JCTC Journal of Chemical Theory and Computation

Toward a Separate Reproduction of the Contributions to the Hartree−Fock and DFT Intermolecular Interaction Energies by Polarizable Molecular Mechanics with the SIBFA Potential[†]

Jean-Philip Piquemal,*[,‡] Hilaire Chevreau,[‡] and Nohad Gresh*[,§]

Laboratoire de Chimie Théorique, UMR 7616, Université P. & M. Curie, Case courrier 137, 4, place Jussieu, F. 75252 Paris, Cedex 05, France, and Laboratoire de Pharmacochimie Moléculaire et Structurale, U648 INSERM, IFR Biomédicale, 45, Rue des Saints-Pères, 75006, Paris, France

Abstract: Following recent refinements of the SIBFA intermolecular potential to the multipolar electrostatic contribution by inclusion of an explicit 'penetration' component, the short-range repulsion term is augmented with a S^2/R^2 component. The SIBFA potential, and the behaviors of its individual contributions encompassing polarization and charge transfer, were evaluated in a diversity of hydrogen-bonded complexes as well as in a model stacked complex by comparisons with results from ab initio quantum-chemical (QC) computations with energy decomposition. Close agreements between SIBFA and QC results are found on both the interaction energies and their contributions. Extensions to computations at the DFT level are also presented.

Introduction

In Anisotropic Polarizable Molecular Mechanics (APMM) procedures,[1−8] representation of the electrostatic contribution of the interaction energy with distributed multipoles constitutes an essential asset. Following the developments pioneered by Stone (see ref 9 and references herein) and Claverie,[10] the multipoles are extracted from the quantum-chemical (QC) wave function of the molecule considered and stored in a library. The electrostatic interaction energy between two interacting molecules is then computed as a sum of multipole−multipole interactions. This enables to faithfully reproduce the anisotropic features of the Coulomb contribution of a corresponding ab initio supermolecule computation.[11,13−15] However, the Coulomb component embodies, in addition to the multipolar component, attractive effects due to charge penetration.[16,17] The corresponding energy term, denoted as E_{pen}, is not explicitly represented in virtually all APMM procedures. Instead, on account of its overlap-dependent character, it is lumped together with the short-range repulsion contribution. The first explicit introduction of E_{pen} was done in the framework of the Effective Fragment Potential (EFP) method.[16] It was recently introduced[17] as well in the SIBFA (Sum of Interactions Between Fragments Ab initio computed)[1,2] procedure and in the GEM (Gaussian Electrostatic Model) approach[4] using different formalisms. The use of energy-decomposition procedures in the ab initio supermolecular approach unravels the weights of the individual component of the Hartree−Fock (or DFT) interaction energy, ΔE_{HF} (or ΔE_{DFT}) between two, or several, interacting molecules and their distance and angular dependencies.[11,13,15,18]

It was found using SIBFA that the multipolar contribution E_{MTP} augmented with by E_{pen}, denoted as E_{MTP*}, could closely reproduce the numerical values of the ab initio Coulomb contribution E_c for diverse representative complexes, such as neutral or ionic H-bonded, or stacking as well as cation−ligand complexes.[17] The Frozen Core term, E_{FC}, from an ab initio energy-decomposition procedure is the sum of the Coulomb, E_c, and the short-range exchange, E_{exch}, contribu-

tions. In this work and by analogy to perturbation theory we will term it E_1. The corresponding term in SIBFA is $E_{MTP} + E_{pen} + E_{rep}$. Since the first two terms should match $E_c(HF)$, E_{rep} should now match as closely as possible E_{exch}, rather than it be calibrated to match the actual $E_{FC}(HF)-E_{MTP}(SIBFA)$ difference. While this could on principle be attained by simply rescaling the multiplicative factor of E_{rep}, we sought for further refinements in its formulation. Such comparisons are carried out parallel to the corresponding comparisons between E_{MTP*} and E_c. In 'classical' molecular mechanics using point charges to compute electrostatics, the repulsion contribution is generally computed as a sum of 'isotropic' $1/R^n$ terms. In the context of APMM, it would be desirable to have a representation of E_{exch} that accounts as closely as possible for both its distance and its angular dependencies. Apart from SIBFA, only a few APMM procedures endow E_{rep} with angular features[3] and with dependencies upon the electronic populations of the interacting atoms:[19,20] that is, the more electron-rich a given atom, the larger its contribution to the repulsion. However to our knowledge, there are scarce reports that do confront the angular dependencies of E_{rep} to those of E_{exch}. Examples of such reports were published using the SIBFA procedure for both hydrogen-bonded[21] and cation−ligand complexes.[22] Since the initial inception of this procedure, and following the early proposals by Murrell et al.,[23] a dependence of E_{rep} upon a functional, S^2, of the square of the overlap between the interacting molecules was sought for. Thus E_{rep} was expressed under the form of a sum of bond−bond, bond−lone pair, and lone pair−lone pair overlaps interactions.[24] Denoting by R the distance between the centroids of the simulated localized orbitals, further developments resorted to a S^2/R formula instead of an S^2 one and explicitly introduced the effects of the hybridization of the orbitals localized on the bonds, not just those localized on the lone pairs.[22] In the present work, we will seek to confer more flexibility to such a representation. Again following the work by Murrell and Teixeira-Dias,[23] we will include an additional term, with an actual S^2/R^2 dependency.[25] This introduction occurs at the cost of only a minor calibration effort, since the amplitude of each of the S^2/R and the S^2/R^2 terms is governed by only two parameters, namely a multiplicative coefficient and the exponent of the exponential. As in our preceding papers, the values of the effective radii of each involved atom depend on its chemical nature and hybridization, are identical for both S^2 components, and are transferable. The present work constitutes a generalization[25] of the one recently published by us[26] which bore on the refinements of Zn(II) representation in this context. The formulation of the two second-order contributions, polarization (E_{pol}) and charge-transfer (E_{ct}), is the same as in our previous papers.[26] The organization of the paper is the following.

A further refinement of the expression of E_{pen} is briefly introduced, that now embodies the effects of penetration on the charge-quadrupole component of E_{MTP}, in addition to those exerted on the charge−charge and charge-dipole components. The formulation of E_{rep*} is then presented. Calibration of both E_{MTP*} and E_{rep*} is subsequently done on a training set constituted by the water dimer in the same

five configurations as investigated in our previous paper.[17] This is followed by tests on water clusters. Water boxes in an energy-minimized icelike arrangement are considered for $n = 12$, 16, and 20 molecules, and, for $n = 16$, an arrangement extracted from a Monte Carlo simulation on a large box of 64 molecules is also considered. Three nonstandard H-bonded chains ($n = 12$) are next considered to probe the behavior of E_{pol} in polarizable potentials against their QC counterpart: these involve bifurcated and transverse bifurcated arrangements as introduced by Giese and York[27] as well as helical as some of us have suggested in a recent paper (Piquemal et al., submitted for publication). We then present several tests of the accuracy of E_{MTP*} and E_{rep} in first-order, and E_{pol} and E_{ct} in second-order, by comparing their distance and/or angular dependencies to those of their respective counterparts E_c, E_{exch}, E_{pol}, and E_{ct}. These tests bear on the following hydrogen-bonded complexes: a neutral H-bond complex, the formamide dimer; an anionic H-bond acceptor, formate, with a neutral H-bond donor, water; a cationic H-bond donor, methylammonium, with a neutral acceptor, water; and an ionic complex, formate−methylammonium. It was finally essential to assess if the refinements in the representation of H-bonded complexes would translate into an improved representation of stacked complexes, such that as many polar atoms of one monomer would 'overlap' those of the other monomer. This was done on the complex between two parallel formamide molecules upon performing rotation of the second monomer around the z-axis. To conclude, we present an extension of this work to interaction energies computed at the correlated DFT level for water using a large basis set. In this context, we also investigate ten hydrogen-bonded complexes that were recently studied in detail by van Duijneveldt et al. to benchmark several molecular mechanics potentials against high level QC computations.[28]

The present work builds up on several previous studies in which SIBFA, ab initio HF, and MP2 as well as DFT calculations were performed in parallel. Several of these were done in close collaboration with Professor Salahub's group and were instrumental in the continued evolution and refinements of this procedure: from bimolecular or small complexes,[21] to intramolecular interactions in di-and oligopeptides,[29,30] and to critical assessment of the handling of cooperativity effects in peptide H-bonded networks[31] and anisotropy in dimerization energies of protein−protein recognition motifs.[32]

Procedure

Ab Initio Calculations. At the Hartree−Fock level, the energy decomposition calculations have been carried using the RVS[11] decomposition scheme at Hartree Fock level, as implemented in GAMESS.[33] The basis sets retained for these computations are the CEP 4-31G(2d)[34] and DZVP2.[35] The SIBFA DZVP2 results required for the SIBFA version dedicated to open shell cation interactions[36] will be discussed if different from the CEP 4-31G(2d) version. The CEP 4-31G(2d) pseudopotential has been shown to provide ΔE values close to those computed with the 6-311G** in models of zinc metalloprotein complexes with inhibitors.[1] At the

826 *J. Chem. Theory Comput., Vol. 3, No. 3, 2007*

Piquemal et al.

DFT level, the chosen functional was B3LYP[37,38] coupled to the aug-cc-PVTZ basis set.[39] All energy decomposition computations at this level have been carried out with a modified version of the CSOV[18] energy decomposition scheme implemented in a modified version of HONDO 95.3[13] enabling the computations of the electrostatic interaction energies in complexes with more than two molecules. Distributed polarizabilities are extracted from a procedure originally due to Garmer and Stevens[40] and implemented at the HF and DFT level in our "in house" version of HONDO 95.3. Jaguar 6.0[41] has also been used to calculate the DFT total interaction energies in water oligomers.

Extended Formulation for E_{MTP*}. E_{MTP*} is calculated using distributed multipoles, up to quadrupoles and derived from the ab initio wave function of the molecule considered. They are distributed on its atoms and the barycenters of its chemical bonds following a procedure due to Vigne-Maeder and Claverie.[10]

E_{MTP*} was computed[17] as a sum of six terms:

$$E_{MTP*} = E_{mono-mono*} + E_{mono-dip*} + E_{mono-quad} + E_{dip-dip} + E_{dip-quad} + E_{quad-quad} \quad (1)$$

To take into account the short-range electrostatic penetration effect, we have in ref 17 modified two terms of the classical E_{MTP}, which are both related to monopole interaction ($E_{mono-mono*}$ and $E_{mono-dip*}$). Even though these modifications gave accurate results, we have extended here the correction to the monopole–quadrupole term.

The expression is grounded on the ab initio formulation of the Coulomb electrostatic interaction energy, E_c (see refs 13 and 14 and references herein for details)

$$E_c = -2 \sum_i \sum_v Z_v \int (|\varphi_i(1)|^2)/r_{1v})d\iota_1 -$$
$$2 \sum_j \sum_\mu Z_\mu \int (|\varphi_i(1)|^2)/r_{2\mu})d\iota_2 +$$
$$4 \sum_i \sum_j \int (|\varphi_i(1)|^2 |\varphi_j(2)|^2)/r_{12})d\iota_1 d\iota_2 + \sum_\mu \sum_v Z_\mu.Z_v/r_{\mu v} \quad (2)$$

where μ and φ_i are the nuclei and the unperturbed molecular orbitals of monomer A and v and φ_j, those of monomer B.

The monopole–monopole energy for two interacting centers i and j is given by

$$E_{mono-mono*} = [Z_i Z_j - \{Z_i(Z_j - q_j)(1 - \exp(-\alpha_j \cdot r)) + Z_j(Z_i - q_i)(1 - \exp(-\alpha_i \cdot r))\} + (Z_i - q_i)(Z_j - q_j)(1 - \exp(-\beta_i \cdot r))(1 - \exp(-\beta_j \cdot r))] *(1/r) \quad (3)$$

where Z_i and Z_j are the number of valence electrons of the two atoms concerned. In the case of the monopoles located on bonds Z is equal to zero. α_i and β_i are parameters depending on effective van der Waals radii r_{vdw} and are given by

$$\alpha_i = \gamma/r_{vdw\,i} \quad \text{and} \quad \beta_i = \delta/r_{vdw\,I} \quad (4)$$

where γ and δ are parameters depending on the basis set/methodology used for reference ab initio calculations. For bond monopoles the r_{vdw} values are taken equal to the arithmetic mean between those of the atoms forming the bond.

The monopole–dipole energy term is given by

$$E_{mono-dip*} = -\mu_j \cdot \xi* \quad (5)$$

with

$$\xi* = \{Z_i - (Z_i - q_i)(1 - \exp(-\eta r))\} \cdot \mathbf{r_{ij}}/r_{ij}^3 \quad (6)$$

and

$$\eta = \chi/((r_{vdw\,i} + r_{vdw\,j}))/2 \quad (7)$$

where χ is a parameter depending on the basis set/methodology used for reference ab initio calculations. At this point, this formulation includes a correction for terms varying like R^{-1} (monopole–monopole correction), for terms varying like R^{-2} (monopole–dipole correction) but does not include correction for terms varying like R^{-3} (dipole–dipole and monopole–quadrupole).

The standard monopole–quadrupole interaction is given by

$$E_{mono-quad} = E_{mono-quad1} + E_{mono-quad2} \quad (8)$$

with

$$E_{mono-quad} = q(Q_a/2r^3)[3(\mathbf{a} \cdot \mathbf{r}/r)^2 - 1] \quad (9)$$

where $E_{mono-quad1}$ and $E_{mono-quad2}$ are respectively the interaction energy of a monopole interacting with an axial quadrupole (the two axial quadrupoles representing the true quadrupole are different for $E_{mono-quad1}$ and $E_{mono-quad2}$). $\mathbf{a}$ defines the unit vector defining the axis, $\mathbf{r}$ is the vector along r, directed from the monopole to axial quadrupole, and Q_a is the corresponding quadrupole magnitude with direction $\mathbf{a}$.

The energy can be refined by modifying the monopole q

$$E_{mono-quad*} = E_{mq1*} + E_{mq2*} \quad (10)$$

with

$$E_{mq1*} = \{Z_i - (Z_i - q_i)(1 - \exp(-\varphi r))\}(Q_a/2r^3)[3(\mathbf{a} \cdot \mathbf{u})^2 - 1] \quad (11)$$

and

$$\varphi = \Omega/((r_{vdw\,i} + r_{vdw\,j}))/2 \quad (12)$$

where Ω is a parameter depending on the basis set/methodology used for reference ab initio calculations.

Thus E_{MTP*} is given by

$$E_{MTP*} = E_{mono-mono*} + E_{mono-dip*} + E_{mono-quad*} + E_{dip-dip} + E_{dip-quad} + E_{quad-quad} \quad (13)$$

where γ, δ, χ, and φ were fit so that E_{MTP*} reproduces E_c on the linear and bifurcated water dimers reported in this study.

Formulation of the Short-Range Repulsion Intermolecular Interaction Energy. Initially, E_{MTP}, lacking the attractive penetration component E_{pen}, was systematically less

attractive than E_c. Thus we had chosen a S^2/R formulation of E_{rep} in order to include E_{pen} in the repulsion since a similar dependency upon in S^2/R is observed for both exchange-repulsion and penetration.[16,42] E_{rep} was calibrated so that the actual sum of $E_{MTP} + E_{rep}$, namely E_1(SIBFA), matches the corresponding sum of $E_c + E_{exch}$, namely E_1(RVS). Since E_{MTP*} should now match E_c, E_{rep} should correspondingly match E_{exch}. This has led us to accordingly search for an improved reformulation of E_{rep}, while still retaining a molecular orbital (MO) overlaplike formulation.

E_{rep}(SIBFA) is now formulated as a sum of bond–bond, bond–lone pair, and lone pair–lone pair interactions under the form of an $S^2/R + S^2/R^2$ formulation. **S** denotes the overlap between localized MOs of the interacting partners expressed under the form of sums of Slater hybrid orbitals around the pair of atoms making up the bonds and as hybrids around the lone pair bearing-heteroatoms. The localized MOs are represented by centers along the chemical bonds and the lone pairs of heteroatoms.

In the case of two interacting molecules A with bonds AB and lone pairs L_α and C with bonds CD and lone pairs $L\gamma$, E_{rep} has the form

$$\mathbf{E_{rep}} = C_1(\sum_{AB}\sum_{CD}\mathbf{rep}(AB,CD) + \sum_{AB}\sum_{L\gamma}\mathbf{rep}(AB,L\gamma) + \sum_{L\alpha}\sum_{CD}\mathbf{rep}(L_\alpha,CD) + \sum_{L\alpha}\sum_{L\gamma}\mathbf{rep}(L_\alpha,L_\gamma)) \quad (14)$$

Each term of this equation depends upon a functional, **S**, of the overlap as

$$\mathbf{rep}(AB,CD) = N_{occ}(AB)N_{occ}(CD)\mathbf{S}^{**2}(AB,CD)/(D_{AB,CD})^n \quad (15)$$

where $n = 1$ or 2, and $N_{occ}(AB)$ and $N_{occ}(CD)$ are the occupation numbers of bonds AB and CD. N_{occ} is equal to 2 for doubly occupied bonds and lone pairs and to 1 for π type orbitals. $D_{AB,CD}$ denotes the distance between the barycenters of bonds AB and CD.

The expression for the overlap term **S** was detailed in a preceding paper.[22] The formulation of **S** includes exponentials of the distance between pair of atoms belonging to the interacting bonds or lone pairs. Such distances are divided by the geometric mean of the effective radii of these atoms. As for E_{MTP*} such radii are atom-type dependent. For any given atom type such as $O(sp^2)$, $O(sp)$, etc., they are also determined on the basis of isodensity contour maps around a representative isolated molecule to which they belong. Dependencies of these radii upon the electronic populations of the interacting atoms are considered following an expression described in ref 43.

Polarization and Charge-Transfer Terms

The formulation and calibration of E_{pol} and E_{ct} are identical to those given in ref 26 for all computations done at the HF level, while new parameters have been calibrated for the DFT computations. E_2 corresponds to the sum of polarization and charge-transfer energies.

Note on Polarization Energies. In the tables, two values of the ab initio E_{pol} energies are given. The first corresponds

to the nonantisymmetrized but fully relaxed Kitaura–Morokuma[44] value and the second to antisymmetrized RVS values. The fully relaxed SIBFA E_{pol} can be compared to the E_{pol}(KM) as only the first iteration of the SIBFA polarization contribution is compared to E_{pol}(RVS). The E_{pol}(RVS) values are given in parentheses in the table as well as the values of E_{pol}(SIBFA) prior to iterating (for details, see Piquemal et al., submitted for publication).

Results and Discussion

Calibration and Tests on Five Model Water Dimers. We have in ref 17 compared the evolutions of E_{MTP*} and of E_c in the five water dimers represented in Figure 1a–e. We have observed that when augmented with E_{pen}, E_{MTP*} could closely match E_c in the range of relevant O–H, O–O, and H–H distances, even in unphysical configurations, such as complexes 1d and 1e. We now compare (Figure 1a–e) the behaviors of E_{rep*} to those of E_{exch}.

The calibration of E_{rep} bore on the effective radii of O and H and the exponents α_1 and α_2 and multiplicative constants C_1 and C_2 of the S^2/R and of the S^2/R^2 terms. They were given in ref 26. It was done in order for E_{rep*} to reproduce the numerical values of E_{exch} in the linear and bifurcated water dimers (complexes 1a and 1c) upon performing variations of the H–O distance. Extension to complexes other than 1a and 1c, such as 1b, 1d, and 1e, have to our knowledge little or no precedents in the development and evaluation of APMM procedures. Such comparisons should be allowed for evaluation if bond–bond, bond–lone pair, and lone pair–lone pair interactions are correctly expressed and balanced within E_{rep}. We note in particular the predominance of lone pair–lone pair repulsion in configuration d and that of bond–bond repulsion in configuration e, while bond–lone pair repulsion should be the dominant repulsive contribution in a and c. To what an extent will these varying weights enable the reproduction of the numerical values of E_{exch} and its radial behaviors?

Figure 1a–e shows close agreements throughout the range of relevant H–O distances (>1.7 Å) as well as O–O and H–H distances (>2.7 and >1.5 Å respectively). Thus, for complexes 1a, 1b, and 1c, the errors amount to 0.17, 0.07, and 0.11 kcal/mol, respectively, at equilibrium distance. For complexes 1d and 1e, they amount to 0.08 and 0.10 kcal/mol at the representative H–H and O–O distances of 2.3 and 2.8 Å, respectively.

The values of ΔE(SIBFA) and ΔE(RVS) and their respective contributions at equilibrium distances for complexes 1a–c and for complexes 1d and 1e are reported in Table 1. It shows that the close agreements between ΔE(SIBFA) and ΔE(RVS) are due to corresponding agreements at the level of the individual contributions.

Water Clusters. An essential objective of APMM procedures is the simulation of very large complexes that are not amenable to QC procedures. It is necessary to ensure that the agreement found at the level of bimolecular complexes will be preserved in multimolecular complexes. A critical issue relates to whether E_{pol} and E_{ct} from SIBFA can reproduce the nonadditive behaviors of their RVS counterparts. We have previously addressed the issues of
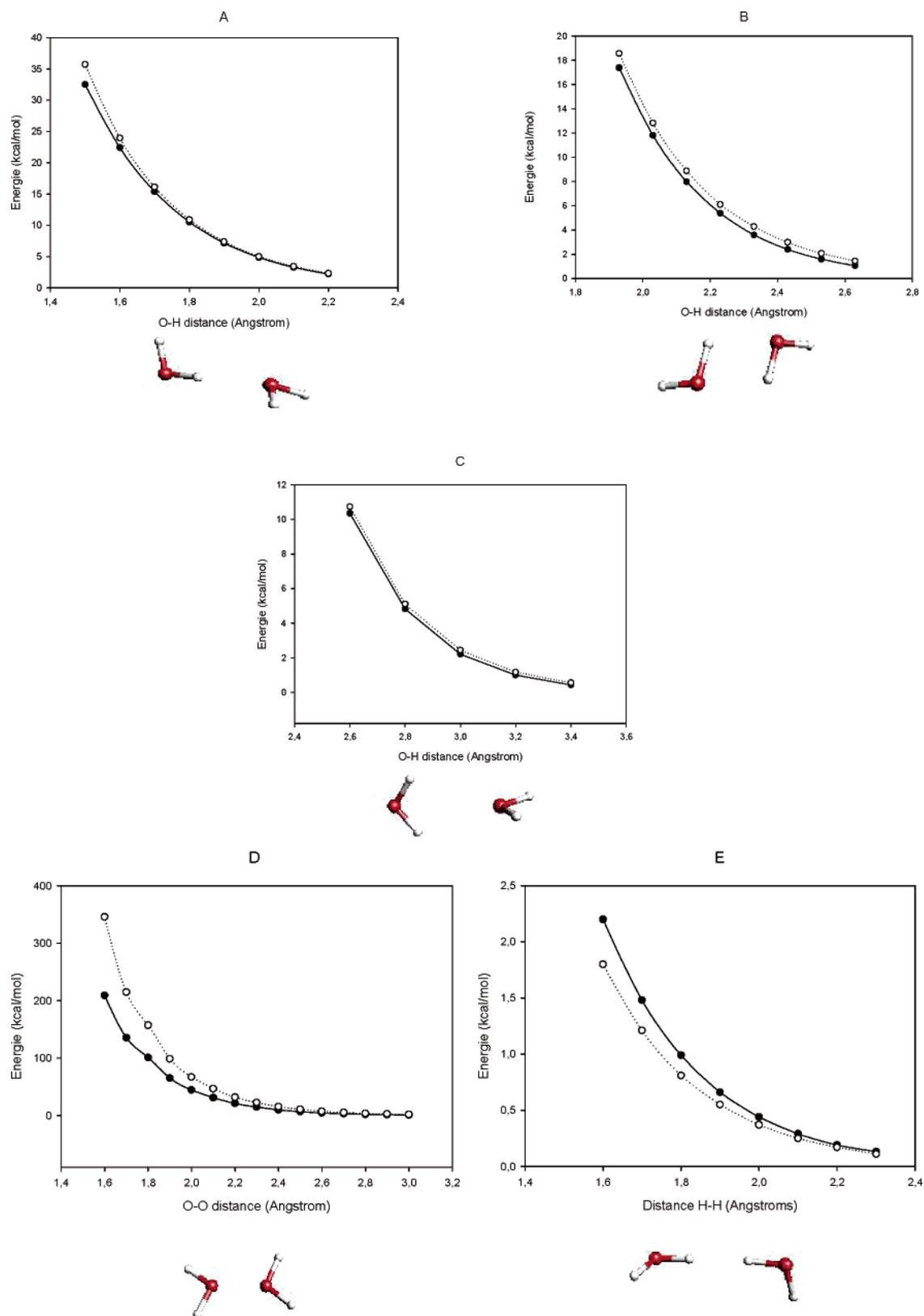
Figure 1. Distance variations of the exchange repulsion component (E_{exch}) (kcal/mol) versus modified SIBFA component (E_{rep}*) for five water dimers at the Hartree−Fock level.

cooperativity in water aggregates[43,45] and peptide H-bonded networks[31] and of anticooperativity in polyligated Zn(II) complexes.[46] We here present results on water aggregates in three-dimensional cubic arrangements with $n = 12, 16,$ and 20 water molecules (shown in Figure 2a for $n = 20$)

which were previously investigated in the framework of the density fitting based GEM force field.[4] Single-point RVS analyses were performed for each of the three energy-minimized structures. We have also performed a similar SIBFA vs RVS comparison in a small aggregate ($n = 16$)

Table 1. RVS and SIBFA Interaction Energies (kcal/mol) for the Water Dimers at the Equilibrium Point or Standard Orientation

energy (kcal/mol)	E_c	E_{exc}	E_1	E_{pol}	E_{ct}	E_2	ΔE
linear dimer (SIBFA)	−5.98	3.42	−2.56	−0.70	−0.60	−1.30	−3.87
linear dimer (RVS)	−5.81	3.26	−2.54	−0.67	−0.63	−1.30	−4.04
cyclic dimer (SIBFA)	−5.42	2.87	−2.55	−0.34	−0.26	−0.60	−3.16
cyclic dimer (RVS)	−4.79	2.19	−2.40	−0.31	−0.29	−0.60	−3.23
bifurcated dimer (SIBFA)	−3.48	1.17	−2.31	−0.20	−0.15	−0.35	−2.67
bifurcated dimer (RVS)	-−3.17	1.00	−2.17	−0.21	−0.19	−0.40	−2.78
H−H dimer (SIBFA)(2.3)	1.84	0.17	2.03	-0.06	0.00	−0.06	1.96
H−H dimer (RVS)	2.11	0.13	2.24	−0.12	−0.12	−0.24	1.88
O−O dimer (SIBFA)(2.8)	1.91	2.41	4.32	−0.31	0.0	−0.31	4.02
O−O dimer (RVS)	2.83	1.96	4.79	−0.38	−0.12	−0.50	4.12
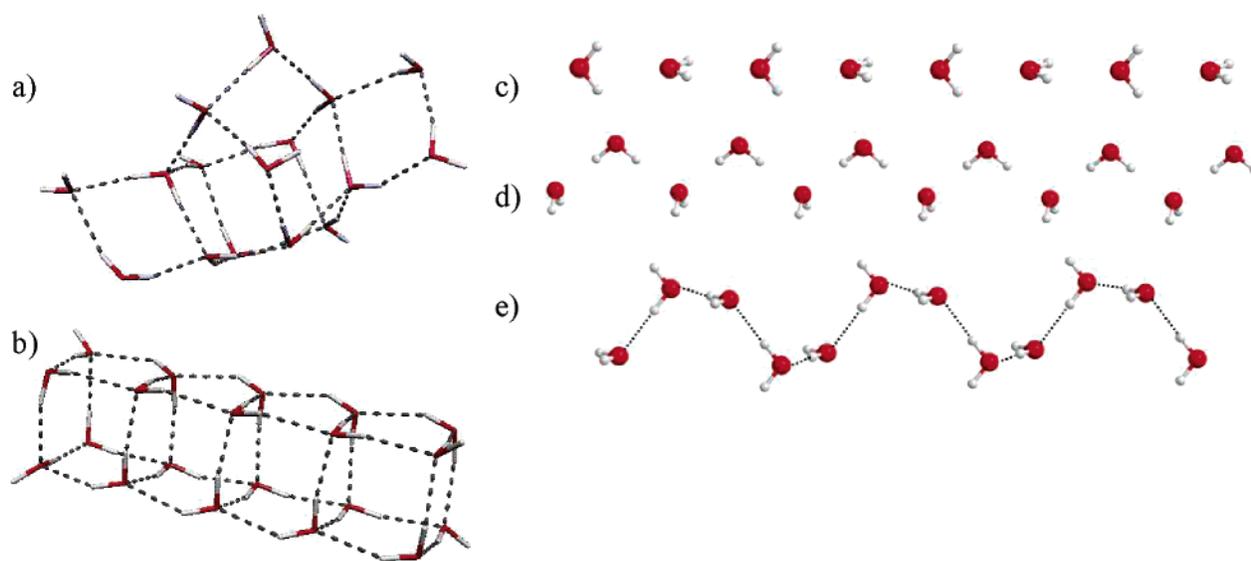
Figure 2. Representation of water aggregates [a), b)] and of sequentially H-bonded 12 water chains [c)−e)]: a), a water aggregate in an energy-minimized ice box with $n = 20$ water molecules; b), a water aggregate with $n = 16$ waters, as extracted from a Monte Carlo simulation on a water box with 64 molecules; c), a bifurcated arrangement; d), a transverse bifurcated; and e), a helical arrangement.

**Table 2.** RVS and SIBFA Interaction Energies (kcal/mol) in Four 12−20 Water Clusters

no. of waters	12		16		16 (MC)		20	
	SIBFA*	RVS	SIBFA*	RVS	SIBFA*	RVS	SIBFA*	RVS
$E_{MTP}*/E_c$	−167.6	−168.5	−230.9	−231.4	−179.5	−179.8	−293.2	−294.3
$E_{rep}*/E_{exch}$	151.9	151.4	207.9	207.5	149.8	149.9	263.6	263.2
E_1	−15.8	−17.1	−23.1	-−23.9	−29.7	−29.9	−30.6	−31.1
$E_{pol}*/E_{pol}$ RVS	−30.6	−34.7	−42.0	−47.8	−32.7	−35.5		
E_{pol}/E_{pol}	−41.3	−44.7	−56.5	−61.7	−44.1	−45.1	−71.3	−78.6
E_{ct}	−22.1	−23.1	−30.2	−31.3	−22.6	−23.1	−37.3	−39.4
ΔE(SIBFA)/ΔE(RVS)	−79.2	−80.1	−109.8	−110.4	−96.4	−94.8	−139.2	−139.1

extracted from an ongoing Monte Carlo (MC) simulation performed on a water box of $n = 64$ molecules (Figure 2b). Thus we wished to evaluate not only the overall accuracy of ΔE(SIBFA) as compared to ΔE(RVS) but also the extent of related agreements of the individual contributions. The results are reported in Table 2. A striking feature of the three cubic arrangements relates to E_{pol}, whose numerical values outweigh those of the summed first-order contribution E_1. This is because the large stabilizing values of E_{MTP*} are opposed by those of E_{rep}, a reflection of the shortening of

the O−O H-bonding distances (in the 2.72−2.90 Å range for $n = 20$) due to cooperativity. In fact, even E_{ct} has larger stabilizing values than E_1 in these three cubiclike structures. The weights of the second-order terms increase with respect to E_1 upon increasing n. E_{pol} also has a greater stabilizing role than E_1 in the MC structure, while E_{ct} is smaller in magnitude than it, a reflection of the relative lengthening of intermolecular O−O distances. For all four complexes, ΔE(SIBFA) reproduces very closely ΔE(RVS), the relative error being contained within 2%. The individual contribu-

Table 3. RVS and SIBFA Interaction Energies (kcal/mol) for Water Chains: Bifurcated Chain (BC), Transverse Hydrogen-Bonded Chain (t-HBC), and Longitudinal Hydrogen-Bonded Chain (l-HBC)

energies (kcal/mol)	E_c	E_{exch}	E_1	E_{pol}	E_{ct}	E_2	ΔE
HBC SIBFA	−81.6	54.0	−27.6	−18.2 (−14.0)	−9.5	−27.7	−55.4
HBC RVS	−81.2	54.3	−26.8	−17.3 (−14.5)	−9.8	−24.3	−53.1
t-HBC SIBFA	−58.8	29.9	−28.8	−9.0 (−7.2)	−3.6	−12.6	−41.4
t-HBC RVS	−53.5	27.3	−26.3	−9.8 (−8.2)	−3.5		−39.2
l-HBC SIBFA	−60.9	54.0	−6.9	−3.9 (−3.6)	−7.5	−11.4	−18.3
l-HBC RVS	−60.5	55.1	−5.4	−5.5 (−4.7)	−7.7		−17.8
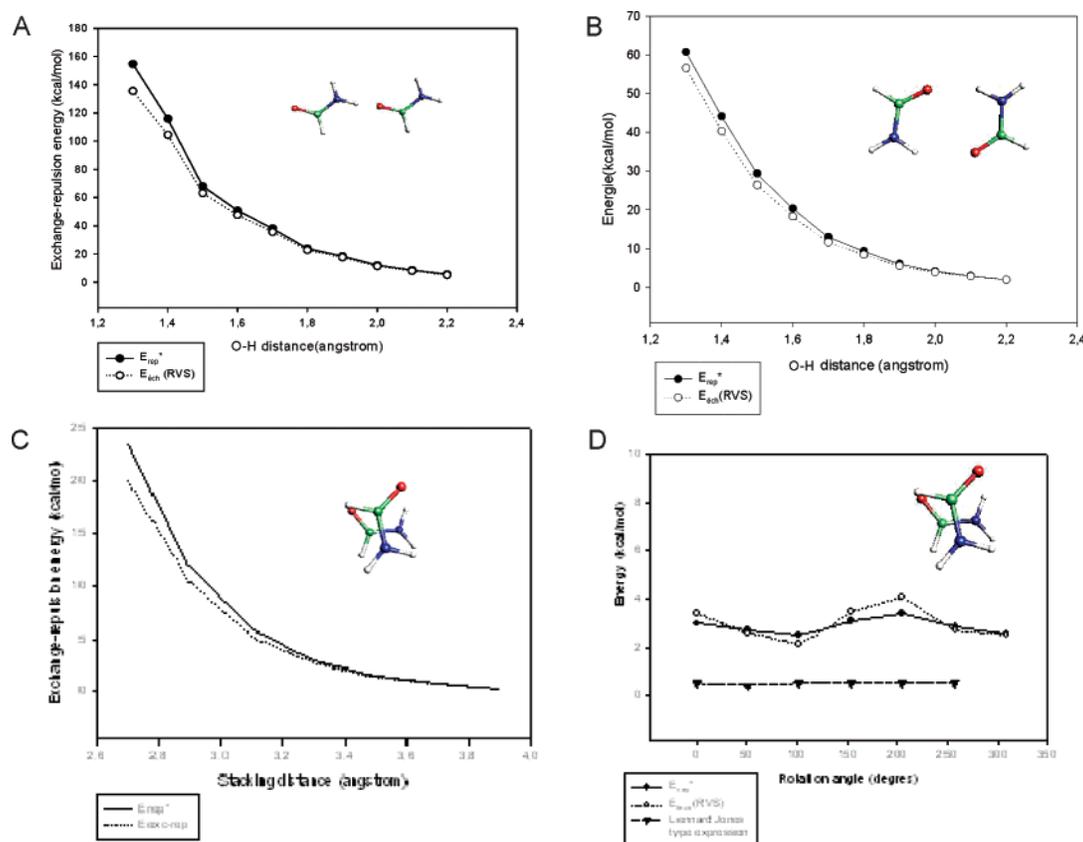
Figure 3. Formamide dimers. *In-plane H-bonded.* a) linear monodentate and b) bridged. Compared evolutions (in kcal/mol) of E_{exch}(RVS) and E_{rep}(SIBFA) as a function of the H−O distance. *Stacked.* Compared evolutions (in kcal/mol) of E_{exch}(RVS) and E_{rep}(SIBFA) as a function of c) interplanar separation and d) rotations around the z-axis at fixed interplanar separation of 3.3 Å.

tions, E_{MTP*} and E_{rep} within E_1 and E_{pol} and E_{ct} in second-order, match their RVS counterparts. As commented in previous papers[43,45,46] there is a good correspondence, on the one hand, between E_{pol*}(SIBFA) which is computed with the field due to the sole permanent multipoles and E_{pol}(RVS) and, on the other hand, E_{pol}(SIBFA) embodying the effect of induced dipoles on the field[43] and E_{pol}(KM) that results from the Kitaura−Morokuma (KM) analysis[44] (for details, see Piquemal et al., submitted for publication).

Water Chains. Further tests on the ability of polarizable potentials to account for nonadditive effects were put forth by Giese and York[27] and Chelli and Procacci.[47,48] They bore on two kinds of H-bonded chains of water molecules, namely bifurcated and transverse (Figure 2c,d). The possible issues of overpolarization (due to the absence of exchange-polarization in some potentials) as opposed to underpolarization (due to the lack of an explicit charge-transfer

contribution) were addressed by these authors. In another chain (Figure 2e), denoted as longitudinal helical, that was recently considered by Chelli and Procacci (Piquemal et al., *J. Phys. Chem. B*, in press), each nonterminal water acts simultaneously as a single H-bond donor and as a single H-bond acceptor (Figure 3c). Such a complex was designed in order to amplify the polarization response. We have recently evaluated the ability of both SIBFA and two Chemical Potential Equalization procedures designed by these authors[48] to give correct E_{pol} values from QC calculations as well as for the average water dipole moment in these chains (Piquemal et al., submitted for publication). As a continuation of this work, we give in Table 3 the results of SIBFA versus RVS analyses on these three dodecamer chains. As in ref 27, O−O H-bond distances are set to 2.97 Å and $n = 12$ water molecules. The analyses were also done at O−O H-bond distances of 2.48 and 3.50 Å and with
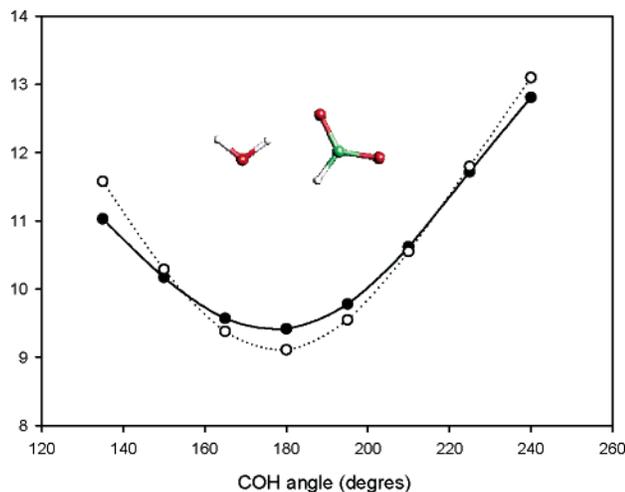
SIBFA Intermolecular Potential

J. Chem. Theory Comput., Vol. 3, No. 3, 2007 **831**

Figure 4. Compared in-plane angular evolutions (in kcal/mol) of E_{exch}(RVS) and E_{rep}(SIBFA) (dashed line) in a formate−water complex, as a function of the $\theta = C-O-H$ angle. The H−O distance is fixed at 1.8 Å.

shorter chains (unpublished). The results from Table 3 confirm the ability of SIBFA to correctly reproduce ΔE(RVS) and its individual contributions in these three chains.

CEP 4-31G(2d) vs DZVP2: Pseudopotential versus All Electrons Basis Set. The same agreement with ab initio is obtained at both level of basis set. Nevertheless, a difference can be noticed concerning the repartition of the penetration correction. While its monopole−quadrupole component is negligible with the CEP 4-31G(2d) basis set, this is no longer true for the full electron calculation where the correction monopole−quadrupole correction is required to obtain the same level of agreement with RVS computations (not shown).

Other H-Bonded Complexes. Formamide Dimer. The study by ab initio SCF methods, of the formamide dimer as a model for the H-bond between peptide units, was pioneered by Dreyfus and Pullman.[49] Such a complex had also lent itself to a study by one of the very first energy-decomposition approaches. Figure 3a,b bears on the linear and on the bridged formamide dimers, respectively. They represent the evolution of E_{rep}(SIBFA) compared to E_{exch}(RVS) as a function of the N−O distance of approach. These two figures illustrate that the parallelism between ΔE(SIBFA) and ΔE(RVS) reflects that between their individual first-order contributions as well as (not shown) second-order ones.

The results concerning the bridged formamide dimer are given in Table 1 of the Supporting Information.

Formate−Water. In order to evaluate the extent to which E_{rep} can account for the anisotropy features of E_{exch}, we have considered a complex between formate and a water molecule acting as an H-bond donor. In this complex, the distance between the donated H and one anionic O is fixed at 1.8 Å, the O−H−O angle is fixed at 180°, and stepwise in-plane variations of the $\theta = C-O-H$ angle are done (see Figure 4). This figure shows both E_{rep}(SIBFA) and E_{exch}(RVS) to have a marked angular behavior, with a maximum at θ at approximately 120°, E_{rep}(SIBFA) having a shape that parallels that of E_{exch}. *Such an angular behavior can only obtain thanks to the explicit introduction of localized lone pairs in*

SIBFA. By contrast, a simplified representation of E_{rep} under the form of atom−atom terms with $1/R^{12}$ dependence gives rise to a flat behavior (not shown). This is explained by the fact in the whole zone of water variations away from the other anionic O, the closest water−formate distance is that between the donated water H and the acceptor anionic O. Such a distance is constant, while, due to the $1/R^{12}$ dependence, the other atoms on both monomers affect only negligibly the behavior of E_{rep} with such a representation.

In Table 2 of the Supporting Information are reported the RVS and SIBFA interaction energies concerning the bridged formate−water complex at equilibrium distance ($d_{O-H} = 1.9$ Å), confirming again the agreement of SIBFA and RVS results at the level of both the total energies and their individual contributions.

Methylammonium−Water. The methylammonium−water complex was previously investigated in three distinct arrangements, the cation approaching water along the external bisector of the HOH angle: (a) in the prolongation of one NH bond; (b) through the external bisector of one HNH angle; and (c) in the prolongation of the CN bond.[21] Concerning the first-order contributions, however, E_1(SIBFA) was then identified to E_1(RVS) without further identification of E_{MTP*} to E_c and of E_{rep} to E_{exch}. The results with the present refinements are reported as Supporting Information (Table 3) concerning complexes $a-c$ at their optimized O−N distances. Good agreements are noticed for the three structures.

Formate−Methylammonium. The formate−methylammonium complex had, similarly, been investigated in ref 21 in two binding modes: (a) bidentate (denoted as 'B'), in which two ammonium protons interact each with one anionic oxygen, the two H−O distances being equal and the HNH plane being coplanar to the formate plane; and (b) mono-dentate (denoted as 'M'), in which one ammonium H binds externally to one anionic O. The N−H−O angle is 180, one HNH plane involving this H is coplanar with the formate plane, and similar to the formate−water complex, stepwise 15 deg variations are done on the $\theta = H-O-C$ angle. Results for the bidentate complex are reported as Supporting Information (see Table 4) at the optimized O−N distances of 'B' and 'M' complexes. A close agreement is found between SIBFA and RVS. Figure 5 confirms this point and shows the importance of a good description of the anisotropy of the exchange-repulsion in the rotations performed in the monodentate mode.

Stacked Formamide Complex. An ubiquitous determinant in molecular recognition concerns stacking interactions with aromatic or conjugated groups. Numerous examples are provided by structural biology, supramolecular chemistry, and solid-state X-ray crystallography.[50−52] An important issue relates to the computation of the van der Waals contribution to the total binding energy, since it is a major contributor to stabilization, yet its accurate evaluation requests beyond-HF calculations and the use of very extended basis sets. We concentrate here on a model complex of two stacked formamide molecules (see Figure 3c,d). This complex has been previously investigated by Sponer and Hobza.[53,54] These

Table 4. CSOV/B3LYP and SIBFA* Energies for Selected Dimers and Water Cluster: aug-cc-pVTZ Basis Set[a]

energies (kcal/mol)	E_c	E_{exch}	E_1	E_{pol}	E_{ct}	E_2	ΔE
linear dimer	−7.68	6.21	−1.44	−1.37	−1.61	−2.98	−4.42
(CSOV)	−7.78	6.30	−1.47	−1.31	−1.58	−2.90	−4.40
cyclic dimer	−5.27	2.90	−2.37	−0.30	−0.45	−0.75	−3.12
(CSOV)	−4.83	2.69	−2.15	−0.42	−0.56	−0.98	−3.13
bifurcated dimer	−4.32	2.30	−2.02	−0.28	−0.44	−0.73	−2.75
(CSOV)	−4.34	2.67	−1.67	−0.38	−0.41	−0.79	−2.46
H−H dimer (2.2)	2.06	0.26	2.32	−0.08	0.00	−0.08	2.25
(CSOV)	2.1	0.26	2.36	−0.23	−0.27	−0.50	1.87
H−H dimer (2.5)	1.48	0.10	1.58	−0.04	0.00	−0.04	1.38
(CSOV)	1.52	0.10	1.63	−0.12	−0.12	−0.24	1.39
O−O dimer (2.8)	1.86	2.23	4.10	−0.26	−0.15	−0.42	3.68
(CSOV)	1.53	2.57	4.10	-0.40	0.00	−0.40	3.68
cluster 16 H$_2$O	−184.0	149.2 (154.5)	−34.8 (−29.5)	−45.9	−41.0	−86.9	−121.7 (−116.3)
ab initio	−186.4	166.5	−19.8	−45.1	NC	NC	−114.0
cluster 20 H$_2$O	−298.2	262.1 (272.3)	−36.0 (−26.1)	−72.3	−75.8	−148.1	−184.1 (−174.0)
ab initio	−309.4	292.2	−17.2	−78.6	NC	NC	−168.1

[a] The values of the ab initio polarization energies are given at the HF/CEP 4-31G(2d) level. The ΔE (DFT) values are BSSE corrected. For the exchange-repulsion contribution, results given in parentheses correspond to SIBFA calculations performed using lone pairs positions as described in ref 4.
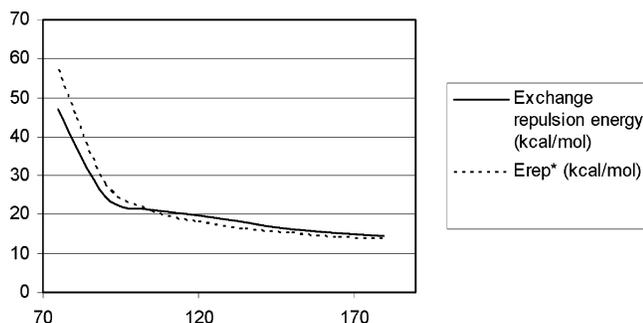

Figure 5. Compared in-plane angular evolutions (in kcal/mol) of E_{exch}(RVS) and E_{rep}(SIBFA) in a formate−monomethyl-ammonium complex, as a function of the $\theta = $ C−O−H angle. The H−O distance is fixed at 1.7 Å.

authors have evaluated the effects on the binding energies of correlation (at both MP2 and CCSD(T) levels) and of different basis sets, up to the cc- and aug-cc-pVDZ/cc-pVDZ ones. The two extreme arrangements, parallel as well as antiparallel, were considered. These studies were undertaken in view of subsequent detailed analyses of stacking interactions in nucleic acid bases, which, similar to formamide, encompass C=O, C−N, and −NH functional groups. This was exemplified in a recent study of stacked cytosine dimers.[55] In the present study, and prior to similar subsequent extensions, we wish to first evaluate the extent to which each of the individual SIBFA contributions can reproduce the distance and the orientation dependencies of its HF counterpart. For that purpose, we start from a position where each monomer overlaps maximally with the other. This is done by giving to each atom of the second monomer the same x and y coordinates as the corresponding atom of the first. ΔE(SIBFA) is then optimized by varying the z coordinate of the second formamide kept parallel to the first. At the optimized value of z (3.3 Å), clockwise rotation of the second formamide is done around the z-axis. Such arrangements are chosen in order to once more critically evaluate E_{rep}(SIBFA). Indeed, the onset of bond−bond, bond−lone pair, and lone

pair−lone pair interactions is maximized, so that this contribution could only reproduce the behavior and numerical values of E_{exch}(RVS) if all three components are properly and consistently formulated and weighted. An additional requisite is that the π lone pairs on the C, N, and O atoms be properly represented. For that purpose, we have proposed in ref 26, a representation of the π system as bent sp hybrids.[25] Their localizations with respect to the atom bearing them as well as their partial occupation numbers were determined by using a Zn(II) cation as a probe over the formamide plane above the C, N, and O atoms and performing parallel SIBFA and RVS computations to optimize the fit of E_{rep}(SIBFA) to E_{exch}(RVS). Here, we justify this choice by drawing electron density maps (see Figure 1 of the Supporting Information). In these figures, the outermost contour corresponds to a density of 0.001 au which has been shown to be a measure of the van der Waals radius[56] in a way that can be directly measured (in Å). As can be seen, the density expansion is greater in the molecular plane than in the plane perpendicular to it and containing the π system. Figure 3c,d compares the evolutions of E_{rep}(SIBFA) and E_{exch}(RVS) as a function of, respectively, the interplanar distance z and of the rotation angle around the z-axis.

E_{rep}(SIBFA) shows the same directional features as E_{exch}(RVS) but slightly underestimates (see Figure 3d). By contrast, a $1/R^{12}$ expression gives rise to a flat behavior.

As can be seen in Figure 6, the agreement of E_{MTP*}(SIBFA), here at the DZVP2 level, can be further improved upon inclusion of E_{pen} compared to previous calculations at the same level.[17] Again, the extra monopole−quadrupole correction is required to obtain such results.

To conclude on the stacked formamide dimers, a very good agreement can be evidenced upon comparing the evolution of ΔE(SIBFA) to that of ΔE(RVS) for rotations around the z-axis (see Figure 7). This demonstrates that the anisotropic character of SIBFA can faithfully mirror that of the ab initio computations for stacked complexes.
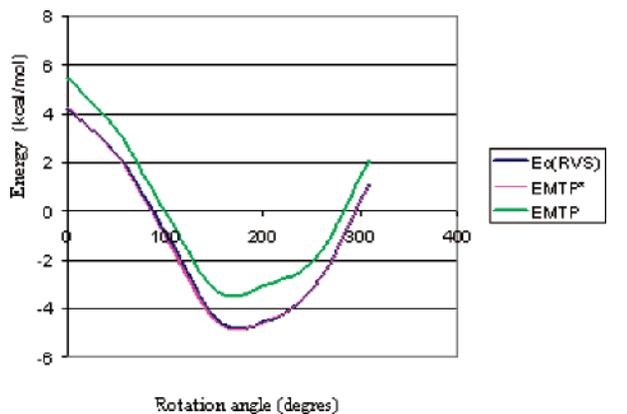
SIBFA Intermolecular Potential

J. Chem. Theory Comput., Vol. 3, No. 3, 2007 **833**

Figure 6. Stacked formamide dimer. DZVP2 computations. Variations (in kcal/mol) of E_c(RVS) and E_{MTP^*}(SIBFA) as a function of angle of rotation around the *z*-axis.

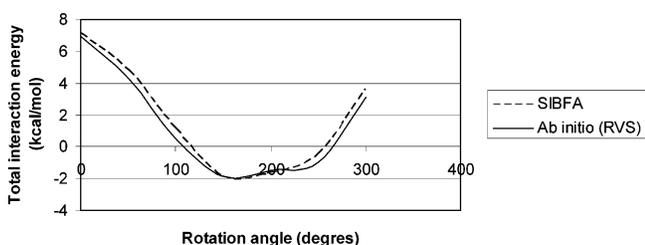**Figure 7.** Stacked formamide dimer. Compared evolutions (in kcal/mol) of ΔE(SIBFA) and ΔE(RVS) for rotations around the *z*-axis.

Inclusion of Electron Correlation. Our approach to handling the effects of electron correlations in APMM procedures is the following: (a) concerning E_{MTP^*} and E_{pol}, by deriving distributed multipoles and polarizabilities from DFT computations, and (b) concerning E_{rep} and E_{ct}, by recalibrating the multiplicative constants so that these contributions reproduce their counterparts from CSOV[13,18] analyses of bimolecular complexes computed at the DFT level. There remains the problem of dispersion since present functionals cannot provide the energy gain due to dispersion effects, namely the actual van der Waals term.[57] Even standard MP2 computations need to be augmented by CCSD(T) computations.[58,59] Initially regarding SIBFA, the calibration of the dispersion contribution E_{disp} was performed in a "HF + dispersion" approximation so that $\Delta E+E_{disp}$ matches in model bimolecular complexes high-level MP2, CCSDT, or Symmetry-Adapted Perturbation Analyses (SAPT)[12] interaction energies. In this work, we propose to use a "DFT+E_{disp}" approach.

DFT Computations. Toward reproduction of DFT calculations, it is essential to first evaluate the extent to which SIBFA can reproduce first the individual components of CSOV not only in the five water dimers presented above but also in more difficult configurations. Indeed, in 2002, van Duijneveldt et al.[28] have reinvestigated using molecular mechanics ten water dimers originally due to Tschumper et al.[60] and occurring as stationary points on the water dimer surface obtained at high level QC (CCSD(T)/large basis set + diffuse functions). Energy-decomposition analysis enabled the evaluation of the relative merits of several polarizable

potentials concerning the representation of the individual QC energy contributions. Recently, some of us proposed a new-generation force field based on density fitting termed GEM (Gaussian Electrostatic Model)[4,65] able to address accurately such a difficult issue. In the present study, these ten complexes are reinvestigated in light of the refinements to the SIBFA first-order contributions that now include correlated multipoles and polarizabilities. The choice of aug-cc-pVTZ is consistent with both van Duijneveldt et al. and the GEM studies. It also enables the evaluation of the SIBFA potential as compared to reference calculations performed using large basis sets with diffuse functions.

The results for the first five dimers (the linear configuration being also included in Tschumper's et al.[60] training set) are reported in Table 4 and appear in good agreement with the CSOV results. Figures 8 and 9 show the behavior of the energy components for scans of the intermolecular O−H distance in the linear dimer configuration. A good agreement of all the different SIBFA components is obtained with their ab initio counterparts even below the equilibrium position ($d_{O\cdots H}$ = 1.95 Å). This also confirms the GEM-0 results[4] that showed the capabilities of the charge-transfer expression to account for the QC charge-transfer energy gain observed upon going from HF to DFT.[13]

At such a difficult level including diffuse functions, more configurations are needed to test the potential energy surface. Therefore, we have performed calculations following van Duijneveldt et al. [28] recommendations. As can be seen in Figure 10 which bears on all 14 dimer configurations tested in this work a good correlation is obtained with ab initio (0.989). Compared to CSOV, the mean error of the SIBFA total interaction energy on the additional nine dimers[28,60] is 0.22 kcal/mol. This error appears larger than the one observed with the more sophisticated GEM-0 force field.[4] Nevertheless, it demonstrated that SIBFA can meet the requirements suggested by van Duijneveldt et al.[28] concerning molecular dynamics potentials, namely that they should be able to reproduce ab initio total interaction energies with errors about 1 kJ in order to stay below kT at room temperature. It is interesting to quote that most of the errors are, following the numbering of ref 28, concentrated in dimers 6−10 (the worst agreement being on dimer 9 with an error of 0.6 kcal/mol) which are only weakly attractive and highly stabilized by the introduction of diffuse functions at the ab initio level.[60] We have also tested this correlated SIBFA potential on water clusters (16 and 20 molecules, see Table 4) for which ab initio data are available from ref 4. While the DFT-derived E_{MTP^*} shows robustness compared to CSOV, a deviation is observed for E_{rep^*} with an error about 10% compared to reference data. To understand the origin of these discrepancies we have changed the positions of the water lone pairs (initially derived at the HF level) according to the position of the centroids of the Boys localized orbitals[61] obtained at the B3LYP/aug-cc-pVTZ that were retained in the GEM-0 approach.[4] Indeed, the errors are decreasing (see Table 4) when a consistent location is chosen. It is important to point out that we had also recomputed the exchange-repulsion energies of the 14 water dimers with this new location and noticed no improvement. This underlines the difficulty of a
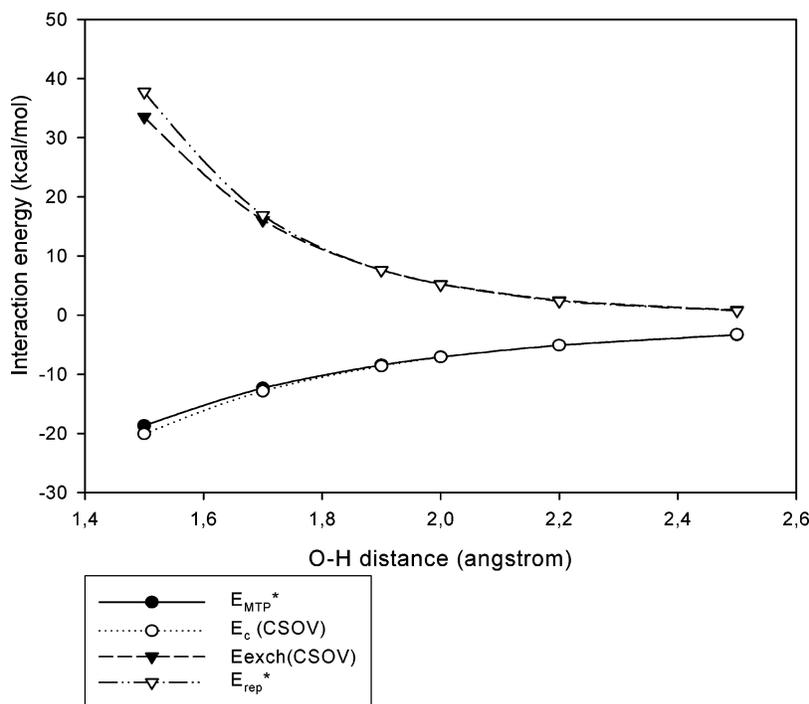
Figure 8. Linear water dimer. Variations (in kcal/mol), as a function of intermolecular distance, of E_c(RVS) and E_{exch}(RVS) calculated at the B3LYP/aug-cc-pVTZ/CSOV level and corresponding variations from modified SIBFA contributions E_{MTP*} and E_{rep*}.
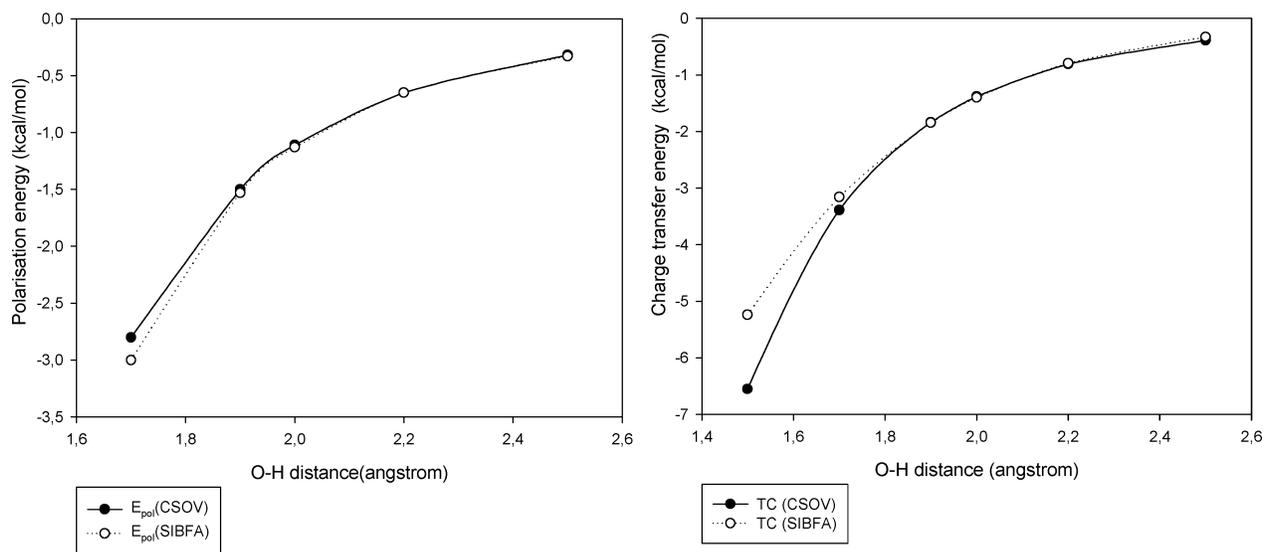

Figure 9. Linear water dimer. Variations (in kcal/mol), as a function of intermolecular distance, of E_{pol}(RVS) and E_{ct}(RVS) calculated at the B3LYP/aug-cc-pVTZ/CSOV level and corresponding variations from modified SIBFA contributions E_{pol} and E_{ct}.

choice of a limited training set of geometries. To conclude on the total energies, we have observed a good agreement with ab initio in the two structures even if some error in compensation between components occurs.

Inclusion of Dispersion: DFT+ disp Approximation vs CCSD(T). We present here some preliminary results using a "DFT+dispersion" approximation. We have modified by a factor of 70% the multiplication coefficient of the dispersion energy contribution, E_{disp}, as formulated in ref 43 so that it matches the values given in ref 28 for the ten water dimers computed at the CCSD(T)/aug-cc-pVTZ level. We observed an average error of 0.22 kcal/mol on the

total training set. Figure 11 reports, concerning the linear water dimer, the evolution of the SIBFA(B3LYP/aug-cc-pVTZ)+E_{disp} interaction energies along with those of the CCSD(T)/aug-cc-pVTZ ones. Starting from an O−H distance of 1.85−2.5 Å, we observed that SIBFA reproduces CCSD(T) computations with an average error limited to 0.2 kcal (0.1 kcal at the equilibrium geometry, $d_{OH} = 1.95$ Å). More detailed explorations of these approximations, supplemented by extensive CCSD(T) calculations, are currently under investigation and will be reported subsequently.

Perspectives. The development of polarizable molecular potentials is the object of intense efforts, as attested since
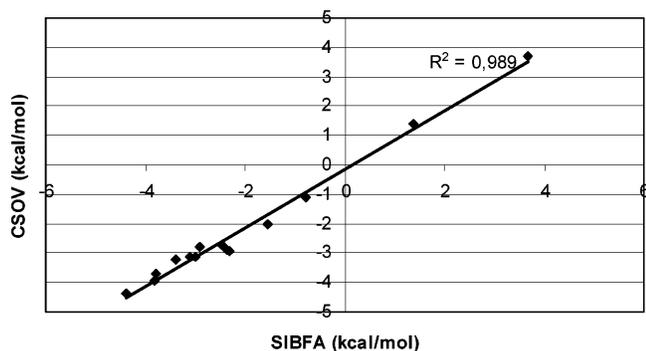
SIBFA Intermolecular Potential

J. Chem. Theory Comput., Vol. 3, No. 3, 2007 **835**

Figure 10. Correlation line between SIBFA and CSOV total interaction energies for fourteen water dimers.
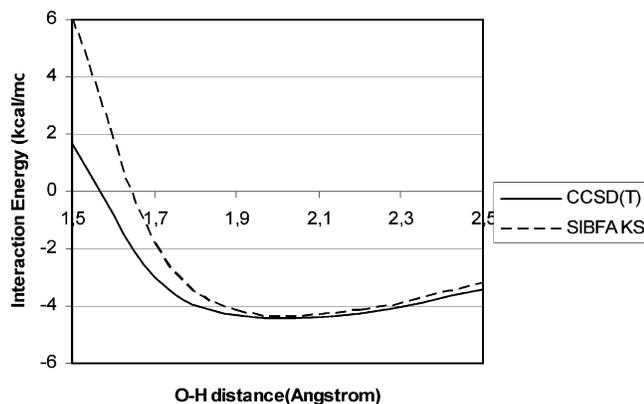

Figure 11. Linear water dimer. Evolution of the SIBFA-(B3LYP/aug-cc-pVTZ)+E_{disp} interaction energies along with that of the CCSD(T)/aug-cc-pVTZ ones.

2001 by review papers published on a nearly yearly basis.[2,62,63] We indicate here some ongoing developments and enrichments of the SIBFA potential.

From APMM toward APMD. The analytical gradients of most of the energy contributions have been coded and checked, while the coding of the remaining gradients is underway. This has enabled preliminary molecular dynamics (MD) simulations to be done in the framework of the SIBFA potential. Extension of this approach including Particle Mesh Ewald[6,67−70] procedures to handle long-range interactions is underway. This should allow for a further advance, namely from APMM toward APMD.

Toward Third-Generation Molecular Mechanics Potentials. A GEM methodology (Gaussian Electrostatic Model)[4,66] was recently developed and is able to provide total intermolecular interactions energies[4] and to handle long-range electrostatic thanks to a generalized PME procedure.[66] In GEM, fitted Gaussian densities are derived from first-order density matrices and used to compute in the framework of quantum chemistry the intermolecular Coulomb and overlap integrals, the latter then enabling an accurate evaluation of the exchange-repulsion interactions. In a series of test cases including water dimers and oligomers, E_c and E_{exch} from CSOV analysis were reproduced by their GEM counterpart with relative errors <1% and a considerable time gain compared to ab initio. It was, furthermore, shown by Piquemal et al.[4] that E_{pol} and E_{ct} could be computed *in the framework of the SIBFA procedure* upon resorting to the

GEM potentials and screened fields as an alternative to those derived from the distributed multipoles. Thus, a molecular mechanics procedure initially formulated and calibrated on the basis of QC can now resort concerning its electrostatic and overlap-depending terms to analytical integrals formulated in the context of ab initio QC. Such a methodology can be considered as a representative of future third-generation molecular mechanics potentials. The recent integration of QM and GEM by Cisneros et al.[70] constitutes an incentive for the next level of integration, that is, toward a QM/GEM/SIBFA procedure.

Conclusions

Along with the developments published in refs 17 and 26, we have elaborated on further refinements of the two first-order contributions, E_{MTP} and E_{rep}, of the SIBFA procedure. This enables term-to-term identifications of both first- and second-order contributions to their counterparts from RVS/KM/CSOV analyses of the HF/DFT intermolecular interaction energy. ΔE(SIBFA) has been validated by comparisons with ΔE(RVS) in several bi- and multimolecular H-bonded complexes, in arrangements significantly different from those used in the calibration. These tests were carried out on each of the four SIBFA contributions against their RVS counterparts. They bore on water clusters and water chains, formamide dimers, and complexes involving one or two ionic molecules. A striking result found in the water clusters related to the predominant weight of the second-order contributions E_{pol} and E_{ct}, particularly for the cubiclike arrangements and for the larger values of *n*. This was fully supported by the RVS analysis, and very close numerical agreements were found for both the total interaction energies and each of its four contributions. Close numerical agreements were also found in three dodecameric water chains which had been originally designed to probe the nonadditivity response of E_{pol} as a function of the H-bonding geometry. The anisotropic behavior of ΔE(SIBFA) and of its contributions were also probed in two illustrative examples namely a formate−water complex and a stacked formamide dimer. In the latter the number of bond−bond, bond−lone pair, and lone pair−lone pair interactions is maximized, and all three terms need to be properly expressed and weighted. The correct reproduction of E_{exch} by E_{rep} throughout the angular rotations proved the correctness of the formulation of E_{rep}.

We showed that it was also possible to account for the effects of correlation on ΔE, by deriving the distributed multipoles and polarizabilities obtained at the DFT level. This approach was validated by several tests on 14 water dimers as well as on two water aggregates. A final extension consisted of the reintroduction of the 'dispersion' contribution. With correlated multipoles and polarizabilities, it was possible to only refit E_{disp} so that it reproduces the difference between CCSD(T) and DFT computations at equilibrium distance. This procedure is being presently generalized and adapted to the GEM[4,66] procedure (Piquemal, et al., manuscript in preparation).

As underlined in refs 2, 4, and 26 the present results illustrate the necessity for APMM procedures to be separate, anisotropic, nonadditive, and transferable. Each of these

836 *J. Chem. Theory Comput., Vol. 3, No. 3, 2007*

Piquemal et al.

facets was addressed here. With these refinements, the SIBFA procedure was recently applied in studies of inhibitor binding to the Zn-metalloenzyme phosphomannoisomerase[71] and to the C-terminal Zn-finger of the HIV-1 nucleocapsid (Miller-Jenkins et al., submitted for publication). Extensions are underway to drug binding to kinases.

Supporting Information Available: RVS and SIBFA interaction energies (Tables 1−4) and density maps (Figure 1). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Antony, J.; Piquemal, J.-P.; Gresh, N. *J. Comput. Chem.* **2005**, *26*, 1131.

(2) Gresh, N. *Curr. Pharm. Des.* **2006**, *12*, 2121.

(3) Millot, C.; Stone, A. *J. Mol. Phys.* **1992**, *77*, 439.

(4) Piquemal, J.-P.; Cisneros, G. A.; Reinhardt, P.; Gresh, N.; Darden, T. A. *J. Chem. Phys.* **2006**, *124*, 104101.

(5) Ren, P.; Ponder, J. W. *J. Phys. Chem. B* **2003**, *107*, 5933.

(6) Piquemal, J.-P.; Perera, L.; Cisneros, G. A.; Ren, P.; Pedersen, L. G; Darden, T. A. *J. Chem. Phys.* **2006**, *125*, 054511.

(7) Gordon, M. S.; Freitag, M; Bandyopadhyay, P.; Jensen, J. H.; Kairys, V.; Stevens, W. J. *J. Phys. Chem. A* **2001**, *105*, 293.

(8) Hagberg, D.; Karlstrom, G.; Roos, B. O.; Gagliardi, L. *J. Am. Chem. Soc.* **2005**, *127*, 14250.

(9) Stone, A. J. *J. Chem. Theory Comput.* **2005**, *1*, 1128.

(10) Vigne-Maeder, F.; Claverie, P. *J. Chem. Phys.* **1988**, *88*, 4934.

(11) Stevens, W. J.; Fink, W. H. *Chem. Phys. Lett.* **1987**, *139*, 15.

(12) Jeziorski, S. B.; Moszynski, R.; Szalewicz, K. *Chem. Rev.* **1994**, *94*, 1887.

(13) Piquemal, J.-P.; Marquez, A.; Parisel, O.; Giessner-Prettre, C. *J. Comput. Chem.* **2005**, *26*, 1052.

(14) Cisneros, G. A.; Piquemal, J.-P.; Darden, T. A. *J. Chem. Phys.* **2005**, *123*, 044109.

(15) te Velde, G.; Bickelhaupt, F. M.; Baerends, E. J.; Fonseca Guerra, C.; van Gisbergen, S. J. A; Snijders, J. G.; Ziegler, T. *J. Comput. Chem.* **2001**, *22*, 931.

(16) Freitag, M. A.; Gordon, M. S.; Jensen, J. H.; Stevens, W. J. *J. Chem. Phys.* **2003**, *112*, 7300.

(17) Piquemal, J.-P.; Gresh, N.; Giessner-Prettre, C. J. *Phys. Chem. A.* **2003**, *107*, 10353.

(18) Bagus, P. S.; Illas, F. *J. Chem. Phys.* **1992**, *96*, 8962.

(19) Claverie, P. In *Intermolecular interactions: from diatomics to biopolymers*; Pullman, B., Ed.; Wiley: 1978; pp 69−305.

(20) Hermida-Ramon, J.; Brdarski, S.; Karlstrom, G.; Berg, U. *J. Comput. Chem.* **2003**, *24*, 161.

(21) Gresh, N.; Leboeuf, M.; Salahub, D. R. In *Modeling the Hydrogen Bond*; ACS Symposium Series, 1994; Vol. 569, pp 82−112.

(22) Gresh, N. *J. Comput. Chem.* **1995**, *16*, 856.

(23) Murrell, J. N.; Teixeiras-Dias, J. J. N. *Mol. Phys.* **1970**, *19*, 521.

(24) Gresh, N.; Claverie, P.; Pullman, A. *Int. J. Quantum Chem.* **1986**, *29*, 101.

(25) Piquemal, J.-P. Ph.D. Thesis, Université Pierre et Marie Curie, Paris, France, 2004.

(26) Gresh, N.; Piquemal, J.-P.; Krauss, M. *J. Comput. Chem.* **2005**, *26*, 1052.

(27) Giese, T. J.; York, D. M. *J. Chem. Phys.* **2005**, *123*, 164108.

(28) van Duijneveldt-van de Ridjt, J. G. C. M.; Mooij, W. T. M.; van Duijneveldt, F. B. *Phys. Chem. Chem. Phys.* **2003**, *5*, 1169.

(29) Gresh, N.; Tiraboschi, G.; Salahub, D. R. *Biopolymers* **1998**, *45*, 405.

(30) Gresh, N.; Kafafi, S. A.; Truchon, J. F.; Salahub, D. R. *J. Comput. Chem.* **2004**, *25*, 823.

(31) Guo, H.; Gresh, N.; Roques, B. P.; Salahub, D. R. *J. Phys. Chem. B* **2000**, *104*, 9746.

(32) Gresh, N.; Guo, H.; Salahub, D. R.; Roques, B. P.; Kafafi, S. A. *J. Am. Chem. Soc.* **1999**, *121*, 7885.

(33) Gordon M. S.; Schmidt M. W. Advances in electronic structure theory: GAMESS a decade later. In *Theory and Applications of Computational Chemistry, the first forty years*; Dykstra, C. E., Frenking, G., Kim, K. S., Scuseria, G. E., Eds.; Elsevier: Amsterdam, 2005.

(34) Stevens, W. J.; Basch, H.; Krauss, M. *J. Chem. Phys,* **1984**, *81*, 6026.

(35) Godbout, N.; Salahub, D. R.; Andzelm, J.; Wimmer, E. *Can. J. Chem.* **1992**, *70*, 560.

(36) Piquemal, J.-P.; Williams-Hubbard, B.; Fey, N.; Deeth, R. J.; Gresh, N.; Giessner-Prettre, C. *J. Comput. Chem.* **2003**, *24*, 1963.

(37) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.

(38) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

(39) Dunning, T. H., Jr. *J. Chem. Phys.*, **1989**, *90*, 1007.

(40) Garmer, D. R.; Stevens, W. J. *J. Chem. Phys.* **1989**, *93*, 8263.

(41) *JAGUAR, Version 6.0*; Schrödinger, LLC: New York, 2005.

(42) Claverie, P. Ph.D. Thesis, Université Pierre et Marie Curie, Paris, France, 1973.

(43) Gresh, N. *J. Phys. Chem. A* **1997**, 101, 8680.

(44) Kitaura, K.; Morokuma, K. *Int. J. Quantum Chem.* **1976**, *10*, 325.

(45) Masella, M.; Gresh, N.; Flament, J.-P. *J. Chem. Soc. Faraday Trans.* **1998**, *94*, 2745.

SIBFA Intermolecular Potential

J. Chem. Theory Comput., Vol. 3, No. 3, 2007 **837**

(46) Tiraboschi, G.; Gresh, N.; Giessner-Prettre, C.; Pedersen, L. G.; Deerfield, D. W. *J. Comput. Chem.* **2000**, *21*, 1011.

(47) Chelli, R.; Schettino, V.; Procacci, P. *J. Chem. Phys.* **2005**, *122*, 234107.

(48) Chelli, R.; Procacci, P. *J. Chem. Phys.* **2002**, *117*, 9175.

(49) Dreyfus, M.; Pullman, A. *Theor. Chim. Acta* **1970**, *19*, 20.

(50) Burley, S. K.; Petsko, G. A. *Science* **1985**, *229*, 23.

(51) Fyfe, M. C. T.; Stoddart, J. F. *Acc. Chem. Rev.* **1997**, *10*, 3393.

(52) Claessens, C. G.; Stoddart, J. F. *J. Phys. Org. Chem.* **1997**, *10*, 254.

(53) Hobza, P.; Sponer, J. *J. Mol. Struct. (THEOCHEM)* **1996**, *388*, 115.

(54) Sponer, J.; Hobza, P. *Chem. Phys. Lett.* **1997**, *267*, 263.

(55) Jurecka, P.; Sponer, J.; Hobza, P. *J. Phys. Chem. B* **2004**, *108*, 5466.

(56) Gillespie, R. J.; Popelier, L. A. In *Chemical bonding and molecular geometry*; Oxford University Press: New York, 2001.

(57) Guo, H.; Salahub, D. R. *Angew. Chem.* **1998**, *37*, 2985.

(58) Jurecka, P.; Sponer, J.; Cerny, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985.

(59) Zierkiewicz, W.; Michalska, D.; Cerny, J.; Hobza, P. *Mol. Phys.* **2006**, *104*, 2317.

(60) Tschumper, G. S.; Leininger, M. L.; Hoffman, B. C.; Valeev, E. F.; Quack, M.; Schaefer, H. F., III. *J. Chem. Phys.* **2002**, *116*, 690.

(61) Foster, J. M.; Boys, S. F. *Rev. Mod. Phys.* **1960**, *32*, 300.

(62) Halgren, T. A.; Damm, W. *Curr. Opin. Struct. Biol.* **2001**, *11*, 236.

(63) Rick, S. W.; Stuart, S. J. *Rev. Comput. Chem.* **2002**, *18*, 89.

(64) Ponder, J. W.; Case, D. A. *Adv. Protein Chem.* **2003**, *66*, 27.

(65) McKerell, J. *J. Comput. Chem.* **2004**, *25*, 1584.

(66) Cisneros, G. A.; Piquemal, J.-P.; Darden, T. A. *J. Chem. Phys.* **2006**, *125*, 184101.

(67) Darden, T. A.; York, D. M.; Pedersen, L. G. *J. Chem. Phys.* **1993**, *98*, 10089.

(68) Essmann, U.; Perera, L.; Bertkowitz, M. L.; Darden, T. A.; Lee, H.; Pedersen, L. G. *J. Chem. Phys.* **1995**, *103*, 8577.

(69) Sagui, C.; Pedersen, L. G.; Darden, T. A. *J. Chem. Phys.* **2004**, *120*, 1630791.

(70) Cisneros, G. A.; Piquemal, J.-P.; Darden,T. A. *J. Phys. Chem. B* **2006**, *110*, 13682.

(71) Roux, C.; Gresh, N.; Perera, L.; Piquemal, J.-P.; Salmon, L. *J. Comput. Chem.* **2007**, *28*, 938.