properties have been tested in one of our Laboratories on the type I (zinc) PMI from *Escherichia coli* using the reported procedure.[25] Referring to 5PAH as compound *a*, the first analog, *b*, 3-deoxy-5PAH, is dehydroxylated at the second C atom proximal to the hydroxamate function. The second, *c*, 2-deoxy-5PAH, is dehydroxylated at the first C atom proximal to it. The third analog, *d*, 5-phospho-pentanohydroxamate, is totally dehydroxylated. Compounds *a–d* are represented in Figure 1 under the hydroxamate form. The experimental results are reported in Table 1. Evaluation of their ligand potencies gave rise to an unexpected finding, in that, whereas removal of just one hydroxyl group, as in *b* and *c*, resulted into a thousand-fold increase in the IC_{50} values (and corresponding decrease in the binding affinities) with respect to *a*, removal of all three hydroxyls gave rise to the same one thousand-fold increase in IC_{50} values. The synthetic protocol and complete characterization of the four compounds are given as Supplementary material S1. The IC_{50} values were done in triplicate for each compound. The pK_a of sugar hydroxamic acid is 8.8–9.6,[26,27] so that in solution at pH 8, in which the experimental tests were done this group should be predominantly (to an amount of 90%) protonated. Accordingly, the solvation of ligands *a–d* in the 64-water bath will consider both forms separately. On the other hand, the close vicinity of a Zn(II), as in the PMI binding site, should lower the pK_a of hydroxamic acid, so that the predominant form should be deprotonated. We will thus consider this sole state in the PMI complexes. The second pK_a phosphate group is 5.7–6.8 in solution,[11,28] so that it should be predominantly doubly deprotonated (i.e., >90%) at pH 8. In the PMI complex, this group is bound to two protonated residues, Arg304 and Lys310. It can thus be assumed that similar to Zn(II), these should favor the dianionic form of the phosphate giving rise to a locally neutral complex.

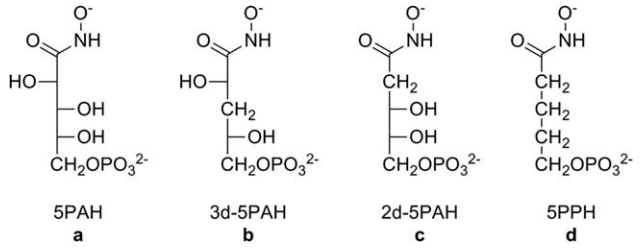

Figure 1. PMI ligands evaluated: (a) 5PAH: 5-phospho-D-arabinonohydroxamic acid, (b) 3d-5PAH: 3-deoxy-5-phospho-D-arabinonohydroxamic acid, (c) 2d-5PAH: 2-deoxy-5-phospho-D-arabinonohydroxamic acid, and (d) 5PPH: 5-phospho-pentanohydroxamic acid.

Accounting for such results should clearly require very long-duration molecular dynamics (MD) simulations and a dependable evaluation of the desolvation energies and entropy losses of the ligands prior to PMI complexation. There is an acknowledged procedure to derive protein–ligand free energies, namely free energy perturbation (FEP)[29] in which a parent compound is progressively mutated into a closely related analog, and the energies are integrated over a long-time trajectory for discrete variations of the mutation parameter λ, from zero to one. FEP was recently applied, in conjunction with an advanced polarizable potential, AMOEBA, to investigate the binding of pyrimidine dicarbonyl inhibitors to a Zn-dependent metalloproteinase, MMP-13.[29] The most secure use of FEP is when the structure of the complex of the "parent" ligand with its target is known experimentally. This is not the case for the complexes of PMIs with any known inhibitor. This might be remedied if accurate PMM potentials enable to predict the most favorable and likely binding structures, selected as the lowest-energy poses from long-duration MD. As appears from the energy analyses reported in this study, a pitfall in the use of lesser-accuracy force-fields could be the preferential stabilization of "wrong" binding pose(s) because of imbalances of the energy contributions making up the inter- or the intramolecular interaction energies. An objective we seek is to use FEP in the context of one advanced PMM procedure, SIBFA, and how much it could be dependable in view of such MD simulations. Thus, the prime purpose of this study to evaluate

Table 1. IC_{50} and K_i values determined for *E. coli* PMI inhibition by compounds *a–d*. K_i values given in italic were estimated from corresponding IC_{50} values using equation $IC_{50} = K_i([S]/K_M + 1)$ for competitive inhibition.

Ligand	IC50 (μM)	K_i (μM)	K_M/K_i
a	0.365 ± 0.004	0.084[c]	8450
b	1900 ± 80	1210 ± 50	0.6
c	440 ± 10	280 ± 6	2.5
d	1900 ± 100	1210 ± 60	0.6

[a] Conditions: Hepes buffer, pH 7.1, 25°C. [b] K_M F6P = 710 ± 50 μM. [c] PhD thesis, Céline Roux, 2005, Orsay (France).

its present accuracy (and possible present shortcomings) by benchmarking it against QC results in four distinct cases:

a. the intermolecular interaction energies of ligands *a–d* in models for the PMI recognition site, extracted from simulations on their complexes with the entirety of the PMI model. These will be computed both with and without the presence of polarized structural water molecules. These complexes total up to 264 atoms;
b. the intermolecular interaction energies of each ligand in a shell of 64 water molecules in a diversity of structures at the outcome of MD and energy-minimization;
c. the conformation energy differences between several ligand conformations, characterized in the course of energy-minimizations of their complexes with the 64 discrete waters;
d. the conformation-dependent continuum solvation energy of the four ligands.

While (a) is clearly the most important and difficult issue, reliable calculations of (b)–(d) could also be critical, since errors in these could distort the energy balances which include the desolvation energy of the ligands and their conformational energy rearrangements.

Methods

QC computations

As in our preceding studies, the QC calculations were performed with the CEP 4-31G(2d) basis set.[30,31] The computations on the 12 model complexes were also redone with the cc-pVTZ(-f) basis set,[32,33] yielding closely parallel trends of ΔE values (see below). Effective core potentials were used on Zn(II).[34] The calculations done at the Hartree–Fock (HF) level were complemented by calculations at the DFT-D levels using the B3LYP-D3[35–37], and B97-D3 functionals.[38] As for the HF calculations, both CEP 4-31G(2d) and cc-pVTZ(-f) basis sets were used. The choice of the B97-D3 was motivated by recent results on a series of monoligated complexes of metal cations which showed this functional to enable closer reproduction of CCSD(T) and MP2 results than B3LYP-D3 [unpublished]. The continuum solvation energies were computed with the Polarized Continuum Model (PCM).[39] These computations used the GAMESS[40] and G09[41] softwares. The DFT-D calculations were done with both the CEP 4-31G(2d) and cc-pVTZ(-f) basis sets. The latter basis set, while lending itself to computations of complexes of 250 atoms and more, can be considered to have a sufficient accuracy for the present benchmarking purposes. Computations on smaller (30 atoms) polyligated Zn(II) complexes with such a basis gave very close ΔE values to those with the aug-cc-pVTZ basis [unpublished].

SIBFA computations

In the SIBFA procedure,[5,42] the intermolecular interaction energy is computed as the sum of five contributions: electrostatic multipolar (E_{MTP*}), short-range repulsion (E_{rep}), polarization (E_{pol}), charge transfer (E_{ct}), and dispersion (E_{disp})

$$\Delta E_{TOT} = E_{MTP} + E_{rep} + E_{pol} + E_{ct} + E_{disp}$$

E_{MTP} is computed with distributed multipoles (up to quadrupoles) derived from the QC molecular orbitals precomputed for each individual molecule. They are derived from a procedure due to Vigné-Maeder and Claverie[43] and distributed on the atoms and bond barycenters. It is augmented with a penetration term.[44] The anisotropic polarizabilities are distributed on the centroids of the localized orbitals (heteroatom lone pairs and bond barycenters) using a procedure due to Garmer and Stevens.[45] The ligands were assembled from their constitutive fragments: anionic hydroxamate, methanol, methane, and dianionic phosphate $HOPO_3^{2-}$. Phosphate was split into two pseudo-fragments, HPO_3 and HOH, with two fictitious H atoms with null multipoles along the junction OP bond to enable for rotations around the PO bond. Similarly methanol was split into two pseudo-fragments, CH_4 and HOH with H atoms having null multipoles along the connecting CO bond. This enables for rotations around the CO junction bonds. Such a representation was found preferable to that resorting to a representation of methanol with actual water and methane fragments. It is important to recall that the fragmentation of a flexible ligand is necessary to ensure the independence of the multipole and polarizability intensities from any particular conformation, thus avoiding the issue of multipole transferability[46] while enabling to account for intramolecular, interfragment polarization and charge-transfer, which due to nonadditivity must be computed consistently and simultaneously with the intermolecular ones.[47] The distributed multipoles and polarizabilities were derived from CEP 4-31G(2d) calculations. Additional validation studies were also resumed using multipoles and polarizabilities derived from augcc-pVTZ(-f) calculations and a novel calibration of SIBFA on the basis of augcc-pVTZ(-f) energy decomposition analyses [El-Khoury et al., manuscript in preparation]. The distributed multipoles were in that case derived using the Generalized Multipole Analysis (GDMA) method by Stone.[48]

E_{rep} and E_{ct}, the two short-range repulsions, are computed using representations of the molecular orbitals localized on the chemical bonds and on localized lone-pairs. E_{disp} is computed as an expansion into $1/R^{[6]}$, $1/R^{[8]}$, and $1/R^{[10]}$, and also embodies an explicit exchange–dispersion term.[49] Energy-minimization used the Merlin package.[50] We use an augmented representation, embodying a lone pair–lone pair exchange–dispersion component following the formulation of Ref. [51] and the corresponding recalibration of E_{disp} on the basis of SAPT results on the water–water dimer. The Zn(II) parameters were calibrated on the basis of MP2, B97-D3, and B3LYP-D3 calculations on monoligated Zn(II) complexes with representative O-, N-, and S-ligands [manuscript in preparation]. Conformation energy changes of the ligand and the protein sidechains are done in the context of the rigid rotor: standard experimental bond lengths and valence angles are adopted and not relaxed. The continuum solvation energies, ΔG_{solv},

were computed with the Langlet–Claverie procedure[52] in which the electrostatic potential generated by the solute on its van der Waals surface is computed with the same distributed multipoles as for SIBFA. ΔG_{solv} is a sum of an electrostatic, polarization, repulsion, dispersion, and cavitation terms.

As in Refs. 10,11, we used the X-ray structure of type I PMI from *Candida albicans* (PDB ID 1PMI).[53] The starting positions for the structural waters were obtained with the "discrete" algorithm,[54] which locates waters around the accessible hydrophilic sites of the solute by energy-minimization with a simplified interaction energy term. They were relaxed by SIBFA EM prior to MD. We have retained 28 "structural" waters. Our previous study bearing on the complexes of PMI with D-mannose 6-phosphate surrogates[11] was limited to 11 waters. The present increased number enables to more completely connect together the Zn-binding site and the entrance of the cavity by saturated and polarizable networks which could embody a second layer of waters, as shown below. Solvation of the PMI-ligand complexes is completed by the L-C continuum procedure. As in Refs. 10,11, based on a 15 Å proximity to Zn(II), a 164-residue of Zn(II) model was constructed retaining the following residues: Asn15-Gly19; Tyr46-His54; Phe97-Lys153; Val205-Asn211; Gly258-Met333; Asp352-Leu361. The PMI backbone was fixed, and EM was done on the torsional angles of the side-chains of the following residues: Asp17, Trp18, Glu48, Lys100, Ser109, Gln111, His113, Lys136, Glu138, His285, Tyr287, Glu294, Asp300, Arg304, and Lys310. All torsional angles of the ligand were simultaneously relaxed, along with the six intermolecular variables defining its approach to PMI and the corresponding six intermolecular variables of each water.

The short-duration SIBFA MD simulations (10–20 ps) were performed with the velocity-Verlet algorithm[55] and temperatures in the range 0–300 K, controlled by the Berendsen algorithm.[56] The waters were constrained to remain bound to the PMI-5PAH complex by a spherical restraint potential centered around Zn(II). Their internal geometries were constrained by the "SHAKE" algorithm.[57] The protein backbone was not relaxed. The PMI side-chains that were relaxed by MD were the same as previously reported.[10] To sample configurations by these MD simulations, the "simplified" potential resorted to scalar polarizabilities without iterations on the induced dipoles and without the charge-transfer contribution. As commented in our recent study bearing on the structural waters around the bimetallic site of superoxide dismutase (SOD),[12] the duration of such MD is very short and could not enable for a realistic simulation of 5PAH relaxation in the binding site of PMI concomitant with the rearrangements of the water layers. It appeared nevertheless sufficient to enable for reorientations of the hydroxamate group of 5PAH in the Zn-binding site, itself leading to concomitant rearrangements of the hydroxyl groups and of the structured waters. These are of a sufficient amplitude to enable addressing the principal objective of this work, namely how dependable could PMM be relative to QC in a series of competing complexes which have significant structural differences in terms of ligand binding, Zn(II) coordination, and water networks?

This study uses as a starting point the model structure of the PMI-5PAH complex, previously energy-minimized using SIBFA and the Langlet–Claverie (L-C) continuum solvation potential.[52] Competing complexes are unraveled by short-duration MD simulations with a simplified SIBFA potential in the presence of a few "discrete" water molecules. Three selected poses are then energy-minimized with the full SIBFA potential and continuum solvation. The resulting structures are in turn used as starting points for energy-minimization (EM) of the PMI complexes with the three other dehydroxylated ligands. A total of 12 complexes is thus considered.

Procedure followed

Consistent with our previous studies of ligand–protein interactions,[9–12,58,59] to address point (a), models of the recognition sites are built upon extracting from each of the 12 energy-minimized complexes the ligand, the Zn(II) cation, the side-chains of the amino-acids involved in Zn- and ligand-binding as well as residues in their neighborhood, and the discrete water molecules. Single-point calculations of the intermolecular interactions are computed by parallel SIBFA and QC.

To address point (b), each ligand is extracted from each of its three PMI complexes, and subjected to MD followed by energy-minimization in a shell of 64 discrete water molecules and in the presence of continuum solvation. The choice of a large number of discrete water molecules is motivated by the extremely strong potentials and fields exerted by the trianionic charge of the ligand on the solvent, structuring up to a second- and possibly a third shell of solvation. Parallel single-point QC and SIBFA computations were next performed on the energy-minimized structures. These are necessary tests, considering the very strong second-order and non-additivity effects coming into play in anion-triggered structures.[60] They would enable to evaluate the accuracy of SIBFA compared with QC upon computing ligand desolvation energy with structured waters. This point needs to be addressed prior to performing the energy balances. The solvated ligands are considered in two possible hydroxamate ionization states, deprotonated and protonated on the nitro oxygen. To address point (c), we compare the SIBFA and QC intramolecular (conformational) energies of the four ligands extracted from their water complexes. To address point (d), we compare the results of the Continuum solvation energy ΔG_{solv} with the Langlet–Claverie procedure to those from the (PCM procedure).[39,61]

Results and Discussion

Inter-molecular interaction energies in the PMI recognition site

The three PMI-5PAH complexes are represented in Figures 2a–2c. We will denote by A, B, and C the three modes of binding found by MD followed by EM which resorted to the full SIBFA potential. We denote by E_1 the summed two first-order contributions E_{MTP} and E_{rep}, and by E_2 the summed two second-order contributions E_{pol} and E_{ct}.

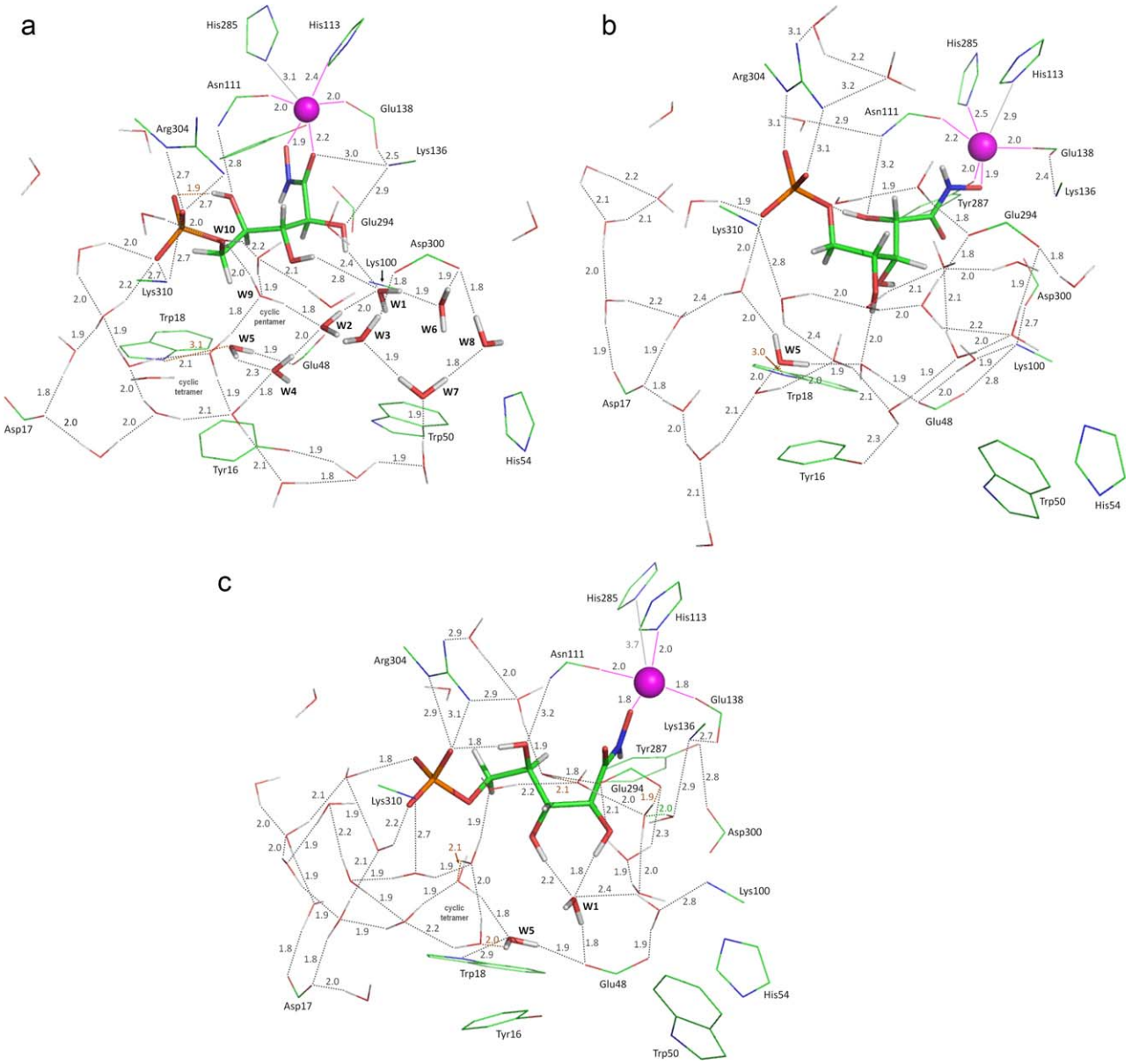

Figure 2. (a–c) Representation of the complexes of 5PAH with the PMI recognition site and 28 water molecules in binding modes A–C, respectively. [Color figure can be viewed at wileyonlinelibrary.com]

In complex A (Fig. 2a), the hydroxamate moiety binds Zn(II) in a bidentate fashion. This occurs at the detriment of Zn-His285 binding, the Zn-Nε distance having increased to 3.1 Å. The hydroxyl proximal to hydroxamate accepts a proton from the side-chain of Lys136. Lys136 was previously also found to donate a proton to a hydroxyl group of the β-D-mannopyranose 6-phosphate substrate.[11,62] This hydroxyl also donates its hydrogen to a water molecule, W1, involved in several interactions: as a proton donor to an Asp300-bound water, W6; as a proton acceptor from both the second 5PAH hydroxyl and a water molecule, W2. W2 is at the base of two networks connecting it to the N-terminal end residues, Trp18 and Asp17 on the one hand, and to Arg304 on the other hand. A more detailed description of the networks is given as Supplementary material S2. The other 5PAH hydroxyl to bind directly to 5PAH is the third one, which accepts a proton from the amine group of Zn-bound Asn111.

In complex B (Fig. 2b), an accented rotation of 5PAH has taken place. The hydroxamate now binds Zn(II) in a monodentate fashion through its nitroxy O (1.9 Å). Its CO bond is now on the opposite direction from Zn(II) enabling Tyr287 to enter the coordination shell (2.0 Å). Interestingly, His285 reenters the coordination shell (2.5 Å) at the detriment of His113 (2.9 Å). The first two hydroxyls now point toward opposite directions, and it is the first one, not the third, that accepts a proton from the Asn111 side-chain. It also donates a proton intramolecularly to the phosphate ester O forming a chair-like six-bond cycle. An extensive water network connects Asp300, Lys100, Glu294, and Glu48 on the one hand, to both Trp18 and Asp17 and to Lys310 and the 5PAH phosphate. Again

some waters, such as Trp18-bound W5, connect together these two networks. The third 5PAH hydroxyl partakes to the network by donating a proton to a water belonging to the network that donates in turn a proton to Glu48.

In complex C (Fig. 2c), the hydroxamate underwent a rotation around the C-C bond connecting it to the hydroxymethyl chain. As in complex B, it binds Zn(II) in a monodentate fashion. The hydroxymethyl chain and phosphate group have also swung. The first two hydroxyls are directed away from Lys136 and the ester phosphate O is directed away from Arg304. On the other hand, the third hydroxyl is H-bonded to the side-chain of Asn111 as in complex A. The first two hydroxyls are solely bound to PMI through the discrete waters. They both donate their proton to a central water, W1. On the one hand, W1 also accepts a proton from a water molecule which in turn accepts a proton from Lys100 and donates one to Glu48. W1 donates a proton to the second anionic O atom of Glu48. Exactly as with complex A, Glu48 accepts a proton from a water molecule, W5, which in turns accepts a proton from Trp18. W5 is additionally H-bonded to two other waters which connect it, on the one hand, to Asp17 at the N-terminal end, and to Lys310 at the entrance of the binding site. A more detailed description of the water networks is given in Supplementary material S2. Clearly, interchanges of the discrete and bulk waters should be anticipated to occur. Nevertheless, the recurrence of two networks are worth noting. The first connects Asp17 and Asp300 at the N and C-terminal ends, and the second connects Lys100 and Asp300 to Arg304 or Lys310 at the entrance of the recognition site. At least one hydroxyl group of the ligand and the imino nitrogen of Trp18 act as relays in such networks. Such networks were previously observed upon studying the complex of a D-mannose malonate ligand with PMI.[11] We plan to monitor the lifetimes and evolutions of such networks using long MD simulations using the SIBFA potential in the context of Tinker-HP.

Only limited rearrangements were found for the three dehydroxylated analogs following EM on the A–C complexes of 5PAH as a starting point. They are illustrated in Supplementary material S3 regarding complexes A.

It is clear that the rearrangements, which could only have been obtained by MD, are substantial regarding Zn(II) coordination, involvements of the hydroxyls, ligand conformation, and structures of the water networks, even though the latter bear several similarities in complexes A–C. Thus a critical point in the perspective of long-duration MD simulations is to resort to QC as a benchmark and evaluate how reliably could PMM account for the differences in the intermolecular interaction energies in these three complexes, as well as upon dehydroxylation. We have thus compared the SIBFA and QC intermolecular interaction energies for the complexation energies of ligands *a–d* in binding modes A–C within the recognition site, totaling 12 complexes. The recognition site is made out of the end-side chains of the following residues: Tyr16, Asp17, Trp18, Glu48, Trp50, His54, Lys100, Ser109, Gln111, His113, Lys136, Glu138, His285, Tyr287, Glu294, Asp300, Arg304, and Lys310.

We focus first on the ΔE values obtained in the presence of the 28 water molecules, the complexes totaling up to 264

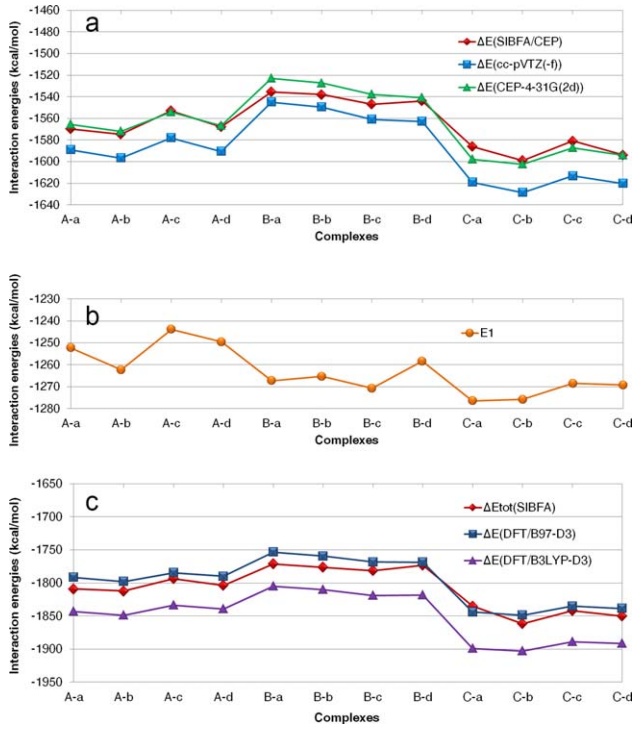

Figure 3. Compared evolutions of the SIBFA and QC intermolecular interaction energies in complexes A-*a* to C-*d*. a) ΔE(SIBFA) without dispersion and ΔE/HF(QC) with the CEP 4-31G(2d) and cc-pVTZ(-f) basis sets; b) corresponding evolution of E_1; c) ΔE_{tot}(SIBFA) including dispersion, ΔE(B97-D3), and ΔE(B3LYP-D3). [Color figure can be viewed at wileyonlinelibrary.com]

atoms. The evolutions of the ΔE values are reported in Figures 3a–3c. Complexes A–C are considered in succession. To gain space, the ΔE values and their energy contributions in SIBFA are reported in Supplementary material S4. Figure 3a compares the evolutions of ΔE(SIBFA) in the absence of the E_{disp} contribution to those of ΔE/HF(QC) with two basis sets: CEP 4-31G(2d) and cc-pVTZ(-f). Figure 3b reports the corresponding evolutions of E_1. Figure 3c compares the evolutions of ΔE_{tot}(SIBFA) including E_{disp} to those of ΔE(DFT) with the B97-D3 and B3LYP-D3 functionals.

At this point, a very commendable feature of the CEP 4-31G(2d) results is noteworthy. Thus at both HF and DFT-D3 levels, the ΔE curve with this basis runs virtually entirely parallel to the cc-pVTZ(-f) curves. At the HF level, to within 1 kcal/mole, there is a constant offset of 20 kcal/mol out of 1520–1600 with respect to the cc-pVTZ(-f) basis. It is in the range of 31–38 kcal/mol out of 1790–1850 at the B97-D3 level. Such parallel evolutions should justify the use of the CEP 4-31G(2d) basis to derive the initial SIBFA library of fragments. For the comparison of the 12 PMI complexes, a further step was taken in this study now resorting to the aug cc-pVTZ(-f) multipoles and polarizabilities. Which accuracy can be expected from the SIBFA computations at both levels? Figure 3a and Supplementary material S4 indicate that ΔE(SIBFA) can match closely ΔE/HF(QC). The most satisfactory agreements are for complexes A and C which are the most stable ones. With respect to the CEP 4-31G(2d) calculations, the numerical errors are $<$10 kcal/mol out of 1500, actually less than the corresponding

differences of ΔE/HF(QC) between this basis and the cc-pVTZ(-f) one. Even for complexes B, the maximum relative errors are $<$15 kcal/mol out of 1500, i.e. 1%. Such relative errors are actually smaller than the 3% ones noted in previous benchmarks of the procedure. The evolutions of ΔE are accounted for overall. This is illustrated in Figure 3a, comparing the ΔE(SIBFA) curve in red to the ΔE/HF(QC-CEP 4-31G (2d)) curve in green. The ΔE(HF) curve with the cc-pVTZ(-f) in blue is 20 kcal/mol lower than with the smaller basis set, but as mentioned above, the two QC curves exhibit a near-perfect parallelism. Such a parallelism could not be matched by the SIBFA curve, nevertheless this curve is virtually encompassed by the two QC curves. This situation is reminiscent of the one found in the complexes of mercaptocarboxylate ligands with the recognition site of metallo-β-lactamase.[9,58] The complexes then investigated were made out of 95 atoms, as contrasted to 264 ones in the present case, i.e. an app. three-fold increase in size. We note that while E_1, the sum of E_{MTP} and E_{rep}, accounts in these complexes for app. 80% of ΔE(SIBFA), its evolution cannot match those of ΔE(SIBFA) and *a fortiori* ΔE(QC) (Fig. 3b). It could also be readily observed that E_1 favors complexes B over complexes A, contrary to the results from ΔE(SIBFA) and ΔE(QC).

Following the inclusion of E_{disp}, the evolution of ΔE_{tot}(SIBFA) appears to match well that of ΔE(B97-D3) (Fig. 3c) with relative errors amounting to 1%. The most uniform agreements are for complexes A. However, a downgrading of the overall shape occurs due to complexes B-*d* and C-*a*, the latter being the only one for which ΔE_{tot}(SIBFA) has a smaller magnitude than ΔE(B97-D3). E_{disp} has comparable magnitudes to $\delta\Delta E_{corr/disp}$, but in a somewhat larger range (from $-$229 to $-$263 kcal/mol) than it (from $-$224 to $-$248 kcal/mol).

The role of the discrete waters on the relative stabilities of the complexes was evaluated upon recomputing ΔE_{nw}(SIBFA) and ΔE_{nw}(HF) following their removal, but retaining the same geometries as in their presence. The results are reported and commented in Supplementary material S5a for the 12 complexes. It was instructive to also report the stabilization energies, ΔE_{stab}, conferred by the discrete waters, namely, the difference between the ΔE and ΔE_{nw} values. The energy results are reported and commented in Supplementary material S5b.

Results with the aug-cc pVTZ(-f) multipoles and polarizabilities. A new library of SIBFA fragments is being constructed which resorts to distributed multipoles and polarizabilities derived from the aug-cc pVTZ(-f) basis set. Its calibration has been performed on the basis of Energy Decomposition Analyses (EDA) with this basis set, also resorting to an improved SIBFA representation of sp and sp2 lone-pairs, partially "smearing" them above and below the molecular plane [El-Khoury et al., manuscript in preparation]. For the purpose of the present comparisons, we resort here to HF multipoles and polarizabilities. A calibration resorting to *correlated* multipoles and polarizabilities is presently underway on the basis of correlated EDA procedures, as enabled by SAPT[63,64] and ALMO[65] methods and will be reported separately. Which agreements could now be reached, now regarding cc-pVTZ(-f) HF and DFT-D3 calculations? The results are reported in Table 2. The values of ΔE(SIBFA) differ by less than 10 kcal/mol out of 1500 from those computed with the CEP 4-31G(2d) calibration, yet enable an improved agreement with respect to the "target" ΔE_{HF}(cc-pVTZ(-f)) results. The magnitudes of E_{MTP} and especially E_{pol} are larger than with the CEP 4-31G(2d) basis set, but are to a large extent compensated by much larger values of E_{rep} (compare Tables 2 and Supplementary material S4). The errors with respect to the target HF results are more uniform between ligands for a given complex, A–C, than for any of the four ligands considered in the three complexes. The largest error is

Table 2. Complexes of ligands *a–d* with the recognition site of PMI and 28 discrete waters. Values (kcal/mol) of the intermolecular interaction energies and their contributions for SIBFA calculations with aug-cc pVTZ (-f) multipoles and polarizabilities and of the QC interaction energies in the 12 model complexes.

Complexes A	A-a	A-b	A-c	A-d
E_{MTP}	−1882.0	−1895.3	−1895.9	−1897.9
E_{rep}	675.7	679.4	694.0	687.1
E_1	−1206.2	−1215.9	−1201.9	−1210.8
E_{pol}	−274.7	−274.8	−262.4	−276.5
E_{ct}	−91.5	−91.0	−91.0	−91.5
ΔE	**−1572.6**	**−1581.8**	**−1555.3**	**−1578.8**
ΔE(cc-pVTZ(-f))	**−1589.2**	**−1596.7**	**−1578.1**	**−1590.7**
E_{disp}(SIBFA)	**−235.3**	**−233.5**	**−237.1**	**−231.1**
ΔE_{tot}(SIBFA)	**−1807.9**	**−1815.3**	**−1792.4**	**−1809.9**
$\delta\Delta E_{(corr/disp)}$	**−238.1**	**−236.5**	**−233.6**	**−233.5**
ΔE(DFT/B97-D3)	**−1823.8**	**−1829.8**	**−1807.5**	**−1821.2**
$\delta\Delta E_{(corr/disp)}$	**−296.8**	**−294.4**	**−298.7**	**−291.2**
ΔE(DFT/B3LYP-D3)	**−1882.5**	**−1887.7**	**−1872.6**	**−1878.6**
Complexes B	B-*a*	B-*b*	B-*c*	B-*d*
E_{MTP}	−1890.4	−1900.2	−1892.7	−1878.0
E_{rep}	647.9	659.0	644.1	641.8
E_1	−1242.5	−1241.2	−1248.6	−1236.2
E_{pol}	−211.1	−220.8	−218.9	−239.1
E_{ct}	−86.8	−86.4	−85.3	−86.1
ΔE	**−1540.4**	**−1548.3**	**−1552.9**	**−1561.3**
ΔE(cc-pVTZ(-f))	**−1545.1**	**−1549.6**	**−1561.0**	**−1562.9**
E_{disp}(SIBFA)	**−233.5**	**−234.3**	**−229.7**	**−227.3**
ΔE_{tot}(SIBFA)	**−1773.9**	**−1782.6**	**−1782.6**	**−1788.6**
$\delta\Delta E_{(corr/disp)}$	**−250.5**	**−250.2**	**−251.1**	**−246.7**
ΔE(DFT/B97-D3)	**−1791.6**	**−1796.0**	**−1807.5**	**−1806.7**
$\delta\Delta E_{(corr/disp)}$	**−303.5**	**−303.8**	**−302.5**	**−296.6**
ΔE(DFT/B3LYP-D3)	**−1844.6**	**−1849.6**	**−1858.9**	**−1856.6**
Complexes C	C-*a*	C-*b*	C-*c*	C-*d*
E_{MTP}	−1955.1	−2011.9	−1977.6	−1992.0
E_{rep}	717.0	783.5	756.8	770.3
E_1	−1238.1	−1228.4	−1220.8	−1221.7
E_{pol}	−267.3	−285.9	−268.1	−285.0
E_{ct}	−93.8	−98.9	−94.9	−97.2
ΔE	**−1599.2**	**−1613.1**	**−1583.7**	**−1604.0**
ΔE(cc-pVTZ(-f))	**−1619.3**	**−1628.7**	**−1613.3**	**−1620.3**
E_{disp}(SIBFA)	**−243.0**	**−254.8**	**−250.1**	**−248.7**
ΔE_{tot}(SIBFA)	**−1842.0**	**−1867.9**	**−1833.8**	**−1852.7**
$\delta\Delta E_{(corr/disp)}$	**−246.9**	**−263.9**	**−267.6**	**−261.1**
ΔE(DFT/B97-D3)	**−1864.9**	**−1887.1**	**−1874.4**	**−1876.2**
$\delta\Delta E_{(corr/disp)}$	**−304.6**	**−324.4**	**−326.2**	**−319.1**
ΔE(DFT/B3LYP-D3)	**−1922.6**	**−1947.6**	**−1933.0**	**−1934.2**

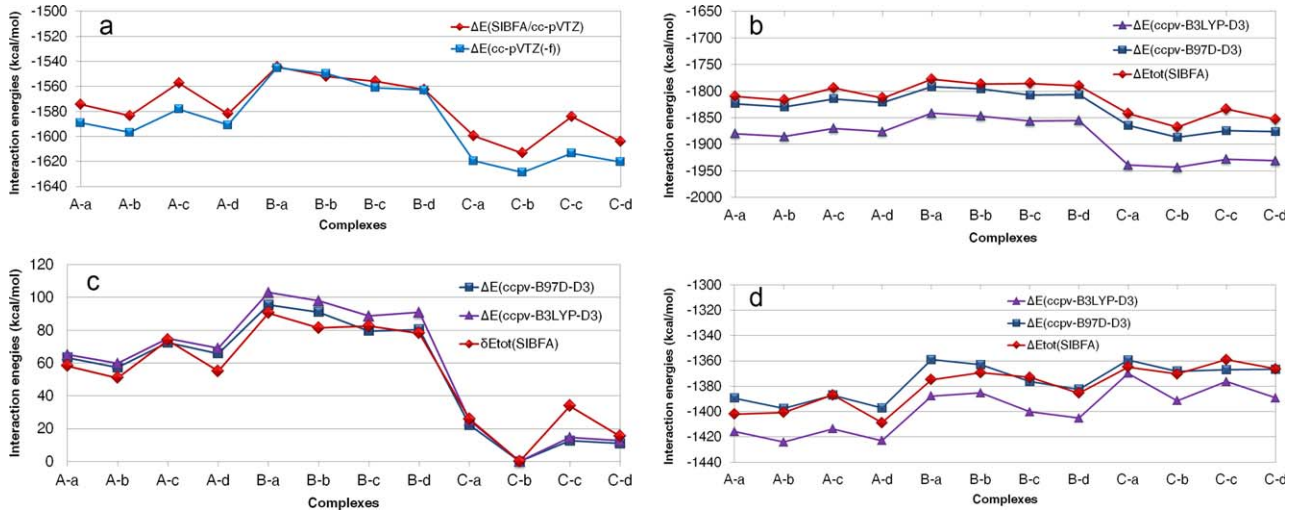

Figure 4. (a–d) Results from SIBFA computations with aug-cc-pVTZ(-f) multipoles and polarizabilities and QC calculations with the cc-pVTZ(-f) basis set. (a) Compared evolutions of ΔE(SIBFA) without dispersion and ΔE/HF(QC) in complexes A-*a* to C-*d*; (b) compared evolutions of ΔE_{tot} (SIBFA), ΔE(B97-D3), and ΔE(B3LYP-D3); (c) compared SIBFA, B97-D3, and B3LYP-D3 evolutions of the relative energy differences with respect to the most stably bound complex; (d) Compared evolutions of ΔE_{tot}(SIBFA) including dispersion, ΔE(B97-D3), and ΔE(B3LYP-D3) in complexes A-*a* to C-*d* n the absence of the discrete waters. [Color figure can be viewed at wileyonlinelibrary.com]

29 kcal/mol out of 1600, i.e. 2%, for complex C-*c*. The other errors are <25 kcal/mol. Accounting for BSSE effects could possibly reduce the energy difference, but this will not be addressed here. Figure 4a compares the evolutions of ΔE_{HF}(cc-pVTZ(-f)) and of ΔE(SIBFA). It shows a good parallelism for complexes A and B, with a smaller offset for complexes B. Parallelism is downgraded in complexes C, because of the larger offset occurring with ligand *c*. We have also compared ΔE_{tot}(SIBFA) with ΔE(B97-D3) and ΔE(B3LYP-D3). Toward a fully consistent comparison with DFT, a library of multipoles and polarizabilities derived at the MP2 and DFT levels is being constructed (Nasim-Khan et al., in preparation). The comparisons at the present stage should nevertheless be instructive. A much better agreement clearly obtains with ΔE(B97-D3) than with ΔE(B3LYP-D3). It can be noted that the B97-D3 and B3LYP-D3 values can differ between themselves by 50–60 kcal/mol along the series. There is no straightforward explanation for such differences since the ΔE values in a series of monoligated complexes with O and N ligands were closely similar with both DFT-D3 methods (unpublished results). Complex C-*c* again downgrades ΔE_{tot}(SIBFA), with an error of 47 kcal/mol out of 1800. With this exception, a good parallelism can be observed in the evolutions of the SIBFA and DFT-D3 curves (Fig. 4b) with no curve-crossing as with the CEP 4-31G(2d) basis set (Fig. 3c). Figure 4c gives the evolutions of the relative interaction energies, computed as the difference with respect to the best bound complex taken as energy zero, namely complex C-*b*. Again with the exception of C-*c*, the SIBFA curve matches satisfactorily the trends of the DFT-D3 curves. This is true regarding not only the B97-D3 curve, but also the B3LYP-D3 one, despite the significantly larger offset in the absolute ΔE values. Each of the 28 discrete water molecule is involved in an average of two-three H-bonds, which are mostly of a cooperative nature. If at the present stage of development even small remaining inaccuracies remained in ΔE_{tot}(SIBFA) in water–water or water–ligand interactions, these could be amplified in the global complexes. We wished to evaluate if this is the case and/or whether there are imbalances in ΔE computations for the different Zn(II) coordinations or phosphate binding modes in complexes A–C. We have thus recomputed the interaction energies in the absence of the waters. The evolutions are reported in Figure 4d. There now is a very close reproduction of ΔE(B97-D3) by ΔE_{tot}(SIBFA), the largest error, now occurring with B-*a*, being of 15 kcal/mol out of 1350, i.e. 1%. It is hoped that the forthcoming refinements to SIBFA should retain such agreements and enable improved ones following inclusion of large shells of discrete water molecules.

Intermolecular interaction energies in a shell of 64 water molecules

The desolvation energy of the ligands prior to complexation is a critical element in the overall energy balances. It requests exhaustive sampling of the potential energy surface and addressing entropy issues as well, but which accuracy could be expected from polarizable molecular mechanics regarding trianionic solutes? Anionic solutes can have an organizing effect on their first hydration shell which could extend to a second shell and possibly beyond. This was demonstrated by Pullman et al. as early as 1975 regarding the dimethylphosphate monoanion, a building block of the main-chain of nucleic acids and of phospholipids.[66] Such an organizing effect could be anticipated *a fortiori* for ligands *a–d*, which bear a net trianionic charge. Prior to assessing whether such an effect could persist over long MD time scales, we deemed it essential to evaluate the accuracy of the SIBFA potential in a diversity of structures. The ranking of solvation energies could result from a fluctuating interplay of first- and second-order contributions and of dispersion within ΔE_{tot}. For such an

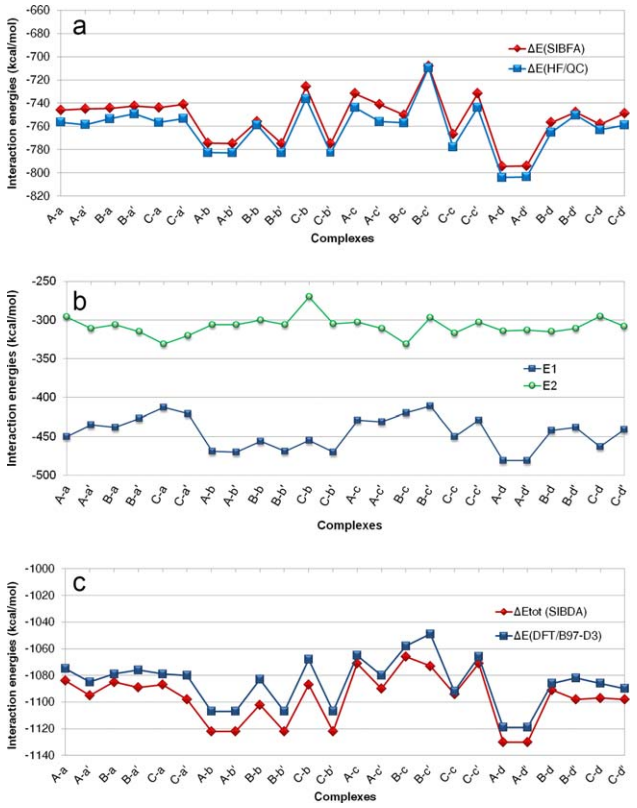

Figure 5. (a) Complexes of ligands *a*–*d* with a shell of 64 waters. Evolutions of (a) ΔE(SIBFA) without dispersion and ΔE(HF/QC), (b) E_1 and E_2, and (c) ΔE_{tot}(SIBFA) including dispersion and ΔE(DFT/B97-D3). [Color figure can be viewed at wileyonlinelibrary.com]

assessment, each ligand was extracted from each of its three PMI complexes and solvated in a shell of 64 water molecules by the "discrete" algorithm.[54] The complexes were subsequently energy-minimized with the SIBFA potential and the "Merlin" minimizer. They were then subjected to two separate molecular dynamics with a spherical shell potential. In the first, the conformation of the ligand was relaxed together with the 64 waters. In the second complexes, denoted by primes, only the waters were relaxed. In both simulations, the temperature was progressively increased from 30 to 300 K for 500,000 steps with a 0.75 fs time-step, i.e. 400 picoseconds, and then cooled to 30 K with the same number of steps. Larger steps caused instabilities in some complexes. Energy-minimization was then resumed relaxing the internal coordinates of both the ligand and the waters. Single-point QC computations were done using the structures derived at the outcome of EM. We have in addition considered an extended conformation of ligand *a* minimizing the electrostatic repulsions between phosphate and hydroxamate. It was similarly solvated with 64 waters by the "discrete" algorithm, then energy-minimized on the ligand and water internal coordinates. For comparative purposes, ligand *d*, devoid of hydroxyl groups, was solvated by the same procedure. We denote these two additional structures as "I," standing for "ice-like." There is thus a total of 26 complexes. The results are illustrated on Figures 5a–5c. Figure 5a compares ΔE(SIBFA) along with E_1 and E_2 (Fig. 5b) to ΔE(HF). Figure 5c compares ΔE_{tot}(SIBFA) with ΔE(DFT/B97-D3). Figure 5a shows ΔE(SIBFA) to match closely the evolutions of ΔE(HF), the relative errors being less than 1%. E_2 accounts for up to 45% of ΔE. Within E_2, the magnitude of E_{ct} is half that of E_{pol}. Again, as previously observed for the water networks in superoxide dismutase,[12] both E_1 and E_2 are necessary to confer to the evolution of ΔE a shape matching that of ΔE(HF). Several complexes have closely similar ΔE values yet differ significantly by the relative magnitudes of E_1 and E_2 which can even display antagonistic trends (Fig. 5b). Thus for ligand *a*, complexes A-*a*′ and C-*a* have ΔE(SIBFA) values that differ by 1.4 kcal/mol out of 740. Yet A-*a* is favored by E_1 by a 21.8 kcal/mol preference, and disfavored by E_2 by 20.4 kcal/mol. Figure 5c shows that at following the addition of E_{disp}, ΔE_{tot}(SIBFA) can still match ΔE(DFT/B97D3) with errors not exceeding 20 kcal/mol out of 1080, i.e. 2%, although a lesser parallelism between the two curves could be observed. There is one relative outlier, namely complex B-*c*′, which is overestimated by 25 out of 1070 kcal/mol, thus predicted to have a more favorable ΔE_{tot}(SIBFA) than, e.g., B-*c*, contrary to ΔE(DFT/B97D3). Yet this complex gave rise to much better agreements at the HF level.

Could "ce-like" structures be transiently stabilized around trianionic complexes? Complex I-*a* was energy-minimized into a cube-like arrangement of 40 water molecules solvating the phosphate, similar to the one previously identified for water clusters of 20 molecules.[67] While this is most likely a consequence of the protocol followed, it has a ΔE value which is the most favorable for the whole series of the solvated complexes of ligand *a*. This is borne out by the QC computations. By contrast, complex I only ranks third in terms of relative ΔE values for the dehydroxylated ligand *d*. At the DFT level, I-*a* is no longer the most favored and is ranked third in terms of ΔE_{tot}(SIBFA), while I-*d* has the least favorable ΔE values along the series of *d* complexes. Complex I-*a* and representative folded complexes A-*a* and B-*a* are represented in Figures 6a–6c, showing the "globular" nature of the water solvation shell in the latter two complexes. In complexes I, the phosphate and hydroxamate-solvating waters should be further enabled to connect by additional waters to attain a globular-like shape. Extensive MD sampling will then be planned to evaluate any possible persistence, or transient stabilization, of ice-like structures such as I within large water shells. Nevertheless the relative stabilization of I for the tri-hydroxylated ligand as compared with the non-hydroxylated one is noteworthy. It implies that the hydroxyl groups could relay, or anchor to, the cooperative network of ice-like waters. Due to their confinement in some protein binding sites, water molecules could undergo an increased ordering, such as in two-dimensional networks. The hydroxyl group proximal to the hydroxamate could act as a relay as well, and this could be observed in complex A-*a* of PMI.

We have recomputed the solvation energy of the ligands with hydroxamic acid replacing hydroxamate. For this purpose, the proton in the N—-OH bond was retained cis to the carbonyl oxygen. We performed energy-minimization on complexes A-*a* to C-*d*′ starting from the corresponding structures in the hydroxamate form. We did not redo MD, as our purpose is to

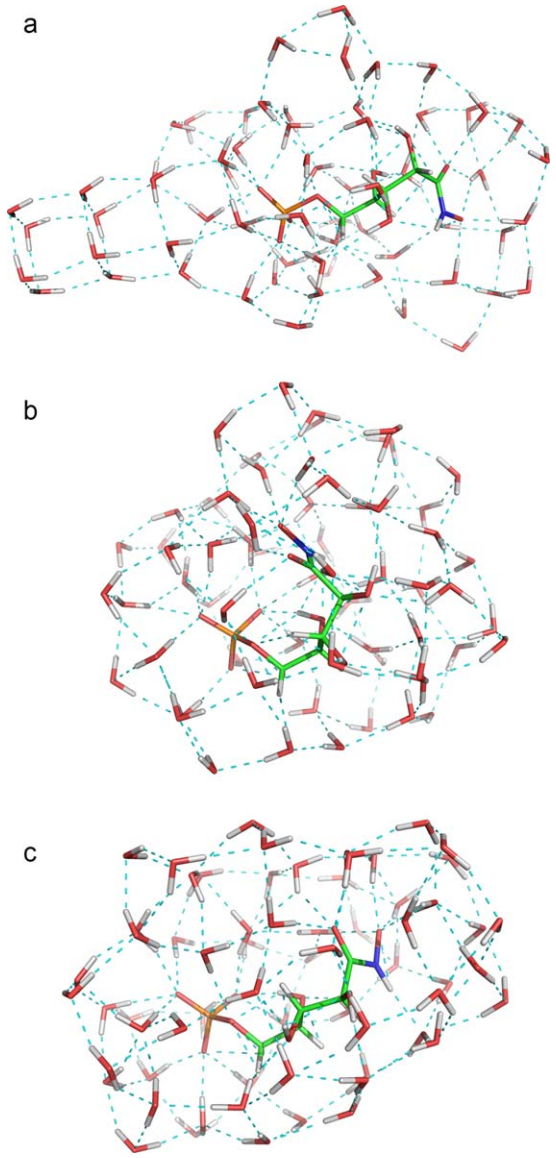

Figure 6. Representation of the complex of a with (a) an "ice-like" arrangement of 64 waters; (b) and (c) two globular arrangements of the 64 waters. [Color figure can be viewed at wileyonlinelibrary.com]

validate SIBFA in this alternative protonation state. Demonstrating an agreement with QC results could also be necessary to validate future predictions of protonation states. The results are reported in Figures 7a and 7b which compare, respectively, ΔE(SIBFA) and ΔE(HF), and ΔE_{tot}(SIBFA), ΔE(B97-D3), and ΔE(B3LYP-D3). An equally good match is observed at either level as the one found with the hydroxamate form.

Conformation energy differences

Another important issue for the energy balances is whether the variations of the SIBFA ligand conformational energies in the protein and in water are properly accounted for as compared with QC. For such an evaluation, we have extracted ligands *a–d* from their energy-minimized complexes with the 64 waters at the outcome of EM in the structures denoted A-*a* to C-*d*′ above. They were energy-minimized in the presence of the Langlet–Claverie Continuum procedure. Single-point computations were done at the outcome of EM by both SIBFA and QC. The rankings of relative conformation energies, δE_{conf}, were compared using the most stable conformer as energy zero. The results are reported in Figure 8a for ligand *a* comparing the evolutions of δE_{conf}(SIBFA) and δE_{conf}(HF) and in Figure 8b comparing $\delta E_{conftot}$(SIBFA) including the dispersion contribution and δE_{conf}(MP2). The six conformers of ligand *a* are represented in Supplementary material S6. These figures illustrate the diversity of likely conformations of this ligand, which can be stabilized by competing H-bonds of the hydroxyl groups with the ester or anionic oxygens of the phosphate, with the oxygens of hydroxamate, or between themselves. The corresponding curves for ligands *b–d* are given in Supplementary material S7. Both δE_{conf}(HF) and δE_{conf}(MP2) can be reproduced closely by δE_{conf}(SIBFA) and $\delta E_{conftot}$(SIBFA), respectively. It is noted that no calibration of the rotation barriers was attempted beforehand, as we retained standard values of 1.5, 2.3, and 0 kcal/mol around the C—-O, C—-C, and P—-O bonds, respectively. The trends are only qualitatively dictated by those of δE_1 even though δE_{conf} is dominated by the mutual electrostatic repulsions of the two anionic groups (not shown). Thus conformers A′-*a* and B-*a* have similar relative stabilities in terms of δE_1, while it is δE_2 that enables to preferentially stabilize A′-*a* with respect to B-*a*. This should lend credence to the procedure followed to compute the intramolecular polarization energy detailed in Ref. [47]. Equally satisfactory agreements between SIBFA and QC conformational energies are found for ligands *b* and *c*. For ligand *d*, however, three high-lying conformers, B-*c*, B′-*d* and C-*d* are more destabilized by SIBFA than by QC.

Conformation-dependent continuum solvation energy of the four ligands

The last determinant is continuum solvation. Previous validations of the L-C procedure were published for models of the binding site of MBL[9] and PMI.[10] In these studies the ligand

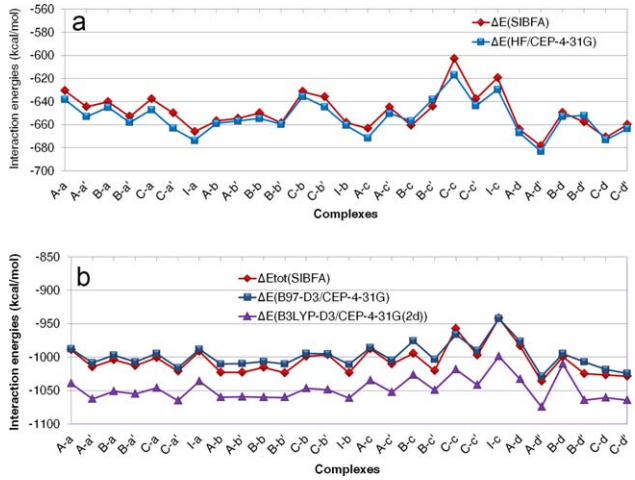

Figure 7. Complexes of ligands *a-d* in the hydroxamic acid form with a shell of 64 waters. Compared evolutions of (a) ΔE(SIBFA) without dispersion and ΔE(HF/QC) and of (b) ΔE_{tot} (SIBFA) including dispersion and ΔE(DFT/B97-D3). [Color figure can be viewed at wileyonlinelibrary.com]

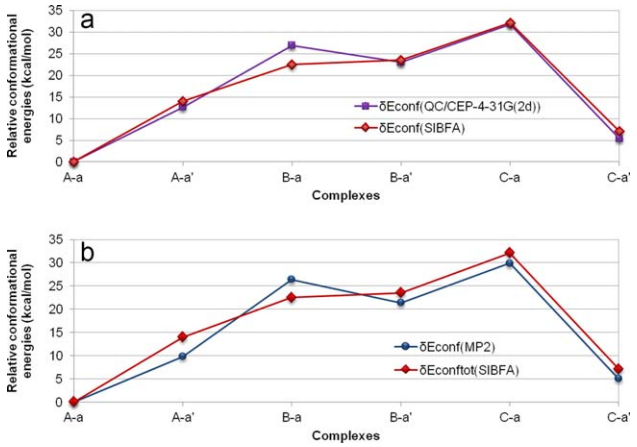

Figure 8. Ligand *a*. Compared evolutions of (a) δE_{conf}(SIBFA) without dispersion and δE_{conf}(HF) and (b) $\delta E_{conftot}$(SIBFA) including dispersion and δE_{conf}(MP2) for a series of solvated conformations. [Color figure can be viewed at wileyonlinelibrary.com]

was to a large extent shielded from the solvent by the side-chains of the PMI recognition site. The present tests are limited to ligands fully exposed to the solvent, but are demanding owing to the presence of two anionic ligands having net charges of −2 and −1, yet separated by only three methylene groups. This results into very strong intra-molecular polarization effects and induced dipoles, which after iterations could contribute to the potential generated by the solute on its van der Waals surface. Note that for the present evaluation we do not take into account the additional iterative polarization between the solute induced dipoles and the solvent. Which accuracy could then be expected? We compare in Figure 9a the evolutions of the electrostatic contribution of ΔG_{solv}(LC) to those of the electrostatic contribution of ΔG_{solv}(PCM). The corresponding evolutions of the complete ΔG_{solv}(LC) and ΔG_{solv}(PCM) are reported in Figure 9b. It is seen that in both cases the LC procedure can match very satisfactorily the PCM calculations, in spite of the very large magnitudes of ΔG_{solv}(PCM) and E_{el}(PCM). Inclusion of the coupling between the solute induced dipoles and those induced on

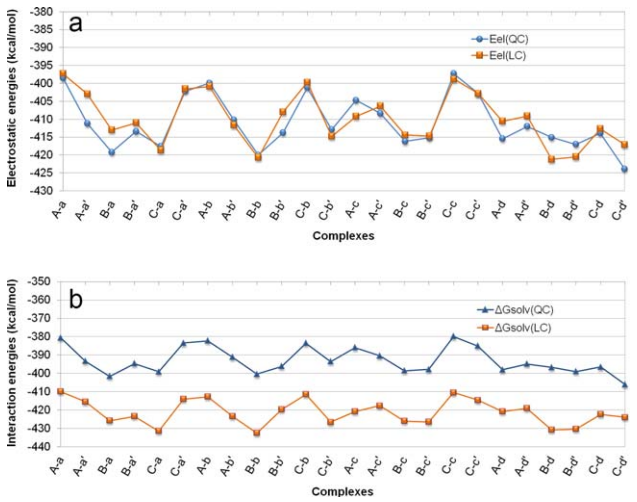

Figure 9. Ligands *a–d*. Compared evolutions of (a) the electrostatic contributions of ΔG_{solv}(PCM) and ΔG_{solv}(LC) and (b) ΔG_{solv}(PCM) and ΔG_{solv}(LC). [Color figure can be viewed at wileyonlinelibrary.com]

the solvent as was recently done in Ref. [68] should further improve the agreements.

Toward relative energy balances. Realistic energy balances to compare the energy balances of complexation of ligands *a–d* to PMI should be envisaged only after a much more exhaustive sampling of the energy surface than in the present context, and inclusion of entropy effects. This clearly appears in Figures 5a–5c and 7a and 7b for the solvation of the hydroxamate and hydroxamic acid forms of ligands *a–c*. There is no convergence toward a single global energy minimum that could be used to quantify the desolvation/conformation energy rearrangement costs of these ligands in the energy balances, and there are no entropy effects. Such samplings, which should be enabled by Tinker-HP, are planned after the porting of the SIBFA potential on this software is completed (Lagardere et al., to be published). The present results indicate that inclusion of continuum solvation and accounting for conformational energy changes of the ligand and of PMI favors complexes A over B and C, owing to the summed second-order contributions, E_2, as well as to ΔG_{solv}(LC). It will be instructive to see if this will hold at the outcome of long-duration MD. Notwithstanding the present limited sampling, the solvation ΔH values do not appear to be favored by hydroxylation, which would have disfavored ligand *a* in the energy balances (see Figures 5 and 7). There remains a possibility that the solvation entropy of ligand *a* is actually smaller than those of *b* and *c*, themselves smaller than that of *d*. This could be due to reduced conformational flexibility due to hydroxylation favoring the formation of transient intra-ligand hydrogen bonds with the phosphate and hydroxamate/hydroxamic acid groups. Reduced flexibility would in turn reduce the thermalization of the first- and second shell waters as previously proposed in Ref. [11] upon comparing the relative affinities for PMI of a phosphate and a malonate derivative of d-mannose. To what an extent could the preferential affinity of ligand *a* for PMI stem from a lesser $T\Delta S_{solv}$ *prior* to complexation, and could this be accounted for by long-duration MD in solution? This should also be addressed as a continuation of this work.

Conclusions and Perspectives

The advent of massively parallel codes for polarizable molecular dynamics now enables MD simulations on a much larger scale than presently affordable, by at least two if not three orders of magnitude. A most topical target for drug design is the class of Zn(II) metalloproteins, which are involved in numerous diseases.[1] Metalloproteins are a most sensitive class of targets in the context of molecular mechanics/dynamics, owing to the onset of large polarization/charge-transfer and non-additivity. Therefore prior to very long duration MD simulations with polarizable potentials, it is essential to validate the accuracy of the energy function by comparisons with *ab initio* QC calculations. We have focused on the complexes of a Zn-metalloenzyme, phosphomannose isomerase, with a lead compound, 5PAH, denoted as *a*, bearing a dianionic phosphate group, a monoanionic hydroxamate one, and three hydroxyls. We compared its binding energies to those of two analogs, *b* and *c*, devoid of one

hydroxyl group, and to one analog, *d,* devoid of all three hydroxyl groups. Three representative poses, A–C, were selected and refined from short-duration MD. In all three, the dianionic phosphate is anchored at the entrance of the binding site by Arg304 and Lys310, while the hydroxamate is bound to Zn(II) with differing coordination motives. Structure A is the only one with a bidentate coordination of hydroxamate to Zn(II), and with a direct interaction of the proximal hydroxyl to Lys136. In all structures, two-dimensional ordered networks of discrete waters were identified, connecting residues of the Zn- as well as phosphate-binding site, to the N-terminal residues, similar to the situation occurring with the complex of a D-mannose malonate ligand with PMI.[11]

The first SIBFA vs. QC tests bore on the 12 A–C complexes of ligands *a–d*, denoted from A-*a* to C-*d*. We compared the SIBFA and QC intermolecular interaction energies values in models of the binding sites, with and without the structural waters, totaling up to 264 atoms. The comparisons were done at the HF and DFT-D3 levels and with two different basis sets, CEP 4-31G(2d) and cc-pVTZ(-f). They showed good agreements regarding the evolution of ΔE along the series, the ranking in terms of relative energies, and the actual magnitudes of ΔE. Regarding the latter, the relative error of ΔE(SIBFA) with respect to ΔE(QC) was $<3\%$. The most satisfactory agreements were those obtained at the most advanced level in this study, namely between ΔE_{tot}(SIBFA) resorting to aug cc-pVTZ(-f) multipoles and polarizabilities and ΔE(B97-D3) with the cc-pVTZ(-f) basis set in the presence as well as in the absence of the 28 discrete water molecules (Figs. 4b and 4d).

The second tests concerned the solvation of the four ligands in a droplet of 64 water molecules. Both anionic and protonated forms of hydroxamic acid were considered. The water molecules undergo the potential and field of the tri- or dianionic ligand, extending beyond the first-shell. These give rise to very large second-order effects, further amplified by the cooperative nature of E_{pol} in ordered water networks.[60,67,69] The comparisons of SIBFA with QC bore on up to 26 different complexes. There was a notable similarity between the SIBFA and the QC curves, even though the first- and second-order SIBFA contributions had in many cases distinct if not antagonistic behaviors (Figs. 5 and 7). This reemphasizes the need for the separability of ΔE_{tot} and a proper balance of first- and second-order terms and control of non-additivity. The fully dehydroxylated ligand *d* but not ligand *a* was the one for which a conformation with the most stable solvation energy was found. This counter-intuitive finding was confirmed by the QC calculation, but more exhaustive sampling should clearly be necessary. Of interest was an "ice-like" structure, I, found in the case of ligand *a*, in which several waters are arranged as in a three-dimensional parallelepiped anchored to the dianionic phosphate. Such a structure ranked among the most stable ones in the case of ligand *a* but not of ligand *d*. This suggests that the hydroxyl groups of *a* vicinal to the phosphate could assist in the local ordering of the waters.

Two other SIBFA/QC comparisons concerned the evolution of the intramolecular energies, δE_{conf}, of ligands *a–d* in a series of energy-minimized conformations on the one hand, and those of their continuum solvation energies ΔG_{solv} on the other hand. With the exception of three high-lying δE_{conf} values of ligand *d*, close agreements were also found between the two approaches. This could not be granted a priori, owing to the very large polarization effects caused by the vicinity of the two ionic groups, and the manifold of competing intramolecular H-bonds involving the hydroxyl groups with either phosphate, hydroxamate, or between themselves.

Is the present accuracy of the procedure satisfactory enough for us to consider large-scale MD as soon as the integration of SIBFA into Tinker-HP is achieved? Two apparently antagonistic features are to be considered. On the one hand, extensive sampling could level off/average out a large part of the errors inherent to the interaction potential. This has been put forth to justify the success of Free Energy Perturbation approaches in the context of "classical," force-field, including non-polarizable ones, as some of them, being highly refined other the years to model specific and well documented (i.e., embodying large experimental data) systems can achieve "chemical accuracy" even though large energy differences with respect to QC could be found in model complexes. On the other hand, there are numerous cases in which several competing ligand binding modes can be generated in the course of MD or by *de novo* drug design algorithms when, as is the case with PMI, there is no available crystal structure for the ligand-bound target. Imbalances in the relative stabilities of competing binding modes could result into MD stabilizing modes erroneously ranked as the most stable ones, and the underrepresentation or obliteration of those which a more refined potential would have selected. All the tests reported in this study indicate that the SIBFA potential should now be considered for large-scale MD and FEP. Nevertheless, regular validations by QC calculations of the most stable MD snapshots, have become affordable and are commendable. Along these lines, there are newly-emerging applications of polarizable molecular mechanics: namely, their use to sample the energy surface and perform, for the best poses, QM/MM or QM/PMM calculations to reconstruct a QC-like trajectory. This has been reported concerning "classical" potentials for the simulation of liquid water,[70] for the solvation energy of a series of flexible ligands,[71] and for ligand-binding free energies.[72] In such approaches, the larger the overlap between the QC and MM energies, the better the sampling.

The present results could also be used to validate other polarizable molecular mechanics, as well as semi-empirical QC procedures. For that purpose, the coordinates of complexes A-*c* to C-*d* are given as Supplementary material S8.

Keywords: Zn-metalloenzymes·hydroxamate ligands·polarizable molecular mechanics/dynamics·quantum chemistry·drug design

How to cite this article: N. Gresh, D. Perahia, B. de Courcy, J. Foret, C. Roux, L. El-Khoury, J.-P. Piquemal, L. Salmon. *J. Comput. Chem.* **2016**, *37*, 2770–2782. DOI: 10.1002/jcc.24503

Additional Supporting Information may be found in the online version of this article.

[1] F. E. Jacobsen, J. A. Lewis, S. M. Cohen, *ChemMedChem* **2007**, *2*, 152.
[2] J. W. Ponder, TINKER Molecular Modeling Package, http://dasher.wustl.edu/tinker/, **2016**.
[3] TinkerHP_Team, Tinker-HP, http://www.ip2ct.upmc.fr/tinkerHP/, **2016**.
[4] F. Lipparini, L. Lagardère, C. Raynaud, B. Stamm, E. Cancès, B. Mennucci, M. Schnieders, P. Ren, Y. Maday, J. P. Piquemal, *J Chem Theory Comput* **2015**, *11*, 623.
[5] N. Gresh, G. A. Cisneros, T. A. Darden, J. P. Piquemal, *J Chem Theory Comput* **2007**, *3*, 1960.
[6] N. Gresh, *J Comput Chem* **1995**, *16*, 856.
[7] G. Tiraboschi, N. Gresh, C. Giessner-Prettre, L. G. Pedersen, D. W. Deerfield, *J Comput Chem* **2000**, *21*, 1011.
[8] N. Gresh, J. P. Piquemal, M. Krauss, *J Comput Chem* **2005**, *26*, 1113.
[9] J. Antony, J. P. Piquemal, N. Gresh, *J Comput Chem* **2005**, *26*, 1131.
[10] C. Roux, N. Gresh, L. E. Perera, J. P. Piquemal, L. Salmon, *J Comput Chem* **2007**, *28*, 938.
[11] N. Gresh, B. de Courcy, J. P. Piquemal, J. Foret, S. Courtiol-Legourd, L. Salmon, *J Phys Chem B* **2011**, *115*, 8304.
[12] N. Gresh, K. El Hage, D. Perahia, J. P. Piquemal, C. Berthomieu, D. Berthomieu, *J Comput Chem* **2014**, *35*, 2096.
[13] R. W. Gracy, E. A. Noltman, *J Biol Chem* **1968**, *243*, 3161.
[14] B. Wu, Y. Zhang, R. Zheng, C. Guo, P. G. Wang, *FEBS Lett* **2002**, *519*, 87.
[15] J. H. Patterson, R. F. Waller, D. Jeevarajah, H. Billman-Jacobe, M. J. McConville, *Biochem J* **2003**, *372*, 77.
[16] D. Shinabarger, A. Berry, T. B. May, R. Rothmel, A. Fialho, A. M. Chakrabarty, *J Biol Chem* **1991**, *266*, 2080.
[17] A. Garami, T. Ilg, *J Biol Chem* **2001**, *276*, 6566.
[18] M. A. Payton, M. Rheinnecker, L. S. Klig, M. DeTiani, E. Bowden, *J Bacteriol* **1991**, *173*, 2006.
[19] E. A. Wills, I. S. Roberts, M. Del Poeta, J. Rivera, A. Casadevall, G. M. Cox, J. R. Perfect, *Mol Microbiol* **2001**, *40*, 610.
[20] D. J. Smith, M. A. Payton, *Mol Cell Biol* **1994**, *14*, 6030.
[21] D. J. Smith, A. E. I. Proudfoot, M. De Tiani, T. N. C. Wells, M. A. Payton, *Yeast* **1995**, *11*, 301.
[22] R. Hardré, C. Bonnette, L. Salmon, A. Gaudemer, *Bioorg Med Chem Lett* **1998**, *8*, 3435.
[23] R. Hardré, L. Salmon, *Carbohydr Res* **1999**, *318*, 110.
[24] C. Roux, J. H. Lee, C. J. Jeffery, L. Salmon, *Biochemistry* **2004**, *43*, 2926.
[25] C. Roux, F. Bhatt, J. Foret, B. de Courcy, N. Gresh, J. P. Piquemal, C. J. Jeffery, L. Salmon, *Proteins* **2011**, *79*, 203.
[26] E. P. Serjeant, B. Dempsey, Ionisation Constants of Organic Acids in Aqueous Solution, Pergamon Press, Oxford; New York, **1979**.
[27] R. Hardré, PhD Thesis, Université Paris-Sud (Orsay, France), **2001**.
[28] J. K. M. Roberts, N. Wade-Jardetzky, O. Jardetsky, *Biochemistry* **1981**, *20*, 5389.
[29] J. Zhang, W. Yang, J. P. Piquemal, P. Ren, *J Chem Theory Comput* **2012**, *8*, 1314.
[30] W. J. Stevens, H. Basch, M. Krauss, *J Chem Phys* **1984**, *81*, 6026.
[31] W. J. Stevens, M. Krauss, H. Basch, P. G. Jasien, *Can J Chem* **1992**, *70*, 612.
[32] T. H. Dunning, *J Chem Phys* **1989**, *90*, 1007.
[33] D. Feller, *J Comput Chem* **1996**, *17*, 1571.
[34] D. Figgen, G. Rauhut, M. Dolg, H. Stoll, *Chem Phys* **2005**, *311*, 227.
[35] A. D. Becke, *Phys Rev A* **1988**, *38*, 3098.
[36] C. Lee, W. Yang, R. G. Parr, *Phys Rev B* **1988**, *37*, 785.
[37] A. D. Becke, *J Chem Phys* **1993**, *98*, 1372.
[38] S. Grimme, J. Antony, S. Ehrlich, H. Krieg, *J Chem Phys* **2010**, *132*, 154104.
[39] J. Tomasi, B. Mennucci, R. Cammi, *Chem Rev* **2005**, *105*, 2999.
[40] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis, J. A. Montgomery, *J. Comput. Chem*. **1993**, *14*, 1347.
[41] Gaussian 09, Revision E.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski, and D. J. Fox, Gaussian, Inc., Wallingford CT, **2009**.
[42] N. Gresh, P. Claverie, A. Pullman, *Theoret Chim Acta* **1984**, *66*, 1.
[43] F. Vigné-Maeder, P. Claverie, *J Chem Phys* **1988**, *88*, 4934.
[44] J. P. Piquemal, N. Gresh, C. Giessner-Prettre, *J Phys Chem A* **2003**, *107*, 10353.
[45] D. R. Garmer, W. J. Stevens, *J Phys Chem A* **1989**, *93*, 8263.
[46] C. H. Faerman, S. L. Price, *J Am Chem Soc* **1990**, *112*, 4915.
[47] N. Gresh, G. B. Shi, *J Comput Chem* **2004**, *25*, 160.
[48] A. J. Stone, *J Phys Chem A* **2011**, *115*, 7017.
[49] S. Creuzet, N. Gresh, J. Langlet, *J Chim Phys Phys Chim Biol* **1991**, *88*, 2399.
[50] G. A. Evangelakis, J. P. Rizos, I. E. Lagaris, I. N. Demetropoulos, *Comput Phys Commun* **1987**, *46*, 401.
[51] N. Gresh, J. E. Sponer, M. Devereux, K. Gkionis, B. de Courcy, J. P. Piquemal, J. Sponer, *J Phys Chem B* **2015**, *119*, 9477.
[52] J. Langlet, P. Claverie, J. Caillet, A. Pullman, *J Phys Chem* **1988**, *92*, 1617.
[53] A. Cleasby, A. Wonacott, T. Skarzynski, R. E. Hubbard, G. J. Davies, A. E. I. Proudfoot, A. R. Bernard, M. A. Payton, T. N. C. Wells, *Nat Struct Biol* **1996**, *3*, 470.
[54] P. Claverie, J. P. Daudey, J. Langlet, B. Pullman, D. Piazzola, M. J. Huron, *J Phys Chem* **1978**, *82*, 405.
[55] W. C. Swope, H. C. Andersen, P. H. Berens, K. R. Wilson, *J Chem Phys* **1982**, *76*, 637.
[56] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, J. R. Haak, *J Chem Phys* **1984**, *81*, 3684.
[57] J. P. Ryckaert, G. Ciccotti, H. J. Berendsen, *J Comput Phys* **1977**, *23*, 327.
[58] J. Antony, N. Gresh, L. Olsen, L. Hemmingsen, C. Schofield, R. Bauer, *J Comput Chem* **2002**, *23*, 1281.
[59] N. Gresh, N. Audiffren, J. P. Piquemal, J. de Ruyck, M. Ledecq, J. Wouters, *J Phys Chem B* **2010**, *114*, 4884.
[60] H. Guo, N. Gresh, B. P. Roques, D. R. Salahub, *J Phys Chem B* **2000**, *104*, 9746.
[61] J. Tomasi, M. Persico, *Chem Rev* **1994**, *94*, 2027.
[62] J. Foret, B. de Courcy, N. Gresh, J. P. Piquemal, L. Salmon, *Bioorg. Med. Chem* **2009**, *17*, 7100.
[63] A. J. Misquitta, R. Podeszwa, B. Jeziorski, K. Szalewicz, *J Chem Phys* **2005**, *123*, 214103.
[64] A. J. Misquitta, *J Chem Theory Comput* **2013**, *9*, 5313.
[65] R. Z. Khaliullin, E. A. Cobar, R. C. Lochan, A. T. Bell, M. Head-Gordon, *J Phys Chem A* **2007**, *111*, 8753.
[66] B. Pullman, A. Pullman, H. Berthod, N. Gresh, *Theoret Chim Acta* **1975**, *40*, 93.
[67] J. P. Piquemal, H. Chevreau, N. Gresh, *J. Chem. Theory Comput*. **2007**, *3*, 824.
[68] F. Lipparini, L. Lagardère, B. Stamm, E. Cancès, M. Schnieders, P. Ren, Y. Maday, J. P. Piquemal, *J Chem Theory Comput* **2014**, *10*, 1638.
[69] N. Gresh, *J Phys Chem* **1997**, *101*, 8680.
[70] M. Del Ben, M. Schönherr, J. Hutter, J. VandeVondele, *J Phys Chem Lett* **2013**, *4*, 3753.
[71] E. C. Dybeck, G. König, B. R. Brooks, M. R. Shirts, *J Chem Theory Comput* **2016**, *12*, 1466.
[72] M. A. Olson, P. Soderhjelm, U. Ryde, *J. Comput. Chem*. **2016**, *37*, 1589.

Received: 22 July 2016
Revised: 31 August 2016
Accepted: 4 September 2016
Published online on 4 October 2016