Computational Enzymology

Article *in* Methods in molecular biology (Clifton, N.J.) · January 2013

DOI: 10.1007/978-1-62703-017-5_4 · Source: PubMed

2 authors, including:

Alessio Lodola
University of Parma
187 PUBLICATIONS **4,830** CITATIONS

SEE PROFILE

Chapter 4

Computational Enzymology

Alessio Lodola and Adrian J. Mulholland

Abstract

Techniques for modelling enzyme-catalyzed reaction mechanisms are making increasingly important contributions to biochemistry. They can address fundamental questions in enzyme catalysis and have the potential to contribute to practical applications such as drug development.

Key words: QM/MM, Enzyme, Catalysis, Protein dynamics, Biomolecular simulation, Quantum mechanics/molecular mechanics

1. Introduction

Molecular modelling and simulations can explore mechanisms of biological catalysts (i.e., enzymes) at a level of detail that cannot be achieved experimentally (1–11). Modelling can unravel the mechanisms of enzyme-catalyzed reactions, identify the origins of catalysis, analyze effects of mutations and genetic variations, and help to develop structure–activity relationships (12–14). Since its origins (15, 16), computational enzymology has grown enormously, particularly in recent years. There has also been a significant improvement in the accuracy of computational methods. For example, it is now possible to achieve an unprecedented level of accuracy in calculations on enzyme-catalyzed reactions with combined quantum mechanics/molecular mechanics (QM/MM) methods (17). In the best cases, calculations can give activation energies that agree extremely well with experiment. High-level quantum chemical methods allow calculations of energy barriers, in the best cases, near "chemical accuracy" (1 kcal/mol) (18). Quantitative predictions at this level were only previously possible for very small molecules. Carefully parameterized empirical molecular simulation approaches also give excellent agreement with experiments for enzyme reactions (19).

Luca Monticelli and Emppu Salonen (eds.), *Biomolecular Simulations: Methods and Protocols*, Methods in Molecular Biology, vol. 924, DOI 10.1007/978-1-62703-017-5_4,

Identifying the chemical mechanisms of enzymes solely from experiments is often difficult. Many mechanisms in the literature are probably wrong in important details, e.g., as more recent experiments and simulations have shown for hen egg-white lysozyme (20, 21). The physical origins of enzyme catalysis also continue to be hotly debated. Recent controversies have centered on "low-barrier" hydrogen bonds (22–25), so-called near-attack conformations (26, 27) enzyme dynamics (28, 29), quantum tunnelling (30–33), and entropic effects (34). The applicability of transition state theory to enzyme reactions has also been questioned (35). Molecular simulations are proving to be crucial in testing these proposals.

Transition states are central to understanding chemical reactivity and catalysis, but experiments cannot directly study them in enzymes because of their extremely short lifetimes, and because of the large size and complexity of enzymes. Molecular modelling can analyze transition states directly and identify interactions involved in catalysis (e.g., a conserved proline residue that specifically stabilizes the transition state for aromatic hydroxylation in the flavin dependent monooxygenases *para*-hydroxybenzoate hydroxylase and phenol hydroxylase (36, 37)). Such interactions may not be apparent (and may not exist) in available experimental structures. This type of knowledge can assist in ligand design, e.g., as a potential route to enhanced affinity. Also, in contrast to some experimental (e.g., structural) studies, which may require mutation of the enzyme or use of alternative (e.g., inefficient) substrates (e.g., to slow down the reaction, and prolong the lifetime of intermediates, to allow their spectroscopic or structural characterization), molecular modelling can study directly the "wild type" reaction (i.e., the reaction as it occurs in the naturally occurring enzyme). Computational enzymology interacts fruitfully with experiments, which can validate modelling approaches, which in turn can interpret experimental findings and suggest new experiments.

2. Materials and Methods

2.1. Choice of Enzyme Structure for Modelling

An enzyme structure from X-ray crystallography (ideally high resolution), is the usual starting point for modelling an enzyme-catalyzed reaction. A crystal structure of an enzyme alone, with no ligand bound at the active site, may not be useful, because it is difficult to predict binding modes and protein conformational changes associated with binding. NMR structure ensembles can give useful complementary information on dynamics and interactions, but generally do not define atomic positions precisely enough for mechanistic modelling. In some cases a homology model may be sufficiently reliable, but should be treated with great caution: the positions, relative orientations and packing of side chains may not

be modelled sufficiently accurately. An example of the successful use of homology modelling for mechanistic studies is an investigation of the substrate binding mode and reaction mechanism of a malarial protease with a novel active site, by automated docking, and molecular dynamics/reaction free energy profile simulations (38).

The structure used for modelling must accurately represent the reacting enzyme complex; a crystal structure of an enzyme-inhibitor complex is often a good choice. The inhibitor should resemble the substrate, product, transition state or an intermediate, in its bound conformation. One should remember that there can be local structural uncertainty due e.g., to protein dynamics, and conformational variability or disorder, even in high-resolution structures. It is usually not possible to determine crystal structures of active enzyme-substrate complexes, unless specialized conditions or variant substrates or enzymes are used. In some cases, structures of several complexes along the reaction pathway may be available (30). Calculations can then model the chemical and structural changes, unstable intermediates and transition state structures, providing a picture of the whole reaction in the enzyme. For many enzymes, the conformational changes that take place during the reaction are small, and so modelling based on a single structure of an enzyme complex can give a reliable picture of the reaction. It may, however, be necessary to consider the effects of conformational fluctuations in the protein (see below).

2.2. Effects of Protein Dynamics

Proteins have many conformational substates, and a single structure may not be truly representative for modelling a mechanism (39). Extensive conformational sampling may be needed to generate a representative ensemble of structures. To calculate free energy profiles, i.e., potentials of mean force (40), a simulation method must be capable of calculating trajectories of many picoseconds at least (or a similarly large number of configurations in a Monte Carlo simulation). Alternatively, molecular dynamics simulations can be used to generate multiple structural models for subsequent mechanistic calculations, to ensure wide sampling of possible enzyme configurations. If multiple different crystal structures of the same enzyme are available, these may be suitable as different starting models, and similarly help to examine the effects of structural variation on the reaction.

Protein dynamics are believed to be important to their biological functions in many cases (41). It is well known that many enzymes undergo large conformational changes during their reaction cycles (42). The possible relationship of dynamics to enzyme catalysis is more controversial. It has been proposed that protein dynamics contribute significantly to enhancing the rates of reaction in enzymes, but simulations indicate that the effect of protein dynamics in determining the rates of chemical reactions in enzymes is relatively small (43). Protein conformational changes

(e.g., involved with substrate binding or product release) can, however, be rate-limiting for the overall reaction in many enzymes (44), and in some cases are coupled to chemical changes (e.g., facilitating product release). Quantum effects such as nuclear tunnelling are important in reactions involving hydrogen transfer (30, 45, 46) and the effects of protein dynamics on reactions involving quantum tunnelling is an area of particularly active debate (5, 10, 28, 47).

2.3. Determining the Mechanism

Determining the chemical mechanism is an essential first step in studying an enzyme-catalyzed reaction. This is not trivial: many "textbook" mechanisms are probably wrong. The first aim is to establish the identities and functions of catalytic residues; many mechanisms in the biochemical literature assign functions to residues which are probably incorrect. Next, any specific interactions that stabilize transition states or reaction intermediates should be identified and analyzed. A typical computational approach to modelling reactions is to optimize the structures of key species (such as transition state structures); preferably entire reaction pathways should be optimized or simulated.

2.4. Analyzing Catalysis

To understand *catalysis*, i.e., to understand why a reaction in an enzyme reaction is faster than an uncatalyzed or nonenzymic reaction, the two reactions should be compared (although deciding on an appropriate "reference" reaction may be difficult (6)). Practical applications often have simpler aims such as predicting the effects of a mutation on activity or on the specificity of an enzyme for alternative substrates. Overall, understanding enzyme mechanism, specificity and catalysis presents a range of challenges, and different types of modelling or simulation methods are needed to investigate different types of question, as outlined below.

2.5. Methods for Modelling Enzymes and Enzyme-Catalyzed Reactions

2.5.1. Molecular Mechanics

"Molecular mechanics" (MM) methods can model protein structure and dynamics well, but standard MM methods cannot be used directly to model chemical reactions, because of their simple functional forms (e.g., harmonic terms for bond stretching, and an inability to model changes in electronic distribution because of the invariant atomic point partial charge model). The simplicity of MM "force fields" (potential functions) allows long timescale (e.g., now up to millisecond) simulations of protein dynamics, and simulations of large proteins. Molecular dynamics simulations can study conformational changes (which are rate-limiting in many enzymes under typical conditions): e.g., simulations of the human scavenger decapping enzyme (DcpS) found a cooperative periodic opening and closing of the dimer, over tens of nanoseconds (48). Molecular dynamics simulations can also investigate substrate conformational behavior, which can help to develop mechanistic ideas,

but simulations of substrate complexes without consideration of the reaction can sometimes be misleading (49).

Computer programs for biomolecular dynamics simulations include AMBER (50), CHARMM (51), GROMOS (52), NAMD (53) and TINKER (54); these should not be confused with force fields, which may have the same or similar names. A force field consists of the energy function and the parameters. Current protein force fields use similar potential energy functions, in which bonds and valence angles are represented by harmonic terms, electrostatic interactions are represented by invariant point charges on atoms, a simple representation that cannot capture the full electrostatic properties of a molecule. Dispersion and exchange repulsion are included by a simple Lennard-Jones function (usually of the 12-6 variety). Current widely used all-atom MM force fields for proteins are OPLS/AA (55, 56), CHARMM22-27 (57), and AMBER (PARM99 (50, 58–60)). Force fields for other types of biological macromolecules (e.g., lipids, nucleic acids (61, 62)) consistent with these protein force fields have also been developed (e.g., the CHARMM27 (63, 64) and AMBER nucleic acid parameters (65) and CHARMM parameters for lipids (66)). Most biomolecular MM force fields have been developed to be compatible with simple point charge models of water, e.g., the TIP3P water model (67). Current standard biomolecular MM force fields only include electronic polarization in an average, invariant way. The next generation of protein MM force fields will probably include electronic polarization explicitly (68, 69).

Standard MM potential functions cannot be applied to model the breaking and making of bonds (and electronic reorganization) in a chemical reaction. Also, MM force field parameters are developed based on the properties of stable molecules, and so will usually not be applicable to transition states and intermediates. MM functions and parameters can be developed specifically for reactions (e.g., using different functional forms, such as Morse curves to allow bond breaking), which has been successful for organic reactions in solution (70). Such parameters are, however, applicable only to a particular reaction, or small class of reactions. Also, the form of the potential function imposes limitations, such as the neglect of electronic polarization.

2.5.2. Empirical Valence Bond Methods

In the empirical valence bond (EVB) method (6), a few resonance structures are chosen to represent the reaction. The energy of each resonance form is given by a simple empirical force field, with the potential energy given by solving the related secular equation. The EVB Hamiltonian is calibrated to reproduce experimental data for a reaction in solution, or ab initio results can be used (71). The surrounding protein and solvent are modelled by an empirical force field, with appropriate treatment of long-range electrostatics. The activation free energy of activation is calculated from free

energy perturbation simulations (72). The free energy surfaces can be calibrated by comparison with experimental data for reactions in solution. The EVB method allows the use of a non-geometrical reaction coordinate, which allows evaluation of nonequilibrium solvation effects (6). A mapping procedure gradually moves the system from the reactants to products. The simplicity of the EVB potential function allows extensive molecular dynamics simulations, giving good sampling (73). The EVB method is now widely used for studying reactions in condensed phases, particularly in enzymes (74–80).

2.5.3. Quantum Chemical Calculations on Small (Active Site) Models

In most enzymes, the chemical changes occurring in the reaction are confined to a relatively small region, the active site of the enzyme. One approach to the study of enzyme-catalyzed reactions is to study only the active site, using quantum chemical methods (this is sometimes called the "supermolecule" or cluster approach). Such models can represent important features of an enzyme reaction, and can identify likely mechanisms. The active site model should contain molecules representing the substrate(s) (and any cofactors) and enzyme residues involved in the chemical reaction or in binding substrate. Important functional groups (such as catalytic amino acid side chains) are represented by small molecules, e.g., acetate can represent an aspartate side chain, imidazole for histidine, etc.). The initial positions of these groups are usually coordinates taken from a crystal structure, or from a molecular dynamics simulation of an enzyme complex.

Quantum chemical calculations (i.e., methods that calculate molecular electronic structure using quantum mechanics, e.g., ab initio molecular orbital or density functional theory calculations) can give excellent results for reactions of small molecules. The best "ab initio" methods (such as CCSD(T)), which include correlation between electrons, can calculate rate constants for reactions involving very few atoms (in the gas phase) with small error bars, similar to experiments on these systems. Such calculations require very large computational resources, however, severely limiting the size of the system that can be treated. More approximate methods, (such as the semiempirical molecular orbital techniques AM1 and PM3), are computationally much cheaper, and can model larger systems (containing of the order of hundreds of atoms). Techniques (e.g., "linear-scaling" methods) have been developed that allow semiempirical electronic structure calculations on whole proteins (81–83). Semiempirical methods are, however, inaccurate for many applications (e.g., typical errors of over 10 kcal/mol for barriers and reaction energies, though specifically parameterized semiempirical methods can give improved accuracy for a particular reaction (40, 47)). Density functional theory (DFT) methods (e.g., applying the B3LYP functional) are considerably more accurate, while also allowing calculations on relatively large systems

(e.g., active site models of the order of 100 atoms), larger than is feasible with correlated ab initio calculations. Many DFT methods, however, lack important physical interactions, such as dispersion, which are important in the binding of ligands to proteins. Dispersion effects can also be important in the calculation of energy barriers (84). DFT often gives barrier heights that are too low by several kcal/mol, and it can be difficult to assess the accuracy of results, because DFT does not offer a route to their systematic improvement or testing.

Calculations on active site models can provide models of transition states and intermediates (see below). This has proved particularly useful for studying metalloenzymes, using DFT methods. In many metalloenzymes, all the important chemical steps take place at one metal center (or a small number of metal ions bound at one site). The metal also holds its ligands in place, limiting the requirement for restraints to maintain the correct active site structure. Calculations on small clusters can give useful mechanistic insight (85, 86): e.g., a mechanism can be ruled out if the calculated barriers for it are significantly higher than the experimentally derived activation energy, based on the accuracy of the computational method. The effects of the environment are usually either omitted, or included only in an approximate way (e.g., by continuum solvation methods, which cannot fully represent the heterogeneous electrostatic environment in an enzyme). It is useful to test the sensitivity of the results to the choice of, e.g., dielectric constant.

To calculate the energy barrier for a reaction in a cluster model, structures of the reactant, transition state, intermediates and products of the reaction should be optimized. Doing this while maintaining the correct orientations of the groups in the protein can be difficult. Small models may also lack some important functional groups. It is important to consider carefully which groups to include, striking a balance between computational feasibility and the desire for a larger, more extensive model. A larger model is not, however, always a better model: a larger model will be susceptible to greater conformational complexity: conformational changes, even outside the active site, may artificially affect relative energies along the reaction path). Also, charged groups can have unrealistically large effects on reaction energies. One should test the sensitivity of the results to the choice of model (and also to factors such as the choice of density functional).

2.5.4. Combined Quantum Mechanics/Molecular Mechanics Methods for Modelling Enzyme Reactions

"Hybrid" methods that combine quantum chemical methods with molecular mechanics allow more extensive calculations, on larger models of enzymes, than is possible with purely quantum chemical techniques. Such QM/MM methods are very important in computational enzymology. The QM/MM approach is simple: a small part of the system, at the active site, is treated quantum mechanically, i.e., by an electronic structure method of one of the types

discussed above, which allows modelling of the electronic rearrangements involved in the breaking and making of chemical bonds. The QM region contains the reacting groups of the enzyme, substrate and any cofactors. The rest of the system is treated by MM. QM/MM calculations can be carried out at ab initio (87) or semiempirical (88) molecular orbital, density functional (89) or approximate density functional (e.g., self-consistent charge density functional tight-binding, SCC-DFTB (90)) levels of QM electronic structure calculation. Different types of coupling between the QM and MM regions are possible (see below).

Many different QM/MM implementations are available, in several widely used programs. Reaction pathways and transition state structures can be optimized (91, 92). Molecular dynamics simulations are possible with cheaper QM/MM methods (such as semiempirical or SCC-DFTB level QM) (93). Free energy differences, such as activation free energies can be calculated, as can quantum effects such as tunnelling and zero-point corrections (5, 30, 40). High-level QM/MM calculations (e.g., ab initio or density functional level QM) are required for some systems and also have an important role in testing more approximate methods. The computational demands of high level (e.g., ab initio, (94)) QM/MM calculations (17) typically limit their application to "single point" energy calculations on structures optimized at lower levels (95). DFT/MM methods can be used for energy minimization/geometry optimization to generate reaction paths.

QM/MM methods can also be used in free energy perturbation simulations (96), e.g., to calculate relative binding affinities, and in molecular docking and scoring (97). QM/MM methods provide several advantages over MM methods in studies of ligands bound to proteins, including potentially a better physical description of a ligand (e.g., including electronic polarization), and avoiding the need for time-consuming MM parameterization for the ligand.

Interactions Between the QM and MM Regions

One of the main differences between various QM/MM methods is the type of QM/MM coupling employed i.e., in how the interactions (if any) between the QM and MM systems are treated (98). The simplest linking of QM and MM methods involves a straightforward "mechanical" embedding of the QM region in the MM environment, treating interactions between the QM and MM regions only by MM (i.e., the QM system is represented by (MM) point charges in its interaction with the MM environment). In calculations of this type, the QM/MM energy of the whole system, $E_{\mathrm{TOTAL}}{}^{\mathrm{QM/MM}}$, is calculated in a simple subtractive scheme. This simple subtractive approach can be applied to all combinations of theory levels (for example combining different levels of QM treatment (QM/QM) as opposed to QM with MM) and forms the basis for the (simplest form of the) multilayer

ONIOM (Our own N-layered Integrated molecular Orbital and molecular Mechanics) method (99). A QM/QM calculation involves a high and a low level of QM theory, with a small region treated by a high level and the *entire* model treated at the low level; polarization is included at the lower level of QM theory.

More intensive QM/MM calculations include polarization of the QM region by the MM environment. This is likely to be important for many enzymes, given their polar nature. QM/MM methods of this type include electrostatic interactions between the QM and MM regions in the QM calculation, thus modelling polarization of the QM system by the MM atoms, by directly including the MM atomic charges of the MM group in the QM calculation. The electronic structure calculation thus includes the effects of the MM atoms. A further level of complexity would involve polarization of the MM region also through the use of a polarizable MM force field, and potentially self-consistent polarization of the MM region through an iterative procedure. Models of this sort are vastly more computationally intensive and may not always yield better results (100). QM/MM methods that include polarization of the MM system have been developed for small molecular systems (101). Current standard MM force fields for biological macromolecules do not model changes in polarization, however.

In typical QM/MM calculations, the energy of the QM atoms, E_{QM}, is given by a molecular orbital or DFT method, and the energy of the atoms in the MM region, E_{MM}, is given by MM. A boundary term, $E_{Boundary}$, is usually necessary to account for the effects of the surroundings, e.g., to include the effects of parts of the protein that are not included in the simulation model. It may also be necessary to scale/reduce charges at the boundary of the simulation system: this represents the effects of dielectric screening in a crude sense, to avoid overestimating the effects of charged groups on the active site (102). The QM/MM interaction energy, $E_{QM/MM}$ typically consists of terms due to electrostatic interactions and van der Waals interactions, and any bonded interaction terms. In many implementations, MM bonding terms (energies of bond stretching, angle bending, torsion angle rotation, etc.) are included for all QM/MM interactions which involve at least one MM atom (88). In an ab initio QM/MM calculation, the MM atomic charges are generally included directly through one-electron integrals. The treatment of QM/MM electrostatic interactions is a little less straightforward when semiempirical molecular orbital methods such as AM1 and PM3 are used, because they treat only valence electrons directly, including the core electrons together with the nucleus as an atomic "core." In semiempirical QM/MM methods such as the AM1/CHARMM method of Field et al., the electrostatic interactions between QM and MM atoms are calculated by treating the MM atoms as if they where semiempirical atomic cores (88).

QM/MM van der Waals interactions (representing dispersion and exchange repulsion interactions between QM and MM atoms) are usually calculated by a molecular mechanics procedure (e.g., through Lennard-Jones terms), exactly as the corresponding interactions would be calculated between MM atoms not interacting through bonding terms. MM van der Waals parameters must therefore be chosen for each QM atom: these interactions are significant at short distances, and are important in determining QM/MM interaction energies and geometries. The van der Waals parameters are important in differentiating MM atom types in their interactions with the QM system, e.g., for MM atoms of the same charge, which would otherwise be indistinguishable to the QM system; van der Waals interactions are also important for interactions of the QM system with nearby MM atoms whose charges are close to zero. Often, standard MM van der Waals (Lennard-Jones) parameters optimized for similar MM groups are used for QM atoms in QM/MM calculations. This is convenient, but it is always important to consider whether the van der Waals parameters provide a reliable description of QM/MM interactions. Where necessary, the (MM) van der Waals parameters for the QM atoms can be optimized to reproduce experimental or high-level ab initio results (e.g., structures and interaction energies) for small molecular complexes (93). One limitation of current QM/MM approaches of this type is that the same van der Waals parameters are typically used for the QM atoms throughout a simulation: in modelling a chemical reaction, the chemical nature of the groups involved (treated by QM) may change, altering their interactions, and so the use of unchanging MM parameters may be inappropriate. Riccardi et al. have investigated the effects of van der Waals parameters in QM/MM (SCC-DFTB/CHARMM22) simulations (103). Different parameter sets gave very different results for gas-phase clusters and solvent structures around the solutes. However, condensed phase thermodynamic quantities (e.g., the calculated reduction potential and potential of mean force) were less sensitive to the van der Waals parameters. These authors concluded that work to improve the reliability of QM/MM methods for condensed phase energetic properties should focus on factors other than van der Waals interactions between QM and MM atoms, such as the treatment of long-range electrostatic interactions.

Treatment of Long-Range Electrostatic Interactions in QM/MM Simulations

To reduce computational requirements, the model may include only a part of the whole protein (for example, a rough sphere around the active site). In simulating a truncated protein system, it is necessary to include restraints or constraints in the boundary region to force the atoms belonging to it to remain close to their positions in the crystal structure. One common approach is the stochastic boundary molecular dynamics method (104, 105), in which the simulation system is divided into a reaction region,

a buffer region and a reservoir region. Typically, the whole simulation system may include all residues with an atom within a distance of e.g., 15–25 Å of an atom in the active site. The buffer region would contain atoms in the outer layer (e.g., outer 5 Å shell). Atoms in the reaction region are treated by standard Newtonian molecular dynamics, and are not subject to positional restraints. The protein heavy atoms in the buffer region are restrained to remain close to their (e.g., crystallographically determined) positions by harmonic forces, while a solvent deformable boundary potential prevents "evaporation" of water. Atoms in the buffer are subject to frictional and random forces (hence the term "stochastic") to represent exchange of energy with the surroundings (reservoir region). Atoms in the reservoir region are usually not included because their presence (as fixed atoms) has been found to cause excessive rigidity of the protein.

Ideally, long-range electrostatic interactions should be included explicitly. Schemes for treatment of long-range electrostatic interactions in QM/MM simulations have been developed, to allow periodic boundary simulations (106). An alternative approach, for QM/MM calculations under spherical boundary conditions (107) is the generalized solvent boundary potential (GSBP) method (108). This retains the practical advantage of treating a truncated system, avoiding having to include the entire macromolecule in a periodic simulation. The effects of the bulk solvent and macromolecule atoms outside the simulation system are included at the Poisson-Boltzmann level. Simulations using the GSBP method were found to be more consistent with experimental data. Conventional stochastic boundary molecular dynamics simulations produced artifacts, depending on the treatment of electrostatic interactions. It was suggested that the commonly used interaction truncation schemes should not be applied if possible in QM/MM simulations, in particular for simulations that may involve extensive conformational sampling.

QM/MM Partitioning Methods and Schemes

Most QM/MM studies of enzymes require partitioning of covalently bonded molecules into QM and MM regions. Typically, some amino acid side chains participate in the reaction, and must therefore be included in the QM region. Other side chains will play binding roles, and a MM representation might be inadequate. Similarly, it may be more practical to treat only the reactive parts of large cofactors or substrates by quantum chemical methods. There are two general QM/MM partitioning techniques that can be employed: firstly special treatment of orbitals to satisfy the valence shell of the QM atom at the QM/MM junction, for example the local self-consistent field (LSCF) method (109, 110) or the generalized hybrid orbital (GHO) method (111). Alternatively a QM atom (or pseudoatom) can be added at the

QM/MM boundary, e.g., using a "link atom" or connection atom method.

The local self-consistent field (LSCF) method (112) uses a strictly localized bond orbital, also often described as a frozen orbital, for the QM atom at the frontier between QM and MM regions. The electron density of the orbital is calculated in advance, using small models, and does not change during the QM/MM calculation. The orbitals must be parameterized for each system and QM method. The LSCF method avoids the need for dummy atoms and provides a reasonable description of the chemical properties of the frontier bond. It has been applied at semiempirical (113) and ab initio (112) QM/MM levels.

The generalized hybrid orbital (GHO) method (114) uses hybrid orbitals as basis functions on the frontier atom of the MM fragment. It does not require extensive specific parameterization, unlike the LSCF method. It uses four hybrid orbitals for an sp^3 carbon atom, one of which is included in the self-consistent field optimization of the QM region, while three are treated as auxiliary orbitals. The parameters for the frontier atom are optimized to reproduce properties of full QM systems. The localized orbitals can be transferred, without specific parameterization of the active orbital for each new system. A similar approach in DFT/MM calculations is to freeze the electron density at the QM/MM junction (115). The GHO method has been applied in QM/MM calculations at ab initio (116), SCC-DFTB (117) and density functional (118) QM levels.

The "dummy junction atom" or link atom method introduces so-called link atoms to satisfy the valence of the frontier atom in the QM system (119). The link atom is usually a hydrogen atom (88), but other atom types have been used, such as a halogen (120). The link atom approach has been criticized, e.g., because it introduces additional degrees of freedom associated with the link atom, and the fact that a C–H bond is clearly not exactly equivalent to a C–C bond. It is, however, simple and is widely used. The results can be sensitive to the positioning of the link atom, and also on exactly which MM atoms are excluded from the classical electrostatic field that interacts with the QM region. Comparison of the LSCF and link atom approaches for semiempirical QM/MM calculations, however, showed that the two methods gave similar results (121). It has been recommended that the link atom should interact with all MM atoms except for those closest to the QM atom to which the link atom is bonded. The link atom method can give good results, with a good choice of the boundary between QM and MM regions, e.g., across a carbon–carbon single bond, far from chemical changes, and also preferably not close to highly charged MM atoms.

Another method for dealing with the QM/MM boundary between covalently bonded atoms is the connection atom method

(110, 122), which uses a monovalent pseudoatom instead of a link atom. The parameters for the connection atom are optimized for the partitioned covalent bond. The connection atoms interact with the other QM atoms as a (specifically parameterized) QM atom, and with the other MM atoms as a standard carbon atom. This avoids the problem of a supplementary atom in the system, as the connection atom and the classical frontier atom are unified. However, the need to reparameterize for each type of covalent bond at a given level of quantum chemical theory is potentially laborious (123). The connection atom method has been developed for AM1 and PM3 (100), and DFT (122) QM/MM calculations. Tests indicated that it is more accurate than the standard link atom approach (100).

To overcome some of the problems that can arise with the single link atom method (e.g., an unphysical dipole), Brooks et al. have proposed a "double link atom" method (124). Also, their Gaussian delocalization method for MM atomic charges could simplify the calculation of energies and forces: e.g., even at short distances, the delocalized Gaussian MM method does not require the MM host atom charge to be excluded from the QM calculation, as would be necessary when treating it as a point charge. The delocalized Gaussian method can be combined with many QM/MM partitioning techniques, such as the link atom, frozen orbital, or pseudopotential methods. Tests of the delocalized Gaussian MM and double link atom methods on small model systems indicated that these methods gave better energetic properties than point atomic MM charge and single link atom methods.

Cui et al. have tested link atom QM/MM partitioning methods, for the SCC-DFTB QM method (125), including all the options available in the CHARMM program, which differ in their treatment of electrostatic interactions with the MM atoms close to the QM/MM frontier. They also proposed a divided frontier charge protocol, in which the partial charge associated with the MM atom bonded to the QM atom is distributed evenly on the other MM atoms in the same MM group. Tests of these various link atom schemes showed that QM/MM proton affinities and deprotonation energies are highly dependent on the particular link atom scheme employed: standard single link atom methods gave errors of up to 15–20 kcal/mol compared to pure QM calculations. Other schemes were found to give better results. Activation barriers and reaction energies were found, however, to be fairly insensitive to the choice of link atom scheme (e.g., within 2–4 kcal/mol) because of cancellation of errors. This is encouraging: the effect of using different link atom schemes in QM/MM simulations was found to be relatively small for chemical reactions in which the total charge does not change. Other technical details, such as the treatment of long-range electrostatics, are likely to play a more significant role in determining energetics generally, and should be treated carefully for reliable results.

2.6. Modelling Enzyme Reactions by Calculating Potential Energy Surfaces

With QM or QM/MM methods, potential energy surfaces of enzyme reaction mechanisms can be explored at a level of accuracy that can enable discrimination between different mechanisms: e.g., if the barrier for a proposed mechanism is significantly larger than that derived from experiment (using transition state theory), within the limits of accuracy of the computational method and experimental error, then that mechanism can be considered to be unlikely. A mechanism with a calculated barrier comparable to the apparent experimental barrier (for that step, or failing that for the overall reaction) is more likely. However, to calculate rate constants also requires reliable methods to calculate enthalpies, energies, and free energies of reaction and activation, given the potential energy surface. Traditional approaches to modelling reactions (e.g., in the gas phase) rely on the identification of stationary points (reactants, products, intermediates, transition states) via geometry optimization, followed by computation of second derivatives to enable relatively simplistic evaluation of zero-point corrections, thermal and entropy terms. Algorithms developed for small molecules are often not suitable for large systems: e.g., direct calculation, storage and manipulation of Hessian matrices becomes extremely difficult. A basic means of modelling approximate reaction paths is the "adiabatic mapping" or "coordinate driving" approach. The energy of the system is calculated by minimizing the energy at a series of fixed (or restrained, e.g., by harmonic forces) values of a reaction coordinate, e.g., the distance between two atoms. This approach has been applied with success to many enzymes (1, 2), but it is only valid if one conformation of the protein can represent the state of the system at a particular value of the reaction coordinate. A single minimum energy structure of this conformation may adequately represent the several closely related structures making up the reacting conformational state. Minimizing the QM/MM potential energy of such a representative conformation along the reaction coordinate should give a reasonable approximation of the enthalpic component of the potential of mean force (the free energy profile) for the reaction. In contrast, simple calculations of potential energy surfaces are likely to be unsuccessful or misleading for enzyme reactions involving large movements of charge or large changes in solvation (e.g., particularly for solvent–exposed sites, where rearrangement of water molecules might involve an unrealistically large energy penalty where adiabatic mapping calculations are used, (126)).

Due to the complexity of protein internal motions, many conformational substates exist, and a single structure might not be truly representative. If this is the case, calculations including extensive sampling of the system to obtain configurationally averaged free-energy changes are needed, as opposed to energy minimizations, which do not include entropic effects and are sensitive to starting geometries (126). A more simple approach to investigating

conformational effects is to use molecular dynamics simulations, and/or to use multiple different crystal structures, to generate multiple models for mechanism (e.g., adiabatic mapping) calculations, to ensure wide sampling of possible enzyme configurations (127), with averaging, or Boltzmann-weighted averaging, of energy barriers.

Despite the limitations and drawbacks of the adiabatic mapping approach, it has been applied in many QM/MM applications. It has the advantage that it is simple to apply, and does not require intensive calculations, such as second derivative evaluations, or simultaneous treatment of several points on a pathway. It can be useful for initial scans of potential energy surfaces, and for generating approximate models of transition states and intermediates, in which some allowance is made for structural relaxation to chemical changes at the active site. It is suitable only for reactions involving small chemical and structural changes, involving a small number of groups. For some enzymes, this type of approach has been validated through a correlation of calculated QM/MM barriers with activation energies derived from experiment (37).

As mentioned before, approaches based purely on calculations of a potential energy surface may not account for significant conformational fluctuations of the protein. Conformational changes, even on a small scale, may be coupled to, or significantly affect, chemical changes. Fluctuations of the active site can greatly affect the energy barrier. In the case of fatty acid amide hydrolase, conformational fluctuations do not affect the general shape of the potential energy surfaces, but consistency between experimental and calculated barriers is observed only with a specific (and rarely occurring) arrangement of the enzyme-substrate complex (49). These findings indicate that investigation of different protein conformations is essential for a meaningful determination of the energetics of enzymic reactions for calculations of potential energy profiles or surfaces.

2.7. Calculating Free Energy Profiles for Enzyme-Catalyzed Reactions

The rate constant of a reaction is actually related not to the potential energy barrier, but to the free energy barrier, according to transition state theory. The techniques above calculate potential energy barriers, for a particular conformation. Techniques that sample configurations along a reaction coordinate give a more sophisticated and extensive description, by taking account of multiple conformations and estimating entropic effects, and can be essential for modelling some types of enzyme reactions. Simulations of this type provide estimates of the free energy profile along a specific (reaction) coordinate, which is often referred to as the potential of mean force. Molecular dynamics and Monte Carlo methods in principle allow such sampling, but do not provide good sampling of high energy regions, such as in the vicinity of transition states. Conformational

sampling of processes of chemical change therefore requires specialized techniques, e.g., to bias the simulation to sample the transition state region. Umbrella sampling is such a method, which is widely used in molecular dynamics simulations e.g., with QM/MM techniques, to model enzymic reactions (27). In this technique, a biasing potential is applied to force the system to remain close to a specific value of a defined reaction coordinate. Often, an umbrella sampling simulation will begin with simulation of a transition state or reactant complex; an umbrella (e.g., harmonic) potential restrains the reaction coordinate to a value corresponding to e.g., the reactants. In other, subsequent simulations, the reference value of the restraint is changed by a small amount to sample other regions of the reaction coordinate. Often, the reaction coordinate is defined in terms of bond lengths, in which case a typical difference between the points would be 0.1–0.2 Å. The neighboring potentials should give overlapping distributions: this can be achieved by choosing an appropriate spacing of reaction coordinate values for different simulations, and an appropriate magnitude of the force constant of the restraint. The number of simulations is a balance between accuracy and efficiency. The reaction coordinate values during the (restrained) simulations are recorded. The effects of the restraining potentials are removed in the analysis and combined, typically by the weighted histogram analysis method (WHAM). This gives the unbiased potential of mean force along the reaction coordinate. It is important also to test for convergence with respect to length (and numbers) of simulations.

QM/MM umbrella sampling simulations are possible with low levels of QM theory, such as semiempirical molecular orbital methods (e.g., AM1 or PM3). Often, such methods are highly inaccurate for reaction barriers and energies. Their accuracy can be improved significantly by reparameterization for a specific reaction. For example, specifically parameterized semiempirical QM/MM methods have been used to investigate model reactions of glutathione-S-transferase (GST) enzymes. QM/MM umbrella sampling molecular dynamics simulations of the reaction of phenanthrene 9,10-oxide in a glutathione-S-transferase, identified a single amino acid as a likely determinant of stereospecificity in the epoxide ring opening (93). Similarly, specifically parameterized QM/MM methods have been applied to model the reaction between glutathione and 1-chloro-2,4-dinitrobenzene. The results of QM/MM umbrella sampling molecular dynamics simulations of this reaction in the M1-1 GST isoenzyme, in mutant enzymes, and in solution, agreed very well with experiment (128). QM/MM molecular dynamics simulations are much more computationally demanding than MM simulations, because of the computational expense of the evaluation of the QM forces. Typical QM/MM umbrella sampling applications have involved trajectories of picoseconds to nanoseconds (multiple simulations of 30 ps each, at each value of the

reaction coordinate, in the case of Bowman et al. (128)), with semiempirical QM methods. Approaches based on Monte Carlo simulations (129, 130) avoid the requirement for force calculations, and are also promising.

3. Notes

The choice of an appropriate method for the particular enzyme and questions of interest is vital. Careful testing and validation is important. Quantitative predictions of reaction rates or the effects of mutation remain very challenging, but for many enzymes, with appropriate methods, useful predictions can be made with some confidence. It is important to validate predictions from modelling by comparisons with experimental data. An example is comparison of calculated barriers for a series of alternative substrates with activation energies derived from experimental rates: demonstration of a correlation can validate mechanistic calculations (36, 131). Some enzymes have become important model systems in the development and testing of computational methods and protocols: these include chorismate mutase (1, 8, 10, 17, 27, 87), citrate synthase (3, 23, 132), $P450_{cam}$ (1, 2, 7, 11, 84), *para*-hydroxybenzoate hydroxylase (2, 17, 36), triosephosphate isomerase (5, 13, 102), fatty acid amide hydrolase (49, 95, 127, 129), and methylamine dehydrogenase (5, 7, 46, 47, 133).

Acknowledgments

AJM is an EPSRC Leadership Fellow.

References

1. Lonsdale R, Ranaghan KE, Mulholland AJ (2010) Computational enzymology. Chem Commun 46:2354–2372
2. Senn HM, Thiel W (2009) QM/MM methods for biomolecular systems. Angew Chem Int Ed Engl 48:1198–1229
3. van der Kamp MW, Mulholland AJ (2008) Computational enzymology: insight into biological catalysts from modelling. Nat Prod Rep 25:1001–1014
4. Mulholland AJ (2005) Modelling enzyme reaction mechanisms, specificity and catalysis. Drug Discov Today 10:1393–1402
5. Garcia-Viloca M, Gao J, Karplus M, Truhlar DG (2004) How enzymes work: analysis by modern rate theory and computer simulations. Science 303:186–195
6. Warshel A (2003) Computer simulations of enzyme catalysis: methods, progress, and insights. Annu Rev Biophys Biomol Struct 32:425–443
7. Friesner RA, Guallar V (2005) *Ab initio* quantum chemical and mixed quantum mechanics/molecular mechanics (QM/MM) methods for studying enzymatic catalysis. Annu Rev Phys Chem 56:389–427

8. Martí S, Roca M, Andrés J, Moliner V, Silla E, Tuñón I, Bertrán J (2004) Theoretical insights in enzyme catalysis. Chem Soc Rev 33:98–107
9. Himo F (2006) Quantum chemical modeling of enzyme active sites and reaction mechanisms. Theor Chem Acc 116:232–240
10. Warshel A, Sharma PK, Kato M, Xiang Y, Liu HB, Olsson MHM (2006) Electrostatic basis for enzyme catalysis. Chem Rev 106:3210–3235
11. Senn HM, Thiel W (2007) QM/MM studies of enzymes. Curr Opin Chem Biol 11:182–187
12. Mulholland AJ, Grant GH, Richards WG (1993) Computer modelling of enzyme catalysed reaction mechanisms. Protein Eng 6:133–147
13. Mulholland AJ, Karplus M (1996) Simulations of enzymic reactions. Biochem Soc Trans 24:247–254
14. Åqvist J, Warshel A (1993) Simulation of enzyme reactions using valence bond force fields and other hybrid quantum/classical approaches. Chem Rev 93:2523–2544
15. Warshel A, Levitt M (1976) Theoretical studies of enzymic reactions: dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. J Mol Biol 103:227–249
16. Scheiner S, Lipscomb WN (1976) Catalytic mechanism of serine proteinases. Proc Natl Acad Sci USA 73:432–436
17. Claeyssens F, Harvey JN, Manby FR, Mata RA, Mulholland AJ, Ranaghan KE, Schultz M, Thiel S, Thiel W, Werner H-J (2006) High accuracy computation of reaction barriers in enzymes. Angew Chem Int Ed 45:6856–6859
18. Mulholland AJ (2007) Chemical accuracy in QM/MM calculations on enzyme-catalysed reactions. Chem Cent J 1:19
19. Braun-Sand S, Olsson MHM, Warshel A (2005) Computer modeling of enzyme catalysis and its relationship to concepts in physical organic chemistry. Adv Phys Org Chem 40:201–245
20. Vocadlo DJ, Davies GJ, Laine R, Withers SG (2001) Catalysis by hen egg white lysozyme proceeds via a covalent intermediate. Nature 412:3835–3838
21. Bowman AL, Grant IM, Mulholland AJ (2008) QM/MM simulations predict a covalent intermediate in the hen egg white lysozyme reaction with its natural substrate. Chem Commun 7:4425–4427
22. Cleland WW, Frey PA, Gerlt JA (1998) The low barrier hydrogen bond in enzymatic catalysis. J Biol Chem 273:25529–25532
23. Mulholland AJ, Lyne PD, Karplus M (2000) *Ab initio* QM/MM study of the citrate synthase mechanism: a low-barrier hydrogen bond is not involved. J Am Chem Soc 122:534–535
24. Schutz CN, Warshel A (2004) The low barrier hydrogen bond (LBHB) proposal revisited: the case of the Asp · His pair in serine proteases. Proteins 55:711–723
25. Molina PA, Jensen JH (2003) A predictive model of strong hydrogen bonding in proteins: the Nδ1–H–Oδ1 hydrogen bond in low-pH α-chymotrypsin and α-lytic protease. J Phys Chem B 107:6226–6233
26. Hur S, Bruice TC (2003) Just a near attack conformer for catalysis (chorismate to prephenate rearrangements in water, antibody, enzymes, and their mutants). J Am Chem Soc 125:10540–10542
27. Ranaghan KE, Mulholland AJ (2004) Conformational effects in enzyme catalysis: QM/MM free energy calculation of the 'NAC' contribution in chorismate mutase. Chem Commun 10:1238–1239
28. Olsson MH, Parson WW, Warshel A (2006) Dynamical contributions to enzyme catalysis: critical tests of a popular hypothesis. Chem Rev 106:1737–1756
29. Hammes-Schiffer S, Watney JB (2006) Hydride transfer catalyzed by Escherichia coli and Bacillus subtilis dihydrofolate reductase: Coupled motions and distal mutations. Phil Trans Roy Soc B 361:1365–1373
30. Masgrau L, Roujeinikova A, Johannissen LO, Hothi P, Basran J, Ranaghan KE, Mulholland AJ, Sutcliffe MJ, Scrutton NS, Leys D (2006) Atomic description of an enzyme reaction dominated by proton tunneling. Science 312:237–241
31. Hatcher E, Soudackov AV, Hammes-Schiffer S (2007) Proton-coupled electron transfer in soybean lipoxygenase: dynamical behavior and temperature dependence of kinetic isotope effects. J Am Chem Soc 129:187–196
32. Limbach HH, Lopez JM, Kohen A (2006) Arrhenius curves of hydrogen transfers: tunnel effects, isotope effects and effects of preequilibria. Phil Trans Roy Soc B 361:1399–1415
33. Nagel ZD, Klinman JP (2006) Tunneling and dynamics in enzymatic hydride transfer. Chem Rev 106:3095–3118

34. Villa J, Strajbl M, Glennon TM, Sham YY, Chu ZT, Warshel A (2000) How important are entropic contributions to enzyme catalysis? Proc Natl Acad Sci USA 97:11899–11904
35. Lonsdale R, Harvey JN, Manby FR, Mulholland AJ (2011) Comment on "A stationary-wave model of enzyme catalysis" by Carlo Canepa. J Comput Chem 32:368–369
36. Ridder L, Harvey JN, Rietjens IMCM, Vervoort J, Mulholland AJ (2003) Ab initio QM/MM modeling of the hydroxylation step in *p*-hydroxybenzoate hydroxylase. J Phys Chem B 107:2118–2126
37. Ridder L, Mulholland AJ, Rietjens IMCM, Vervoort J (2000) A quantum mechanical/molecular mechanical study of the hydroxylation of phenol and halogenated derivatives by phenol hydroxylase. J Am Chem Soc 122:8728–8738
38. Bjelic S, Åqvist J (2004) Prediction of structure, substrate binding mode, mechanism, and rate for a malaria protease with a novel type of active site. Biochemistry 43:14521–14528
39. Zhang YK, Kua J, McCammon JA (2003) Influence of structural fluctuation on enzyme reaction energy barriers in combined quantum mechanical/molecular mechanical studies. J Phys Chem B 107:4459–4463
40. Gao JL, Truhlar DG (2002) Quantum mechanical methods for enzyme kinetics. Annu Rev Phys Chem 53:467–505
41. Karplus M, Gao YQ, Ma JP, van der Vaart A, Yang W (2005) Protein structural transitions and their functional role. Phil Trans R Soc A 363:331–355
42. Fersht A (1999) Structure and mechanism in protein science. A guide to enzyme catalysis and protein folding. Freeman, New York
43. Olsson MHM, Warshel A (2004) Solute solvent dynamics and energetics in enzyme catalysis: the S(N)2 reaction of dehalogenase as a general benchmark. J Am Chem Soc 126:15167–15179
44. Wolf-Watz M, Thai V, Henzler-Wildman K, Hadjipaylou G, Eisenmesser EZ, Kern D (2004) Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair. Nat Struct Mol Biol 11:945–949
45. Kohen A, Cannio R, Bartolucci S, Klinman JP (1999) Enzyme dynamics and hydrogen tunneling in a thermophilic alcohol dehydrogenase. Nature 399:496–499
46. Sutcliffe MJ, Masgrau L, Roujeinikova A, Johannissen LO, Hothi P, Basran J, Ranaghan KE, Mulholland AJ, Leys D, Scrutton NS (2006) Hydrogen tunnelling in enzyme-catalysed H-transfer reactions: flavoprotein and quinoprotein systems. Phil Trans Roy Soc B 361:1375–1386
47. Ranaghan KE, Mulholland AJ (2010) Computer simulations of quantum tunnelling in enzyme-catalysed hydrogen transfer reactions. Interdiscip Sci 2:78–97
48. Pentikäinen U, Pentikäinen OT, Mulholland AJ (2008) Cooperative symmetric to asymmetric conformational transition of the apoform of scavenger decapping enzyme revealed by simulations. Proteins 70:498–508
49. Lodola A, Mor M, Zurek J, Tarzia G, Piomelli D, Harvey JN, Mulholland AJ (2007) Conformational effects in enzyme catalysis: reaction via a high energy conformation in fatty acid amide hydrolase. Biophys J 92: L20–L22
50. Case DA, Cheatham TE III, Darden T, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling C, Wang B, Woods R (2005) The AMBER biomolecular simulation programs. J Comput Chem 26:1668–1688 (see http://amber.scripps.edu/)
51. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M (1983) CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. J Comput Chem 4:187–217, (See also http://www.charmm.org)
52. Scott WRP, Hunenberger PH, Tironi IG, Mark AE, Billeter SR, Fennen J, Torda AE, Huber T, Kruger P, van Gunsteren WF (1999) The GROMOS biomolecular simulation program package. J Phys Chem A 103:3596–3607, (see http://www.igc.ethz.ch/gromos/gromos.html)
53. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K (2005) Scalable molecular dynamics with NAMD. J Comput Chem 26:1781–1802, (see http://www.ks.uiuc.edu/Research/namd/)
54. Ponder JW, Richards FM (1987) An efficient newton-likemethod for molecular mechanics energy minimization of large molecules. J Comput Chem 8:1016–1024, (see http://dasher.wustl.edu/tinker/)
55. Jorgensen WL, Maxwell DS, Tirado-Rives J (1996) Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. J Am Chem Soc 118:11225–11236
56. Kaminski GA, Friesner RA, Tirado-Rives J, Jorgensen WL (2001) Evaluation and reparametrization of the OPLSAA force field for proteins via comparison with accurate quantum chemical calculations on peptides. J Phys Chem B 105:6474–6487

57. MacKerell AD Jr, Bashford D, Bellott M, Dunbrack RL Jr, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE III, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiórkiewicz-Kuczera J, Yin D, Karplus M (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B 102:3586–3616
58. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995) A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic molecules. J Am Chem Soc 117:5179–5197
59. Mackerell AD (2005) Empirical force fields for proteins: current status and future directions. Ann Rep Comp Chem 1:91–111
60. Ponder JW, Case DA (2003) Force fields for protein simulations. Adv Protein Chem 66:27–75
61. Cheatham TE III (2005) Molecular modeling and atomistic simulation of nucleic acids. Ann Rep Comp Chem 1:75–90
62. Cheatham TE III (2004) Simulation and modeling of nucleic acid structure, dynamics and interactions. Curr Opin Struct Biol 14:360–367
63. Foloppe N, MacKerell AD (2000) All-atom empirical force field for nucleic acids: I parameter optimization based on small molecule and condensed phase macromolecular target data. J Comput Chem 21:86–104
64. MacKerell AD, Banavali NK (2000) All-atom empirical force field for nucleic acids: II application to molecular dynamics simulations of DNA and RNA in solution. J Comput Chem 21:105–120
65. Cheatham TE, Cieplak P, Kollman PA (1999) A modified version of the Cornell et al. force field with improved sugar pucker phases and helical repeat. J Biomol Struct Dyn 16:845–862
66. Feller SE, Yin DX, Pastor RW, MacKerell AD (1997) Molecular dynamics simulation of unsaturated lipid bilayers at low hydration: parameterization and comparison with diffraction studies. Biophys J 73:2269–2279
67. Jorgensen WL, Tirado-Rives J (2005) Potential energy functions for atomic-level simulations of water and organic and biomolecular systems. Proc Natl Acad Sci USA 102:6665–6670
68. Kim BC, Young T, Harder E, Friesner RA, Berne BJ (2005) Structure and dynamics of the solvation of bovine pancreatic trypsin inhibitor in explicit water: a comparative study of the effects of solvent and protein polarizability. J Phys Chem B 109:16529–16538
69. Shaw KE, Woods CJ, Mulholland AJ (2010) Compatibility of quantum chemical methods and empirical (MM) water models in quantum mechanics/molecular mechanics liquid water simulations. J Phys Chem Lett 1:219–223
70. Lim D, Jenson J, Repasky MP, Jorgensen WL (1999) Solvent as catalyst: computational studies of organic reactions in solution. In: Truhlar DG, Morokuma K (eds) Transition state modeling for catalysis. American Chemical Society, Washington DC
71. Bentzien J, Muller RP, Florian J, Warshel A (1998) Hybrid *ab initio* quantum mechanics molecular mechanics calculations of free energy surfaces for enzymatic reactions: the nucleophilic attack in subtilisin. J Phys Chem B 102:2293–2301
72. Warshel A (1997) Computer modeling of chemical reactions in enzymes and solutions. John Wiley & Sons, New York
73. Villa J, Warshel A (2001) Energetics and dynamics of enzymatic reactions. J Phys Chem B 105:7887–7907
74. Warshel A, Sharma PK, Chu ZT, Åqvist J (2007) Electrostatic contributions to binding of transition state analogues can be different from the corresponding contributions to catalysis: phenolates binding to the oxyanion hole of ketosteroid isomerase. Biochemistry 46:1466–1476
75. Bjelic S, Åqvist J (2006) Catalysis and linear free energy relationships in aspartic proteases. Biochemistry 45:7709–7723
76. Trobro S, Åqvist J (2005) Mechanism of peptide bond synthesis on the ribosome. Proc Natl Acad Sci USA 102:12395–12400
77. Sharma PK, Xiang Y, Katom M, Warshel A (2005) What are the roles of substrate-assisted catalysis and proximity effects in peptide bond formation by the ribosome? Biochemistry 44:11307–11314
78. Hammes-Schiffer S (2004) Quantum-classical simulation methods for hydrogen transfer in enzymes: a case study of dihydrofolate reductase. Curr Opin Struct Biol 14:192–201
79. Liu HB, Warshel A (2007) Origin of the temperature dependence of isotope effects in enzymatic reactions: the case of dihydrofolate reductase. J Phys Chem B 111:7852–7861

80. Sharma PK, Chu ZT, Olsson MHM, Warshel A (2007) A new paradigm for electrostatic catalysis of radical reactions in vitamin B12 enzymes. Proc Natl Acad Sci USA 2007 (104):9661–9666
81. Van derVaart A, Gogonea V, Dixon SL, Merz KM (2000) Linear scaling molecular orbital calculations of biological systems using the semiempirical divide and conquer method. J Comput Chem 21:1494–1504
82. Khandogin J, York DM (2004) Quantum descriptors for biological macromolecules from linear-scaling electronic structure methods. Proteins 56:724–737
83. Khandogin J, Musier-Forsyth K, York DM (2003) Insights into the regioselectivity and RNA-binding affinity of HIV-1 nucleocapsid protein from linear-scaling quantum methods. J Mol Biol 330:993–1004
84. Lonsdale R, Harvey JN, Mulholland AJ (2010) Inclusion of dispersion effects significantly improves accuracy of calculated reaction barriers for cytochrome p450 catalyzed reactions. J Phys Chem Lett 1:3232–3237
85. Himo F, Siegbahn PE (2003) Quantum chemical studies of radical-containing enzymes. Chem Rev 103:2421–2456
86. Siegbahn PE, Himo F (2009) Recent developments of the quantum chemical cluster approach for modeling enzyme reactions. J Biol Inorg Chem 14:643–651
87. Woodcock HL, Hodoscek M, Sherwood P, Lee YS, Schaefer HF, Brooks BR (2003) Exploring the quantum mechanical/molecular mechanical replica path method: a pathway optimization of the chorismate to prephenate Claisen rearrangement catalyzed by chorismate mutase. Theor Chem Acc 109:140–148
88. Field MJ, Bash PA, Karplus M (1990) A combined quantum-mechanical and molecular mechanical potential for molecular-dynamics simulations. J Comput Chem 11:700–733
89. Lyne PD, Hodoscek M, Karplus M (1999) A hybrid QM–MM potential employing Hartree–Fock or density functional methods in the quantum region. J Phys Chem A 103:3462–3471
90. Cui Q, Elstner M, Kaxiras E, Frauenheim T, Karplus M (2001) A QM/MM implementation of the self-consistent charge density functional tight binding (SCC-DFTB) method. J Phys Chem B 105:569–585
91. Marti S, Moliner V (2005) Improving the QM/MM description of chemical processes: a dual level strategy to explore the potential energy surface in very large systems. J Chem Theory Comput 1:1008–1016
92. Prat-Resina X, Bofill JM, Gonzalez-Lafont A, Lluch JM (2004) Geometry optimization and transition state search in enzymes: different options in the microiterative method. Int J Quantum Chem 98:367–377
93. Ridder L, Rietjens IMCM, Vervoort J, Mulholland AJ (2002) Quantum mechanical/molecular mechanical free energy simulations of the glutathione S-transferase (M1-1) reaction with phenanthrene 9,10-oxide. J Am Chem Soc 124:9926–9936
94. Ranaghan KE, Ridder L, Szefczyk B, Sokalski WA, Hermann JC, Mulholland AJ (2004) Transition state stabilization and substrate strain in enzyme catalysis: ab initio QM/MM modelling of the chorismate mutase reaction. Org Biomol Chem 2:968–980
95. Lodola A, Mor M, Hermann JC, Tarzia G, Piomelli D, Mulholland AJ (2005) QM/MM modelling of oleamide hydrolysis in fatty acid amide hydrolase (FAAH) reveals a new mechanism of nucleophile activation. Chem Commun 35:4399–4401
96. Riccardi D, Schaefer P, Cui Q (2005) pK_a calculations in solution and proteins with QM/MM free energy perturbation simulations: a quantitative test of QM/MM protocols. J Phys Chem B 109:17715–17733
97. Raha K, Merz KM (2004) A quantum mechanics-based scoring function: study of zinc ion-mediated ligand binding. J Am Chem Soc 126:1020–1021
98. Bakowies D, Thiel W (1996) Hybrid models for combined quantum mechanical and molecular mechanical approaches. J Phys Chem 100:10580–10594
99. Svensson M, Humbel S, Froese RDJ, Matsubara T, Sieber S, Morokuma K (1996) ONIOM: A multilayered Integrated MO + MM Method for geometry optimizations and single point energy predictions. A test for Diels – Alder reactions and Pt(P(t-Bu)$_3$)$_2$ + H_2 oxidative addition. J Phys Chem 100:19357–19363
100. Antes I, Thiel W (1999) Adjusted connection atoms for combined quantum mechanical and molecular mechanical methods. J Phys Chem A 103:9290–9295
101. Jensen L, van Duijnen PT (2005) The first hyperpolarizability of p-nitroaniline in 1,4-dioxane: a quantum mechanical/molecular mechanics study. J Chem Phys 123:Art. No. 074307
102. Cui Q, Karplus M (2002) Quantum mechanics/molecular mechanics studies of triosephosphate isomerase-catalyzed reactions: effect of geometry and tunneling on proton-transfer

rate constants. J Am Chem Soc 124:3093–3124

103. Riccardi D, Li GH, Cui Q (2004) Importance of van der Waals interactions in QM/MM simulations. J Phys Chem B 108:6467–6478
104. Poulsen TD, Garcia-Viloca M, Gao JL, Truhlar DG (2003) Free energy surface, reaction paths, and kinetic isotope effect of short-chain acyl-coa dehydrogenase. J Phys Chem B 107:9567–9578
105. Brooks CL III, Karplus M, Pettitt BM (1988) Proteins, a theoretical perspective of dynamics, structure and thermodynamics. Wiley, New York
106. Nam K, Gao JL, York DM (2005) An efficient linear-scaling ewald method for long-range electrostatic interactions in combined QM/MM calculations. J Chem Theory Comput 1:2–13
107. Schaefer P, Riccardi D, Cui Q (2005) Reliable treatment of electrostatics in combined QM/MM simulation of macromolecules. J Chem Phys 123:Art. No. 014905
108. Im W, Berneche S, Roux B (2001) Generalized solvent boundary potentials for computer simulations. J Chem Phys 114:2924–2937
109. Monard G, Loos M, Théry V, Baka K, Rivail J-L (1996) Hybrid classical quantum force field for modelling very large molecules. Int J Quantum Chem 58:153–159
110. Assfeld X, Rivail J-L (1996) Quantum chemical computations on parts of large molecules: The *ab initio* local self consistent field method. Chem Phys Lett 263:100–106
111. Gao J, Amara P, Alhambra C, Field MJ (1998) Method for the treatment of boundary atoms in combined QM/MM calculations. J Phys Chem A 102:4714–4721
112. Ferre N, Assfeld X, Rivail J-L (2002) Specific force field parameters determination for the hybrid ab initio QM/MM LSCF method. J Comput Chem 23:610–624
113. Antonczak S, Monard G, Ruiz-Lopez MF, Rivail J-L (1998) Modeling of Peptide Hydrolysis by Thermolysin. A Semiempirical and QM/MM Study. J Am Chem Soc 120:8825–8833
114. Garcia-Viloca M, Gao JL (2004) Generalized hybrid orbital for the treatment of boundary atoms in combined quantum mechanical and molecular mechanical calculations using the semiempirical parameterized model 3 method. Theor Chem Acc 111:280–286
115. Wesolowski TA, Warshel A (1993) Frozen density functional approach for *ab-initio* calculations of solvated molecules. J Phys Chem 97:8050–8053
116. Pu JZ, Gao JL, Truhlar DG (2004) Generalized hybrid orbital (GHO) method for combining ab initio hartree—fock wave functions with molecular mechanics. J Phys Chem A 108:632–650
117. Pu JZ, Gao JL, Truhlar DG (2004) Combining self-consistent-charge density-functional tight-binding (SCC-DFTB) with molecular mechanics by the generalized hybrid orbital (GHO) Method. J Phys Chem A 108:5454–5463
118. Pu JZ, Gao JL, Truhlar DG (2005) Generalized hybrid-orbital method for combining density functional theory with molecular mechanicals. Chemphyschem 6:1853–1865
119. Amara P, Field MJ (2003) Evaluation of an *ab-initio* quantum mechanical molecular mechanical hybrid-potential link-atom method. Theor Chem Acc 109:43–52
120. HyperChem Users Manual (2002) HyperCube, Inc: Waterloo, Ontario, Canada
121. Reuter N, Dejaegere A, Maigret B, Karplus M (2000) Frontier bonds in QM/MM methods: a comparison of different approaches. J Phys Chem A 104:1720–1735
122. Zhang Y, Lee T-S, Yang W (1999) A pseudobond approach to combining quantum mechanical and molecular mechanical methods. J Chem Phys 110:46–54
123. Monard G, Prat-Resina X, Gonzalez-Lafont A, Lluch JM (2003) Determination of enzymatic reaction pathways using QM/MM methods. Int J Quant Chem 93:229–244
124. Das D, Eurenius KP, Billings EM, Sherwood P, Chatfield DC, Hodoscek M, Brooks BR (2002) Optimization of quantum mechanical molecular mechanical partitioning schemes: Gaussian delocalization of molecular mechanical charges and the double link atom method. J Chem Phys 117:10534–10547
125. Konig PH, Hoffmann M, Frauenheim T, Cui Q (2005) A critical evaluation of different QM/MM frontier treatments with SCC-DFTB as the QM method. J Phys Chem B 109:9082–9095
126. Klahn M, Braun-Sand S, Rosta E, Warshel A (2005) On possible pitfalls in ab initio quantum mechanics/molecular mechanics minimization approaches for studies of enzymatic reactions. J Phys Chem B 109:15645–15650
127. Lodola A, Sirirak J, Fey N, Rivara S, Mor M, Mulholland AJ (2010) Structural fluctuations in enzyme-catalyzed reactions: determinants

of reactivity in fatty acid amide hydrolase from multivariate statistical analysis of quantum mechanics/molecular mechanics paths. J Chem Theor Comput 6:2948–2960

128. Bowman AL, Ridder L, Rietjens IMCM, Vervoort J, Mulholland AJ (2007) Molecular determinants of xenobiotic metabolism: QM/MM simulation of the conversion of 1-chloro-2,4-dinitrobenzene catalyzed by M1-1 glutathione S-transferase. Biochemistry 46:6353–6363
129. Acevedo O, Jorgensen WL (2010) Advances in quantum and molecular mechanical (QM/MM) simulations for organic and enzymatic reactions. Acc Chem Res 43:142–151
130. Woods CJ, Manby FR, Mulholland AJ (2008) An efficient method for the calculation of quantum mechanics/molecular mechanics free energies. J Chem Phys 128:Art. No. 014109
131. Ridder L, Mulholland AJ (2003) Modeling biotransformation reactions by combined quantum mechanical/molecular mechanical approaches: from structure to activity. Curr Top Med Chem 3:1241–1256
132. van der Kamp MW, Zurek J, Manby FR, Harvey JN, Mulholland AJ (2010) Testing high-level QM/MM methods for modeling enzyme reactions: acetyl-CoA deprotonation in citrate synthase. J Phys Chem B 114:11303–11314
133. Ranaghan KE, Masgrau L, Scrutton NS, Sutcliffe MJ, Mulholland AJ (2007) Analysis of classical and quantum paths for deprotonation of methylamine by methylamine dehydrogenase. Chemphyschem 8:1816–1835