Regular article

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

Yohann Moreau, Pierre-François Loos, Xavier Assfeld

Equipe de Chimie et de Biochimie théoriques, UMR CNRS-UHP 7565, BP 239, 54506
Vandoeuvre-lès-Nancy Cedex, France

Abstract. The construction of the three-layer hybrid local self-consistent field/molecular mechanics/self-consistent reaction field method is detailed. This method is specifically devoted to the study of the reactivity of large chemical systems in solution. The solvent, modeled by a polarizable continuum, surrounds the whole solute molecule. Solute–solvent interactions are taken into account by means of the self-consistent reaction field approach. The solute system is treated by both quantum and molecular mechanics, the former being principally applied to the reactive part, i.e., the part undertaking bond forming or breaking, the latter being reserved for the ancillary encumbering groups. The connection between the molecular mechanics and the quantum mechanics part is accomplished by a strictly localized bond orbital that remains frozen within the local self-consistent field framework. As a test system, the asymmetric Diels–Alder reaction between cyclopentadiene and (−)-menthyl acrylate is studied for the first time with steric interactions and electrostatic solvent effects taken into account simultaneously. The results indicate that the coupling of both interactions leads to conclusions that could not have been guessed from separate calculations.

Keywords: Solvent effect – QM/MM – Diels-Alder – Ab initio – combined method

Introduction

For a long time theoretical chemistry was mainly devoted to explaining experimental results by means of methods based on crude approximations of the Schrödinger equation and/or of the chemical systems under study. Nowadays, one can ask for predictions made on accurate values obtained for realistic chemical systems. This change was made possible, of course, by the enormous increase in computer power and by the development of better algorithms, and also by new types of methods. Among the latter, hybrid quantum mechanics (QM)/molecular mechanics (MM) methods have had wide success in very different areas of chemistry, like biochemistry and homogeneous and heterogeneous catalysis. In a few words, one can say that QM/MM methods take advantage of the very local character of chemical processes. A small fragment of the large system is described by QM, to accurately represent electronic phenomena, and the remaining part of the system is treated by means of MM. Evidently, the computational time savings are tremendous if compared with full QM methods. However, owing to the very large number of degrees of freedom one faces the problem of getting free energies instead of potential energies, specifically for solvation processes.

Our aim is then to develop a tool able to correctly model the chemical reactivity of relatively large molecular systems in solution. Our proposal consists in a three-layer hybrid method combining QM for the treatment of the reactive part, MM for the explicit description of encumbering substituents, and an accurate solvent model that gives free energies of solvation. As detailed later, all the tools we need have already been developed by Rivail and coworkers.

All through his carrier, Rivail has concentrated his research activities on introducing surroundings effects in quantum chemical calculations. Among others, the first ever published paper—30 years ago!—combining the Schrödinger and Poisson equations to take solvent effects into account in quantum chemical calculations [1], leading to the famous self-consistent reaction field

(SCRF) method, is certainly his most renowned achievement. Furthermore, one has to point out his more recent interest in macromolecular systems for which he developed the first hybrid QM/MM method at the ab initio level [2], known as the local self-consistent field (LSCF) method. Indeed, Rivail has provided us with the apparatus we need to investigate solvent effects on large chemical systems by means of theoretical chemistry.

In this article we outline the coupling of the ab initio LSCF/MM approach [2, 3, 4, 5, 6] with the SCRF model [1, 7], resulting in the three-layer LSCF/MM/SCRF hybrid method. It is illustrated by studying the modification induced by the solvent on the asymmetric Diels–Alder reaction between cyclopentadiene and (−)-menthyl acrylate. We chose this reaction, already studied 10 years ago [8] with methyl acrylate as a model molecule, since it was the first ever localized transition state (TS) in solution, of course with Rivail's model. The fact that solvent effects and substituent effects have never been studied simultaneously is another good reason to choose this reaction as the first candidate on which our freshly developed method can be applied.

It has been shown [8] that solvent effects, specifically those due to electrostatic solute–solvent interactions, play a significant role in diastereofacial selectivity. Nevertheless, the encumbering menthyl functional group was modeled by the hindrance of the free methyl group. Since the electronic inductions of both groups are certainly close to each other, one can hope that the methyl fragment correctly mimics the electronic effects of the menthyl substituent. However, the methyl group can by no means model the steric hindrance of the menthyl moiety, and one really needs to explicitly describe all atoms of the reactants to obtain accurate geometries of the various diastereoisomers. One has also to notice that it was assumed that the isopropyl functional group of the menthyl fragment, presenting sufficient steric repulsion with the approaching cyclopentadiene, governs the diastereoselectivity of the reaction. However, it has been shown, using a QM/MM approach, that depending on the conformation of the isopropyl group, the steric interactions can be neglected [9]. The latter conclusions are based on gas-phase calculations, and it is known that solvent effects can have a quite important influence. For example, water, as a "green chemistry" solvent, presents numerous features, like enhancing the chemical selectivities (endo/exo or diastereoselectivity) and increasing the reaction rates [10, 11]. One invokes generally, to explain these peculiar behaviors, the three following reasons: solvophobic interactions, solvent polarity and hydrogen-bonding interactions. Owing to the Curtin–Hammet principle [12], to study selectivity, one is concerned merely with the relative stability of the various TSs. One can hope that for the system under investigation, the solvophobic and hydrogen-bonding variations from one TS to the other are less important than the polarization term [8].

One has to mention that some other works [13, 14, 15] have shown the usefulness of QM/MM methods to study Diels–Alder reactions, although semiempirical Hamiltonians were used.

Solvent effects and the effects of steric hindrance on the asymmetric Diels–Alder reaction between cyclopentadiene and (−)-menthyl acrylate are discussed in Sect. 4 with the new three-layer LSCF/MM/SCRF method described in Sect. 2. Moreover, also in Sect. 4, the LSCF method is used in a qualitative manner to analyze the influence of the solvent on the wave function of the TSs. Computational details are collected in Sect. 3.

Methodology

The LSCF formalism

Although the LSCF formalism was not developed a priori for this purpose[2], it appears that it is particularly well suited to the optimization of constrained wave functions, within the orbital approximation. The constrains applied to the total wave function that the LSCF formalism is able to handle are that some predefined orbitals are known and frozen (not allowed to vary). The mathematical details of this method have already been given [2, 3, 4, 5, 6]. Nevertheless, to ease the reading of this article, some essential points are detailed in the following.

One starts with the expression of the orbitals that must remain constants. Let us call them frozen orbitals (FOs):

$$\theta_l = \sum_{\mu}^{K} b_{\mu l} \phi_{\mu}, \tag{1}$$

where θ_l is the FO number l, ϕ_{μ} is a basis function belonging to a set of K functions, and $b_{\mu l}$ are the usual coefficients that satisfy the normality condition

$$\langle \theta_l | \theta_l \rangle = \sum_{\mu}^{K} \sum_{\nu}^{K} b_{\mu l}^* S_{\mu \nu} b_{\nu l} = 1, \tag{2}$$

with $S_{\mu \nu}$ the overlap matrix element between functions φ_{μ} and φ_{ν}.

Our aim is to find the remaining orbitals of the system that minimize the electronic energy, as the usual Hartree–Fock (HF) method does, for example. To avoid the factorial number of terms one has to compute with nonorthogonal orbitals and to take the Pauli repulsion into account we ask that the variationally optimized molecular orbitals be orthogonal to the FOs. Using the traditional method with Lagrange multipliers one can find modified Roothaan equations [3, 6].[1] An easy way to solve these complicated equations is to proceed as follows.

---

[1] The original basis set can already possess linear dependencies

For practical reasons, the L FOs are mutually orthogonalized by means of the Löwdin symmetric method [17], leading to orthogonal FOs (OFOs) ϕ_l:

$$\varphi_l = \sum_{\mu}^{K} a_{\mu l} \phi_\mu. \tag{3}$$

Each basis function, $|\phi_\mu\rangle$, is then projected out of the subspace span by the OFOs, leading to a set of K functions, $|\hat{\phi}_\mu\rangle$, orthogonal to the OFOs, but containing at least L linear dependencies:

$$|\tilde{\phi}_\mu\rangle = N_\mu \left(1 - \sum_l^L |\varphi_l\rangle\langle\varphi_l|\right) |\phi_\mu\rangle, \tag{4}$$

N_μ being a normalization factor. The linear dependencies are removed by means of a canonical orthogonalization procedure [18], leading to a set of $(K–L)$ functions mutually orthogonal and orthogonal to the OFOs. These two last steps can be combined in a unique matrix operating on the original basis set to give the final set. This rectangular matrix is then used in place of the usual orthogonalization matrix in a self-consistent-field calculation.

One has to note that no restriction has been put on the occupation number of the FOs. Neither have we imposed a restriction to work within the restricted HF (RHF) or unrestricted HF formalisms. Then our method can be applied for RHF and unrestricted HF calculations with FOs containing no, one or two electrons [4]. The only point one has to take care of is the density matrix associated with the OFOs.

Straightforwardly, LSCF calculations are feasible within density functional theory as long as one uses the Kohn–Sham formalism [5]. Moreover, any type of post-HF methods, based on orbitals, can be applied to LSCF wave functions [4].

The QM/MM coupling

There are two kinds of QM/MM methods: those where the interactions between the MM and the QM parts are essentially physicals, i.e., intermolecular in essence, as is the case for a solute in a solvent, and those where they are chemicals, i.e., intramolecular, as in a macromolecule. The latter kind are the more difficult to define because covalent bonds need to be cut to isolate the QM center from the surroundings, leading to the well-known dangling bond problem. Many solutions have been proposed to circumvent this problem, and LSCF is one of those. Some features are discussed hereafter.

Each covalent bond that is cut to separate the MM and the QM subsystems is replaced by a strictly localized bond orbital (SLBO). This SLBO, containing two electrons, is of course frozen in the LSCF formalism, and is obtained from a separate calculation on a model mole-
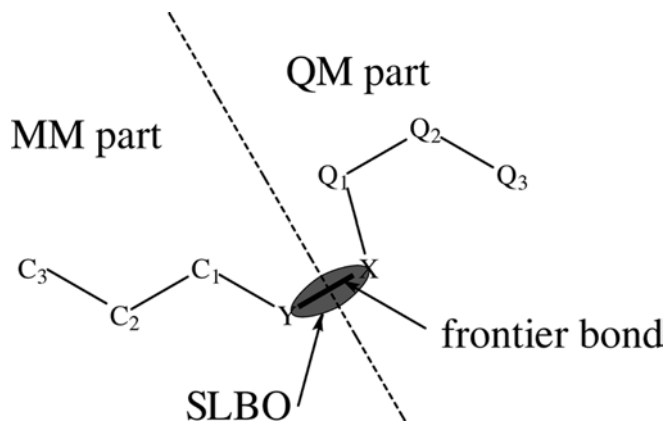

Fig. 1. Quantum mechanics (QM)/molecular mechanics (MM) frontier. The atoms are labeled with respect to their classical connectivity from the $X–Y$ bond

Table 1. List of nonelectrostatic quantum mechanics (QM)/molecular mechanics (MM) interactions

Bond potential	$X–Y$
Angle potential	$X–Y–C_1$
Torsion potential	$Q_1–X–Y–C_1$, $X–Y–C_1–C_2$
van der Waals potential	$X–C_i$ with $i > 1$, $Q_i–Y$ with $i > 1$, $Q_i–C_i$

cule owing to the transferability principle. Various orbital localization procedures are available and details of how to obtain the SLBO can be found elsewhere [3]. We note that the polarity of any type of covalent bond is by definition correctly reproduced by the SLBO, even dative-type bonds.

The frontier atom denoted Y involved in the SLBO is on the MM side. This atom deserves very particular attention since it is both a QM and a MM atom. As a QM atom it possess atomic orbitals (i.e., basis functions) and contributes generally with one electron to the electronic system.[2] Hence, its nuclear charge is $+1$ au. It is noteworthy that dative bonds can easily be described with the SLBO if the nuclear charge of the Y atom is put to 0 or 2 depending on whether it acts as an acceptor or as a donor atom.

From the notations used in Fig. 1, several contributions to the nonelectrostatic interaction energy between the QM and the MM parts need to be defined, and these are gathered in Table 1. Except for the $X–Y$ bond for which we have developed specific force field parameters [3], all other parameters, bonded or van der Waals, are taken from the standard force field, in order to keep our method as general as possible.

In addition to these interactions, electrostatic contributions play a very specific role. Since atoms belonging to the MM part possess a partial atomic charge, they interact directly with electrons and nuclei, instead of interacting with atoms as is the case in MM

---

[2] Except for a dative bond, where it can contribute with zero or two electrons

calculations. Hence, the corresponding Fock operator contains the classical charges—Coulombic contributions of electrons. The $\mu\nu$th element of the associated matrix expressed in the atomic orbitals basis (φ_μ, φ_ν) is then

$$F_{\mu\nu} = F_{\mu\nu}^0 + \sum_i^{MM} \left\langle \phi_\mu \left| \frac{Q_i}{r_i} \right| \phi_\nu \right\rangle, \qquad (5)$$

where $F^0_{\mu\nu}$ is the unperturbed Fock matrix element, without the surrounding classical charges, and Q_i is the partial atomic charge of MM atom i distant from the electron by r_i. Of course, if the force field used is polarizable, one has to take into account the interaction between the electrons and the induced electric dipole moments of the MM atoms.

This ab initio QM/MM method, based on the LSCF formalism, is implemented in our local version of Gaussian98 package [19], where the three force fields AMBER [20], DREIDING [21], and UFF [22] are available.

The SCRF formalism

The continuum models of solvation are numerous [7, 23, 24, 25] and are still widely used in computational quantum chemistry although the first related paper was published 30 years ago [1]. Since then, many improvements have been done to include as much physics in it as possible, to make the algorithm faster and more computer-efficient [7]. Among all these models, we chose the one developed in our group because it is very robust for wave function or geometry optimization and because it is fast [7]. Anyway, all these models have common features and what is detailed later for the SCRF formalism can be transferred, with slight modifications, to the other models.

The solvent is represented by a dielectric medium, characterized by the macroscopic relative dielectric constant (ϵ_0). The solute molecule, described by QM, is placed inside a cavity created in the continuum. The charge distribution (nuclei and electrons) of the solute molecule polarize the continuum, creating an electric field. This field, in turn, polarizes the charge distribution of the solute. This process lasts until the equilibrium is reached.

The shape of the cavity is not a crucial parameter since it has been shown that geometries or wave functions arising from different cavity shapes—spherical, ellipsoidal or molecular—are very similar as long as the volumes of the cavities are equivalent. However, when the volumes of different cavities differ, molecular shape [26] cavities are highly recommended.

Kirkwood [27] has shown that the solute–solvent free energy of interaction can be expressed, at a given center, with a Taylor expansion of the electrostatic potential created by the solute charge distribution in the continuum. This can be easily generalized to the multi-center case; however, we will only discuss the mono-center multipolar expansion here to simplify the reading. Further details can be found elsewhere [7]. An electric multipole of rank l has $2l+1$ components and is written M_l^m, where m goes from $-l$ to l. The reaction field component which can interact with this multipole is noted R_l^m. In the linear response approximation, it can be evaluated as

$$R_l^m = \sum_{l'} \sum_{m'} f_{ll'}^{mm'} M_{l'}^{m'}, \qquad (6)$$

where $f_{ll'}^{mm'}$ are the so-called reaction field factors. They only depend on the cavity shape and on the macroscopic relative dielectric constant of the continuum. The electrostatic free energy of solvation comes naturally as

$$\Delta G_{solv}^{elec} = -\frac{1}{2} \sum_l \sum_m R_l^m M_l^m. \qquad (7)$$

Of course, the Fock operator needs to be modified accordingly, and the $\mu\nu$th element of the associated matrix expressed in the atomic orbitals (φ_μ, φ_ν) basis is

$$F_{\mu\nu} = F_{\mu\nu}^0 + \sum_l \sum_m R_l^m \left\langle \phi_\mu \left| \hat{M}_{l,e}^m \right| \phi_\nu \right\rangle, \qquad (8)$$

where $F_{\mu\nu}^0$ is the unperturbed Fock matrix element—without the continuum—and $\hat{M}_{l,e}^m$ the electronic part of the multipole operator.

The three-layer LSCF/MM/SCRF method

Coupling the SCRF method with the hybrid LSCF/MM method is quite straightforward. One has, however, to specify some special points.

The cavity created in the continuum contains the whole QM/MM cluster, such that the dielectric medium is polarized by both the charge distribution of the QM part and by the partial atomic charges of the MM subsystem. As a counterpart, the geometry of the whole macromolecule is polarized by the reaction field. It is, however, clear that only the QM fragment is electronically polarized since the MM force field available does note yet contain polarization. This deficiency will soon be removed in forthcoming work. We note that at least two approaches are possible. Either one uses explicitly polarizable force fields with atomic polarizabilities, induced dipole moments and so on, or one can take advantage of the less expensive continuum model. One can ask the MM subunit to be inside a dielectric medium characterized by the electronic part of the relative dielectric constant. Anyway, the Fockian of the whole system needs to be modified accordingly as detailed in Eqs. (5) and (8):

$$F_{\mu\nu} = F_{\mu\nu}^0 + \sum_i^{MM} \left\langle \phi_\mu \left| \frac{Q_i}{r_i} \right| \phi_\nu \right\rangle$$
$$+ \sum_l \sum_m R_l^m \left\langle \phi_\mu \left| \hat{M}_{l,e}^m \right| \phi_\nu \right\rangle. \tag{9}$$

Moreover, the total multipole operator, entering the calculation of R_l^m, has to be changed to take the action of the MM portion on the continuum into account:

$$M_{l,tot}^m = \sum_A^{nuclei} Z_A M_{l,A}^m + \sum_\mu \sum_\nu P_{\mu\nu}^T \left\langle \phi_\mu \left| \hat{M}_{l,e}^m \right| \phi_\nu \right\rangle$$
$$+ \sum_C^{MM} Q_C M_{l,C}^m. \tag{10}$$

The actual three-layer method presented here has exactly the same options the original SCRF has (spherical, ellipsoidal or molecular shape cavity; mono-center or multicenter multipolar expansion, etc.).

The analytical first derivatives of the total energy with respect to the nuclear coordinates are of course also available, allowing us to optimize the geometry of the LSCF/MM cluster in solution. This method is implemented in our modified version of the Gaussian98 [19] package.

Computational details

All calculations concerning either the LSCF, the LSCF/MM or the LSCF/MM/SCRF methods were carried out with the modified local version of Gaussian98, on an IBM RISC/6000 397 workstation. Some calculations were performed with the Amsol [28] package to obtain nonelectrostatic solute–solvent interaction terms.

The cyclopentadiene molecule was described by means of QM and the dienophile molecule by the hybrid LSCF/MM method. The QM/MM partitioning of (−)-menthyl acrylate is illustrated in Fig. 2. Four conform-
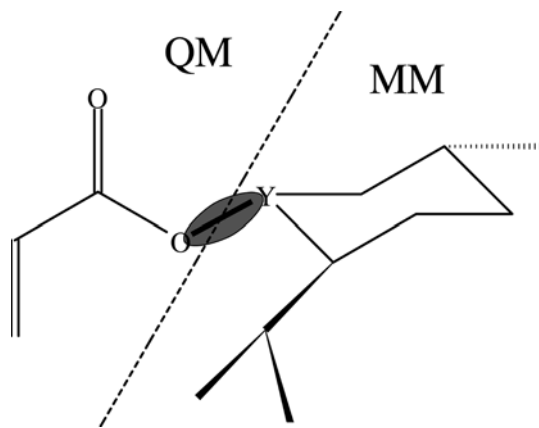

Fig. 2. QM/MM partitioning of the (−)-menthyl acrylate molecule

ers of the acrylate were considered in this study and the atomic arrangements are depicted in Fig. 3. We studied 16 different pathways. They differ by the conformation of the dienophile, by the endo or exo attack on cyclopentadiene, and by the Re or Si prochiral addition faces of (−)-menthyl acrylate.

In this study, we are primarily interested in relative energies since we want to discuss the influence of the solvent on the diastereomeric excess, and also we want to test our method against known results. For these reasons, the QM level of theory used is RHF/6-31G*. Moreover, it has been shown that relative energy values are practically insensitive to the size of the basis set, as long as it is at least of double-ζ polarized quality, and to electronic correlation [8]. The DREIDING force field was chosen since it is able to handle such systems.

In order to perform QM/MM calculations, we need to specify some special parameters as the partial atomic charges of the MM atoms. We tested different sets of partial atomic charges coming from various ways to fit the electrostatic potential of the (−)-menthyl acrylate molecule at the RHF/6-31G* level of theory (Table 2). Atomic labels are given in Fig. 4. We compared the Mulliken atom charges [33] of the dienophile obtained at the LSCF/MM level with those arising from the usual RHF/6-31G* calculation. Among all the methods [29, 30, 21, 32], the Merz–Kollman [29] one gives satisfactory results. The highest occupied molecular orbital and lowest unoccupied molecular orbital energies (Table 2) compared with the full QM calculations are quite well reproduced although they are peculiarly sensitivity to the set of atomic charges. In order to get a better agreement between the hybrid LSCF/MM calculations and the standard RHF/6-31G* computation, these atomic charges were multiplied by a common factor of 1.1. This point is very important for our study since the positions of the frontier orbitals are crucial for chemical reactivity. In contrast, the Mulliken atomic charges of the acrylate moiety are less sensitive to the chosen MM set of charges. One has to note that except for optimally partitioned electric property [32] charges all sets of partial charges give satisfying trends. To maintain the electroneutrality of the MM subunit, the classical charge of the frontier Y atom was adjusted.

As previously recommended [3] the Weinstein–Pauncz localization criterion [34] was applied to the methanol molecule, used as the model molecule for the C–O bond. The third atom, needed to uniquely define the frame of rotation, is of course the carbonyl carbon atom [3].

The parameters used for the continuum model are quite standard. A molecular shape cavity built with atomic van der Waals radii scaled by a factor of 2, instead of the usual 1.3084 scaling factor of the original method. This was done to ease the convergence of the wave function's optimization, avoiding some oscillating behavior. Of course, the magnitude of the electrostatic free energies of solvation will be reduced, and henceforth slighter than those obtained with other methods using
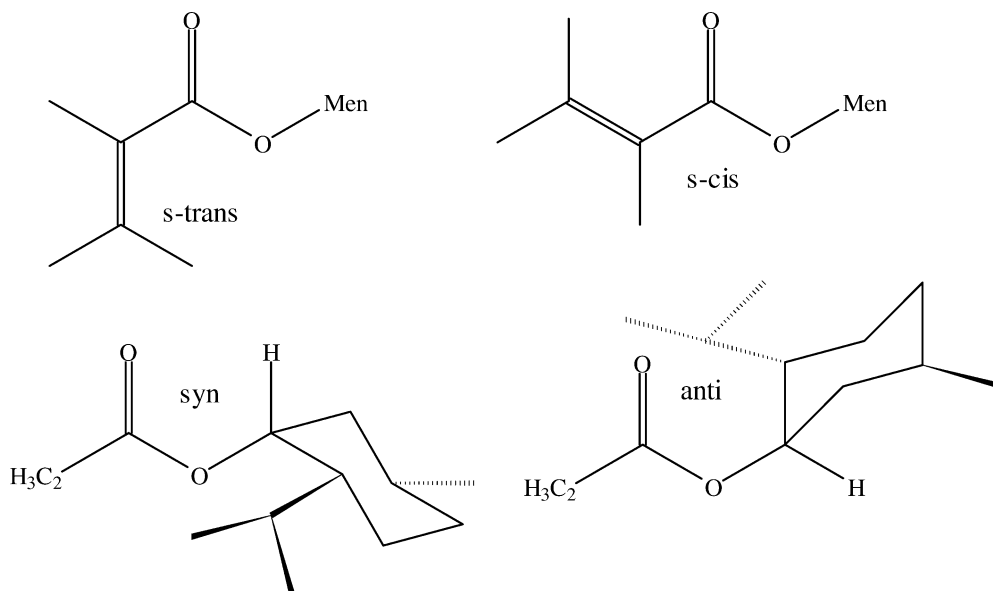
Fig. 3. Denomination of the conformers of (–)-menthyl acrylate

Table 2. Energies (atomic units) of the frontier orbitals of (–)-menthyl acrylate and Mulliken atomic charges) at various hybrid levels of theory. Only the name of the method used to fit the atomic point charges of the MM part to the electrostatic potential at the RHF/6-31G* level of theory, which serves as the reference (QM), is given for each column. See Fig. 4 for the atomic labels

|	QM	MK	MK*1.1	Chelp	ChelpG	OPEP
ϵ_{HOMO}	−0.3913	−0.3830	−0.3829	−0.3938	−0.3879	−0.3729
ϵ_{LUMO}	0.1172	0.1137	0.1166	0.0265	0.0839	0.1405
C_1	−0.354	−0.349	−0.349	−0.346	−0.348	−0.352
C_2	−0.250	−0.251	−0.251	−0.249	−0.251	−0.254
H_3	0.218	0.221	0.221	0.224	0.222	0.218
H_4	0.197	0.193	0.193	0.198	0.194	0.185
C_5	0.807	0.791	0.791	0.793	0.792	0.787
H_6	0.223	0.224	0.224	0.230	0.225	0.214
O_7	−0.584	−0.587	−0.587	−0.572	−0.585	−0.614
O_8	−0.672	−0.488	−0.488	−0.494	−0.490	−0.457
C_9	0.163	0.271	0.274	0.399	0.296	−0.143
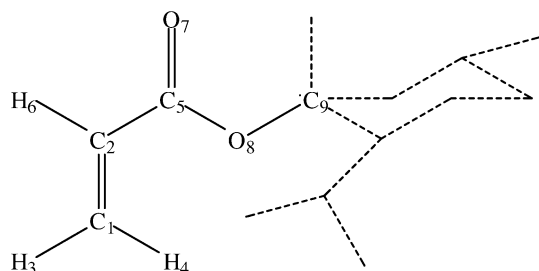
Fig. 4. Labeling of the atoms of the dienophile molecule described by QM

smaller cavities, but one can hope that relative quantities will still be discriminating. The maximal order for the multipolar expansion is equal to 6 since it is enough to converge the SCRF energy. For example, the contributions at orders 0–6 are 0.000, −0.416, −0.429, −0.273, −0.061, 0.033 and 0.001 kcal mol^{-1}, respectively, for the TS corresponding to an endo attack of the cyclopentadiene on the Re face of the stabler conformer of the

s-cis menthyl acrylate. Finally, the relative dielectric constant is equal to 78.4 to model water.

One could argue that our model does not fully take into account hydrogen-bond interactions, which have been shown to play a crucial role [35, 36, 37, 38], principally to explain the chemical rate acceleration. Hydrogen-bond interactions mainly have two components: the electrostatic and the charge-transfer interactions. The electrostatic interactions are correctly handled by our SCRF model, although it lacks the charge-transfer energy contribution and its directionality. However, since the hydrogen-bonded group of the solute is the carbonyl oxygen atom and given that it is always accessible for small molecules such as water, it seems reasonable to think that differences in hydrogen-bond interactions, between different TSs, will not be as large as the polarization energy differences. In fact differences in hydrogen-bonding interactions have been found to be less than 1 kcal mol^{-1} [35], which correspond mainly to the difference of electrostatic part. Hence, since we are dealing with relative energies, we consider that hydrogen

Table 3. Relative energies of the transition state (*TS*) (kilo calories per mole) in the gas phase and in solution. Solvent effects are composed of the electrostatic (*elec.*) and of the nonelectrostatic (cavitation–dispersion–specific, *CDS*, interactions) parts of the free energy of solvation. Δ corresponds to the difference between the relative energies in solution (elec. + CDS) and the relative energies in the gas phase

| Face	TS	Gas phase	Solution		Δ	Product
			Elec.	Elec. + CDS		
Free	endos-cis Re syn	0.00	0.00	0.00	0.00	1S-2S
Hindered	endos-cis Re anti	7.86	7.97	7.89	0.03	1S-2S
Hindered	endos-cis Si syn	0.81	0.96	1.07	0.26	1R-2R
Free	endos-cis Si anti	6.30	6.27	6.30	0.00	1R-2R
Hindered	endos-trans Re syn	2.65	2.22	2.12	−0.53	1S-2S
Free	endos-trans Re anti	8.28	7.92	8.01	−0.27	1S-2S
Free	endos-trans Si syn	1.35	1.00	1.04	−0.31	1R-2R
Hindered	endos-trans Si anti	2.08	1.87	1.88	−0.20	1R-2R
Free	exos-cis Re syn	0.54	0.60	0.64	0.10	1R-2S
Hindered	exos-cis Re anti	8.27	8.37	8.47	0.20	1R-2S
Hindered	exos-cis Si syn	1.68	1.95	1.89	0.21	1S-2R
Free	exos-cis Si anti	6.94	6.98	7.03	0.09	1S-2R
Hindered	exos-trans Re syn	3.33	3.28	3.09	−0.24	1R-2S
Free	exos-trans Re anti	8.33	7.96	7.93	−0.40	1R-2S
Free	exos-trans Si syn	2.02	1.69	1.76	−0.26	1S-2R
Hindered	exos-trans Si anti	10.08	9.63	9.68	−0.40	1S-2R
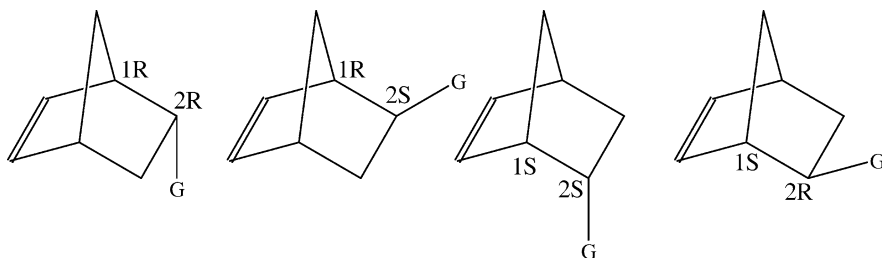
Fig. 5. Denomination of the four diastereoisomer products with respect to the absolute configuration of the carbon atoms. The *G* functional group is a shorthand notation for the remaining part of the (−)-menthyl acrylate moiety

bonds are very similar for every TS, and thus we ignore them in this study. To enforce this assumption, we performed semiempirical single-point calculations with the AM1 Hamiltonian [39] and the SM2 solvent model [40], using the AMSOL package [28], in order to get the nonelectrostatic contributions to the free energy of solvation (i.e., the cavitation–dispersion-specific term). The hydrogen-bond contribution is already included in this type of model, at least implicitly.

Since the analytical second derivatives of the LSCF energy with respect to the nuclear coordinates are not yet available, the TS localizations were carried out with initial force constants arising from an RHF/3-21G calculation on the whole system. Every stationary point found was characterized, afterward, by calculation of numerical frequencies.

Results and discussion

Energetic and selectivity

To discuss selectivity, we are primarily interested in relative energies. Total energies both from gas-phase calculations and from SCRF calculations are available as supplementary material. The relative energies in kilocalories per mole of the 16 TSs considered in this study computed with the gas-phase results, with the SCRF results, and with the nonelectrostatic solute–solvent interactions included are contained in Table 3. To ease the reading, the product to which each TS leads is also indicated and it is noted whether the attack is on the free or on the hindered face of the dienophile. The four diastereoisomers formed are sketched in Fig. 5. One has to bear in mind that, experimentally, the 1R-2R product is formed preferentially and its formation is enhanced by the polarity of the solvent.

It is noteworthy that, as previously found with methyl acrylate [8], the stablest TS does not correspond to the experimentally more favored one. This simply shows that the level of theory chosen here is not high enough to give reliable absolute energies within a few tens of a kilocalorie per mole. The energetic order of the TSs is roughly that of the reactants (the s-cis syn isomer is stabler than the s-trans syn conformer by about 1 kcal mol^{-1} and the anti reactants are 6–7 kcal mol^{-1} higher). Hence, we will focus our attention on variations induced by the solvent rather than on the absolute values, as was previously done [8].

One can remark that the nonelectrostatic contributions (cavitation, dispersion, and some specific solute–solvent interactions) do not modify significantly the relative energetic order of the TSs as expected, and they

Table 4. Energy differences (kilo calories per mole) for the TSs between the two prochiral faces (hindered–free), between the two attacks (endo–exo), and between the two configurations (anti–syn). In this A–B notation, a positive number means that conformer B is stabler than conformer A

Hindered–free	Gas phase	Solution	exo–endo	Gas phase	Solution	anti–syn	Gas phase	Solution
endos-cis syn	0.81	1.07	s-cisRe syn	0.54	0.64	endos-cis Re	7.86	7.89
endos-cis anti	1.56	1.59	s-cisRe anti	0.41	0.58	endos-cis Si	5.49	5.23
endos-trans syn	1.30	1.08	s-cis Sisyn	0.87	0.82	endos-trans Re	5.63	5.89
endos-trans anti	−6.20	−6.13	s-cisSi anti	0.64	0.73	endos-trans Si	0.73	0.84
exos-cis syn	1.14	1.25	s-trans Re syn	0.68	0.97	exos-cis Re	7.73	7.83
exos-cis anti	1.33	1.44	s-transRe anti	0.05	−0.08	exos-cis Si	5.26	5.14
exos-trans syn	1.31	1.33	s-transSi syn	0.67	0.72	exos-trans Re	5.00	4.84
exos-trans anti	1.75	1.75	s-transSi anti	8.00	7.80	exos-trans Si	8.06	7.92

will not be discussed any further. Since our method accounts for both steric hindrance and electrostatic solvent effects we will examine them more closely starting with the latter.

One sees form the Δ values of Table 3 that the s-trans conformers are stabilized by the solvent with respect to the s-cis ones. This is in perfect agreement with the results obtained on methyl acrylate [8], and correctly represents the experimental behavior [9]. Except for the s-cis/s-trans cleavage no other general trend of solvent effects can be found from our results. Surprisingly, the stabilization by the solvent of the endo TSs with respect to the exo ones that was previously found with methyl acrylate is not present here, at least not for all conformers. However, considering the endo and exo attacks for the TSs corresponding to the s-cis Re syn and s-trans Si syn conformers (free face, and stablest conformer), we remark that the endo TSs are stabilized by the solvent compared with the corresponding exo ones. Thus, when almost no steric hindrance—both intermolecular hindrance due to the isopropyl group and intramolecular hindrance due to the anti conformation—is present, our model recover the results obtained on a simpler model molecule (methyl acrylate), and corroborates experimental preferences. However, when steric repulsions exist no general trend can be derived and calculations must be performed. This point illustrates perfectly that solvent effects and steric repulsions are coupled, and it justified completely our approach.

In order to analyze the strength of the steric hindrance, the energy differences are reported in Table 4. Let us first examine the effect of the repulsive isopropyl functional group, reflected by the energy difference between two TSs differing only by the prochiral addition

face (hindered or free). The values are about 1–2 kcal mol^{-1} in favor of the free face, in agreement with a previous study [9], for both the gas phase and solution, and for all TSs but the endo s-trans anti isomers for which the addition on the hindered face is favored by more than 6 kcal mol^{-1}. This fact deserves particular attention. As already pointed out [9], the isopropyl group of menthyl acrylate can adopt various conformations (Fig. 6). We found the axial conformation to be the preferred one for the reactant. For this reason, all TSs were localized with that particular conformation. However, the steric crowding is such that in the endo s-trans Si anti TS it produces a rotation of the isopropyl group from the axial conformation to the stacked one. All attempts to localize this TS with the axial conformation failed. Hence, this stabilization of more than 6 kcal mol^{-1} does not reflect a preference for the hindered face, but rather for the stacked conformation. Since we have not consider the stacked conformation for the remaining TSs, we will not draw any conclusions on the stability of the endo s-trans Si anti TS with respect to the others.

The energy differences between TSs differing only by the type of attack (endo or exo) indicate that the endo conformers are slightly stabler (less than 1 kcal mol^{-1}) than the corresponding exo conformers. This result is consistent with previously published data [8], either theoretical or experimental, although in this study no systematic additional stabilization due to the solvent is found, as already mentioned. An interesting feature concerns the two s-trans Re anti TSs. Since the Re face corresponds to the free face for s-trans anti isomers, no repulsion due to the isopropyl group can be invoked in that case. In fact, both transition structures have about
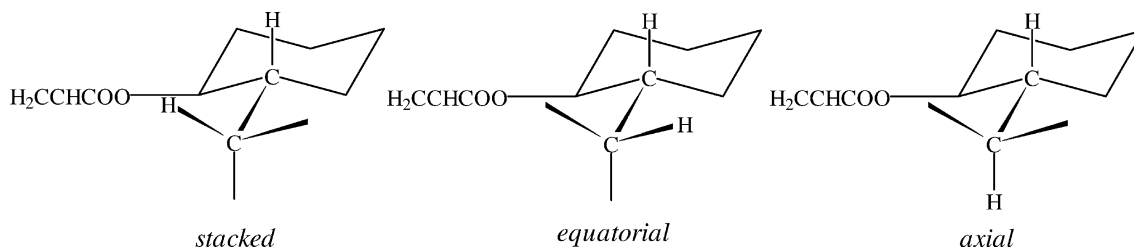

Fig. 6. Definition of the H–C–C–H dihedral angle used to characterize the conformation of the isopropyl group, here sketched in the axial conformation

the same energy, with a very slight preference for the exo TS in solution. This result could not have been guessed form the calculations performed on methyl acrylate, which always predict an endo stabilization and which cannot accounts for the syn/anti effect. This remark is again in favor of our three-layer method. Note also, that because of the rotation of the isopropyl group for the endo s-trans Si anti TS, the endo conformer is outrageously stabilized compared with the corresponding exo TS.

Finally, one notes that the syn conformers are much stabler than the corresponding anti conformers by about 5–6 kcal mol^{-1}, for both the gas-phase and the solution computations. This result is in agreement with experimental data [9] and is in qualitative agreement with a previous semiempirical study [9]. In fact, at the AM1 level, the energy difference is close to 1 kcal mol^{-1}. This discrepancy could be attributed to the Lewis acid chelating the oxygen atom in the semiempirical calculation. Again, owing to the rotation of the isopropyl group the energy of the endo s-trans Si anti TS is very close in energy to that of the corresponding syn conformer.

Diabatic states analysis

Apart from the selectivity features discussed previously, we would like to analyze the TS wave function in terms of similitude with respect to the reactant or to the product wave function. This property could be used to predict which factor will stabilize or destabilize the TS knowing only the effect it has on either the reactant or the product. Here we specifically would like to rationalize the influence of the solvent on this property. From our calculations, (for details see the supplementary material) we have found that each barrier height is increased by the solvent and that the exothermicity of the reaction is decreased by solvation. Consequently, the barrier heights of the reverse reaction are sometimes increased and sometimes decreased by the solvent. Hence, it is quite delicate to judge whether the TS looks more like the reactants or more like the product from these energetic general considerations. Moreover, the answer will not be global for the 16 TSs. For these reasons, we decided to use the LSCF method to decompose the wave function of the TS into its reactant and product contributions.

Since the LSCF formalism allows one to freeze orbitals, we can look at the way the energy, of the completely frozen electronic configuration of a species, transforms as the geometry is changed. In that way, we can define the kinds of diabatic surfaces based on the reactants or on the product electronic configurations. Of course, to make the orbitals follow the geometric modifications keeping their entire characteristics one needs to use both one-center orbitals, for the core and lone-pair electrons, and two-center orbitals like SLBOs, for the remaining electrons. Three or more center orbitals cannot be adapted to any other geometry than the one

they were determined with, without changing their properties.

To build the diabatic surfaces, we proceed as follows:

- Starting from the TS an intrinsic reaction cordinate (IRC) calculation is performed to obtain geometries along the reaction path.
- The reactant (product) wave function is optimized at the reactant (product) equilibrium geometry.
- Using a global localization criterion, the whole electronic system of the reactant (product) is transformed into localized orbitals.
- The coefficients of atomic functions that do not belong to the center bearing the core or lone-pair electrons are zeroed, giving strictly localized core orbitals or strictly localized lone-pair orbitals.
- The same zeroing operation is performed for the coefficients of atomic orbitals belonging to other atoms than the two defining the bond bearing the localized orbital, yielding SLBOs.
- This set of strictly localized functions is adequately rotated to adjust them to the given geometry corresponding to a point of the IRC pathway.
- An LSCF-type calculation is then performed for each IRC point.

To illustrate this technique the reaction path leading to the endo s-cis Re syn product, both in the gas phase and in solution, was chosen. One should note that the same QM/MM partitioning is observed for menthyl acrylate. Ten points on each side of the TS were found by the IRC calculations. The Pipek–Mezey [41] localization criterion was used. The results are displayed in Fig. 7 for both the gas-phase calculations and for the SCRF computations. One can readily see that the reactant configuration curve crosses the product curve on the reactant side of the reaction path for the gas-phase calculations, but that the crossing occurs on the product side in solution. This is easily rationalized if one remembers that the reaction is less exothermic in solution than in the gas phase, i.e., the reactant configuration is stabilized more by the solvent. From these curves, one can decompose solvent effects into two parts, the electronic part and the geometric part. First of all, we can confirm that the electronic configuration of the reactants is largely stabilized by the solvent compared with that of the product. Second, the geometries on the reactant side give rise to more solvated systems than those on the product side. This effect is clearly seen with the reactant curves, and also exists, although with a smaller magnitude, with the product curves. Hence, the fact that the barrier heights are increased by the solvent arises from combined electronic and geometric effects.

If we consider that the wave function of the TS can be expressed as a linear combination of the two diabatic states—the reactant state and the product state—then it is mainly composed of the lowest state at the TS geometry; hence, the TS wave function presents a cer-
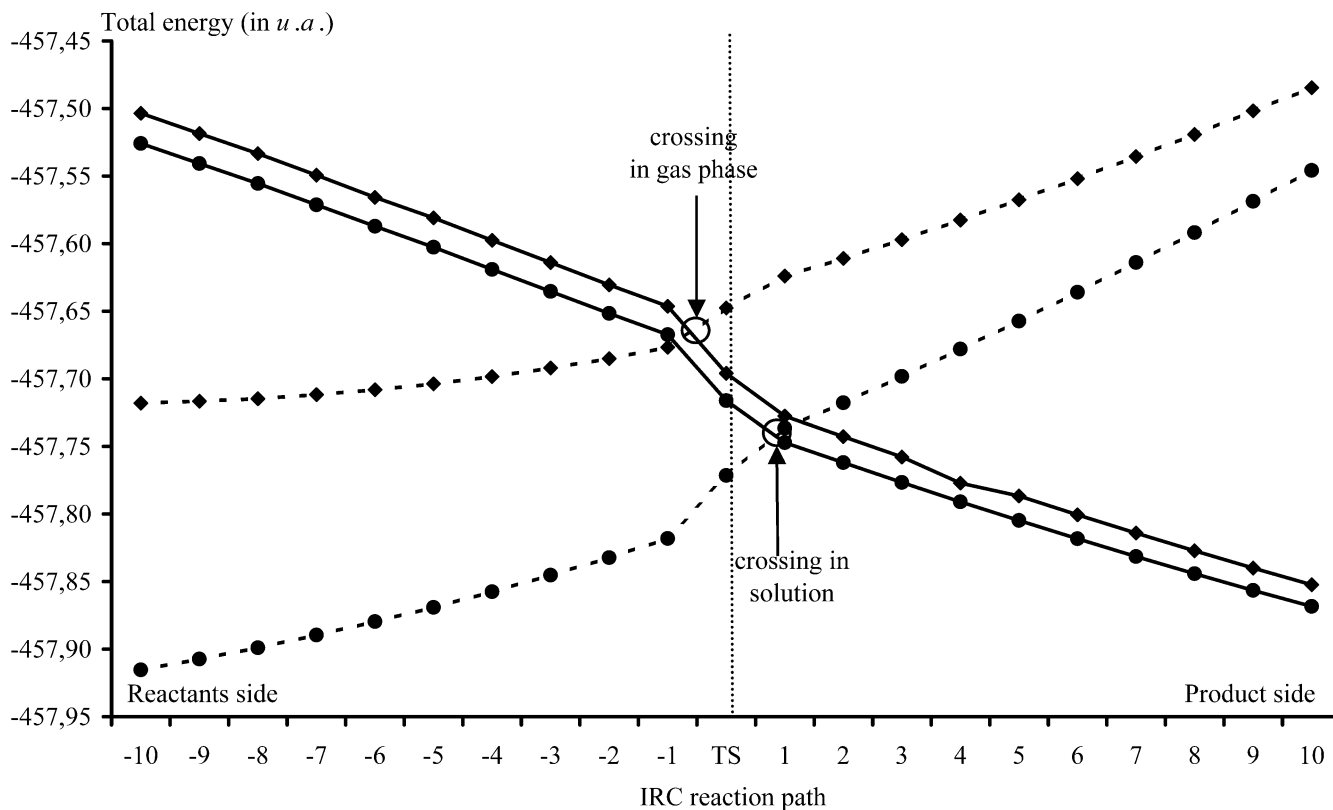
Fig. 7. Diabatic reactant configuration curve (*dashed line*) and diabatic product configuration curve (*solid line*) along the 21 points resulting from an intrinsic reaction coordinate (IRC) calculation for the endo s-cis Re syn reaction path, in the gas phase (*diamonds*) and in solution (*circles*)

tain "productlike" character in the gas phase and a "reactant-like" character in solution. If we do not want to talk about absolute character, at least we can say that the solvent makes the TS more "reactant-like". The displacement of the crossing is due to the higher stabilization of the reactant electronic configuration. Since solvent effects decrease the exothermicity of the 16 reactive channels studied here, we can conclude that this feature is common to the 15 other TSs. Hence, we can conclude that modifications of the reactants, to improve the selectivity, for example, are certainly more important in polar solvents than in the gas phase.

Conclusions

The three-layer LSCF/MM/SCRF hybrid method has been described. It is designed for the study of large molecular systems in solution. The reactive part, generally small, is treated by means of quantum chemistry approaches, within the LSCF framework, while the remaining parts of the molecule, for example, the encumbering ancillary groups, are described by employing MM force fields. The solvent is modeled as a polarizable continuum characterized by the macroscopic relative dielectric constant of the solvent in order to easily access macroscopic thermodynamic quantities such as solvation free enthalpies. One can note that

explicit solvent molecules can be added to the QM/MM system either at the QM or at the MM level depending on what one is interested in.

The method was applied to the study of the asymmetric Diels–Alder reaction between cyclopentadiene and (−)-menthyl acrylate, to ascertain its validity (Table 5). This reaction had already been studied at the AM1/MM3 level but with solvent effects missing [9, 13] or using a model molecule with electrostatic solvent effects [8], hence lacking the steric repulsion. For the first time, steric hindrance and solvent effects have been considered simultaneously.

All the conclusions found here are consistent with those previously found, removing the ambiguities in the conclusions based on incomplete systems. For example, we have found that the endo preference induced by the solvent is only valid for species free of steric hindrance. When intermolecular or intramolecular (for the anti conformer) steric repulsion is considered, solvent effects stabilize either the endo or the exo attack. This clearly shows that our new three-layer hybrid method is an adequate and complete tool to study chemical reactivity of large molecular systems in solution.

In a future study, it will be interesting to investigate more precisely the role of the solvent on the chemical rate acceleration by incorporating explicit water molecules preferentially at the MM level to minimize the computational cost, if QM/MM hydrogen-bond inter-

Table 5. Total energies (in *a.u.*) of the reactants, the 16 transition states (TS) and the 16 adducts, in the gas phase and in solution (electrostatic solute-solvent interactions only) at the RHF/6-31G*(DREIDING) level of theory

Reactants	Gas phase		Solution	
cyclopentadiene	-192.791721		-192.792094	
(-)-menthyl acrylate *s-cis syn*	-265.635994		-265.638409	
(-)-menthyl acrylate *s-cis anti*	-265.626600		-265.628145	
(-)-menthyl acrylate *s-trans syn*	-265.634763		-265.636373	
(-)-menthyl acrylate *s-trans anti*	-265.625626		-265.627303	
	Gas phase		Solution	
Reaction path	TS	products	TS	products
endo s-cis Re syn	-458.372459	-458.462569	-458.374028	-458.464971
endo s-cis Re anti	-458.359942	-458.450848	-458.361324	-458.452363
endos-cis Si syn	-458.371177	-458.461493	-458.372496	-458.462784
endos-cis Si anti	-458.362421	-458.452973	-458.364042	-458.454459
endos-trans Re syn	-458.368242	-458.460110	-458.370488	-458.462436
endos-trans Re anti	-458.359272	-458.452600	-458.361402	-458.45386
endos-trans Si syn	-458.370311	-458.462895	-458.372429	-458.464087
endos-trans Si anti	-458.369145	-458.451856	-458.371045	-458.453197
exos-cis Re syn	-458.371597	-458.463278	-458.373068	-458.464512
exos-cis Re anti	-458.359288	-458.452802	-458.360687	-458.454204
exos-cis Si syn	-458.369793	-458.461507	-458.370915	-458.463560
exos-cis Si anti	-458.361399	-458.453217	-458.362901	-458.454566
exos-trans Re syn	-458.367164	-458.459632	-458.368802	-458.460953
exos-trans Re anti	-458.359193	-458.450519	-458.361341	-458.452194
exos-trans Si syn	-458.369250	-458.46194	-458.371335	-458.464274
exos-trans Si anti	-458.356403	-458.453217	-458.358680	-458.454571

actions can be correctly estimated. Moreover, the polarization of the menthyl group by the solvent is certainly something one would have to look into in the near future.

Solvent interactions were analyzed in terms of electronic and geometric effects, by means of frozen electronic configuration curves along the IRC reaction path. This new application of the LSCF approach shows that in addition to the quantitative information one expects, the LSCF method provides a qualitative tool to rationalize the numbers.

References

1. Rivail J-L, Rinaldi D (1973) Theor Chim Acta 32:57
2. Assfeld X, Rivail J-L (1996) Chem Phys Lett 263:100
3. Ferré N, Assfeld X, Rivail J-L (2002) J Comput Chem 23:610
4. Ferré N, Assfeld X (2002) J Chem Phys 117:4119
5. Ferré N, Assfeld X (2003) J Mol Struct (THEOCHEM) 632:83
6. Assfeld X, Rivail J-L(2003) Theor Chem Acc (in press)
7. Rinaldi D, Bouchy A, Rivail J-L, Dillet V (2003) J Chem Phys
8. Ruiz-Lopez MF, Assfeld X, Garcia JI, Mayoral JA, Salvatella L (1993) J Am Chem Soc 115:8780
9. Salvatella L, Mokrane A, Cartier A, Ruiz-López MF (1998) J Org Chem 63:4664
10. Otto S, Engberts JBFN (2000) Pure Appl Chem 72:1365
11. Cativiela C, Garcia JI, Mayoral JA, Salvatella L (1996) Chem Soc Rev 25:209
12. Seeman JI (1986) J Chem Educ 63:42
13. Salvatella L, Mokrane A, Cartier A, Ruiz-López MF (1998) Chem Phys Lett 296:239
14. Bakowies D, Thiel W (1996) J Phys Chem 100:10580
15. de Pascual-Teresa B, Gonzalez J, Asensio A, Houk KN (1995) J Am Chem Soc 117:4347
16. Philipp D, Friesner RA (1999) J Comput Chem 20:1468
17. Löwdin PO (1950) J Chem Phys 18:365
18. Szabo A, Ostlund NS (1989) Modern quantum chemistry. McGraw-Hill, New York, p 144
19. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Zakrzewski VG, Montgomery JA Jr, Stratmann RE, Burant JC, Dapprich S, Millam JM, Daniels AD, Kudin KN, Strain MC, Farkas O, Tomasi J, Barone V, Cossi M, Cammi R, Mennucci B, Pomelli C, Adamo C, Clifford S, Ochterski J, Petersson GA, Ayala PY, Cui Q, Morokuma, D. K. Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Cioslowski J, Ortiz JV, Baboul AG, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Gomperts R, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Gonzalez C, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Andres JL, Gonzalez C, Head-Gordon M, Replogle ES, Pople JA (1998) Gaussian 98, revision A.7. Gaussian, Pittsburgh, PA
20. Cornell W, Cieplak P, Bayly C, Gould I, Merz K Jr, Fergusson D, Spellmeyer D, Fox T, Cladwell J, Kollman P (1995) J Am Chem Soc 117:5179
21. Mayo SL, Olafson BD, Goddard WA (1990) J Phys Chem 94:8897
22. Rappé AK, Casewit CJ, Colwell KS, Goddard WA III, Skiff WM (1992) J Am Chem Soc 114:10024
23. Cossi M, Barone V, Cammi R, Tomasi J (1996) Chem Phys Lett 255:327
24. Klamt A, Schüürmann GJ (1993) J Chem Soc Perkin Trans 2 5:799
25. Li J, Zhu T, Hawkins GD, Winget P, Liotard DA, Cramer CJ, Truhlar DG (1999) Theor Chem Acc 103:9
26. Silla E, Tuñón I, Pascual-Ahuir JL (1991) J Comput Chem 12:1077
27. Kirkwood JG (1934) J Chem Phys 2:351
28. Hawkins GD, Giesen DJ, Lynch GC, Chambers CC, Rossi I, Storer JW, Li J, Rinaldi D, Liotard DA, Cramer CJ, Truhlar DG (1997) AMSOL version 6.5.2. University of Minnesota, Minneapolis, based in part on AMPAC version 2.1 (Liotard DA, Healy EF, Ruiz JM, Dewar MJS)

29. Besler BH, Merz KM Jr, Kollman PA (1990) J Comput Chem 11:431
30. Chirlian LE, Francl MM (1987) J Comput Chem 8:894
31. Breneman CM, Wiberg KB (1990) J Comput Chem 11:361
32. Ángyán JG, Chipot C, Dehez F, Hättig C, Jansen G, Millot C OPEP: a tool for the optimal partitioning of electric properties. Université Henri Poincaré, Vandoeuvre-Nancy Cedex, France
33. Mulliken RS (1949) J Chim Phys 46:497
34. Weinstein H, Pauncz R (1968) Symp Faraday Soc 2:23
35. Blake JF, Lim D, Jorgensen WL (1994) J Org Chem 59:803
36. Chandrasekhar J, Sharillskul S, Jorgensen WL (2002) J Phys Chem B 106:8078
37. Furlani TR, Gao J (1996) J Org Chem 61:5492
38. Kong S, Evanseck JD (2000) J Am Chem Soc 122:10418
39. Dewar MJS, Zoebisch EF, Healy EF, Stewart JJP (1985) J Am Chem Soc 107: 3902
40. Cramer CJ, TruhlarDG (1992) Science 256:213
41. Pipek J, Mezey P (1989) J Chem Phys 90:4916