Theor Chem Acc (1998) 100:307–313
DOI 10.1007/s002149800m68

Theoretical
Chemistry Accounts

Regular article

Implementation of the IMOMM methodology for performing combined QM/MM molecular dynamics simulations and frequency calculations

Tom K. Woo[1], **Luigi Cavallo**[2], **Tom Ziegler**[1]

[1]Department of Chemistry, University of Calgary, 2500 University Drive, N.W., Calgary, Alberta, T2N 1N4, Canada
[2]Dipartimento di Chimica, Universitá "Federico II" di Napoli, Via Mezzocannone 4, I-80134 Napoli, Italy

Received: 11 May 1998 / Accepted: 14 August 1998 / Published online: 16 November 1998

Abstract. A technique for implementing the integrated molecular orbital and molecular mechanics (IMOMM) methodology developed by Maseras and Morokuma that is used to perform combined quantum mechanics/molecular mechanics (QM/MM) molecular dynamics simulations, frequency calculations and simulations of macromolecules including explicit solvent is presented. Although the IMOMM methodology is generalized to any coordinate system, the implementation first described by Maseras and Morokuma requires that the QM and MM gradients be transformed into internal coordinates before they are added together. This coordinate transformation can be cumbersome for macromolecular systems and can become ill-defined during the course of a molecular dynamics simulation. We describe an implementation of the IMOMM method in which the QM and MM gradients are combined in the cartesian coordinate system, thereby avoiding potential problems associated with using the internal coordinate system. The implementation can be used to perform combined QM/MM molecular dynamics simulations and frequency calculations within the IMOMM framework. Finally, we have examined the applicability of thermochemical data derived from IMOMM framework. Finally, we have examined the applicability of thermochemical data derived from IMOMM frequency calculations.

Key words: Macromolecular systems – Molecular dynamics – Quantum mechanics – Molecular mechanics – Frequencies

1 Introduction

The combined quantum mechanics and molecular mechanics (QM/MM) method [1–7] has recently received significant attention [5] because it holds the promise of allowing simulations of bond breaking and forming processes within large molecular systems. The combined QM/MM approach involves partitioning the system into QM and MM regions whereby the molecular potential is determined partially by a QM electronic structure calculation and partially by a MM force field.

Although the combined QM/MM method is conceptually simple, there are some substantial practical issues to contend with. This is particularly true, if the QM/MM partition occurs within the same molecule as illustrated in Fig. 1a. Here, the difficulty lies in the fact that at least one covalent bond will involve an atom from the QM region and one from the MM region, and therefore the electronic system of the QM region must in some way be truncated as the QM/MM boundary is crossed along these bonds. A simple solution to this truncation problem, first proposed by Singh and Kollman [2], involves capping the electronic system of the QM region with "capping" atoms. In this way, the electronic structure calculation is performed on what is referred to as the QM model system, as shown in Fig. 1b. Compared to other truncation approaches [1, 8, 9] the primary advantage of the capping atom approach is that no special treatment of the electronic system is necessary, thereby allowing the QM/MM methodology to be easily implemented within existing electronic structure codes. In this paper we shall refer to the "real system" as the system consisting of all atoms (Fig. 1a), QM or MM, not including the capping atoms.

Within the capping atom framework, the covalent bonds that cross the QM/MM boundary involve three atoms that we will label the QM-link atom, the capping atom, and the MM-link atom as shown in Fig. 2. The QM- and MM-link atoms make up the covalent bond in the real system that links the QM and MM regions. The capping atom is introduced to satisfy the valences of the QM system and is not part of the "real system". The relationship between the coordinates of the two link

Correspondence to: T. Ziegler
e-mail: ziegler@acs.ucalgary.ca

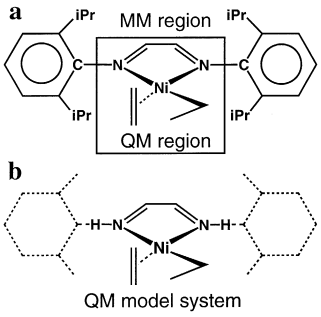

Fig. 1. **a** Example of the partitioning of quantum mechanics (QM) and molecular mechanics (MM) regions within a single molecule. **b** The QM model system for which the electronic structure calculation is performed in the capping atom QM/MM approach

atoms and the capping atom, is different for different QM/MM approaches [2, 4, 6].

In the popular capping atom prescription of Singh and Kollman [2], all three atoms in the link bond exist as independent variables. The combined QM/MM total energy in this framework can be defined as:

$$E_{QM/MM} = E_{QM} + E_{MM} \tag{1}$$

where E_{QM} is simply the energy of the QM model system with the capping atom included, and E_{MM} is the sum of all of the MM energy terms of the real system that contain at least one MM atom (for example, the bond stretching potential between the QM-link atom and the MM-link atom is handled by the appropriate MM stretching potential) [1].

In this approach the total energy of Eq. (1) is optimized with respect to all the degrees of freedom of the real system as well as the three degrees of freedom of the capping atom. In most cases this approach is appropriate and the method has been successfully applied to numerous systems [3, 10–12]. However, in cases where a geometry optimization of $E_{QM/MM}$ leads to a structure where the three atoms (QM-link, MM-link and capping) deviate from a collinear arrangement, the approach becomes somewhat dissatisfying, because some geometric distortions involving the link bond will not be manifested in the electronic structure of the QM model system. This situation is exaggerated in Fig. 2a. Another feature of the method is that compared to the real system, the QM/MM system possesses extra degrees of freedom due to the capping atoms.

In the IMOMM approach of Maseras and Morokuma [4], the position of the MM-link atom is no longer an independent variable and a strict relationship between the coordinates of the link atoms and the capping atom is enforced. Specifically, the MM-link atom is positioned along the QM-link atom – capping atom bond vector at a distance $R_{MM\text{-}link}$ from the QM-link atom as shown in Fig. 2b. Thus, the bond distance, $R_{MM\text{-}link}$; the bond angle, $\theta_{MM\text{-}link}$ and the dihedral angle,

[1] Most authors define the total combined QM/MM energy expression with a component that expresses the interaction energy between the QM and MM regions. In Eq. (1), E_{QM} and E_{MM} include these interaction energies

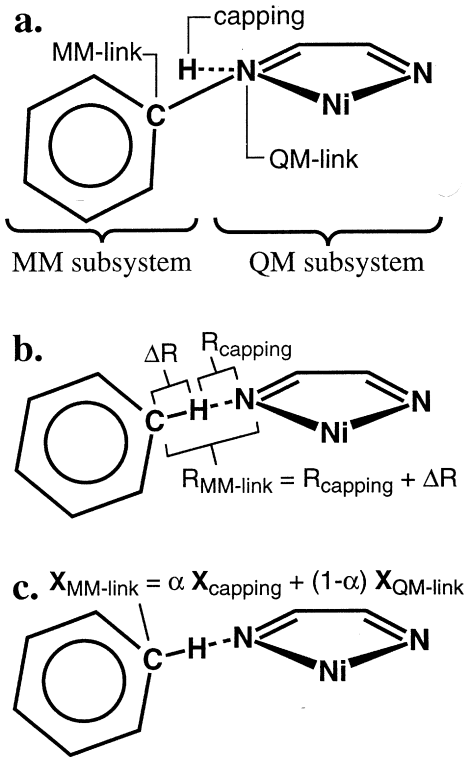

Fig. 2a-c. Schematic description of the relationship between the three atoms (QM-link atom, capping atom and MM-link atom) involved in covalent bonds that cross the QM/MM boundary for the methods of a) Singh and Kollman, [2] b) Maseras and Morokuma [4] and c) the current adaptation.

$\phi_{MM\text{-}link}$ used to define the MM-link atom are related to the bond distance, $R_{capping}$; the bond angle, $\theta_{capping}$ and the dihedral angle, $\phi_{capping}$, used to define the capping atom within the QM model system as:

$$R_{MM\text{-}link} = R_{capping} + \Delta R \ , \tag{2}$$

$$\theta_{MM\text{-}link} = \theta_{capping} \ , \tag{3}$$

$$\phi_{MM\text{-}link} = \phi_{capping} \ , \tag{4}$$

where ΔR is a constant. As a result of the relations enforced in Eqs. (2–4), geometric distortions involving the link bond are exhibited in the electronic structure of the QM model system. Again the total energy $E_{QM/MM}$ is given by an expression similar to that of Eq. (1). However, the MM terms included in the total energy expression (Eq. 1) are different from those of the Kollman scheme [2]. In the IMOMM approach, MM potentials are only included if they depend on atoms which do not have a corresponding atom in the QM model system. For example, the MM bond stretching potential inolving the QM-link and the MM-link atoms is not included since this potential is assumed to be adequately handled by the correspnding bond in the QM model system involving the capping atom and the QM-link atom (see Ref. [4] for more details and exceptions).

Since the the relationship defining the position of the MM-link atom is expressed in internal coordinates,

the original IMOMM implementations [2] requires that the QM and MM gradients first be transformed into internal coordinates before being added [4]. The combined QM/MM gradients in internal coordinates are simply:

$$\frac{\partial E_{QM/MM}}{\partial R_i} = \frac{\partial E_{QM}}{\partial R_i} + \frac{\partial E_{MM}}{\partial R_i} \ , \tag{5}$$

$$\frac{\partial E_{QM/MM}}{\partial \theta_i} = \frac{\partial E_{QM}}{\partial \theta_i} + \frac{\partial E_{MM}}{\partial \theta_i} \ , \tag{6}$$

$$\frac{\partial E_{QM/MM}}{\partial \phi_i} = \frac{\partial E_{QM}}{\partial \phi_i} + \frac{\partial E_{MM}}{\partial \phi_i} \ , \tag{7}$$

where i runs over all degrees of freedom of the real system, except that the degrees of freedom of the MM-link atoms are replaced by the degrees of freedom of the capping atom through Eqs. (2–4). Thus the gradient that characterizes the capping atom is defined with the help of Eqs. 2–4 as:

$$\begin{aligned}\frac{\partial E_{QM/MM}}{\partial R_{capping}} &= \frac{\partial E_{QM}}{\partial R_{capping}} + \frac{\partial E_{MM}}{\partial R_{MM\text{-}link}} \frac{\partial R_{MM\text{-}link}}{\partial R_{capping}} \\ &= \frac{\partial E_{QM}}{\partial R_{capping}} + \frac{\partial E_{MM}}{\partial R_{MM\text{-}link}} \ ,\end{aligned} \tag{8}$$

$$\frac{\partial E_{QM/MM}}{\partial \theta_{capping}} = \frac{\partial E_{QM}}{\partial \theta_{capping}} + \frac{\partial E_{MM}}{\partial \theta_{MM\text{-}link}} \ , \tag{9}$$

$$\frac{\partial E_{QM/MM}}{\partial \phi_{capping}} = \frac{\partial E_{QM}}{\partial \phi_{capping}} + \frac{\partial E_{MM}}{\partial \phi_{MM\text{-}link}} \ . \tag{10}$$

Equations (8–10) reveal that, although the position of the MM-link atoms are not free variables, the forces acting on these atoms are passed onto other atoms via the chain rule in the equations.

For some types of applications, the transformation of the MM gradients between coordinate systems or even the definition of the internal coordinates can present practical problems. For macromolecular systems such as enzymes which typically contain 10^3–10^4 atoms, the coordinate transformation can be cumbersome since it involves the inversion of a $3N \times 3N$ matrix where N is the number of atoms in the "real system". Moreover, when explicit solvent molecules are considered, the transformation is likely to become ill-defined during the course of an optimization or molecular dynamics simulation (unless efforts are made to redefine the internal coordinates to prevent the evolution of linear or near-linear angles). For these reasons, the original IMOMM implementation that requires a coordinate transformation can be impractical for these important kinds of simulations.

In this paper we describe an adaptation of the IMOMM methodology of Maseras and Morokuma where no coordinate transformation is required, thereby avoiding the potential difficulties associated with using internal coordinates. The implementation allows for practical IMOMM simulations of macroscopic systems, explicitly solvated systems, and it allows for energy conserving molecular dynamics simulations and frequency calculations to be performed. To demonstrate this, a combined QM/MM molecular dynamics simulation with explicit solvent is presented. Finally, we explore the applicability of utizing the IMOMM method for performing frquency calculations and deriving thermochemical data.

2 Methodology

The combined QM/MM implementation described here is based on the original IMOMM scheme of Maseras and Morokuma [4]. All of the rules for determining which MM potentials to accept or discard are the same as in the original IMOMM scheme. We will deal exclusively with defining the relationship between the atoms associated with a QM/MM link bond and how the QM and MM based gradients are combined.

We start by defining the relationship between the link atoms and the capping atoms that are involved in a covalent bond that crosses the QM/MM boundary. In the IMOMM scheme, the position of the MM-link atom is not an independent variable. Instead, the MM-link atom is always placed along the bond vector of the QM-link atom and the capping atom bond. Expressed in cartesian coordinates, this relationship is defined by Eq. (11) and rearranged in Eq. (12):

$$\mathrm{X}_{MM\text{-}link} = \mathrm{X}_{QM\text{-}link} + \alpha(\mathrm{X}_{capping} - \mathrm{X}_{QM\text{-}link}) \ , \tag{11}$$

$$\mathrm{X}_{MM\text{-}link} = (1-\alpha)\mathrm{X}_{QM\text{-}link} + \alpha\mathrm{X}_{capping} \ . \tag{12}$$

If α is defined in the following way,

$$\alpha = \frac{\|\mathrm{X}_{capping} - \mathrm{X}_{QM\text{-}link}\| + \Delta R}{\|\mathrm{X}_{capping} - \mathrm{X}_{QM\text{-}link}\|} = \frac{R_{capping} + \Delta R}{R_{capping}} \tag{13}$$

then the orginal implementation of Maseras and Morokuma is recovered in cartesian coordinates where Eq. (2) is satisfied. In our implementation we define α as a constant parameter. In this way the bond distance, R_{MM}, between the QM- and MM-link atoms is defined as a constant factor of the QM-link – capping atom bond distance, $R_{capping}$, such that Eq. (14) is satisfied:

$$R_{MM\text{-}link} = \alpha R_{capping} \ . \tag{14}$$

The relationship defined in Eq. (14) is analogous to Eq. (2) of the original scheme, and α for each link bond is chosen in a similar fashion as ΔR. Comparison of the relationships (Eq. 2, 14) reveals that they are similar in that both allow changes in $R_{capping}$ to be reflected in the bond distance $R_{MM\text{-}link}$. There is no clear advantage in using either relationship in terms of the physical model of the QM/MM link bond. However, by defining α as a constant, combining the MM and QM forces in cartesian coordinates is simplified.

By applying the relationship defined in Eq. (12) where α is a constant and applying the chain rule, the combined QM/MM gradients on the QM-link atom and the capping atom are expressed by Eqs. (15) and (16), respectively.

$$\frac{\partial E_{QM/MM}}{\partial \mathrm{X}_{QM\text{-}link}} = \frac{\partial E_{QM}}{\partial \mathrm{X}_{QM\text{-}link}} + \frac{\partial E_{MM}}{\partial \mathrm{X}_{QM\text{-}link}} + (1-\alpha)\frac{\partial E_{MM}}{\partial \mathrm{X}_{MM\text{-}link}} \tag{15}$$

$$\frac{\partial E_{QM/MM}}{\partial \mathrm{X}_{capping}} = \frac{\partial E_{QM}}{\partial \mathrm{X}_{capping}} + \alpha\frac{\partial E_{MM}}{\partial \mathrm{X}_{MM\text{-}link}} \ . \tag{16}$$

Equations (15) and (16) are valid for the cartesian coordinate system and there is no need to add the QM and MM forcess in an internal coordinate system [3].

[2] It should be noted that the IMOMM approach is generalized to any coordinate system by Maseras and Morokuma, but the authors emphasize that the implementation is restricted to an internal coordinate system

[3] Equations equivalent to Eqs. (15) and (16) can also be easily derived with α defined by Eq. (13). This would recover the original implementation of Maseras and Morokuma where the constraints in Eqs. (2–4) are satisfied

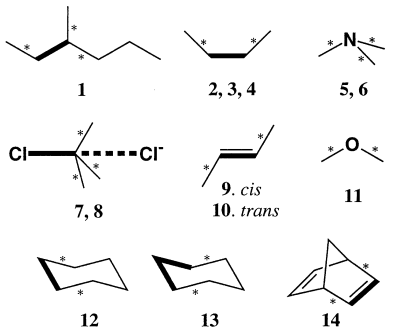

Fig. 3. List of structures showing the partitioning of QM and MM regions. *Thick bonds* represents the QM region, while *thin bonds* represent the MM region. Covalent bonds labelled with *asterisks* denote the QM/MM link bonds. All of the link bonds have been capped with hydrogen atoms in the QM model system

3 Results and discussion

3.1 IMOMM molecular dynamics

The IMOMM methodology described above has been implemented within the Projector Augmented Wave (PAW) [13] Car-Parrinello [14] ab initio molecular dynamics package. In this section we demonstrate the suitability of this scheme for performing combined QM/MM molecular dynamics and simulations with explicit solvent molecules.

A 4000 time-step combined QM/MM molecular dynamics simulation of solvated 3-methylhexane, **1**, has been performed with a time step of $\Delta t = 3.0$ a.u. Figure 3 illustrates the partitioning in **1** where the bold region represents the QM subsystem and the bonds labelled with asterisks denote the QM/MM link bonds. In this simulation the three link bonds are capped with hydrogen atoms in the QM model system such that the electronic structure calculation is performed on ethane. The system is solvated by 14 isobutane molecules in a 20 Å cubic cell for which periodic boundary conditions [15] have been applied with a 10 Å non-bonded cutoff. Jorgensen's OPLS-all atom MM force field [16] was used for both the solvent and the MM components of the solute. Non-bonded electrostatic interactions between the QM and MM regions were treated in the familiar MM fashion. For each of the three link bonds, an α (as defined in Eq. 11) of 1.38 was utilized. Masses of the capping atoms were rescaled to those of the corresponding MM-link atom. In other words, the capping hydrogen atoms were assigned a mass of 12.0 amu. For the calculation of the QM model system, the wave function was expanded in plane waves up to an energy cutoff of 30 RY within a periodic cell spanned by the lattice vectors [0.0 6.5 6.5], [6.5 0.0 6.5], [6.5 6.5 0.0] (in Å). The local-density approximation according to the parameterization of Perdew and Zunger [17], with gradient corrections due to Becke [18] and Perdew [19, 20], were utilized. The trajectory presented was pre-equilibrated and pre-thermostated [21] at 300 K for 2000 time-steps. Although the thermostat was turned off, an average temperature of 300 K was still retained over the course of the simulation presented.

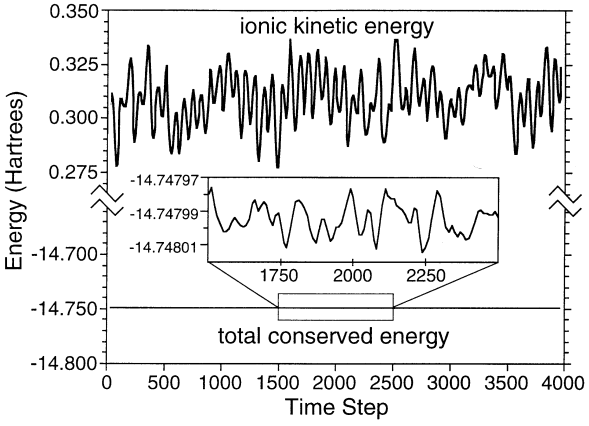

Fig. 4. Plot of the kinetic energy of the nuclei (QM and MM) and the total conserved energy during a combined QM/MM molecular dynamics simulation of 3-methylhexane, **1**. The total energy is plotted at the same scale as the kinetic energy. The inset reveals the scale at which the total energy fluctuates during the molecular dynamics. During the simulation the average temperature of the system is 300 K

Figure 4 illustrates the energy conservation of the QM/MM molecular dynamics simulation of **1**. In the upper portion of Fig. 4, the kinetic energy of the nuclei (QM, MM and capping) is plotted. Since the MM-link atoms are not free dynamic variables in the IMOMM scheme, they have no kinetic energy associated with their motion. Thus, the kinetic energy that is plotted is that of the 3N nuclear degrees of freedom where N is the number of atoms in the real system. Here, nine of the degrees of freedom correspond to those of the three capping atoms where the masses of the capping hydrogen atoms have been rescaled to 12.0 amu. Plotted in the lower half of the plot is the total energy having the same scale as the kinetic energy. It is clear that there is no significant drift in the total conserved energy during the simulation. The inset in Fig. 4 reveals the scale at which the total conserved energy oscillates. The exceptional energy conservation [22] demonstrates that the adapted IMOMM scheme can be used to perform QM/MM molecular dynamics (including systems with explicit solvent molecules).

3.2 IMOMM frequencies

Using the IMOMM framework, we have calculated normal-mode vibrational frequencies and thermodynamic properties for a number of minimum and transition state structures, **2**–**14**, which are shown in Fig. 3. The combined QM/MM frequencies and properties have been compared to results generated from both pure QM and pure MM potential surfaces.

The adapted IMOMM method has been implemented within the Amsterdam density functional (ADF) program system [23], developed by Baerends and Ros. QM calculations for both the pure QM and IMOMM results reported were performed with the local exchange-correlation potential of Vosko et al. [24] for the gradients, while the energies reported were obtained with Becke's

Table 1. Comparison of integrated molecular orbital and molecular mechanics (IMOMM) and pure quantum mechanics (QM) frequencies and zero-point energies (ZPEs)

No[a]	Species[b]	RMS deviation in frequencies (cm^{-1})		ZPE (kcal/mol)		
		$< 1000\ cm^{-1}$	all	IMOMM	pure QM	%diff
2	*n*-butane C-C-C-C = 180°	25	85	82.6	80.0	+3.1
3	*n*-butane C-C-C-C = 60°	36	85	82.6	80.1	+3.1
4	*n*-butane C-C-C-C = 0° [TS]	49	85	82.5	80.2	+2.8
5	$N(CH_3)_3$	45	130	77.0	73.1	+5.4
6	$N(CH_3)_3$ inversion [TS]	60	152	76.1	72.0	+5.6
7	$ClCMe_3 + Cl^-$ complex	56	128	78.9	74.4	+6.1
8	$Cl–CMe_3–Cl^-$ [TS]	108	152	78.4	73.4	+6.9
9	*cis*-2-butene	80	99	68.0	65.6	+3.8
10	*trans*-2-butene	51	88	67.8	65.4	+3.6
11	dimethyl ether	36	113	50.6	48.6	+4.0
12	cyclohexane	34	116	107.8	103.3	+4.4
13	cyclohexane	36	117	107.4	103.3	+4.0
14	norbornadiene	100	167	81.9	78.5	+4.4
	Average	55	117			+4.4

[a] QM/MM partitioning illustrated in Fig. 3
[b] Species marked with the [TS] label represent transition states

exchange [18] and Perdew's correlation [19, 20] corrections to the LDA (local-density approximation) energies as a perturbation of the LDA charge density. Double -ζ STO basis sets for hydrogen (1*s*), carbon, nitrogen, oxygen and chlorine, augmented with a single 3*d* (2*p* for H) polarization function, were used. The inner shells on the carbon, nitrogen, oxygen and chlorine were treated within the frozen-core approximation. For the MM potential, the AMBER-95 [25] force field was utilized. Van der Waals parameters for chlorine were taken from Rappe et al's UFF (universal force-field) [26]. The pure QM, pure MM and QM/MM vibrational frequencies were evaluated from a Hessian constructed from the numerical differentiation of analytical energy gradients at the corresponding optimized geometry. When the link and capping atoms are displaced during the numerical differentiation procedure, the MM-link atoms are correspondingly displaced according to Eq. (12). Thermodynamic properties were evaluated according to standard textbook procedures [27, 28]. Normal modes characterized by frequencies of less than $200\ cm^{-1}$, and which correspond to torsions, were treated as pure rotations [28]. All the reported analyses were performed at 298 K. Masses for the capping hydrogen atoms in structures **2**–**14** were rescaled to that of the corresponding MM-link atom. In this way capping hydrogen atoms were assigned masses of 12.0 amu. For all calculations masses were placed at the position of the capping atom, except for the moment of inertia, for which the masses and positions of the "real system" were utilized.

Table 1 compares the frequencies and zero-point energy (ZPE) corrections derived from the IMOMM and pure QM potential surfaces. The root-mean-squared (RMS) difference between the IMOMM frequencies and the pure QM frequencies is shown for frequencies of less than $1000\ cm^{-1}$ and for all normal-mode frequencies. The RMS difference, averaged over all structures is 55 cm^{-1} for frequencies of less than $1000\ cm^{-1}$, and $117\ cm^{-1}$ for all frequencies. For the three transition states considered, that is the eclipsed conformation of *n*-butane (**4**), the planar trimethylamine inversion transition state (**6**) and chloride S_N2 transition state (**8**), value of the imaginary frequency corresponding to the transition vector compare reasonably. The IMOMM values for these vibrations are, 189i, 261i and 199i cm^{-1} for structures **4**, **6** and **8**, respectively, whereas they are 226i, 308i, and 322i cm^{-1}, respectively, for pure QM calculations.

Since our test suite contains complexes with larger MM regions than QM regions, we have examined the same frequencies calculated from a pure MM force field calculation. Vibrational frequencies based on the pure MM (AMBER) potential surface were calculated for the same set of structures with the exception of structures **7** and **8** for which the MM force field is inappropriate. The RMS difference between the pure MM and pure QM vibrational frequencies averaged over all of the sturctures was determined to be $60\ cm^{-1}$ for frequencies less than $1000\ cm^{-1}$, and $73\ cm^{-1}$ for all frequencies. Thus, for the low frequency vibrations the IMOMM and the pure MM frequencies deviate by approximately the same amount from the pure QM frequencies. In the high-frequency range, the IMOMM frequencies deviate significantly more than the pure MM frequencies. Inspection of the frequencies reveals that the IMOMM frequencies are systematically higher than the pure QM frequencies [4].

The source of this systematic deviation results from the nature of the link bond in the IMOMM scheme and the relationship between the link atoms and the capping atoms defined in Eq. (14). In other combined QM/MM schemes, such as the scheme of Kollman [2], the strength (force constant) of the link bond is primarily determined by the appropriate MM force constant. In constrast, the strength of the link bond in the IMOMM scheme is primarily determined by the strength of the corre-

[4] This is also evidenced by the systematic overestimation of the ZPEs shown in Table 1

Table 2. Comparison of IMOMM and pure QM relative thermochemical data at 298 K

No[a]	Species	$\Delta E^{b,c}$		$\Delta H_{vib}^{b,d}$		$T\Delta S^{b,e}$		ΔG^{b}	
		IMOMM	QM	IMOMM	QM	IMOMM	QM	IMOMM	QM
2	n-butane C-C-C-C = 180°	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	n-butane C-C-C-C = 60°	0.53	0.90	−0.03	0.04	−0.12	−0.09	0.62	1.03
4	n-butane C-C-C-C = 0° [TS]	4.53	5.75	−0.18	0.06	−0.26	−0.27	4.51	6.08
5	$N(CH_3)_3$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
6	$N(CH_3)_3$ inversion [TS]	4.93	7.97	−2.04	−2.06	−1.71	−1.54	4.60	7.45
7	$ClCMe_3 + Cl^-$ complex	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
8	$Cl–CMe_3–Cl^-$ [TS]	9.90	22.30	−1.64	−1.90	−2.02	−1.41	10.28	21.81
9	*cis*-2-butene	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10	*trans*-2-butene	−0.49	0.82	−0.09	−0.01	0.18	0.18	−0.40	0.99

[a] QM/MM partitioning illustrated in Fig. 3
[b] kcal/mol
[c] Internal potential energy
[d] Includes ZPE correction
[e] Includes ΔS_{rot}, ΔS_{vib} and ΔS_{trans}

sponding capping atom bond in the QM model system. Thus, if the real link bond is a C—C bond and the corresponding bond in the model QM system is a C—H bond, then the bond-stretching frequency of this C—C bond will correspond to that of a much higher C—H bond stretching frequency [5]. One can minimize this effect, as we have, by rescaling the mass of the capping atom to that of the corresponding MM-link atom. In this way, the difference in the bond-stretching frequency is not due to the light mass of the proton, but due to the difference in the bond-stretching-force constants. For the above example, a $C(sp^3)$-H bond-stretching-force constant is only about 10% larger than the corresponding $C(sp^3)$-$C(sp^3)$force constant [28, 34], and, therefore, the bond-stretching frequency of the C—C link bond in the IMOMM scheme will be about 4% higher than a typical C—C bond[6].

The results presented in Table 1 reveal that the IMOMM method may not be appropriate for determining absolute frequencies. Although this may not seriously limit the applications of the IMOMM method, it may present a problem when determining those properties derived form the frequencies such as finite temperature, ZPE, and entropic corrections. This aspect of utilizing the IMOMM method will be examined next.

3.3 IMOMM ZPEs and finite temperature corrections

Table 1 compares the pure QM and IMOMM ZPE corrections for structures **2**–**14**. The IMOMM ZPEs are higher for all structures compared to the pure QM ZPEs. The percent difference between the QM and IMOMM ZPEs averaged over all structures is only +4.4%. Since the differences are all in the positive direction, this again reveals that the frequencies generated from the IMOMM potential surface are systematically higher than the pure QM frequencies. Moreover, within each of the conformational groups(**2**–**4**, **5**–**6**, **7**–**8**, and **9**–**10**), the percent difference in the ZPE corrections is approximately constant. For example, for the conformations of *n*-butane, **2**–**4**, the percent difference are all roughly 3.0% and for the trimethylamine complex, **5**, and inversion transition state, **6**, the percent difference are both about 5.5% .

Table 2 compares the relative free energies, ΔG, and selected components determined from the IMOMM and pure QM potential surfaces. More specifically, a decomposition of the relative free energy for various conformations and/or transition states of *n*-butane (**2**–**4**), 2-butene (**9**, **10**), trimethylamine inversion (**5**, **6**) and for a chloride S_N2 reaction (**7**, **8**) are reported. Table 2 reveals that for components of ΔG that depends on the normal-mode vibrational analysis (ΔH_{vib} and $T\Delta S$), the IMOMM results compare exceptionally well to the pure QM values. The differences in ΔH_{vib} and $T\Delta S$ are of the order of 0.5 kcal/mol .The large deviations in the relative free energy, ΔG, between the two methods can be attributed to large differences in the pure QM and the combined IMOMM potential energies, ΔE.[7] However, the QM/MM partitioning in these systems is severe and the electronic structure of QM model systems is a poor representation of that in the real systems. In practice, the partitioning of the QM and MM regions is generally chosen much more judiciously as to minimize the problem of charge transfer effects across the link bond. It is not our intent to highlight this aspect of the QM/MM approach since it has been thoroughly studied elsewhere [3, 4, 30–33]

The approximations in treating the QM/MM link bonds inherent to the IMOMM scheme preclude the precise calculation of vibrational frequencies with the scheme without additional modification to the potential energy expression. However, the results presented in

[5] The same effect also exists for bond angles and torsions that are approximated by the QM model system in the IMOMM scheme

[6] 'Off-diagonal' elements are also effected, such that normal mode vibrations other than those that can be assigned to the stretcing vibrations of the link bonds are also shifted

[7] The reason for the unusually large disparity in the pure QM and QM/MM reaction barriers for the S_N2 reaction is due to the fact that electrostatic interactions between the QM and MM regions were neglected for the results presented in Tables 1 and 2

Table 2 indicate that without further modification or parameterizations, the IMOMM scheme is capable of evaluating the relative thermodynamic properties, ΔH_{vib}, ΔH_{ZPE} and ΔS_{vib}, adequately. We have found that this is true even for systems were the partitioning of QM and MM regions is so severe that the relative potential energies are not well represented by the QM/MM method. In other words, the fashion in which the link bonds are treated in the IMOMM scheme does not adversely effect the calculation of these properties which are based on the vibrational frequencies.

4 Conclusions

We have adapted the IMOMM scheme of Maseras and Morokuma to allow for practical energy conserving molecular dynamics simulations, simulations of mixed solute-solvent systems and for the calculation of harmonic vibrational frequencies. The extension is simple, but does not require a coordinate transformation from cartesians to internals which can limit the application of the original IMOMM implementation to the aforementioned types of calculations. The methodology is demonstrated by an energy conserving IMOMM CarParrinello/OPLS-AA molecular dynamics simulation of solvated 3-methylhexane where three covalent bonds cross the QM/MM boundary. We have applied the scheme to calculate normal-mode vibration frequencies and thermochemical data on a number of minimum and transition state structures. Although the absolute value of the frequencies generated by the IMOMM scheme can deviate significantly from those determined from the corresponding pure QM potential surface, the deviations are systematic in nature and of a known origin. This is an essential feature since the cancellation of errors allows for the reliable IMOMM calculations of relative thermochemical properties, namely, ΔH_{vib}, ΔH_{ZPE} and $T\Delta S$. The adaptation, albeit simple, is important in the sense that it allows free energy surfaces to be examined with the IMOMM method via molecular dynamics simulations and frequency calculations.

Acknowledgements. The authors thank Dr. P.E. Blöchl, Dr. R. Schmid and Dr. P. Margl for helpful discussions. This work has been supported by the National Sciences and Engineering Research Council of Canada (NSERC), as well as by the donors of the Petroleum Research Fund, administered by the American Chemical Society (ACS-PRF No. 31205-AC3) and by Novacor Research and Technology Corporation (NRTC) of Calgary. L.C. thanks the Ministero della Ricerca Scientifica e Technologica of Italy. T.K.W. wishes to thank NSERC, the Alberta Heritage Scholarship Fund and the Izaak Walton Killam Memorial Foundation.

References

1. Warshel A, Levitt M (1976) J Mol Biol 103: 227
2. Singh UC, Kollman PA (1986) J Comp Chem 7: 718
3. Field MJ, Bash PA, Karplus M (1990) J Comp Chem 11: 700
4. Maseras F, Morokuma K (1995) J Comp Chem 16: 1170
5. Gao J (1996) In: Lipkowitz KB, Boyd DB (eds:) Reviews in computational chemistry, vol 7. VCH, New York, p 119
6. Humbel S, Sieber S, Morokuma K (1996) J Chem Phys 105: 1959
7. Svensson M, Humbel S, Froese RDJ, Matsubara T, Sieber S, Morokuma K (1996) J Phys Chem 100: 19357
8. Théry V, Rinaldi D, Rivail JL, Maigret B, Ferenczy GG (1994) J Comp Chem 15: 269
9. Bersuker IB, Leong MK, Boggs JE, Pearlman RS (1997) J Comp Chem 63: 1051
10. Bash PA, Field MJ, Davenport R, Ringe D, Petsko G, Karplus M (1991) Biochemistry 30: 5826
11. Lyne PD, Mulholland AJ, Richards WG (1995) J Am Chem Soc 117: 11345
12. Hartsough DS, Merz KM Jr (1995) J Phys Chem 99: 11266
13. Blöchl PE (1994) Phys Rev B 50: 17953
14. Car R, Parrinello M (1985) Phys Rev Lett 55: 2471
15. Allen MP, Tildesley DJ (1987) Computer simulation of liquid. Oxford University Press, Oxford
16. Jorgensen WL, Maxwell DS, Tirado-Rives J (1996) J Am Chem Soc 118: 11225
17. Perdew JP, Zunger A (1981) Phys Rev B 23: 5048
18. Becke A (1988) Phys Rev A 38: 3098
19. Perdew JP (1986) Phys Rev B 33: 8822
20. Perdew JP (1986) Phys Rev B 34: 7406
21. Nosé S (1984) Mol Phys 52: 255
22. Remler DK, Madden PA (1990) Chem Phys
23. Baerends EJ, Ros P (1973) Chem Phys 2: 52
24. Vosko SH, Wilk L, Nusair M (1980) Can J Phys 58: 1200
25. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz Jr. KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995) J Am Chem Soc 117: 5179
26. Rappé AK, Casewit CJ, Colwell KS, Goddard III WA, Skiff WM (1992) J Am Chem Soc 114: 10024
27. McQuarrie DA (1973) Statistical thermodynamics. Harper and Row, New York
28. Hehre WJ, Radom L, Schleyer PR, Pople JA (1986) Ab initio molecular orbital theory. Wiley, New York
29. Silverstein RM, Bassler GC, Morril TC (1991) Spectrometric identification of organic compounds, 5th edn. Wiley, New York
30. Stanton RV, Hartsough DS, Merz Jr. KM (1995) Comput Chem 16: 113
31. Bakowies D, Thiel W (1996) J Phys Chem 100: 10580
32. Matsubara T, Sieber S, Morokuma K (1996) Int J Quantum Chem 60: 1101
33. Matsubara T, Maseras F, Koga N, Morokuma K (1996) J Phys Chem 100: 2573