ACCOUNTS
of chemical research

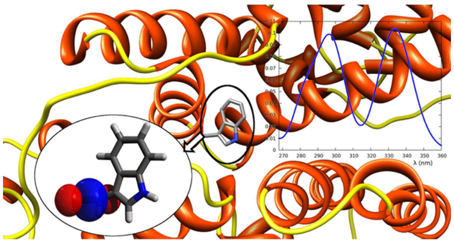
Theoretical Modeling of Large Molecular Systems. Advances in the Local Self Consistent Field Method for Mixed Quantum Mechanics/Molecular Mechanics Calculations

ANTONIO MONARI,[†,‡] JEAN-LOUIS RIVAIL,[†,‡] AND
XAVIER ASSFELD*,[†,‡]

[†]*Théorie Modélisation Simulation, Université de Lorraine, SRSMC UMR 7565,*
Vanduvre-lès-Nancy F-54506, France, and [‡]*Théorie Modélisation Simulation,*
CNRS, SRSMC UMR 7565, Vanduvre-lès-Nancy F-54506, France

RECEIVED ON SEPTEMBER 21, 2012

CONSPECTUS



Molecular mechanics methods can efficiently compute the macroscopic properties of a large molecular system but cannot represent the electronic changes that occur during a chemical reaction or an electronic transition. Quantum mechanical methods can accurately simulate these processes, but they require considerably greater computational resources. Because electronic changes typically occur in a limited part of the system, such as the solute in a molecular solution or the substrate within the active site of enzymatic reactions, researchers can limit the quantum computation to this part of the system. Researchers take into account the influence of the surroundings by embedding this quantum computation into a calculation of the whole system described at the molecular mechanical level, a strategy known as the mixed quantum mechanics/molecular mechanics (QM/MM) approach.

The accuracy of this embedding varies according to the types of interactions included, whether they are purely mechanical or classically electrostatic. This embedding can also introduce the induced polarization of the surroundings. The difficulty in QM/MM calculations comes from the splitting of the system into two parts, which requires severing the chemical bonds that link the quantum mechanical subsystem to the classical subsystem. Typically, researchers replace the quantoclassical atoms, those at the boundary between the subsystems, with a monovalent link atom. For example, researchers might add a hydrogen atom when a C−C bond is cut.

This Account describes another approach, the Local Self Consistent Field (LSCF), which was developed in our laboratory. LSCF links the quantum mechanical portion of the molecule to the classical portion using a strictly localized bond orbital extracted from a small model molecule for each bond. In this scenario, the quantoclassical atom has an apparent nuclear charge of $+1$. To achieve correct bond lengths and force constants, we must take into account the inner shell of the atom: for an sp^3 carbon atom, we consider the two core 1s electrons and treat that carbon as an atom with three electrons. This results in an LSCF$+3$ model. Similarly, a nitrogen atom with a lone pair of electrons available for conjugation is treated as an atom with five electrons (LSCF$+5$). This approach is particularly well suited to splitting peptide bonds and other bonds that include carbon or nitrogen atoms.

To embed the induced polarization within the calculation, researchers must use a polarizable force field. However, because the parameters of the usual force fields include an average of the induction effects, researchers typically can obtain satisfactory results without explicitly introducing the polarization. When considering electronic transitions, researchers must take into account the changes in the electronic polarization. One approach is to simulate the electronic cloud of the surroundings by a continuum whose dielectric constant is equal to the square of the refractive index. This Electronic Response of the Surroundings (ERS) methodology allows researchers to model the changes in induced polarization easily. We illustrate this approach by modeling the electronic absorption of tryptophan in human serum albumin (HSA).

Published on the Web 12/18/2012 www.pubs.acs.org/accounts
10.1021/ar300278j

1. Introduction

The understanding of the role of biomacromolecules (polypeptides, enzymes, nucleic acids) usually requires the knowledge of their electronic structure, of their geometry, and often of their reactivity. This knowledge can only be reached by means of quantum chemical computations. The main challenge, from a quantum chemist's point of view, is their size: often thousands of atoms and their treatments by means of standard quantum chemical codes are still unrealistic. Special computational strategies have to be devised to get the information of interest in such very large systems. This field is at the present time the object of numerous methodological approaches which witness the importance of the question, and several reviews or general papers, focused on various aspects of the problem, have been published.[1−5]

Some recent methodological developments, known as linear scaling,[6−9] made quantum computations affordable on moderately large molecular systems. However, macromolecular systems or solutes are also characterized by an inherent flexibility, showing the possibility to assume different conformations with almost equal weights. Such systems therefore are usually treated by means of molecular dynamics, following the Born−Oppenheimer or the Car−Parrinello[10,11] schemes.

The knowledge of the electronic properties of the whole molecule is not always required to understand the chemical behavior of a large molecular system. The chemically interesting part can be a solute in a solution, a local defect, the active site of an enzyme, or a local chromophore. The rest of the system cannot be discarded, as it interacts with the center of interest by means of mechanical constraints, electrostatic, and induction interactions. The simplest idea is to perform an embedding of the interesting chemical part into its surroundings treated by means of a less demanding method, that is, molecular mechanics. This gives rise to a family of treatments known under the generic name of Quantum Mechanical/Molecular Mechanical (QM/MM) methods.[1] The empirical force fields of molecular mechanics (MM)[12−14] allow the computation of the energy variations of systems following geometry changes. Some more elaborate force fields include a polarizability term which takes explicitly into account the contribution of the induction energy.[15,16]

Within the framework of this methodology, the wave function, energy, and energy derivatives are obtained with the help of a general Hamiltonian which can be written as

$$H = H_{\mathrm{MM}} + H_{\mathrm{QM}} + H_{\mathrm{QM\text{-}MM}} \qquad (1)$$

The usual Fock or Kohn−Sham operators derive from this Hamiltonian.

The interactions occurring can be divided in two kinds of contributions: steric (geometric constrains) and electrostatic. The steric contributions contain the nonbonded van der Waals interactions between MM and QM atoms and, when the MM and the QM parts are chemically bonded, the variation of the bonded terms. When the $H_{\mathrm{QM\text{-}MM}}$ operator represents only this kind of contributions, one speaks of *mechanical embedding*.[2] The electrostatic contributions are mainly the Coulomb interactions between the point charges defined on the MM atoms by the force-field and the electrons and nuclei in the QM region. This constitutes the so-called *electrostatic embedding*.[2] Most of the QM/MM methods, even the modified IMOMM one,[17] consider this embedding to be enough for a correct description of the system.

If the electrostatic properties of the MM part are allowed to change under the influence of the changes of local electric field, one faces the *polarization embedding*.[2] One has to note that the MM point charges are derived in order to reproduce condensed phase properties; that is, they are adjusted in such a way that one can consider that they contain an average of the induction interaction. This is the reason why the electrostatic embedding is generally considered without the polarization refinement. However, when the QM part undergoes a severe electronic rearrangement such as an electronic transition, the charge distribution of the MM fragment should be allowed to adapt itself to this new situation; in this case, the polarization embedding becomes crucial. To treat such effects, one could describe the MM part by using the so-called polarizable force fields that require a more important computational cost. An alternative solution, particularly efficient in the case of electronic transitions, will be described later in this Account.

In such a framework, apart from the case of solutions in which the solute can be considered as not chemically bonded to the solvent, the use of QM/MM methods poses the problem of the treatment of the chemical bonds between the quantum and the classical subsystems. A large variety of strategies have been proposed to allow such dichotomy to be as chemically meaningful as possible, and the detailed expression of the term may vary from one method to another. In this Account, a particular emphasis on the Local Self Consistent Field (LSCF) method developed in our group is given together with original illustrative results.

2. Hybrid Methods for Electronic Properties of the Active Part of a Macromolecule

2.1. The Link or Connection Atom. The simplest idea to separate a chemically significant group of atoms from the macromolecule is to cut covalent carbon−carbon single bonds and to saturate each dangling bond by a hydrogen atom called *link atom*. The first attempts to perform QM/MM computations were built upon this link atom process.[18,19]

Special care is required to parametrize these bonds as well as the electrostatic properties of the classical atoms bonded to the quantum part.[18] Energy and energy derivatives are then readily computed, and the method allows geometry optimizations and molecular dynamics.

The use of hydrogen link atoms is quite popular, especially when the separation between the subsystems is done by cutting carbon−carbon single bonds, such as the $C_\alpha-C_\beta$ bonds in proteins. The presence of link atoms introduces some distortions in the electric field (overpolarization) and some unphysical degrees of freedom which have to be corrected,[20] especially when energy derivatives are needed, for geometry optimizations or molecular dynamics.

To overcome the difficulty introduced by the hydrogen link atom which is located in between the actual carbon atoms of the bond to be cut, the use of an adjusted connection atom, located at the place of the first classical atom, has been proposed.[21] The properties of this connection atom are adjusted in such a way that the bond with a neighboring carbon atom belonging to the quantum subsystem mimics an actual C−C single bond. Regarding the classical subsystem, this atom is assumed to behave like any other similar carbon atom. Several other strategies have been introduced to overcome the inconveniences of the original link atom scheme.[22−24]

An alternative ab initio methodology takes advantages of the pseudopotential theory by adjusting the properties of a pseudocarbon or pseudofluorine atom which replaces the actual carbon atom at the boundary position on the MM subsytem;[23,24] another related approach deals with "effective group potentials",[25] or balanced redistributed charges.[26]

2.2. Frozen Density Approaches. As early as 1976, Warshel and Levitt[27] proposed a semiempirical scheme involving a hybrid orbital engaged in the bond between the quantum and the classical parts of the system. The bond between the quantum and the classical part can be described by a strictly localized bond orbital (SLBO), that is, a localized orbital free of localization tails[28] hosting two electrons. In the LSCF scheme,[29] the localized orbital is extracted from a model small molecule, computed at the same level of theory, and incorporated into the QM/MM scheme. This tool, initially developed at the semiempirical level,[30] has demonstrated its efficiency in several applications such as the study of a peptide enzyme hydrolysis.[31]

The localized molecular orbital representing the bond joining the subsystems can be analyzed as a combination of two hybrid atomic orbitals, one on the quantum atom, one on the classical one, each pointing toward the other atom. Starting from this remark, Gao and co-workers[32,33] developed a method called generalized hybrid orbital (GHO) in which four hybrid atomic orbitals are defined on the MM (quantoclassical) frontier atom. One of the crucial points is the parametrization of the boundary MM atom to allow an integrated force field. This point has been revisited later by introducing charges and dipoles along the classical bonds of the boundary atom.[34]

2.3. The Nonempirical Local Self Consistent Field (LSCF). The first QM/MM method using an ab initio Hamiltonian was based on the LSCF *ansatz*.[35] Solving the LSCF problem means optimizing a monodeterminantal wave function in the orbital approximation knowing that some orbitals are externally imposed, and that these "external" orbitals should remain constant during the optimization procedure, that is, frozen. One has to note that the type of orbitals to be frozen is completely free: they can be monatomic, diatomic, or polyatomic.[36] It is also noteworthy that these frozen orbitals can be occupied or empty.[37,38] Of course, for QM/MM purposes, they are doubly occupied SLBOs. The molecular orbitals of the quantum subsystem have to be kept orthogonal to each SLBO. This can be achieved by an orthogonalization of the basis set to the SLBOs, but owing to the fact that some functions of the set enter the SLBOs, linear dependencies appear between the orthogonalized functions. This inconvenience can be overcome by means of a canonical orthogonalization[39] which yields a set of orthogonal, linearly independent basis functions which can be used to develop the molecular orbitals of the quantum subsystem.

The only other modification one has to take care of is the construction of the total density matrix to build the Fock (or Kohn−Sham) matrix elements. The contribution arising from the frozen orbitals has to be added to the usual density matrix built over the variational orbitals.

This work has been improved by a careful determination of the transformations of the SLBO when its orientation changes, and of its derivatives to allow a full geometry optimization.[40] Nevertheless, it appears that, during this optimization, the length of the frontier bond, that is, the bond linking the quantum to the classical system, is systematically found too short and the shape of the potential energy surface (PES) around the minimum is different from the one obtained by a full QM calculation, whatever the method used to localize the orbital. This defect is interpreted as the underestimation of the interaction between nuclei, since the nuclear charge (*Z*) of the quantoclassical atom, only contributing for one electron to the SLBO, is replaced by a charge $+1$.

The treatment of the inner shell electrons by means of frozen core orbitals[41,42] gives an elegant solution to this problem. It is possible to switch the nuclear charge to $+3$ and add two 1s electrons in the QM system. The core electrons are no longer frozen, and the description of the frontier bond is still correct. This modified LSCF method is called LSCF+3 by contrast to the initial approach called LSCF+1. This scheme has been tested[41] considering the energy variations of the ethane molecule when the C−C bond length is varied in a representation in which one CH_3 group is quantum and one is classical. The LSCF+3 scheme reproduces well the position of the minimum as well as the curvature obtained with the full quantum results, the error on the minimum distance being of the order of 0.1 Å. This scheme has been successfully applied to the CN peptide bond where the C atom is at the MM frontier. It has also been extended to the CN peptide bonds in which the quantoclassical atom is the nitrogen atom.[43] In this case, in addition to the electron involved in the frontier bond and the 1s electrons of the inner shell, one must consider two extravalence electrons in the orbital conjugated with the C=O π orbital. One deals therefore with a LSCF+5 method which allows a symmetric description of the amino acid residues without the arbitrariness of adjusting the classical point charges to obtain an integer value. This procedure is particularly attractive, since it allows one to directly cut through a peptide bond. It has to be noted that link atom approaches which use hydrogen atoms to saturate the dangling bond cannot reproduce this feature.

In Figure 1, we represent the variation of the energy when the distance of the peptide bond is varied in the case of a simple diglycine system. The full QM results, obtained at DFT level with the B3LYP functional and the 6-311G basis set, are compared with the case where one of the peptides is treated at MM level (CHARMM force field). In particular, we compare the situation in which the carbon (LSCF+3) or the nitrogen (LSCF+5) atom is represented as the quantoclassical atom, respectively. Whatever cutting scheme used, the results are in good agreement with the full QM case, both for the position of the minimum and for the curvature.

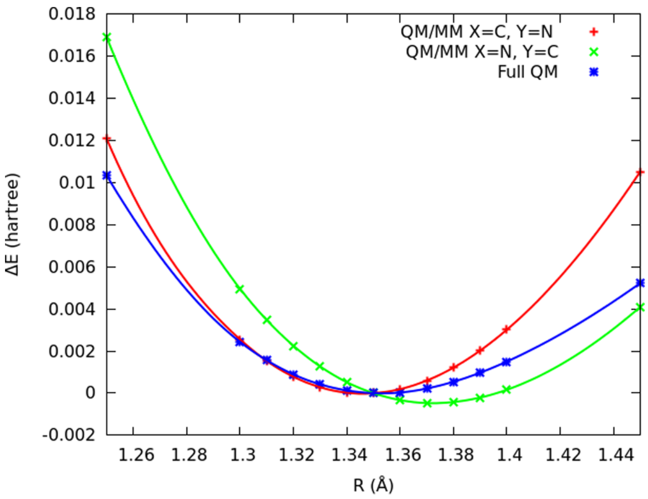


FIGURE 1. Variation of the relative energy of a capped diglycine peptide with the C−N interatomic distance for a full DFT (B3LYP/6-311G) computation and for the QM/MM (CHARMM force field). X represents the quantum and Y the quantoclassical frontier atom.

The maximum deviation of the equilibrium distance from the full QM results is less than 0.02 Å, the force constant being also well reproduced. The previous results can all be considered comparable to the uncertainty introduced by using a classical force field. The same trends are confirmed by the QM/MM optimization of the polypeptide composed of three glycine amino acids where only the central residue is treated at the QM level. Note that in this case the QM/MM carbon frontier atom is treated with the LSCF+3 and the nitrogen one with the LSCF+5 scheme. The equilibrium distance values for both peptide bonds are exactly the same as the ones obtained for the previous case, and the values of the dihedral angles around the amide bond are about 179° for both schemes. This implies that the planar structure of the peptide bond, due to the partial double bond character, is also correctly reproduced.

These results justify cutting the peptide bond to perform QM/MM geometry optimization and even a QM/MM molecular dynamics study. The extension to other heteroatoms would probably require the consideration of more electrons on the quantoclassical atom. We recall that a $+11$ nuclear charge was already used for Si atoms together with a pseudopotential to describe the 10 core electrons.[44]

The LSCF scheme rests on the use of strictly localized molecular orbitals which are obtained by using a standard localization procedure[45] after removing the usual tails of orthogonalization and delocalization, and renormalizing. The use of different localization methods has been considered[46] and compared with the use of extremely localized molecular orbitals (ELMOs).[47,48] This study confirms the reliability of the method, since the different localization methods behave quite similarly, even if the ELMOs give the best results.

There are still few possible improvements: As an application of mixing bonding and antibonding bond orbitals,[28] an Optimized LSCF method in which each SLBO is allowed to mix with its corresponding Strictly Localized Anti Bonding Orbital has been developed successfully.[49]

2.4. Polarization Embedding. All the above methods are developed with usual molecular mechanics softwares, which do not include polarizability. Nowadays several force fields add induced polarization effects to the usual electrostatic interactions,[15,16] as does the X-Pol method,[50] and can be, if necessary, easily used in the previous QM/MM schemes. Since most of the standard force fields are now perfectly validated for modeling ground state properties, there are only few attempts to take polarization embedding into account. There is however a situation in which electronic polarization is expected to play an important role: the study of vertical electronic transitions in chromophores or solutes.[51−53] The induced electronic moments of the surrounding are able to follow the fast changes of the solute's properties, contrarily to electrostatic interactions due to the charges or permanent moments on the atoms which can be considered as frozen during the electronic transition (Franck−Condon principle). A simple alternative to the use of a polarizable force field consists in modeling the induced polarization by a nonpolar, polarizable continuum characterized by a dielectric constant extrapolated at infinite frequency. The quantum subsystem is assumed to be placed in a cavity surrounded by this continuum. The use of this simple model is only justified if one wants to evaluate the variations of the induction energy when the geometry of the system does not change. Vertical spectroscopic transitions fulfill these conditions. It would, however, not be justified to claim that one reaches the absolute value of the induction energy, in particular, because this energy is already taken into account, at least partly, by using a modern force field. This methodology named "Electronic Response of the Surroundings" (ERS)[54] has been applied successfully to the modeling of absorption spectra and circular dichroism of biological macromolecules,[55−58] showing that inclusion of electronic polarizability has non-negligible effects on these spectroscopic properties. It is also noteworthy that, for the ground state, our QM/MM calculations within the EE or PE schemes give quite similar results. Indeed, as shown by Very et al.,[57] if one considers the QM atomic charges, the variation induced by ERS is only 0.1%, while the effect of changing the classic force field induces a variation of about 5%.

2.5. Absorption Spectrum of Tryptophan in Solution and Embedded in a Protein. To illustrate the influence of electrostatic and polarization effects on the computation of absorption spectra of chromophores embedded in macromolecular environments, we report the QM/MM computation of the UV/vis spectrum of tryptophan in aqueous solution and in the human serum albumin (HSA) protein. Tryptophan is known to have an intense absorption band around 270 nm and is widely studied[59,60] due to its emission spectrum allowing its use as a fluorescence probe.

In the case of the chromophore in solution, we considered *N*-acetyl-tryptophan. The solute was embedded in a water box of 30 Å side. A molecular dynamics trajectory of 2 ns was performed using periodic boundary conditions in the constant temperature and pressure ensemble (NPT). The system was modeled using the amber99 force field. Independent snapshots were extracted, and UV/vis spectra were computed by QM/MM. *N*-Acetyl-tryptophan was treated at Time Dependent-DFT (TDDFT) using the PBE0 functional and the 6-311++G(d,p) basis set. Mechanical embedding (ME), electrostatic (mainly electrostatic plus ground state polarization) embedding (EE), and polarization embedding (using the ERS methodology) (PE) were considered. In Figure 2, we report the spectrum arising from the convolution of the vertical transition energies for all the snapshots, the convolution having been done with Gaussian functions of fixed half-length-width of 0.2 eV.

The three embeddings give quite similar results, and are all able to represent the absorption at 270 nm, as well as the less intense one at 250 nm. PE has an important effect on intensity that appears, in that case, higher. The relatively small importance of electrostatic and polarization effects is probably due to the average over different solvent conformation so that the individual different electrostatic and polarization effects tend to cancel in the final average. Our results confirm very well the experimental ones,[59] both for the position of the peaks and for the general shape of the bands, hence validating the QM/MM procedure.

Much more striking effects can be seen on the absorption spectrum of HSA. Indeed, this quite important protein (Figure 3) has only one tryptophan in its primary sequence. The chromophore is part of an α-helix, and because of the combination of secondary and tertiary structures it is found in an internal region of the protein. Moreover, it has four charged amino acids in its vicinity (Figure 3): two lysines, one arginine, and one aspartic acid. Other more polarizable amino acids such as phenylalanine and leucine are also found in the vicinity of the chromophore.

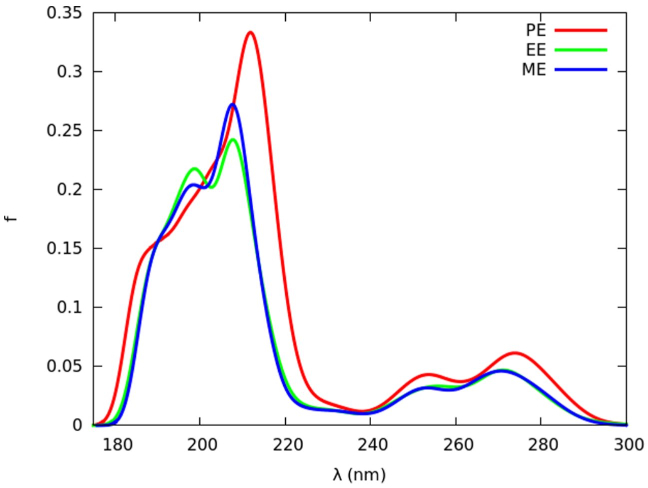


FIGURE 2. Convoluted QM/MM spectra of *N*-acetyl-tryptophan in water solution.

Due to the relative rigid nature of the protein, we performed a QM/MM optimization starting from the NMR resolved structure. The QM region, composed of the lateral chain of tryptophan, has been treated at the DFT level with the 6-311++G(d, p) basis set and the PBE0 functional. The rest of the protein has been treated at the MM level with the amber99 force field. The frontier has been represented by a SLBO via the LSCF approach. Once the geometry is optimized, the UV/vis spectrum is obtained at TDDFT level as vertical transitions.

The spectrum, convoluted with Gaussian functions, is reported for ME, EE, and PE models in Figure 4. Contrarily to the previous case, here the intensity of the electrostatic and polarization effects is certainly noteworthy. When using the ME model, the spectrum bears a strong resemblance with the spectrum in solution; when including the electrostatic effects, one experiences an impressive red-shift with now two very well-defined, and intense, maxima at 301 and 345 nm. The inclusion of ERS, and therefore of polarization effects on the electronic transition, shifts again the maxima, appearing at 297 and 333 nm. Electrostatic effects account for a shift of almost 100 nm, and are probably due to the charged amino acids close to the chromophore. Moreover, α-helices are known to have a very strong dipole moment[61] that adds to the electrostatic effects due to the charges. Polarization effects go in the opposite direction from electrostatic ones and contribute to diminish the total red-shift. This occurrence cannot be generalized, since many cases are observed in which polarization effects actually go in the same direction as electrostatic ones.

The previous results are difficult to compare to the experimental ones, since in the UV/vis spectrum of HSA the effects of the tryptophan residue are masked by the presence of 10 tyrosines that absorb in the same region of the spectrum. Moreover, quite often the available experimental spectra either are of low resolution or are referred to the denaturated protein.[59] Anyway, our results seem qualitatively coherent with the red-shift observed for instance by Goswani et al.;[60]

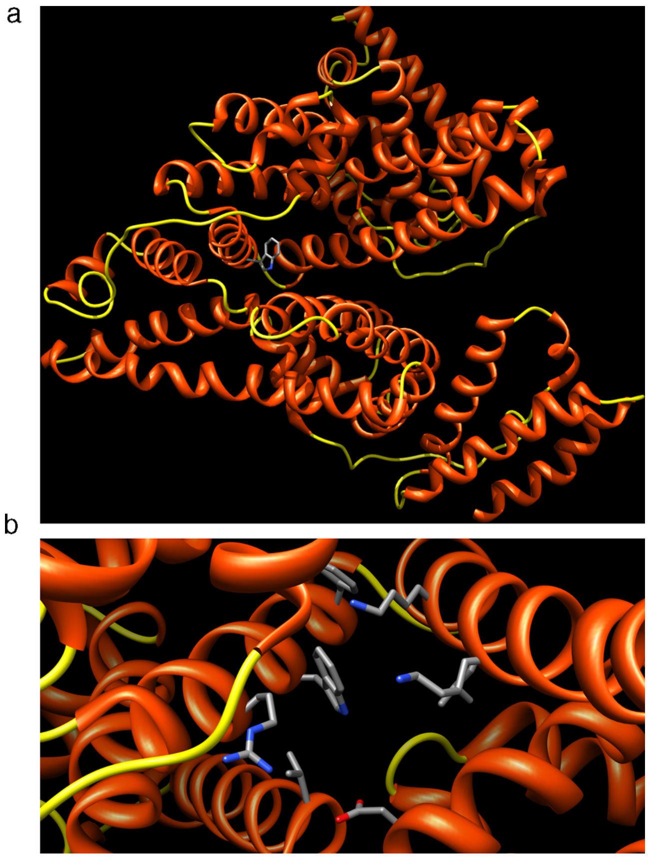


FIGURE 3. (a) Secondary and tertiary structure of HSA. (b) Zoom of the chromophore region.

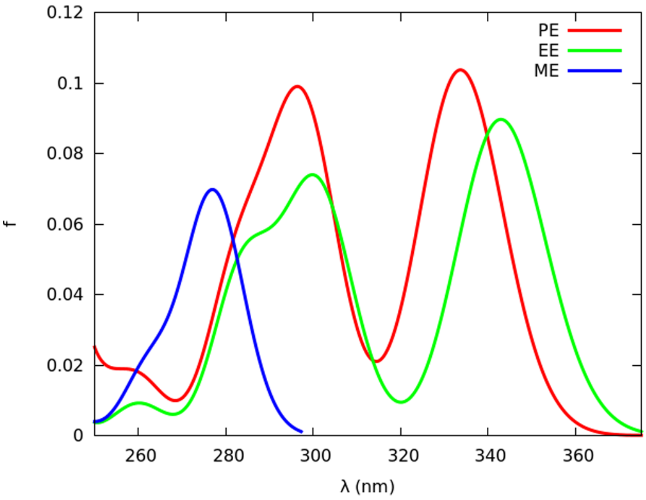


FIGURE 4. Convoluted QM/MM spectrum of tryptophan embedded in the HSA.

the red-shift is also coherent to the fact that fluorescence spectra are often observed upon excitation between 300 and 350 nm. This study clearly confirms that all the three effects have to be taken into account in order to have a reliable description of the absorption spectra. ERS methodology appears as a cheap and efficient way to take polarization into account in the QM/MM calculation of vertical excitation energies.

3. Perspectives

The QM/MM methods have been developed mainly to address the important problems of structure and reactivity of biological macromolecular systems and condensed-phase small-molecule systems. They also have interesting applications to other important chemical or physical problems such as surface reactivity of solids[62] or reactivity in zeolites.[44] Some other physical properties of macromolecules can be addressed by means of this methodology. An example is provided by electronic conductivity of peptide sequences.[63] Finally, one can mention a particularity of the LSCF approach in which the localized electronic states allow the study of core ionization and core electronic excitation in macromolecules.[38] As a conclusion, one can anticipate that the QM/MM methods will remain the adequate methodology for many kinds of applications concerning macromolecular systems.

BIOGRAPHICAL INFORMATION

Antonio Monari, born in 1976 in Bologna (Italy), obtained is Ph. D. degree in theoretical chemistry in 2007 from the University of Bologna with a thesis on the computation of polarizabilities and intermolecular forces. During his postdoc fellowships at the universities of Bologna and Toulouse, he was mainly interested in the development and applications of wave-function methods to study quasi-degenerate excited states and electron transfer phenomena. Since 2010, he is assistant professor at the Université de Lorraine in Nancy (France). His main interest since then has been devoted to mixed QM/MM methods to model environment effects on excited states and environment controlled light induced phenomena. The main applications are related to the domain of photo- and electron-active systems, dye sensitized solar cells, or to the interaction of organic and inorganic chromophores with biological macromolecules.

Jean-Louis Rivail, born in 1937 in la Mure (France), is emeritus professor of theoretical chemistry at the Université de Lorraine since 2003. As early as 1973, he published a pioneer work on the use of continuum models in the treatment of solvent effects in quantum chemistry and generalized this approach by introducing the concept of Self Consistent Reaction Field whose several variants are now widely used in many computational codes. The Local Self Consistent Field method for the Quantum Mechanical/Classical Mechanical treatment of very large molecular systems was introduced, first at the semiempirical level and extended to the ab initio and DFT approaches which received several improvements as described in this Account.

Xavier Assfeld, born in 1966 in Nancy (France), is full professor at the Université de Lorraine since 2003. He obtained his Ph.D. degree in Theoretical Chemistry in 1994 from the Université Henri Poincaré. His thesis was dedicated to chemical reactivity in solution. He published the first article dealing with the localization of nontrivial transition states in solution at the ab initio level in 1992. During his postdoctoral stay at the University of Minnesota, he worked with Professor Donald Truhlar on degeneracy corrected perturbation theory. He obtained an assistant professor position in 1995 at the University Henri Poincaré and started working on the development of hybrid QM/MM methods at the ab initio level, especially the Local Self-Consistent Field method in collaboration with Professor Jean-Louis Rivail. Lately, his research interests are turned toward describing excited states (core or valence) of macromolecules taking into account the electronic response of the surroundings as detailed in this Account.

FOOTNOTES

*To whom correspondence should be addressed. E-mail: Xavier.Assfeld@univ-lorraine.fr.

The authors declare no competing financial interest.

REFERENCES

1 Monard, G.; Merz, K. M., Jr. Combined Quantum Mechanical/Molecular Mechanical Methodologies Applied to Biomolecular Systems. *Acc. Chem. Res.* **1999**, *32*, 904–911.

2 Lin, H.; Truhlar, D. G. QM/MM: what we have learned, where are we, and where do we go from here? *Theor. Chem. Acc.* **2007**, *117*, 185–199.

3 Senn, H. M.; Thiel, W. QM/MM Methods for biomolecular systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198–1229.

4 Gordon, M. S.; Fedorov, D. G.; Pruitt, S. R.; Slipchenko, L. V. Fragmentation Methods: A Route to Accurate Calculations on Large Molecular Systems. *Chem. Rev.* **2012**, *112*, 632–672.

5 Senn, H. M.; Thiel, W. QM/MM Methods for biomolecular systems. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198–1229.

6 Dixon, S. L.; Merz, K. M., Jr. Semiempirical Molecular-Orbital Calculations with Linear System Size Scaling. *J. Chem. Phys.* **1996**, *104*, 6643–6649.

7 Daniels, A. D.; Millam, J. M.; Scuseria, G. E. Semiempirical Methods with Conjugate Gradient Density Matrix Search to Replace Diagonalization for Moleculas Systems Containing Thousand of Atoms. *J. Chem. Phys.* **1997**, *107*, 425–431.

8 Yang, W.; Lee, T. S. A Density Matrix Divide-and-Conquer Approach for Electronic Structure Calculations of Large Molecules. *J. Chem. Phys.* **1995**, *103*, 5674–5678.

9 Monard, G.; Bernal-Uruchurtu, M. I.; van der Vaart, A.; Merz, K. M., Jr.; Ruiz-López, M. F. Simulation of Liquid Water Using Semiempirical Hamiltonians and the Divide and Conquer Approach. *J. Phys. Chem. A* **2005**, *109*, 3425–3432.

10 Car, R.; Parrinello, M. Unified Approach for Molecular Dynamics and Density-Functional Theory. *Phys. Rev. Lett.* **1985**, *55*, 2471–2474.

11 Carloni, P.; Rothlisberger, U.; Parrinello, M. The Role and Perspective of Ab Initio Molecular Dynamics in the Study of Biological Systems. *Acc. Chem. Res.* **2002**, *35*, 455–464.

12 van Gunsteren, W. F.; Berendsen, H. J. C. *Groningen Molecular Simulation (GROMOS) Library Manual*; Biomos: Groningen, 1987 .

13 Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E., III; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate structural and energetic properties of molecules. *Comput. Phys. Commun.* **1995**, *91*, 1–41.

14 MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, W.; Wiórkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

15 Dehez, F.; Chipot, C.; Millot, C.; Ángyán, J. G. Fast and accurate determination of induction energies: reduction of topologically distributed polarizability models. *Chem. Phys. Lett.* **2001**, *338*, 180–188.

16 Dupuis, M.; Aida, M.; Kawashima, Y.; Hirao, K. A polarizable mixed Hamiltonian model of electronic structure for micro-solvated excited states. I Energy and gradients formulation and application to formaldehyde ($^1A_2$). *J. Chem. Phys.* **2002**, *117*, 1242–1255.

17 Maseras, F.; Morokuma, K. IMOMM, a new integrated ab initio + molecular mechanics geometry optimization scheme of equilibrium structure and transition states. *J. Comput. Chem.* **1995**, *16*, 1170–1179.

18 Singh, U. C.; Kollman, P. A. A combined ab initio quantum-mechanical and molecular mechanical method for carrying out simulations on complex molecular-systems. Applications to the $CH_3Cl + Cl^-$ exchange-reaction and gas-phase protonation of polyethers. *J. Comput. Chem.* **1986**, *7*, 718–730.

19 Eurenius, K. P.; Chatfield, D. C.; Brooks, B. R. Enzyme Mechanisms with Hybrid Quantum and Molecular Mechanical Potentials. I. Theoretical Considerations. *Int. J. Quantum Chem.* **1996**, *60*, 1189–1200.

20 Ferré, N.; Olivucci, M. The amide bond: pitfalls and drawback of the link atom scheme. *J. Mol. Struct.: THEOCHEM* **2003**, *632*, 71–82.

21 Antes, I.; Thiel, W. Adjusted Connection Atoms for Combined Quantum Mechanical and Molecular Mechanical Methods. *J. Phys. Chem. A* **1999**, *103*, 9290–9295.

22 Swart, M. AddRemove: A New Link Model for Use in QM/MM Studies. *Int. J. Quantum Chem.* **2003**, *91*, 177–183.

23 DiLabio, G. A.; Hurley, M. M.; Christiansen, P. A. Simple one-electron quantum capping potentials for use in hybrid QM/MM studies of biological molecules. *J. Chem. Phys.* **2002**, *116*, 9578–9584.

24 Xiao, C.; Zhang, Y. Design-atom approach for quantum mechanical/molecular mechanical covalent boundary: A design-carbon atom with five valence electrons. *J. Chem. Phys.* **2007**, *127*, 124102.

25 Poteau, R.; Ortega, Y.; Alary, F.; Ramirez Solis, A.; Barthelat, J.-C.; Daudey, J. P. Effective Group Potentials. 1. Method. *J. Phys. Chem. A* **2001**, *105*, 198–205.

26 Wang, B.; Truhlar, D. G. Combined Quantum Mechanical and Molecular Mechanical Methods for Calculating Potential Energy Surfaces: Tuned and Balanced Redistributed-Charge Algorithm. *J. Chem. Theor. Comput.* **2010**, *6*, 359–369.

27 Warshel, A.; Levitt, M. Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* **1976**, *103*, 227–249.

28 Ferenczy, G. G.; Rivail, J.-L.; Surján, P. R.; Náray-Szabó, G. NDDO fragment self-consistent-field approximation for large electronic systems. *J. Comput. Chem.* **1992**, *13*, 830–837.

29 Théry, V.; Rinaldi, D.; Rivail, J.-L.; Maigret, B.; Ferenczy, G. G. Quantum mechanical computations on very large molecular systems: The local self-consistent field method. *J. Comput. Chem.* **1994**, *15*, 269–282.

30 Monard, G.; Loos, M.; Théry, V.; Baka, K.; Rivail, J. L. Hybrid Classical Quantum Force Field for Modeling Very Large Molecules. *Int. J. Quantum Chem.* **1996**, *58*, 153–159.

31 Antonczak, S.; Monard, G.; Ruiz-López, M. F.; Rivail, J.-L. Modeling of Peptide Hydrolysis by Thermolysin. A semiempirical and QM/MM Study. *J. Am. Chem. Soc.* **1998**, *120*, 8825–8833.

32 Gao, G.; Amara, P.; Alhambra, C.; Field, M. J. A Generalized Hybrid Orbital (GHO) Method for the Treatment of Boundary Atoms in Combined QM/MM Calculations. *J. Phys. Chem. A* **1998**, *102*, 4714–4721.

33 Pu, J.; Gao, J.; Truhlar, D. G. Combining Self-Consistent-Charge Density-Functional Tight-Binding (SCC-DFTB) with Molecular Mechanics by the Generalized Hybrid Orbital (GHO) Method. *J. Phys. Chem. A* **2004**, *108*, 5454–5463.

34 Lin, H.; Truhlar, D. G. Redistributed Charge and Dipole Schemes for Combined Quantum Mechanical and Molecular Mechanical Calculations. *J. Phys. Chem. A* **2005**, *109*, 3991–4004.

35 Assfeld, X.; Rivail, J.-L. Quantum chemical computations on parts of large molecules: the ab initio local self consistent field method. *Chem. Phys. Lett.* **1996**, *263*, 100–106.

36 Moreau, Y.; Loss, P.-F.; Assfeld, X. Solvent effects on the asymmetric Diels−Alder reaction between cyclopentadiene and (−)-menthyl acrylate revisited with the three-layer hybrid local self-consistent field/molecular mechanics/self-consistent reaction field method. *Theor. Chem. Acc.* **2004**, *112*, 228–239.

37 Ferré, N.; Assfeld, X. Application of the local self-consistent-field method to core-ionized and core-excited molecules, polymers, and proteins: True orthogonality between ground and excited states. *J. Chem. Phys.* **2002**, *117*, 4119–4125.

38 Loos, P. F.; Assfeld, X. Core-ionized and core-excited states of macromolecules. *Int. J. Quantum Chem.* **2007**, *107*, 2243–2252.

39 Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry*; McGraw Hill: New York, 1989, p 144.

40 Ferré, N.; Assfeld, X.; Rivail, J.-L. Specific Force Field Determination for the Hybrid Ab Initio QM/MM LSCF Method. *J. Comput. Chem.* **2002**, *23*, 610–624.

41 Fornili, A.; Loos, P.-F.; Sironi, M.; Assfeld, X. Frozen core orbitals as an alternative to specific frontier bond potential in hybrid Quantum Mechanics/ Molecular Mechanics methods. *Chem. Phys. Lett.* **2006**, *427*, 236–240.

42 Loos, P.-F.; Fornili, A.; Sironi, M.; Assfeld, X. Removing extra frontier parameters in QM/MM methods: a tentative with the Local Self-Consistent Field Approach. *Comput. Lett.* **2007**, *4*, 473–486.

43 Loos, P.-F.; Assfeld, X. On The Frontier Bond Location In The QM/MM Description Of Peptides and Proteins. *AIP Conf. Proc.* **2007**, *963*, 308–315.

44 Ferré, N.; Assfeld, X. A new three-layer hybrid method (LSCF/MM/Madelung) devoted to the study of chemical reactivity in zeolites. Preliminary results. *J. Mol. Struct.: THEOCHEM* **2003**, *632*, 83–90.

45 Jensen, F. *Introduction to Computational Chemistry*; J. Wiley & Sons: Chichester, 2001; pp 227−229.

46 Fornili, A.; Moreau, Y.; Sironi, M.; Assfeld, X. On the Suitability of Strictly Localized Orbitals for Hybrid QM/MM Calculations. *J. Comput. Chem.* **2006**, *27*, 515–523.

47 Sironi, M.; Famulari, A. An orthogonal approach to determine extremely localised molecular orbitals. *Theor. Chem. Acc.* **2000**, *103*, 417–422.

48 Genoni, A.; Ghitti, M.; Pieraccini, S.; Sironi, M. A novel extremely localized molecular orbitals based technique for the one-electron density matrix computation. *Chem. Phys. Lett.* **2005**, *415*, 256–260.

49 Loos, P. F.; Assfeld, X. Self-Consistent Strictly Localized Orbitals. *J. Chem. Theory Comput.* **2007**, *3*, 1047–1053.

50 Xie, W.; Song, L.; Truhlar, D. G.; Gao, J. The variational explicit polarization potential and analytical first derivative of energy: Towards a next generation force field. *J. Chem. Phys.* **2008**, *128*, 234108.

51 Marenich, A. V.; Cramer, C. J.; Truhlar, D. G.; Guido, C. A.; Mennucci, B.; Scalmani, G.; Frisch, M. J. Practical computation of electronic excitation in solution: vertical excitation model. *Chem. Sci.* **2011**, *2*, 2143–2161.

52 Jacquemin, D.; Mennucci, B.; Adamo, C. Excited-state calculations with TD-DFT: from benchmarks to simulations in complex environments. *Phys. Chem. Chem. Phys.* **2011**, *13*, 16987–16998.

53 Mennucci, B.; Cappelli, C.; Guido, C. A.; Cammi, R.; Tomasi, J. Structures and properties of electronically excited chromophores in solution from the polarizable continuum model coupled to the TD-DFT. *J. Phys. Chem. A* **2009**, *113*, 3009–3020.

54 Jacquemin, D.; Perpète, E. A.; Laurent, A. D.; Assfeld, X.; Adamo, C. Spectral properties of self-assembled squaraine-tetralactam: A theoretical assessment. *Phys. Chem. Chem. Phys.* **2009**, *11*, 1258–1262.

55 Laurent, A. D.; Assfeld, X. Effect of the Enhanced Cyan Fluorescent Proteinic Framework on the UV/visible absorption spectra of some chromophores. *Interdisc. Sci. Comput. Life Sci.* **2010**, *2*, 38–47.

56 Monari, A.; Very, T.; Rivail, J.-L.; Assfeld, X. Effects of mutations on the absorption spectra of copper proteins: a QM/MM study. *Theor. Chem. Acc.* **2012**, *131*, 1221.

57 Very, T.; Despax, S.; Hébraud, P.; Monari, A.; Assfeld, X. Spectral properties of polypyridyl Ruthenium complex intercalated in DNA: theoretical insights on the surrounding effects for $[Ru(dppz)(bpy)_2]^2$. *Phys. Chem. Chem. Phys.* **2012**, *14*, 12496–12504.

58 Chantzis, A.; Very, T.; Monari, A.; Assfeld, X. Improved Treatment of Surrounding Effects: UV/vis Absorption Properties of a Solvated Ru(II) Complex. *J. Chem. Theory Comput.* **2012**, *8*, 1536–1541.

59 Lakowicz, J. R. *Principle of Fluorescence Spectroscopy*; Springer: 2006.

60 Goswani, N.; Makhal, A.; Kumar Pal, S. Toward an Alternative Intrinsic Probe for Spectroscopic Characterization of a Protein. *J. Phys. Chem. B* **2010**, *114*, 15236–15243.

61 Dumont, E.; Loos, P.-F.; Laurent, A. D.; Assfeld, X. Huge disulfide-linkage's electron capture variation induced by α-helix orientation. *J. Chem. Theory Comput.* **2008**, *4*, 1171–1173.

62 DiLabio, G. A.; Wolkow, R. A.; Johnson, E. R. Efficient silicon surface and cluster modelling using quantum capping potentials. *J. Chem. Phys.* **2005**, *122*, 044708.

63 Rodriguez-Ropero, F.; Zanuy, D.; Assfeld, X.; Alemán, C. Modeling an Electronic Conductor Based on Natural Peptide Sequences. *Biomacromolecules* **2009**, *10*, 2338–2343.