Enzyme Mechanisms with Hybrid Quantum and Molecular Mechanical Potentials. I. Theoretical Considerations

KIRSTEN P. EURENIUS, DAVID C. CHATFIELD, AND BERNARD R. BROOKS*

Division of Computer Research and Technology, Laboratory of Structural Biology, National Institutes of Health, Bethesda, Maryland 20892

MILAN HODOSCEK

National Institute of Chemistry, Hajdrihova 19, 61115 Ljubljana, Solvenia

Received October 16, 1995; revised manuscript received December 13, 1995; accepted December 27, 1995

ABSTRACT

The application of hybrid quantum mechanical and molecular mechanical (QM/MM) potentials to the study of chemical reactions in enzymes is outlined. The discussion is general and addresses the difficulties encountered in an enzyme QM/MM study. First, general criteria for determining whether a particular enzyme is an appropriate candidate for a QM/MM approach are outlined. Methods for obtaining starting structures are detailed. The importance of choosing appropriate levels of ab initio or semiempirical theory is emphasized. Approaches for interfacing the QM and MM regions are briefly discussed, with greater detail given to describing our CHARMM–GAMESS interface. Techniques for partitioning the system into QM and MM regions are explored. Link atom placement, as distant from reacting atoms as possible within the confines of computational efficiency, is examined in some detail. Methods for determining reaction paths are also discussed.

* To whom correspondence should be addressed.

International Journal of Quantum Chemistry, Vol. 60, 1189–1200 (1996)

CCC 0020-7608 / 96 / 061189-12

Introduction

The use of molecular dynamics (MD) simulation to study mechanisms of chemical reactions in enzymes is relatively new. This is largely because low-cost molecular mechanical potentials are incapable of describing covalent bond formation and dissociation. The recent technique of combining a quantum mechanical (QM) treatment of moieties directly involved in catalysis with a molecular mechanical (MM) treatment of the bulk of a chemical system promises to be an important theoretical tool for studying enzyme mechanisms. A number of groups have made progress in applying combined QM/MM methods to chemical reactions in solution [1–11]. However, less progress has been made in applying such methods to reactions in enzymes [12–16]. Reasons for this include: (1) Enzymes are dynamic, complex structures and their catalytic processes may involve a wide range of multistep mechanisms; (2) determining reactant, intermediate, and transition-state structures for enzyme–substrate systems is difficult; (3) the region requiring a quantum mechanical treatment is often quite large, making ab initio calculations at an adequate QM level expensive; (4) interfacing the QM and MM regions appropriately is complicated; (5) methods for treating a QM/MM boundary that bisects a covalent bond have not been well documented.

This article presents a general discussion of the application of QM/MM methods to enzyme mechanisms. We address several issues: What types of enzymatic reactions can be successfully studied with QM/MM methods, and what must be known before a QM/MM study can be considered? How are structures appropriate for reaction initiation obtained? How is the QM/MM Hamiltonian defined, and how are the QM and MM regions partitioned and interfaced? What level of theory is needed for the QM portion of the calculation? How are reaction paths calculated? What can be learned from QM/MM reaction path studies?

Selecting an Enzyme for QM/MM Study

Enzymes are remarkably complex. They are dynamic structures comprising a large number of atoms and may have bound substrates or cofactors. Added to their structural complexity is their mechanistic variety. Mechanisms may involve closed- or open-shell states in a range of multiplicities. Metal cofactors frequently undergo oxidation-state changes. Global conformational changes may occur during the catalytic process. Probing enzyme mechanisms, therefore, is a daunting task. When initiating a QM/MM enzyme study, the limitations of the QM/MM method in describing such complex systems should be considered.

The ability of current QM methods to model the reaction must be determined for each system of interest. Basis sets should be evaluated to determine which level is necessary for reasonable geometries and energies. Ab initio and semiempirical methods have been applied to numerous systems, and their limitations are well documented [17]. Excited states, diradicals, radicals, metals, and any mechanism where mixing of electronic states is possible are often poorly described with a single-configuration self-consistent field (SCF) approach. Alternatives include multiconfigurational SCF (MCSCF), coupled-clusters (CC), or configuration interaction (CI) calculations, any of which can be costly. Density functional theory (DFT) offers an alternative to conventional QM methods. Enzyme reactions which required a high level of theory may prove too expensive for extensive QM/MM modeling.

The mechanisms under study should be described as a well-defined series of simple steps. Steps which are rate-determining and those which will clearly differentiate among possible mechanisms must be explored. Experimental data which illuminate topological and conformational changes during the reaction will suggest structural parameters useful in modeling transition structures and reactive intermediates. In the absence of such data, modeling these structures may become haphazard.

Large-scale conformational changes during a reaction complicate exploration of a mechanism. Reaction path studies are often a series of constrained minimizations along a proposed reaction coordinate and provide a static, not dynamic, picture. Mechanisms which include significant conformational changes may be poorly explored with such an approach. Molecular dynamics simulations of selected structures can provide insight into questions of a dynamic nature.

The QM region size is an important factor in determining the CPU cost and hence the viability of a QM/MM application. One approach to minimizing CPU expenditure is to study large QM

regions at the lowest theoretical level that yields reasonable geometries. Subsequently, single-point energy calculations for these structures or further energy minimization using a smaller QM region may be performed at a higher level of theory.

Enzymes which have been chosen for QM/MM study to date include HIV-1 protease, trypsin, phospholipase A, triose phosphate isomerase, and acetyl cholinesterases [12–16]. These enzymes tend to employ mechanisms involving proton transfers and simple nucleophilic attacks. Reactions such as these are suitable for QM/MM study as they involve simple steps without complex electronic rearrangements which would require expensive levels of QM theory for an adequate description.

In summary, enzymatic catalysis which is suitable for QM/MM study must be well described by a level of QM theory which is not expensive, be easily described as a series of simpler, well-defined steps, not undergo global conformational fluctuations during catalysis, and have a QM region of a tractable size. The choice of theory level, system and QM region size, and simulation technique must be explored and weighed against available resources.

Starting Structures

Quantum mechanical and molecular mechanical enzyme simulations should be initiated with well-characterized starting structures. Studies have illustrated the large variation in QM reaction profiles obtained with different starting conformations of the same proton transfer system [18]. Thus, a poor starting structure may easily lead to misleading QM/MM results.

Well-characterized structures can be obtained from nuclear magnetic resonance (NMR) or x-ray crystallographic coordinates of enzyme–inhibitor complexes. These complexes are particularly useful if the inhibitor resembles a natural substrate. Reactant or intermediate structures may be generated by computationally mutating the inhibitor to a model substrate [12]. When the only available structure is of an uncomplexed enzyme, starting geometries may be generated by computationally docking a model substrate with the resolved enzyme structure [19]. As docking requires a "guessed" substrate position, however, the substrate–enzyme conformation must be thoroughly explored in order to warrant confidence in the resultant starting structure. Restrained minimization or simulated annealing will relax the mutated or docked structure into a low-energy conformation.

Although crystal structures may have well-resolved substrate and enzyme coordinates, waters and counterions may be less well characterized. Positioning counterions is difficult as they have limited mobility during simulation and their initial position has a dominant effect on where they end up [20]. Missing or misplaced water molecules and counterions near the enzyme active site may have a significant affect on QM/MM simulations. Cavities, in particular, suggest the inclusion of more waters. The issue of water and counterion placement must be resolved before a QM/MM enzyme starting structure is determined. Where uncertainty remains, parallel simulations may be appropriate.

A single, static structure of the enzyme–substrate complex, whether obtained directly from experiment or by computational mutation or docking, is not necessarily an appropriate starting point for a QM/MM reaction path study. It may be useful to perform classical MD simulations or restrained energy minimizations to sample other accessible conformation. Those conformations with interatomic distances consistent with the initiation of the reaction mechanisms that also have reasonable energies may provide useful starting coordinates for QM/MM studies. The use of unrestrained dynamics was employed to generated starting structures for a QM/MM study of HIV-1 protease [21]. The criteria for selecting productive conformations from a MD trajectory involved examining linear combinations of selected distances that are important for the step of the proposed mechanism, and selecting those frames with optimal values. For the first step of the proposed HIV protease mechanism, this metric was the sum of the nucleophile (either a water molecule or an aspartate side chain) and peptide carbonyl carbon distance and the distance between the carbonyl oxygen and nearby polar hydrogens.

The QM/MM Method

The effective QM/MM Hamiltonian is the sum of terms representing the QM region, the MM region, and the interaction between them [3(b)].

Further terms describing boundary effects and restraints can be included and are well described in classical simulation literature. Table I details the QM/MM atom classification scheme:

$$\hat{H}_{\text{eff}} = \hat{H}_{\text{QM}} + \hat{H}_{\text{MM}} + \hat{H}_{\text{QM/MM}} + \hat{H}_{\text{Boundary}} + \hat{H}_{\text{Restraints}}. \quad (1)$$

The $\hat{H}_{\text{QM}}$ term describes all the interactions within the quantum region. Any ab initio or DFT package can be used to generate the appropriate electronic description. $\hat{H}_{\text{MM}}$ details the classical interactions and can be generated using any standard molecular simulation or molecular mechanical program. In our approach, the molecular simulation program CHARMM [22] was merged with the quantum package GAMESS [23], allowing flexibility in the choice of ab initio and classical treatments.

The $\hat{H}_{\text{QM/MM}}$ Hamiltonian may be described as the sum of three terms:

$$\hat{H}_{\text{QM/MM}} = \hat{H}_{\text{QM/MMbonded}} + \hat{H}_{\text{QM/MMvdw}} + \hat{H}_{\text{QM/MMelec}}. \quad (2)$$

TABLE I
Definition of atoms in QM / MM calculations.

Definitions	Abbreviation
MM Atom Any atom within the classical region.	MMA
QM Atom Any atom within the quantum region.	QMA
Link Atom An atom added to the QM region which serves to cap the QM electron density when a QM/MM boundary bisects a covalent QMA-MMA bond. The link atom should not distort the MM region, nor introduce large interactions in the QM region.	LA
Link Atom Host The MM atom of the bisected QMA-MMA bond.	LAH
Link Atom Bond Partner The QM atom of the bisected QMA-MMA bond.	LABP

Classical computation of selected bonded QM/MM terms is required to maintain reasonable geometries. Several researchers treat all bonded terms between at least one MM and at least one QM atom classically [2, 3,15]. Our approach is a refinement of this treatment and avoids duplicating interactions computed quantum mechanically. We have devised a series of rules detailing which QM/MM bonded terms should be treated classically:

1. All bond terms involving one QM atom and one MM atom are computed classically to maintain bonds across the interface.
2. Angles and dihedral angles with one or two central MM atoms are included in the MM treatment. Those with QM central atoms are assumed to be treated by $\hat{H}_{\text{QM}}$.
3. Improper dihedrals with atoms i and l in the QM region are treated quantum mechanically while those with outer MM atoms need to be computed classically [24].

Of the two nonbond QM/MM terms, only the van der Waals interactions are computed classically. These interactions are handled as if both atoms in the atom pair were MM, allowing the QM region to respond to size effects of the classical atoms. For this purpose, classical van der Waals parameters are assigned to the atoms in the QM region. Van der Waals interactions between two QM atoms are, of course, treated within the QM region.

The QM/MM electrostatic interactions are described by treating the MM atoms as point charges in an electrostatic field which interacts with the QM atoms through the one-electron integrals. Other approaches include treating the MM atom charges as Gaussians instead of point charges, allowing a more flexible interface between polarizable and non polarizable models [25]. The selection of MM atoms to be included in the electrostatic field is discussed below.

An important issue when selecting a QM/MM method is how the QM electron density is terminated at a QM/MM boundary. If the QM region comprises an entire molecule, this issue is avoided. When the boundary bisects a covalent bond, however, the electron density along the covalent bond is terminated abruptly at the end of the QM region. Several groups have used "link" atoms to

gracefully cap the electron density [2, 3, 15]. A link atom is inserted along the bond between the QM atom, to be called the link atom bond partner (LABP), and the first MM atom, referred to here as the link atom host (LAH). Link atoms may be any atom type in the QM region, and are often hydrogen atoms. Ostlund's approach, however, uses an "effective halogen" atom to cap the QM region which coincides exactly with the LAH [26]. Link atoms are not intrinsic components of the system being modeled and thus should not distort the MM region. In our approach, a classical angle term, link atom–LABP–LAH, is computed to maintain the link atom position along the LABP–LAH bond vector. This parameter is assigned the force constant of the comparable H–LABP–LAH angle and an optimal angle of 0°. No other classical energy terms involving link atoms are computed. The link atom does, however, feel the full effect of the MM electrostatic field along with the rest of the QM region. In this way the link atom is influenced by, and can respond to, the MM atoms without introducing new interactions in the classical region.

While link atoms do not disrupt the MM region, their effect on the QM region can be complex. Link atoms introduce van der Waals and electrostatic interactions with the other QM atoms and with each other. It may be possible to estimate the size of some of the link atom–QM atom or link atom–link atom interactions through classical calculations of the QM region with the link atom treated as a standard hydrogen atom type. Model studies are necessary to explore the placement of link atoms, whose bonds are suitable for interfacing QM and MM regions, and the interactions introduced by the partitioning.

Other approaches do not require the addition of link atoms. Warshel and Levitt used a single hybrid orbital on the MM atom of the QM–MM bond. They include all valence electrons with hybrid atomic orbitals in their semiempirical method [1a]. Rivail replaces the atomic orbitals of the atoms in the QM–MM bond with four parameterized hybrid orbitals defined on the QM atom and one on the MM atom [27]. A strictly localized bond orbital (SLBO) is constructed from hybrid orbitals between the QM atom and the MM atom. The three remaining QM hybrid orbitals, which are orthogonal to the SLBO, contribute to molecular orbitals describing the electrons of the entire QM region. Good agreement with QM results was seen in a proton transfer study. Their current implementation is at the semiempirical level, though extensions to the ab initio level of theory are underway.

If link atoms are used, the LAH should be excluded from the electrostatic field. Inclusion may lead to collapse of the link atom onto the associated LAH charge. However, exclusion of the LAH may result in an unbalanced electrostatic field. Studies have show that removing an α-carbon LAH while including the charges of its hydrogens grossly underestimates proton affinities (Fig. 1). Reasonable results were obtained when the excluded atom list was expanded to include the LAH and its hydrogens, an overall neutral group. Expanding the excluded atom list in essence shifts the charge of the LAH outward onto its neighboring atoms. Removal of the LAH and other atoms, however, may lead to the neglect of critical electrostatic interactions.

Table II summarizes the rules developed for our QM/MM method and outlines the steps we take when partitioning a covalent bond between QM and MM regions. Several differences exist between these guidelines and those proposed by other researchers. Kollman, Bash, and Payling treat QM/MM bonded energy terms classically when they involve at least one QM and at least one MM atom [2, 3, 15]. This approach allows the potential duplication of classical interactions treated quantum mechanically.

Nonbond van der Waals terms between QM and MM atoms are treated classically by many groups [2, 3, 15]. Morokuma, however, neglects any QM/MM van der Waals interactions between the LAH and the QM region in an effort to avoid double counting this interaction [8]. Most research groups treat the QM/MM electrostatics by modeling the MM atoms as point charges in an electric field [2, 3, 15]. There is greater variance in the selection of which MM atoms are included. Payling excludes the LAH from the electrostatic field [15]. Kollman, however, excludes any MM atom within two bonds of any QM atoms [2]. Issues of balanced electrostatic field and neglected crucial electrostatic interactions are not addressed. Model studies should be performed to evaluate the effects of excluding particular atoms from the electrostatic field. Morokuma, on the other hand, calculates electrostatics between QM atoms and MM atoms (excluding LAH) classically [8]. This does not allow for any polarization of the QM region.

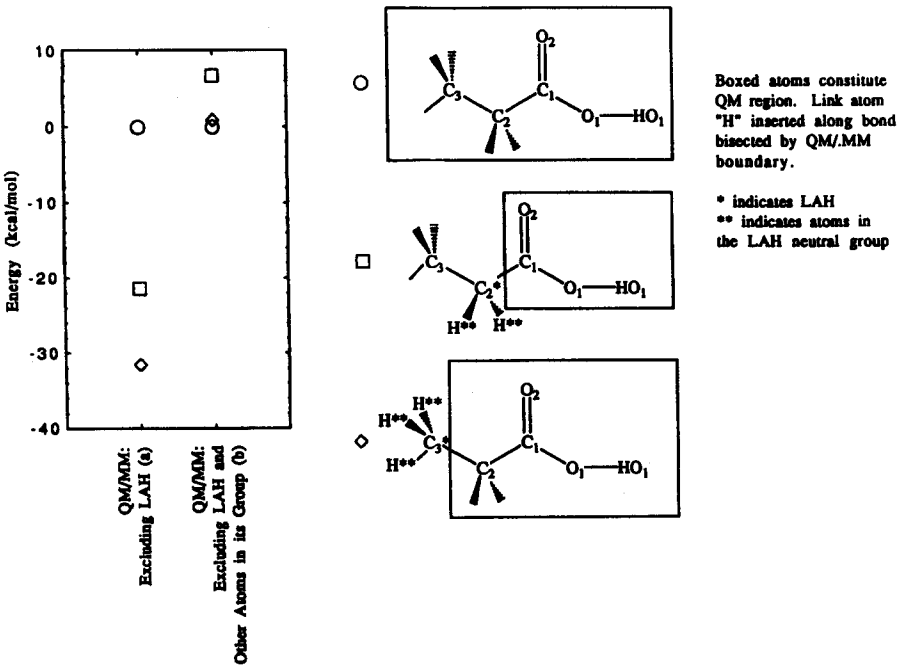

FIGURE 1. Relative QM/MM propionic acid proton affinities calculated with the RHF/STO-3G basis set. Both propionic acid and propionate anion were subjected to rigorous energy minimization. (a) QM/MM calculation excludes LAH from the MM electrostatic field, resulting in an overall partially positive electrostatic field. (b) QM/MM calculation excludes LAH and the hydrogens attached to the LAH from the electrostatic field, resulting in overall neutral group exclusion.

TABLE II
Classically computed QM / MM terms.

Interaction	# of Atoms	Classical Computed When:
Bonded Terms:		
• Bond	2	QMA-MMA
• Angle	3	X-MMA-X
• Dihedral	4	X-MMA-X-X
• Improper Dihedral	4	MMA-X-X-X
Nonbond Terms		
• Van der Waals	2	QMA-MMA
• Electrostatics	2	NEVER MMA atoms treated as point charges and resultant electrostatic field interacts with the QMAs through their one-electron integrals. *LAH is excluded from electrostatic field- other atoms may be excluded as needed.*
Link Atom Terms		
• Angle	3	LA-LABP-LAH Term added to maintain link atom position along the LABP-LAH bond vector.
• Van der Waals	2	NEVER Link atom should be invisible to MM region.
• Electrostatics	2	NEVER Link atom feels the MMA electrostatic field through its one-electron integrals.

A further difference between Morokuma's approach and those of most of other QM/MM groups lies in the treatment of the link atom. In Morokuma's method, the LABP–link atom and LABP–LAH bond distances are both fixed. Other approaches tend to impose far less constraint on the link atom position.

In the CHARMM–GAMESS approach, minimization is performed by the MM side of the program. The QM energies and forces are calculated in the presence of the electrostatic field for each minimization step. The forces exerted by and on the QM atoms are fed back to CHARMM. A classical energy and force calculation is performed, and the QM and MM forces are combined. The combined set of forces determines the structural change using standard techniques (MD simulations, minimization, etc.).

Partitioning into QM and MM Regions

The size of the QM region largely determines the cost of the simulation calculation for a particular level of ab initio theory. Choosing the smallest QM region that yields reasonable results is a prerequisite for efficient QM/MM calculations. The following guidelines may aid in selecting a QM region of appropriate size.

Any atom which undergoes a hybridization change during the course of the reaction should be included in the QM region. Unfortunately, this may require the inclusion of an entire aromatic or conjugated system, potentially an expensive partitioning scheme. If it is necessary to divide across conjugation, partitioning effects should be well characterized.

Proton affinities may be highly dependent on which atoms are assigned to the QM region, especially if neighboring groups are strongly electron-donating or electron withdrawing. The selection of atoms excluded from the electrostatic field is of concern. Our proton affinity and transfer studies have shown a strong dependence upon which MM atoms are included in the electrostatic field which interacts with the QM region (Fig. 1). An unbalanced electrostatic field, generated by excluding a partial positive or negative charge, may grossly over- or underestimate proton affinities. Our remedy of choice for an unbalanced electrostatic field is to extend the excluded atom list to include any atoms necessary for the exclusion of an overall neutral charge. The results in Figure 1 show the improved proton affinities that this method generates. Excluding a larger group of atoms, however, may result in further neglect of crucial electrostatic interactions, thus biasing the energy calculations.

An additional concern is the choice of link atom sites. As noted above, link atoms may influence electrostatic and van der Waals interactions in the QM region, distorting the QM/MM results. These interactions combined with the neglected LAH electrostatic interactions may significantly effect the QM/MM energies (Fig. 2). To avoid distortion from link atoms, it may be necessary to select a link atom site further from the region where changes are taking place.

Multiple link atom placement has been explored in several small molecule systems [19]. Though most partitioning schemes perform well, significant problems arise when two link atoms are positioned 1,4 to each other. Butane rotation about the central C–C bond, modeled with the two outer carbons treated as QM methanes and the inner two carbons MM, yields energies significantly larger than experimental values (Fig. 3). Positioning link atoms more closely, 1,3 to each other, or further apart, 1,5 and up, in 2,2-dimethyl butane rotation yields no major difficulties, beyond the neglect of LAH electrostatic interactions. A further discussion of link atom placement may be found in our related work [28].

Reaction Path Determination

A number of reaction path methods based on the Hessian of potential energy have been developed for small molecules, but these are prohibitively expensive for systems of more than a few tens of atoms or for certain types of QM calculations [20]. Several methods developed for large systems, including the travel method [29], may be useful for QM/MM treatment of enzymes. However, the number of energy evaluations required by these methods may be impractical. Even at a low level of theory, QM/MM energy evaluations are much more expensive than the MM energy evaluations, for which these methods were developed.

A method useful for initial mapping of the energy surface involves the use of general re-

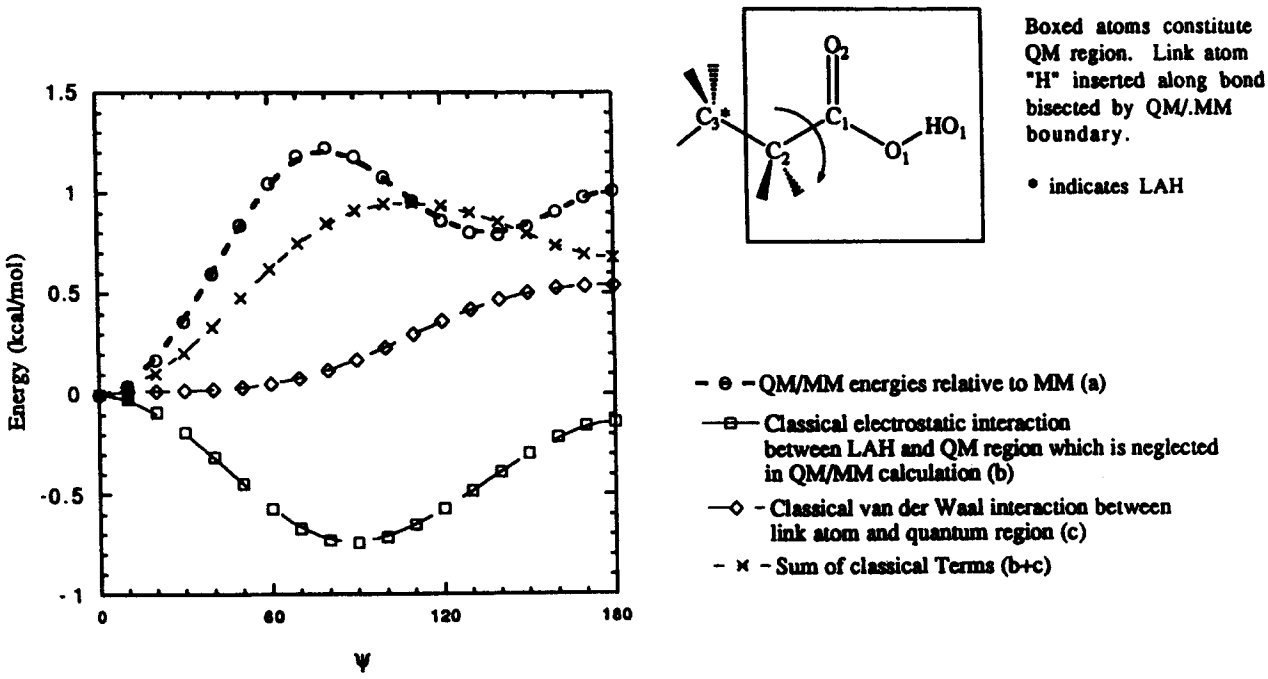

FIGURE 2. QM/MM rotational barrier about dihedral angle O1–C1–C2–C3 for propanoic acid calculated at the RHF/STO-3G level. Complete energy minimization was performed at each constrained value of the dihedral angle. Subsequent classical calculations were performed on frozen QM/MM geometries. (a) QM/MM rotational profile relative to MM energies. (b) Classical calculation of neglected LAH electrostatic interactions with the QM region. (c) Classical calculation of link atom electrostatic interactions with the QM region. (d) Sum of (b) and (c).

straints in which a reaction coordinate is followed by placing large restraints on degrees of freedom most critical to the proposed mechanism. For example, if the reaction is a proton transfer from atom A to atom B, a harmonic restraint of the form $\frac{1}{2}k(R_{H-B} - R_{H-A} + d)^2$ is applied, where k is a force constant, R_{H-B} is the distance from the proton to atom B, and d is a distance parameter which is systematically modified. Energy minimization with a suitably large k brings R_{H-A} very close to d_0. A reaction path is mapped out by decreasing d in small increments from its initial value and energy minimizing at each step until R_{H-B} is roughly 1.0 Å. The larger the value of k, the greater the weight given to change in the degree of freedom R_{H-B} during minimization. When the relaxation of other degrees of freedom is expected to be small compared to the change in R_{H-B}, this is an efficient way to map out a reaction coordinate.

Multiple reaction steps can be treated simultaneously by generalizing this approach. Such a method, named RESD (REStrained Distances), has been incorporated into CHARMM, and we describe it here. To the energy is added a restraint term of the form

$$E_{\text{RESD}} = \sum_{j}^{N} \frac{1}{2} k_J \left(\sum_{I}^{M_J} a_{IJ} |R_{1IJ} - R_{2IJ}| - d_J \right)^2 \quad (3)$$

where $R_{1IJ} - R_{2IJ}$ is an interatomic distance and a_{IJ} is its weighting factor which may be negative. For example, Figure 4 depicts model first steps for two possible mechanisms consisting of a nucleophilic attack of atom O on atom C and a proton transfer to O_C. To calculate a reaction path with RESD, one might set $a_1 = a_2 = 1$ and initially set d to 5.0 Å, the sum of the initial $C - O$ and $O_C - H$ distances. The reaction path is calculated by alternately decrementing d and energy minimizing. The barrier maximum locates the transition state. The order in which R_{C-O} and R_{O_C-H} change is likely to indicate the order of the reaction steps; a simultaneous change indicates a concerted mecha-

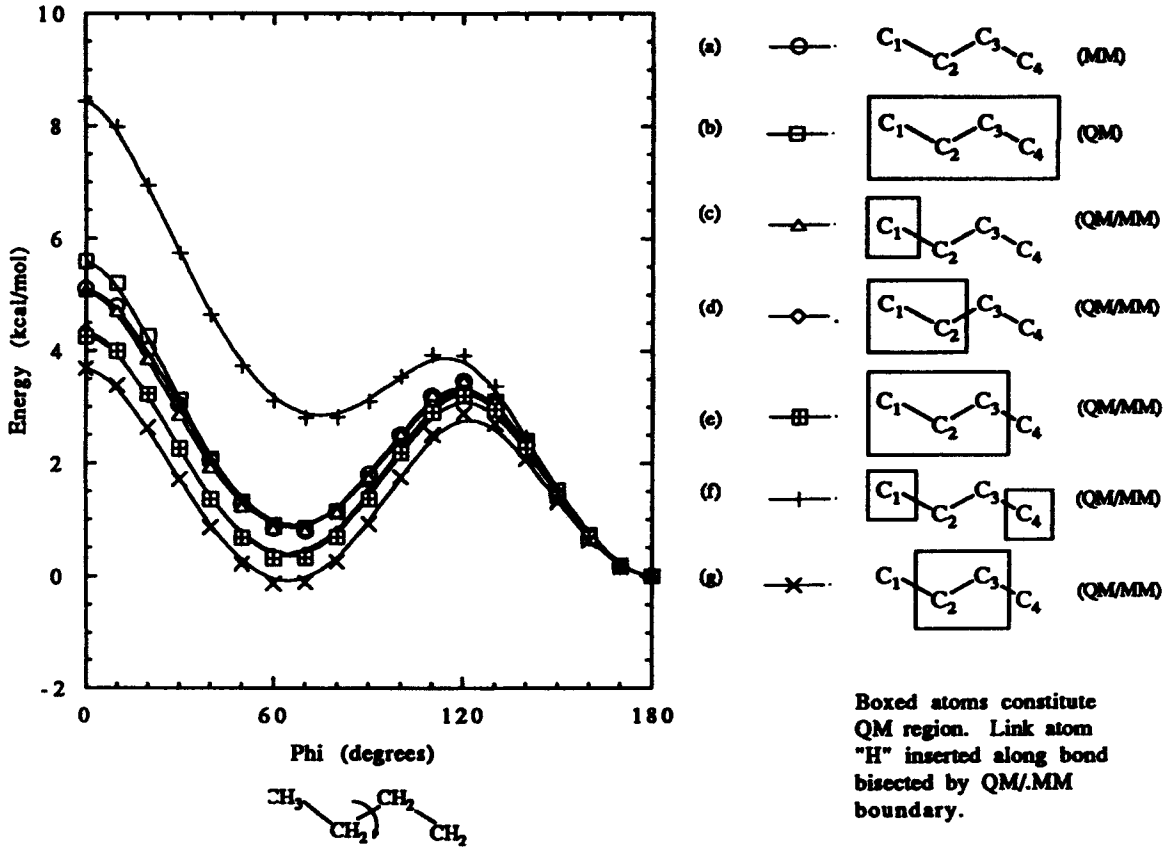

FIGURE 3. Butane rotational barriers about the C1–C2–C3–C4 dihedral angle. Complete energy minization was performed at each constrained value of the dihedral angle. QM energies are calculated at the RHF/STO-3G level. Link atoms are inserted along bonds bisected by a QM/MM boundary.

nism. A model reaction barrier and plot of R_{C-O} vs. R_{O_C-H} are shown in Figures 5 and 6. More complicated mechanisms can be modeled with RESD by including additional J or I terms in expression (3). It is worth noting that this approach, and the other reaction path methods discussed in this section, make no estimate of entropic contributions and merely determine a path based upon the potential energy.

Promise and Problems of QM / MM Studies of Enzymes

Applying QM/MM methods requires judicious screening of possible systems. The enzyme active site should not require a QM region of intractable size nor large-scale conformational changes to model the mechanism. A solution or crystal structure of the enzyme complexed with a related inhibitor, with well-resolved water molecules and counterions, can provide suitable starting structures. Link atom placement is a delicate issue balancing computational efficiency and adequate electronic representation. Further consideration must be given to the interface between the QM and MM regions. Our QM/MM results show a strong dependence upon which MM atoms comprise the electrostatic field which interacts with the QM region.

Hybrid QM/MM calculations should be useful in elucidating mechanisms of enzymatic catalysis. Such calculations can help to distinguish the most likely of several hypothetical mechanisms. How-

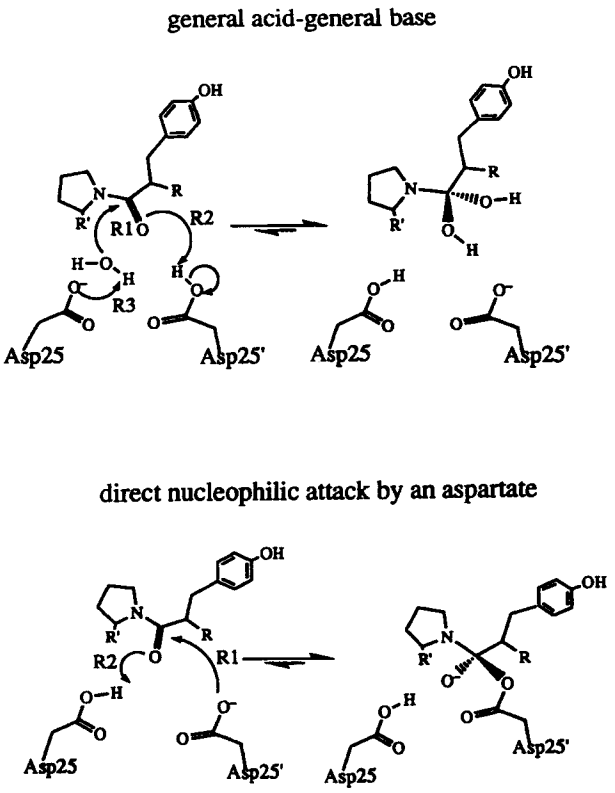

FIGURE 4. The first steps of two possible reaction mechanisms for proteolytic cleavage by HIVPR: (a) general acid–general base with neutral intermediate and (b) general acid–general base with zwitterion intermediate.

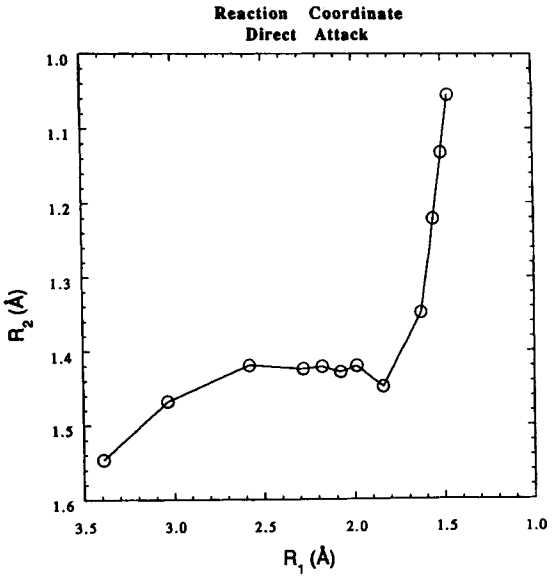

FIGURE 5. Potential energy surface for the first step for direct attack mechanism using the RESD approach. The distance R_1 is the distance between the asp side chain (acting as a nucleophile) and the carbonyl carbon. Distance R_2 is that between the carbonyl oxygen and a nearby polar hydrogen. The nucleophile attacks the carbonyl carbon before the proton is transferred, indicating a stepwise mechanism.

ever, several factors can influence QM/MM energies and must be considered carefully. The computational method and basis set will have paramount influence on the energies calculated. The QM/MM partitioning scheme will influence the energetics. Furthermore, the interactions of link atoms with each other and with the other atoms of the QM region can distort the energies being explored. Each of these issues should be addressed by the QM/MM study.

Determining ground and transition-state structures is another of simulation's goals. One difficulty lies in how a starting structure is generated. A poor starting structure may drastically alter subsequent QM/MM geometries. Additionally, while several methods can approximate transition structures, it remains somewhat difficult to locate a true QM/MM transition state. Thus, while QM/MM calculations may provide insight into reaction path structures, exact geometric details may remain elusive.

A particularly promising application of the QM/MM method is the systematic exploration of site-specific mutations. The wealth of mechanistic information generated by such studies would provide a framwork to assess the ability of the QM/MM method to reproduce systematic experimental evidence. Furthermore, these studies would provide a chance to explore the mechanistic implications of mutations. A further use of the QM/MM approach is to combine the use of hybrid potentials with the calculation of kinetic isotope effects or NMR shifts. A marriage of the QM/MM method with other methods which elucidate structural details may provide the strongest support for use of a QM/MM approach in enzymatic mechanistic studies.

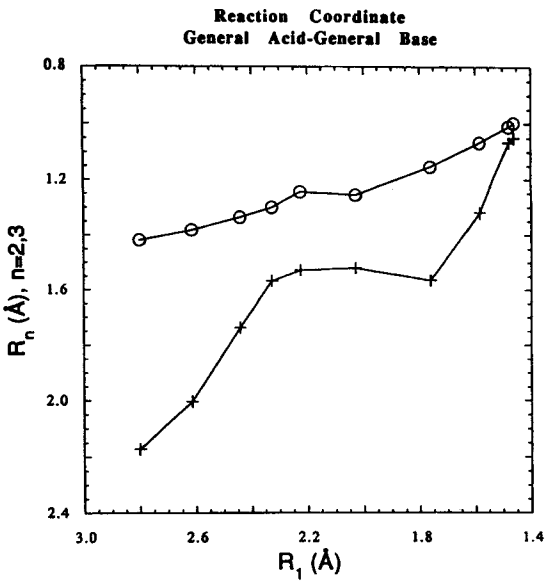

FIGURE 6. Potential energy surface for the first step of the general acid–general base mechanism using the RESD approach. The distance R_1 is the distance between the water oxygen (acting as a nucleophile) and the carbonyl carbon. Distance R_2 is that between the carbonyl oxygen and a nearby polar hydrogen, while R_3 is the distance between the asp side chain oxygen and the closest water hydrogen. There is a nearly simultaneous proton transfer along r_2 (+ symbols) as well as along r_3 (∘ symbols) with nucleophilic attack.

References

1. (a) A. Warshel and M. Levitt, J. Mol. Biol., **103**, 227 (1976); (b) J.-K. Hwang, G. King, S. Creighton and A. Warshel, J. Am. Chem. Soc. **110**, 5297 (1988); (c) A. Warshel, *Computer Modeling of Chemical Reactions in Enzymes and Solutions* (Wiley, New York, 1991); (d) A. Warshel and V. Luzhkov, J. Comput. Chem. **13**, 199 (1992).
2. U. C. Singh and P. A. Kollman J. Comput. Chem. **7**, 718 (1986).
3. (a) P. A. Bash, M. J. Field, and M. Karplus, J. Am. Chem. Soc. **109**, 8092 (1987); (b) M. J. Field, P. A. Bash, and M. Karplus, J. Comput. Chem. **11**, 700 (1990).
4. (a) J. Gao, J. Phys. Chem. **96**, 537 (1992); (b) J. Gao and X. Xia, Science **258**, 631 (1992); (c) J. Gao and J. J. Pavelites, J. Am. Chem. Soc. **114**, 1912 (1992); (d) J. Gao, J. Phys. Chem. **96**, 6432 (1992); (e) J. Gao, J. Am. Chem. Soc. **116**, 1563 (1994); (f) J. Gao, J. Am. Chem. Soc. **116**, 9324 (1994); (g) J. Gao and M. Freindorf, J. Comput. Chem., to appear; (h) J. Gao, in *Methods and Applications of Combined Quantum Mechanical and Molecular Mechanical Potentials*, D. B. Boyd, Ed., Reviews in Computational Chemistry, Vol. 7 (VCH, New York, 1996).
5. (a) R. V. Stanton, D. S. Hartsough, and K. M. Merz, J. Phys. Chem. **97**, 11868 (1993); (b) R. V. Stanton, D. S. Hartsough, and K. M. Merz, J. Comput. Chem. **16**, 113 (1995); (c) D. S. Hartsough and K. M. Merz, J. Phys. Chem. **99**, 384 (1995); (d) R. V. Stanton, L. R. Little, and K. M. Merz, J. Phys. Chem. **99**, 483 (1995); (e) R. V. Stanton, S. L. Dixon, and K. M. Merz, J. Phys. Chem. **99**, 10701 (1995); (f) D. S. Hartsough and K. M. Merz, J. Phys. Chem. **99**, 11266 (1995).
6. H. Liu and Y. Shi J. Comput. Chem. **15**, 1311 (1994).
7. H. Lie, F. Muller-Plathe, and W. F. van Gunsteren, J. Chem. Phys. **102**, 1722 (1005).
8. F. Maseras and K. Morokuma, J. Comput. Chem. **16**, 1170 (1995).
9. M. A. Thompson, E. D. Glendening, and D. Feller, J. Phys. Chem. **98**, 10465 (1994).
10. V. V. Vasilyev, A. A. Bliznyuk, and A. A. Voityuk Int. J. Quantum Chem. **44**, 897 (1992).
11. D. Wei and D. R. Salahub, Chem. Phys. Lett. **224**, 291 (1994).
12. (a) P. A. Bash, M. J. Field, R. C. Davenport, G. A. Petsko, D. Ringe, and M. Karplus, Biochemistry **30**, 5826 (1991); (b) M. Karplus, J. D. Evanseck, D. Joseph, P. A. Bash, and M. J. Field, Faraday Discuss. **93**, 239 (1992).
13. D. S. Hartsough and K. M. Merz, J. Phys. Chem. **99**, 11266 (1995).
14. M. Levitt and A. Warshel, J. Mol. Bio. **103**, 227 (1976).
15. (a) B. Waszkowycz, I. H. Hillier, N. Gensmantel, and D. W. Payling, J. Chem. Soc. Perkin Trans. **2**, 225 (1991); (b) B. Waszkowycz, I. H. Hillier, N. Gensmantel, and D. W. Payling, J. Chem. Soc. Perkin Trans. **2**, 1819 (1991); (c) B. Waszkowycz, I. H. Hillier, N. Gensmantel, D. W. Payling, and J. Chem. Soc. Perkin Trans. **2**, 2025 (1991).
16. V. V. Vasilyev, J. Mol. Struct. (Theochem) **304**, 129 (1994).
17. (a) W. J. Henre, L. Radom, P. v. R. Schleyer, and J. A. Pople, *Ab Initio Molecular Orbital Theory* (Wiley, New York, 1986); (b) I. N. Levine, *Quantum Chemistry*, 4th ed. (Prentice-Hall, Englewood Cliffs, NJ, 1991).
18. (a) S. Scheiner, J. Am. Chem. Soc. **103**, 315 (1981); (b) S. Scheiner and L. B. Harding, J. Am. Chem. Soc. **103**, 2169 (1981); (c) S. Scheiner, J. Phys. Chem. **77**, 4039 (1982).
19. Yong S. Lee, personal communication.
20. W. F. van Gunsteren, H. J. C. Berendsen, R. G. Geurtsen, and H. R. J. Zwinderman, Ann. New York Acad. Sci. **482**, 287 (1986).
21. D. C. Chatfield and B. R. Brooks, J. Am. Chem. Soc. **117**, 5561 (1995).
22. B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, J. Comp. Chem. **4**, 187 (1983).
23. M. W. Schmidt, K. K. Baldridge, J. A. Boatz, J. H. Jensen, S. Koseki, M. S. Gordon, K. A. Nguyen, T. L. Windus, and S. T. Elbert, QCPE Bull. **10**, 52 (1990).
24. In CHARMM, improper dihedrals are defined for four atoms, *i, j, k, l*, three of which are bonded to the central atom *i*. The improper dihedral is the angle between the planes defined by atoms *j-i-l* and atoms *k-i-l*. The dihedral is specified *i-j-k-l*, where *i* and *l* are termed the "outer" atoms. An example is a carbonyl carbon with two substitu-

tions. The *i* atom is the carbonyl carbon, the *l* atom is the carbonyl oxygen, and the realtionship between the two substitutions is defined by the improper dihedral.

25. I. Logunov and K. Schulten, in *Proceedings for the Ecole de Physique des Houches on Quantum Mechanical Simulation Methods for Studying Biological Systems*, (Springer-Verlag, Paris, 1995).
26. HyperChem, Hypercube Inc., Cambridge, ON.
27. V. Thery, D. Rinaldi, J.-L. Rivail, B. Maigret, and G. C. Ferenczy, J. Comput. Chem. **15**, 269 (1994).
28. K. P. Eurenius and B. R. Brooks, unpublished results.
29. S. Fischer and M. Karplus, Chem. Phys. Lett. **194**, 252 (1992).