Simulating Enzyme Reactions: Challenges and Perspectives

MARTIN J. FIELD
Laboratoire de Dynamique Moléculaire, Institut de Biologie Structurale, Jean-Pierre Ebel, 41 Rue Jules Horowitz, F-38027 Grenoble Cedex 01, France

Received 5 February 2001; Accepted 20 June 2001

Abstract: Elucidating how enzymes enhance the rates of the reactions that they catalyze is a major goal of contemporary biochemistry, and it is an area in which computational and theoretical techniques can make a major contribution. This article outlines some of the processes that need to be investigated if enzyme catalysis is to be understood, reviews the current state-of-the-art in enzyme simulation work, and highlights challenges for the future.

© 2002 John Wiley & Sons, Inc. J Comput Chem 23: 48–58, 2002

Key words: enzymes; reactions; catalysis; molecular simulations

Introduction

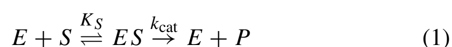
Enzymes are remarkable catalysts[1] that can provide, in those cases where comparisons are possible, rate enhancements of many orders of magnitude over the equivalent uncatalyzed reactions. Despite having been the object of intense experimental and theoretical scrutiny for over a century, Blow argued in a recent review article that there is still no quantitative understanding of how enzymes work. In particular, he blames the lack of progress upon the difficulties of applying "rigorous quantum mechanical and thermodynamic analyses" to the simulation of "active sites in huge molecules surrounded by water."[2] Is this view justified and, if so, what is to be done?

The aim of this article is to address these questions by reviewing the current state of the enzyme simulation field. The outline of the article is as follows. The next section recapitulates the basic notions of enzyme catalysis, then the following section presents an overview of techniques that are designed to simulate enzyme reactions; the Recent Simulation Studies section examines examples of recent simulation studies and the last section concludes.

Enzyme Catalysis

Macroscopic Equations

A fundamental feature of enzyme catalysis is that to effect the reaction that it catalyzes the enzyme must bind the substrate. This is enshrined in the Michaelis–Menten mechanism, which is the basic model used to describe much of enzyme catalysis.[1] It postulates the following scheme:



$$E + S \overset{K_S}{\rightleftharpoons} ES \overset{k_{cat}}{\rightarrow} E + P \tag{1}$$

where E denotes the enzyme and S and P are the substrate and the product of the reaction, respectively. There are two steps in the mechanism. First, there is a binding of the substrate to the enzyme with an equilibrium constant of K_S, which is defined as:

$$K_S = \frac{[E][S]}{[ES]} \tag{2}$$

and, second, there is the reaction of the substrate bound to the enzyme with a rate constant of k_{cat}. With this model, the initial rate of formation of products v is k_{cat} $[ES]$, which can be rewritten as:

$$v = \frac{k_{cat}}{K_S}[E][S] \tag{3}$$

$$= \frac{k_{cat}[E]_0[S]}{K_S + [S]} \tag{4}$$

using eq. (2) and defining the total enzyme concentration $[E]_0$ as $[E] + [ES]$. Equation (4) is appropriate for the simple scheme outlined in eq. (1). An equation of the same form is often valid, however, for more complex mechanisms but with a different definition for the equilibrium constant K_S. In the general case the notation K_M is used instead of K_S, a convention that shall be adhered to in the subsequent discussion.

It has been common to help interpret enzyme catalysis using the framework of transition state theory (TST). TST assumes the existence of a species, the transition state (TS), that is in thermodynamic equilibrium with the reactant species and that decays at a constant rate to products. The advantage of TST is that it introduces the notion of an equilibrium, and so permits the stabilization of the various species along the reaction pathway to be described

Correspondence to: M. J. Field; e-mail: mjfield@ibs.fr

Contract/grant sponsors: Institut de Biologie Structural–Jean-Pierre Ebel, the Commissariat à l'Energie Atomique, and the Centre National de la Recherche Scientifique

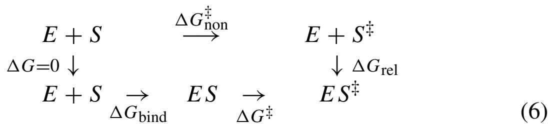
© 2002 John Wiley & Sons, Inc.

with free energies. Thus, for the scheme of eq. (1), one obtains:

$$E + S \overset{\Delta G_{\text{bind}}}{\rightarrow} ES \overset{\Delta G^{\ddagger}}{\rightarrow} ES^{\ddagger} \rightarrow E + P \quad (5)$$

where $S^{\ddagger}$ denotes the TS and the ΔGs are the free-energy differences associated with the various steps of the reaction. The total free-energy barrier is made up of two contributions—one due to binding, and one due to the chemical step that occurs when the substrate is bound to the enzyme.

In many cases, it is of interest to be able to compare the rate of the enzyme reaction to the same reaction in solution. This can be conveniently done with a simple cycle first introduced by Kurz,[3] and which can be written in terms of free energies as:



$$\begin{array}{ccccc} E + S & & \overset{\Delta G^{\ddagger}_{\text{non}}}{\longrightarrow} & & E + S^{\ddagger} \\ \Delta G{=}0 \downarrow & & & & \downarrow \Delta G_{\text{rel}} \\ E + S & \underset{\Delta G_{\text{bind}}}{\rightarrow} & ES & \underset{\Delta G^{\ddagger}}{\rightarrow} & ES^{\ddagger} \end{array} \quad (6)$$

The rate of reaction in the enzyme will be faster than the noncatalyzed reaction if $\Delta G_{\text{bind}} + \Delta G^{\ddagger}$ is less than ΔG_{non}. This implies that ΔG_{rel} is negative or that the TS $S^{\ddagger}$ is more stable when it is bound to the enzyme than when it is in solution. Although the rate of the enzyme reaction ($\equiv k_{\text{cat}}/K_M$) is maximized by having the enzyme bind the TS as strongly as possible, this does not imply that the enzyme must also bind the substrate strongly. Indeed, it can be shown that it will in general be preferable if the substrate is bound weakly (high K_M).[1]

Microscopic Mechanisms

A multitude of different mechanisms have been proposed to explain enzyme catalysis at a microscopic level, and while there may be broad agreement over the general factors responsible there is controversy over their relative importance. This diversity of views is apparent from some recent reviews (refs. 4 – 13). The currently favored theories will not be discussed separately here. Instead, the list below considers the general events that occur during an enzyme reaction, and some plausible explanations why they might lead to rate-enhancing effects:

1. The enzyme must be catalytically ready. Simpler enzymes will automatically be active, but for more complicated enzymes covalent modification, cofactor, or effector binding or a conformational change may be required to attain the active state.
2. The substrate or substrates and enzyme must approach each other for binding. This will happen by diffusion, but the probability of an encounter may be enhanced. One way of doing this is by “channeling,” in which the free volume available for the diffusion is reduced. Channeling can be produced either structurally, as occurs in multisubunit complexes, or virtually, in which case, for example, electrostatic interactions are used to guide the substrate to the active site.
3. The substrate binds to the enzyme’s active site and, in the process, the substrate is desolvated and water is forced out of the active site. Binding has two principal effects. First, it brings together the substrates, cofactors, and the catalytic groups of the enzyme in an arrangement that is suitable for reaction and, second, it provides an environment that is very different from that in solution.
   As emphasized by Page and Jencks,[14, 15] the alignment of the reacting groups in a specific orientation is entropically unfavorable because it restricts the conformations available to the substrate, but it means that the subsequent reaction will be entropically more likely because the substrate will have to explore fewer conformations to reach the TS. The entropy loss on binding is compensated for by the enthalpy of binding between the substrate and the enzyme and the enthalpy and entropy coming from the liberation of water from within the active site and around the substrate. This entropy effect has been used to convincingly explain the rate enhancements for certain reactions such as those observed between uni- and bimolecular reactions in solution, but there is less certainty about its role in enzyme reactions (see, e.g., ref. 16).
   The active site can have very different dielectric properties from those in solution. In general, proteins have dielectric constants much lower than that of water, and so the further the active site is buried within the protein the lower its dielectric constant will become. This can significantly affect the properties of the substrate, such as the pK_as of its charged groups. It should be noted that even though the active site may have a low dielectric constant, it can still be highly polar due to the presence of electric fields created by groups within the protein.
   Substrate binding may be accompanied by a conformational change in the enzyme, which can significantly affect the substrate’s environment. This conformational change requires energy, but it may be necessary to exclude solvent from the active site region or to bring groups that are catalytically active or important for binding into contact with the substrate.
4. The reaction takes place to give products. The enzyme can accelerate the reaction by stabilizing the TS for the reaction more than that in solution (note that it is assumed that the mechanisms are the same—it may be that the enzyme can use an alternative mechanism, which is either less favorable or simply not possible in solution). Stabilization is achieved by the enzyme providing a “preorganized” environment that is complementary to the TS or, in other words, the shape and the interactions of the active site will be optimal for the TS structure. It is probable that a variety of interactions aid in stabilization of the TS, but Warshel has highlighted the importance of electrostatic interactions.[5, 10] A consequence of preorganization is that the reorganization energy associated with the reaction in the enzyme is small. In contrast, in solution, reorganization energies are often large because the solvent must change its structure substantially to accommodate the TS structure. Preorganization is not free because it stresses the enzyme structure, but this price is paid for by the enzyme when it is synthesized and not during catalysis.[17]
   Although the idea of TS stabilization is widely accepted, some workers have argued that it is not the whole story. Thus, for example, Bruice and Benkovic[4] and Menger[13] propose that formation of enzyme–substrate complexes in which the substrate is in a conformation ready to react is crit-

ical. These conformations, termed near-attack conformers (NACs) by Bruice and coworkers, have the reacting groups of both the substrate and enzyme in close contact and resemble the TS structure. Experiments and calculations on series of intramolecular reactions have shown that rates of reaction are proportional to the probability of the reactant being in a NAC. This data, together with the results of simulations of several enzyme systems, has led to the suggestion that it is the ability of enzymes to bind substrates in NACs that is central to their catalytic ability.

Many other mechanisms have been put forward as playing a role in catalysis. These include dynamical effects (although the weight of evidence at the moment seems to indicate that there is not a great diffference between enzyme and solution), quantum effects, such as tunneling, and low-barrier hydrogen bonds (LBHBs). Some of these will be discussed in more detail below.

The above scheme is by no means intended to be a comprehensive survey of all the factors that are important for enzyme catalysis but will serve to focus the discussion of later sections.

Simulation Techniques

It should be clear from the previous section that an understanding of enzyme reactions requires the investigation of a variety of processes occurring on a variety of scales of time and of distance. In such circumstances, no single simulation technique will be appropriate for studying all the events occurring during the catalysis, and it will be necessary to employ a range of theoretical approaches. Most commonly-used protein-modeling tools can be fruitfully applied to looking at some aspect or other of the enzyme catalytic cycle but, due to space limitations, the discussion in this section will concentrate on recent developments in techniques specifically designed for simulating reactions in complex systems.

Quantum Mechanical Methods

The chemical, bond-breaking and -forming steps lie at the heart of the study of all enzyme reactions, whether these steps are rate-limiting or not. To explain these events, it is crucial to have quantum mechanical (QM) methods that can accurately predict the structures and the energetics of reacting groups in large molecular systems. Undoubtedly, the methods that currently best meet these requirements are those based upon density functional theory (DFT), which have revolutionized quantum chemistry in the last decade or so.[18] In contrast to methods that attempt to solve the Schrödinger equation and for which the wave function is the basic variable, DFT methods express the electronic energy of the system as a functional of the electron density. In principle, the advantages of this approach are obvious because the wave function is a function of $3n$ space variables, n being the number of electrons in the system, whereas the electron density is a function of only three space variables.[19] In practice, the picture is a little more complicated as the quality of the solutions of wave function-based methods can be improved in a systematic way,[20] whereas the accuracy of DFT methods lies in a quantity called the exchange-correlation functional, whose analytic form is unknown.

DFT techniques can be implemented efficiently and are comparable in cost to molecular orbital (MO) Hartree–Fock (HF) methods. They do, however, provide energetic and structural results for ground and excited state systems that are often as precise as much more expensive correlated MO-based methods. But there are limitations to their accuracy, notably for applications to hydrogen-bonded and weakly bound systems and for the determination of the barrier heights for some reactions. These problems are likely to become less marked though as better exchange-correlation functionals are developed.[18]

Whereas DFT methods may be the choice for calculations on large systems, wave function-based methods are still desirable, if only to provide benchmark results against which the accuracy of the DFT techniques can be compared. Thus, for example, a typical approach would be to obtain structures for the reactant, saddle point, and product species in a reaction using a DFT method and then perform single point calculations to get the relative energies of the same structures at a higher level of theory. This is valid because DFT methods often predict geometries that are very close to those obtained with the more accurate methods. Such a strategy was adopted in a recent study of the reaction of cyclopentadiene with atomic oxygen in its ^{3}P state by Grossman et al.[21] What was notable in this article, though, was the use of the quantum Monte Carlo (QMC) method, which has the ability to produce results of very high accuracy at a much lower cost than methods of equivalent precision such as the configuration-interaction and coupled-cluster algorithms.[22] Unfortunately, the use of this promising technique has been limited due to the absence of any widely available general-purpose QMC programs.

Although *ab initio* quantum chemistry methods are to be preferred when studying reactions, they are costly to apply, and there are instances in which quicker, albeit cruder, QM methods are useful. Thus, for example, it is often important to be able to carry out calculations in the initial stages of an investigation rapidly so that a range of different possible pathways can be examined. Equally, extensive molecular dynamics (MD) or Monte Carlo (MC) calculations on systems of any reasonable size are currently impractical with *ab initio* techniques and so the determination of free energies, for example, requires other approaches.

The most popular alternatives to *ab initio* quantum chemical methods are the semiempirical methods. There is definitely still a need for such techniques although the most commonly used ones, such as the AM1 and MNDO methods of Dewar et al.[23–25] and the PM3 parametrization of the AM1 Hamiltonian of Stewart[26] are becoming dated. Even so, for many problems, these methods have a precision comparable to or better than *ab initio* DFT or HF calculations performed with small basis sets (up to double-ζ) despite being substantially less expensive.[27] As yet, there are no generally available replacements for the MNDO-like methods, which are based upon the MO HF theory, but Thiel and coworkers have published refinements to their formalism and preliminary parametrizations for some elements.[28, 29]

There are, of course, other ways of formulating semiempirical methods. A recently introduced approach that has been claimed to give improvements over the Dewar methods is the self-consistent charge tight-binding (SCCTB) method of Elstner, Porezag, Frauenheim, and coworkers.[30–32] In this method, the energy of the system, E, can be written as the sum of two terms. The first, E_0, is

the typical tight-binding energy expression:

$$E_0 = \sum_i n_i \langle \psi_i | \hat{H} | \psi_i \rangle + E_{\text{rep}} \quad (7)$$

where $\hat{H}$ is the tight-binding Hamiltonian for the system, ψ_i is the ith orbital with occupation number n_i, and E_{rep} is the repulsion energy between atomic cores. The second term, E_1, is:

$$E_1 = \frac{1}{2} \sum_{\alpha\beta} \gamma_{\alpha\beta} q_\alpha q_\beta \quad (8)$$

where q_α is the partial charge on atom α and $\gamma_{\alpha\beta}$ is a matrix element for the interaction between two charges. The orbitals for the system are expanded in terms of atom-centered basis functions and the optimum orbitals are obtained by minimizing the total energy expression with respect to the orbital expansion coefficients. As the atomic charges are derived from a population analysis of the system's wave function, they depend upon the orbitals and so solution of the equations must be performed in a self-consistent fashion.

An interesting aspect of tight-binding methods is that they can be shown to arise from DFT by expanding the Kohn–Sham equations in terms of fluctuations in the atomic density.[33] In addition to providing a firm theoretical foundation, such a link gives expressions for many of the parameters appearing in the model in terms of quantities that can be determined from *ab initio* DFT calculations. This, in principle, greatly simplifies the parametrization procedure. Finally, it is worth remarking that the tight-binding approaches, widely used in solid-state physics, are equivalent to the Hückel methods of quantum chemistry and that the self-consistent determination of charges has elements in common with the iterative Hückel methods developed by Hoffmann and other workers.[34, 35]

As well as the general-purpose semiempirical methods discussed above, other more specific QM methods have been used to represent the potential energy surfaces of reacting systems in simulation studies. These include the well-known empirical valence bond (EVB) method of Warshel and coworkers,[36] more recent refinements of this technique that use MO methods to calculate the wave functions of the different valence bond resonant states[37] and approaches that reparametrize existing general-purpose semiempirical methods to give results that are accurate for a particular reacting system.[38, 39] All these methods have their place and can give useful results when carefully employed. Unfortunately, this often requires considerable experience, which has limited the routine application of these methods to problems of arbitrary complexity, although future work may change this.[40]

The cost of calculations with the general-purpose *ab initio* and semiempirical methods mentioned above scale formally as the cube or a higher power of the size of the system, i.e. $\geq O(N^3)$ where N is the number of atoms in the system. Depending upon the type of algorithm being employed, the two most common bottlenecks in electronic structure calculations are the evaluation of the interelectronic interaction terms and the diagonalization of the Hamiltonian matrix to find the electronic orbitals. Thus, the evaluation of the two-electron integrals in a Hartree–Fock calculation that uses Gaussian or Slater basis functions scales formally as $O(N^4)$ [although it is relatively straightforward to reduce this scaling to between $O(N^2)$ and $O(N^3)$], whereas diagonalization scales as $O(N^3)$.

It is evident that such scalings will limit substantially the types of system to which electronic structure calculations can be applied and so much effort has gone into developing algorithms whose cost scales linearly with the size of the system. Much progress has been made and linear-scaling approaches now exist for most of the common *ab initio* and semiempirical DFT and HF methods. A nice review of this field has been given by Goedecker.[41] Specific applications to protein systems have mostly been limited to semiempirical methods and include those by Yang, York, and coworkers,[42] by Merz and coworkers,[43] and by Gready and coworkers.[44] Although linear-scaling algorithms permit extended systems to be studied with QM methods, they have the disadvantage at present that they only become competitive with the traditional methods for relatively large numbers of atoms. Thus, calculations on such systems are expensive which means, for example, that linear-scaling QM calculations of proteins can be used for energy evaluations and restricted geometry optimizations but not for prolonged molecular dynamics simulations.

Hybrid Potential Methods

In the not-so-distant future algorithmic enhancements and improved computer hardware will allow systems of thousands of atoms to be studied using purely QM techniques. Currently, however, this is not practicable and so other methods have had to be devised to treat reactions in condensed-phase systems. The most successful class of approaches are those based upon hybrid potentials, in which potentials of differing accuracy are used to treat different regions of the system. Thus, for example, chemical reactions are typically studied by treating the reacting atoms and those immediately surrounding them with a QM potential, and using a simpler method for the atoms of the remainder of the system. When partitioning the system in this way, the assumption is made that the process of interest is localized in the QM region—this will be reasonable for many reactions but it will not be valid in some instances, such as, for example, when the reaction is coupled to an electron-transfer event.

Many types of hybrid potential have been implemented (for reviews, see refs. 45–47). They differ in the number of regions into which the system is divided, the types of potential used to treat the different regions and the ways in which the interfaces between the potentials are handled. Probably the most common hybrid potentials are those for the analysis of solvent effects in which only the solute is treated explicitly and the solvent is modeled as a continuum with a dielectric constant of the appropriate value.[48, 49] These potentials have also been used to study reactions in enzymes (see, e.g., ref. 50) but such models will, in general, be inappropriate due to the heterogeneity of the active site structure, and it will be necessary to include more detailed information about the environment. This is accomplished using hybrid QM/MM potentials in which the atoms of the protein are included explicitly in the calculation and whose potential is modeled with a molecular mechanics (MM) or an empirical energy function.

QM/MM potentials have been employed to treat a wide range of enzyme reactions (for a partial list, see ref. 46), since their introduction by Warshel and Levitt in the 1970s.[51] The great majority

of studies have been done with EVB or MNDO-type semiempirical QM methods to describe the reacting atoms, a standard nonpolarizable protein force field to represent the rest of the protein and an interface in which there are nonbonding electrostatic and Lennard-Jones interactions between the atoms of the two regions. The way of dealing with the case where there are covalent bonds between the QM and MM regions has depended upon the semiempirical method being used. For EVB potentials, the bonds are treated with standard MM terms because most EVB parametrizations have represented the energy of each valence bond resonant structure in an MM-like manner. For the MNDO semiempirical methods, it has been usual to employ the link-atom approximation in which additional atoms, one per bond between the QM and MM atoms, are introduced into the QM region to replace the MM atom of the broken bonds in the quantum calculation. As each fictitious atom serves to sate a valence of one of the QM atoms bonded covalently to an MM atom, it is represented by a univalent atom, normally a hydrogen but sometimes a halogen.[52]

Whereas most hybrid potential applications to enzymes have been performed up until now with semiempirical QM methods, there are more and more applications with *ab initio* QM/MM potentials. This is, in large part, due to the increasing availability of programs that implement hybrid potentials and is a trend that will continue as commercial products, from such companies as Gaussian[53] and Schrödinger,[54] push these techniques into the mainstream. The use of *ab initio* QM methods in QM/MM potentials places more demands on the precision of the coupling between the atoms of the QM and MM regions and, because of this, most of the recent innovative developmental work on QM/MM potentials has concerned them. This is in contrast to the case of semiempirical QM/MM potentials for which the basic formalism of the coupling and its accuracy is probably satisfactory.

Most of the work on coupling has concentrated on how to describe covalent bonds between QM and MM atoms. As mentioned above, the link atom scheme has been the most popular with semiempirical QM/MM potentials, although other algorithms have been proposed. These include methods based upon hybrid orbitals[51, 55, 56] and those that use pseudoatoms.[57] Although these methods are more elegant in that they dispense with the fictitious link atoms in the QM region, they have not yet been shown to be consistently better than the newer link atom approximations,[58, 59] which are more versatile and, in comparison to the hybrid orbital methods, simpler to implement. The situation is less clear with *ab initio* QM/MM potentials for which link atom,[60, 61] hybrid orbital,[62–65] and pseudoatom[66] methods have all been developed. The most impressive results have been obtained by Murphy et al.,[65] with a frozen orbital method that has been parametrized for amino acid groups and that reproduces conformational energetics to very high accuracy (to within $\sim$5 kJ mol^{-1}). The formalism of the method though is rather complex, and it has to be reparametrized every time there is a covalent bond between QM and MM atoms in a different environment.

Understandably, most work on QM/MM methods has gone into dealing with covalent bonds at the QM and MM interface. With the increasing use of *ab initio* methods, however, it is likely that other aspects of the coupling will need to be improved, and that the treatment of the interactions between the QM and MM regions with simple electrostatic and Lennard-Jones terms will no longer be sufficient. Little work appears to have been done in this area, although Murphy et al. have performed a thorough parametrization of the QM atom Lennard-Jones parameters and experimented with terms to correct hydrogen-bond interactions between QM and MM atoms,[65] and Day et al. have developed an effective fragment model in which the charge distribution on the MM atoms (the "spectator groups") is treated with a multipole expansion and there are explicit terms for charge penetration, exchange-repulsion, and polarization effects.[67]

An alternative way of improving the boundary in hybrid potential methods is to employ several QM methods of varying precision for the core region of the simulation system instead of a single one. Thus, a high-level *ab initio* method could be used for the reacting atoms, a lower level QM method for the atoms in their immediate environment and an MM method for the remainder of the system. In this way, the problem of treating the boundaries between different regions is less acute as the differences between the methods on either side of the boundary are smaller. Methods of this type have been pioneered by Morokuma and coworkers with their IMOMO/ONIOM series of models (see, e.g., refs. 68 and 69), but others have also developed similar techniques, including Gogonea et al., who have introduced a combined *ab initio* DFT/semiempirical QM algorithm,[70] and Hong et al., who describe a DFT method in which the electronic density of the environment is either frozen or constrained.[71]

To finish this section it is perhaps worth restating the advantage that QM/MM methods have over pure QM methods for treating large molecular systems. This point was made vividly in the recent work of Titmuss et al., who used both linear-scaling semiempirical QM and hybrid QM/MM potentials to determine the reaction pathways for hydride transfer in dihydrofolate reductase (DHFR) and found that the hybrid potential calculations were about 500 times quicker.[44]

Miscellaneous Developments

As stated earlier, a variety of simulation approaches will be appropriate for looking at enzyme reactions. The previous sections discussed in detail the most crucial aspect of any study—how to represent accurately the potential energy surface for the reacting atoms—but in this section developments in other areas that are likely to benefit enzyme simulation work will be briefly considered.

Although QM potentials are of primary concern for reactions, work on MM potentials is important for improving the description of the environment in hybrid potentials as well as the quality of normal force-field simulation studies. Apart from the push to better parametrize existing force fields, the most significant change happening in this area at the moment is the gradual introduction of general-purpose potentials that are polarizable. Whereas the representation of a system's charge distribution by fixed charges on the atoms captures the essential features of its electrostatic interactions, the inclusion of polarization is necessary for a more detailed treatment.

The use of polarization terms in MM potentials has a long history (see, e.g., ref. 72), and can be done in a number of different ways. Much recent work, however, has gone into the elegant charge equilibration or fluctuating charge schemes that were first used in conjunction with MM potentials by Rappé and Goddard.[73] These

methods, whose formalisms are derivable from density functional theory,[74, 75] no longer treat the charge distributions of the atoms as fixed, but allow them to change in response to their environment. There are a number of different varieties of charge equilibration algorithms, which differ principally in the way that the changing charge distributions are modeled. The simplest schemes, like that of Rappé and Goddard,[73] use charges centered on the atoms, but other representations in terms of dipoles[76] or arbitrary basis sets[77] are also possible. In addition to the development of the models themselves, it has been shown how these schemes may be incorporated straightforwardly into hybrid potentials.[78, 79]

There is, of course, much more to a simulation than the potential because the potential must be used appropriately if meaningful results are to be obtained. In general, the whole gamut of simulation methodologies can be formulated for use with any given potential, but the principal limiting factor dictating which method can be exploited with which potential is the time required for the evaluation of the energy (or related quantities such as the forces on the atoms) for a single configuration of the system. Thus, *ab initio* QM or QM/MM potentials can be combined effectively with methods that search the potential energy surface locally, such as geometry minimization or saddle-point location algorithms, whereas semiempirical QM or QM/MM and MM potentials can be employed with methods that sample the system's phase space more comprehensively, such as molecular dynamics or procedures that determine free energies.

It is unnecessary to review the range of simulation methodologies here as most can be adapted, with little change, to studying enzyme systems. There is one aspect that needs highlighting though and that concerns simulations with *ab initio* hybrid potentials in which there is a large disparity in cost between the QM and MM portions of the calculation. In these cases, it pays to find algorithms that minimize the number of QM calculations that are required. Several such schemes have been developed, including geometry optimization methods,[65, 80] in which extensive minimizations in the MM subspace are coupled to less frequent geometry updates of the QM region, and methods for the calculation of free energies, in which the sampling of the QM part of the potential energy surface is enhanced by the use of approximate replacement potentials, of either EVB-[36, 81] or of fluctuating charge-type.[82]

To end this section, a new class of methods will be considered that are having a large impact in the investigation of reaction processes, although they do not yet appear to have been applied to enzyme systems. Traditionally, calculations of the rate of a condensed-phase reaction have employed a transition-state or activated-dynamics approach (see, e.g., ref. 83), which typically consists of three steps: (i) the definition of a reaction coordinate; (ii) the evaluation of the activation free energy by calculation of the potential of mean force (PMF) along the reaction coordinate; (iii) the estimation of dynamical factors, such as the transmission coefficient, that correct the transition-state rate expression. This scheme, while fairly robust, relies upon being able to identify a reaction coordinate. This is normally done either by selecting the geometrical variables, which are thought to be important for the description of the reaction, or by locating a minimum-energy path between reactants and products on the potential energy surface of the system using one of the standard techniques. For simple reactions both approaches are probably satisfactory, but in complex systems there are two problems. First, it can be difficult to identify appropriate variables or to locate a pathway in the high-dimensional space available to the system and, second, the hypothesis of a single, dominant reaction coordinate may not be a good one.

These problems are circumvented in recently introduced statistical algorithms that generate ensembles of transition paths between stable states and dispense with the notion of a single, well-defined reaction path entirely. The original method was due to Pratt,[84] but his ideas have been elaborated upon by Chandler and coworkers,[85–87] and by Woolf and coworkers.[88] These groups have formulated various recipes for determining the transition paths between reactants and products and have shown how to use this information to estimate the reaction rate. Chandler *et al.* have applied their methods to reactions of biomolecular relevance using both *ab initio* QM and MM potentials, the former to a proton transfer in a water cluster[86] and the latter to the isomerization of the alanine dipeptide in solution.[87]

Recent Simulation Studies

In this section, examples of enzyme simulation studies are discussed. The section starts with a general presentation followed by the elaboration of two specific topics. Other recent reviews of various aspects of the simulation of enzyme reactions may be found in refs. 89–92.

General Studies

It is impractical to review here all the enzymes that have been investigated with simulation techniques and so only a couple of examples will be mentioned. However, to give readers a flavor of what has been done, recent work (among much other) has been performed on the reactions catalyzed by acetyl cholinesterase,[93] aldose reductase,[94] carbonic anhydrase,[95–97] catechol *O*-methyltransferase,[98–100] chorismate mutase,[101, 102] citrate synthase,[103–106] dihydroxyfolate reductase,[107–109] enolase,[110] formate dehydrogenase,[111] glyoxalase I,[112, 113] haloalkane dehydrogenase,[114] HhaI methyltransferase,[115] orotidine 5′-monophosphate decarboxylase (ODCase),[116, 117] protein tyrosine phosphatase,[118, 119] the GTPase reaction of p21 RAS,[120] ribonuclease A,[121] and trypsin.[122, 123] Additional references, to pre-1999 work on enzymes using hybrid potentials, may be found in ref. 46. Other aspects of enzyme function that have received attention include calculations to determine the dielectric properties of enzyme active sites,[124] molecular dynamics simulations to estimate the transmission coefficient correction factor to the transition state rate for a proton transfer in the triosephosphate isomerase (TIM) reaction,[125] simulations of the diffusional encounter between enzymes and substrates,[126] and calculations on intra- and inter-molecular anhydride formation that act as models for enzyme-catalyzed reactions.[127–129]

The studies mentioned above employed the whole range of theoretical methodologies from high-level *ab initio* QM calculations on model systems to the determination of the free-energy profile for the reaction in the enzyme using molecular dynamics simulations. In many cases, several techniques have been applied to the same enzyme so as to obtain different insights into the behavior of

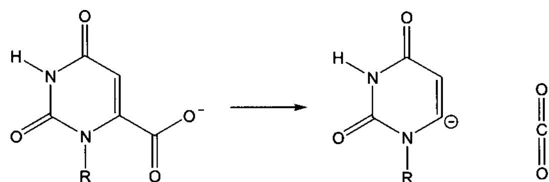


Figure 1. The decarboxylation of orotidine 5′-monophosphate to form a carbanion and carbon dioxide. "R" stands for a ribose-phosphate moiety, and the 4-position of the orotate ring is at the top.

the reaction. A nice example of this concerns the enzyme ODCase that catalyzes the decarboxylation of orotidine 5′-monophosphate to uridine 5′-monophosphate (see Fig. 1). It is currently the most catalytically proficient enzyme known with a ΔG_{rel} of about $-130\ \mathrm{kJ\,mol^{-1}}$ (see earlier).[130] In an early study of this reaction, before X-ray crystallographic structures of the enzyme were available, Lee and Houk employed *ab initio* QM calculations of model systems to postulate a mechanism in which decarboxylation occurred simultaneously with protonation of the orotate ring in the 4-position by a weak acid in the protein to produce a carbene intermediate.[50] Such a mechanism was shown to be consistent with the observed rate acceleration produced by ODCase as long as the enzyme active site was assumed to have a low dielectric constant.

Since then, four crystal structures of different ODCases complexed with substrate and substrate analogs have appeared that have not supported this mechanism. Instead, the active site has been shown to contain a novel Lys–Asp–Lys–Asp charge network, with one of the aspartates juxtaposed against the carboxylate group of the orotate ring. This has led to the suggestion that the driving force for the reaction is due to destabilization of the substrate in the enzyme–substrate complex because the electrostatic repulsion between the negatively charged aspartate and carboxylate will be relieved in the transition state. The energy required to destabilize the reacting groups is paid for by the favorable interactions between the ribose and phosphate groups of the substrate and the enzyme.[131] This scheme has been supported by simulations with a semiempirical QM/MM hybrid potential that determined the free-energy profiles for the enzyme and solution-phase reactions and the free energy of transfer of the substrate between enzyme and solution.[117].

Warshel and coworkers have also carried out simulations on this reaction using their EVB methodology.[116] Like the previous workers, they agree that electrostatic effects are responsible for the catalysis but they argue against ground-state destabilization because they estimate that the interactions between the ribose and phosphate of the substrate and the enzyme that pay for the destabilization are too small, and that such a large destabilization energy would in any case induce a change in the protonation state of the charged groups at the active center. In contrast, they propose that the charged groups in the active site are preorganized so as to preferentially stabilize the charge distribution of the transition state structure.

Another enzyme reaction that has been subjected to a range of different theoretical techniques is that of citrate synthase, which catalyzes the condensation of oxaloacetate and an acetyl group attached to acetyl-coenzyme A. One of the principal mechanistic uncertainties in this reaction has been the identity of the interme-

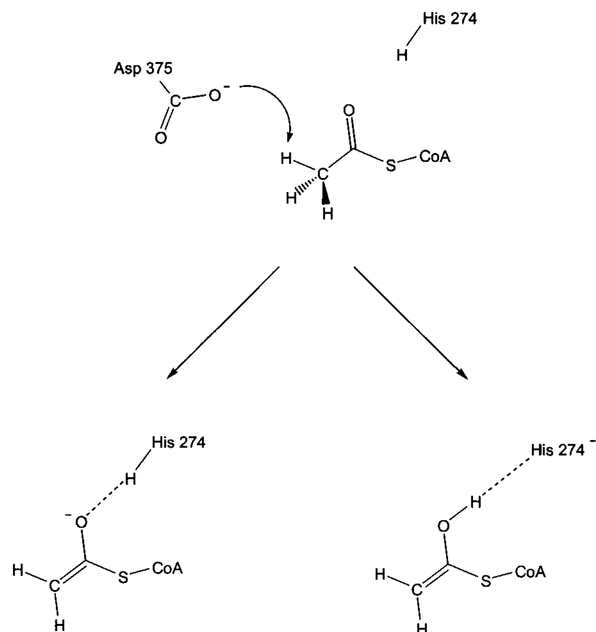


Figure 2. The initial hydrogen abstraction step of the citrate synthase reaction giving either an enolate (left) or an enol (right) intermediate. Co-A stands for coenzyme A, and the amino acid numbering corresponds to that of the pig enzyme.[105]

diate after the abstraction of the proton from the terminal methyl group of acetyl-coenzyme A by the side chain of an aspartate residue in the enzyme (see Fig. 2). It had been argued that the enolate-imidazole form would be unstable, and that an alternative intermediate would be preferred. This has been postulated to be an enol-imidazolate pair or a form intermediate between enol and enolate that involves stabilization by short, strong hydrogen bonds (or LBHBs).

Traditional hydrogen bonds with nitrogen and oxygen donors and acceptors have distances between heavy atoms of the order of 2.8 Å and energies of approximately $20\ \mathrm{kJ\,mol^{-1}}$. LBHBs are significantly shorter and stronger, with distances often around 2.5 Å and energies of up to $80\ \mathrm{kJ\,mol^{-1}}$. The main difference between the two types of hydrogen bond is that LBHBs have a high covalent character, whereas the interactions in traditional hydrogen bonds are primarily electrostatic. As the hydrogen-bond donor–acceptor distance shortens, the barrier to proton transfer decreases and the hydrogen becomes more diffusely distributed between the two. Short hydrogen bonds have been observed in small molecule, enzyme–inhibitor and enzyme–substrate analog crystal structures,[8] and have been characterized in *ab initio* QM calculations of hydrogen-bonded systems.[132–134] LBHBs have been postulated to play an important role in enzyme catalysis because formation of a LBHB in a TS would provide stabilization and a reduced barrier to proton transfer. Examples of enzyme mechanisms that have been proposed to be LBHB-assisted include those of the serine proteases, of TIM, and of citrate synthase.[8] These ideas are not universally accepted, however. Thus, Warshel and collaborators think that LBHBs are likely to be anticatalytic because their more diffuse charge distribution will result in a smaller electrostatic stabilization of a TS structure by an enzyme.[135, 136]

In early work on the citrate synthase reaction, Mulholland and coworkers tested various mechanistic hypotheses employing semiempirical QM/MM potentials in conjunction with geometry optimization and reaction-path location techniques.[103, 104] They found that the initial hydrogen abstraction step was indeed effected by an aspartate, and that the enolate intermediate was more stable than the alternative enol form. Due to doubts about the accuracy of the semiempirical methods, they have more recently performed calculations with *ab initio* QM/MM potentials and obtained equivalent results. They see no evidence of involvement of LBHBs.[105]

Donini et al. have reached similar conclusions in their simulations of citrate synthase but by a different route.[106] Rather than using QM/MM potentials to examine the reaction directly in the enzyme, Kollman and coworkers have adopted a slightly simpler strategy that they name the quantum-mechanical free-energy (QM-FE) approach.[137] This involves three steps: (1) the characterization of the reaction in a model system in the gas phase with high-level *ab initio* calculations; (2) the use of this data to parametrize a MM model for the reaction in the enzyme; (3) standard molecular dynamics and free-energy perturbation simulations with the MM model to determine the free-energy profiles for the reaction or the free-energy differences between various reaction intermediates. The QM-FE method has the advantage that molecular dynamics and free-energy simulations are much less expensive than with QM/MM potentials, but the disadvantage that the reaction mechanism of the model system is assumed not to be significantly perturbed in the enzyme environment.

Metalloproteins

A significant number of enzymes contain metals that participate passively or actively in the catalysis. A passive participation is due to electrostatic effects in which, for example, the metal cation serves to align a negatively charged substrate in the active site or to stabilize a negative charge that forms along the reaction pathway. A more active participation is involved when the redox state of the metal changes during the course of the reaction or when its coordination is altered.

If the metal's influence is essentially electrostatic, then it can be modeled in a reasonable way using either MM or semiempirical QM potentials. There is a caveat, though, which is that the groups coordinating the metal can be highly polarized when the charge on the metal is large, and these effects are often either neglected or poorly reproduced by the commonly used MM and semiempirical QM methods. With this reservation, a number of studies of metalloproteins have been performed using semiempirical QM/MM potentials, including those of the enzymes acetohydroxyisomeroreductase[138] and D-xylose isomerase[139] (magnesium), of carbonic anhydrase[96] (zinc) and of the metal-binding protein, metallothionein[140] (cadmium and zinc).

When the metal participates directly in the reaction of the enzyme as is the case, for example, with the hydrogenases (iron and nickel) or enzymes that use coenzyme B_{12} (cobalt), there is no alternative but to use an *ab initio* QM method if a reasonable description of the reaction is to be obtained. The most practical methods, for looking at reactions involving transition metals are the DFT methods which have a cost comparable to that of HF calculations but much higher precision. Quite a few DFT studies of metalloenzymes have been performed, but all of them, with only a couple of exceptions, have neglected the protein environment and employed truncated models of the protein's active site. Notable examples of such work are the investigation of the soluble methane monooxygenase (MMO) by Dunietz et al.[141] and of a number of metalloproteins, including cytochrome *c* oxidase (iron and copper) and the water oxidizing complex in photosystem II (manganese), by Blomberg, Siegbahn, and coworkers.[142, 143] The article of Dunietz et al. on MMO represents the state-of-the-art in *ab initio* studies of models of enzyme active sites. Their approach differed substantially from previous ones because they included a large number of atoms in the model (~100), not only the two irons of the active center and their immediate coordinating atoms but also all the atoms of the coordinating groups. In addition, they used large basis sets (triple-ζ with polarization functions) to determine the energetics of the structures they had previously geometry optimized with smaller basis sets.

There have been very few DFT studies of metalloenzymes that include the environment. The first was performed by Amara et al. on the nickel-iron hydrogenase from *Desulfovibrio gigas*.[144] The aim of this work was not the determination of the reaction mechanism per se, but the elucidation of the nature of the redox states taking part in the catalytic cycle. The protein was modeled with a DFT QM/MM hybrid potential with approximately 30 atoms of the active site in the quantum region and 15,000 atoms, protein, and water, in the MM region. The results of the simulations underlined the importance of representing the environment properly as the structures of the nickel-iron group in the different, redox states were distinctly different depending upon whether the protein was present or not. A more recent study by Himo et al. of galactose oxidase (GO) also employed a hybrid DFT QM/MM potential,[145] but, unlike the hydrogenase work, only a very limited portion of the protein was included in the MM region. GO converts alcohols to aldehydes with a simultaneous reduction of molecular oxygen to hydrogen peroxide, and its mechanism involves a copper atom and tyrosine radical intermediates.

Quantum Effects

Up until now, in this review, a lot has been spoken of QM methods but these have always referred to methods that calculate the potential energy surface for the atoms in the system. The use that has been made of this potential energy surface—for example, the finding of reaction paths and the calculation of free energies—has involved algorithms that are based upon classical mechanics. There is increasing evidence though that quantum dynamical effects, and, in particular, tunneling, play a role in determining the rates of enzyme reactions.[146–148]

For a reaction to occur within a classical picture, the system must pass over the potential barrier that separates reactants from products. In the quantum world, this is no longer the case because there is a certain probability that the system will be able to tunnel through the barrier and allow the reaction to occur. Tunneling is more pronounced the lighter the particle, and so is important in reactions involving electron transfer, but it will also take place when there are transfers of hydrogen and, to a lesser extent, its isotopes deuterium and tritium. There are several different descriptions of tunneling reactions. In the simplest version, the barrier through

which the tunneling takes place is static, and the quantum contribution to the reaction will be independent of temperature. In more complicated theories, there is a temperature dependence because the dynamics of the protein environment serves to align the reactant and product wells in a configuration that is suitable for the hydrogen transfer. Both these types of tunneling appear to have been observed in practice.[148]

Simulations of enzyme reactions that try and account for quantum effects have been relatively rare, primarily because the algorithms available for performing quantum dynamical calculations are immature compared to their classical counterparts. Probably the earliest workers in this area were Warshel, Hwang, and coworkers who estimated quantum corrections to the rates of hydrogen and deuterium transfer reactions in enzymes and solution given by classical TST.[149–151] They did this by assuming that the classical and quantum TST rate expressions were of the same form except that the classical activation free energy, $\Delta G^{\ddagger}_{cl}$, was replaced by its quantum equivalent, $\Delta G^{\ddagger}_{qu}$, thus giving a quantum correction factor of $\exp[-(\Delta G^{\ddagger}_{qu} - \Delta G^{\ddagger}_{cl})/RT]$. The classical free energies were calculated in the normal way, whereas the quantum free energies were determined using a path-integral simulation algorithm.[152, 153] Path-integral simulations are straightforward to implement and to perform (see, e.g., ref. 59 for details), but they are computationally expensive and so further applications of this technique to enzyme reactions have been limited, although a study of a hydrogen transfer in the reaction catalyzed by the flu virus neuraminidase has also been published.[154]

A number of techniques, other than path-integral methods, have been used to investigate quantum dynamical effects in enzyme reactions. Thus, in two recent articles, Alhambra et al. employed semiempirical QM/MM potentials in conjunction with semiclassical variational TST to calculate the rates and the kinetic isotope effects (KIEs) for the proton transfer in the yeast enolase reaction[155] and the hydride transfer in liver alcohol dehydrogenase (LADH).[156] The hydride transfer in LADH has also been studied by Hammes–Schiffer and coworkers,[157, 158] but, in contrast to Alhambra et al., they solved the time-independent Schrödinger equation to obtain wave functions for the transferring hydrogen along various paths between the reactant and product states. Both groups of workers found that tunneling played a significant contribution in the transfer, and needed to be included if the experimental secondary KIEs were to be reproduced. Bala et al. have also looked at a proton transfer in an enzyme reaction, this time in phospholipase A_2, but they were able to observe the quantum dynamics of the transfer directly because they employed an algorithm that coupled the solution of the time-dependent Schrödinger equation for the quantum particle with a classical molecular dynamics simulation of the remaining atoms.[159]

Summary

So, is Blow correct?[2] Well, in a strict sense, yes. There are few (if any) enzymes for which a fully quantitative understanding obtained from theory is available (and upon which all the workers in the field would agree!). The reason, as Blow so rightly states, is due to the difficulty of accurately modeling the interactions and the dynamics that determine reaction processes in systems that are as large as enzymes. Having said this though, there has been a huge amount of high quality theoretical work that has led to a real understanding of certain aspects of certain enzyme reactions.

The situation is only likely to get better, as there has been substantial progress in the last few years. It is now possible to investigate with DFT QM methods reactions in systems of between 100 to 200 atoms and, although not consistently capable of providing chemical accuracy, they will be adequate in many instances. Likewise, robust implementations of *ab initio* QM/MM potentials are appearing that extend the applicability of the QM methods to larger systems, and improvements in simulation algorithms mean that transition paths and thermodynamic quantities can be reliably determined for reactions in complex systems. All in all, the outlook is bright—new algorithms combined with increased computer power mean that precise simulations of realistic models of enzyme reactions are likely very soon. This is fortunate, and indeed necessary, if "computational enzymology" is to contribute fully to the interpretation of the function of the large number of enzyme structures to be generated by the structural genomics initiatives[160] and in the use of this information for the rational design of drugs[161] and new enzyme functions.[162]

References

1. Fersht, A. Enzyme Structure and Mechanism; W. H. Freeman and Company: New York, 1985.
2. Blow, D. Structure 2000, 8, R77.
3. Kurz, J. L. J Am Chem Soc 1963, 85, 987.
4. Bruice, T. C.; Benkovic, S. J. Biochemistry 2000, 39, 6268.
5. Warshel, A. Theroet Chem Acc 2000, 103, 337.
6. Bruice, T. C.; Lightstone, F. C. Acc Chem Res 1999, 32, 127.
7. Neet, K. E. J Biol Chem 1998, 273, 25527.
8. Cleland, W. W.; Frey, P. A.; Gerlt, J. A. J Biol Chem 1998, 273, 25529.
9. Cannon, W. R.; Benkovic, S. J. J Biol Chem 1998, 273, 26257.
10. Warshel, A. J Biol Chem 1998, 273, 27035.
11. Warshel, A.; Florián, J. Proc Natl Acad Sci USA, 1998, 95, 5950.
12. Cannon, W. R.; Singleton, S. F.; Benkovic, S. J. Nat Struct Biol 1996, 3, 821.
13. Menger, F. M. Acc Chem Res 1993, 26, 206.
14. Page, M. I.; Jencks, W. P. Proc Natl Acad Sci USA, 1971, 68, 1678.
15. Jencks, W. P. Adv Enzymol 1975, 43, 219.
16. Villà, J.; Strajbl, M.; Glennon, T. M.; Sham, Y. Y.; Chu, Z. T.; Warshel, A. Proc Natl Acad Sci USA 2000, 97, 11899.
17. Shoichet, B. K.; Baase, W. A.; Kuroki, R.; Matthews, B. W. Proc Natl Acad Sci USA, 1995, 92, 452.
18. Koch, W.; Holthausen, M. C. A Chemist's Guide to Density Functional Theory; Wiley VCH: New York, 2000.
19. Kohn, W. Rev Mod Phys 1999, 71, 1253.
20. Pople, J. A. Rev Mod Phys 1999, 71, 1267.
21. Grossman, J. C.; Lester, W. A., Jr.; Louie, S. G. J Am Chem Soc 2000, 122, 705.
22. Ceperley, D. M.; Mitas, L. Adv Chem Phys 1996, XCIII, 1.
23. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. J Am Chem Soc 1985, 107, 3902.
24. Dewar, M. J. S.; Thiel, W. J Am Chem Soc 1977, 99, 4899.
25. Thiel, W.; Voityuk, A. A. Theor Chim Acta 1992, 81, 391.
26. Stewart, J. J. P. J Comput Chem 1989, 10, 209; J Comput Chem 1989, 10, 221.
27. Dewar, M. J. S.; O'Connor, B. M. Chem Phys Letts 1987, 138, 141.

28. Kolb, M.; Thiel, W. J Comput Chem 1993, 14, 775.
29. Weber, W.; Thiel, W. Theor Chem Acc 2000, 103, 495.
30. Porezag, D.; Frauenheim, Th.; Köhler, Th.; Seifert, G.; Kaschner, R. Phys Rev B 1995, 51, 12947.
31. Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, Th.; Suhai, S.; Seifert, G. Phys Rev B 1998, 58, 7261.
32. Frauenheim, Th.; Seifert, G.; Elstner, M.; Hajnal, Z.; Jungnickel, G.; Porezag, D.; Suhai, S.; Scholz, R. Phys Stat Sol B 2000, 217, 41.
33. Foulkes, W. M. C.; Haydock, R. Phys Rev B 1989, 39, 12520.
34. Dixon, S. L.; Jurs, P. C. J Comput Chem 1994, 15, 733.
35. Vela, A.; Gázquez, J. L. J Phys Chem 1988, 92, 5688.
36. Warshel, A. Computer Modeling of Chemical Reactions in Enzymes and Solutions; John Wiley and Sons: New York, 1991.
37. Mo, Y.; Gao, J. J Comput Chem 2000, 21, 1458.
38. Rossi, I.; Truhlar, D. G. Chem Phys Letts 1995, 233, 231.
39. Bash, P. A.; Ho, L. L.; MacKerell, A. D., Jr.; Levine, D.; Hallstrom, P. Proc Natl Acad Sci USA, 1996, 93, 3698.
40. Kim, Y.; Corchado, J. C.; Villà, J.; Xing, J.; Truhlar, D. G. J Chem Phys 2000, 112, 2718.
41. Goedecker, S. Rev Mod Phys 1999, 71, 1085.
42. York, D. M.; Lee, T.-S.; Yang, W. Phys Rev Letts 1998, 80, 5011.
43. Nadig, G.; Van Zant, L. C.; Dixon, S. L.; Merz, K. M., Jr. J Am Chem Soc 1998, 120, 5593.
44. Titmuss, S. J.; Cummins, P. L.; Bliznyuk, A. A.; Rendell, A. P.; Gready, J. E. Chem Phys Letts 2000, 320, 168.
45. Gao, J. In Reviews in Computational Chemistry; Lipkowitz, K. B.; Boyd, D. B., Eds.; VCH: New York, 1995; p. 119, vol. 7.
46. Amara, P.; Field, M. J. In Computational Molecular Biology; Leszczynski, J., Ed.; Elsevier: Amsterdam, 1999; p. 1.
47. Monard, G.; Merz, K. M. Acc Chem Res 1999, 32, 904.
48. Tomasi, J.; Persico, M. Chem Rev 1994, 94, 2027.
49. Simonson, T. Curr Opin Struct Biol 2001, 11, 243.
50. Lee, J. K.; Houk, K. N. Science 1997, 276, 942.
51. Warshel, A.; Levitt, M. J Mol Biol 1976, 103, 227.
52. Field, M. J.; Bash, P. A.; Karplus, M. J Comput Chem 1990, 11, 700.
53. Gaussian, Inc.: Carnegie Office Park, Building 6, Suite 230, Carnegie, PA 15106, USA.
54. Schrödinger, Inc.: 1500 S.W. First Avenue, Suite 1180, Portland, OR 97201, USA.
55. Monard, G.; Loos, M.; Thery, V.; Baka, K.; Rivail, J.-L. Int J Quantum Chem 1996, 58, 153.
56. Gao, J.; Amara, P.; Alhambra, C.; Field, M. J. J Phys Chem A 1998, 102, 4714.
57. Antes, I.; Thiel, W. J Phys Chem A 1999, 103, 9290.
58. Reuter, N.; Dejaegere, A.; Maigret, B.; Karplus, M. J Phys Chem A 2000, 104, 1720.
59. Field, M. J.; Albe, M.; Bret, C.; Proust-De Martin, F.; Thomas, A. J Comput Chem 2000, 21, 1088.
60. Eichinger, M.; Tavan, P.; Hutter, J.; Parrinello, M. J Chem Phys 1999, 110, 10452.
61. Lyne, P. D.; Hodoscek, M.; Karplus, M. J Phys Chem A 1999, 103, 3462.
62. Assfeld, X.; Rivail, J.-L. Chem Phys Letts 1996, 263, 100.
63. Phillip, D. M.; Friesner, R. A. J Comput Chem 1999, 20, 1468.
64. Murphy, R. B.; Phillip, D. M.; Friesner, R. A. Chem Phys Letts 2000, 321, 113.
65. Murphy, R. B.; Phillip, D. M.; Friesner, R. A. J Comput Chem 2000, 21, 1442.
66. Zhang, Y.; Lee, T.-S.; Yang, W. J Chem Phys 1999, 110, 46.
67. Day, P. N.; Jensen, J. H.; Gordon, M. S.; Webb, S. P.; Stevens, W. J.; Krauss, M.; Garmer, D.; Basch, H.; Cohen, D. J Chem Phys 1996, 105, 1968.
68. Vreven, T.; Morokuma, K. J Comput Chem 2000, 21, 1419.
69. Maseras, F.; Morokuma, K. J Comput Chem 1995, 16, 1170.
70. Gogonea, V.; Westerhoff, L. M.; Merz, K. M., Jr. J Chem Phys 2000, 113, 5604.
71. Hong, G.; Strajbl, M.; Wesolowski, T. A.; Warshel, A. J Comput Chem 2000, 21, 1554.
72. Warshel, A.; Russel, S. T. Q Rev Biophys 1984, 17, 283.
73. Rappé, A. K.; Goddard, W. A., III. J Phys Chem 1991, 95, 3358.
74. Itskowitz, P.; Berkowitz, M. L. J Phys Chem A 1997, 101, 5687.
75. York, D. M.; Yang, W. J Chem Phys 1996, 104, 159.
76. Stern, H. A.; Kaminski, G. A.; Banks, J. L.; Zhou, R.; Berne, B. J.; Friesner, R. A. J Phys Chem B 1999, 103, 4470.
77. Bret, C.; Field, M. J.; Hemmingsen, L. Mol Phys 2000, 98, 751.
78. Field, M. J. Mol Phys 1997, 91, 835.
79. Bryce, R. A.; Vincent, M. A.; Burton, N. A. J Chem Phys 1998, 109, 3077.
80. Zhang, Y.; Liu, H.; Yang, W. J Chem Phys 2000, 112, 3483.
81. Bentzien, J.; Muller, R. P.; Florián, J.; Warshel, A. J Phys Chem B 1998, 102, 2293.
82. Iftimie, R.; Salahub, D.; Schofield, J. J Chem Phys 2000, 113, 4852.
83. Karplus, M. J Phys Chem B 2000, 104, 11.
84. Pratt, L. R. J Chem Phys 1986, 85, 5045.
85. Dellago, C.; Bolhuis, P. G.; Csajka, F. S.; Chandler, D. J Chem Phys 1998, 108, 1964.
86. Geissler, P. L.; Dellago, C.; Chandler, D.; Hutter, J.; Parrinello, M. Chem Phys Letts 2000, 321, 225.
87. Bolhuis, P. G.; Dellago, C.; Chandler, D. Proc Natl Acad Sci USA, 2000, 97, 5877.
88. Zuckerman, D. M.; Woolf, T. B. J Chem Phys 1999, 111, 9475.
89. Bruice, T. C.; Kahn, K. Curr Opin Chem Biol 2000, 4, 540.
90. Kazlauskas, R. J. Curr Opin Chem Biol 2000, 4, 81.
91. Tantillo, D. J.; Chen, J.; Houk, K. N. Curr Opin Chem Biol 1998, 2, 743.
92. Friesner, R. A.; Beachy, M. D. Curr Opin Struct Biol 1998, 8, 257.
93. Fuxreiter, M.; Warshel, A. J Am Chem Soc 1998, 120, 183.
94. Várnai, P.; Warshel, A. J Am Chem Soc 2000, 122, 3849.
95. Toba, S.; Colombo, G.; Merz, K. M., Jr. J Am Chem Soc 2000, 121, 2290.
96. Merz, K. M., Jr.; Banci, L. J Am Chem Soc 1997, 119, 863.
97. Lu, D.; Voth, G. J Am Chem Soc 1998, 120, 4006.
98. Lau, E. Y.; Bruice, T. C. J Am Chem Soc 1998, 120, 13387.
99. Zheng, Y.-J.; Bruice, T. C. J Am Chem Soc 1997, 119, 8137.
100. Kuhn, B.; Kollman, P. A. J Am Chem Soc 2000, 122, 2586.
101. Kharyin, N. A.; Synder, J. P.; Menger, F. M. J Am Chem Soc 1999, 121, 11831.
102. Martí, S.; Andrés, J.; Moliner, V.; Silla, E.; Tuñón, I.; Bertrán, J.; Field, M. J. J Am Chem Soc 2001, 123, 1709.
103. Mulholland, A. J.; Richards, W. G. Proteins Struct Funct Genet 1997, 27, 9.
104. Mulholland, A. J.; Richards, W. G. J Phys Chem B 1998, 102, 6638.
105. Mulholland, A. J.; Lyne, P. D.; Karplus, M. J Am Chem Soc 2000, 122, 534.
106. Donini, O.; Darden, T.; Kollman, P. A. J Am Chem Soc 2000, 122, 12270.
107. Radkiewicz, J. L.; Brooks, C. L., III. J Am Chem Soc 2000, 122, 225.
108. Castillo, R.; Andrés, J.; Moliner, V. J Am Chem Soc 1999, 121, 12140.
109. Cummins, P. L.; Gready, J. E. J Phys Chem B 2000, 104, 4503.
110. Liu, H.; Zhang, Y.; Yang, W. J Am Chem Soc 2000, 122, 6560.
111. Torres, R. A.; Schiott, B.; Bruice, T. C. J Am Chem Soc 1999, 121, 8164.
112. Feierburg, I.; Cameron, A. D.; Aqvist, J. FEBS Letts 1999, 453, 90.
113. Feierburg, I.; Luzhkov, V.; Aqvist, J. J Biol Chem 2000, 275, 22657.

114. Lau, E. Y.; Kahn, K.; Bash, P. A.; Bruice, T. C. Proc Natl Acad Sci 2000, 97, 9937.
115. Lau, E. Y.; Bruice, T. C. J Mol Biol 1999, 293, 9.
116. Warshel, A.; Strajbl, M.; Villà, J.; Florián, J. Biochemistry 2000, 39, 14728.
117. Wu, N.; Mo, Y.; Gao, J.; Pai, E. F. Proc Natl Acad Sci 2000, 97, 2017.
118. Kolmodin, K.; Nordlund, P.; Aqvist, J. Proteins Struct Funct Genet 1999, 36, 370.
119. Kolmodin, K.; Aqvist, J. FEBS Letts 1999, 456, 301.
120. Glennon, T. M.; Villà, J.; Warshel, A. Biochemistry 2000, 39, 9641.
121. Glennon, T. M.; Warshel, A. J Am Chem Soc 1998, 120, 10234.
122. Peräkylä, M.; Kollman, P. A. J Am Chem Soc 2000, 122, 3436.
123. Stanton, R. V.; Peräkylä, M.; Bakowies, D.; Kollman, P. A. J Am Chem Soc 1998, 120, 3448.
124. Simonson, T.; Archontis, G.; Karplus, M. J Phys Chem B 1999, 103, 6142.
125. Neria, E.; Karplus, M. Chem Phys Letts 1997, 267, 23.
126. Zhou, H.-X.; Wong, K.-Y.; Vijayakumar, M. Proc Natl Acad Sci 1997, 94, 12372.
127. Lightstone, F. C.; Bruice, T. C. J Am Chem Soc 1996, 118, 2595.
128. Lightstone, F. C.; Bruice, T. C. J Am Chem Soc 1997, 119, 9103.
129. Peräkylä, M.; Kollman, P. A. J Phys Chem A 1999, 103, 8067.
130. Radzicka, A.; Wolfenden, R. Science 1995, 267, 90.
131. Begley, T. P.; Appleby, T. C.; Ealick, S. E. Curr Opin Struct Biol 2000, 10, 711.
132. Schiøtt, B.; Iversen, B. B.; Madsen, G. K. H.; Bruice, T. C. J Am Chem Soc 1998, 120, 12117.
133. Schiøtt, B.; Iversen, B. B.; Madsen, G. K. H.; Larsen, F. K.; Bruice, T. C. Proc Natl Acad Sci 1998, 95, 12799.
134. Kumar, G. A.; McAllister, M. A. J Am Chem Soc 1998, 120, 3159.
135. Warshel, A.; Papazyan, A.; Kollman, P. A. Science 1995, 269, 103.
136. Warshel, A.; Papazyan, A. Proc Natl Acad Sci 1996, 93, 13665.
137. Kollman, P. A.; Kuhn, B.; Donini, O.; Peräkylä, M.; Stanton, R. V.; Bakowies, D. Acc Chem Res 2001, 34, 72.
138. Proust-De Martin, F.; Dumas, R.; Field, M. J. J Am Chem Soc 2000, 122, 7688.
139. Hu, H.; Liu, H.; Shi, Y. Proteins Struct Funct Genet 1997, 27, 545.
140. Berweger, C. D.; Thiel, W.; van Gunsteren, W. F. Proteins Struct Funct Genet 2000, 41, 299.
141. Dunietz, B. D.; Beachy, M. D.; Cao, Y.; Whittington, D. A.; Lippard, S. J.; Friesner, R. A. J Am Chem Soc 2000, 122, 2828.
142. Siegbahn, P. E. M.; Blomberg, M. R. A. Chem Rev 2000, 100, 421.
143. Siegbahn, P. E. M.; Blomberg, M. R. A. Annu Rev Phys Chem 1999, 50, 221.
144. Amara, P.; Volbeda, A.; Fontecilla-Camps, J.; Field, M. J. J Am Chem Soc 1999, 121, 4468.
145. Himo, F.; Eriksson, L. A.; Maseras, F.; Siegbahn, P. E. M. J Am Chem Soc 2000, 122, 8031.
146. Sutcliffe, M. J.; Crutton, N. S. Trends Biochem Sci 2000, 25, 405.
147. Scrutton, N. S.; Basran, J.; Sutcliffe, M. J. Eur J Biochem 1999, 264, 666.
148. Kohen, A.; Klinman, J. P. Chem Biol 1999, 6, R191.
149. Hwang, J.-K.; Warshel, A. J Am Chem Soc 1996, 118, 11745.
150. Hwang, J.-K.; Warshel, A. J Phys Chem 1993, 97, 10053.
151. Hwang, J.-K.; Chu, Z. T.; Yadav, A.; Warshel, A. J Phys Chem 1991, 95, 8445.
152. Feynman, R. P.; Hibbs, A. R. Quantum Mechanics and Path Integrals; McGraw-Hill: New York, 1965.
153. Chandler, D.; Wolynes, P. G. J Chem Phys 1981, 74, 4078.
154. Thomas, A.; Jourand, D.; Bret, C.; Amara, P.; Field, M. J. J Am Chem Soc 1999, 121, 9693.
155. Alhambra, C.; Gao, J.; Corchado, J. C.; Villà, J.; Truhlar, D. G. J Am Chem Soc 1999, 121, 2253.
156. Alhambra, C.; Corchado, J. C.; Sánchez, M. L.; Gao, J.; Truhlar, D. G. J Am Chem Soc 2000, 122, 8197.
157. Webb, S. P.; Agarwal, P. K.; Hammes-Schiffer, S. J Phys Chem B 2000, 104, 8884.
158. Agarwal, P. K.; Webb, S. P.; Hammes-Schiffer, S. J Am Chem Soc 2000, 122, 4803.
159. Bala, P.; Grochowski, P.; Nowinski, K.; Lesyng, B.; McCammon, J. A. Biophys J 2000, 79, 1253.
160. See, for example, the articles in the supplement to Nat Struct Biol 2000, 7.
161. Gane, P. J.; Dean, P. M. Curr Opin Struct Biol 2000, 10, 401.
162. Cedrone, F.; Ménez, A.; Quéméneur, E. Curr Opin Struct Biol 2000, 10, 405.