Hybrid Quantum Mechanics/Molecular Mechanics Approaches

Paul Sherwood

Modern Methods and Algorithms of Quantum Chemistry,
Proceedings, Second Edition, J. Grotendorst (Ed.),
John von Neumann Institute for Computing, Jülich,
NIC Series, Vol. 3, ISBN 3-00-005834-6, pp. 285-305, 2000.

HYBRID QUANTUM MECHANICS/MOLECULAR MECHANICS APPROACHES

PAUL SHERWOOD
CLRC Daresbury Laboratory
Warrington, WA4 4AD
United Kingdom
E-mail: p.sherwood@dl.ac.uk

An overview is provided of the range of approaches to hybrid QM/MM (quantum mechanics/molecular mechanics) calculations. The factors considered include the choice of QM and MM methods, the construction of the total QM/MM energy expression, the nature of the QM/MM coupling, and the treatment of bonds between QM and MM regions. The practical issues associated with handling the increased conformational complexity of macromolecular systems, and the construction of QM/MM codes are discussed.

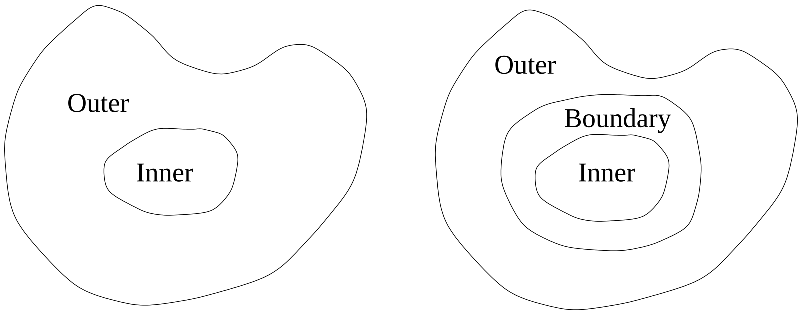
1 Introduction

The modelling of complex chemical systems is still a daunting challenge. We have at our disposal sophisticated first-principles methods for simulating reactions and electronic processes to high accuracy but these are limited by their computational cost to small molecules. The systems of chemical interest in computational biology and catalysis are often condensed phase systems with many thousands of participating atoms. While significant progress is being made in the development of quantum chemical approaches applicable to large systems[1], it is clear that to treat complex biological and catalytic systems we still need to be able to integrate a range of computational chemistry methodologies with differing accuracies and cost. By embedding a quantum mechanics calculation in a classical molecular mechanics model of the environment, the hybrid QM/MM schemes attempt to incorporate environmental effects at an atomistic level, including such influences as mechanical constraints, electrostatic perturbations and dielectric screening. Since the first published example from the field of computational enzymology[2] many QM/MM schemes have been implemented and applied in a wide variety of chemical applications.

The subject has steadily developed, and the last couple of years in particular have seen rapid increase in the rate of publication of QM/MM applications. The availability of implementations within commerical packages (e.g. CHARMM and Gaussian98), and improvements in the available computational resources will doubtless contribute to a continued increase in popularity. Recent reviews of the subject include those by Gao[3] and Mordasini and Thiel[4].

2 Terminology

The first, trivial, step is to divide the entire system (E), into inner (I) and outer (O) regions (Figure 1a). The objective in all cases is to use the QM calculation to model processes in the inner region and to use MM to model the outer region. As always



a) Inner and Outer Regions b) Boundary Region

Figure 1. Terminology for regions used in the QM/MM scheme

the devil is in the detail, and it is rarely possible to simply write down the total energy in terms of two, non-overlapping subsystems, Very often the interactions between the systems are sufficiently strong (the obvious example being the presence of a chemical bond) to ensure that a QM calculation on the inner region alone is not sensible. Some form of termination, or treatment of the boundary is required. For the purpose of this article, we will classify the approaches to this QM termination into two groups;

1. those based on **link atoms**[5,6], additional centres added to the QM calculation but not present in the entire system(E). The position of centres in the region (L) are either viewed as independent variables or as a function of the positions of atoms in both (I) and (O) regions (*vide infra*). Link atoms are generally invisible to the MM calculation.

2. those having a **boundary region**, a subset of the centres of the systems which will feature in both QM and MM calculations (Figure 1b). In the QM calculation these centres can have a range of roles, ranging from a re-parameterised semi-empirical Hamiltonian to an *ab-initio* pseudopotential or a frozen hybrid orbital.

3 Overview of QM/MM Schemes

In this section we discuss some of the choices that distinguish the range of possible QM/MM schemes.

3.1 The QM/MM Energy Expression

Given the definitions given in Section 2 we can group QM/MM schemes into two broad classes:

3.1.1 Additive Schemes

This general classification is applied to schemes for which the QM and MM energies are considered complementary, the total energy of the system is obtained by adding them together and applying any coupling terms or corrections.

For link atom based schemes the total energy expression takes the form of Equation 1 (E = entire, I = inner, O = outer and L = link);

$$E^{E}_{QM/MM} = E^{O}_{MM} + E^{I,L}_{QM} + E^{I,O}_{QM-MM} - E^{I,L}_{Corr} \tag{1}$$

Here the term $E^{I,O}_{QM-MM}$ is QM-MM coupling term includes all terms that couple the two regions, for example MM-style bonding and van der Waals interactions and modifications to the QM Hamiltonian to reflect the influence of some or all of the atoms in the outer region.

The correction term $E^{I,L}_{Corr}$ represents terms designed to reduce the dependency of the total energy on the centres in the link atom region[7]. How this term is handled depends on the the choice of link atom coordinates, as discussed below in Section 3.5.2. However, this term is often neglected.

Additive schemes are probably the most widely adopted approaches to QM/MM calculations, particularly in the biomolecular area, with the AMBER[5,8,9] and CHARMM [6,10,11,12] based implementations being important examples.

The main problems with the scheme arise because it is difficult to compute the coupling term $E^{I,O}_{QM-MM}$ accurately in the presence of the link atoms[13], particularly if electrostatic perturbations to the QM Hamiltonian are included.

For boundary-region based methods, which are usually additive in type, the energy total may be written

$$E^{E}_{QM/MM} = E^{O,B}_{MM} + E^{I,B}_{QM} + E^{I,O,B}_{QM-MM} \tag{2}$$

There is no need for any link atom corrections, but since the boundary atoms are treated by both QM and MM methods it is important that the classical energy expression be modified to avoid multiple counting of interactions.

This class of scheme has been adopted most widely for studies involving strongly ionic materials where the boundary region is treated by model- or pseudo-potentials. However, as discussed below in Section 3.5.1, a number of treatments designed for more covalent systems also eliminate link atoms, placing a re-parameterised atomic description or a frozen orbital at the site of the first MM atom.

3.1.2 Subtractive Schemes

Here the entire system is treated by MM, and a third calculation on the inner region at the MM level is performed to eliminate multiple counting[7]. The approach is generally applied to link-atom based schemes, in which case the total energy may be written:

$$E^{E}_{QM/MM} = E^{E}_{MM} + E^{I,L}_{QM} - E^{I,L}_{MM} \tag{3}$$

The coupling term $E^{I,O,L}_{QM-MM}$ is no longer required as all interactions between inner and outer regions are handled at the MM level of theory, in the E^{E}_{MM} term. The handling of link atom corrections here occurs implicitly as a result of the subtraction. It is necessary that the forces in the link region arising from the difference between the QM and MM representations ($E^{I,L}_{QM} - E^{I,L}_{MM}$) remain small for all reasonable positions of the link atoms. It is therefore particularly valuable in this case to use a forcefield designed to reproduce forces at the particular QM approximation used for the inner region.

If the process under investigation involves changes in chemical bonding it will become more difficult to provide a suitable forcefield. However, if the inner region is large enough, it is possible to ensure that the contribution to the total energy from atoms in the interior of the inner region completely cancel when the subtraction $E^{E}_{MM} - E^{I,L}_{MM}$ is performed, and there is therefore no requirement for the forcefield to model the energetics of the reaction. Nevertheless, the forcefield must be able to compute the interaction between the reacting centre and the outer region at the MM level of theory, which requires, for example, partial charges for the former. Since the charge density of the inner region may change during the course of the reaction this can be a demanding requirement. Subtractive schemes are clearly not suitable for cases in which the electronic structure of the QM region is expected to be significantly perturbed by interaction with the environment. However, in most application areas explored so far this approximation has not proved problematic[14]. Where good quality forcefields are available the approach can be very accurate since there are no problems with interactions between the link atom region and the classical environment.

To date the main applications have been in the areas of organometallic[15] and zeolite chemistry[16,14,17] This scheme is actually quite general and can be used to coupling different levels of QM theory, as exemplified by the IMOMO [18,19] scheme of Morokuma.

The subtractive and additive models as defined here are closely related, as discussed by Bakowies and Thiel[7], where the subtractive model is used as a starting point for the derivation of link-atom corrected additive model.

3.2 The Choice of QM model

The choice of QM method will not be dealt with in detail here as it does not a fundamentally affect the design of a QM/MM scheme and will largely be governed by the same critera that apply to pure QM calculations.

Since the first Warshel and Levitt[2], study schemes based on semi-empirical methods have dominated the field for biological applications, and for reasons of computational cost such schemes are likely to remain important for applications incorporating molecular dynamics.

Approaches incorporating parameterised Hamiltonians include the MOLARIS implementation[20] (using an empirical valence bond (EVB) scheme) and the MM-VB scheme of Bernardi, Robb and co-workers[21,22]

A large number of *ab-initio* schemes based on Hartree-Fock[5,10,23,24] and density functional[9,25,26,11,16] approaches have been implemented. Recently a number of approaches based on Car-Parrinello DFT codes have been reported[27,28,30].

3.3 The Choice of MM model

The significance of the choice of MM scheme depends on whether the additive or subtractive schemes are chosen, since within the subtractive scheme any forcefield can be used (see the discussion in Section 3.1.2).

Within additive schemes, type of MM model can have significant influence on the treatment of the boundary, since different classical approaches differ markedly with respect to the handling of both bonded and non-bonded interactions. The most important distinction is that between:

- valence force fields, exemplified by the biomolecular force fields (CHARMM[31], AMBER[32]) and a number of more general purpose forcefields including MM3[33] and the consistent force field (CFF[34]) constructed from energy terms such as bond stretches, angle bends *etc.*

- ionic forcefields in which the principal terms are the electrostatic and short-range (van der Waals) forces, exemplified by force fields based on the shell model[35,36,14].

The choice affects the construction of the MM model in two major respects:

1. The choice of forcefield influences the atomic partial charges thus affecting the long-range QM/MM interactions, as for most current implementations the same charges are used for MM...MM and MM...QM interactions (in principle it would be possible to construct schemes based on two sets of MM charges but this approach has not been widely adopted). For a given material, (for example the zeolites[37,38]) the shell model forcefields tend to be based on larger charges than those generated by fitting to the electrostatic potential. In many case the ionic forcefields employ formal ionic charges.

2. Handling of bonding and close interionic contacts between QM and MM regions will generally follow the same approach as treatment of similar interactions within the MM region. In the valence forcefield case it is easy to identify the terms involved, typically bond-stretch, angle-bend and torsion terms that are needed, and it is simple to delete those that correspond to terms handled by the QM interaction. For ionic force-fields the short-range QM..MM attractive terms will come from the presence of the MM charges in the QM Hamiltonian, and cannot readily be separated from the long-range interactions.

For these reasons, additive schemes based on link atoms are easier to construct with valence forcefields. The ionic class of forcefields can be used in boundary-region additive schemes, but only if the forcefield charges can generate the correct electrostatic potential in the QM region and thus the correct interionic forces. This is more likely to be true for highly ionic materials, unless significant parameterisation of the boundary region is carried out.

3.4 Handling of the QM/MM non-bonded coupling terms

Within most classical modelling schemes, the non-bonded interactions comprise electrostatic and short-range (or van der Waals) forces. In the context of the additive QM/MM schemes, the same decomposition is applied to the non-bonded interaction between QM and MM centres.

3.4.1 Short-range or Lennard-Jones terms

The treatment of the short-range QM/MM interaction generally follows the model used in the MM calculation. Re-fitting of the non-bonded parameters is often carried out, particularly in the case of solvation studies[39,40,41] or where the details of non-bonded contacts are particulary important[42].

3.4.2 Electrostatic terms

Bakowies and Thiel[7] defined three ways of treating QM/MM electrostatic interaction, labelled A-C, as follows.

A mechanical embedding, in which the QM calculation is essentially performed in the gas phase, without electronic coupling to the environment. The electrostatic interaction between QM and MM regions is either omitted or performed by the MM code, using a classical point charge model for the QM charge distribution (*e.g.* a potential derived charge model).

B electrostatic embedding, in which the classical partition appears as an external charge distribution (*e.g.* a set of point charges) in the QM Hamiltonian. The polarisation of the QM region by the MM charge distribution thus occurs as part of the QM electronic structure calculation. The partial charges used to describe the MM distribution are frequently taken to be those used in the forcefield [5,6,43], relying on the use of electrostatic properties in the forcefield charge derivation. When using an aluminosilicate CFF forcefield[37] for electrostatic QM/MM modelling of zeolites[44] it was found necessary to replace the original MM charges with those derived by fitting to electrostatic potentials. Charge equilibration schemes, which determine the MM charges as a function of geometry have also been employed[45,46].

In *ab-initio* schemes it is clear that the electrostatic embedding scheme should be implemented, at least at long range, by adding the contribution of the MM point charges to the 1-electron Hamiltonian. However, within the semi-empirical formalism the definition of the electrostatic potential is more ambiguous as a result of the overlap approximations used, and alternative formulations for the 1-electron integral terms have been suggested[6,47,45].

C polarised embedding, in which the polarisation of the MM region in response to the the QM charge distribution is also included. Intuitively this makes most sense when the forcefield incorporates polarisation as unpolarised forcefields implicitly incorporate MM polarisation in their parameterisation, and care must be taken to ensure such implicit contributions to not occur in

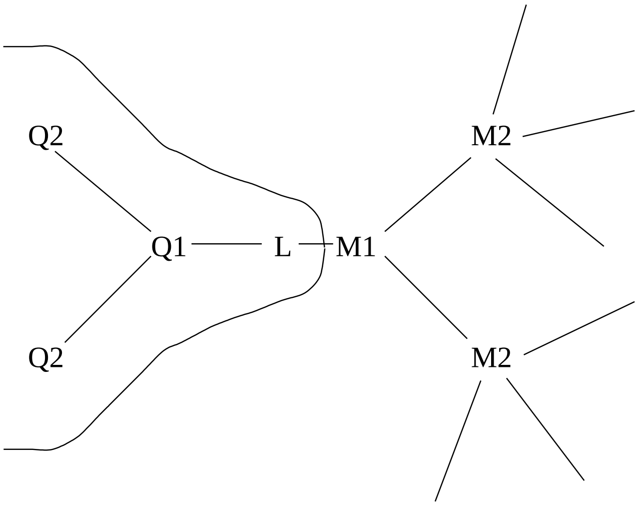


Figure 2. Labelling of atoms in the link atom region

the QM/MM potential. A variety of models for the classical polarisation are available, including the shell model[35], and coupled distributed atomic polarizabilities[48]. Polarisation of MM atoms close to the QM region (*e.g.* those connected by link-atom terminated bonds) were found to be unphysically large, leading to the suggestion that these atoms be treated as unpolarizable[7].

In model C the result of the change in MM charge distribution from the classical polarisation is not propagated to the QM calculation resulting in a non variational total energy.

D While not part of the original definition, model D[49] is defined as an extension to model C where QM and MM polarisations are made self-consistent, either by iterative solution of the SCF and polarizability problems[50], or by matrix inversion techniques, as exemplified by the the Direct Reaction Field (DRF) model[51,52,53,54,55].

3.5 Termination of the QM region

In those cases where there are bonds or strong ionic interactions between the QM and MM regions it is necessary to introduce some termination of the QM calculation, either through the link atom or boundary region approaches. Figure 2 indicates the labelling adopted in the discussion below for those models incorporating link atoms.

3.5.1 Chemical Nature of the Termination

For termination of sites where a covalent bond has been broken addition of a link atom is the most popular approach. An extra nuclear centre is introduced, together with basis functions and electrons required to form a covalent bond to the QM region that will mimic the bond to the MM region. The simplest and still most popular choice is the use of a hydrogen atom[5].

There are clearly chemical differences between hydrogen and the chemical group it replaces. Within empirical and semi-empirical schemes, an obvious enhancement is to adjust the link atom parameterisation to mimic more closely the modelled group. Recently, Antes and Thiel[57] have described the semi-empirical Adjusted Connection Atom (ACA) scheme, in which the link atom is replaced with a boundary atom with parameters chosen to model a methyl group. A related approach within *ab-initio* based QM/MM methodologies is to place a pseudopotential at the MM site to mimic the electronic properties of the replaced bond[58].

For embedding treatments of highly ionic materials the main requirement of the termination is to stop the unphysical polarisation of the charge density of the QM cluster by the adjacent cations. "Leakage" of charge from the cluster will clearly occur to an increasing extent as the QM basis set is extended. Most treatments therefore include a pseudo- or model- potential at the cationic sites, without any valence electrons or basis functions[59,60]. Modelling the anions by model potentials[24,61] is a further refinement, which may be more important for materials with a more significant covalent character. Here the use of bare anionic charges might be expected to lead to poor structural predictions which can be ameliorated by adjustments to the potential.

All the approaches listed above assume that once the termination potential has been set up a full SCF calculation is performed on the resulting cluster. An alternative approach is to constrain the SCF solution to reflect the influence of the bonds that have been omitted. The local self-consistent field (LSCF) scheme[62,63] of Rivail and coworkers involves the preliminary calculation of the localised bond orbital, which is then frozen during the calculation. The generalised hybrid orbital (GHO) scheme of Gao *et al.*[64] constrains the hybridisation of the terminating centre but allows adjustment of the QM-MM bond itself, allowing changes in local geometry to be handled.

Similar approaches based on *ab-initio* wavefunctions have been less widely used to date, but include adaption of the LSCF scheme [65] and a recent implementation by Philipp and Friesner[43] which incorporates geometry changes at the boundary.

While the frozen hybrid orbital approaches promise to mimic the electronic properties of the extended system, problems with the calculation of the $E^{I,O}_{QM-MM}$ term in Equation 1 still remain. As described by Philipp and Friesner[43] it is probably necessary to combine the definition of the terminating orbitals with reparameterisation of the link atom region to obtain accurate conformational energies.

3.5.2 Coordinates and forces for the termination sites

In contrast to the boundary atom schemes, the appearance of additional QM centres in the link atom approaches leads to additional variability in the definition of the coordinates and forces. Inital coordinates can be chosen by placing the link atom on the bond that is being terminated, but once a geometry optimisation or molecular dynamics run has been started there are a number of different ways of updating the coordinates and handling the forces.

1. Optimised Link Atoms

   In these approaches, the link atom coordinates are added to the atom list used in the geometry optimisation or molecular dynamics scheme, and the coordinates are free to vary[6].

   Sometimes some additional forcefield terms may be added to favour the positioning along the M1-Q1 bond[66] This is particularly important when link atom corrections or (*vide supra*) lead to small and unphysical forces on the link atoms.

2. Constrained Link Atom Procedures

   If the link atom coordinates can be written as a function of the real atom coordinates it is possible to eliminate them from the set of coordinates used in the optimisation or dynamics. Such an elimination is particularly desirable for molecular dynamics and the evaluation of vibrational frequencies, which would be modified by coupling to any independent link atom motions.

   Since the link atoms have non-zero forces it is necessary add a term to the real atom forces to account for the changes in the link atom position resulting from movement of the real atoms. In the IMOMM scheme of Maseras and Morokuma[68] this was implicitly performed by using working in internal coordinates such that the same internal coordinates were used to position the link atoms and to define the corresponding real atom (M1) coordinates. The internal coordinate force can then be obtained by adding the QM and MM contributions.

   When working in pure cartesian coordinates the same effect can be realised by using the chain rule to establish the contribution to the forces on the real atoms X_i.

   $$\frac{\partial E}{\partial X_i} = \frac{\partial E}{\partial X_i} + \frac{\partial E}{\partial X_l} \cdot \frac{\partial X_l}{\partial X_i} \tag{4}$$

   where the derivative $\frac{\partial X_l}{\partial X_i}$ is a 3x3 matrix describing the coupling of the link atom and real atom motions as a function of the constraint term. There is a term for each atom i which appears in the definition of the position of the link atom l, For the case of a link atom placed at a fixed position along the QM–MM bond there will be a correction term for the atom at each end of the bond in question (M1 and Q1 in Figure 2). This approach is used by the ChemShell implementation[69] and by the QM-Pot scheme of Sauer *et al.*[16]. A similar adaptation to the IMOMM scheme has been published by Woo *et al.*[70].

   When using constrained link atoms in additive QM schemes some adjustment to the forcefield is required. The MM force constant for the angle bend Q2-Q1-M1 is in effect supplemented by the bending potential for the group Q2-Q1-H, since the restoring forces acting on H are transferred to atom M1 as described in equation 4. The approach adopted in the ChemShell scheme is to delete the MM term on the basis that the link atom bending potential will replace it. Greater accuracy could be achieved by fitting a modified MM parameter to reflect the presence of the link atom. Similar considerations apply to the torsion angles of the form Q3-Q2-Q1-M1. As a general rule, the Q1-H distance is kept fixed. It has been suggested that variation of the Q1-H distance provides a simple way to tune the electronic characteristics of the termination[69].

Similar considerations apply to the boundary atom schemes. It is generally assumed, or considered as a requirement for the boundary atom parameterisation, that the QM calculation incorporating a model potential will generate a suitable geometry for the short QM...MM contacts.

3.6 Modifications to the Classical charge distribution at the boundary

Within all of the polarised QM/MM schemes (models B-D in Section 3.4.2) problems can be expected when point charges modelling the MM region closely approach the QM region. In the absence of link atoms, close approach is usually prevented by the non-bonded interaction potential which is repulsive at short range. However in the region of bonds across the QM-MM boundary some adjustment to the classical charge distribution is essential, as the nearest point charges to the QM region will be at most a single bond distance away. In the case of terminating link atom (e.g. hydrogen) the link atom will be almost superimposed on the first classical atom (M1 in Figure 2).

Clearly this problem will be more severe when large basis sets are used, and in fact it is possible to disregard it in the semi-emprical case[6]. Antes and Thiel[71,49] have discussed a variety of approaches to the problem and suggested the L1-L3 classification included below.

3.6.1 Selective deletion of one-electron integral terms

For QM calculations with small basis sets the leading spurious interaction is that of the basis functions on the link atom with the classical MM charges. Since the link atom is an artefact of the QM/MM scheme it has been suggested that the model can be improved simply by deleting the 1-electron Hamiltonian contributions involving link atom basis functions and the full set of MM charges (scheme L1[49]).

Antes and Thiel also defined a further scheme, denoted L3, in which all 1-electron terms are included with the exception of those of the that involve basis functions on the link atom *and* the charge on centre M1.

3.6.2 Deletion of selected atomic charges

Perhaps the simplest way of dealing with the charge on the nearby classical centres is to delete them from the Hamiltonian. The QUASI Gaussian/AMBER scheme[5] omitted any MM charges less than 4 bonds removed from any QM atom. Waszkowycz *et al.*[8] modified this approach so that only a single MM charge at the M1 site was excluded.

The danger of these schemes is that simply deleting charges according to the connectivity will often result in the remaining MM atoms, as experienced by the QM region, appearing to have a net charge. Such an artefact will have particularly serious effect on computed energies for processes in which the total charge of the QM region is modified, such as protonation reactions.

Many biochemical forcefields have the feature that sets of neighbouring atomic charges can be grouped together such that the total group charge is an integer (usually zero)[72]. This is a convenient feature for a forcefield as it enables a molecule built by combining these charge groups to have an integral charge without any adjustment of the atomic charges being necessary. In the L2 scheme[49] the charges on the charge group containing atom M1 (in Figure 2) are neglected when building the QM Hamiltonian. While this will clearly remove some significant physical interactions, the fact that the charges removed will sum to zero will ensure that the total MM charge experienced by the QM calculation is correct. The leading term that is missing will be the dipole moment of the first charge group.

A series of tests on protonation reactions were used to evaluate the schemes L1 to L3[49], using semi-empirical, DFT and MP2 wavefunctions. The differences between the schemes were observed to be more pronounced for the *ab-initio* wavefunction, as expected from an analysis of the influence of the integral approximations[49]. L1 was recommended only at the semi-empirical level, for *ab-initio* wavefunctions the selective integral deletion was considered to lead to problems of imbalance in the electrostatics for the link atom region. The L2 scheme was found to be robust for all types of wavefunctions, and for *ab-initio* studies using forcefields based on neutral groups it appears the obvious choice. The results for L3 were less consistent and this scheme was not recommeded.

3.6.3 Charge Shifting Schemes

Within our calculations on zeolites[44,73,69] a different scheme has been adopted. Since the aluminosilicate forcefield used[37] does not consist of charge groups, there is no simple subset of MM atoms that can be deleted without associating an artificial total charge with the local MM environment.

The first stage of the approach is to adjust the charge on centre M1 to account for the deletion of the M1-Q1 bond. For an aluminosilicate forcefield with silicon charges of $+2x$ and oxygen charges of $-x$, each Si-O bond can be considered to contribute a charge of magnitude $0.5x$ to the atom at each end, since there are 2 bonds to each O and 4 to each Si. The idea of deriving electrostatic models from a sum of dipolar contributions from the bonds to each atom has been used for more general chemical sytems, for example the MM3 forcefield[33] uses an electrostatic model including bond dipole terms. For each QM-MM bond each atom M1 is involved in, the charge on M1 is reduced by $0.5x$, thus ensuring that the defect created in the MM lattice is electrically neutral.

Since the M1 centres still have finite charge, further adjustments are necessary. The approach adopted sets M1 to zero, but for each of the connected atoms M2 a charge adjustment is made to conserve the total charge of the MM system. As M1 and M2 will usually have charges of opposite sign the M2 atoms will have reduced charges. To compensate for the dipole that has been created by the charge shift a pair of equal and opposite point charges are added close to each M2 centre along the M1-M2 bond direction. The resulting MM charge distribution therefore has the same charge and dipole moment to that of the MM defect created by the first stage, but the charge distribution close to the link atoms has been eliminated.

3.6.4 Gaussian Blur

Brooks has suggested[66] that the problems of close approach of the MM charges to the QM region are largely the result of the representation of this charge distribution by a point charge model. The "Gaussian blur" approach replaces the point charges for selected MM centres with a Gaussian charge distribution. An implementation of this approach[67] is included in the coupling of GAMESS-UK[84] and CHARMM[31].

4 The Issue of Conformational Complexity

One reason for the slow uptake of the *ab-initio* based QM/MM methods has been the computational costs. Although the cost of each energy and gradient evaluation is similar to that for a calculation on a small-molecule cluster model, the number of energy and gradient evaluations needed is likely to be much greater as a result of the larger number of degrees of freedom of the system.

In favourable cases it is possible to concentrate on a particular conformation for the environment, perhaps by simulated annealing or a related conformational search technique, using a cheaper energy function. The QM/MM study can then proceed without any attempt to repeat the conformational search, and relaxation of the MM coordinates can be repeated for each geometry change of the QM core. An example is the IMOMM scheme[68] which restricts the main geometry optimisation loop to those coordinates describing the QM geometry. At each step of this optimisation the remaining (MM) coordinates were optimised. Since IMOMM is a mechanical embedding scheme this relaxation does not change the QM energy and the QM calculation can therefore be skipped. By ensuring that the derivative of the energy with respect to the MM coordinates was zero at all times, standard optimisation algorithms, including transition state (TS) searching could be used. In favorable cases the number of QM calculations required will be similar to a pure QM calculation.

A related approach, but based on cartesian coordinates and polarised embedding was adopted by Turner *et al.*[12] for geometry optimisation and transition state determination for reactions in enzymes and solution. The MM relaxation was performed in the field of a classical model for the QM region, obtained by fitting to the QM contribution to the forces on the MM atoms at each QM geometry. As well as the savings in computational cost that arise from the reduced number of QM evaluations, the division of the coordinate set also proves useful for TS optimisations. If the TS search can be restricted to a reduced number of coordinates the stability of algorithms such as P-RFO[74] can be improved. This is especially true when the MM coordinate set is characterised by many soft eigenmodes and incipient transion states, as the case for solvated systems.

Recently Billeter *et al.*[75] have developed a geometry optimisation scheme based on delocalised internal co-ordinates[76]. The decomposition of the system into two or three sets of coordinates[12] is extended to an arbitrary number of residues, each subject to optimisation in turn. This “divide and conquer” approach enables linear scaling properties with system size to be obtained.

In cases where a number of QM/MM energies and forces are required for geometries with similar QM structures it is possible to reduce the number of QM energy evaluations required, interpolation approaches have been suggested for this purpose[77].

5 Software Implementation

In many respects the issues governing implementation of QM/MM computer codes are similar to those associated with the individual QM and MM methods. Most of the coupling terms are readily computed using the machinery present in either the QM or MM packages. However, it is worth giving brief consideration to a couple of implementational issues.

5.1 Program architecture

Given that the starting point is working QM and MM codes, QM/MM implementations can be considered to fall into three groups:

(i) those based on classical modelling packages, with a QM code integrated as a force-field extension

(ii) those based on a QM packages, incorporating the MM environment as a perturbation

(iii) modular schemes in which a central control program is provided and a choice of both QM and MM methods is left open.

Probably the most popular approach to date has been (i), for the sound reason that the modelling tools present in a typical MM/MD package are well suited for manipulation of large, complex chemical systems. A good example of this approach is the series of QM implementations within the CHARMM package [6,10,11,67]. However, one area in which the functionality of a traditional macromolecular program may need enhancing is the search for transition states. Turner *et al.*[12] extended the QM/MM capabilities of CHARMM by adding a driver package (GRACE) with capabilities including eigenvector-following search methods, and on-the-fly charge model fitting.

Option (ii) is particulary well suited when the tools required are those associated with small-molecule quantum chemistry, for example internal (Z-matrix coordinates) and Hessian-based transition-state searching using active coordinates which involve only QM atoms, By keeping the environment stationary, or fully relaxed at each step, the conformational complexity is hidden and the problem resembles a QM optimisation. Morokuma’s IMOMM scheme[68] has the appearance of being designed in this way. A coupling of GAMESS-UK and AMBER[8] followed a similar philosophy. Breuer *et al.*[78] used the framework of the MSI Cerius-2 package to introduce a mechanically-coupled MM model for the environment in an implementation exploiting the RI-DFT and redundant internal coordinate optimisation capabilities of TURBOMOLE[79].

We will discuss option (iii) in a little more detail, not because it is intrinsically superior, but because the QM/MM development work at Daresbury has been concentrated on this approach, and we therefore have more experience with the benefits and pitfalls of the approach.

The benefits can be summarised as follows

(a) A modular construction offers the greatest flexibility, and is particularly valuable if the same program system is to be used for investigation of an number of different types of chemical systems, each requiring different classes of forcefields. A range of QM schemes can readily be supported within a common environment, and commercial packages can be exploited in some of the simpler models.

(b) If care is taken to ensure that the component packages are modified to a minimal extent this approach also promises to provide the best hope for substituting up-to-date versions of the QM and MM packages when they become available.

The drawbacks are also significant, including

(a) additional programming complexity is introduced by the need to try and generalise the interfaces

(b) the need to provide, as part of the system a set of tools (*e.g.* structure manipulation, geometry optimisation, MD) which can be used with any choice of QM and MM methods.

(c) there are efficiency implications of attempting to keep the QM and MM calculations independent, particulary with respect to data transfer between the programs.

Examples of systems of modular design include the coupling of AMBER[32] to CADPAC[81] and MNDO[82] by Antes and Thiel[49] and the ChemShell package from our Laboratory[83]. The ChemShell package currently has a range of interfaces, including GAMESS-UK[84], DL_POLY[86], MNDO94[82], TURBOMOLE[79], GULP[85], CHARMM[31], GROMOS[87] and MOPAC[80] Interfaces with a variety of codes including Gaussian94[88], CADPAC[81] and AMBER[32] are in various stages of development. Program control is provided by a scripting language (Tcl[89]), which can be used to construct dynamics simulation protocols and geometry optimisation algorthms (using a suite of standard 2nd-order algorithms). Geometry optimisation using the HDLCOpt scheme[75] is also under development within the QUASI[90] project.

5.2 HPC and parallel implementation

The approach The efficient exploitation of HPC resources provides a challenge to the construction of QM/MM schemes, particularly the those of modular design.

The significant overheads associated with loading programs onto a large parallel computer force the requirement that all of the composite programs be linked into a single executable image, an awkward process since many of the components are themselves very large and complex packages, with different parallelisation models. HPC developments of the ChemShell package are taking place as part of the QUASI project[90], and include integrating the parallel version of a number of the component codes.

6 Summary and Outlook

This brief review of the QM/MM approach has emphasised the variety of ways that QM and MM calculations can be combined. It is by no means exhaustive, and the subject is now expanding rapidly.

The subtractive schemes have the advantage of relative simplicity of implementation, and the fact that there is no need to validate the QM/MM interaction. The increased reliance on the forcefield is a potential problem but this is increasingly being addressed by forcefields derived from *ab-initio* data.

Additive approaches based on link atoms, whether simple implementations with rather *ad hoc* treatments of boundaries, or more complex schemes incorporating extensive QM/MM parameterisation, are likely to suffer for some time from a lack of user confidence. As may be clear from the number of variations that are possible it will probably be difficult to get exactly the same answer with two separarate implementations, and, like the forcefields themselves, the methodology will gradually gain acceptance on the basis of experience. Cross-checks between methods and validation studies (*e.g.* comparison with pure QM results) will continue to be important for some time to come.

One of the chief challenges will be exploration of conformational space, particularly in those applications areas for which semi-empirical methods are not applicable. Large scale parallel computing can be expected to play an increasingly important role here.

References

1. See for example the article of Head-Gordon in this volume.
2. A. Warshel and M. Levitt, Theoretical Studies of Enzymic Reactions: Dielectric, Electrostatic and Steric stabilisation of the carbonium ion in the reaction of Lysozyme. J. Mol. Biol. **103**, 227 (1976).
3. J. Gao, Methods and applications of combined quantum mechanical and molecular mechanical potentials, p.119-185 in *Reviews in Computational Chemistry*, VCH Publishers, Inc., New York (1996).
4. T.Z. Mordasini and W. Thiel, Combined quantum mechanical and molecular mechanical approaches, Chimia **52(6)**, 288-291 (1998).
5. U.C. Singh and P.A. Kollman, 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. Comp. Chem. **7(6)**, 718-730 (1986).

6. M.J. Field, P.A. Bash, and M. Karplus, A Combined Quantum Mechanical and molecular mechanical potential for molecular dynamics simulation, J. Comp. Chem. **11(6)**, 700-733 (1990).
7. D. Bakowies and W. Thiel, Hybrid models for combined quantum mechanical and molecular mechanical approaches., J. Phys. Chem. **100(25)**, 10580-10594 (1996).
8. B. Waszkowycz, I.H. Hillier, N. Gensmantel, and D.W. Payling, Combined quantum mechanical-molecular mechanical study of catalysis by the enzyme phospholipase-A2 – an investigation of the potential-energy surface for amide hydrolysis, J. Chem. Soc., Perkin Trans. 2 , 2025-2032 (1991).
9. R.V. Stanton, D.S. Hartsough, and K.M. Merz, Calculation of solvation free energies using a density functional/molecular dynamics coupled potential, J. Phys. Chem. **97**, 11868-11870 (1993).
10. Y.S. Lee, M. Hodoscek, B.R. Brooks, P.F. Kador, Catalytic mechanism of aldose reductase studied by the combined potentials of quantum mechanics and molecular mechanics, Biophys. Chem. **70(3)**, 203-216 (1998).
11. P.D. Lyne, M. Hodoscek and M. Karplus, A hybrid QM-MM potential employing Hartree-Fock or density functional methods in the quantum region, J. Phys. Chem. A **103(18)**, 3462-3471 (1999).
12. A.J. Turner, V. Moliner, and I.H. Williams, Transition-state structural refinement with GRACE and CHARMM: Flexible QM/MM modelling for lactate dehydrogenase, Phys. Chem. Chem. Phys. **1(6)**, 1323-1331 (1999).
13. K.P. Eurenius, D.C. Chatfield, B.R. Brooks, and M. Hodoscek, Enzyme mechanisms with hybrid quantum and molecular mechanical potentials .1. Theoretical considerations, Int. J. Quant. Chem. **60(6)**, 1189-1200 (1996).
14. M. Sierka and J. Sauer, Structure and reactivity of silica and zeolite catalysts by a combined quantum mechanics-shell-model potential approach based on DFT, Faraday. Discuss. **106**, 41-62 (1997).
15. M. Svensson, S. Humbel, R.D.J. Froese, T. Matsubara, S. Sieber, and K. Morokuma, ONIOM: A Multilayered Integrated MO+MM Method for geometry optimizations and single point energy predictions. A test for Diels-Alder reactios and $Pt(P(t\text{-}Bu)_3)_2 + H_2$ oxidative addition, J. Phys. Chem. **100**, 19357-19363 (1996).
16. U. Eichler, C.M. Kölmel, and J. Sauer, Combining *ab initio* techniques with analytical potential functions for structure predictions of large systems: Method and application to crystalline silica polymorphs, J. Comp. Chem. **18(4)**, 463-477 (1997).
17. D. Nachtigallova, P. Nachtigall, M. Sierka, and J. Sauer, Coordination and siting of Cu^+ ions in ZSM-5: A combined quantum mechanics interatomic potential function study, Phys. Chem. Chem. Phys **1(8)**, 2019-2026 (1999).
18. S. Humbel, S. Seiber, and K. Morokuma, The IMOMO method: Integration of different levels of molecular orbital approximations for geometry optimisation of large systems: Test for n-butane conformation and $S_n'2$ reaction: $RCl+Cl^-$, J. Chem. Phys. **105(5)**, 1959-1967 (1996).
19. R.D.J. Froese, S. Humbel, M. Svensson, and K. Morokuma, IMOMO(G2MS): A new high-level G2-like method for large molecules and its application to

Diels-Alder Reactions, J. Phys. Chem. A **101**, 227-233 (1997).

20. A. Warshel and S. Creighton, MOLARIS, in *Computer Simulation of Biomolecular Systems* W.F. van Gunsteren, P.K. Weiner (Eds), ESCOM, Leiden, 1989, vol 2, pp. 120.
21. F. Bernardi, M. Olivucci, and M.A. Robb, Simulation of MC-SCF results on covalent organic multibond reactions - molecular mechanics with valence bond (MM-VB), J. Am. Chem. Soc. **114(5)**, 1606-1616 (1992).
22. M.J. Bearpark, P. Celani, F. Jolibois, M. Olivucci, M.A. Robb, and F. Bernardi, Characterization of the indacene S-0/S-1 conical intersection: an MMVB and CASSCF study, Molec. Phys. **96**, 645-652 (1999).
23. J.H. Harding, A.H. Harker, P.B. Keegstra, R. Pandey, J.M. Vail and C. Woodward, Hartree-Kock cluster computations of defect and perfect ionic-crystal properties, Physica B **131(1-3)**, 151-156 (1985).
24. Z. Barandiaran and L. Seijo, The *ab-initio* model potential representation of the crystalline environment. Theoretical study of the local distortion on $NaCl^{-}Cu^{+}$, J. Chem. Phys. **89(9)**, 5739-5746 (1988).
25. D. Wei and D.R. Salahub, A combined density functional and molecular dynamics simulation of a quantum water molecule in aqueous solution, Chem. Phys. Lett. **224**, 291 (1994).
26. X.P. Long, J.B. Nicholas, M.F. Guest, and R.L. Ornstein, A combined density functional theory molecular mechanics formalism and its application to small water clusters, J. Mol. Struct. **412(1-2)**, 121-133 (1997).
27. T.K. Woo, P.M. Margl, P.E. Blöchl, and T. Zeigler, A Combined Car-Parrinello QM/MM implementation for *ab-initio* molecular dynamics simulations of extended systems: Application to transition metal catalysis, J. Phys. Chem. **101(40)**, 7877-7880 (1997).
28. U. Rothlisberger, Hybrid Car-Parrinello simulations of enzymatic reactions. Presentation[29], for abstract see J. Mol. Graphics. Mod. **16(4-6)**, 275-276 (1998).
29. Molecular Graphics and Modelling Society meeting: *MM/QM methods and applications*, Southampton, 14-16 April, 1999.
30. M. Eichinger, P. Tavan, J. Hutter, and M. Parrinello, A hybrid method for solutes in complex solvents: Density functional theory combined with empirical force fields, J. Chem. Phys **110(21)**, 10452-10467 (1999).
31. B.R. Brooks, R.E. Bruccoleri, B.D. Olafson, D.J. States, S. Swaminathan, and M. Karplus CHARMM - A program for macromolecular energy, minimization, and dynamics calculations, J. Comp. Chem. **4(2)**, 187-217 (1983).
32. W.D. Cornell, P. Ciepak, C.I. Bayly, I.R. Gould, K. Merz, D.M. Ferguson, D.C. Spellmeyer, T. Fox, J.W. Caldwell, and P. Kollman, A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic-molecules, J. Am. Chem. Soc. **117**, 5179 (1995).
33. The MM3 force-field for amides, polypeptides and proteins, J.H. Lii and N.L. Allinger, J. Comp. Chem. **12(2)**, 186-199 (1991).
34. J.R. Maple, M.-J. Hwang, T.P. Stockfish, U. Dinur, M. Waldman, C.S. Ewig, and A.T. Hagler, Derivation of class-II force-fields .1. Methodology and quantum force-field for the alkyl functional-group and alkane molecules, J. Comp.

Chem. **15(2)**, 162-182 (1994).

35. B.G. Dick and A.W. Overhauser, Chem. Rev. **112**, 90 (1958).
36. R.A. Jackson and C.R.A. Catlow, Molecular Simulation **1**, 207 (1988).
37. J.R. Hill and J. Sauer, Molecular mechanics potential for silica and zeolite catalysts based on *ab initio* calculations 2. Aluminosilicates, J. Phys. Chem. **99**, 9536-9550 (1995).
38. K.P. Schroder and J. Sauer, Potential functions for silica and zeolite catalysts based on ab initio calculations .3. A shell model ion pair potential for silica and aluminosilicates, J. Phys. Chem. **100(26)**, 11043-11049 (1996).
39. J. Gao and X. Xia, A Priori evaluation of aqueous polarization effects through Monte Carlo simulations, Science **258**, 631 (1992).
40. P.L. Cummins and J.E. Gready, Coupled semiempirical molecular orbital and molecular mechanics model (QM/MM) for organic molecules in aqueous solution, J. Comp. Chem. **18(12)**, 1496-1512 (1997).
41. Y.Q. Tu and A. Laaksonen, On the effect of Lennard-Jones parameters on the quantum mechanical and molecular mechanical coupling in a hybrid molecular dynamics simulation of liquid water, J. Chem. Phys. **111(16)**, 7519-7525 (1999).
42. G. Ujaque, F. Maseras, and O. Eisenstein, Different van der Waals radii for organic and inorganic halogen atoms: a significant improvement in IMOMM performance, Theor. Chem. Acc. **96**, 146-150 (1997).
43. D.M. Philipp and R.A. Friesner, Mixed *ab initio* QM/MM modeling using frozen orbitals and tests with alanine dipeptide and tetrapeptide, J. Comp. Chem **20(14)**, 1468-1494 (1999).
44. P. Sherwood, A.H. de Vries, S.J. Collins, S.P. Greatbanks, N.A. Burton, Vincent, and I.H. Hillier, Computer simulation of zeolite structure and reactivity using embedded cluster methods, Faraday. Discuss. **106**, 79-92 (1997).
45. D. Bakowies and W. Thiel, Semiempirical treatment of electrostatic potentials and partial charges in combined quantum mechanical and molecular mechanical approaches, J. Comp. Chem. **17(1)**, 87-108 (1996).
46. R.A. Bryce, M.A. Vincent, and I.H. Hillier, Binding energy of $F(H_2O)(^-)$ and the simulation of fluoride water clusters using a hybrid QM/MM (fluctuating charge) potential, J. Phys. Chem. A **103(20)**, 4094-4100 (1999).
47. P.L. Cummins and J.E. Gready, The electrostatic potential in the semiempirical molecular orbital approximation, Chem. Phys. Lett. **225**, 11-17 (1994).
48. B.T. Thole, Chem. Phys. **59**, 341 (1981).
49. I. Antes and W. Thiel, In *Combined Quantum Mechanical and Molecular Mechanical Methods*; J. Gao and M.A. Thompson (eds), ACS Symp. Ser., vol. **712**, American Chemical Society, Washington, DC 1998, p. 50.
50. M.A. Thompson, QM/MMpol: A consistent model for solute/solvent polarization. Application to the aqueous solvation and spectroscopy of formaldehyde, acetaldehyde, and acetone, J. Phys. Chem. **100(34)**, 14492-14507 (1996).
51. B.T. Thole and P.Th. van Duijnen, On the quantum mechanical treatment of solvent, Theor. Chim. Acta. **55**, 307 (1980).
52. A.H. de Vries, P.Th. van Duijnen, A.H. Juffer, J.A.C. Rullmann, J.P. Dijkman, H. Merenga, and B.T. Thole, Implementation of reaction field methods in

quantum chemistry computer codes, J. Comp. Chem. **16(1)**, 37-55 (1995), correction in J. Comp. Chem. **16(11)**, 1445-1446 (1995).

53. P.Th. van Duijnen and A.H. de Vries, Direct reaction field force field: A consistent way to connect and combine quantum-chemical and classical descriptions of molecules, Int. J. Quant. Chem. **60(6)**, 1111-1132 (1996).
54. A.H. de Vries, P.Th. van Duijnen, Solvatochromism of the $\pi^* \leftarrow n$ transition of acetone by combined quantum mechanical classical mechanical calculations, Int. J. Quant. Chem. **57(6)**, 1067-1076 (1996).
55. P.T. van Duijnen, F. Grozema, and M. Swart, Some applications of the direct reaction field approach, J. Mol. Struct. (THEOCHEM) **464(1-3)**, 191-198 (1999).
56. F.C. Grozema, R.W.J. Zijlstra, and P.T. van Duijnen, Many-body interactions calculated with the direct reaction field model, Chem. Phys. **246(1-3)**, 217-227 (1999).
57. I. Antes and W. Thiel, Adjusted connection atoms for combined quantum mechanical and molecular mechanical methods, J. Phys. Chem. A **103(46)**, 9290-9295 (1999).
58. Y.K. Zhang, T.S. Lee, W.T. Yang, A pseudobond approach to combining quantum mechanical and molecular mechanical methods, J. Chem. Phys. **110(1)**, 46-54 (1999).
59. N.W. Winter and R.M. Pitzer, Configuration-Interaction calculation of the electronic-spectra of MgF_2-V^{+2}, J. Chem. Phys. **89**, 446 (1988).
60. G. Pachioni, A.M. Ferrari, and G. Ierano, Cluster model calculations of oxygen vacancies in SiO_2 and MgO, formation energies, optical transitions and EPR spectra, Faraday. Discuss. **106**, 155-172 (1997).
61. M.A. Nygren, L.G.M. Pettersson, Z. Barandiaran, and L. Seijo, Bonding between CO and the MgO(001) surface - a modified picture, J. Chem. Phys. **100(3)**, 2010-2018 (1994).
62. V. Thery, D. Rinaldi, J.-L. Rivail, B. Maigret, and G.G. Ferenczy, Quantum mechanical computations on very large molecular systems: The local self-consistent field method, J. Comp. Chem. **15(3)**, 269-282 (1994).
63. G. Monard, M. Loos, V. Thery, K. Baka, and J.L. Rivail, Hybrid classical quantum force field for modeling very large molecules, Int. J. Quant. Chem. **58(2)**, 153-159 (1996).
64. J. Gao, P. Amara, C. Alhambra, and M.J. Field, A Generalised Hybrid Orbital (GHO) Method for the treatment of boundary atoms in combined QM/MM calculations, J. Phys. Chem. A **102**, 4714-4721 (1998).
65. X. Assfeld and J.L. Rivail, Quantum chemical computations on parts of large molecules: The *ab initio* local self consistent field method, Chem. Phys. Lett. **263(1-2)**, 100-106 (1996).
66. B.R. Brooks, personal communication.
67. P. Sherwood, E. Billings, and B.R. Brooks, unpublished.
68. F. Maseras and K. Morokuma, IMOMM: A new integrated *ab-initio* + molecular mechanics geometry optimisation scheme of equilibrium structures and transition states, J. Comp. Chem. **16(9)**, 1170-1179 (1995).
69. A.H. de Vries, P. Sherwood, S.J. Collins, A.M. Rigby, M. Rigutto, and

G.J. Kramer, Zeolite structure and reactivity by combined quantum-chemical-classical calculations, J. Phys. Chem. B **103(29)**, 6133-6141 (1999).

70. T.K. Woo, L. Cavallo, T. Ziegler, Implementation of the IMOMM methodology for performing combined QM/MM molecular dynamics simulations and frequency calculations, Theor. Chem. Acc **100(5-6)**, 307-313 (1998).
71. I. Antes, Ph.D Thesis, University of Zurich (1998).
72. P.E. Smith and W.F. van Gunsteren in *Computer Simulation of Biomolecular Systems*, W.F. van Gunsteren, P.K. Weiner, and A.J. Wilkinson (Eds), ESCOM, Leiden, 1993, vol. **2**, pp. 182-212.
73. P.E. Sinclair, A. de Vries, P. Sherwood, C.R.A. Catlow, and R.A. van Santen, Quantum-chemical studies of alkene chemisorption in chabazite: A comparison of cluster and embedded-cluster models, J. Chem. Soc., Faraday Trans **94(22)**, 3401-3408 (1998).
74. J. Baker An Algorithm for the location of transition states, J. Comp. Chem. **7(4)**, 385-395 (1986), and references therein.
75. S.R. Billeter, A.J. Turner and W. Thiel, Linear Scaling Geometry Optimisation and Transition State Search in Hybrid Delocalized Internal Coordinates, *Phys. Chem. Chem. Phys.*, in press.
76. J. Baker, A. Kessi and B. Delley, The generation and use of delocalised internal coordinates in geometry optimisation, J. Chem. Phys **105(1)**, 192-206 (1996).
77. C.D. Berweger, W.F. van Gunsteren, F. Müller-Plathe, Finite element interpolation for combined classical/quantum mechanics dynamics simulations, J. Comp. Chem. **18(12)**, 1484-1495 (1997).
78. K. Breuer, J.H. Teles, D. Demuth, H. Hibst, A. Schafer, S. Brode, and H. Domgorgen, Zinc silicates: Very efficient heterogeneous catalysts for the addition of primary alcohols to alkynes and allenes, Angew. Chem. Int. Ed. Engl. **38(1))**, 1401-1405 (1999).
79. R. Ahlrichs, M. Bär, M. Häser, H. Horn and C. Cölmel, Chem. Phys. Lett **162**, 165 (1989).
80. MOPAC 5, J.J.P. Stewart.
81. CADPAC5.2, The Cambridge Analytic Derivatives Package, issue 5.2, Cambridge, 1992. A suite of quantum chemistry programs developed by R.D. Amos with contributions from I.L. Alberts, S.M. Colwell, N.C. Handy, D. Jayatilaka, P.J. Knowles, R. Kobayashi, N. Koga, K.E. Laidig, P.E. Maslen, C.W. Murray, J.E. Rice, J. Sanz, E.D. Simandiras, A.J. Stone, and M.-D. Su.
82. W. Thiel, Program MNDO94; University of Zürich, 1995.
83. ChemShell Users Manual, P. Sherwood and A.H. de Vries, CLRC Daresbury Laboratory 1997.
84. GAMESS-UK, see `http://www.dl.ac.uk/CFS`
85. J.D. Gale, J. Chem. Soc. Faraday Trans. **93**, 629 (1997), J.D. Gale, *Program GULP (General Utility Lattice Program)* Royal Institution/Imperial College, London, 1992.
86. W. Smith and T. Forester, DL_POLY 2.0: A general purpose parallel molecular dynamics simulation package, J. Mol. Graph. **14**, 136 (1996).
87. W. F. van Gunsteren and S. R. Billeter, A. A. Eising, P. H. Hünenberger, P. Krüger, A. E. Mark, W. R. P. Scott and I. G. Tironi, *Biomolecular Sim-*

ulation: The GROMOS96 Manual and User Guide, Biomos b.v., Zürich and Groningen, VdF Hochschulverlag, ETH Zürich, 1996, and W. R. P. Scott and P. H. Hünenberger, I. G. Tironi, A. E. Mark, S. R. Billeter, J. Fennen, A. E. Torda, T. Huber and P. Krüger and W. F. van Gunsteren, The GROMOS Biomolecular Simulation Program Package, J. Phys. Chem. A **103(19)**, 3596-3607 (1999).

88. Gaussian 94, See `http://www.gaussian.com`
89. J.K. Ousterhout, *Tcl and the Tk toolkit*, Addison Wesley, 1994.
90. QUASI - Quantum Simulation in Industry, EU project 25047. For a summary of the project please see `http://www.cse.clrc.ac.uk/Activity/QUASI`