J. Chem. Theory Comput. **2007,** *3,* 1047−1053

JCTC Journal of Chemical Theory and Computation

Self-Consistent Strictly Localized Orbitals

Pierre-François Loos and Xavier Assfeld*

Equipe de Chimie et Biochimie Théoriques, UMR 7565 CNRS-UHP,
Institut Jean Barriol (FR CNRS 2843), Faculté des Sciences et Techniques,
Nancy-Université, B.P. 239, 54506 Vandœuvre-lès-Nancy Cedex, France

Received October 30, 2006

Abstract: Among all the Quantum Mechanics/Molecular Mechanics (QM/MM) methods available to describe large molecular systems, the Local Self-Consistent Field/MM (LSCF/MM) one uses frozen doubly occupied Strictly Localized Bonding Orbital (SLBO) to connect the QM fragment to the one treated at the MM level. This approach is correct as long as the QM part is large enough to minimize the artifacts that could arise because of the fixed SLBO. If one wants to decrease the size of the QM subsystem, one clearly needs to help the SLBO to relax according to the variations of the global wave function. Also, the SLBO have to adjust itself according to the modification of the surrounding if we want to improve the method. Here, we present a modification of the original LSCF method called Optimized LSCF (OLSCF) where each SLBO is allowed to mix with its corresponding Strictly Localized Anti Bonding Orbital (SLABO) resulting in an adjustment of the two-electron bond described by a self-consistent SLBO (SCSLBO). We test the new methodology against the modification of the QM part (internal perturbation) and against the variation of the surroundings (external perturbation) represented either by a dielectric continuum or by a classical point charge. In each case the initial SLBO is the symmetric C−C SLBO of the ethane molecule. It is shown that the optimized SCSLBO presents a final polarity in perfect agreement with what could be expected as the result of a reaction to the internal or external perturbation.

1. Introduction

During the past decade, the interest of theoreticians for macromolecular systems, or more globally large systems, has grown enormously. This is a direct consequence of the tremendous improvement of computers, both from the storage (memory and/or hard drive) and the CPU points of view. However, one has to bear in mind that studying chemical reactions in such systems (i.e. locating transition states) is still out of reach for the casual methods of quantum chemistry, despite the computer enhancement. To be able to treat such large molecular systems, the community of quantum chemists has developed new methods that combine different levels of theory. The total system is divided in several parts, each one described at a given level of theory. The levels of theory are chosen such that, for each part, the most important physical phenomenon is correctly handled. Various schemes (combinations) are possible. For example, the part of the macromolecule undergoing the chemical reaction (generally few atoms) needs to be treated with high level (for example, correlated ab initio) techniques. This is called Quantum Mechanics (QM) level. The neighboring region, larger than the QM one, where consecutive reactions can take place—for example proton transfer rearrangement like in the bacteriorhodopsin molecule—can be described with a less time demanding method, typically a semiempirical one. Let us call this level QM′ to avoid confusion with the former QM level. The remaining atoms of the macromolecule, which play a nontrivial and nonisotropic role by means of electrostatic and mechanic interactions, are often handled with Molecular Mechanics (MM) force fields. Finally the surrounding solvent can be included explicitly via a classical force field (MM′) or implicitly using a

---

* Corresponding author e-mail: Xavier.Assfeld@cbt.uhp-nancy.fr.

10.1021/ct6003214 CCC: $37.00

polarizable continuum (Self-Consistent Reaction Field, SCRF). The combination of these four levels of theory gives rise to a QM/QM′/MM:MM′ or QM/QM′/MM:SCRF method, if one uses the "/" symbol to represent the separation between two parts connected by a chemical bond and the ":" character when no chemical bond connects the two parts. One could of course think of more complicated combinations, but the most widely known and used schemes of hybrid method are certainly the QM/MM and the QM:MM ones.[1]

Let us focus our attention on the methods that need the "/", i.e. when covalent chemical bonds are formally cut to divide the system. Many solutions have been proposed and applied to circumvent the so-called dangling bond problem.[2−13] If used carefully each scheme will correctly handle the connection between the two parts. We refer the interested reader to a recent review for further details.[14] In this article we will restrain our discussion on the Local Self-Consistent Field combined with MM (LSCF/MM)[15−18] way to treat such connection.

The connecting bond is described by a frozen doubly occupied Strictly Localized Bonding Orbital (SLBO). By frozen we mean that the expansion coefficients of the basis functions in the SLBO are kept constant during the wave function optimization (during the Self-Consistent Field (SCF) procedure). Trivially the SLBO is readjusted by rotation and renormalization if the geometry is changed during either a geometry optimization or a molecular dynamics step. Thanks to the transferability principle, the SLBO is generally obtained from a preliminary calculation on a small model molecule that possesses the chemical bond of interest. Freezing the SLBO implies that it is the same whatever the state of the QM part is. This can only be correct if the size of the QM part is large enough to reduce the interaction between the frozen SLBO and the global wave function to something negligible. It is a severe limitation if one is willing to perform reactive dynamic calculations, since the required CPU time will be enormous and unreachable with normal computer equipment. The corollary is that to have the size of the QM part reduced, it is necessary that the SLBO can relax somehow. In order to keep the QM size as small as possible we propose a modification of the LSCF method—based on quantum mechanics only (not of the hybrid LSCF/MM method yet)—called Optimized LSCF (OLSCF), in which the SLBOs can readjust themselves according to the variation of the global wave function (called internal perturbation) and to the variation of the surroundings (called external perturbation). The LSCF method needs SLBO, the OLSCF one produces Self-Consistent SLBO (SCSLBO).

This paper is organized as follows. Section 2 deals with the general principle of the method. The basic equations of the OLSCF method are established in section 3. Finally, some illustrative calculations are gathered in section 4 and discussed in the fifth section. Our aim is to prove that the OLSCF method is able to correctly represent the polarity variation of the SCSLBO when the surroundings change and that these changes are consistent with the general chemical sense, before adapting the new methodology to the QM/MM framework.

2. Principle

The LSCF method allows the optimization of a wave function at the Hartree−Fock level, or equivalently at the Kohn−Sham level in the Density Functional Theory (DFT) framework, knowing that some predefined (spin-)orbitals must remain unchanged. The predefined orbitals can be of any shape and occupied or not. Any post-HF method can follow the LSCF optimization. All orbitals are orthogonal.

Our aim is to allow SLBOs, expanded on the basis functions of the two atoms defining the bond, to remain strictly localized but to have some degrees of freedom to adjust themselves according to the surroundings. We decide to mix the doubly occupied SLBO to its corresponding empty Strictly Localized Anti-Bonding Orbital (SLABO). Hence, each connecting bond will be described with a doubly occupied SCSLBO that is a linear combination of the initial SLBO and SLABO, i.e. we give one electronic degree of freedom to the bond. To avoid the delocalization of the SCSLBO, we do not mutually orthogonalize the SLBOs nor the SLABOs. The variational Molecular Orbitals (MO) are mutually orthogonal and orthogonal to each {SLBO, SLABO} pair.

The mixing of the SLBO and the SLABO is not new. Many studies have been performed in the 1980s mainly by P. Surján,[19−24] and even before in the 1960s and 1970s with the PCILO method of J.-P. Malrieu,[25−27] and related methods.[28−31] Here, we are less ambitious since only a few bonds, those connecting two parts described at different levels of theory, will be treated like that.

3. Theory

SLBO. Let us consider a closed shell system of $2n$ electrons treated with the spin restricted formalism, for simplicity. The generalization to the unrestricted case is trivial. We use a set of atom centered basis functions $\{|\phi_\mu\rangle\}_{\mu=1}^{K}$. Let say that L SLBOs $\{|l_i\rangle\}_{i=1}^{L}$ are considered, with $L < K$. They are developed onto the basis functions of the two atoms X_i and Y_i defining the bonds number i.

$$|l_i\rangle = \sum_{\mu \in X_i} a_{\mu i} \,|\, \phi_\mu \rangle + \sum_{\nu \in Y_i} a_{\nu i} \,|\, \phi_\nu \rangle \tag{1}$$

These SLBOs are obtained from usual localization procedures[32−37] on small model molecules. The strict localization is attained by zeroing the coefficients of basis functions belonging to atom different from X_i or Y_i, if necessary.[38]

They can be written in a shorter form introducing normalized hybrid orbitals (HO)

$$|l_i\rangle = a_{X_i i} \,|\, h_i^{X_i} \rangle + a_{Y_i i} \,|\, h_i^{Y_i} \rangle \tag{2}$$

where $h_i^{X_i}$ is a HO belonging to atom X_i and used in the ith SLBO l_i.

SLABO. The corresponding SLABOs, noted $\{|l_i^*\rangle\}_{i=1}^{L}$, can be obtained with different manners, and we list four of them below.

(1) They can be obtained the same way SLBOs are created with the usual localization techniques[32−37] applied on the virtual orbitals of a small model molecule.

(2) However, as the size of the basis set increases, the number of virtual orbitals increases also, and the convergence of the localization is more difficult to achieve. In that case one can just project the SLABO found with a smaller basis (SB) set on the large basis (LB) set

$$|l_i^*\rangle_{\mathrm{LB}} = \sum_{\mu,\nu}^{LB} |\,\phi_\mu\rangle\,(S^{-1})_{\mu\nu}\langle\,\phi_\nu\,|\,l_i^*\,\rangle_{\mathrm{SB}} \tag{3}$$

where $S_{\mu\nu}$ is the overlap integral between the functions ϕ_μ and ϕ_ν of the large basis set.

(3) SLABOs can also be defined with the elaborate method proposed by M. Head-Gordon and co-workers.[39]

(4) Finally, it is also possible to derive the antibonding orbital directly from the bonding one.[20,40] The sign of the coefficients of the basis functions of one atom is reversed, and the resulting function is renormalized.

$$\begin{aligned} |l_i^*\rangle &= N_i^*\,(a_{X_i\mathrm{i}}|h_i^{X_i}\rangle - a_{Y_i\mathrm{i}}\,|\,h_i^{Y_i}\rangle) \\ &= a_{X_i\mathrm{i}}^*\,|\,h_i^{X_i}\rangle + a_{Y_i\mathrm{i}}^*\,|\,h_i^{Y_i}\rangle \end{aligned} \tag{4}$$

One has to note that for the first three possibilities, the SLBO and the corresponding SLABO are not developed over the same hybrid orbitals, whereas the fourth method expands the SLBO and the SLABO over the same two HOs.

In this article the two first procedures are used since they are easy to get and because they give more reliable results, based on preliminary calculations.

MO. To determine the variational MOs we use the same procedure as the one we developed for the LSCF method 10 years ago.[15] The only exception is that the predefined orbitals (all pairs of SLBO and SLABO) are not orthogonal. First, we recall the steps followed in the LSCF method:

- Requirement: the MOs are orthogonal to the frozen orbitals (FOs). Hence, we expand the MO over basis functions that are already orthogonal to the FOs.

- Basis functions orthogonal to the FOs: each basis function ϕ_μ from the original basis set $\{|\phi_\mu\rangle\}_{\mu=1}^{K}$ is projected out of the subspace defined by the FOs. To have a simple expression of the projection operator, the FOs are mutually orthogonalized, generally by means of the Löwdin procedure

$$|\tilde{\phi}_\mu\rangle = [1 - \sum_i^L {R_{\mu i}}^2]^{-1/2}\,[|\phi_\mu\rangle - \sum_i^L |ol_i\rangle\langle ol_i\,|\phi_\mu\rangle] \tag{5}$$

where $R_{\mu i}$ is the overlap integral between the basis function ϕ_μ and the orthonormalized FO ol_i, $\langle\phi_\mu|ol_i\rangle$.

- Removing the linear dependencies: since L FOs where predefined, the new set of K basis functions has at least L linear dependencies. They are removed by means of the canonical orthogonalization procedure.[41] The set of $(K - L)$ resulting basis functions is exactly what is needed to expand the MOs.

The only difference with the new OLSCF method is that to avoid the delocalization of the SLBOs and SLABOs we keep them strictly localized and thus possibly nonorthogonal. The projection operator is then modified

$$|\tilde{\phi}_\mu\rangle = [1 - \sum_i^{2L}\sum_j^{2L} T_{\mu i}\,(D^{-1})_{ij}\,T_{\mu j}]^{-1/2}\,[|\phi_\mu\rangle - \sum_i^{2L}\sum_j^{2L} |\,l_i\,\rangle\,(D^{-1})_{ij}\langle l_j\,|\,\phi_\mu\rangle] \tag{6}$$

with D_{ij} being the overlap integral between the FOs l_i and l_j, $\langle l_i|l_j\rangle$, and $T_{\mu i}$ is the overlap integral between the basis function ϕ_μ and the FO l_i, $\langle\phi_\mu|l_i\rangle$. One has to note that here i and j run over the SLBOs and the SLABOs and that l_i can be a SLBO or a SLABO. This transformation can be done with the matrix $\boldsymbol{M}$ whose elements are given by

$$M_{\mu\nu} = [1 - \sum_i^{2L}\sum_j^{2L} T_{\mu i}\,(D^{-1})_{ij}\,T_{\mu j}]^{-1/2}\,[\delta_{\mu\nu} - \sum_i^{2L}\sum_j^{2L}\sum_\eta a_{\nu i}\,(D^{-1})_{ij}\,a_{\eta j}\,S_{\eta\mu}] \tag{7}$$

where $\delta_{\mu\nu}$ is the Kronecker symbol, and $S_{\eta\mu}$ is the overlap integral between functions ϕ_η and ϕ_μ.

This projection is followed by the canonical orthogonalization to remove the $2L$ linear dependencies. The orthogonalization is performed with a matrix named $\boldsymbol{X}$. The conjunction of these two steps lead to the definition of a matrix, named $\boldsymbol{B} = \boldsymbol{MX}$, that transforms the original basis set of K nonorthogonal functions into a set of $(K - 2L)$ functions mutually orthogonal and orthogonal to the SLBOs and SLABOs. This $\boldsymbol{B}$ matrix plays, for the OLSCF method, exactly the same role the Löwdin matrix ($\boldsymbol{X} = \boldsymbol{S}^{-1/2}$) plays in the SCF procedure.

The other point one has to take care of in the OLSCF (or LSCF) method is the definition of the density matrix. In the LSCF method, the density matrix ($\boldsymbol{P}^{\boldsymbol{L}}$) due to the occupied frozen orbitals must be added to the one resulting from the variational MOs ($\boldsymbol{P}^{\boldsymbol{Q}}$) to give the total density matrix $\boldsymbol{P}^{\boldsymbol{T}} = \boldsymbol{P}^{\boldsymbol{L}} + \boldsymbol{P}^{\boldsymbol{Q}}$. Here, thanks to the orthogonality of the MOs (mutually orthogonal and orthogonal to the strictly localized orbitals (SLOs)) the total density matrix can also be split into the two usual components $\boldsymbol{P}^{\boldsymbol{Q}}$ and $\boldsymbol{P}^{\boldsymbol{L}}$. However, due to the nonorthogonality of the SLOs the density matrix due to the occupied predefined orbital requires special attention. One matrix element of the $\boldsymbol{P}^{\boldsymbol{L}}$ matrix is written as

$$P_{\mu\mathrm{n}}^L = 2\sum_j^L a_{\mu j}\,\tilde{a}_{\nu j} \tag{8}$$

where $\tilde{a}_{\nu j}$ are the expansion coefficients of the dual orbitals (DOs).[24]

$$\tilde{a}_{\nu j} = \sum_i^{2L} a_{\nu i}\,(D^{-1})_{ij} \tag{9}$$

The DOs $\{|\tilde{l}_i\rangle\}_{i=1}^{2L}$ are orthogonal to the set of SLO $\{|l_i\rangle\}_{i=1}^{2L}$ ($\langle l_i\,|\,\tilde{l}_j\rangle = \delta_{ij}$), but the two sets are not mutually orthogonal $\langle l_i\,|l_j\rangle = D_{ij}$ and $\langle\tilde{l}_i\,|\,\tilde{l}_j\rangle = \tilde{D}_{ij}$). The SLOs and the DOs form a biorthogonal basis set.[42] Although it is possible to express the energy and the Roothaan equation as functions of the initial nonorthogonal orbitals,[43] it requires a lot of modifications of the existing code. However, the DOs propose a very simple form very close to the usual one. With the definition of the density matrix given above (eq 8) the Fock matrix elements and the electronic energy can be obtained with the traditional formula (the nonorthogonality is hidden in the density matrix)

$$F_{\mu\nu} = H^{c}_{\mu\nu} + \sum_{\lambda,\sigma} P^{T}_{\lambda\sigma}\left[(\mu\nu|\lambda\sigma) - \frac{1}{2}(\mu\sigma|\lambda\nu)\right] \qquad (10)$$

$$E = \sum_{\mu,\nu} P^{T}_{\mu\nu}H^{c}_{\mu\nu} + \frac{1}{2}\sum_{\mu,\nu}\sum_{\lambda,\sigma} P^{T}_{\mu\nu}P^{T}_{\lambda\sigma}\left[(\mu\nu|\lambda\sigma) - \frac{1}{2}(\mu\sigma|\lambda\nu)\right] \qquad (11)$$

where we have used the usual notation for the various integrals. Equation 11 is equivalent to half the trace of the matrix product ($\boldsymbol{P^T}(\boldsymbol{H}+\boldsymbol{F})$). At this point we have all the tools to obtain the MOs.

SCSLBO. For each pair (l_i, l_i^*) composed of a SLBO and its corresponding SLABO expanded over the basis functions of the two atoms defining the connecting bond, we build a 2 × 2 Fock matrix

$$\begin{pmatrix} \langle l_i|\hat{F}|l_i\rangle & \langle l_i^*|\hat{F}|l_i\rangle \\ \langle l_i|\hat{F}|l_i^*\rangle & \langle l_i^*|\hat{F}|l_i^*\rangle \end{pmatrix} \qquad (12)$$

with

$$\langle l_i|\hat{F}|l_i\rangle = \sum_{\mu,\nu} a_{\mu i}a_{\nu i}F_{\mu\nu} \qquad (13)$$

where $F_{\mu n}$ comes from eq 10.

This matrix is transformed in a basis where the SLBO is orthogonal to the SLABO, with either a Löwdin or a Gram-Schmidt orthogonalization and further diagonalized. The two eigenvectors correspond to the new SLBO and SLABO that we call SCSLOs. From these new functions a new density matrix $\boldsymbol{P^L}$ can be formed and new MOs can be sought. The whole process is conducted until convergence is reached. It is summarized in the algorithm given below.

Algorithm. We have implemented this method in our modified version of the package Gaussian 03.[44]

(1) Determine the SLBOs and SLABOs on model molecules. Orthogonalize the SLABO to its SLBO, with the Löwdin or the Gram-Schmidt method.

(2) Orthogonalize the basis functions with respect to all SLOs, i.e. build the matrix $\boldsymbol{M}$.

(3) Perform the canonical orthogonalization of the new basis set with the matrix $\boldsymbol{X}$.

(4) Get an initial guess of the density matrix.

(5) Build in the Fock matrix $\boldsymbol{F}$.

(6) Transform $\boldsymbol{F}$ in the orthogonal basis: $\boldsymbol{F}' = \boldsymbol{B}^t\,\boldsymbol{F}\,\boldsymbol{B}$.

(7) Diagonalize $\boldsymbol{F}'$: $\epsilon = \boldsymbol{C}'^t\,\boldsymbol{F}'\,\boldsymbol{C}'$ where ϵ is the diagonal $(K-2L)\times(K-2L)$ matrix of the eigenvalues.

(8) Backtransfrom in the original basis: $\boldsymbol{C} = \boldsymbol{B}\,\boldsymbol{C}'$.

(9) Compute the density matrix due to the MOs, $\boldsymbol{P^Q}$.

(10) Diagonalize the 2 × 2 Fock matrix, for each (l_i, l_i^*) pair to get the SCSLOs.

(11) Compute the density matrix due to the SCSLBOs, $\boldsymbol{P^L}$.

Table 1. Half Mulliken Electronic Population of the C−C Bond Orbitals (in Electron) of the Carbon Atom Bonded to the X Ligand in the CH_3CX_3 Molecules (i.e. Bond Polarity)[c]

	Pipek-Mezey			Boys-Foster	
X	HF/ 6-31G*	HF/ 6-311G**[a]	HF/ 6-311G**[b]	HF/ 6-31G*	HF/ 6-311G**[b]
			SCSLBO		
H	0.500	0.500	0.500	0.500	0.500
Li	0.479	0.481	0.472	0.490	0.484
BH_2	0.521	0.519	0.529	0.504	0.507
CH_3	0.529	0.524	0.539	0.509	0.516
NH_2	0.559	0.553	0.576	0.524	0.539
OH	0.586	0.581	0.612	0.538	0.560
F	0.592	0.594	0.621	0.545	0.571
			SLBO		
H	0.500	0.500	0.500	0.500	0.500
Li	0.415	0.460	0.460	0.430	0.475
BH_2	0.518	0.544	0.544	0.519	0.548
CH_3	0.498	0.529	0.529	0.498	0.529
NH_2	0.527	0.580	0.580	0.524	0.575
OH	0.546	0.579	0.579	0.544	0.574
F	0.523	0.559	0.559	0.515	0.553

[a] SLABO obtained at the HF/6-311G**. [b] SLABO obtained at the HF/6-31G* and projected on the 6-311G** basis set. [c] The SCSLBOs result from an OLSCF calculation. Their polarities are given in the upper part of the table. The SLBOs are obtained as usual (see the Theory section) after a traditional SCF calculation. Their polarities are given in the lower part of the table.

(12) Compute $\boldsymbol{P^T} = \boldsymbol{P^Q} + \boldsymbol{P^L}$.

(13) Exit test. If not satisfied go back to step 5.

4. Test Results

In order to show that with our new formalism a given SLBO can adjust itself with respect to either an internal modification or to the surroundings variation we perform three series of calculations.

For the first one, we consider the H_3C-CX_3 molecules (X=H, Li, BH_2, NH_3, OH, F) for which the initial SLBO and SLABO are determined on the ethane molecule. Two levels of theory (HF/6-31G* and HF/6-311G**) and two localization schemes (Boys-Foster,[32] BF, and Pipek-Mezey,[33] PM) are considered. The polarity (P) of the optimized SCSLBO is represented as half the electronic population (q) due to this SCSLBO on the carbon atom connected to the X substituents (C_X) computed with the Mulliken partitioning.[45]

$$P = \frac{q_{C_X}}{2} = \sum_{\mu}^{\in C_X}\sum_{\nu}^{AO} c_{\mu i}c_{\nu i}S_{\mu\nu} \qquad (14)$$

For comparison, the polarity of the SLBO obtained with the BF or PM a posteriori criterion on a SCF optimized wave function is also provided. The geometries were optimized at the SCF level. The polarities are gathered in Table 1.

The second series is concerned with the influence of the solvent. The solvent is represented by a polarizable continuum with the help of the SCRF method developed previously in our group.[46,47] The initial SLBO is determined on the ethane molecule at the B3LYP/6-311++G** level of theory with the help of the PM localization scheme and

Table 2. Half Mulliken Electronic Population (in Electron) of the Carbon Atom Bearing the NH_2 Group in the β-Alanine Molecule (i.e. Bond Polarity), Due to the SCSLBO and to the SLBO(PM)[a]

form	phase	SCSLBO	SLBO(PM)
neutral	gas	0.495	0.501
	solvent	0.495	0.493
Δ		+0.0004	−0.0079
zwitterionic	gas	0.553	0.637
	solvent	0.528	0.552
Δ		−0.0246	−0.0854

[a] Δ is the difference of polarity between the solvent and the gas phases.

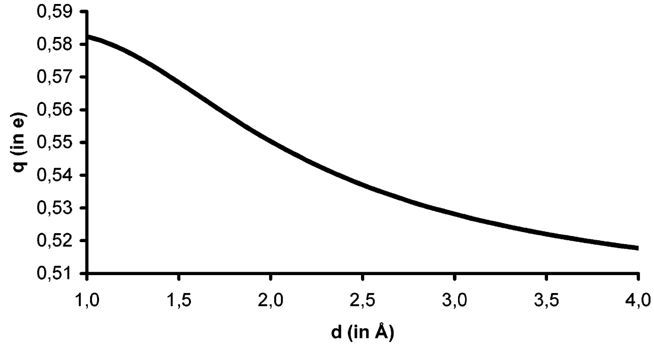


Figure 1. Half Mulliken electronic population (q, in electron) of the carbon atom nearest to the classical point charge due to the SCSLBO (i.e. bond polarity), as a function of the distance (d, in Å) between the positive point charge and the nearest carbon atom of the ethane molecule. The point charge is placed on the line defined by the two carbon atoms.

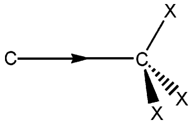
transferred on the central C−C bond of the β-alanine molecule in its neutral and zwitterionic forms, $H_2N-CH_2-CH_2-COOH$ and $H_3N^+-CH_2-CH_2-COO^-$, respectively. Calculations are performed in the gas phase and in solution at the B3LYP/6-311++G** level of theory. The geometries are optimized at this same level in solvent phase only and without frozen orbital. The polarities of the SCSLBO are given in Table 2. For comparison purposes, pure B3LYP/6-311++G** calculations are carried out and followed by a PM localization procedure to obtained SLBOs. The polarity of the SLBO of the central C−C bond is given in Table 2. Polarities correspond to half the electronic population, arising from the SLO, of the carbon atom bearing the NH_2 group.

Finally, the effect of a surrounding classical point charge ($q = +|e|$) is studied in the third series, to anticipate QM/MM calculations. The chosen molecule is ethane. The starting SLBO comes from a preceding calculation on the C_2H_6 molecule at the HF/6-311G** level of theory with the PM criterion. The point charge is placed on the line defined by the two carbon atoms. The distance between the point charge and the nearest carbon atom is varied from 1 to 4 Å. The polarity of the SCSLBO, represented as half the electronic population due to this SCSLBO of the C atom nearest to the point charge, is sketched in Figure 1.

5. Discussion

From the results presented in Table 1, one can readily extract that the polarity (i.e. half the electronic population, due to the SCSLBO only, of the carbon atom connected to the X substituents) increases when the electronegativity of the substituent increases. This result is in perfect agreement with the usual "chemical sense" (see Scheme 1). Although not exactly identical, this trend is the same whatever the level of theory is (double- or triple-ζ), whatever the localization criterion is (BF or PM), and whatever the SLABO is (projected from double-ζ or not). However, it seems that the PM SLOs give a larger polarization than the BF ones.

Scheme 1. Polarization of the C−C SLBO by the Electronegative Substituents X



To have a point of comparison, we run traditional SCF calculations on all molecules, and we determine the C−C Localized Molecular Orbital (LMO) for each molecule with either the PM or the BF criterion at all levels of theory. These LMOs were further transformed in SLBOs, and their polarities are given in the second part of Table 1. If one compares the polarity of the SCSLBO with the one obtained for the SLBO, one can remark that the values are quite similar although not identical. The increase of the electronic population with the increasing electronegativity of the ligands, for the SLBO, is not as straight as it is for SCSLBO. For SLBO, the polarity seems to drop down for the CH_3 and the F ligand. This is due to the fact that these SLBOs are obtained by rotation of all the occupied molecular orbitals of the molecule and thus reflect the global effect of the ligands on the total wave function. Of course this effect cannot be fully reproduced by the polarity of just one function even if it is a SCSLBO. Anyway, we can conclude that the SCSLBOs correctly show the expected trend and that they adequately adjust themselves according to the chemical modification of the molecule.

The values given in Table 2 show that the SCSLBO and the SLBO are not much modified by the solvent effects for the neutral form of the β-alanine molecule. The difference of electronic population between the gas and the solvent phase is small in both cases. For the zwitterionic molecule, the electronic polarization induced by the solvent is at least ten times stronger than for the neutral isomer. We can put that in parallel with the electrostatic solvation free energy that is also small for the neutral form, 14.5 kcal•mol^{-1}, compared to the one of the zwitterionic form, 59.7 kcal•mol^{-1}. The polarization of the SCSLBO is less than the one of the SLBO, albeit both present the same tendency. Again, this discrepancy can be attributed to the global character of the SLBO and to the local nature of the SCSLBO. Anyway, it is shown that the SCSLBO reacts correctly with respect to the perturbation of an external electric field (here produced by the dielectric continuum).

Finally, before the adaptation of the OLSCF scheme to the QM/MM formalism, we decide to focus our attention to the action of a classical charge on the polarity of the SCSLBO. One can see (Figure 1) that the stronger the interaction, i.e. the shorter the distance between the classical charge and the molecule, the stronger the polarization of the SCSLBO. That is exactly what could be expected and is encouraging for the future implementation in a QM/MM framework.

6. Conclusion

In this first step, we propose a modification of the LSCF method—based on quantum mechanics only (not the LSCF/MM hybrid method)—which allows the relaxation of strictly localized bonding orbitals according to the variation of the surroundings and/or of the quantum state. The new method is named Optimized LSCF (OLSCF). For each SLBO, a SLABO is defined, and the mixing of these two functions gives the electronic degree of freedom responsible for the relaxation of the orbital. The new methodology is tested against the modification of the molecule itself, against the influence of the surrounding solvent (modeled by means of a polarizable continuum), and against the presence of a classical point charge. In all three cases investigated here, it is shown that the SCSLBOs properly adjust themselves to the situation. According to these encouraging results, we will further apply our new technique to the QM/MM formalism and to the QM/QM′ method in forthcoming papers.

Acknowledgment. P.-F.L. and X.A. are deeply indebted to Professor Jean-Louis Rivail for the initial idea leading to this new development and for many fruitful discussions.

References

(1) Moreau, Y.; Assfeld, X. In *Lecture Series on Computer and Computational Sciences*; Maroulis, G., Simos, T., Eds.; Koninklijke Brill NV: Leiden, The Netherlands, 2005; Vol. 3, pp 1−9.

(2) Ranganathan, S.; Gready, J. E. *J. Phys. Chem. B* **1997**, *101*, 5614−5618.

(3) Cummins, P. L.; Gready, J. E. *J. Phys. Chem. B* **2000**, *104*, 4503−4510.

(4) Das, D.; Eurenius, K. P.; Billings, E. M.; Sherwood, P.; Chatfield, D. C.; Hodoscek, M.; Brooks, B. R. *J. Chem. Phys.* **2002**, *117*, 10534−10547.

(5) Maseras, F.; Morokuma, K. *J. Comput. Chem.* **1995**, *16*, 1170−1179.

(6) Zhang, Y.; Lee, T.-S.; Yang, W. *J. Chem. Phys.* **1999**, *110*, 46−54.

(7) Antes, I.; Thiel, W. *J. Phys. Chem. A* **1999**, *103*, 9290−9295.

(8) DiLabio, G. A.; Hurley, M.; Christiansen, P. A. *J. Chem. Phys.* **2002**, *116*, 9578−9584.

(9) Gao, J.; Amara, P.; Alhambra, C.; Field, M. J. *J. Phys. Chem. A* **1998**, *102*, 4714−4721.

(10) Pu, J.; Gao, J.; Truhlar, D. G. *J. Phys. Chem. A* **2004**, *108*, 632−650.

(11) Lin, H.; Truhlar, D. G. *J. Phys. Chem. A* **2005**, *109*, 3991−4004.

(12) Kairys, V.; Jensen, J. H. *J. Phys. Chem. A* **2000**, *104*, 6656−6665.

(13) Murphy, R. B.; Philipp, D. M.; Friesner, R. A. *J. Comput. Chem.* **2000**, *21*, 1442−1457.

(14) Lin, H.; Truhlar, D. G. *Theor. Chem. Acc.* **2007**, *117*, 185−199.

(15) Assfeld, X.; Rivail, J.-L. *Chem. Phys. Lett.* **1996**, *263*, 100−106.

(16) Ferré, N.; Assfeld, X.; Rivail, J.-L. *J. Comput. Chem.* **2002**, *23*, 610−624.

(17) Moreau, Y.; Loos, P.-F.; Assfeld, X. *Theoret. Chem. Acc.* **2004**, *112*, 228−239.

(18) Fornili, A.; Loos, P.-F.; Sironi, M.; Assfeld, X. *Chem. Phys. Lett.* **2006**, *427*, 236−240.

(19) Surján, P.; Mayer, I. *Theor. Chim. Acta* **1981**, *59*, 603−607.

(20) Mayer, I.; Surján, P. *J. Chem. Phys.* **1984**, *80*, 5649−5658.

(21) Surján, P.; Révész, M.; Mayer, I. *J. Chem. Soc., Faraday Trans. 2* **1981**, *77*, 1129−1131.

(22) Surján, P. *J. Mol. Struct. (THEOCHEM)* **1988**, *169*, 95−104.

(23) Surján, P. *Chem. Phys. Lett.* **1985**, *117*, 386−388.

(24) Surján, P. *Chem. Phys. Lett.* **1982**, *92*, 483−485.

(25) Diner, S.; Malrieu, J.-P.; Clavery, P. *Theor. Chim. Acta* **1969**, *13*, 1−17.

(26) Malrieu, J.-P.; Clavery, P.; Diner, S. *Theor. Chim. Acta* **1969**, *13*, 18−45.

(27) Malrieu, J.-P. In *Semiempirical methods in electronic structure calculation;* Segal, G. A., Ed.; Plenum: New York, 1977.

(28) Kvasnika, V. *Theor. Chim. Acta* **1974**, *34*, 61−65.

(29) Cullen, J. M.; Zerner, M. C. *Int. J. Quantum Chem.* **1982**, *22*, 497−535.

(30) Boča, A. *Theor. Chim. Acta* **1982**, *61*, 179−192.

(31) Barone, V.; Douady, J.; Ellinger, Y.; Subra, R. *J. Chem. Soc., Faraday Trans. 2* **1979**, *75*, 1597−1611.

(32) Foster, J. M.; Boys, S. F. *Rev. Mod. Phys.* **1960**, *32*, 300−302.

(33) Pipek, J.; Mezey, P. *J. Chem. Phys.* **1989**, *90*, 4916−4926.

(34) Weinstein, H.; Pauncz, R.; Cohen, M. In *Advances in Atomic and Molecular Physics*; Academic Press: New York, 1971; Vol. 7, p 97.

(35) Magnasco, V.; Perico, A. *J. Chem. Phys.* **1972**, *47*, 971−981.

(36) Fornili, A.; Moreau, Y.; Sironi, M.; Assfeld, X. *J. Comput. Chem.* **2006**, *27*, 515−523.

(37) Fornili, A.; Sironi, M.; Raimondo, M. *J. Mol. Struct. (THEOCHEM)* **2003**, *632*, 157−172.

(38) Functions obtained with the ELMO method[37] are already strictly localized.

(39) Subotnik, J. E.; Head-Gordon, M. *J. Chem. Phys.* **2005**, *122*, 034109.

(40) Reed, A. E.; Curtiss, L. A.; Weinhold, F. *Chem. Rev.* **1988**, *88*, 899−926.

(41) Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Structure Theory*; Dover Publications Inc.: Mineola, NY, 1989; p 144.

(42) McDouall, J. J. W. *Theor. Chim. Acta* **1992**, *83*, 339−350.

(43) Stoll, H.; Wagenblast, G.; Preuss, H. *J. Am. Chem. Soc.* **1978**, *100*, 7742−7743.

(44) *Gaussian 03, Revision C.03*; Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. Gaussian, Inc.: Wallingford, CT, 2004.

(45) Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1833−1840.

(46) Rinaldi, D.; Bouchy, A.; Rivail, J.-L.; Dillet, V. *J. Chem. Phys.* **2004**, *120*, 2343−2350.

(47) Rinaldi, D.; Bouchy, A.; Rivail, J.-L. *Theor. Chem. Acc.* **2006**, *116*, 664−669.

CT6003214