Progress Towards Accurate Molecular Modeling of Metal Complexes Using Polarizable Force Fields

R. Chaudret[a,b], S. Ulmer[a,b], M-C van Severen[a,b], N. Gresh[c], O. Parisel[a,b], G. A. Cisneros[d], T. A. Darden[d] and J-P Piquemal*[a,b]

[a] *UPMC Univ Paris 06, UMR 7616, Laboratoire de Chimie Théorique, case courrier 137, 4 place Jussieu, F-75005, Paris, France.*

[b] *CNRS, UMR 7616, Laboratoire de Chimie Théorique, case courrier 137, 4 place Jussieu, F-75005, Paris, France.*

[c] *Laboratoire de Pharmacochimie Moléculaire et Cellulaire, U648 INSERM, UFR Biomédicale, Université Paris Descartes, 45, rue des Saints-Pères, 75006 Paris, France.*

[d] *Laboratory of Structural Biology, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina 27709 (USA)*

Abstract We present refinements of the SIBFA (Sum of Interaction Between Fragments ab initio) and GEM (Gaussian electrostatic Model) polarizable molecular mechanics procedure to represent the intermolecular interaction energies of metal cations. Improved forces fields for closed-shell, open-shell and heavy metals are discussed. Some perspectives towards a multiscale SIBFA-GEM approach using density fitting techniques are presented.

Keywords: metals, cations, molecular mechanics, density fitting, multiscale approaches.
PACS: 31.15.A, 31.15.B, 31.15.E, 31.15.xv

INTRODUCTION

Divalent cations such as Zn(II) are of fundamental importance in structural and molecular biology, acting both as a cofactor of numerous metalloenzymes and as a key structural element in the architecture of Zn-.finger and related proteins. It is also used in the construction of several supramolecular structures. Although *ab initio* quantum chemistry (QC) is the most accurate procedure for the computation of intermolecular interactions, it could not be applied presently to systematically investigate very large complexes (200 atoms) and perform detailed investigations of the potential energy hypersurfaces. QM/MM approaches at varying levels of theory are being used for metalloprotein simulations in which the cation-binding site is computed by QC while its periphery is computed by standard molecular mechanics (MM) approaches. Anisotropic polarizable molecular mechanics (APMM) methods [1] could be a viable alternative to QM/MM if the energy potential were of sufficient accuracy. In this contribution, we present some recent results obtained using the SIBFA (Sum of

CP1102, *Theory and Applications of Computational Chemistry – 2008*
edited by D.-Q. Wei and X.-J. Wang

Interaction Between Fragments ab initio)[1] and GEM (Gaussian electrostatic Model) [1, 2] polarizable molecular mechanics procedures. After a short methodological section, we will show the capabilities of such methods in the reproduction of quantum short-range effects within the electrostatic and exchange-repulsion energies in the case of transition and heavy metals. As focus will then be given to the reproduction of ab initio results on open-shell cations [3, 4] such as Cu(II), some preliminary results on Zn(II) using a multiscale SIBFA/GEM approach [5] dedicated to the treatment of metals are presented.

COMPUTATIONAL DETAILS

Formulation of the SIBFA Procedure.

The SIBFA intermolecular interaction energy is formulated as a sum of five contributions:

$$\Delta E_{int} = E_{MTP^*} + E_{rep^*} + E_{pol} + E_{ct} + E_{disp} + E_{LF} \quad (1)$$

denoting respectively the multipolar short-range penetration corrected electrostatic (E_{MTP^*}) [6], short-range repulsion (E_{rep^*})[7], polarization (E_{pol}), charge-transfer (E_{ct}), and dispersion (E_{disp}) contributions. The analytical forms of these contributions are given in the original papers and we only recall here their essential features (see reference 1 and references therein).

E_{MTP^*} is computed with multipoles (up to quadrupoles) that are distributed on the atoms and bond barycentre following a procedure developed by Vigné-Maeder and Claverie[8]. In its latest refinements [7], E_{MTP^*} has been augmented with an explicit penetration term, $E_{pen.}$. This was shown to afford for a closer match to the Coulomb contribution, $E_{C,}$ which obtains from energy-decompositions analyses of the ab initio intermolecular interaction energies. E_{rep} is formulated as a sum of bond-bond, bond-lone pair, and lone pair-lone pair interactions. An S^2/R representation has been used since 1994 (see reference [7] and references therein). **S** denotes an approximation of the overlap between localized MO's (LMO's) of the interacting partners. Hybridization is on chemical bonds as well as on the lone pairs. R is the distance between the LMO centroids. Following the E_{MTP^*} refinements with inclusion of the E_{pen} term, E_{rep} is augmented with an S^2/R^2 term [7].

In E_{pol} the polarizing field is computed with the same permanent multipoles as E_{MTP}. The field is screened by a Gaussian function of the distance between the two interacting centres. Such a screening embodies part of short-range effects including exchange-polarization (see for example [9]). The contribution of the induced dipoles to the field is computed by a self-consistent iterative procedure. Since 1991, the polarizabilities are tensors that are distributed on the bond barycenters and on the heteroatom lone pairs and are derived from the LMO's of the considered molecule or molecular fragment by a procedure due to Garmer and Stevens [10] which has been coded into our local code [11]. As such, both distributed multipoles and polarizabilities can be obtained from one ab initio computation performed on a

molecule or constitutive molecular fragment. Each molecular entity is stored in the SIBFA library of fragment and used for subsequent assembly of molecules or molecular complexes.

E_{ct} [12] is derived from the development of a formula due to Murrell et al. [13]. A coupling with electrostatics is present. That is, the ionization potential, I_A, of the electron donor on the one hand, and the electron affinity, A_M, and 'self-potential', V_M, of the electron acceptor on the other hand, are modified by the electrostatic potential that each undergoes in the complex. These include the effect of the induced dipoles along with those of the permanent multipoles, thereby introducing a coupling with polarization. Such modifications of I_A, A_M, and V_M were essential to account for the very strong anticooperative character of E_{ct} in polycoordinated complexes of divalent cations.

E_{disp} is computed as a sum of $1/R^6$, $1/R^8$, and $1/R^{10}$ terms. Directionality effects are accounted for by the introduction of lone-pairs under the form of fictitious atoms. An exchange-dispersion term was also introduced. For H-bonded complexes, E_{disp} was initially calibrated on the basis of Symmetry-Adapted Perturbation Theory (SAPT, [14]) energy-decomposition analyses.

E_{LF} is the ligand-field energy contribution [3, 4] that can be introduced when open-shell cation are present.

Formulation of the Gaussian Electrostatic Model: towards a multiscale SIBFA/GEM approach.

GEM [1, 2] uses the formalism of the variational density fitting method [15], an approach which is usually devoted to a fast approximation of the Coulomb interaction.

This method relies on the use of an auxiliary Gaussian basis set (ABS) to fit the molecular electron density obtained from a relaxed one-electron density matrix using a linear combination of atomic orbitals.

$$\tilde{\rho} = \sum_{k=1}^{N} xkk(r) \approx \rho = \sum_{\mu\nu} P_{\mu\nu}\phi_\mu(r)\phi_\nu^*(r) \qquad (2)$$

The determination of the coefficients requires the use of a modified singular value decomposition (SVD) procedure in which the inverse of an eigenvalue is set to zero if it is below a certain cutoff.

Using the fitted electronic densities, it has been shown [16] that it is possible to accurately compute the intermolecular Coulomb interaction energy (see Equation 3) from frozen monomer densities in the direct spirit of ab initio energy decomposition schemes (see for example references [17]).

$$Ecoulomb = \frac{Z_A Z_B}{r_{AB}} - \int \frac{Z_A \tilde{\rho}^B(\mathrm{r_B})}{r_{AB}} dr - \int \frac{Z_B \tilde{\rho}^A(\mathrm{r_A})}{r_{AB}} dr + \int \frac{\tilde{\rho}^A(\mathrm{r_A})\tilde{\rho}^B(\mathrm{r_B})}{r_{AB}} dr \qquad (3)$$

By using density fitting, both long-range multipolar and short-range penetration electrostatic energies (missing in a distributed multipole treatment) are included, the errors being relatively small compared to reference ab initio data using the same density matrices.

All the required integrals (electron-electron and electron-nuclear) were computed based on the McMurchie-Davidson recursions [18] enabling the use of higher angular

moment Gaussian functions if required. It is important to point out that the formalism also enables an accurate representation of both electrostatic potentials and fields enabling a direct link between GEM and SIBFA through a multiscale approach [5].

RESULTS

I. Improving electrostatics and short-range repulsion: from transition metal to heavy metals

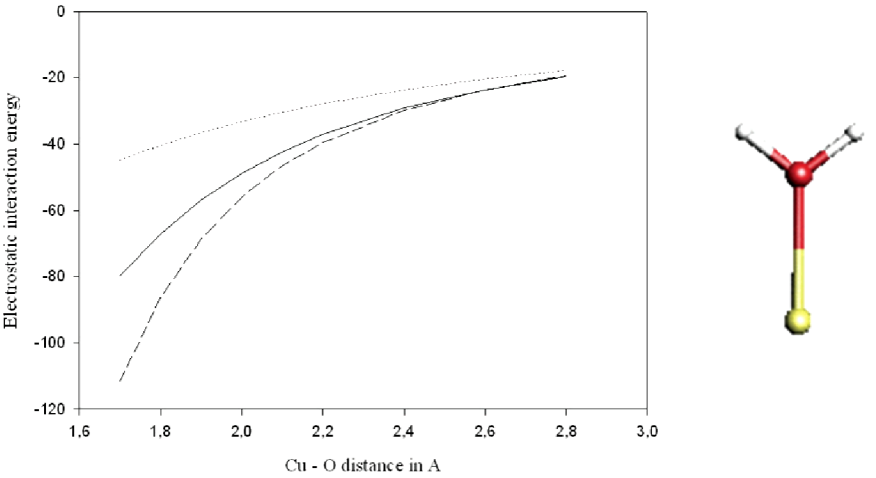

FIGURE 1. Variations of the Cu(I)-water electrostatic energy (kcal/mol) as a function of the Cu-O distance (A) computed at the Hartree-Fock level (E_c, DZVP2 basis set), using distributed multipoles (E_{MTP}) and short-range penetration corrected distributed multipoles (E_{MTP}*)

E_c -------- E_{MTP} ..... E_{MTP*} ————

Figure 1 displays the variations of the Cu(I)-water electrostatic energy (kcal/mol) as a function of the Cu-O distance. As we can see, the EMTP* formulation improves considerably the agreement with ab initio results as penetration energy is efficiently included. Such behaviour can be obtained for any metal cation from transition to heavy metals.

Concerning heavy metals, Figure 2 illustrates, for example, the accuracy of the short-range repulsion (denoted Erep*) in reproducing its ab initio exchange-repulsion Hartree-Fock counterpart in Pb(II)-water complex.[19] The $S/R+S^2/R^2$ formulation of Erep* enables a faithful reproduction of exchange-repulsion by the SIBFA APMM.

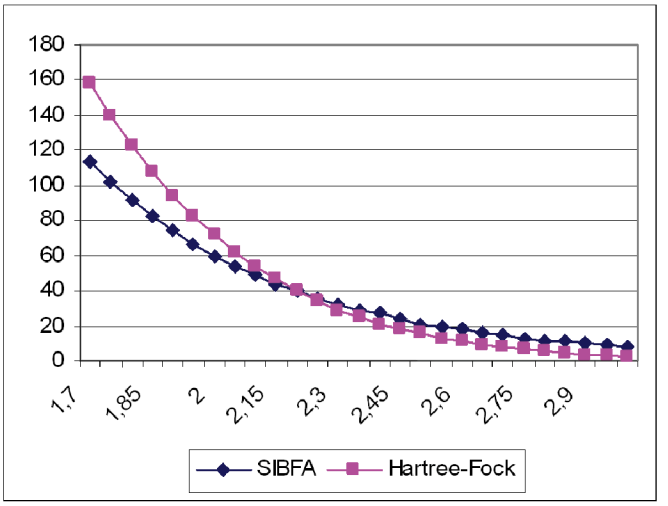

FIGURE 2. Variations of the Pb(II)-water exchange-repulsion energy (kcal/mol) as a function of the Pb-O distance (A) computed at the Hartree-Fock level (SDD (Pb)-DZVP2) and SIBFA level.

2. Modeling Cu(II) complexes with SIBFA: ligand field and short-range electrostatic and repulsion energies.

a) Cu(II)-water complex

Table I displays the SIBFA results for the Cu(II)-water complex.

TABLE I. SIBFA and MP2 (DZVP2) electrostatic, exchange-repulsion and total interaction energies for the Cu(II)-water complex at equilibrium distance (kcal/mol).
lines.

Ligand	SIBFA	Ab initio	SIBFA*-LF
$E_{es/MTP/MTP^*}$	-73.1	-93,0	-94.0
$E_{ex/rep}$	25.8	50,8	51.1
$E_{es/MTP/MTP^*} + E_{ex/rep}$	47,3	-42,2	-42,9
ΔE_{int}	-87,7	-83,2	-87,8

SIBFA*-LF denotes the latest version of SIBFA [4] that includes all short-range effects in both electrostatic and repulsion. As we can see, such corrections coupled to ligand field (LF) corrections [3] closely match the ab initio MP2 results.

b) Cu(II)-Poly-aqua complexes

Table II displays the SIBFA results for the Cu(II)-$(H_2O)n$ – complexes, respectively with 6 water molecules in the first solvation shell, 4 in the first shell and 1 in the second (5+1); 4 in the first shell and 2 in the second (4+2).

TABLE II. SIBFA and MP2 (DZVP2) total interaction energies for the Cu(II)-$(H_2O)n$ complexes at equilibrium distance (kcal/mol).

	SIBFA-LF		SIBFA		MP2	
Ligands	ΔE	d(Cu—O)	ΔE	d(Cu—O)	SIBFA-LF	SIBFA
$(H_2O)_6$	-340,7	1,99/2,17	-349,4	2,06	-346,4	-347,1
$(H_2O)_5$ — H_2O	-335,8	1,98-2,0/2,21	-340,9	1,99	-343,5	-342,2
$(H_2O)_4$ — $(H_2O)2$	-332,7	1,98/1,99	331,5	197/1,99	-339,6	-334,2

Here also, SIBFA*-LF is clearly able to energetically rank the 4 different complexes against MP2.

3. From GEM to a multiscale SIBFA/GEM integrated scheme.

a) **Modeling Metals using GEM.**

The GEM methodology has been recently successfully applied to water complexes as first results on metals were shown for electrostatic and exchange-repulsion. We present here some preliminary results [5] on polarization energies computed using GEM densities for both metal and water coupled to the SIBFA polarization scheme [1, 9]. Figure 3 displays the polarization energy of a water molecule interacting with a Ca(II). As we can see, the GEM approach offers a quasi-exact match of the ab initio polarization energy (here computed at the B3LYP level), even at very short-range when coupled to the SIBFA damping scheme [1, 9].

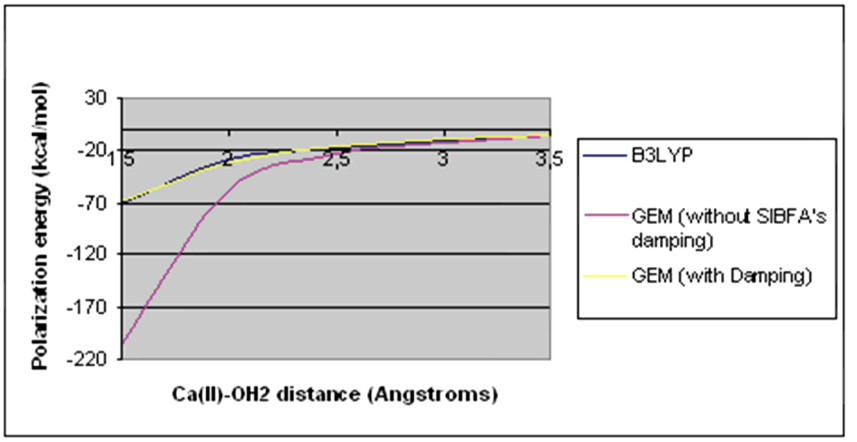

FIGURE 3. Variations of the Ca(II)-water exchange-repulsion energy (kcal/mol) as a function of the Pb-O distance (A) computed at the Hartree-Fock level (SDD (Pb)-DZVP2) and GEM levels (with or without damping).

b) **Towards a multiscale SIBFA/GEM approach.**

Of course as GEM requires the computations of integrals, one would like to be able to limit the approach to a subsystem like in QM/MM (see Figure 4).

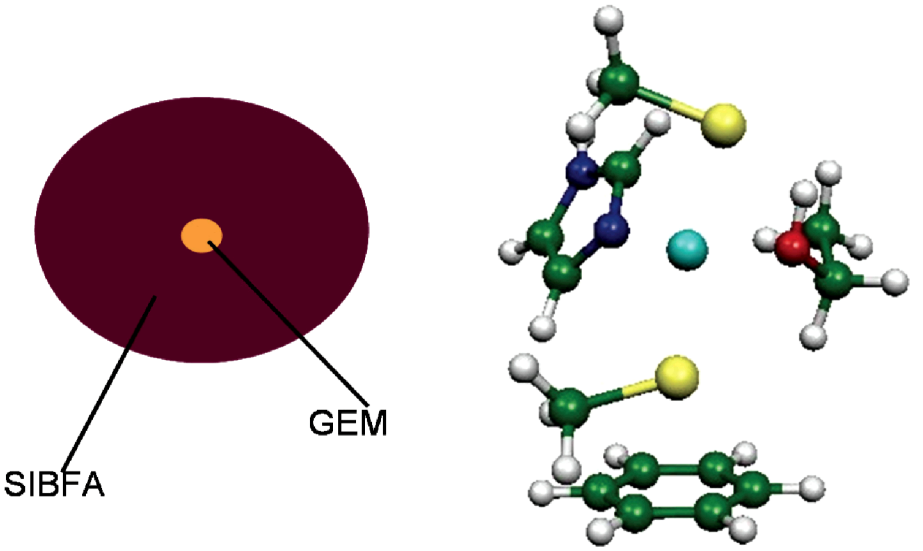

FIGURE 4. (left) Schematics of the multiscale SIBFA/GEM approach; (right) in this model of the alcoolate deshydrogenase (ADH) active site, GEM is limited to the treatment of the Zn(II) cation.

We present here a scheme where GEM is only applied to the Zn(II) metal cation and limited to the computation of the polarization and charge transfer energies.

Table III displays results on complexes of Zn(II) with formamide and imidazol. It also displays a more complex system: the alcoolate deshydrogenase (ADH) active site (see Figure 4 and reference [20] for details). As we can see, the scheme offers a high level of accuracy when compared to QC.

TABLE III. SIBFA, SIBFA-GEM and ab initio results (HF, SBK basis set) for selected complexes. The ADH geometry is extracted from reference [20].

Complexe	Ab initio		SIBFA		SIBFA/GEM	
	E_{pol}	E_{ct}	E_{pol}	E_{ct}	E_{pol}	E_{ct}
$Zn(formamide)^{2+}$	-61,6	-12,2	-61,7	-9,6	-59,8	-9,2
$Zn(imidazol)_3^{2+}$	-132,5	-37,7	-128,2	-29,2	-136,9	-33,5
$Zn(imidazol)_4^{2+}$	-132,6	-37,6	-127,2	-29,6	-133,7	-33,0
ADH	-84,5	-43,0	-93,0	-41,3	-83,7	-41,7

CONCLUSION

As we have seen the SIBFA and GEM methods can be applied to a vast range of molecules and metals, from transition metals to heavy metals. The development of an integrated SIBFA-GEM multiscale scheme should enable use to perform computations on very large system with high accuracy. Such an interface with GEM which can itself be coupled to QM should give rise in the near future to a multilevel QM/GEM/SIBFA methodology since GEM offers a direct connection between multipoles and densities.

ACKNOWLEDGMENTS

We thank the computer centers of CINES (Montpellier, France), CRIHAN (Rouen, France), CCRE (Paris, France), NIEHS (RTP, NC, USA) and of the Advanced Biomedical Computer Center, NCI-FCRDC, USA. This research was supported in part by the French ANR (project Lasihmodo), CNRS (France), the Ligue Nationale contre le Cancer (Comité Ile-de-France) and by the Intramural Research Program of the NIH and NIEHS (USA).

REFERENCES

1. N. Gresh, G. A. Cisneros, T. A. Darden and J-P Piquemal, J. Chem. Theory. Comput. **3**, 1960 (2007).
2. a) J-P. Piquemal, G. A. Cisneros, P. Reinhardt, N. Gresh and T. A. Darden, J. Chem. Phys. **124**, 104101 (2006); b) G. A. Cisneros, J-P. Piquemal and T. A. Darden, J. Chem. Phys. **125**, 184101 (2006); c) G. A. Cisneros, D. Elking, J-P Piquemal and T. A. Darden, J. Phys. Chem. A **111**, 12049 (2007).
3. J-P. Piquemal, B. Williams-Hubbard, N. Fey, R. J. Deeth, N.Gresh and C. Giessner-Prettre, J. Comput. Chem. **24**, 1963 (2003).
4. S. Ulmer, N. Gresh, R. J. Deeth, H. Gérard, B. de Courcy and J-P Piquemal, submitted (2008).
5. R. Chaudret, N. Gresh, O. Parisel, G. A. Cisneros, T. A. Darden and J-P Piquemal, submitted (2009).
6. J-P. Piquemal, N. Gresh and C. Giessner-Prettre, J. Phys. Chem A **107**, 10353 (2003).
7. a) J-P. Piquemal, Phd Thesis, Université Pierre et Marie Curie, France (2004); b) J-P Piquemal, H. Chevreau and N. Gresh, J. Chem. Theory. Comput. **3**, 824 (2007); c) N. Gresh, J-P. Piquemal and M. Krauss, J. Comput. Chem. **26**, 1113 (2005).
8. F. Vigné-Maeder, P. Claverie, J. Chem Phys. **88**, 4934 (1988).
9. J-P Piquemal, R. Chelli, P. Procacci and N. Gresh, J. Phys. Chem. A **111**, 8170 (2007).
10. D. R. Garmer, W. J. Stevens, J. Phys. Chem. A **93**, 8263 (1989).
11. J-P Piquemal, Int. J. Quant. Chem., submitted (2009).
12. N. Gresh, C. Claverie, A. Pullman , Int. J. Quantum Chem. **29**, 101 (1986).
13. J. N. Murrell, M. Randic, D. R. Williams, Proc. Roy. Soc. London Ser. A **284,** 566 (1966).
14. B. Jeziorski, R. Moszynski, K. Szalewicz, *Chem. Rev.* **94**, 1887 (1994).
15. S. F. Boys, I. Shavitt, A Fundamental Calculation of the Energy Surface *for the System of Three Hydrogens Atoms* NTIS, Springfield, VA, 1959, AD212985; b) B. I. Dunlap, J. W. D. Connoly, J. R. Sabin, J. Chem. Phys. **71**, 4993 (1979).
16. G. A. Cisneros, J-P. Piquemal and T. A. Darden, J. Chem. Phys. **123**, 044109 (2005).
17. J-P. Piquemal, A. Marquez, O. Parisel and C. Giessner-Prettre, J. Comput. Chem. **26**, 1052 (2005).
18. L. E. McMurchie, E. R. Davidson, J. Comput. Phys. *26*, 218 (1978).
19. M-C van Severen, J-P Piquemal, N. Gresh, R. Chaudret, O. Parisel, submitted (2009).
20. B. de Courcy, J-P Piquemal and N. Gresh, J. Chem. Theo. Comput. **4**, 1659 (2008).