# A Combined Quantum Mechanical and Molecular Mechanical Potential for Molecular Dynamics Simulations

Martin J. Field, Paul A. Bash, and Martin Karplus

*Department of Chemistry, 12 Oxford Street, Harvard University, Cambridge, Massachusetts 02138*

A combined quantum mechanical (QM) and molecular mechanical (MM) potential has been developed for the study of reactions in condensed phases. For the quantum mechanical calculations semiempirical methods of the MNDO and AM1 type are used, while the molecular mechanics part is treated with the CHARMM force field. Specific prescriptions are given for the interactions between the QM and MM portions of the system; cases in which the QM and MM methodology is applied to parts of the same molecule or to different molecules are considered. The details of the method and a range of test calculations, including comparisons with *ab initio* and experimental results, are given. It is found that in many cases satisfactory results are obtained. However, there are limitations to this type of approach, some of which arise from the AM1 or MNDO methods themselves and others from the present QM/MM implementation. This suggests that it is important to test the applicability of the method to each particular case prior to its use. Possible areas of improvement in the methodology are discussed.

## INTRODUCTION

Chemistry is concerned with the properties of molecules and their reactions. The most widely studied class of chemical processes occur in the condensed phase where interactions with the solvent play a significant role in influencing the behavior of the molecules and their reactions. Both solution reactions and those in enzymes are of great interest. The molecular dynamics simulation method is well suited for the investigation of such processes.[1] It assumes that the motion of the atoms is determined by an effective potential arising from interactions between the electrons and nuclei of the system and that Newton's equations (or their semiclassical or quantum analogues) can be used for the dynamics. Both time-dependent and time-independent properties of a system can be determined with such simulations. Examples include the calculation of reactive trajectories for determining a rate constant, the examination of a phase transition and the computation of the free energy difference between two species in solution. The method is most useful for studying microscopic processes that occur on the picosecond or nanosecond timescales. Activated processes, that have small macroscopic rate constants due to the presence of

barriers in the potential, can be treated by special molecular dynamic simulation methods.

Most potential energy functions used in molecular dynamics calculations are of the molecular mechanics type. This means that they do not allow bonds to break or form and so cannot be applied to reactive processes because they have the wrong analytic form. Moreover, they are usually parametrized with experimental data which provides little information about parts of the potential energy surface that are far away from stable or equilibrium configurations.[2]

Both these problems can be overcome if information from quantum mechanical calculations is introduced. They are able, in principle, to describe reactive processes, including the making and breaking of bonds, electron transfer processes, and excited state isomerization.[3] Quantum mechanical potential surface calculations can be introduced into molecular dynamics simulations in two ways. The first, and most common, is to parameterize a suitable analytic form for the potential surface of a system (e.g., one that includes the making and breaking of specific bonds) to augment the normal MM force field.[4] Alternatively the QM and MM methods could be combined so that the reacting portions of the system are treated quantum mechanically and the rest, the environment, are treated by standard molecular mechanics. We have adopted the second approach.

Potential energy surfaces have long been used to study the dynamics of simple reacting sys-

tems. The first full reaction dynamics calculation for a gas phase system was performed over 20 years ago; classical trajectories were calculated for the hydrogen atom/hydrogen molecule reaction with a semiempirical quantum mechanical potential energy surface.[5] Since then progress has been rapid and now surfaces are being calculated with highly sophisticated theoretical treatments for a wide range of simple reactions.[6] Since it is difficult to generate analytic potential energy surfaces systematically for reactions with more than a few degrees of freedom, it is necessary to restrict the conformational space that is investigated for large systems. Fukui[7] and Miller, Handy and Adams[8] have advocated an intrinsic reaction coordinate or reaction path approach in which the minimum energy path between an initial and a final state on the potential energy surface is determined and dynamical calculations are performed along the path while taking account approximately of some additional degrees of freedom. The method simplifies the description of the reaction process, but may by assumption eliminate some of the complexities of the actual dynamics. Applications of the method to gas phase reactions have been published.[9] Also of interest is the work of Houk and coworkers[6b] that makes use of a MM2 type force field which is modified on the basis of ab initio calculations to model transition states.

An extension to reactions in solution has been made by Chandrasekar et al.[4] and Jorgensen et al.[10] The first reaction to be studied in this way is the $S_N2$ reaction $Cl^- + CH_3Cl \rightarrow ClCH_3 + Cl^-$. They determine the gas phase reaction path using high level ab initio calculations and then calculate relative free energies along the same path for the reaction in solution. The interaction between the solvent and reactive solute atoms was determined beforehand by parameterizing suitable analytic expressions from the results of quantum mechanical calculations on systems comprising a small number of solvent molecules (one or two) and the solute species. Corresponding potentials have been employed by Hynes and Wilson[11] in classical trajectory studies of the dynamics of the chloride/methyl chloride exchange reaction in solution.

The reaction path method is useful but suffers from the limitation that assumptions have to be made as to the mechanism of the reaction process. An alternative approach is to make no assumptions and to evaluate the energy and gradients for a given nuclear configuration quantum mechanically as the need arises during a dynamics simulation. An early study of this type considered the insertion of $H_2$ into $CH_2$ to form $CH_4$ using classical dynamical trajectories; the energy and forces for the system were calculated

quantum mechanically with the CNDO method.[12] Extensions to larger systems were hampered by computational limitations and so hybrid schemes were developed.[13] An example is provided by conjugated molecules, for which the energies of the $\sigma$-bonded part of the system were determined from a molecular mechanics force field and a quantum mechanical calculation was performed to obtain the energies for the $\pi$-electron system. The ground and excited electronic state potential surfaces of conjugated molecules and transitions between them were studied for a number of gas phase and crystal systems.[14] Dynamical simulations of the photoisomerization of ethylene and retinal were performed with these or related potentials.[15]

A number of groups[16] have combined ab initio quantum mechanical programs with molecular mechanics programs so that calculations can be performed on groups of reacting atoms in the field of the surrounding molecular mechanics atoms which represent the environment. Although such implementations are straightforward and, in principle, of general applicability, no dynamical simulations have been done using such combined potentials. Only static potential surface explorations are published.[17] The major problem is that, even with supercomputers, the time required for ab initio calculations is so great that calculations of the forces for a dynamical simulation of even a relatively small number of atoms is prohibitive.

The aim of the present work is to implement a generally applicable method by which a molecular system can be studied with geometry optimization, molecular dynamical and Monte Carlo methods. To do this we synthesize a range of techniques for calculating potential energy surfaces. Molecular mechanics force fields are fast to calculate but require extensive parameterization. Quantum mechanical methods require more computer time but they are much more flexible and can be used to study reactions. Moreover, the ab initio and semiempirical formulations require less parametrization than molecular mechanics methods; e.g., the ab initio treatment has as "parameters" only the form and types of basis functions satisfactory for given calculations, while semiempirical treatments require, in addition, parametered expressions for certain of the molecular integrals. To obtain the calculational speed required for large dynamic simulations, we have combined the CHARMM program,[18] which uses an empirical force field, with the MOPAC program[19] which incorporates the AM1[20] and MNDO[21] semiempirical molecular orbital procedures. By such a synthesis it is possible to treat systems including 50 quantum mechanical and 2000 molecular mechanics atoms and do simula-

tions on the order of 20 ps or longer. A related approach is that of Warshel et al.,[22] who used an empirical valence-bond representation of reactive processes, combined with a molecular mechanics program. In the development of a combined quantum mechanical/molecular mechanical (QM/MM) potential surface and force program, the crucial aspect is how the interactions between the different parts of the system are determined. It is necessary to develop a method for the treatment of the QM atoms in the field of the MM atoms and vice versa. If they are in different molecules this is relatively straightforward. However, if a given molecular entity is treated partly by QM and partly by MM, the connections between the two have to be formulated rather carefully. Moreover, because the MM and QM procedures are approximate and the errors are difficult to estimate a priori, it is essential to calibrate the method by doing tests on systems for which there are experimental and/or *ab initio* data. Examples of calibration procedures for semiempirical and *ab initio* methodologies have been described.[23,24] However, even with all the tests included here, caution is still required in application of the methodology. Both semiempirical QM and empirical MM potential surfaces represent "interpolation" formulas. In each new application, it is necessary to check their validity.

The method is described in the next two sections. The general methodology and implementation is given in the second section and the method for treating molecules that are partly QM and partly MM is outlined in the third. The fourth and fifth present a series of tests to validate the method. The results are summarized and some conclusions about the usefulness of the method are given in the sixth section. An appendix lists the molecular mechanics parameters for all the molecules used in the calculations.

## METHODOLOGY

In this section we present the implementation of the method and the assumptions made in deriving it. The combined QM/MM potential is described in the first three subsections and its application to geometry optimizations and molecular dynamics simulations is considered briefly in the fourth.

### Partitioning of the System

We assume that the molecular system to be studied can be divided into three parts — a QM region, a MM region, and a boundary region — as illustrated in Figure 1. The QM and MM regions contain all the atoms that are explicitly treated

in the calculation while the boundary region is included so as to account for the surroundings that are neglected. As formulated either periodic boundaries or the stochastic boundary method can be used.

Atoms in the QM region are represented as nuclei and electrons. The potential surface for these atoms is determined within the Born–Oppenheimer approximation[25] and energy derivatives as a function of the nuclear positions are obtained. The QM region contains all those atoms that are involved in the reaction process or for which for other reasons a QM description of the system is necessary (e.g., if excited states are involved).

The MM region contains the remaining atoms of the system. They are represented as atoms, rather than electrons and nuclei and their interactions are determined from an empirical potential energy function. They constitute the immediate environment for the QM atoms and are included because their interactions and dynamics will have a significant influence on the QM atoms and so require a detailed treatment. These atoms cannot be involved in the making or breaking of bonds.

The boundary region is a standard feature of many QM and MM calculations. A number of approximations are possible and these are described below.

### The System Hamiltonian

To calculate the energies and forces on each of the atoms in the system an effective Hamiltonian, $\hat{H}_{\text{Eff}}$, is constructed and the time-independent Schrödinger equation is solved for the energy of
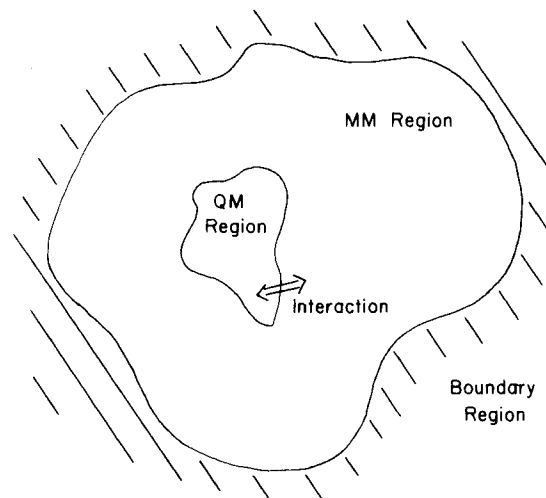


**Figure 1.** Partitioning of the system into quantum mechanical (QM), molecular mechanical (MM) and boundary regions.

the system, $E$, and the wavefunction of the electrons of the QM atoms, $\Psi$:

$$\hat{H}_{Eff}\Psi(\mathbf{r}, \mathbf{R}_\alpha, \mathbf{R}_M) = E(\mathbf{R}_\alpha, \mathbf{R}_M)\Psi(\mathbf{r}, \mathbf{R}_\alpha, \mathbf{R}_M)$$
(1)

The wavefunction is a function of the coordinates of the electrons, $\mathbf{r}$, and parametrically depends on the positions of both the quantum mechanical nuclei, $\mathbf{R}_\alpha$, and the molecular mechanical atoms $\mathbf{R}_M$. The energy is calculated for each geometry of the QM nuclei and the MM atoms and generates the Born–Oppenheimer potential surface for the system. It is given by the expectation value of the wavefunction over the Hamiltonian

$$E = \langle\Psi|\hat{H}_{Eff}|\Psi\rangle/\langle\Psi\,|\,\Psi\rangle$$
(2)

The forces on the QM nuclei, $\mathbf{F}_\alpha$, and the MM atoms, $\mathbf{F}_M$, are obtained by differentiating eq. (2) with respect to the nuclear or atomic Cartesian coordinates; that is,

$$F_\alpha = -\frac{\partial E}{\partial \mathbf{R}_\alpha} \quad \text{and} \quad F_M = -\frac{\partial E}{\partial \mathbf{R}_M}$$
(3)

The effective Hamiltonian for the partitioned system is written as the sum of four terms,

$$\hat{H}_{Eff} = \hat{H}_{QM} + \hat{H}_{MM} + \hat{H}_{QM/MM} + \hat{H}_{Boundary}$$
(4)

and the total energy is the sum of the expectation values of each term in the Hamiltonian; i.e.,

$$E = E_{QM} + E_{MM} + E_{QM/MM} + E_{Boundary}$$
(5)

Since the wavefunction depends only parametrically on the $\mathbf{R}_M$, the term $E_{MM}$ can be taken outside the integral in eq. (2) and we have

$$E = \langle\Phi|\hat{H}_{QM} + \hat{H}_{QM/MM} + \hat{H}_{Boundary(QM)}|\Psi\rangle/\langle\Psi\,|\,\Psi\rangle$$
$$+ E_{MM} + E_{Boundary(MM)}$$
(6)

where we have also separated the boundary term into QM and MM parts.

The first term in eq. (4) is the Hamiltonian describing the quantum mechanical particles (electrons and nuclei) of the system and their interactions with each other. It is a standard nonrelativistic electronic Hamiltonian and, in atomic units, is written as

$$\hat{H}_{QM} = -\frac{1}{2}\sum_i \nabla_i^2 + \sum_{ij}\frac{1}{r_{ij}} - \sum_{i\alpha}\frac{Z_\alpha}{r_{i\alpha}} + \sum_{\alpha\beta}\frac{Z_\alpha Z_\beta}{R_{\alpha\beta}}$$
(7)

where $i,j$ and $\alpha,\beta$ indicate electronic and nuclear coordinates respectively, $r$ is an electron/electron or electron/nuclear distance, $R$ is a nuclear/nuclear distance, and $Z_\alpha$ is a nuclear charge. Simplifications are introduced into the QM Hamiltonian for the semi-empirical treatments (see below). The MM part of the Hamiltonian, $\hat{H}_{MM}$, describes the MM region. The interaction Hamiltonian,

$\hat{H}_{QM/MM}$, describes the interaction between the QM and MM atoms and the interaction of the QM and MM regions with the environment is determined by the boundary Hamiltonian, $\hat{H}_{Boundary}$. Each of the terms in the Hamiltonian is considered in the following subsections.

## The QM Method

We use a Hartree–Fock (HF) linear combination of atomic orbitals (LCAO), molecular orbital (MO) method as the QM method[3]; electron correlation can be introduced via configuration interaction or many body perturbation theory. This class of methods is the most extensively tested and documented that is available. Other QM schemes such as density functional[26] or quantum Monte Carlo schemes[27] were considered; the latter is too restricted and time consuming in its present implementation, while the former is a possibility for the future. There exists also a related density functional scheme where both the quantum (electronic) and classical (nucleic) degrees of freedom are treated by classical mechanics.[28]

To obtain a HF or LCAO MO scheme that combines reasonable accuracy with the speed required for molecular dynamics, we focused on semiempirical approaches. Of the various implementation that exist,[29] the AM1/MNDO semiempirical HFSCF MO method of Dewar et al.[20,21] is best suited for the applications of primary interest to us; other semiempirical approaches (e.g., the extended PPP method[13] for studying the ground and excited states of conjugated molecules) could be substituted as appropriate. The AM1 and MNDO semiempirical procedures are parameterized HF–SCF MO methods that have been extensively tested by Dewar and his group on organic and inorganic systems. For many classes of systems[20,30] the calculated properties appear to be of an accuracy equivalent to *ab initio* calculations with the small split-valence basis sets 3-21G[31] and 4-31G[32] while being up to 1000 times faster. In some cases the results are better because of the empirical parameterizations. However, in other cases, including calculations on clusters of waters,[33] problems do arise. The AM1 and MNDO methods have been parameterized[20,21] for calculations on hydrogen, the first row elements plus silicon, phosphorus, sulphur and the halogens. The MNDO method has additional parameters for aluminium, phosphorus, chromium, germanium, tin, mercury and lead. Parameters for other elements are being developed.[34]

The method uses a valence-electron, minimum basis set approximation so that the QM system is comprised of fixed cores made up of the nuclei and the inner shell of electrons, plus the valence shell electrons which are treated quantum me-

chanically. The valence electrons occupy MOs made up of linear combinations of Slater-type atomic orbitals. Only a minimum basis set for each atom is used; i.e., a 1s orbital for hydrogen, 2s and 2p orbitals for the first row atoms, 3s and 3p orbitals for second-row atoms and so on.

In the NDDO (neglect of diatomic differential overlap) approximation that forms the basis for the AM1/MNDO procedure the assumption is made that all integrals involving diatomic differential overlap can be neglected and that the overlap integrals are omitted from the normalization of the MOs in the SCF procedure. Thus, all three and four center integrals are neglected. The nonzero integrals remaining are the one-center one-electron kinetic energy and core attraction integrals, $U_{\mu\mu}$, the one-center two-electron repulsion integrals, $(\mu\mu \mid \nu\nu)$ and $(\mu\nu \mid \mu\nu)$, two-center one-electron core attraction integrals, $V_{\mu\nu,B}$, two-center resonance integrals, $\beta_{\mu\lambda}$, one and two center two-electron repulsion integrals, $(\mu\nu \mid \lambda\sigma)$ and the overlap integrals, $S_{\mu\nu}$, which is used to evaluate $\beta_{\mu\lambda}$. Only the overlap integrals are evaluated analytically, the rest are approximated by empirical forms. Full details and formulae for the integrals can be found in the original papers of Dewar et al.[20,21,37]

The AM1/MNDO methods are implemented in the MOPAC program package.[19] Closed and open-shell restricted HF and unrestricted HF SCF and configuration interaction (CI) calculations can be performed. First derivatives are calculated analytically for closed-shell RHF and UHF wavefunctions but by finite differences for open-shell restricted HF and CI wavefunctions. CI or other correlated wavefunctions are particularly important for certain types of transition states and reaction paths[38] and for excited states. It should be noted that the AM1/MNDO parametrization attempts to reproduce the geometries and heats of formation for compounds in the electronic ground state at 25 °C. Thus, effects on the energy due to correlation and molecular vibrations are implicitly included in the procedure.

## The MM Hamiltonian

The molecular mechanics Hamiltonian is a function of the positions of the molecular mechanics atoms. It does not depend upon the coordinates of the electrons and nuclei of the QM atoms and so it is a number, i.e.,

$$\hat{H}_{MM} = E_{MM} \tag{8}$$

A standard MM force field is used to determine the MM energy; details are given in reference 18. The MM atoms are represented by partial charges and van der Waals spheres centered at the atom centers, through which the atoms interact at long

range. They are linked by harmonic bond and other internal coordinate terms through which they interact at short range and which determine the connectivity of the molecule. The force field and parameterization used in the present test calculations are based on the CHARMM force fields.[18,40,41]

## The QM/MM Hamiltonian

The interaction Hamiltonian $\hat{H}_{QM/MM}$ describes how the QM and MM atoms interact. The MM atoms in the normal force field are represented by point charges and van der Waals parameters and so the simplest interaction Hamiltonian would be written, in atomic units, as

$$\hat{H}_{QM/MM} = -\sum_{iM} \frac{q_M}{r_{iM}} + \sum_{\alpha M} \frac{Z_\alpha q_M}{R_{\alpha M}} + \sum_{\alpha M} \left\{ \frac{A_{\alpha M}}{R_{\alpha M}^{12}} - \frac{B_{\alpha M}}{R_{\alpha M}^{6}} \right\} \tag{9}$$

where the subscripts $i$ and $\alpha$ correspond to the QM electrons and nuclei respectively and M to the MM atoms. There are two electrostatic terms, in which the MM atoms interact with the QM electrons and nuclei respectively, and one van der Waals term. The latter is necessary because some MM atoms possess no charge and so would be invisible to the QM atoms. In addition the van der Waals terms on the MM atoms often provide the only difference in the interactions of one atom type versus another; e.g., chloride and bromide ions both have unit negative charges and only differ in their van der Waals terms.

The interaction Hamiltonian given in eq. (9) is the exact Hamiltonian for the interaction of QM nuclei and electrons with MM atoms represented by partial charges and van der Waals spheres. The last two terms in the equation are constant (they do not involve electronic coordinates) and can be calculated and added directly into the total energy. The first term involves electronic positions and so must be introduced into the HF–SCF procedure. To do this one-electron integrals over the operator $1/r_{iM}$ are calculated and then added into the one-electron matrix, $H$. The integrals have the form:

$$I_{\mu\nu} = q_M \int \eta_\mu(r) \frac{1}{r_{iM}} \eta_\nu(r) \, dr \tag{10}$$

where $\eta_\mu$ and $\eta_\nu$ are basis functions. The HF equations can be solved in the usual way except that the system now corresponds to one that depends on the external field due to the MM atom partial charges. This type of treatment would be appropriate for the *ab initio* case.

Since the AM1/MNDO method uses modified forms for the electrostatic interactions,[20,21,37] $\hat{H}_{QM/MM}$ has to be modified from eq. (9) to yield consistent results. In choosing the appropriate

form, it is required that the existing parameterization for the QM and MM terms in the Hamiltonian should not be affected. Both the AM1/MNDO methods and the MM interactions have been calibrated by extensive testing so that any change would require a reevaluation of the approximations parameters. Further, a minimum number of new QM/MM parameters should be introduced so that the parameterization is relatively simple and can be derived in a straightforward way from existing QM and MM parameters. In accord with the AM1/MNDO method, no parameters dependent upon two atom types should be required. Finally, the balance between attractive and repulsive forces must be preserved and the QM/MM interactions must be of the correct magnitude with respect to the separate QM and MM contributions.

Two expressions for $\hat{H}_{QM/MM}$ were examined. The first is given by eq. (9) with the one-electron integrals evaluated analytically using the expressions for one-electron integrals between Slater orbitals[42] and the second in eq. (9) replaced by

$$E_{QM/MM}^{core} = \sum_{A,M} \frac{Q_A q_M}{R_{AM}} \qquad (11)$$

where $A$ corresponds to the cores of the QM atoms and $Q_A$ is the core charge. Test calculations showed that analytic evaluation of the interactions was inappropriate because the interaction energy and forces produced were proportionately too large.

The second approach, which was adopted, used the same form of Hamiltonian as is employed in the semiempirical method for the corresponding interactions. The one-electron integrals between the basis functions $\mu$ and $\nu$ on atom Q and the MM atom partial charge, $q_M$ take the form

$$I_{\mu\nu} = -q_M(\mu^Q\nu^Q | s^M s^M) \qquad (12)$$

where $s^M$ is a notional $s$ orbital on a MM atom. The core-core terms for the MNDO Hamiltonian become

$$E_{QM,MM}^{core} = \sum Q_Q q_M(s^Q s^Q | s^M s^M)$$
$$+ |Q_Q q_M| (s^Q s^Q | s^M s^M)\{e^{-\alpha_Q R_{QM}} + e^{-\alpha_M R_{QM}}\} \qquad (13)$$

and there is an additional term in AM1 of the form

$$\sum_i \frac{Q_Q q_M}{R_{QM}} K_i^Q \exp\{-L_i^Q(R_{QM} - M_i^Q)^2\}$$
$$+ \sum_j \frac{Q_Q q_M}{R_{QM}} K_j^M \exp\{-L_j^M(R_{QM} - M_j^M)^2\} \qquad (14)$$

The terms in eqs. (12) to (14) contain parameters that depend on the QM atoms and five MM atom-dependent parameters; they are $\alpha_M, \rho_0^M, K_i^M, L_i^M$, and $M_i^M$. The QM parameters were left unchanged and only the MM parameters were

varied. Investigations were performed in which various parts of the expression for the core-core energies were left out and different scalings of both expressions were tried. Tests to determine the optimum values of the adjustable parameters are described later.

The nonbonded MM interactions (both electrostatic and van der Waals) are truncated to save time by including contributions only from atoms or groups that are within a certain cutoff of each other (typically 8 to 15 Å). The QM interactions in general are expected to extend over a small enough number of atoms that all interactions are included. For the QM/MM interactions a procedure similar to that employed for the MM atoms was employed. A QM/MM nonbonded interaction list was generated each time a normal MM nonbonded list update was performed; this is usually done every 10 to 20 steps. Updating is done on an atom by atom or a group by group basis as for the MM atoms. All of the van der Waals and electrostatic QM/MM interaction terms in eq. (9) or (11)–(14) were multiplied by a shifting or switching function in accord with what was used for the MM nonbonded potential, so that there are no discontinuities in the energies or the values of the integrals at the cutoff.[39] Formulas for the shifting and switching functions can be found in reference 18.

## Boundary Terms

Because any condensed phase or extended system must be simulated as a finite number of atoms, boundary terms are introduced to compensate for truncation of the system. They mimic the rest of the system that cannot be treated explicitly. Two methods are commonly used in standard MM calculations; they are the periodic boundary[43] and the stochastic[44] boundary approaches.

In the periodic boundary method the system is duplicated by translating it periodically in three dimensions so that the full system consists of an infinite lattice of exact copies of the central box or group of primary atoms. The energy and forces of interaction of the primary and image atoms are determined and summed into the total energy and forces for the system. Exactly the same procedure is followed in the calculation on a combined QM/MM system. The energy and forces on the MM atoms are evaluated in the usual way and the image atoms interact with the QM electrons and nuclei through the same Hamiltonian as do the MM atoms inside the box. The only complication is that the images contain copies of the QM atoms and that their charge distribution changes during a calculation.

There are two ways to implement the periodic system for the present case. The image QM

atoms can be treated as MM atoms in which case the charges on the image QM atoms must be determined by performing a population analysis each time a new wavefunction is calculated. Alternatively, the periodic cells are chosen such that the QM atoms and their images are far enough apart so that they do not interact; i.e., they are separated by more than the nonbond cutoff distance so that the interaction energy and forces between the QM and MM atoms can be determined exactly by use of the translational symmetry of the system. In most problems, the latter approach, which is to be preferred, can be used.

The other approach is to use a boundary force which acts on all atoms within the system. The system is represented by a sphere or other shape and the boundary potential is chosen to mimic the potential that a solvent or lattice would produce. A number of different approximations of varying complexity have been used.[44,45] No difficulties arise in implementing this approach with the QM/MM potential because the energy and interaction forces on the QM atoms due to the boundary are calculated in exactly the same way as for the MM atoms.

## Implementation of the Method

We have implemented the QM/MM potential as part of the CHARMM program,[18] which is a large multipurpose molecular mechanics and dynamics package. It is modular in form and so new features can be added relatively easily. There are a wide range of options and extensive analysis facilities making it especially attractive as the framework within which to implement the QM/MM potential. The approach described here could be introduced into other molecular mechanics programs by a similar procedure.

The QM part of the potential was added by incorporating parts of the MOPAC program[19] into the energy and analysis routines. The QM, MM, QM/MM and boundary energies and forces are calculated straightforwardly using the procedures described in the earlier parts of this section.

For a small number of particles numerical derivatives of the energy with respect to the nuclear coordinates were comparable in cost to analytic gradient evaluation. However, for the QM/MM interactions, where large numbers of derivatives have to be calculated, numerical derivative evaluation is impractical and an analytic method is required. This is because the forces on both the QM and MM atoms are required and so both QM and MM atoms have to be displaced in turn if finite difference methods are being used. The latest version of MOPAC (v 4.0) has analytic first derivatives that are incorporated into the program.

At present the program is implemented so that each time an energy or force evaluation is required an HF-SCF calculation is performed. The electronic energy is then added to the MM energy to give the total energy for the system. The QM, QM/MM, and MM gradients are added to give the forces on each particle. The density matrix is saved between energy calculations and used as a starting point for the next SCF calculation. Once the SCF calculation had converged for the first step of a minimization or dynamics run, convergence in subsequent steps, for which the geometry of the system is generally very similar, took only 2–3 SCF interactions. In cases with a relatively large number of QM atoms, it may be advantageous to use hybrid schemes. For example, the QM energy and forces could be calculated every few steps and for intermediate points an estimate of the energy could be made using an extrapolation method or perturbation theory. In our applications the semiempirical QM method is sufficiently fast because many fewer two-electron integrals have to be evaluated than for the *ab initio* case and they are much faster to evaluate because they are simple analytic functions. The SCF equations are quickly solved once the SCF calculation for the first step has converged. For *ab initio* QM or density functional methods, some method to minimize the number of quantum mechanical calculations may be needed.

Because of the use of a cutoff list and because the number of QM atoms present in any particular calculation is limited to 50 or so, the total number of QM/MM interactions is relatively small. The calculation of the one-electron integrals adds very little to the cost of the HF calculation over that for the isolated QM system. Because there are relatively few QM/MM interactions, the gradient evaluation is also not significantly slower than in the pure QM case.

## Application of the QM/MM Potential

The combined QM/MM potential can be used in the same way as a normal MM empirical energy function. The energy and forces may be employed to perform geometry optimizations, MD and Monte Carlo simulations and thermodynamic perturbation theory calculations on combined QM/MM systems. The difference is that reactions (or excited states) can be described for the QM region and that information can be obtained about the electron distribution surrounding the QM nuclei. Thus, charge analyses can be performed during an MD simulation to gain an idea of how the charge cloud distorts with time or along a reaction path under the influence of the environment.[46]

There are no simple rules on how to partition the system into QM and MM regions. Clearly the part of the system directly involved in the reaction will have to be treated quantum mechanically. Beyond that, the partitioning is somewhat arbitrary and will depend upon the nature of the system and the computational resources available. If possible, different partitionings should be tried so that the dependence of the results of minimization or dynamics runs on partitioning can be ascertained.

With current computational resources it is feasible to perform MD simulations in which up to 50 atoms are treated with the AM1 or MNDO semiempirical approximations and several thousand with MM. A QM/MM simulation of this size takes only two to three times longer than a corresponding MM simulation.

## LINK ATOMS

In some systems it is desirable to partition different parts of the same molecule into the QM and MM regions. For example, in an enzyme reaction some residues have important catalytic roles and must be treated quantum mechanically while the rest of the protein does not; e.g., the transfer of a proton between the substrate and a glutamate residue in the reaction catalyzed by the enzyme, triosephosphate isomerase.[47] In these cases there are bonds between the QM and MM atoms. Because the electrons of MM atoms are not treated explicitly, it is necessary to devise a scheme in which the QM electron density along the QM/MM bonds can be terminated satisfactorily. The effect of the interactions of the bonded MM atoms with the QM atoms and their electron density should resemble, as closely as possible, the results of a QM calculation on the complete system. As long as the QM and MM atoms are in different molecules no such problems arise.

There is no exact way of introducing such a truncation, although many schemes of varying sophistication have been proposed.[48,49] We use a simple approximate method in which "dummy" or "link" QM atoms are introduced along the broken QM/MM bond. The link atoms are treated exactly like QM hydrogen atoms in the AM1/MNDO scheme and they are invisible to the MM atoms because no interactions between the link atoms and the MM atoms are calculated. This method has been used before. For example, Christoffersen et al.[48] used it to satisfy the extra valencies of groups in a fragment molecular orbital approach, Allen et al.[49] used it in QM studies of inorganic complexes and enzymes and Singh and Kollman[16] used it in their implementation of a combined *ab initio* QM/MM potential.

The procedure employed in the present version of the program is:

1. Partition a molecule into the QM and MM regions.
2. Define the link atoms for each molecule containing both QM and MM atoms. One must be defined for each bond that crosses the boundary between the two regions. In the CHARMM program[18] they have atom type QQ and are defined as part of the appropriate residue in the topology file. They have no charge, no van der Waals parameters and participate in no MM internal coordinate energy terms. The link atom coordinates are determined automatically to place them at 1.0 Å from the QM atoms along the bonds between the QM and MM atoms. The exact distance does not matter since the position of the link atoms is optimized in subsequent calculations.
3. Delete connectivity amongst the QM atoms. All MM internal coordinate energy terms that involve only QM atoms are deleted. All other terms are retained, even if only one MM atom is involved. For example, in Figure 2 the $\beta$ and $\gamma$ carbons, plus the $\delta$ oxygens of the glutamate residue are treated quantum mechanically, and the other atoms are treated molecular mechanically. Terms involving only MM atoms would be included as would be the following terms involving one or more QM
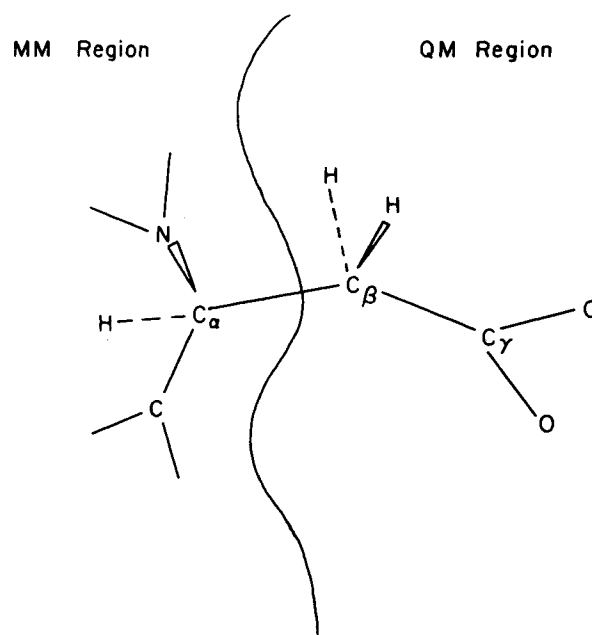


**Figure 2.** An example of how to partition parts of the same molecule into QM and MM regions. The figure shows a glutamate sidechain in a protein. The terminal $CH_2CO_2^-$ group is treated quantum mechanically and the backbone atoms molecular mechanically.

atoms: the $C_\alpha$-$C_\beta$ bond term; the angle terms $N$-$C_\alpha$-$C_\beta$, $C$-$C_\alpha$-$C_\beta$, $H_\alpha$-$C_\alpha$-$C_\beta$, $C_\alpha$-$C_\beta$-$H_{\beta1}$ and $C_\alpha$-$C_\beta$-$H_{\beta2}$; and the proper dihedral terms of which there are nine for the $C_\alpha$-$C_\beta$ torsion, two for the $C_\alpha$-$C_\beta$-$N$-$X$ torsion, two for the $C_\beta$-$C_\alpha$-$C$-$X$ torsion and two for the $C_\alpha$-$C_\beta$-$C_\gamma$-$X$ torsion, where $X$ represents an atom bonded to the preceding atom in the dihegral angle term.

4. Create the nonbond lists. The MM nonbond list is generated as usual and contains a list of pairs involving only MM atoms for which electrostatic and van der Waals interactions must be computed. The list is generated on an atom-by-atom or a group-by-group basis and its length is determined by the size of the nonbonded cutoff which gives the range over which nonbonded interactions are to be computed.

There are two QM/MM lists. One is the van der Waals interaction pair list and it contains the QM/MM atom pairs for which van der Waals interactions must be computed. It is generated in exactly the same way and with the same truncation criteria as used for the MM lists. The second QM/MM list is the electrostatic list and it is created in the same way as the other lists if no link atoms are present. If link atoms occur then the list is generated so as to include *all* interactions — even those that are normally excluded in the other lists; i.e., electrostatic terms are included between MM and QM atoms that are separated by only one, two, or three bonds.

It is to be noted that the link atoms have no interactions with the MM atoms and appear only in the QM portion of the calculations. They have no entries in any of the nonbond lists.

5. Calculate the MM energies and forces. The MM/QM internal coordinate terms are included in the MM internal coordinate energy term lists.

6. Calculate the QM/MM van der Waals interactions using the QM/MM van der Waals lists.

7. Calculate the QM and QM/MM electrostatic interaction energy and forces. The link atoms feel forces from the other QM atoms and exert forces on them. This is done because exclusion of the "link atom energy" from the QM energy is not possible and exclusion of the forces, which is feasible, would result in inconsistencies, i.e., the energies and forces would not correspond.

The approach outlined here permits the partitioning of molecules into separate QM and MM regions. The QM calculation determines the energy and electron distribution of the QM atoms subject to the perturbation caused by the electrostatic potential due to the neighboring MM group while the MM calculation preserves the integrity of the rest of the molecule and ensures the QM atoms are in the correct orientation with respect to it. It is to be noted that the inclusion of the "link atom energy" means that it is only possible to compare the energies of systems with the same number and type of link atoms. Comparisons of the absolute energies for the same system calculated with different QM/MM approximations are meaningless. The fifth section presents tests of the link truncation scheme described here.

## RESULTS AND DISCUSSION

In this section the results of tests of the QM/MM calculations are presented. The tests have been designed to determine the effectiveness of the QM/MM Hamiltonian. Of particular importance for applications are the interaction energies between groups and the consequences of dividing a molecule into a mixed QM/MM system. For ease of comparison with experimental and *ab initio* results, all tests were performed in the gas phase. Calculations on condensed phase systems have been published.[46] Since the tests are on small systems, an infinite nonbonded cutoff was used so that all interactions are included and no shifting or switching functions are applied. A dielectric constant of unity was used for the electrostatic interactions; this is appropriate for both isolated and solvated systems. The molecular mechanics parameters used for systems that were examined are given in appendix A.

The purpose of the method that we have developed is to mimic the interactions between real atoms and molecules with reasonable accuracy and speed. Obviously not all forms of interactions will be capable of being described by the present approach. Since the MM atoms are modeled as point charges and van der Waals spheres bonding between QM and MM atoms cannot occur (unless it is already part of the initial description). Thus, all atoms for which making and breaking of bonds is important must be included in the QM set. Also, the minimal basis description of the QM atoms suggests that strong charge cloud polarization effects will be poorly reproduced. However, for many problems involving molecules separated by distances typical of those in the condensed phase or the interior of biomolecules the necessary interactions are likely to be accurately modeled by this approach.

In presenting the results, the notation QM(AM1 or MNDO/QM atoms)/MM(MM atoms) is used. The parentheses following QM indicate

the semiempirical Hamiltonian used (AM1 or MNDO) and the atoms that are treated quantum mechanically; the parentheses after MM contains a list of the molecular mechanics atoms.

## Charge/Ion and Charge/Molecule Calculations

As described earlier the synthesis of the QM and MM potentials is based on independently developed QM and MM potentials and introducing as few QM/MM dependent parameters as possible. The QM and MM atoms interact only via electrostatic and van der Waals terms. The van der Waals parameters for both the QM and MM atoms are obtained from the MM force field and so no extra QM/MM parameters have to be determined. If the assumption is made that the electrostatic part of the QM/MM Hamiltonian is independent of the MM atom type, it is possible to treat separately the two types of interactions and calibrate the best form for the electrostatic Hamiltonian from comparison with accurate ab initio calculations on point charge/ion and molecule systems. For example, the difference in the way a QM water molecule interacts with a MM chloride, bromide or iodide anion is determined solely by the van der Waals interactions. The electrostatic interactions are the same and can be investigated equally well by calculations on a water molecule and a unit negative charge. These calculations have the advantage that no van der Waals interactions are involved and, moreover, that no compensation for bonding effects has to be made.

The first set of calculations was performed on fluoride anion/point charge and lithium cation/point charge systems. Point charges with values of 1.0, 0.5, −0.5, and −1.0 were used. Potential curves for the interactions were calculated at the ab initio level (MP2/6−311++G**),[24] with the combined QM/MM methods and with a point charge approximation.

The form of the interaction Hamiltonian is described earlier and is given by eqs. (12) to (14). There are two adjustable parameters, $\alpha^M$ and $\rho_o^M$, for the MNDO method and three additional parameters, $K_i^M$, $L_i^M$, and $M_i^M$, for the AM1 method. A search was performed to find the optimum values of the parameters by calculating interaction curves for charge/ion systems and comparing them with the ab initio results. It was found that omission of the extra AM1 core term had an insignificant effect on the interaction energies and so the $K$, $L$, and $M$ parameters were set to zero for the MM atoms. (For convenience, the $M$ superscript is dropped.) The best fit to the ab initio curves was found with the $\alpha$ and $\rho_0$ parameters equal to 5.0 and 0.0, respectively. Representative examples are given in Figure 3 for the fluoride anion/+1.0 system. Ab initio and three QM/MM curves with differing $\{\alpha, \rho_0\}$ values are plotted. Comparable results were found for the other fluoride anion/charge calculations and for the lithium cation case.

In the MOPAC program,[19] "sparkle" elements are defined which are designed to represent ionic charges of values 1.0, 2.0, −1.0, and −2.0. These are the nearest equivalent to MM atoms and they use a $\rho_0$ value of 1.0 and an $\alpha$ value of 1.5.
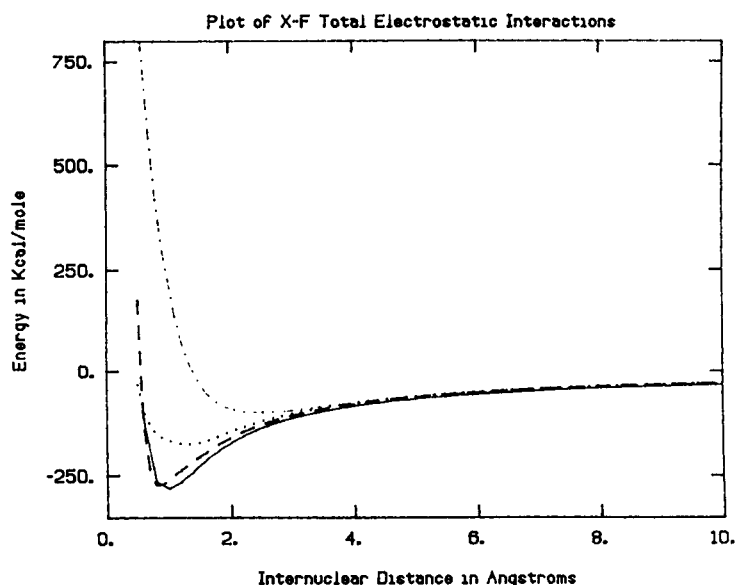


Figure 3. Potential energy interaction curves for the fluoride anion/+1.0 charge system. The ab initio curve (solid line) is compared with the curves produced by using the semi-empirical AM1 Hamiltonian and three different sets of $\{\alpha, \rho_0\}$ parameters, $\{\infty, 1.0\}$ (dotted line), $\{5.0, 0.0\}$ (dashed line) and $\{1.5, 1.0\}$ (dot-dash line).

As can be seen from Figure 3, these parameters do not give a good fit to the *ab initio* results.

To show one of the limitations of the electrostatic QM/MM Hamiltonian, we performed calculations for positive and negative unit charges interacting with a water molecule. The calculations were carried out with an *ab initio* approximation (MP2/6−311++G**) and with the QM/MM potential. The charge approached the hydrogen atom along one of the H-O bonds. The results are shown in Figures 4 and 5. It is clear from the curves that for the negative charge, in particular, the semiempirical QM description is not flexible enough to permit the large charge cloud distortion effects (essentially a charge transfer) that occur for the *ab initio* calculations. This is true because a minimal basis for the valence electrons only is used and no orbital was included in the charge. In most systems of interest, such extreme polarization effects do not occur. The interaction between atoms at short range are dominated by repulsive effects. In the present model they are determined by a combination of the $R^{-12}$ term in the Lennard–Jones

nonbonding potential and the repulsive term present in the electrostatic QM/MM interaction (eq. 13).

## Ion/Molecule and Molecule/Molecule Complexes

From the above, it appears that an electrostatic interaction of the form $\{\alpha, \rho_0\}$ should be adequate for most cases of interest. In this section, we performed a series of test calculations on ion/molecule and molecule/molecule complexes. We compare the QM/MM results using values $\{\alpha = \infty, \rho_0 = 1.0\}$ and $\{\alpha = 5.0, \rho_0 = 0.0\}$ with *ab initio* and MM calculations.

### Fluoride Anion/Water Interaction Curves

The fluoride ion was chosen for study because of the possibility of doing high level *ab initio* calculations. Interaction energy curves were calculated for the approach of the fluoride anion towards a water molecule along one of its H-O bonds. The water geometry was fixed at the opti-
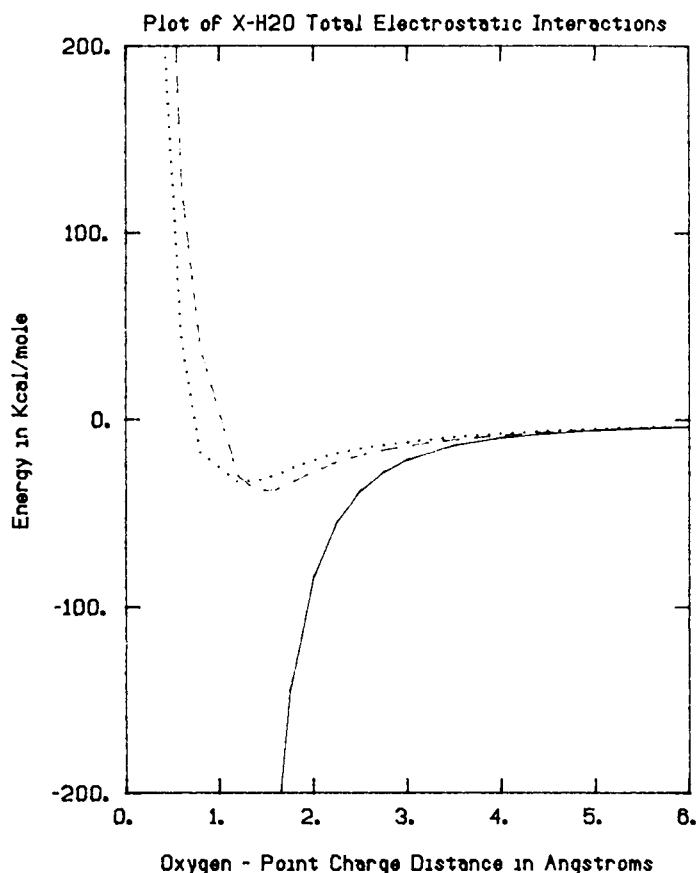


Figure 4. Interaction curves for the approach of a unit negative charge to water along one of its H—O bonds (Fig. 4). The *ab initio* curve (solid line), is compared to two semi-empirical curves using the AM1 Hamiltonian and the parameter sets $\{\alpha, \rho_0\} = \{\infty, 1.0\}$ (dotted line) and $\{5.0, 0.0\}$ (dashed line).
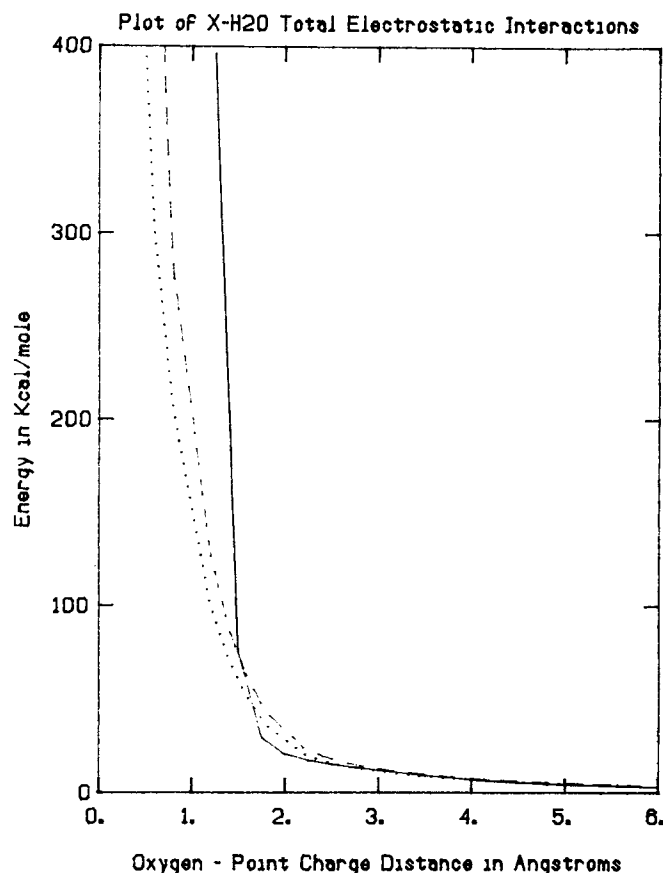
**Figure 5.** Interaction curves for the approach of a unit positive charge to water along one of its H—O bonds (Fig. 4). The ab initio curve (solid line), is compared to two semi-empirical curves using the AM1 Hamiltonian and the parameter sets $\{\alpha, \rho_0\} = \{\infty, 1.0\}$ (dotted line) and $\{5.0, 0.0\}$ (dashed line).

mized structure corresponding to the model being used. Curves were calculated at the *ab initio* (MP2/6−311++G**) level,[24] at the same level with corrections made for basis set superposition error (BSSE) by the counterpoise method,[50] with the AM1 semiempirical method and at the MM level. The results are presented in Figure 6. It can be seen that the BSSE-corrected *ab initio* curve is less attractive and rises more quickly than the full *ab initio* curve and that the AM1 curves follow the BSSE-corrected curves. The MM results have a well that is somewhat less deep.

Figure 7 shows the results for the system obtained with the combined QM/MM potential. Four curves are shown which represent calculations using the AM1 approximation with two sets of $\{\alpha, \rho_0\}$ parameters. In two of the cases the fluoride anion was treated as a QM particle and the water molecule by MM while in the other two the water was treated by QM and the fluoride anion by MM. QM/MM van der Waals interactions were calculated for all the curves.

The curves with $\{\alpha\,\rho_0\}$ parameters $\{\infty, 1.0\}$ are less attractive than the corresponding curves with parameters $\{5.0, 0.0\}$ and are in worse agreement with the *ab initio* results. They are equally repulsive at short range because of the van der Waals $R^{-12}$ interaction. The best QM/MM results are compared with the *ab initio* and AM1 semiempirical curves in Figure 8. The QM/MM curves are too shallow by between 5 and 10 kcal/mol⁻¹.

*Chloride Anion/Water Complexes*

The chloride anion/water complexes are known from *ab initio* calculation at the HF/6-31G* level to have a minimum energy form of $C_s$ symmetry (Fig. 9).[51] A $C_{2v}$ symmetry form, in which the Cl⁻ lies on the bisector of the HOH angle, is a saddle point on the reaction path between the two equivalent $C_s$ minima. Experimental[52] and representative *ab initio* values for the binding energy and some geometrical parameters are given in Table I.
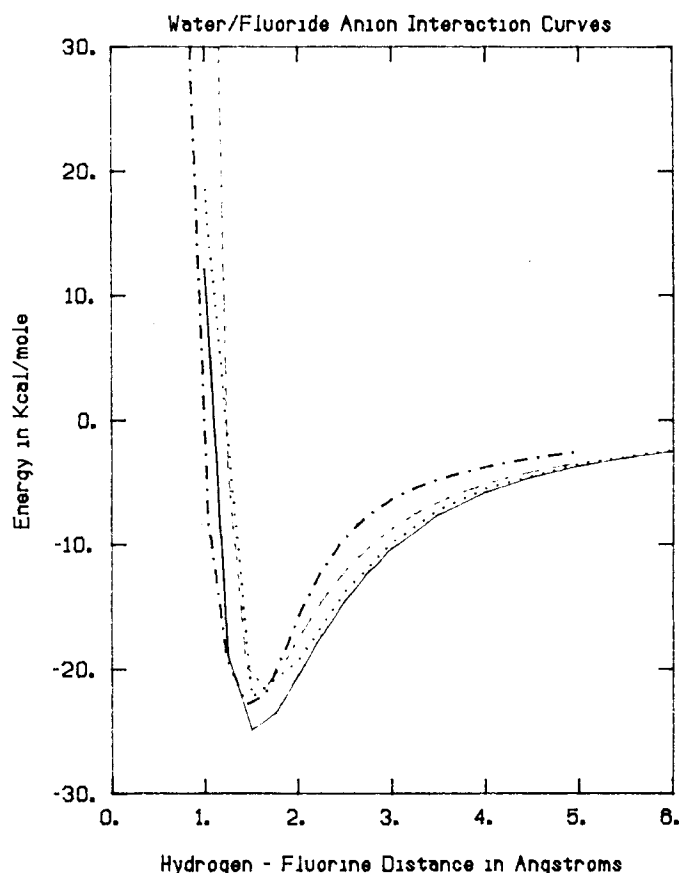
**Figure 6.** Interaction curves for the fluoride anion/water system. Curves are shown that have been calculated using an ab initio approximation (solid line), the same ab initio curve corrected for basis set superposition error (dotted line), a MM model (dashed line) and the AM1 semiempirical method (dot-dash line).

Calculations were performed for the complex with the AM1, MNDO and MM methods (see Table I). The AM1 method gives $C_s$ and $C_{2v}$ complexes with energies in good agreement with experiment and ab initio theory. The $C_s$ symmetry form is a minimum and there is a 1.4 kcal/mol$^{-1}$ energy difference between it and the symmetric $C_2$ form. However, the Cl-O distances are significantly too short. The MNDO method gives a $C_{2v}$ symmetry form as the minumum with a binding energy of 8.0 kcal/mol$^{-1}$. MM calculations give reasonble energies and geometries except for the fact that $C_s/C_{2v}$ complex energy difference is too small.

Studies were also carried out with the QM/MM potential, for the QM(Cl$^-$)/MM(H$_2$O) and QM(H$_2$O)/MM(Cl$^-$) systems. Results were obtained with the AM1 Hamiltonian, with and without van der Waals interactions and with the two sets of $\{\alpha, \rho_0\}$ parameters. All lowest energy complexes, except one, are of $C_{2v}$ symmetry. The energy of $C_s$ structures similar to those discussed above are between 0.6 and 17.8 kcal higher. The $_s$ calculations without van der Waals interactions

all gave binding energies that are much too large except for the case with the chloride anion treated by QM with $\{\alpha, \rho_0\} = \{\infty, 1.0\}$; however, the Cl-O distance is much too short. Of the calculations with van der Waals interactions, the results with $\{\alpha, \rho_0\} = \{5.0, 0.0\}$ are the most satisfactory. The binding energy of the QM(Cl$^-$)/ MM(H$_2$O) complex is in very good agreement with experiment although the summetry of the complex is $C_{2v}$ while the geometry of the QM(H$_2$O)/MM(Cl$^-$) complex has the correct symmetry but the binding energy is about 3.0 kcal mol$^{-1}$ too small; for the latter, the $C_{2v}$ structure is less stable by only a fraction of a kcal/mol$^{-1}$. The $C_{2v}/C_s$ structure energy difference for the QM(Cl$^-$)/MM(H$_2$O) complex is also very small; it is 0.7 kcal/mol$^{-1}$ if the Cl-H-O angle for the $C_s$ complex is fixed at 160° and all the other internal degrees of freedom are allowed to vary.

*The Water Dimer*

The water dimer has been the subject of much theoretical research.[53] Calculations agree that a
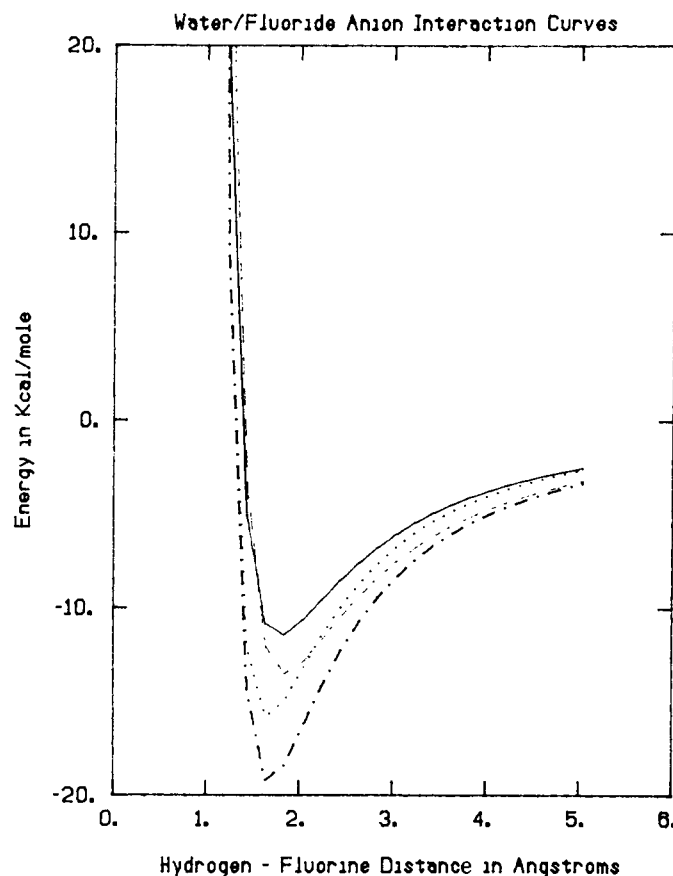
**Figure 7.** Interaction curves for the fluoride anion/water system. Four curves are shown that have been calculated with the combined QM/MM potential. Van der Waal's interactions between the QM and MM atoms are included. They are for the approximation QM(AM1/ $H_2O$)/MM($F^-$) with $\{\alpha, \rho_0\} = \{\infty, 1.0\}$ (solid line) and $\{5.0, 0.0\}$ (dotted line) and for the approximation QM(AM1/$F^-$)/MM($H_2O$) with $\{\alpha, \rho_0\} = \{\infty, 1.0\}$ (dashed line) and $\{5.0, 0.0\}$ (dot-dash line).

$C_s$ symmetry form is the most stable (Fig. 10). This structure also corresponds to experiment.[54] The geometries of $C_s$ and $C_{2v}$ symmetry complexes and the experimental and representative ab initio binding energies and geometries are given in Figure 10 and Table II.

It has been found that the MNDO semiempirical method does not reproduce hydrogen-bonding patterns very well;[20] in fact, AM1 was a reparametrization and extension of the method that had as one of its objectives improvement in the treatment of hydrogen bonds. The AM1 paper by Dewar et al. gives a dimerization energy of 3.3 kcal $mol^{-1}$ for the water dimer which corresponds to a $C_s$ symmetry structure (Fig. 10(a)). Other work[33] has shown that this is not the most stable structure and that one with $C_{2v}$ symmetry, (Fig. 10(b)) has a binding energy that is 1.5 kcal $mol^{-1}$ larger and a cyclic bifurcated complex of approximately $C_{2h}$ symmetry (Fig. 10(c)) is the most stable with a binding energy 0.3 kcal $mol^{-1}$ lower still. The $C_s$ and $C_{2v}$ symmetry structures are saddle points with two and one imaginary frequencies, respec-

tively. These results have suggested that AM1 is unsuitable for modeling water hydrogen bonds. However, a recent ab initio study[56] has found that a bifurcated hydrogen-bond structure similar to the global AM1 minimum, although with larger hydrogen bond lengths, is only 0.2 kcal $mol^{-1}$ higher in energy than the $C_s$ symmetry structure at the level of theory being used. It is clear that considerable caution must be exercized in applying AM1 to bulk water.

The MM and combined QM/MM potentials (one $H_2O$ QM and the other MM) all give minima of $C_s$ symmetry. The $C_{2v}$ forms are between 0.1 and 0.8 kcal $mol^{-1}$ less stable. The ab initio complex $C_s/C_{2v}$ energy difference is about 0.5 kcal $mol^{-1}$.[33] The geometries of the complexes agree with the ab initio and experimental results, although the hydrogen bonding distance for the molecular mechanics complex is too short by about 0.25 Å. The binding energies using the combined potential are all somewhat too small with the parameter set $(\alpha, \rho_0) = \{5.0, 0.0\}$ giving the best results. The MM binding energy is
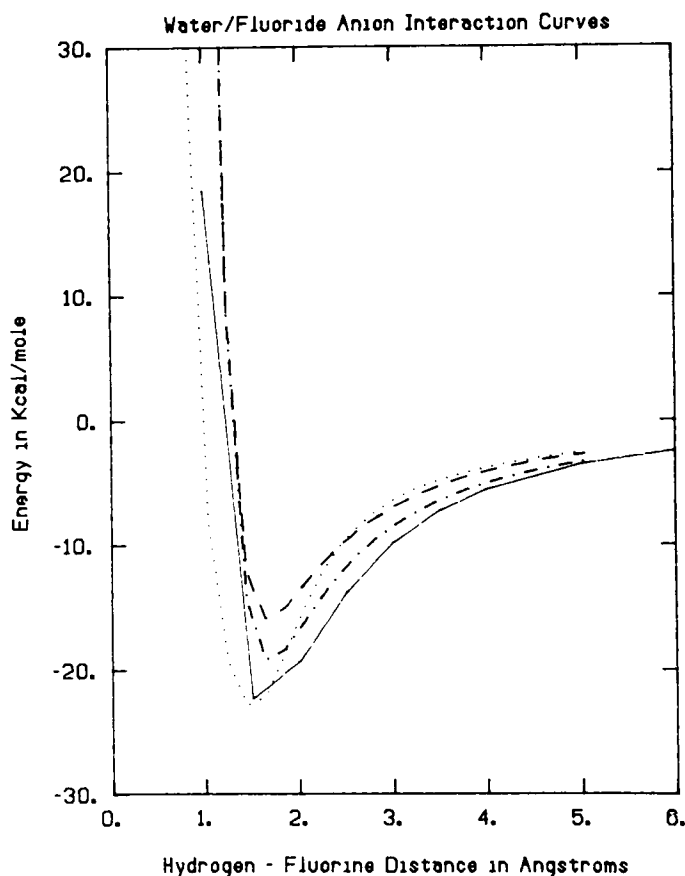
**Figure 8.** Interaction curves for the fluoride anion/water system. Here the BSSE-corrected ab initio (solid line), AM1 (dotted line) and the best two QM/MM curves are compared. The QM/MM curves are for the approximation $QM(AM1/H_2O)/MM(F^-)$ with $\{\alpha, \rho_0\} = \{5.0, 0.0\}$ (dashed line) and $QM(AM1/F^-)/MM(H_2O)$ with $\{\alpha, \rho_0\} = \{5.0, 0.0\}$ (dot-dash line).

slightly too large; this is due to the fact that the parametrization was chosen to include the polarization of water in the liquid state.

### Formate Anion/Water Complexes

The formate anion/water complexes were studied by Lukovits et al.[57] and Jorgensen et al.[58] Studies were carried out using fixed internal geometries for the formate anion and water molecules and for the cases where all internal degrees of freedom were allowed to change subject to the con-

straint that all atoms lie in the same plane. The three most stable complexes have all atoms in a plane and (1) the water hydrogens interacting symmetrically with the formate oxygens (complex 1) or (2) the water making a linear $C—O..H—O$ hydrogen bond and the second hydrogen cis (complex 2) or trans (complex 3) to the other formate oxygen. Their energies and hydrogen bonding distances at the ab initio level and when calculated with the AM1, MM and, combined QM/MM potentials are given in Table III.

In our studies using the AM1, MM, and QM/MM potentials only the $C_{2v}$ symmetry geometry (complex 1), was stable when all internal degrees of freedom were allowed to change; for the fixed internal geometry case, complex 2 and 3 were local minima. All full optimizations collapsed to the $C_{2v}$ symmetry form. Thus, the second set ("relaxed") of results for complexes 2 and 3 for these potentials are for the cases where all internal degrees of freedom are allowed to vary but the $(O^- ... HO)$ hydrogen bond is forced to remain linear.
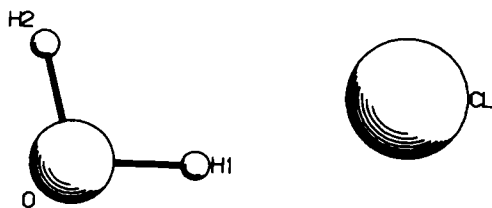


**Figure 9.** Geometry of the $C_s$ symmetry form of the chloride anion/water complex.

**Table I.** Binding energies and geometries from calculations on the chloride anion/water complexes. Energies are in kcal mol$^{-1}$, distances in Å and angles in degrees.

| Approximation | | Symmetry | Binding energy | Cl—O distance | Cl—H—O angle |
|---|---|---|---|---|---|
| Experiment[52] | | | 13.4 | | |
| Ab initio[51] | | $C_s$ | 12.0 | 3.32 | 158.1 |
| | | $C_{2v}$ | 11–12 | | |
| AM1 | (minimum) | $C_s$ | 13.4 | 2.94 | 156.7 |
| | (saddle point) | $C_{2v}$ | 12.6 | 2.97 | 112.1 |
| MNDO | (minimum) | $C_{2v}$ | 8.0 | 3.44 | 112.3 |
| MM | (minimum) | $C_s$ | 14.2 | 3.21 | 143.3 |
| | (saddle point) | $C_{2v}$ | 14.1 | 3.23 | 116.6 |
| QM(AM1/chloride)/MM(water) | | | | | |
| vdw, {∞, 1.0} | | $C_{2v}$ | 11.0 | 3.43 | 116.0 |
| no vdw, {∞, 1.0} | | $C_{2v}$ | 13.6 | 2.72 | 112.2 |
| vdw, {5.0, 0.0} | | $C_{2v}$ | 13.5 | 3.29 | 116.5 |
| no vdw, {5.0, 0.0} | | $C_{2v}$ | 26.5 | 2.18 | 111.3 |
| QM(AM1/water)/MM(chloride) | | | | | |
| vdw, {∞, 1.0} | | $C_{2v}$ | 9.1 | 3.41 | 116.3 |
| no vdw, {∞, 1.0} | | $C_{2v}$ | 36.8 | 1.03 | 65.2 |
| vdw, {5.0, 0.0} | | $C_s$ | 10.4 | 3.32 | 138.9 |
| no vdw, {5.0, 0.0} | | $C_{2v}$ | 39.0 | 1.51 | 91.3 |

The results of all the calculations show the same trends for the binding energies. Complex 1 is the most stable and complex 2 is more stable



**Figure 10.** Geometries of the water dimer: (a) $C_s$ symmetry form; (b) $C_{2v}$ symmetry form; (c) cyclic-bifurcated complex.[33]

than complex 3. The AM1 energies for complexes 2 and 3 are about half the ab initio values. The MM and QM/MM results are in better agreement with the ab initio values, though there are deviations of several kilocalories. For the QM(AM1/H$_2$O)/MM(HCO$_2^-$) model the energy of complex 1 is low too so the energy differences between the complexes are in approximate agreement with the ab initio results while for the QM(AM1/HCO$_2^-$)/MM(H$_2$O) model complex 1 is between 7 and 8 kcal mol$^{-1}$ more stable than complexes 2 and 3, as compared with the ab initio difference of 3 to 4 kcal mol$^{-1}$.

The hydrogen bond distances for the complexes using the combined potential are in approximate agreement with the ab initio values. The AM1 values are too long by about 0.2 to 0.3 Å and the MM values are too short by up to 0.3 Å.

*Formaldehyde/Water Complexes*

Four formaldehyde/water complexes were studied by Madura and Jorgensen with ab initio theory.[59] Some of their geometries are illustrated in Figure 11. Complexes 1 (Fig. 11(a)) and 2 (see Fig. 11(b)) have $C_s$ symmetry and complexes 3 and 4 are of $C_{2v}$ symmetry; complex 3 is planar with the CO group bisecting the HOH angle and complex 4 has the water oxygen bisecting the HCH angle with the water hydrogens perpendicular to the formaldehyde plane. In the ab initio study only the intermolecular geometry was optimized subject to the symmetry constraints. The energies and geometrical parameters for the

**Table II.** Binding energies and geometries of water dimer complexes. Energies are in kcal mol$^{-1}$, distances in Å, and angles in degrees. O...H is the hydrogen-bond distance, O...H—O is the angle between the hydrogen bond and the donor O—H bond and $X$—O...H is the angle between the hydrogen bond and the bisector of the acceptor's H—O—H angle.

| Approximation | Binding energy | O...H | O...H—O | $X$—O...H |
|---|---|---|---|---|
| Experiment[55] | 5.4 ± 0.7 | | | |
| Ab initio[53] | 5.4 | 1.95 | 173 | 125 |
| AM1 $C_s$ symmetry[33] | 3.3 | 2.09 | 161.8 | |
| $C_{2v}$ symmetry | 5.0 | 2.16 | 159.5 | 107.8 |
| MM $C_s$ symmetry | 6.8 | 1.75 | 174.2 | 156.8 |
| $C_{2v}$ symmetry | 4.9 | 2.38 | 111.7 | 161.9 |
| QM(AM1)/MM | | | | |
| QM donor, {∞, 1.0} | 2.6 | 2.28 | 140.8 | 158.8 |
| QM acceptor, {∞, 1.0} | 3.0 | 2.15 | 163.8 | 152.5 |
| QM donor, {5.0, 0.0} | 3.3 | 2.02 | 159.5 | 158.8 |
| QM acceptor, {5.0, 0.0} | 4.6 | 1.99 | 169.0 | 151.1 |

**Table III.** Binding energies ($E$) and hydrogen-bond distances ($R$) of the formate anion/water complexes. Energies are in kcal mol$^{-1}$ and distances are in Å.

| Approximation | Complex 1 Fixed $E$ | Fixed $R$ | Relaxed $E$ | Relaxed $R$ | Complex 2 Fixed $E$ | Fixed $R$ | Relaxed $E$ | Relaxed $R$ | Complex 3 Fixed $E$ | Fixed $R$ | Relaxed $E$ | Relaxed $R$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Ab initio[58] | 22.0 | 3.23 | 19.5 | 1.72 | 18.2 | 1.73 | 18.2 | 3.19 | 15.3 | 1.81 | 14.6 | 1.82 |
| AM1 | 16.0 | 3.20 | 16.1 | 3.22 | 8.7 | 2.04 | 8.8 | 2.03 | 8.0 | 2.05 | 8.1 | 2.04 |
| MM | 21.8 | 3.09 | 25.0 | 3.06 | 16.5 | 1.63 | 18.1 | 1.54 | 14.6 | 1.64 | 16.1 | 1.56 |
| QM(AM1/HCO2)/MM(H2O) | 19.1 | 3.21 | 20.9 | 3.22 | 13.0 | 1.83 | 13.8 | 1.80 | 11.9 | 1.85 | 12.6 | 1.68 |
| QM(AM1/H2O)/MM(HCO2) | 15.5 | 3.22 | 15.9 | 3.22 | 11.6 | 1.72 | 12.0 | 1.81 | 10.3 | 1.73 | 10.6 | 1.69 |

complexes are given in Table IV along with the corresponding values calculated using the AM1, MM, and combined QM/MM potentials.
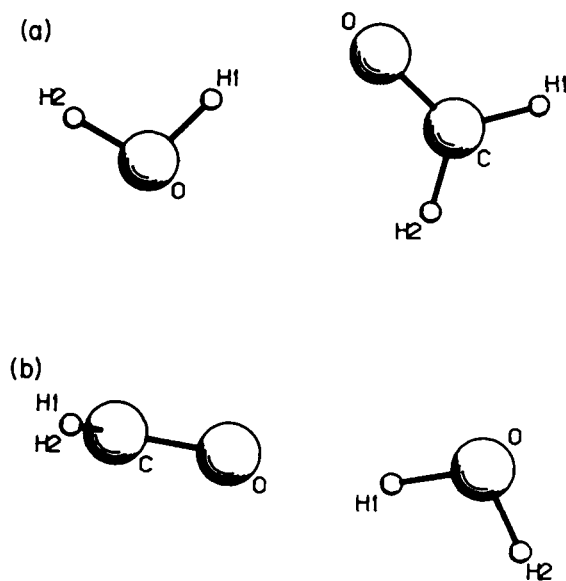


**Figure 11.** Geometries of some formaldehyde/water complexes: (a) complex 1; (b) complex 2.

The MM and the QM(AM1/H$_2$CO)/MM(H$_2$O) models reproduce the ab initio ordering of the various stability complexes. Other than for complex 4, the QM(AM1/H$_2$O)/MM(H$_2$CO) model's energies are too small, so that the energy separation for all complexes is 0.2 kcal or less. The AM1 results have the most disagreement with the ab initio energy bond distances; the O—H distances are too long and the O—O and O—C distances too short. The AM1 complex 2 is, in fact, a $C_{2v}$ symmetry complex with a bifurcated hydrogen bond. The other structural results are in good agreement with the ab initio values, being within 0.08 Å except for two cases — the MM O—H distance for complex 2 is too short by 0.15 Å and the QM(AM1/H$_2$CO)/MM(H$_2$O) O—C distance for complex 4 is too long by 0.21 Å.

## A Formaldehyde/Hydrogen Chloride Complex

The formaldehyde/hydrogen chloride complex has been investigated theoretically using high levels of theory[60] and experimentally.[61] The geometry and binding energy are given in Figure 12 and Table V.

**Table IV.** Binding energies and geometries of the formaldehyde/water complexes. The geometrical parameters are the hydrogen bond distances for complexes 1 and 2, the oxygen-oxygen distance for complex 3, and the water oxygen-carbonyl carbon distance for complex 4. Energies are in kcal mol$^{-1}$ and distances are in Å.

| Complex | Energies | | | | Geometrical Parameters | | | |
| | 1 | 2 | 3 | 4 | 1(O—H) | 2(O—H) | 3(O—O) | 4(C—O) |
|---|---|---|---|---|---|---|---|---|
| Ab initio[59] | 4.7 | 3.8 | 3.4 | 2.5 | 2.09 | 2.15 | 3.09 | 3.26 |
| AM1 | 2.7 | 3.8 | 4.0 | 2.1 | 2.21 | 2.32 | 2.71 | 2.82 |
| MM | 4.1 | 4.1 | 3.3 | 3.0 | 2.00 | 2.00 | 3.06 | 3.24 |
| QM(AM1/H2CO)/MM(H2O) | 4.2 | 4.0 | 3.5 | 1.7 | 2.05 | 2.08 | 3.03 | 3.47 |
| QM(AM1/H2O)/MM(H2CO) | 2.5 | 2.6 | 2.5 | 2.4 | 2.16 | 2.23 | 3.15 | 3.29 |

The AM1, MM, and QM/MM results are given in Table VI. Calculation with the AM1 method predicts a complex with a binding energy of 3.1 kcal mol$^{-1}$, exactly that of the corrected ab initio result, and with a similar geometry, although the O—H bond length is larger by about 0.2 Å and the C—O—H angle is 15° greater. The MM and QM/MM calculations give binding energies which are lower by about 1 kcal mol$^{-1}$ than the ZPE-corrected ab initio value. Further, the complexes are linear and have hydrogen bonding distances that are about 0.4 Å too long. The errors are due to the van der Waals parameters being inappropriate for hydrogen chloride. The results for the geometries can be improved if these are adjusted.

## Ammonium Ion/Water Complexes

Jorgensen and Gao studied ammonium ion/water complexes in the same article as formate anion/water complexes and using the same procedures discussed earlier.[58] Three different approaches of the water to the cation were tried, each with $C_s$ symmetry. Two have linear O...H—N hydrogen bonds, with one of the NH$_4^+$ hydrogens perpendicular to the H—O—N plane (complex 1) and eclipsing one of the water hydrogens in the other (complex 2); the third is shown in Figure 13. Only the intermolecular degrees of freedom were optimized, which because of symmetry means only the N—O distance.

The energies and NO distances are given in Table VI along with the results of calculations
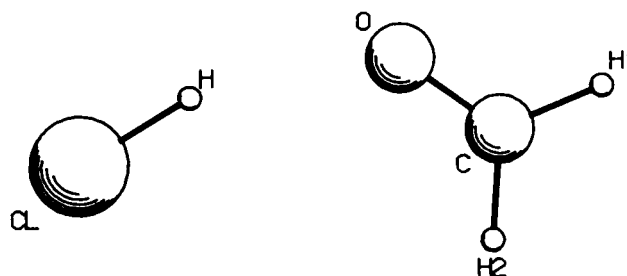


**Figure 12.** Geometry of the formaldehyde/hydrogen chloride complex.

**Table V.** Binding energies and geometries of the formaldehyde/hydrogen chloride complex. The *ab initio* results are corrected for the zero-point vibrational energy difference between the complex and the separated molecules. Energies are in kcal mol$^{-1}$, distances in Å, and angles in degrees.

| Approximation | Binding energy | O—H distance | C—O—H angle | O—H—Cl angle |
|---|---|---|---|---|
| Ab initio[60] | 3.1 | 1.95 | 109 | 163 |
| AM1 | 3.1 | 2.14 | 134 | 159 |
| MM | 1.8 | 2.45 | 179 | 179 |
| QM(AM1/H2CO)/MM(HCl) | 1.8 | 2.44 | 168 | 178 |
| QM(AM1/HCl)/MM(H2CO) | 2.3 | 2.40 | 179 | 179 |

**Table VI.** Binding energies and nitrogen-oxygen distances for the ammonium cation/water complexes. Energies are in kcal mol$^{-1}$ and distances are in Å.

| Complex | Energies | | | N—O Distance | | |
|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 1 | 2 | 3 |
| Ab initio[58] | 21.2 | 18.0 | 17.0 | 2.77 | 2.73 | 2.68 |
| AM1 | 14.0 | 14.0 | 14.2 | 2.97 | 2.97 | 2.67 |
| MM | 20.4 | 20.4 | 15.7 | 2.67 | 2.67 | 2.81 |
| QM(AM1/NH4+)/MM(H2O) | 17.7 | 17.7 | 14.5 | 2.74 | 2.74 | 2.85 |
| QM(AM1/H2O)/MM(NH4+) | 15.9 | 15.9 | 13.0 | 2.81 | 2.81 | 2.83 |

using the MM, AM1, and QM/MM methods. These were carried out in the same way as the ab initio calculations; i.e., the intramolecular parameters were held fixed at the optimized values appropriate to the theory being used.

In the ab initio calculations, the complexes have different energies in order $1 > 2 > 3$, while the other methods all yield $1 > 2 = 3$, except for AM1 which gives $1 = 2 \approx 3$. The MM and QM(AM1/NH$_4^+$)/MM(H$_2$O) are relatively close to the ab initio results in energy and distances, although the NO distance order is different. The AM1 and QM(AM1/H$_2$O)/MM(NH$_4^+$) results are least satisfactory.

*Formamide/Water Complexes*

The formamide/water and formamide/methanol complexes have been studied experimentally by Lovas et al.[62] using microwave spectroscopy. In a theoretical study Jaisen and Stevens[63] performed calculations for the pairs of molecules in four complexes. The microwave structure data supports the complex predicted to be the most stable theoretically, although the hydrogen-bonding distances are shorter by between 0.05 and 0.1 Å. Experimental binding energies are not available.

The energies and geometries of the formamide/water complexes obtained from ab initio, AM1, MM, and QM/MM calculations are given in Table VII and Figure 14. Estimates of the BSSE were made for the ab initio energies by Jaisen

and Stevens and were judged to be about the same for each complex so that the relative stabilities remained unchanged. The absolute value of the BSSE varied depending upon basis set, but was about 25 kcal mol$^{-1}$ for the best basis they used.

The geometries of the complexes were determined by imposing $C_s$ symmetry but allowing all remaining degrees of freedom to vary. Optimized structures were found for each of the complexes examined. All structures had the same general form as the ab initio results. In the AM1 and QM(AM1/HCONH$_2$)/MM(H$_2$O) cases for complex 4, the water hydrogens were 2.23 and 2.60 Å away, respectively, from the formamide oxygen indicating that additional hydrogen bonding is present.

The ab initio result is that complex 1, which is cyclic and has two hydrogen bonds, is about 3 kcal mol$^{-1}$ more stable than the other three complexes. The QM(AM1/HCONH$_2$)/MM(H$_2$O) model predicts the correct order for the binding energies for the complexes. In the AM1 case complex 4 is the most stable while with the QM(AM1/H$_2$O)/MM(HCONH$_2$) model complexes 3 and 4 are of equal stability and have binding energies only 0.2 kcal mol$^{-1}$ greater than complex 1. Complex 2 is predicted to be the least stable for the AM1, MM and QM(AM1/H$_2$O)/MM(HCONH$_2$) models.

The geometries of the complexes are mostly in satisfactory agreement with the ab initio results. The MM model has a tendency to produce hydrogen bonds which are somewhat too short. The combined potential results for complex 1 show that although two hydrogen bonds are produced one is much shorter than the other. The favored interaction is different depending again whether the formamide or water is quantum mechanical. For the QM(AM1/HCONH$_2$)/MM(H$_2$O) results on complex 4 the hydrogen bond is much longer than the *ab initio* results because of the presence of the additional bifurcated hydrogen bonds that were mentioned above. The differences in the angles in each of these cases versus the ab initio results is due to the presence of these much longer hydrogen bonds.
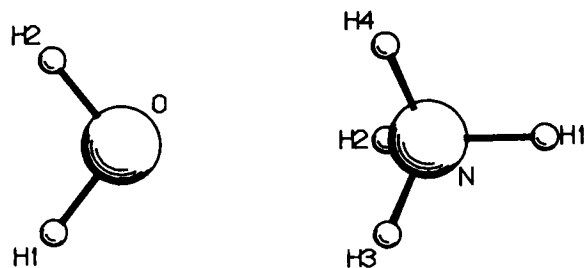


**Figure 13.** Geometry of ammonium cation/water complex 3.

**Table VII.** Binding energies and geometries of the formamide/water complexes. The geometrical parameters are defined in Figure 16. $X$ refers to the bisector of the water molecule H—O—H angle. Energies are in kcal mol$^{-1}$, distances in Å, and angles in degrees.

| Property | | Ab initio[63] | AM1 | MM | QM(AM1/HCONH2) /MM(H2O) | QM(AM1/H2O) /MM(HCONH2) |
|---|---|---|---|---|---|---|
| Binding | 1 | 9.5 | 5.9 | 8.0 | 6.7 | 5.0 |
| Energy | 2 | 6.7 | 3.1 | 4.4 | 5.7 | 2.9 |
| | 3 | 6.0 | 3.5 | 7.4 | 4.0 | 5.2 |
| | 4 | 6.2 | 7.5 | 7.0 | 5.2 | 5.2 |
| Complex 1 | ROH1 | 2.16 | 2.20 | 1.90 | 2.45 | 2.07 |
| Geometry | ROH2 | 2.06 | 2.11 | 2.15 | 2.03 | 2.55 |
| | θHOH | 83.7 | 85.8 | 73.6 | 67.3 | 84.9 |
| | θOHN | 138.6 | 139.1 | 155.9 | 136.4 | 160.6 |
| | θOHO | 143.3 | 139.4 | 146.8 | 161.3 | 126.9 |
| | θCOH | 110.1 | 112.6 | 108.2 | 112.5 | 112.3 |
| Complex 2 | ROH | 2.03 | 2.10 | 2.00 | 2.01 | 2.23 |
| Geometry | θCOH | 118.8 | 120.6 | 129.7 | 124.4 | 141.4 |
| | θOHO | 169.0 | 139.0 | 173.0 | 175.6 | 150.0 |
| Complex 3 | ROH | 2.12 | 2.13 | 1.79 | 2.01 | 1.99 |
| Geometry | θOHN | 175.9 | 176.8 | 177.2 | 182.9 | 176.1 |
| | θHOX | 178.5 | 163.6 | 175.4 | 182.7 | 173.3 |
| Complex 4 | ROH | 2.11 | 2.20 | 1.83 | 2.46 | 2.03 |
| Geometry | θOHN | 177.1 | 171.3 | 169.4 | 137.8 | 164.1 |
| | θHOX | 121.5 | 98.7 | 112.8 | 82.7 | 109.6 |

*Formamide-Methanol Complexes*

The results of the calculations on the formamide/methanol complexes are shown in Figure 15 and Table VIII. As for the formamide complexes, the cyclic system is significantly more stable than the others for the ab initio results. Complex 1 is also most stable for all levels of approximation, in contrast to the water complexes. No method predicts the exact ordering of the ab initio re-
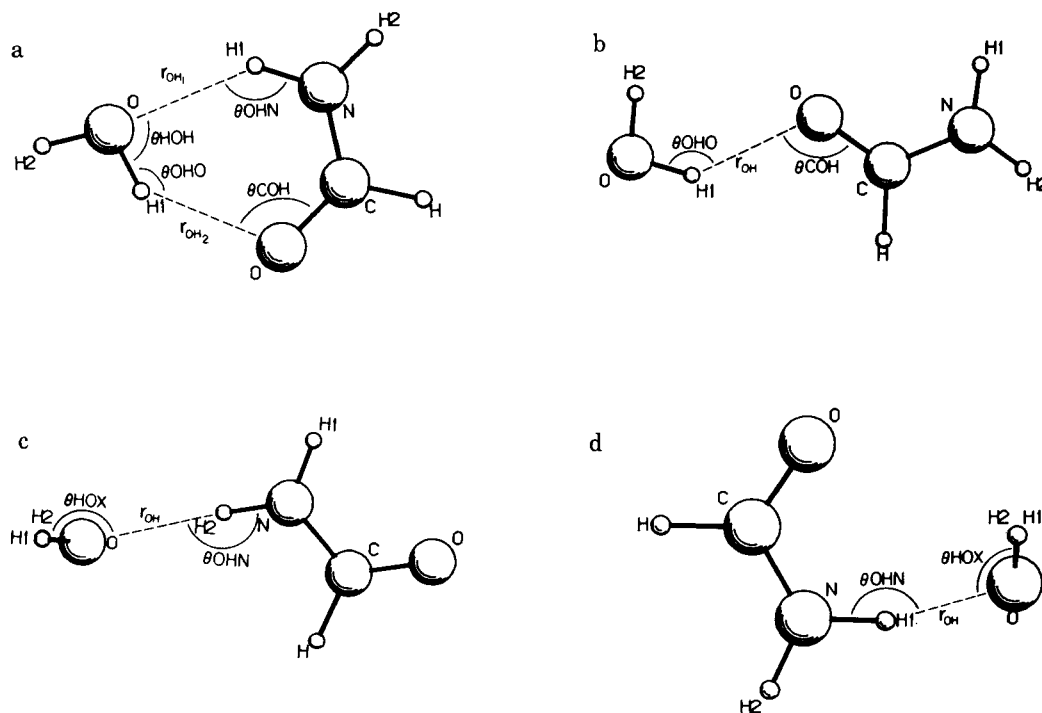


**Figure 14.** Geometries of the formamide/water complexes: (a) complex 1; (b) complex 2; (c) complex 3; (d) complex 4.
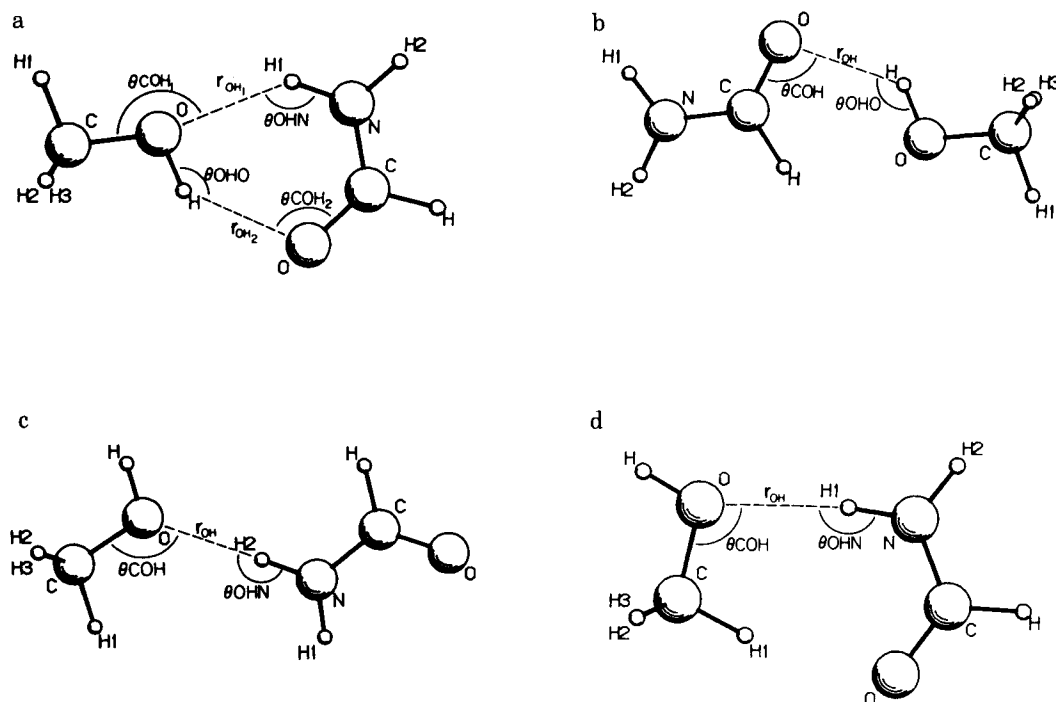
**Figure 15.** Geometries of the formamide/methanol complexes: (a) complex 1; (b) complex 2; (c) complex 3; (d) complex 4.

sults. In complex 1, there are two hydrogen bonds with one significantly shorter than the other in the AM1, QM/MM and MM results, whereas the ab initio complex has two bonds of almost the same length. The geometry for complex 4 with the AM1 and QM(AM1/HCONH$_2$)/MM(CH$_3$OH) method is in better agreement with the ab initio results, because methanol does not

permit the extra bifurcated hydrogen bonding to occur as happens in the water complexes.

## Discussion of the Results on the Intermolecular Interactions

In this section the form of the electrostatic interaction Hamiltonian was investigated and the

**Table VIII.** Binding energies and geometries of the formamide/methanol complexes. The geometrical parameters are defined in Figure 15. Energies are in kcal mol$^{-1}$, distances in Å, and angles in degrees.

| Property | | Ab initio[63] | AM1 | MM | QM(AM1/HCONH2)/MM(CH3OH) | QM(AM1/CH3OH)/MM(HCONH2) |
|---|---|---|---|---|---|---|
| Binding | 1 | 7.6 | 5.5 | 7.0 | 6.4 | 6.1 |
| Energy | 2 | 5.6 | 3.2 | 4.7 | 5.6 | 2.9 |
| | 3 | 4.5 | 3.1 | 5.5 | 3.1 | 5.6 |
| | 4 | 4.7 | 4.1 | 4.9 | 2.8 | 5.8 |
| Complex 1 | ROH1 | 2.08 | 2.21 | 1.96 | 2.32 | 1.97 |
| Geometry | ROH2 | 2.05 | 2.11 | 2.20 | 2.17 | 2.42 |
| | ΘCOH1 | 165.0 | 163.0 | 164.9 | 166.2 | 161.8 |
| | ΘCOH2 | 109.8 | 113.5 | 110.6 | 115.0 | 109.9 |
| | ΘOHO | 144.1 | 137.8 | 135.2 | 139.3 | 128.2 |
| | ΘOHN | 137.3 | 138.8 | 148.4 | 139.3 | 153.7 |
| Complex 2 | ROH | 1.93 | 2.17 | 2.15 | 2.13 | 2.32 |
| Geometry | ΘOHO | 173.0 | 130.8 | 179.2 | 168.3 | 150.5 |
| | ΘCOH | 136.3 | 95.7 | 167.5 | 141.6 | 97.5 |
| Complex 3 | ROH | 2.00 | 2.15 | 1.78 | 1.98 | 1.91 |
| Geometry | ΘOHN | 174.5 | 179.2 | 176.7 | 169.9 | 177.3 |
| | ΘCOH | 127.3 | 126.3 | 85.8 | 156.8 | 125.6 |
| Complex 4 | ROH | 1.95 | 2.16 | 1.87 | 2.02 | 1.93 |
| Geometry | ΘOHN | 176.6 | 171.3 | 158.1 | 162.1 | 174.5 |
| | ΘCOH | 118.4 | 101.8 | 91.2 | 91.3 | 108.5 |

combined QM/MM potential was used to study a range of intermolecular complexes. In all cases comparisons were made between the results of calculations using the AM1, MM and, QM/MM potentials and high level ab initio or experimental results. The combined QM/MM potential reproduces the geometries and energies of the various complexes to the same accuracy as the AM1 and MM models.

It is apparent from the results that the semiempirical and empirical methods do have their shortcomings, so that tests are required prior to their application. In fact, the AM1 Hamiltonian in spite of improvements over MNDO still does not describe water correctly.[33] Thus, as was found in the various tests, MM or QM/MM with MM water tends to give the best results; i.e., although the AM1 method does lead to an attractive water dimer, the MM and QM/MM potentials are more nearly correct.

It is to be emphasized that this is a "zeroth order" combination of a QM and a MM method. The parameters for both the QM and MM methods were used essentially on an as is basis and no attempt was made to optimize them for the present application. The QM(AM1, MNDO) parameters are from the work of Dewar et al.[20,21] and the MM parameters are from the CHARMM force fields[40,41] or taken from Jorgensen et al., if they were not present in CHARMM. The QM and MM atoms interact in exactly the same way as the MM atoms interact with each other, i.e., through electrostatic and van der Waals nonbonding terms. Only two additional parameters were included in the interaction Hamiltonian (i.e., $\alpha$ and $\rho_o$) and they were chosen to be the same for all MM atoms.

Now that the general applicability and some of the limitations of the combined force field are known, it is possible to incrementally improve the methodology. Two directions are in progress. An analysis is being performed for a combined ab initio QM/MM potential[64] and the influence of including polarizability terms on the forcefield is being investigated.[65] The ab initio QM method will allow a "parameter-free" implementation of the combined potential and permits a more rigorous assessment of the deficiencies in the interaction Hamiltonian. Polarizabilities are being added to enhance the accuracy of the MM force field because, at present, there is no way for the charge distribution of the MM atoms to change in response to an external field except by alterations in geometry. Since polarizability effects and other nonadditive terms may be important in some cases for describing correctly molecular interactions,[66] an investigation of their possible role is important. Finally, more detailed treatments of QM/MM interaction terms, including pseudo-potentials are being examined.[67]

## LINK ATOM APPROXIMATION

The link atom approximation is used so that the size of the part of the system to be treated quantum mechanically can be reduced; i.e., a molecule can be partitioned into QM and MM regions. Each MM group was chosen to be neutral because spurious results can arise from the effects of charge separation.[68] In applications of the link atom approximation, it is essential to determine the effects of partitioning on the properties of the molecule. In this way the consequences on calculated results can be evaluated and the appropriate choice can be made as to where partitioning would have least effect. In the following subsections, tests of the link atom approximation on a rotational barrier, proton affinities, deprotonation enthalpies, dipole moments, and ionization potentials are presented.

As in the previous section the calculations were performed by including all nonbonded interactions; i.e., no switching or shifting functions or other truncation methods were applied. The QM/MM electrostatic interaction list was generated as outlined in the third section.

To investigate the relative importance of the MM charge/QM atom interactions arising from the MM charges close to the link atom, different scaling schemes were used; i.e., the interactions between nearest neighbors (1–2) or next nearest neighbors (1–3) were reduced. All the tests presented in this section were performed with different scale factors and it was found that the best results, without exception, were obtained when all MM charge/QM atom interactions were fully included in the calculations. Only these results are presented below.

### Rotational Barriers

The rotational barrier of ethane[69] is 2.93 ± 0.03 kcal mol$^{-1}$. In a recent MM study performed when parameterizing the new CHARMM allatom force field[70] the rotational barrier was found to be dominated by the internal torsional term; because of the high symmetry, van der Waals and electrostatic interactions together contributed only about 0.1 kcal mol$^{-1}$.

A combined QM/MM calculation was performed on the ethane system treating one of the methyl groups by QM and the other by MM. All degrees of freedom except the dihedral angle were allowed to rely in the calculations. The barrier was found to be 2.93 kcal/mol. The torsional term accounted for 2.83 kcal mol$^{-1}$, while the QM/MM van der Waals and QM energy accounted for 0.16 and −0.08 kcal mol$^{-1}$, respectively (whatever $\{\alpha, \rho_0\}$ values were used); the bond and angle energies made up the rest. The AM1 and MNDO procedures give 1.25 and 1.01 kcal mol$^{-1}$, respec-

tively, for the adiabatic barrier in much poorer agreement with experiment. In other examples (e.g., butane and methylamine) rotation barriers are reproduced to within 1 kcal mol$^{-1}$ by the AM1[20] method. In general, the AM1 method appears to underestimate rotational barriers.

## Proton Affinities and Deprotonation Enthalpies

Calculations were performed to determine the gas phase proton affinities and deprotonation enthalpies of a number of compounds. In all cases the group holding the site of protonation or deprotonation was treated quantum mechanically and one link atom was introduced between the QM and MM portions of the molecule. All calculations were performed by fully optimizing the geometry of the cationic or anionic and neutral forms of the molecule.

The proton affinity for the protonation of a base, $B$, is given by

$$PA = \Delta H_f(H^+) + \Delta H_f(B) - \Delta H_f(HB^+) \quad (15)$$

and the deprotonation enthalpy for deprotonation of an acid, $HB$, by

$$DP = \Delta H_f(H^+) + \Delta H_f(B^-) - \Delta H_f(HB) \quad (16)$$

where $\Delta H_f(x)$ is the enthalpy of formation of $x$.

The heat of formation of the proton, $\Delta H_f(H^+)$ is taken as the experimental value of 367.2 kcal mol$^{-1}$,[71] as in standard MNDO and AM1 calculations, and is defined by the process

$$\frac{1}{2}H_2 = H^+ + e^- \quad (17)$$

Experimental and calculated values for some proton affinities are given in Table IX. A series of calculations were performed as indicated in the table. The experimental values and the values calculated quantum mechanically are enthalpies,[72] while the MM values are energies. The appropriate corrections are small (less than 1 kcal mol$^{-1}$) and are not made to the calculated results in the table. The experimental and AM1 calculated values for the proton affinities of ammonia and water are given for comparison. The link atom for each of the calculations was placed along the appropriate carbon-carbon, carbon-nitrogen or carbon-oxygen bond but was free to move in the subsequent optimizations. In all calculations the link atom deviates by, at most, only a few degrees from the direction of the bond along which it was originally placed because the QM atom to which it is bound is constrained by MM bond length and bond angle terms.

The results for both the amines and the alcohols are generally satisfactory; in both cases the

**Table IX.** Proton affinities (in kcal mol$^{-1}$) for some alcohols and primary amines. The partitioning into QM and MM groups is defined in the table.

| Molecule | | Experiment[72] | AM1[72] | Combined potential |
|---|---|---|---|---|
| MM | QM | | | |
| | H2O | 166.5 | 164.5 | — |
| CH3 | OH | 181.9 | 171.9 | 172.5 |
| CH3CH2 | OH | 188.3 | 178.7 | 173.5 |
| CH3 | CH2OH | | | 175.2 |
| (CH3)3C | OH | 193.7 | 187.8 | 174.7 |
| | NH3 | 204.0 | 209.3 | — |
| CH3 | NH2 | 214.1 | 211.1 | 216.6 |
| CH3CH2 | NH2 | 217.0 | 213.4 | 218.1 |
| CH3 | CH2NH2 | | | 212.8 |
| (CH3)3C | NH2 | 220.8 | 220.4 | 219.5 |
| CH3(CH2)3 | NH2 | | | 216.2 |
| CH3(CH2)2 | CH2NH2 | 218.4 | 214.3 | 212.9 |
| CH3CH2 | (CH2)2NH2 | | | 214.5 |
| CH3 | (CH2)3NH2 | | | 212.4 |
| NH2(CH2)4 | NH2 | | | 215.8 |
| NH2(CH2)3 | CH2NH2 | N/A | 212.9 | 212.5 |
| NH2(CH2)2 | (CH2)2NH2 | | | 214.2 |
| NH2CH2 | (CH2)3NH2 | | | 214.1 |
| NH2(CH2)3 | OH | | | 172.1 |
| NH2(CH2)2 | CH2OH | N/A | 176.9 | 173.2 |
| NH2CH2 | (CH2)2OH | | | 179.7 |
| NH2 | (CH2)3OH | | | 179.9 |

correct trends in the values of the proton affinities are observed as a function of chain length. The corresponding AM1 errors are similar. When an extra methylene group in the molecule is treated by QM agreement is improved slightly for ethanol while for ethylamine the proton affinity becomes closer to the full AM1 value and further from experiment.

The calculations for butamine, but-1-4-diamine and 3-aminepropane-1-ol with the combined QM/MM potential compare well with the full AM1 results;[72] partitioning alters the proton affinity by up to about 3.8 kcal mol$^{-1}$ for the amine and 7.5 kcal mol$^{-1}$ for the alcohols. The butamine proton affinity is about 4 kcal mol$^{-1}$ less than experiment for both the AM1 and QM(AM1)/MM calculations. No experimental value is available for comparison with the all-trans form of the diamine or the aminoalcohol. However, the partitioned results for the first species, on average, are higher when using the QM(AM1)/MM approximation and for the aminoalcohol the QM(AM1)/MM results span the AM1 results — two are lower and two are higher.

In comparing the butamine and but-1,4-diamine, we see that the presence of the extra amine group does not have a marked effect on the proton affinity of the other amine. Corresponding behavior is found for the 3-aminepropane-1-ol relative to the alcohol.

The experimental and calculated deprotonation enthalpies for a series of compounds are given in Table X.[72] Equivalent sets of calculations were performed on the primary amines and alcohols as for the proton affinities (except for the t-butylamine, for which there is no experimental value). In addition, calculations were carried out upon a series of carboxylic acids and three calculations were performed on propan-1-ol to test the effect of the partitioning on the result.

The errors for the amine and alcohols are very large (up to 30 kcal mol$^{-1}$) when only the amine or hydroxyl group is treated by QM. This is due to the known failure of AM1 (and MNDO) to adequately describe anions in which the negative charge is localized; e.g., the errors for the AM1 deprotonation enthalpies of ammonia and water are over 20 kcal mol$^{-1}$ in each case. However, the agreement improves substantially for ethylamine, ethanol and propanol if one carbon is treated by QM as well. The errors in these cases are about 10 kcal mol$^{-1}$ and are comparable to the errors of the AM1 method itself. For the acids the agreement is very good; i.e., the error for ethanoic acid is less than 2 kcal mol$^{-1}$ and for propanoic acid less than 9 kcal mol$^{-1}$ if an extra methylene group is treated quantum mechanically.

## Ionization Potentials

To test the effect of partitioning on molecular orbital energies the ionization potentials were calculated for a series of molecules using Koopman's theorem. The results are presented in Table XI. The experimental and AM1 results are taken from reference 20.

Overall, the results with all partitionings are very good; in every case, at least the polar moiety was treated by QM. For acetaldehyde and acetone, the combined potential results are actually in better agreement with experiment than

**Table X.** Deprotonation enthalpies (in kcal mol$^{-1}$) for some alcohols and primary amines. The partitioning into QM and MM groups is defined in the table.

| Molecule | | Experiment[72] | AM1[72] | Combined potential |
|---|---|---|---|---|
| MM | QM | | | |
| | H2O | 390.8 | 412.3 | — |
| CH3 | OH | 379.2/381.4 | 385.4 | 422.8 |
| CH3CH2 | OH | 376.1/378.3 | 384.1 | 422.5 |
| CH3 | CH2OH | | | 387.1 |
| (CH3)3C | OH | 373.3/375.5 | 384.6 | 424.1 |
| | NH3 | 403.6 | 427.0 | — |
| CH3 | NH2 | 403.2 | 407.7 | 434.2 |
| CH3CH2 | NH2 | 399.4 | 409.7 | 435.2 |
| CH3 | CH2NH2 | | | 411.0 |
| CH3(CH2)2 | OH | | | 421.9 |
| CH3CH2 | CH2OH | 374.7/376.9 | 384.1 | 387.3 |
| CH3 | (CH2)2OH | | | 386.0 |
| | HCOOH | 345.2 | 354.6 | — |
| CH3 | COOH | 348.5 | 354.2 | 356.9 |
| CH3CH2 | COOH | 347.3 | 352.9 | 357.8 |
| CH3 | CH2COOH | | | 355.9 |

**Table XI.** First ionization potentials. Values are in electron-volts.

| Molecule | | Experiment[20] | AM1[20] | Combined potential |
|---|---|---|---|---|
| MM | QM | | | |
| | H2O | 12.61 | 12.46 | — |
| CH3 | OH | 10.96 | 11.13 | 11.97 |
| CH3CH2 | OH | N/A | 10.05 | 11.96 |
| CH3 | CH2OH | ⏐ | ⏐ | 11.07 |
| | NH3 | 10.85 | 10.42 | — |
| CH3 | NH2 | 9.45 | 9.76 | 10.24 |
| CH3CH2 | NH2 | N/A | 9.68 | 10.19 |
| CH3 | CH2NH2 | ⏐ | ⏐ | 9.74 |
| | H2CO | 10.89 | 10.78 | — |
| CH3 | CHO | 10.21 | 10.72 | 10.60 |
| (CH3)2 | CO | 9.72 | 10.67 | 10.51 |
| | HCOOH | 11.51 | 11.82 | — |
| CH3 | COOH | N/A | 11.63 | 11.63 |
| CH3CH2 | COOH | N/A | 11.52 | 11.59 |
| CH3 | CH2COOH | ⏐ | ⏐ | 11.45 |

the AM1 results although the difference between them is only about 0.1 eV. For the amines and alcohols the QM/MM results are somewhat worse than the full AM1. For ethylamine and ethanol the agreement is greatly improved if the adjoining methylene group is treated quantum mechanically as well.

The higher ionization potentials of the molecules are much more poorly reproduced because the lower energy orbitals involve the atoms that are missing in the combined QM/MM model. Clearly, it is important to consider the form of the orbital when ionization potentials are being determined. For example, if the first ionization potential of acetaldehyde was being calculated but the CHO group was treated by MM and the methyl group by QM, the results would be inappropriate since the highest energy occupied orbitals are predominantly on the carbonyl carbon and oxygen.

## Dipole Moments and Atomic Charges

The experimental and calculated dipole moments of the molecules in the previous subsection are given in Table XII. The experimental and AM1 dipole moments were obtained from reference 20. The combined QM/MM dipole moments were calculated as the sum of the moments on the MM and QM portions of the molecule. The MM dipole, $\mu_M$, is:

$$\mu_M = \sum_M q_M(\mathbf{R}_M - \mathbf{R}_C) \qquad (18)$$

**Table XII.** Dipole moments. The units are Debyes.

| Molecule | | Experiment[20] | AM1[20] | Combined potential |
|---|---|---|---|---|
| MM | QM | | | |
| | H2O | 1.85 | 1.86 | — |
| CH3 | OH | 1.70 | 1.62 | 1.97 |
| CH3CH2 | OH | 1.69 | 1.55 | 1.92 |
| CH3 | CH2OH | ⏐ | ⏐ | 1.53 |
| | NH3 | 1.47 | 1.85 | — |
| CH3 | NH2 | 1.31 | 1.49 | 1.96 |
| CH3CH2 | NH2 | 1.22 | 1.53 | 1.94 |
| CH3 | CH2NH2 | ⏐ | ⏐ | 1.63 |
| | H2CO | 2.33 | 2.32 | — |
| CH3 | CHO | 2.69 | 2.69 | 2.51 |
| (CH3)2 | CO | 2.88 | 2.92 | 2.77 |
| | HCOOH | 1.41 | 1.48 | — |
| CH3 | COOH | 1.74 | 1.89 | 1.69 |
| CH3CH2 | COOH | 1.75 | 1.95 | 1.66 |
| CH3 | CH2COOH | ⏐ | ⏐ | 1.90 |

where $\mathbf{R}_C$ is the dipole center. The QM dipole is calculated as the sum of a charge term comparable in form to eq. (18) and a hybridization term; details are given in reference 73. In the calculations, the charge on the link atom has to be included; otherwise the dipole is not origin independent for a neutral molecule.

From the table it can be seen that the dipole moments for the alcohols and the amines are too high when calculated with the combined QM/MM potential and only the amine and hydroxyl groups are treated by QM. The agreement improves substantially when an extra methylene group is included in the QM calculation.

The contributions of the QM and MM portions to the dipole moments for the QM/MM results are given in Table XIII. The values for the MM parts are nearly constant for all species as they consist of methyl or ethyl groups. The QM parts are very similar for the different amines, alcohols and carbonyls when only the amine and hydroxyl groups are treated by QM; as soon as a methylene group is included in the QM part, the value changes significantly.

The charges on the atoms are given in Table XIV. They are calculated as the sum of the diagonal density matrix elements, $P$, for each orbital on the atom, in accord with the Mulliken approximation,[74] i.e.,

$$Q_i = \sum_{\mu_i} P_{\mu\mu} \qquad (19)$$

The methyl carbon and hydrogen partial charges for acetaldehyde and acetone are in good agreement with the ones that are used in the MM model. For ethylamine and propanol the agreement for the methyl atom and methylene atom

charges that are not adjacent to the heteratom group is also good. The charges for ethanol or for methylene groups adjacent to an amine or hydroxyl group are not in agreement with the MM charges used for aliphatic carbons and hydrogens. Agreement is better if that group is included in the QM calculations.

The link atom charges (atom type QQ) have magnitudes from 0 to 0.2. In all cases the charge of the link atom can be reduced by increasing the size of the QM portion of the molecule. For example, for propanol the link atom charge goes down from 0.14 to 0.0 as an extra methylene group is made quantum mechanical and stays low (at $-0.02$) as yet another methylene group is treated by QM. That the link atom charge be near zero is a useful criterion for choosing the lower limits of the QM portion of the molecule.

## Comments on the Link Atom Tests

The link atom method developed earlier is rather simple. Nevertheless, the tests presented in this section show that it works well, as long as care is exercised when dividing the molecule into QM and MM parts. It is apparent that changing the partitioning can have significant effects on results. Thus, it is essential that a number of different partitionings be tried to determine the effects on the properties of the system.

There are no definitive rules about how to partition a molecule into separate QM and MM regions but a few guidelines can be given. As the link atoms are hydrogens, they will be joined to their QM atoms by $\sigma$-bonds. Therefore, partitioning should not break $\pi$-bonds or bonds in which conjugation effects will be important. For example, if a histidine side-chain in a protein is to be treated quantum mechanically all the imidazole ring atoms should be QM atoms. Partitioning of other systems should be done correspondingly. As long as conjugation, delocalization, inductive, or other electronic effects are of importance, that part of the molecules has to be treated quantum mechanically.

Work is underway to improve the model.[75] The introduction of polarizabilities for the MM atoms[65] will enable the charge distribution in the molecule to be described more realistically. It is planned to investigate pseudopotential methods,[67] which permit a more rigorous treatment of the bonding interactions of the MM and QM atoms than the simple point charge/van der Waals model used here. Fragment molecular orbital methods[76] are also being examined since they remove the need for the introduction of an unchemical link atom into the system for each QM/MM bond that is broken. However, even without these refinements, the link atom ap-

**Table XIII.** The QM and MM contributions to the total dipole moments for some molecules calculated using the combined QM/MM potential. The units are Debyes.

| Molecule | | QM | MM |
|---|---|---|---|
| MM | QM | | |
| CH3 | OH | 1.68 | 0.49 |
| CH3CH2 | OH | 1.67 | 0.46 |
| CH3 | CH2OH | 1.81 | 0.47 |
| CH3 | NH2 | 1.80 | 0.48 |
| CH3CH2 | NH2 | 1.79 | 0.46 |
| CH3 | CH2NH2 | 1.16 | 0.47 |
| CH3 | CHO | 2.26 | 0.50 |
| (CH3)2 | CO | 2.24 | 0.53 |
| CH3 | COOH | 1.38 | 0.49 |
| CH3CH2 | COOH | 1.34 | 0.49 |
| CH3 | CH2COOH | 2.02 | 0.47 |

**Table XIV.** Atomic partial charges on the QM atoms. Where there are methyl or methylene hydrogens attached to the same carbon in different environments the average value of the charge for all the hydrogens is given.

| Molecule MM | QM | Atom | AM1 | Combined potential |
|---|---|---|---|---|
| CH3 | OH | C | −0.07 | −0.27 |
| | | H(C) | 0.07 | 0.09 |
| | | O | −0.33 | −0.31 |
| | | H(O) | 0.20 | 0.18 |
| | | QQ | — | 0.13 |
| CH3CH2 | OH | C(CH3) | 0.19 | −0.27 |
| | | H(CH3) | 0.00 | 0.09 |
| | | C(CH2) | −0.34 | −0.18 |
| | | H(CH2) | 0.09 | 0.09 |
| | | O | −0.23 | −0.31 |
| | | H(O) | 0.21 | 0.17 |
| | | QQ | — | 0.13 |
| CH3 | CH2OH | C(CH3) | 0.19 | −0.27 |
| | | H(CH3) | 0.00 | 0.09 |
| | | C(CH2) | −0.34 | 0.06 |
| | | H(CH2) | 0.09 | 0.04 |
| | | O | −0.23 | −0.33 |
| | | H(O) | 0.21 | 0.19 |
| | | QQ | — | −0.01 |
| CH3 | NH2 | C | −0.13 | −0.27 |
| | | H(C) | 0.07 | 0.09 |
| | | N | −0.35 | −0.27 |
| | | H(N) | 0.14 | 0.11 |
| | | QQ | — | 0.05 |
| CH3CH2 | NH2 | C(CH3) | −0.21 | −0.27 |
| | | H(CH3) | 0.08 | 0.09 |
| | | C(CH2) | −0.08 | −0.18 |
| | | H(CH2) | 0.06 | 0.09 |
| | | N | −0.35 | −0.27 |
| | | H(N) | 0.14 | 0.11 |
| | | QQ | — | 0.05 |
| CH3 | CH2NH2 | C(CH3) | −0.21 | −0.27 |
| | | H(CH3) | 0.08 | 0.09 |
| | | C(CH2) | −0.08 | 0.02 |
| | | H(CH2) | 0.06 | 0.07 |
| | | N | −0.35 | −0.36 |
| | | H(N) | 0.14 | 0.14 |
| | | QQ | — | −0.08 |
| CH3(CH2)2 | OH | C(CH3) | −0.21 | −0.27 |
| | | H(CH3) | 0.07 | 0.09 |
| | | C(CH2) | −0.16 | −0.18 |
| | | H(CH2) | 0.09 | 0.09 |
| | | C(CH2) | −0.02 | −0.18 |
| | | H(CH2) | 0.06 | 0.09 |
| | | O | −0.33 | −0.31 |
| | | H | 0.20 | 0.17 |
| | | QQ | — | 0.14 |
| CH3CH2 | CH2OH | C(CH3) | −0.21 | −0.27 |
| | | H(CH3) | 0.07 | 0.09 |
| | | C(CH2) | −0.16 | −0.18 |
| | | H(CH2) | 0.09 | 0.09 |
| | | C(CH2) | −0.02 | 0.05 |
| | | H(CH2) | 0.06 | 0.05 |
| | | O | −0.33 | −0.34 |
| | | H | 0.20 | 0.19 |
| | | QQ | — | 0.00 |
| CH3 | (CH2)2OH | C(CH3) | −0.21 | −0.27 |
| | | H(CH3) | 0.07 | 0.09 |
| | | C(CH2) | −0.16 | −0.07 |
| | | H(CH2) | 0.09 | 0.07 |
| | | C(CH2) | −0.02 | −0.03 |

**Table XIV.** Continued

|  |  |  |  |  |
|---|---|---|---|---|
|  |  | H(CH2) | 0.06 | 0.06 |
|  |  | O | −0.33 | −0.33 |
|  |  | H | 0.20 | 0.20 |
|  |  | QQ | — | −0.02 |
| CH3 | CHO | C(CH3) | −0.29 | −0.27 |
|  |  | H(CH3) | 0.10 | 0.09 |
|  |  | C | 0.18 | 0.26 |
|  |  | H | 0.08 | 0.06 |
|  |  | O | −0.29 | −0.29 |
|  |  | QQ | — | −0.02 |
| (CH3)2 | CO | C(CH3) | −0.27 | −0.27 |
|  |  | H(CH3) | 0.10 | 0.09 |
|  |  | C | 0.22 | 0.43 |
|  |  | O | −0.29 | −0.34 |
|  |  | QQ | — | −0.04 |
| CH3 | COOH | C(CH3) | −0.22 | −0.27 |
|  |  | H(CH3) | 0.12 | 0.09 |
|  |  | C | 0.31 | 0.38 |
|  |  | O | −0.36 | −0.37 |
|  |  | O(OH) | −0.32 | −0.33 |
|  |  | H | 0.24 | 0.24 |
|  |  | QQ | — | 0.09 |
| CH3CH2 | COOH | C(CH3) | −0.22 | −0.27 |
|  |  | H(CH3) | 0.09 | 0.09 |
|  |  | C(CH2) | −0.15 | −0.18 |
|  |  | H(CH2) | 0.12 | 0.09 |
|  |  | C | 0.30 | 0.37 |
|  |  | O | −0.36 | −0.37 |
|  |  | O(OH) | −0.32 | −0.33 |
|  |  | H | 0.24 | 0.24 |
|  |  | QQ | — | 0.10 |
| CH3 | CH2COOH | C(CH3) | −0.22 | −0.27 |
|  |  | H(CH3) | 0.09 | 0.09 |
|  |  | C(CH2) | −0.15 | −0.07 |
|  |  | H(CH2) | 0.12 | 0.10 |
|  |  | C | 0.30 | 0.30 |
|  |  | O | −0.36 | −0.36 |
|  |  | O(OH) | −0.32 | −0.33 |
|  |  | H | 0.24 | 0.02 |
|  |  | QQ | — | 0.10 |
| CH3 | COO— | C(CH3) | −0.27 | −0.27 |
|  |  | H(CH3) | 0.05 | 0.09 |
|  |  | C | 0.32 | 0.39 |
|  |  | O | −0.60 | −0.62 |
|  |  | QQ | — | −0.14 |
| CH3CH2 | COO— | C(CH3) | −0.18 | −0.27 |
|  |  | H(CH3) | 0.07 | 0.09 |
|  |  | C(CH2) | −0.22 | −0.18 |
|  |  | H(CH2) | 0.06 | 0.09 |
|  |  | C | 0.33 | 0.38 |
|  |  | O | −0.63 | −0.62 |
|  |  | QQ | — | −0.13 |
| CH3 | CH2COO— | C(CH3) | −0.18 | −0.27 |
|  |  | H(CH3) | 0.07 | 0.09 |
|  |  | C(CH2) | −0.22 | −0.12 |
|  |  | H(CH2) | 0.06 | 0.03 |
|  |  | C | 0.33 | 0.31 |
|  |  | O | −0.63 | −0.60 |
|  |  | QQ | — | −0.06 |

proach, when applied appropriately, seems to have the same accuracy as the AM1 method itself.

## CONCLUDING DISCUSSION

In the present article a combined semiempirical QM/MM potential has been described and the

results of tests to illustrate its usefulness and limitations have been presented. The method is not exact and careful testing and calibration are required to gauge the accuracy of the results and to ensure that the correct behavior of the physical system under investigation is reproduced. This is no different from other applications of electronic structure calculations, though in many cases no such tests have been performed.

To indicate how the method could be used, we consider a condensed phase reaction. The following procedure could be used:

1. Investigate the reaction in the gas phase. If the system is small enough, treat the full system quantum mechanically; otherwise partition the system into QM and MM regions. Determine the minima, saddle points and reaction path for the reaction and compare the results with experiment and high level ab initio calculations. If the system is too big to perform expensive ab initio calculations, use smaller model systems. If link atoms are being used, test that the partitioning and the MM and QM/MM parameters reproduce to sufficient accuracy the conformational energetics and other properties of the full system.

2. Analyze the interactions of the reacting molecules with one or a small number of solvent molecules. Compare the QM/MM results with results from ab initio calculations or experiment.

3. Perform QM/MM calculations on the full system. The results of calculations in points 1 and 2 should enable some estimates of the accuracy of the calculations to be made. Other effects — such as the nonadditivity of interaction energies and forces between solute and solvent[51,66] — that were not accounted for during the initial analysis and parameterization of the reduced systems may have to be tested for and estimates made of their importance.

We have shown in this methodological article and in an application[46] that the QM/MM method is useful for studying systems for which purely molecular mechanical methods are inappropriate. We view it as a logical step in the extension of theoretical methods to problems involving reactions of chemical and biological systems. The QM/MM potential, although relatively computer intensive, is more flexible and less laborious to use than MM force fields which are based on multidimensional analytic representations of reacting surfaces that have been fitted by use of ab initio calculations on model systems.

Many of the results given by the combined potential for the binding energies and geometries of the complexes could be improved by altering the QM and MM atom van der Waals parameters

or the MM atom charges. In most cases only slight changes are necessary because the binding of the complexes depends upon very subtle effects; e.g., in the MM model of the $Cl^-/H_2O$ complex it is necessary only to alter the van der Waals parameters on the oxygen or to add very small ones on the hydrogens to make the stable complex have $C_{2v}$ symmetry instead of $C_s$.

Work is currently underway to improve the accuracy of the potential. The QM part of the potential is being improved by the addition of an ab initio MO method.[64] Although it is much slower than the semiempirical approach it will enable higher accuracy calculations to be performed on combined QM/MM systems or as checks of the semiempirical QM/MM method. Polarizabilities have been added to the potential and their effects assessed.[65] The link atom approximation and alternatives to it are being investigated.[75] It is hoped in this way to achieve more generally applicable and more consistently accurate QM/MM potentials.[79]

## APPENDIX: THE MOLECULAR MECHANICS PARAMETERS

The parameters for the MM atoms in the calculations were taken from the all-atom CHARMM 21 force field[41] and supplemented with others where required. The parameters, including atomic partial charges, are listed for each molecule below. The combination rules used for the van der Waal's parameters are (eq. (9)):

$$A_{ij} = \varepsilon_{ij}\sigma_{ij}^{12} \quad \text{and} \quad B_{ij} = 2\varepsilon_{ij}\sigma_{ij}^6 \quad (A.1)$$

with $\varepsilon_{ij}$ and $\sigma_{ij}$ given by:

$$\varepsilon_{ij} = (E_{\min_i}E_{\min_j})^{1/2} \quad \text{and}$$

$$\sigma_{ij} = \frac{1}{2}(R_{\min_i} + R_{\min_j}) \quad (A.2)$$

All energies, lengths, and angles are in kcal $mol^{-1}$, Angstroms, and degrees respectively. Force constants are in kcal $mol^{-1}$ $Å^{-2}$ for bonds, kcal $mol^{-1}$ $degrees^{-2}$ for angles, and improper dihedrals and kcal $mol^{-1}$ for proper dihedrals. Partial charges are in units of the electron charge and the van der Waal's $E_{\min}$ and $R_{\min}$ parameters are in kcal $mol^{-1}$ and Å respectively. All possible bond angles and proper dihedral angles defined by the connectivity are included in the calculation of the internal energy and Urey–Bradley terms are included for hydrocarbon groups.

## Water

The water model used is the TIPS3P model of Jorgensen et al.[77] modified so as to include internal degrees of freedom. In the CHARMM modified model commonly used van der Waal's radii have been added to the hydrogen atoms. This, however, produces the wrong geometry for the $Cl^-/H_2O$ complex (a $C_{2v}$ rather than a $C_s$-symmetry form) so the original van der Waal's parameters are used throughout in this work.

|   | Charge | $E_{min}$ | $R_{min}$ |
|---|---|---|---|
| O | -0.834 | -0.152 | 1.768 |
| H | 0.417 | -0.0 | 0.0 |

The internal coordinate terms are as follows:

|   | Force Constant | Equilibrium Value |
|---|---|---|
| O—H bond | 450.0 | 0.9572 |
| H—O—H angle | 55.0 | 104.52 |

## Fluoride and Chloride Anions

The parameters for the halogens are derived from the ones obtained by Jorgensen et al.[78]

|   | Charge | $E_{min}$ | $R_{min}$ |
|---|---|---|---|
| F | -1.0 | -0.720 | 1.534 |
| Cl | -1.0 | -0.118 | 2.479 |

## Formate Anion

The nonbond parameters are:

|   | Charge | $E_{min}$ | $R_{min}$ |
|---|---|---|---|
| C | 0.5 | -0.105 | 2.105 |
| O | -0.8 | -0.210 | 1.661 |
| H | 0.1 | -0.030 | 1.300 |

The internal coordinate terms are:

|   | Force Constant | Equilibrium Value |
|---|---|---|
| H—C bond | 330.0 | 1.110 |
| C—O bond | 450.0 | 1.260 |
| H—C—O angle | 50.0 | 118.0 |
| O—C—O angle | 50.0 | 128.0 |
| C improper | 100.0 | 0.0 |

## Formaldehyde

The nonbond parameters are:

|   | Charge | $E_{min}$ | $R_{min}$ |
|---|---|---|---|
| C | 0.18 | -0.105 | 2.105 |
| O | -0.38 | -0.250 | 1.796 |
| H | 0.10 | -0.030 | 1.300 |

The internal coordinate terms are:

|   | Force Constant | Equilibrium Value |
|---|---|---|
| H—C bond | 330.0 | 1.110 |
| C—O bond | 595.0 | 1.215 |
| H—C—O angle | 50.0 | 121.7 |
| H—C—H angle | 50.0 | 118.0 |
| C improper | 89.0 | 0.0 |

## Hydrogen Chloride

The nonbond parameters are:

|   | Charge | $E_{min}$ | $R_{min}$ |
|---|---|---|---|
| H | 0.18 | 0.0 | 0.0 |
| Cl | -0.18 | -0.118 | 2.479 |

There is one bond term which has an equilibrium distance of 1.275 Å and a force constant of 400.0 kcal mol$^{-1}$ Å$^{-2}$.

## Ammonium Cation

The nonbond parameters are:

|   | Charge | $E_{min}$ | $R_{min}$ |
|---|---|---|---|
| N | -0.40 | -0.170 | 1.824 |
| H | -0.35 | 0.0 | 0.0 |

The internal coordinate terms are as follows:

|   | Force Constant | Equilibrium Value |
|---|---|---|
| N—H bond | 405.0 | 0.98 |
| H—N—H angle | 40.0 | 109.5 |

## Formamide

The nonbond parameters are:

|   | Charge | $E_{min}$ | $R_{min}$ |
|---|---|---|---|
| C | 0.280 | -0.105 | 2.105 |
| HC | 0.100 | -0.0045 | 1.468 |
| O | -0.380 | -0.250 | 1.796 |
| N | -0.830 | -0.170 | 1.824 |
| HN | 0.415 | 0.0 | 0.0 |

The internal coordinate terms are:

|   | Force Constant | Equilibrium Value |
|---|---|---|
| H—C bond | 330.0 | 1.110 |
| C—O bond | 595.0 | 1.215 |
| C—N bond | 403.0 | 1.305 |
| N—H bond | 405.0 | 0.98 |
| H—C—O angle | 50.0 | 121.7 |
| H—C—N angle | 41.0 | 120.0 |
| O—C—N angle | 60.0 | 120.6 |
| C—N—H angle | 35.3 | 120.0 |
| H—N—H angle | 40.0 | 109.5 |
| C improper | 5.75 | 0.0 |
| N improper | 5.75 | 0.0 |

There are four dihedral angle terms about the C—N bond. Each term has a force constant of 5.0 kcal mol$^{-1}$, a periodicity of 2, and a phase of 180.0°.

## Methanol

The nonbond parameters are:
s5274min

|     | Charge | $E_{min}$ | $R_{min}$ |
|-----|--------|-----------|-----------|
| C   | −0.27  | −0.0903   | 1.800     |
| HC  | 0.09   | −0.0045   | 1.468     |
| O   | −0.40  | −0.1591   | 1.600     |
| HO  | 0.40   | −0.0498   | 0.800     |

The internal coordinate terms are:

|                | Force Constant | Equilibrium Value |
|----------------|----------------|-------------------|
| H—C bond       | 317.13         | 1.100             |
| C—O bond       | 375.0          | 1.420             |
| O—H bond       | 450.0          | 0.970             |
| H—C—H angle    | 36.0           | 108.5             |
| H—C—O angle    | 50.7           | 110.0             |
| C—O—H angle    | 57.6           | 106.0             |

There are three dihedral angle terms about the C—O bond. Each term has a force constant of 0.5 kcal mol$^{-1}$, a periodicity of 3, and a phase of 0.0°.

## Ethane

The nonbond parameters are:

|   | Charge | $E$     | $R_{min}$ |
|---|--------|---------|-----------|
| C | −0.27  | −0.0903 | 1.800     |
| H | 0.09   | −0.0045 | 1.468     |

The internal coordinate terms are:

|                          | Force Constant | Equilibrium Value |
|--------------------------|----------------|-------------------|
| H—C bond                 | 317.13         | 1.100             |
| C—C bond                 | 229.63         | 1.530             |
| H—C—H angle              | 36.00          | 108.5             |
| H—C—C angle              | 33.43          | 109.5             |
| H—C—H Urey-Bradley       | 5.20           | 1.771             |
| H—C—C Urey-Bradley       | 22.53          | 2.056             |

There are nine dihedral angle terms about the C—C bond. Each term has a force constant of 0.157 kcal mol$^{-1}$, a periodicity of 3, and a phase of 0.0°.

## Amines and Alcohols

The same parameters are used for the cationic, neutral, and anionic forms of the molecules. The hydroxyl and amine groups do not have any partial charges if they are being treated quantum

mechanically although there are internal coordinate and van der Waal's terms associated with them. The amine charges listed are for the amine group in but-1,4-diamine in which one amine group is treated molecular mechanically. It is always neutral. The carbon charges vary depending upon the group they are in. Methyl and methylene group carbons have charges of −0.27 and −0.18 respectively. Tertiary carbons have zero charge.

The nonbond parameters are:

|     | Charge | $E_{min}$ | $R_{min}$ |
|-----|--------|-----------|-----------|
| C   | varies | −0.0903   | 1.800     |
| HC  | 0.09   | −0.0045   | 1.468     |
| N   | −0.60  | −0.2384   | 1.600     |
| HN  | 0.30   | −0.0498   | 0.800     |
| O   |        | −0.1591   | 1.600     |
| OH  |        | −0.0498   | 0.800     |

The internal coordinate terms are:

|                          | Force Constant | Equilibrium Value |
|--------------------------|----------------|-------------------|
| H—C bond                 | 317.13         | 1.100             |
| C—C bond                 | 229.63         | 1.530             |
| H—C—H angle              | 36.00          | 108.5             |
| H—C—C angle              | 33.43          | 109.5             |
| C—C—C angle              | 58.35          | 113.0             |
| H—C—H Urey-Bradley       | 5.20           | 1.771             |
| H—C—C Urey-Bradley       | 22.53          | 2.056             |
| C—C—C Urey-Bradley       | 11.16          | 2.558             |
| H—C—N angle              | 50.0           | 108.9             |
| C—C—N angle              | 50.0           | 101.0             |
| C—N—H angle              | 50.0           | 109.47            |
| H—N—H angle              | 50.0           | 104.5             |
| H—C—O angle              | 50.7           | 110.0             |
| C—C—O angle              | 50.0           | 101.5             |
| C—O—H angle              | 57.6           | 106.0             |

There are nine dihedral angle terms about each C—C bond. Each term has a force constant of 0.157 kcal mol$^{-1}$, a periodicity of 3, and a phase of 0.0°. There are also dihedral terms for the C—N and C—O bonds if the carbon of the bond is being treated molecular mechanically—the number of dihedrals depends upon the protonation state of the nitrogen or oxygen. Each dihedral has a periodicity of 3, a phase of 0.0°, and a force constant of 0.158 or 0.5 kcal mol$^{-1}$ for nitrogen and oxygen respectively.

## Carboxylic Acids

No charges are given for the acid group as they are always treated quantum mechanically. The nonbond parameters are:

|  | Charge | $E_{min}$ | $R_{min}$ |
|---|---|---|---|
| Methyl C | −0.27 | −0.0903 | 1.800 |
| H(CH3) | 0.09 | −0.0045 | 1.468 |
| C |  | −0.0903 | 1.800 |
| O(carbonyl) |  | −0.6469 | 1.300 |
| O(OH) |  | −0.1591 | 1.600 |
| H(OH) |  | −0.0498 | 0.800 |

The aliphatic internal coordinate terms are the same as for the amines and alcohols. MM terms between only QM atoms are omitted too. The rest are:

|  | Force Constant | Equilibrium Value |
|---|---|---|
| C—C(methyl-acid) bond | 187.00 | 1.530 |
| C(CH3)—C(CH2)— C(acid) angle | 51.8 | 110.0 |
| C(CH2)—C— O(carbonyl) angle | 40.0 | 125.0 |
| C(CH2)—C— O(OH) angle | 35.0 | 110.0 |
| H(CH2)—C(CH2)— C(carbonyl) angle | 55.3 | 109.5 |
| Carbonyl carbon improper torsion | 100.0 | 0.0 |

The dihedral terms for the aliphatic carbons are the same as for the amines and alcohols. There are also six dihedrals about the methyl or methylene carbon—carbonyl carbon bond. Each has a force constant of 0.163 kcal mol$^{-1}$, a periodicity of 3 and an equilibrium angle of zero. One term is included for the C—C(carbonyl)—O(OH)—H(OH) dihedral too. It has a force constant, periodicity and phase of 0.4, 2, and 0.0 respectively.

### Acetaldehyde and Acetone

In these molecules the methyl, HCO group for acetaldehyde and CO group for acetone are treated as groups for the scaling scheme. No MM charges are given for the HCO or CO groups because they are always treated quantum mechanically. The nonbond parameters are:

|  | Charge | $E_{min}$ | $R_{min}$ |
|---|---|---|---|
| Methyl C | −0.27 | −0.0903 | 1.800 |
| HC | 0.09 | −0.0045 | 1.468 |
| C |  | −0.0903 | 1.800 |
| H |  | −0.0498 | 0.800 |
| O |  | −0.1591 | 1.600 |

The aliphatic internal coordinate terms are the same as for the amines and alcohols. The rest are:

|  | Force Constant | Equilibrium Value |
|---|---|---|
| C—C bond | 187.00 | 1.530 |
| C—O bond | 595.00 | 1.215 |
| C—H(carbonyl) bond | 330.00 | 1.110 |
| Methyl H—C—C angle | 55.3 | 109.5 |
| Carbonyl H—C—C angle | 50.0 | 116.0 |
| C—C—O | 64.6 | 122.3 |
| H—C—O | 50.0 | 121.7 |
| Methyl C —C— methyl C | 50.0 | 115.4 |
| Carbonyl C improper dihedral | 10.0 | 180.0 |

There are no proper dihedral angle terms because the force constant for a methyl carbon-carbonyl carbon dihedral MM term is zero.

### References

1. M. P. Allen and D. J. Tildesley, *Computer Simulations of Liquids*, Oxford University Press, London, 1987. J. A. McCammon and S. Harvey, *Dynamics of Proteins and Nucleic Acids*, Cambridge University Press, Cambridge, 1987. C. L. Brooks III, M. Karplus, and B. M. Pettitt, *Proteins: A Theoretical Perspective of Dynamics, Structure and Thermodynamics*, Adv. Chem. Phys. Vol. LXXI, J. Wiley and Sons, New York, 1988.
2. U. Burkert and N. L. Allinger, *Molecular Mechanics*, A.C.S. Monographs 177, A.C.S. Washington, D.C., 1982.
3. See, for example: *Methods in Electronic Structure Theory*, H. F. Schaefer III, Ed., Modern Theoretical Chemistry, Vols. 3 and 4, Plenum Press, New York, 1977.
4. See, for example: J. Chandrasekhar, S. F. Smith, and W. L. Jorgensen, *J. Am. Chem. Soc.,* **107**, 154 (1985).
5. M. Karplus, R. N. Porter and R. D. Sharma, *J. Chem. Phys.,* **40**, 2033 (1964).
6. See, for example: a. R. B. Walker and J. C. Light, *Ann. Rev. Phys. Chem.,* **31**, 401 (1980). b. A. E. Dorigo and K. N. Houk, *J. Org. Chem.,* **53**, 1650 (1988).
7. K. Fukui, *J. Phys. Chem.,* **74**, 4161 (1970). K. Fukui, *Acc. Chem. Res.,* **14**, 364 (1981).
8. W. H. Miller, N. C. Handy, and J. E. Adams, *J. Chem. Phys.,* **72**, 99 (1980).
9. See, for example: K. Ishida, K. Morokuma, and A. Komornicki, *J. Chem. Phys.,* **66**, 2153 (1977). S. Kato and K. Morokuma, *J. Chem. Phys.,* **72**, 206 (1980). M. W. Schmidt, M. S. Gordon, and M. Dupuis, *J. Am. Chem. Soc.,* **107**, 2585 (1985). W. H. Miller, *J. Phys. Chem.,* **87**, 3811 (1983). S. K. Gray, W. H. Miller, Y. Yamaguchi, and H. F. Schaefer III, *J. Chem. Phys.,* **73**, 2733 (1980).
10. W. L. Jorgensen and J. K. Buckner, *J. Phys. Chem.,* **90**, 4651 (1986). J. F. Blake and W. L. Jorgensen, *J. Am. Chem. Soc.,* **109**, 3856 (1987). J. D. Madura and W. L. Jorgensen, *J. Am. Chem. Soc.,* **108**, 2517 (1986).

11. J. P. Bergsma, J. R. Reimers, K. R. Wilson, and J. T. Hynes, *J. Chem. Phys.*, **85**, 5625 (1986); J. P. Bergsma, B. J. Gertner, K. R. Wilson, and J. T. Hynes, *J. Chem. Phys.*, **86**, 1356 (1987); B. J. Gertner, J. P. Bergsma, K. R. Wilson, S. Lee, and J. T. Hynes, *J. Chem. Phys.*, **86**, 1377 (1987).

12. I. S. Wang and M. Karplus, *J. Am. Chem. Soc.*, **95**, 8160 (1973).

13. A. Warshel and M. Karplus, *J. Am. Chem. Soc.*, **94**, 5612 (1972). N. L. Allinger and J. T. Sprague, *J. Am. Chem. Soc.*, **95**, 3893 (1973).

14. P. Dauber, M. Brith, E. Huler, and A. Warshel, *Chem. Phys.*, **7**, 108 (1975). A. Warshel and E. Huler, *Chem. Phys.*, **6**, 463 (1974). A. Warshel and M. Karplus, *Chem. Phys. Letts.*, **17**, 7 (1972).

15. A. Warshel and M. Karplus, *Chem. Phys. Letts.*, **32**, 11 (1965).

16. E. Clementi, *Computational Aspects for Large Chemical Systems,* Lecture Notes in Chemistry, Springer, New York, 1980. B. T. Thole and P. Th. van Duijnen, *Theoret. Chim. Acta (Berlin)*, **55**, 307 (1980). B. T. Thole and P. Th. van Duijnen, *Chem. Phys.*, **71**, 211 (1982). J. H. McCreery, R. E. Christoffersen, and G. C. Hall, *J. Am. Chem. Soc.*, **98**, 7191 (1976). O. Tapia and O. Goscinski, *Mol. Phys.*, **29**, 1653 (1975); O. Tapia, F. Sussman, and E. Poulain, *J. Theor. Biol.*, **71**, 49 (1978). O. Tapia and G. Johannin, *J. Chem. Phys.*, **754**, 3624 (1981). U. C. Singh and P. A. Kollman, *J. Comput. Chem.*, **7**, 718 (1986).

17. J. A. C. Rullmann, M. N. Bellido, and P. Th. van Duijnen, *J. Mol. Biol.*, **206**, 101 (1989); G. Alagona, P. Desmeules, C. Ghio, and P. A. Kollman, *J. Am. Chem. Soc.*, **106**, 3623 (1984). S. J. Weiner, U. C. Singh, and P. A. Kollman, *J. Am. Chem. Soc.*, **107**, 2219 (1985). S. J. Weiner, G. L. Seibel, and P. A. Kollman, *Proc. Nat. Acad. Sci.*, **83**, 649 (1986).

18. B. R. Brooks, R. E. Burccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, *J. Comp. Chem.*, **4**, 187 (1983).

19. MOPAC: A general purpose MNDO and MINDO/3 program. J. J. P. Stewart, Quantum Chemistry Program Exchange 455, 1986, Vol. 6, No. 391.

20. M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, and J. J. P. Stewart, *J. Am. Chem. Soc.*, **107**, 3902 (1985).

21. M. J. S. Dewar and W. Thiel, *J. Am. Chem. Soc.*, **99**, 4899 (1977).

22. A. Warshel and R. M. Weiss, *J. Am. Chem. Soc.*, **102**, 6218 (1980).

23. M. J. S. Dewar, *J. Phys. Chem.*, **89**, 2145 (1985).

24. W. J. Hehre, L. Radom, P. v. R. Schleyer, and J. A. Pople, *Ab Initio Molecular Orbital Theory,* Wiley, New York, 1986.

25. See, for example: M. Born and K. Huang, *Dynamical Theory of Crystal Lattices,* Oxford University Press, London, 1954. B. T. Sutcliffe, "Fundamentals of Computational Quantum Mechanics" in *Computational Techniques in Quantum Chemistry,* G. H. F. Diercksen, B. T. Sutcliffe and A. Veillard, Eds., Reidel, Boston, 1975.

26. See, for example: E. Wimmer, H. Krakauer, M. Weinert, and A. J. Freeman, *Phys. Rev.*, **B24**, 864 (1981). M. Weinert, E. Wimmer, and A. J. Freeman, *Phys. Rev.*, **B26**, 4571 (1982).

27. D. Ceperley and B. Alder, *Science*, **231**, 555 (1986).

28. A. Selloni, P. Carnevali, R. Car, and M. Parinello, *Phys. Rev. Letts.*, **59**, 823 (1987).

29. R. Daudel and C. Sandorfy, *Semiempirical Wave Mechanics in Polyatomic Molecules,* Yale University Press, New Haven, CT, 1971. *Semiempirical*

*Methods of Electronic Structure Calculation,* G. A. Segal, Ed., *Modern Theoretical Chemistry,* Vols. 7 and 8, Plenum Press, New York, 1977.

30. M. J. S. Dewar and W. Thiel, *J. Am. Chem. Soc.*, **99**, 4907 (1977).

31. J. S. Binkley, J. A. Pople, and W. J. Hehre, *J. Am. Chem. Soc.*, **102**, 939 (1980).

32. R. Ditchfield, W. J. Hehre, and J. A. Pople, *J. Chem. Phys.*, **54**, 724 (1971).

33. W. C. Herndon and T. P. Radhakrishnan, *Chem. Phys. Letts.*, **148**, 492 (1988).

34. AM1 and MNDO parameters. Full details and references to the parameterization can be found in the MOPAC program and its documentation. Some references are given here: MNDO: H, C, O, N — M. J. S. Dewar and W. Thiel, *J. Am. Chem. Soc.*, **99**, 4907 (1977). Li — W. Thiel from QCPE program #438, Vol. 2, 63 (1982). Be, F — M. J. S. Dewar and H. S. Rzepa, *J. Am. Chem. Soc.*, **100**, 777 (1978). B — M. J. S. Dewar and M. L. McKee, *J. Am. Chem. Soc.*, **99**, 5231 (1977). Al — L. P. Davis et al., *J. Comp. Chem.*, **2**, 433 (1981). Si — M. J. S. Dewar et al., *Organometallics*, **5**, 375 (1986). P — M. J. S. Dewar, M. L. McKee and H. S. Rzepa, *J. Am. Chem. Soc.*, **100**, 3607 (1978). S — M. J. S. Dewar and C. H. Reynolds, *J. Comp. Chem.*, **2**, 140 (1986). Cl — M. J. S. Dewar and H. S. Rzepa, *J. Comp. Chem.*, **4**, 158 (1983). Br — M. J. S. Dewar and E. F. Healy, *J. Comp. Chem.*, **4**, 542 (1983). Sn — M. J. S. Dewar, G. L. Grady, and J. J. P. Stewart, *J. Am. Chem. Soc.*, **106**, 6771 (1984). I — M. J. S. Dewar, E. F. Healy and J. J. P. Stewart, *J. Comp. Chem.*, **5**, 358 (1984). Hg — M. J. S. Dewar et al., *Organometallics*, **4**, 1964 (1985). AM1: H, C, O, N — M. J. S. Dewar et al., *J. Am. Chem. Soc.*, **107**, 3902 (1985).

35. C. C. J. Roothaan, *Rev. Mod. Phys.*, **23**, 69 (1951).

36. C. C. J. Roothaan, *Rev. Mod. Phys.*, **32**, 179 (1960).

37. M. J. S. Dewar and W. Thiel, *Theoret. Chim. Acta.*, **46**, 89 (1977).

38. S. Schroder and W. Thiel, *J. Am. Chem. Soc.*, **107**, 4422 (1985); S. Schroder and W. Thiel, *J. Am. Chem. Soc.*, **108**, 7985 (1986).

39. J. J. Nicolas, K. E. Gubbins, W. B. Street, and D. J. Tildesley, *Mol. Phys.*, **37**, 1429 (1979). D. J. Adams, E. H. Adams, and G. J. Hills, *Mol. Phys.*, **38**, 387 (1979). T. A. Andrea, W. C. Swope, and H. C. Andersen, *J. Chem. Phys.*, **79**, 4576 (1983). P. H. Berens, D. H. J. Mackay, G. M. White and K. R. Wilson, *J. Chem. Phys.*, **79**, 2375 (1983). C. L. Brooks III, M. Karplus and B. M. Pettitt, *J. Chem. Phys.*, **83**, 5897 (1985).

40. The CHARMM 19 polar-hydrogen force field. (W. Reiher and M. Karplus, unpublished work.)

41. The new CHARMM 21 all-atom force field. Y. Arata, M. Bellott, P. Derreumaux, S. Fischer, M. J. Field, D. A. Giammona, H. Guo, K. Kuczera, J. Kuriyan, D. Nguyen, W. E. Reiher III, B. Roux, J. C. Smith, J. E. Straub, N. Summers, H. A. Yu and M. Karplus, to be published.

42. C. C. J. Roothaan, *J. Chem. Phys.*, **19**, 1445 (1951). K. Rudenberg, *J. Chem. Phys.*, **19**, 1459 (1951).

43. M. Born and Th. Von Karman, *Physik. Z.*, **13**, 297 (1912).

44. C. L. Brooks III and M. Karplus, *J. Chem. Phys.*, **79**, 6312 (1983).

45. A. T. Brünger, R. Huber and M. Karplus, *Biochemistry*, **26**, 5153 (1987).

46. P. A. Bash, M. J. Field, and M. Karplus, *J. Am. Chem. Soc.*, **109**, 8192 (1987).

47. T. Alber, D. W. Banner, A. C. Bloomer, G. A. Petsko, D. Phillips, P. S. Rivers, and I. A. Williams,

*Phil. Trans. R. Soc. London,* **B293,** 159 (1981). P. A. Bash, M. J. Field, R. Davenport, and M. Karplus, to be published.

48. R. E. Christoffersen and G. M. Maggiora, *Chem. Phys. Lett.,* **3,** 419 (1969). T. D. Davis, G. M. Maggiora and R. E. Christoffersen, *J. Am. Chem. Soc.,* **96,** 7878 (1974).

49. L. C. Allen, *Annals of the N.Y. Acad. Sci.,* **367,** 383 (1981).

50. J. H. van Lenthe, J. G. C. M. van Duijneveldt-van de Rijdt, and F. B. van Duijneveldt, "Weakly Bound Systems," in *Ab Initio Methods in Quantum Chemistry, Advs. Chem. Phys.,* LXIX, K. P. Lawley, Ed., J. Wiley and Sons, Ltd., 1987.

51. J. Gao, D. S. Garner and W. L. Jorgensen, *J. Am. Chem. Soc.,* **108,** 4784 (1986). J. B. Foresman and C. L. Brooks III, *J. Chem. Phys.,* **87,** 5892 (1987).

52. R. Yamdagni and P. Kebarle, *J. Am. Chem. Soc.,* **93,** 7139 (1971).

53. K. Szalewicz, S. J. Cole, W. Kolos and R. J. Bartlett, *J. Chem. Phys.,* **89,** 3662 (1988). M. J. Frisch, J. E. Del Bene, J. S. Binkley and H. F. Schaefer III, *J. Chem. Phys.,* **84,** 2279 (1986). J. L. Finney, J. E. Quinn, and J. O. Baum in *Water Science Research,* Vol. 1, S. F. Franks, Ed., Cambridge University Press, 1985.

54. T. R. Dyke and J. S. Muenter, *J. Chem. Phys.,* **60,** 2929 (1974).

55. L. A. Curtis, D. J. Fruid, and M. Blander, *J. Chem. Phys.,* **71,** 2703 (1973).

56. M. Mezei and J. J. Dannenburg, *J. Phys. Chem.,* **92,** 5860 (1988).

57. I. Lukovits, A. Karpfen, H. Lischka, and P. Schuster, *Chem. Phys. Letts.,* **63,** 151 (1979).

58. W. L. Jorgensen and J. Gao, *J. Phys. Chem.,* **90,** 2174 (1986).

59. J. D. Madura and W. L. Jorgensen, *J. Am. Chem. Soc.,* **108,** 2517 (1986).

60. J. E. Rice, T. J. Lee and N. C. Handy, *J. Chem. Phys.,* **88,** 7011 (1988).

61. G. T. Fraser, G. W. Gillies, J. Zozom, F. J. Lovas, and R. D. Suenram, *J. Mol. Spectrosc.,* **126,** 200 (1987).

62. F. J. Lovas, R. D. Suenram, G. T. Fraser, C. W. Gillies, and J. Zozom, *J. Chem. Phys.,* **88,** 722 (1987).

63. P. G. Jaisen and W. J. Stevens, *J. Chem. Phys.,* **84,** 3271 (1986).

64. M. J. Field and M. Karplus, manuscript in preparation.

65. M. J. Field and M. Karplus, work in progress.

66. D. V. Belle, I. Couplet, M. Prevost and S. J. Wodak, *J. Mol. Biol.,* **198,** 721 (1987).

67. J. C. Barthelat, P. Durand, and A. Serafini, *Mol. Phys.,* **33,** 159 (1977). Ch. Teichteil, J. P. Malrieu, and J. C. Barthelat, *Mol. Phys.,* **33,** 181 (1977). J. Schnitker and P. J. Rossky, *J. Chem. Phys.,* **86,** 3462 (1987).

68. M. J. Field and M. Karplus, unpublished results.

69. S. Weiss and G. E. Leroi, *J. Chem. Phys.,* **48,** 962 (1968).

70. J. C. Smith and M. Karplus, manuscript in preparation.

71. D. R. Stull, J. Prophet, JANAF Thermochemical Tables, NSRDS-NBS37 (1971).

72. M. J. S. Dewar and K. M. Dieter, *J. Am. Chem. Soc.,* **108,** 8075 (1986).

73. J. A. Pople and D. L. Beveridge, *Approximate Molecular Orbital Theory,* McGraw-Hill, New York, 1970.

74. R. S. Mulliken, *J. Chem. Phys.,* **23,** 1833 (1955).

75. M. J. Field and M. Karplus, work in progress.

76. R. E. Hammer and K. M. Sando, *Chem. Phys. Letts.,* **142,** 48 (1987).

77. W. L. Jorgensen, *J. Am. Chem. Soc.,* **103,** 335 (1981).

78. J. Chandresekhar, D. C. Spellmeyer, and W. L. Jorgensen, *J. Am. Chem. Soc.,* **106,** 903 (1984).

79. The QM/MM facility described in this article will be available in the CHARMM program when it is ready for distribution. Interested parties should contact M. Karplus, Department of Chemistry, Harvard University.