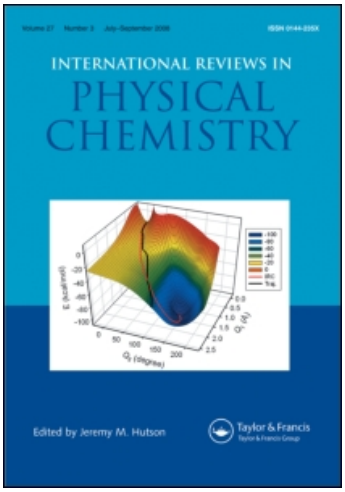

International Reviews in Physical Chemistry
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713724383

Investigations of enzyme-catalysed reactions with combined quantum mechanics/molecular mechanics (QM/MM) methods

Kara E. Ranaghan[a]; Adrian J. Mulholland
[a] School of Chemistry, Centre for Computational Chemistry, University of Bristol, Bristol BS8 1TS, UK

Online publication date: 29 January 2010

To cite this Article Ranaghan, Kara E. and Mulholland, Adrian J.(2010) 'Investigations of enzyme-catalysed reactions with combined quantum mechanics/molecular mechanics (QM/MM) methods', International Reviews in Physical Chemistry, 29: 1, 65 — 133
To link to this Article: DOI: 10.1080/01442350903495417
URL: http://dx.doi.org/10.1080/01442350903495417

International Reviews in Physical Chemistry
Vol. 29, No. 1, January–March 2010, 65–133

Taylor & Francis
Taylor & Francis Group

Investigations of enzyme-catalysed reactions with combined quantum mechanics/molecular mechanics (QM/MM) methods

Kara E. Ranaghan and Adrian J. Mulholland*

School of Chemistry, Centre for Computational Chemistry, University of Bristol, Bristol BS8 1TS, UK

(*Received 22 September 2009; final version received 16 November 2009*)

Combined quantum mechanics/molecular mechanics (QM/MM or QM-MM) methods are an excellent approach for modelling the mechanisms of enzyme-catalysed reactions. QM/MM methods allow detailed modelling of reactions in enzymes by coupling quantum chemical calculations on the active site with a simpler, empirical 'molecular mechanics' treatment of the rest of the protein. Possible reaction mechanisms can be compared and catalytic interactions analysed. QM/MM calculations can now be carried out for enzyme-catalysed reactions with quantum chemical methods of potentially very high accuracy. More approximate QM methods can allow extensive molecular simulations (e.g. molecular dynamics or Monte Carlo simulations). In this review, QM/MM techniques are outlined and some recent applications to enzyme-catalysed reactions are discussed.

Keywords: chorismate mutase; quantum tunnelling; cytochrome P450 compound I; lysozyme; β-lactamases; HIV; computational enzymology

Contents

*Corresponding author. Email: Adrian.Mulholland@bristol.ac.uk

ISSN 0144–235X print/ISSN 1366–591X online

DOI: 10.1080/01442350903495417
http://www.informaworld.com

1. Introduction

The complexity of enzymes and the broad spectrum of reactions they catalyse has led to many long and heated debates about their mechanisms and the underlying catalytic principles. The technical difficulties of studying enzyme reactions often make these arguments difficult to resolve by experiment alone. Combined quantum mechanics/molecular mechanics (QM/MM, QM-MM or MM/QM) methods are a central tool in the growing area of computational enzymology [1–9]. The QM treatment of the electronic structure of a small active site region allows chemical reactions to be modelled, including the effects of the environment through an empirical MM treatment. Enzyme mechanisms can be tested, transition states (TSs) and reactive species identified and characterised and the effects of mutations investigated. The dynamics of proteins can also be simulated, identifying conformational changes that may be associated with reaction. There have been many controversial proposals seeking to explain catalysis in some or all enzymes, including 'low-barrier' hydrogen bonds [10–12], the so-called 'near attack conformations' [13–17], the role of enzyme dynamics in catalysis [18,19], quantum tunnelling [20–23], reorganisation energy [5] and entropic effects [24]. QM/MM methods can help provide evidence to support these proposals beyond that provided by experimental techniques.

The work of many groups has been influential in the development of QM/MM techniques and their application. Warshel and Levitt's [25] study of lysozyme was pioneering in this area. The early *ab initio* QM/MM implementation by Singh and Kollman [26] is notable. The semiempirical QM/MM method developed by Field *et al.* [27] was a significant development: based on a semiempirical molecular orbital QM method (AM1 [28]), it made the study of enzyme reactions practical being significantly less computationally demanding than *ab initio* methods, and implemented in the widely used CHARMM [29] simulation package. Developments in quantum chemical methods, work on the implementation of QM/MM methods and computer resources in the past two decades have led to great advances in the field. QM/MM calculations on enzyme-catalysed reactions at the *ab initio* and density functional levels are nowadays much more feasible. It is now possible to achieve an unprecedented level of accuracy in QM/MM calculations on enzyme-catalysed reaction: high-level correlated *ab initio* quantum chemical methods can now be used in QM/MM calculations to study enzymic reactions allowing the calculations of energy barriers for enzyme-catalysed reactions, in the best cases, of potentially near 'chemical accuracy' ($1\,\mathrm{kcal\,mol^{-1}}$) [30]. The number of publications in this area has also increased significantly over the past two decades. A search of the ISI Web of Science (v4.6: http://wok.mimas.ac.uk) using the criteria: Topic = (qm/mm or mm/qm or qm-mm or mm-qm or qm:mm or mm:qm or quantum mechanics/molecular mechanics or quantum mechanical/molecular mechanical) and Topic = (enzyme) yields 560 records, growing from 6 publications in 1991 to 82 in 2008 and 64 in 2009 thus far (Figure 1). Combined, these papers have been cited 10,649 times with an average of 19.02 citations each. The number of citations per year of papers in this area has also grown significantly from ~200 in 2000 to ~1800 so far in 2009, showing the increasing popularity of QM/MM methods.

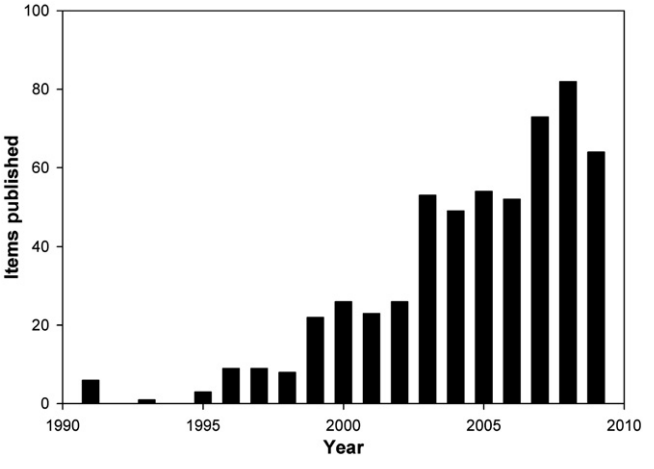

Figure 1. Graph indicating the results of a search on ISI Web of Science (v4.6: http://wok.mimas.ac.uk) using the criteria: Topic = (qm/mm or mm/qm or qm-mm or mm-qm or qm:mm or mm:qm or quantum mechanics/molecular mechanics or quantum mechanical/molecular mechanical) and Topic = (enzyme) sorted by publication year.

In this review, the theoretical basis of QM/MM methods is outlined, and important practical aspects of model preparation and simulation are discussed. Recent examples of QM/MM studies are used to illustrate how QM/MM methods have advanced our understanding of enzyme-catalysed reactions. Topics discussed include how QM/MM studies have contributed to understanding enzymic reactions, including calling into question 'textbook' mechanisms; QM/MM modelling has provided an insight into the basis of catalysis (e.g. in chorismate mutase), the role of quantum tunnelling in enzyme reactions, predictions of drug metabolism and detoxification and features of enzymes important in human diseases. Quantum chemical modelling of small active site models [31–33] and empirical valence bond (EVB) techniques [3,18,34,35] have also contributed significantly to the understanding of enzyme-catalysed reactions, but are beyond the scope of this review.

2. Combined QM/MM methods

The theoretical basis of QM/MM methods has been discussed in detail, e.g. in [9,25,27, 36,37], and only the main aspects are outlined here. The system is divided into a (small) QM region and an (usually much larger) MM region (Figure 2). The total energy of a QM/MM system can be expressed as

$$E_{\mathrm{TOTAL}} = E_{\mathrm{QM}} + E_{\mathrm{MM}} + E_{\mathrm{QM/MM}} + E_{\mathrm{Boundary}} \quad (1)$$

E_{QM} and E_{MM} are the energies of the QM and MM regions, respectively, calculated in a standard way at those levels. For example, the MM energy is defined by the potential of the force field used. Current standard MM force fields for proteins treat atoms as point charges with van der Waals radii, which together determine the non-bonded interactions; force fields also include terms to describe interactions between bonded atoms, e.g. bond, angle and dihedral terms. $E_{\mathrm{QM/MM}}$ describes the interaction between the QM and MM regions, and can be treated in various ways, as described further. E_{Boundary} is a term to account for the fact that the system may have to be truncated for practical reasons (e.g. QM/MM calculations are computationally demanding, and it may not be possible or desirable to include the whole of an enzyme molecule). This term is included to represent the effects of any omitted parts of the system and bulk solution. Periodic boundary [38], stochastic boundary [39,40] or generalised solvent boundary potential (GSBP) [41,42] conditions can be applied. Broadly speaking, there are two particularly important considerations in the interaction between the QM and MM regions: (i) the treatment of non-covalent interactions between the QM and MM regions and (ii) where covalent bonds exist between atoms in the QM and MM regions, the treatment of bonds at the frontier between the two regions. These aspects of QM/MM methods are discussed below.

2.1. *Coupling between the QM and MM region*

Different QM/MM methods can differ not only in the QM and MM methods applied, but also in the treatment of interactions between the QM and MM regions. A useful classification scheme for QM/MM methods has been proposed by Bakowies and Thiel [43], defining four model types A–D, which show increasing levels of sophistication.

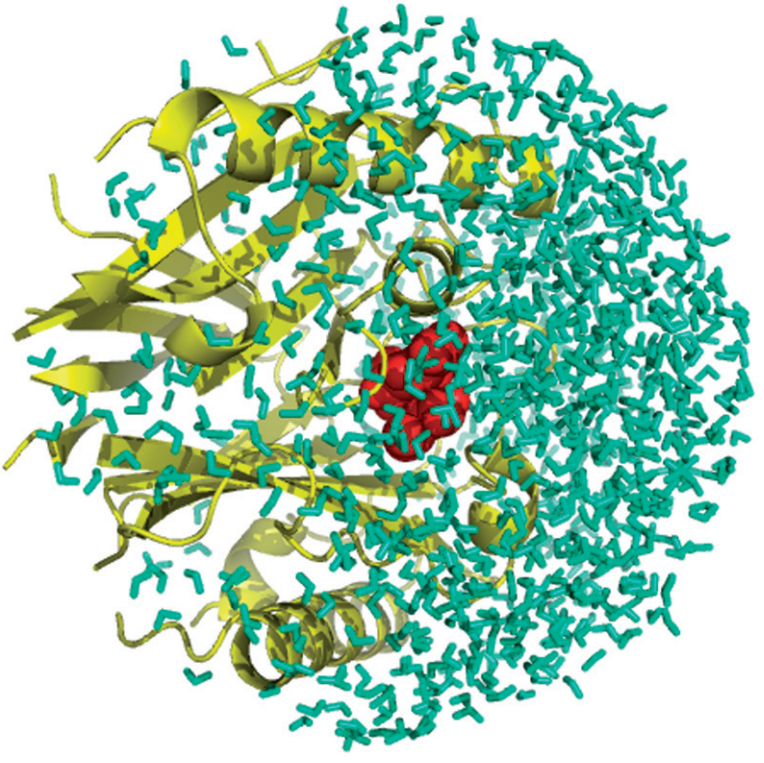

Figure 2. An enzyme (chorismate mutase [210]) partitioned into a QM region (shown as red spheres) and an MM region consisting of protein (yellow cartoons) and solvent (cyan sticks). Here the system is truncated to an approximate sphere (in this case with radius 25 Å), typical of the approach used in many QM/MM simulations of enzyme-catalysed reaction mechanisms (e.g. applying 'stochastic boundary' [39,40] MD).

Type A: QM/MM implementations of type A use the simplest linking of QM and MM methods, involving a straightforward 'mechanical' embedding of the QM region in the MM environment. The interactions between the QM and MM regions are treated purely classically by MM, i.e. the QM system is represented by (MM) point charges in its interaction with the MM environment, meaning the polarisation of the QM region is not included.

Type B: Electrostatic interactions between the QM and MM regions are included in type B models using invariant atomic partial point charges in the MM region (from the MM force field) that enter the core QM Hamiltonian. Thus, the QM region is polarised by its MM environment.

Type C: Models of type C extend beyond type B by also including some polarisation of atoms in the MM region by the electric field generated by the QM region (compare to type D models below).

Type D: Models of type D are the most refined and sophisticated class of QM/MM method. They include the iterative, self-consistent mutual polarisation of the MM and QM regions.

Type A is the most straightforward implementation of a combined QM/MM model, but in this type of QM/MM approach the QM and MM regions do not interact in the quantum chemical calculation, all interactions are treated at the low (MM) level. The QM/MM energy of the whole system, $E_{\mathrm{TOTAL}}^{\mathrm{QM/MM}}$, is calculated in a simple subtractive scheme:

$$E_{\mathrm{TOTAL}}^{\mathrm{QM/MM}} = E_{\mathrm{TOTAL}}^{\mathrm{MM}} + E_{\mathrm{QM\ region}}^{\mathrm{QM}} - E_{\mathrm{QM\ region}}^{\mathrm{MM}} \quad (2)$$

where $E_{\mathrm{TOTAL}}^{\mathrm{MM}}$ is the MM energy of the whole system, $E_{\mathrm{QM\ region}}^{\mathrm{QM}}$, is the QM energy of the QM region and $E_{\mathrm{QM\ region}}^{\mathrm{MM}}$ is the MM energy of the isolated QM region. This type of simple subtractive approach can be applied to all combinations of theory levels, i.e. QM/QM methods. It forms the basis for the multi-layer 'Our own N-layered Integrated molecular Orbital and molecular Mechanics method' (ONIOM) method [44,45]. This method allows the partitioning of a system into several layers, which may include one or more QM regions treated by different quantum chemical methods and with the rest of the system described by MM. The ONIOM method has also been developed to allow the polarisation of the QM region, through 'electrostatic embedding' [46,47].

The most widely used QM/MM methods are of type B allowing the polarisation of the QM region by the MM region, including the charges of the MM groups in the QM calculation. The electronic structure calculation therefore includes the effects of the MM atoms. MM force fields developed to describe biological molecules such as PARM99 [48] implemented in the AMBER molecular modelling program [49], whereas CHARMM22 [50] and OPLS-AA [51,52] do not allow for changes in charge distribution and the more sophisticated coupling schemes of type C and D, where the MM region is also polarised by the QM region, remain difficult to implement and are a subject of ongoing research [53–64].

The interaction between the QM and MM regions can be split into bonded and non-bonded interactions. Bonded interactions between QM and MM atoms are treated at the lower (MM) level, and are applied whenever at least one QM atom is involved in the bond, bond angle or dihedral. If a molecule itself is divided into QM and MM parts, then the frontier bond must be terminated using a suitable method, e.g. the use of a link atom (see below). Non-bonded interactions can be further separated into electrostatic and non-electrostatic interactions. The electrostatic interactions are accounted for by the point charges of the MM atoms, and the non-electrostatic interactions are included with a van der Waals term. In *ab initio* QM/MM methods, the MM point charges are included through one-electron integrals and the interaction of the classical charges with the nuclei of the QM system. Semiempirical molecular orbital methods such as AM1 or PM3 treat only valence electrons explicitly, while the inner electrons of an atom are combined with the nucleus in an invariant positive 'core'. They require a different treatment of the charge-core interactions, and several expressions have been applied. For example, the AM1-CHARMM method of Field *et al.* [27] treats QM/MM electrostatic interactions by including the point charges of MM atoms as atomic 'cores'. The resulting core–core interactions are treated by the same formalism used in the MNDO method [65]. Thus, all MM atoms need two QM parameters: α_{M} used in the core–core interaction term and ρ_0 the monopole distribution of the MM term. Values of $\alpha_{\mathrm{M}} = 5$ and $\rho_0 = 0.0$ were found to be suitable for all MM atoms [27], so only the partial charges and van der Waals parameters differentiate the different types of MM atom in their interaction with the

QM region. The QM/MM van der Waals interactions are included using Lennard–Jones terms, in the same way as if they were between MM atoms, and as a result the QM atoms must also be assigned MM van der Waals parameters. Differentiation of MM atom types by their van der Waals parameters is important for atoms such as halogen ions, which otherwise would be indistinguishable to the QM region as they have the same charge, and also important for distinguishing MM atoms with charges close to zero. Van der Waals terms are important at close range and are an important factor in determining interaction energies and geometries. 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.

The QM/MM interaction, therefore, depends on the choice of the MM point charges, the choice of the van der Waals parameters and the way that the MM charges polarise the QM region. The first two factors depend on the MM force field, as it is common practice to use the standard parameter set associated with the chosen force field. The charge definitions in an empirical force field are defined to reproduce electrostatic potentials; however, the charges of particular atom types can vary quite considerably between different force fields. MM parameters can be developed to reproduce the results of higher level calculations or experimental results, but this is not a trivial task.

2.2. *Partitioning of the QM and MM regions*

In most enzyme reactions, the definition of a suitably sized QM region requires the 'breaking' of a covalent bond on the frontier between the QM and MM regions. This requires special consideration, as the QM part treats electrons explicitly whereas the MM region does not, thus the QM atom will be left with an unpaired electron. Different methods have been developed to overcome this problem, which generally involves either the addition of an atom or a pseudo-atom to fill the valence of the QM atom, e.g. the link atom method [66], or a special treatment for the frontier bond orbital, e.g. the generalised hybrid orbital (GHO) method [67] or local self-consistent field (LSCF) method [68–70] (Figure 3). The details of the different treatments are described below.

The link atom approach [66] is the simplest method for termination of the QM region. This method involves the addition of an atom, known as a link or dummy atom, along the X–Y bond to be broken. The link atom is usually a hydrogen atom, but pseudo-halogen atoms are also available in some programs [71]. Hall *et al.* [72] have explored different treatments for link atoms, and Reuter *et al.* [73] have compared the link atom and LSCF approaches, finding that there is little difference in results at the semempirical level. There is some debate as to whether or not the link atom should interact with the MM region [73–75], and it is now quite widely accepted that the link atom should interact with the MM region, with the exception of the electrostatic group closest to it, to avoid the unwanted polarisation of the boundary [73,74,76–79].

Das *et al.* [80] have proposed a 'double link atom' method. The double link atom method extends beyond the standard single link atom method to overcome some of the problems of electrostatic interactions that can arise with the single link atom method. For example, typical single link atom methods of QM/MM partitioning can give an unphysical overall charge or dipole. These same authors have also developed a Gaussian

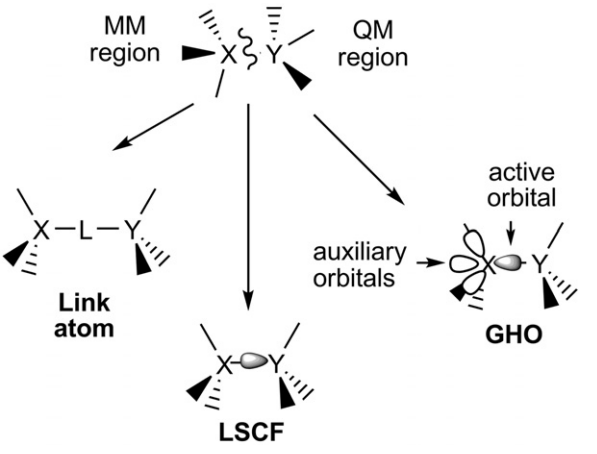

Figure 3. Different approaches to the termination of bonds at the QM/MM frontier: the link atom [66], GHO [67] and LSCF [68–70] methods.

delocalisation method for molecular mechanical charges in QM/MM calculations [80]. This delocalised Gaussian MM charge method aims to provide an empirical means of treating the delocalisation of the electron density that should physically be present for atoms in the MM region. This can have important effects for QM/MM electrostatic interactions. These authors also suggest that this approach may simplify the rules that have to be used to decide which molecular mechanical interactions to include in the calculation of energies and forces. For example, even at short distances between QM and MM atoms, the delocalised Gaussian MM method does not require the MM host atom charge to be excluded from the QM calculation, as is sometimes preferable or necessary when treating it simply as a point charge. The delocalised Gaussian MM method can potentially be combined with a wide variety of QM/MM partitioning methods, such as the link atom, frozen orbital or pseudopotential methods. Das *et al.* tested the delocalised Gaussian MM and double link atom methods on several small model systems chosen to represent key features of reactions in enzymes, such as proton affinities, torsional rotation barriers and deprotonation energies. These methods were found to give better energetic properties for the model compounds studied compared to point atomic MM charge and single link atom QM/MM methods.

Konig *et al.* [81] have tested a number of different QM/MM partitioning methods, based on the link atom approach, for the self-consistent charge density functional tight binding (SCC-DFTB) QM method. An important feature of this work is that it included studies of enzyme reactions, not just the analysis of the properties of gas-phase molecules. Among the partitioning schemes tested were all the options then currently available in the CHARMM program for SCC-DFTB/CHARMM QM/MM simulations. These differ in their treatment of electrostatic interactions with the MM atoms close to the QM/MM frontier. These workers also developed a new method, which they describe as the divided frontier charge approach. In this method, the partial charge associated with the MM atom bonded to the QM atom is evenly distributed to the other MM atoms in the same molecular mechanical group. These various QM/MM link atom schemes were tested for

the calculation of properties such as deprotonation energies, dipole moments, proton affinities and energetics of proton transfer reactions. As also found in earlier work, Konig *et al*. [81] established that QM/MM calculated proton affinities and deprotonation energies are highly dependent on the particular link atom scheme employed. They also found that the standard single link atom approach often gives errors of the order of 15–20 kcal mol^{-1} compared to pure QM calculations. Other schemes were found to give better results and to be generally comparable. It was found that both activation barriers and reaction energies for proton transfer reactions are fairly insensitive to the particular link atom scheme (e.g. to within a range of 2–4 kcal mol^{-1}) because of the cancellation of errors, for reactions in the gas-phase and in enzymes. 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. It is reasonable to expect that the general overall behaviour would be similar at other (higher) levels of QM/MM calculation. These authors emphasise that other technical details, such as the treatment of long-range electrostatics, are likely to play a more significant role in determining energetics generally, and stress that they must be treated carefully for reliable results to be obtained.

A 'connection atom' method for terminating the QM system has also been developed [75]. In this approach the MM atom at the frontier interacts in two different ways. It interacts in the standard way with other MM atoms, but its interaction with the QM region is defined by a new set of parameters developed to reproduce the behaviour of the X–Y bond being broken. Antes and Thiel [75] developed AM1 and PM3 parameters for a pseudoatom mimicking the behaviour of a methyl group. A similar approach has also been adopted by Zhang *et al*. [82] to develop a DFT pseudopotential for the termination of frontier bonds. The main advantages of this approach are that it avoids adding an extra atom into the system, and reproduces the behaviour of the X–Y bond more accurately, but it is necessary to develop parameters for each type of covalent bond being broken for each level of QM theory.

Warshel and Levitt [25] first introduced the idea of using a fixed orbital to terminate the QM region, using an sp^2 orbital at the QM/MM junction. Rivail and co-workers [68–70] developed the LSCF method, in which the electrons of the frontier bond are described by a strictly localised bond orbital (SLBO). The electron density of the SLBO is taken from calculations on small models, and as a result, are not transferable to other systems. The electronic properties of the orbital are then assumed constant during the reaction. The representation is frozen and the molecular orbitals describing the rest of the QM system, which are orthogonal to the SLBOs, are then generated using a local self-consistent procedure. This method has been implemented at both the semiempirical [68] and *ab initio* levels [69]. The GHO method for the termination of frontier bonds was developed by Gao *et al*. [67]. The classical frontier atom is described by a set of orbitals divided into two sets: active and auxiliary. The active set is included in the SCF calculation, while the auxiliary orbitals generate an effective core potential for the frontier and, unlike the LSCF method, the hybrid orbitals are transferable. This approach has been implemented at the semiempirical (for the AM1 [67], PM3 [83], SCC-DFTB methods [84]), density functional [85] and *ab initio* [86] levels of QM/MM theory. Rodriguez *et al*. [87] have compared the link atom and GHO treatments for the termination of QM/MM boundary for the acylation process in Hepatitis C virus NS3 protease. Umbrella sampling simulations were used to calculate the free energy (FE) barrier for the acylation reaction at

the AM1-CHARMM27 level of theory. They found that the two methods lead to similar chemical behaviour, but the agreement was not quantitative. Murphy *et al.* [88] have also developed a frozen orbital approach for the termination of frontier bonds in QM/MM calculations at the density functional level of theory.

2.3. *Implementations of QM/MM methods*

A number of QM/MM implementations have been reported in the literature; some are available for academic or commercial use. Singh and Kollman [26] combined a modified version of the Gaussian 80 UCSF code [89] with AMBER [90] to form the QM/MM program QUEST (quantum energies for simulation and trajectories). The QM/MM method of Field *et al.* [27] was implemented in the CHARMM software package [29], originally at the semiempirical level. Later developments have allowed *ab initio* and density functional QM methods to be implemented through the coupling of CHARMM [29] with GAMESS-US [91], GAMESS-UK [92] and CADPAC [93]. The semiempirical density functional-based SCC-DFTB method has also been implemented in CHARMM [94]. QSite [95] is a QM/MM program developed by Schrödinger Inc., using the Jaguar program [96]. Programs such as ChemShell [97,98] and QoMMMa [99] allow the coupling of many widely available quantum chemical and MM packages, enabling high-level QM/MM calculations. In QoMMMa [99], QM packages such as Molpro [100], Gaussian [101,102] or Jaguar [96] can be coupled with Tinker [103] to perform QM/MM calculations using a variety of MM force fields.

3. Modelling enzyme reactions

3.1. *Model preparation*

In order to model an enzyme with QM/MM methods, a detailed structure of the enzyme at, or mimicking, some point along the reaction pathway is required [40]. High-resolution X-ray crystal structures are usually the starting point, but there have also been examples of a well-validated homology model being used as a starting structure [104]. The crystal structure of the enzyme alone with no ligands bound in the active site is often not a good choice of starting structure as there may be conformational changes associated with binding that cannot be predicted without some indication of the binding mode of a ligand [40]. The efficiency of enzyme catalysis means that it is difficult to obtain structures of reactive substrate complexes from experiments because the turnover rate is too fast, but techniques involving cryogenic trapping, modified substrates, oxidation or reduction or selective mutation of active site residues have allowed the study of such reactive complexes in some cases [20,105]. Time-resolved crystallography can sometimes be used to obtain crystal structures of different complexes along a reaction pathway, and provides important information about any structural changes that occur along the path. The combination of multiple structures with theoretical modelling of the high-energy intermediate species that link them is an ideal method for studying an enzyme reaction [20,39].

Crystal structures are the best starting point for an enzyme model, but they do not contain all the information needed to represent the enzyme system [39]. Hydrogen atoms are not usually resolved in protein crystal structures due to their low electron density, and as a result they must be built into the structure using a theoretical algorithm. The lack

of information about the position of hydrogens can make it difficult to assign the protonation states of titratable residues, or to distinguish between the different tautomers of residues such as histidines. Incorrect assignment of the protonation states of residues, especially in and around the active site, can have a significant effect on the energetics of a reaction. Different methods are available to check the suitability of the chosen protonation states, such as initial calculations of pK_as by finite difference Poisson–Boltzmann methods [106] or using an empirical structure-based method for pK_a prediction developed by Jensen and co-workers (PROPKA: www.propka.ki.ku.dk) [107,108]. Molecular dynamics (MD) simulations can also be used to check the protonation states of a small number of residues by simulating the different charged states for a residue to test which is most consistent with the crystal structure [109,110].

Solvation of the model is also an important part of model preparation. The position of some ordered water molecules may be apparent in the X-ray crystal structure, but it is usually necessary to include more water to reproduce the effects of bulk solution. Many MM models have been developed to describe the properties of liquid water [111]. One commonly used model is TIP3P [112,113], a rigid model that represents the water molecule by point charges on each atom and Lennard–Jones parameters for the oxygen atoms only. However, the TIP3P model does not accurately reproduce the second solvation shell observed experimentally for liquid water [114–116]. The CHARMM22 [29] and CHARMM27 [50] force fields include a modified version of the TIP3P model, TIPS3P [117], which includes Lennard–Jones terms for the hydrogen atoms. Four and five point water models (TIP4P [29] and TIP5P [114]) do show a clear second solvation shell. Shaw *et al*. [118] have tested the compatibility of these MM water models with QM methods in QM/MM free energy calculations. Monte Carlo replica exchange thermodynamic integration FE simulations along a lambda coordinate were used to evaluate the free energy difference between treating a water molecule by a QM and an MM method. If the QM and MM methods are compatible, it should be possible to replace a water molecule in bulk solution with a QM water molecule without any significant changes in energy or structure. The TIP3P, TIPS3P, TIP4P and TIP5P water models were simulated using the HF, BLYP and MP2 QM methods (all using the aug-cc-pVDZ basis set) in the QM/MM calculations. Radial distribution functions (RDFs) were calculated to examine the structural effects of replacing the MM water by a QM water molecule. For calculations involving the BLYP QM method, the FE change for the perturbation to an MM water molecule was 0.5 ± 0.3, 0.5 ± 0.3, -0.1 ± 0.1, 10.4 ± 1.9 kcal mol^{-1} for the TIP3P, TIPS3P, TIP4P and TIP5P water models, respectively. The other QM methods indicated the same trend, showing the TIP5P model to be least compatible with all the QM methods. The RDFs calculated for TIP3P and TIPS3P models showed a second solvation shell as the QM character of the perturbed water increased. Only the TIP4P model showed the second solvation shell over all values of lambda. The TIP4P model was found to be the most compatible of the MM water models with all the QM methods, in terms of both structural and energetic considerations, showing particularly good compatibility with the BLYP method.

As QM/MM calculations are more computationally expensive than equivalent MM simulations, it is necessary to find a balance between the numbers of atoms included in the QM region, and the level of theory used to treat them. QM/MM calculations can be carried out at the semiempirical (e.g. MNDO, AM1, PM3 and SCC-DFTB), *ab initio*,

or density functional QM levels. Higher level calculations are becoming more common with improvements in computer technology and QM codes. Semiempirical methods have well-known limitations, but can sometimes give better results than, for example, Hartree-Fock level QM/MM calculations, and are still often the most practical method for QM/MM dynamics simulations.

All groups that directly participate in the chemical reaction (or reaction step) must be included in the QM region, but it may also be important to include residues that play a role in binding or stabilisation [119]. However, at the semiempirical QM level, some interactions (such as hydrogen bonds) may be better represented when one of the groups interacting is MM, so increasing the size of the QM region does not always give better results [27,43]. It has also been shown that semiempirical methods sometimes treat the conformation of peptide bonds poorly [120,121], but an MM correction can improve the results [122]. Polarisation in QM/MM hydrogen bonds has been tested by the comparison of B3LYP/6-311 + G(d,p)-CHARMM27 optimisation of biologically relevant hydrogen-bonded systems with gas-phase results [123]. It was found that these key interactions were treated well by QM/MM with good-quality levels of QM theory. Optimisation of van der Waals parameters for both the QM and MM atoms can help improve the interaction energies and geometries predicted by QM/MM methods [124–129] but this is not a key determining feature of accuracy in QM/MM condensed phase simulations [130]. The boundary between the QM and MM regions should be as far as possible from the site(s) of chemical change, but should not disrupt any conjugated systems or be in close proximity to highly charged groups [73]. The Cα–Cβ bond of an amino acid is a good choice, as it is relatively non-polar and allows the (reactive) side chain of the amino acid to be included in the QM region.

3.2. *Reaction pathway modelling*

Potential energy surfaces for enzymes are complex. Finding a minimum energy path (MEP) from the reactants to the products (i.e. a potential energy profile) for a reaction step of interest is therefore not trivial. Special techniques are needed to optimise TSs and/or reaction pathways. The algorithm must be able to treat the large number of degrees of freedom in an efficient manner. Some commonly used techniques for reaction path modelling in enzyme reactions are outlined below.

Adiabatic mapping [74,131] is a simple approach that has been used to calculate potential energy profiles in many QM/MM studies. In some cases, where large changes of structure or solvation, or movement of charge are associated with reaction, potential energy profiles may be misleading; however, they have been shown to be useful for a number of enzymes (e.g. by correlations with experiment [132,133]). The method requires the definition of a geometric coordinate to model the reaction, such as a bond length, dihedral angle or the combination of bond lengths. The model system is then optimised at fixed values of the reaction coordinate from reactants to products. The reaction coordinate is restrained, e.g. by using a harmonic potential with a force constant large enough to maintain the desired value of the reaction coordinate. The maximum in the potential energy along the path corresponds to the approximate TS location. By scanning along two-reaction coordinates a three-dimensional potential energy surface can be obtained, which can be useful when reactions are of a more concerted nature [134,135]. The adiabatic

mapping method is most suited to modelling reactions involving only small chemical and structural changes. Adiabatic mapping may overestimate barriers if atom movements connected with the reaction are not included in the reaction coordinate, but can underestimate the barrier in some cases [40]. Even if there are no protein conformational changes associated with the reaction, the fact that enzymes exhibit significant structural variability needs to be taken into account; numerous slightly different configurations can lead to reaction. Multiple pathways should be calculated using different snapshots of the protein from MD simulations to test and account for the effects of protein and solvent structural variability. With current computational resources, it is possible to obtain potential energy profiles for enzyme reactions with *ab initio* or density functional QM/MM methods (the use of *ab initio* methods that include electron correlation is generally still limited to 'single-point' energy calculations [30], not energy *minimisation*), which are potentially highly accurate.

Ideally, one would like to find the true TS (i.e. saddle point) structure for a reaction rather than an approximate TS structure, and then find the reactant and product minima that are connected by the TS. The lowest energy path that connects a TS to the reactant and product minima is known as the MEP or intrinsic reaction coordinate (IRC) [136]. Some optimisation methods use information about the second derivative of the energy in the optimisation process. To make these methods more tractable for large systems such as proteins, the system is often partitioned into 'core' and 'environment' regions [137]. Different algorithms can be used to treat the two regions with, for example, a quasi-Newton method used for the 'macro-iterations' of the core and a less intensive method used for the 'micro-iterations' of the environment. In the simplest case, only the core region is allowed to move and the rest of the system is held fixed, e.g. in variational transition state theory (VTST) calculations [138]. The number of core atoms is limited by the ability to store and manipulate their Hessian matrix, but any of the standard gas-phase TS searching methods can be used. When the environment is also optimised, convergence can be achieved in different ways: either the environment is fully relaxed at each core optimisation step or core and environment optimisations are alternately carried out until convergence. The separation of the system into core and environment regions has important implications for the calculation of the QM/MM electrostatic interactions during optimisations. For efficiency, the QM density is often represented by a point charge description during the micro-iterations of the environment [99,139,140]. Calculation of the QM/MM electrostatic interaction using a lower level of QM theory has also been proposed [141,142].

The methods for obtaining MEPs mentioned above are examples of local methods; a single conformation is adjusted to locate one stationary point at a time. They assume that there is only one saddle point along the path between the two minima. These methods work best for potential energy surfaces with simple topologies, and as a result are not always suited to the problem of TS location in enzyme systems [40]. Other methods exist that treat the entire path as a series of separate points in conformational space. The points are minimised simultaneously (under constraints to keep them falling into nearby minima), with the aim of finding all minima and barriers along the path. One approach, developed by Elber and Karplus [143] based on the elastic band method, is to minimise the average energy of a series of points along an approximate interpolated path between the initial starting structures often called images or replicas. The images are connected with springs,

which have an equilibrium length equal to an average separation between images. The definition of the spring forces in this way can result in the aggregation of images in lower-energy regions of the path. One improvement to the elastic band method involves the addition of a repulsive term between images to improve sampling near the TS [144]. The nudged elastic band (NEB) method [145–147] avoids some of the problems of the original elastic band method, as the use of certain components of the force ensures equal spacing of the images along the path and avoids 'corner' cutting near turns in the path. The NEB method shows initial fast convergence and can handle reactions with multiple steps but it does not locate the TS as it is a first-order method. In an alternative implementation, the climbing image NEB method, a TS can, in theory, be reached as the image with the highest energy is pushed uphill towards the maximum [148]. The parallel path optimiser method developed by Liu *et al*. [149] extends the method Ayala and Schlegel [150] for small molecules to large systems. For each point, the energy and gradients are calculated and empirical Hessians are obtained. The highest energy point is chosen to optimise to the closest TS and the remaining points are optimised to lie on the steepest descents path, but this method is limited to single-step reactions. Smart [151] has developed the conjugate peak refinement (CPR) method to locate TSs and optimise reaction paths in proteins. This method can determine the entire reaction pathway, possibly involving several energy barriers. The CPR method does not treat all the points on the path equally at all times, so that non-optimal points can be eliminated [152]. With all of these methods a pathway can be calculated from suitable reactant and product geometries or the algorithm can be used to refine an initial guess of the pathway, e.g. structures from adiabatic mapping. Calculation time can also be reduced significantly using these methods on parallel architectures.

3.3. *Free energy simulations*

To get a full picture of structural fluctuations and changes, more sampling of conformational space associated with the reaction is needed (so that a free energy profile rather than a potential energy profile can be obtained). Free energy barriers rather than potential energy barriers are the quantity needed for comparison with experimental values (Section 3.4). One popular approach to achieve this, pioneered by Carlson and Jorgensen [153], is to perform extensive sampling of the MM region at several points along the reaction, while the QM region is kept fixed. Such sampling can be performed with classical MD or Monte Carlo simulations. Free energy profiles can then be calculated as a perturbation (FEP, free energy perturbation) of the underlying potential energy surface obtained using QM/MM (or QM) energy minimisation [154,155]. A limitation of this QM/MM-FE approach is that important, different structures of the QM region may not be sampled. This limitation can be overcome by QM/MM simulations with MD or Monte Carlo sampling of both the MM and QM regions [156], or umbrella sampling/FEP simulations with the EVB approach (and subsequently using, for example, a single perturbation to get QM/MM free energies at important points of the profile) [157]. In some implementations of the QM/MM-FE approach, frequency calculations (within the harmonic approximation) are also carried out at several points along the MEP, for example, to include effects of vibrational quantisation of the QM region [158].

Lu and Yang [159] have developed a reaction path potential scheme for the generation of free energy profiles for enzyme-catalysed reactions. The first step is to generate the MEP using QM/MM methods and techniques such as NEB [145–147]. Energy, gradient, Hessian and charge response matrices are calculated at each point along the path and used to generate the reaction path potential. The reaction path potential is then used to carry out constrained MD simulations at each point along the MEP and the free energy profiles can be generated by thermodynamic integration. In this technique, the QM region undergoes dynamics coupled with the MM region and QM charges fluctuate in response to changes in the QM and MM regions. Another development from the Yang group is the QM/MM minimum free energy path method that allows the optimisation of the MEP on the free energy surface and generation of the free energy profiles from a single simulation [160]. The free energy of the system is expressed as a function of QM conformation. The free energy difference between two QM conformations is calculated using QM/MM-FEP methods and the free energy gradients are calculated from MD simulations at given QM conformations. The free energy gradients are used to search for the free energy path using NEB [145–147] or parallel path optimisation [149] methods without the need to specify a reaction coordinate.

Free energy profiles can be calculated by QM/MM MD (or Monte Carlo) simulations along a reaction coordinate. In order to sample regions of higher energy, a biasing potential is typically used (umbrella sampling) [161,162]. The introduced bias can subsequently be removed using statistical methods [162,163] so that a free energy profile or 'potential of mean force' is obtained. To arrive at statistically meaningful results requires the sampling of a large ensemble of conformations. The computational expense of QM/MM calculations makes this a challenging task. The use of semiempirical or similarly efficient QM methods may therefore be necessary, which might limit the accuracy of the results. It is, however, possible to use specifically parameterised semiempirical QM methods [79,164,165] or corrections with higher levels of QM theory [166] to achieve higher levels of accuracy for energy barriers.

The *ab initio* MD technique first proposed by Car and Parrinello around 20 years ago [167,168] has also been applied using QM/MM techniques. For example, Eichinger *et al*. [169] have developed a scheme for Car-Parrinello MD (CPMD) simulations with a QM/MM method, with the CPMD and EGO programs. Using these interfaced programs, efficient and consistent QM/MM Car-Parrinello simulations of large systems can be performed, including the steric and electrostatic effects of the protein and its solvent environment explicitly [168,170–173].

Other techniques, for example, using multiple steering MD [174–176], transition path sampling [177–182] and metadynamics techniques [183–188], exist for free energy calculations and have been used within a QM/MM framework to study enzyme reactions. The multiple steering MD approach, originally proposed by Jarzynski [174], relates the change in the free energy upon changing between two states to the change in the external work performed on a system subject to an external time-dependent perturbation. The free energy profile is calculated by performing a number of finite-time transformations, collecting the work done at each time step and then averaging over an ensemble of MD trajectories. Transition path sampling, developed by Chandler and co-workers [177–179], uses Monte Carlo importance sampling to generate unbiased ensembles of transition path trajectories that pass between reactant and product states. This method does not require

the definition of a reaction coordinate, but some properties such as bond lengths that can be used to separate reactant and product configurations are useful for identifying TSs. Metadynamics techniques [183–186] aim to generate the free energy profile for a reaction by using a set of collective variables to separate different states of the system. Small Gaussian-like potentials are used to form a history-dependent bias potential to prevent sampling regions of space that have already been visited. The free energy surface is then constructed from the biasing potential as the negative of the sum of the Gaussian potential terms.

3.4. *Comparison of calculated barriers with experiment*

Calculated energy barriers for reaction can be related to kinetic data. Often, an experimental apparent first-order reaction rate constant (k_{cat}) is available. Using TS theory, this rate constant can be converted to an (apparent) activation free energy for the overall enzymatic reaction [189]. In some cases the temperature dependence of the experimental reaction rate is known, so that the enthalpy of activation (which is similar to the potential energy barrier) and activation entropy can also be determined. In principle, the barriers obtained from experiment and computation for a single (e.g. rate limiting) step should be the same, providing all significant factors affecting the reaction rate are included in the modelling procedure (and the kinetic data unambiguously identify a single step). In practice, this agreement will depend on the accuracy of the modelling methods used, the quality of the model (influenced by the original structural data and QM/MM model setup) and also the experimental kinetic data – apart from the consideration of errors in kinetic data, it should always be borne in mind that the overall rate may depend on other steps (e.g. conformational changes or product release). Qualitative agreement between computation and experiment, however, is often sufficient to distinguish between alternative mechanisms. Recent developments have made it possible to predict barriers for enzymatic reactions using QM/MM methods capable of near chemical accuracy in favourable cases [30], bringing into view the prospect of reliable quantitative predictions. Kinetic isotope effects (KIEs) (i.e. the differences in rate upon substitution of hydrogen with the heavier isotopes of deuterium or tritium) are another useful experimental probe of reaction mechanism and are often used to indicate non-classical behaviour, i.e. quantum tunnelling. QM/MM methods can be used to calculate KIEs for enzymatic reactions and the inclusion of quantum tunnelling where necessary [138,142,190].

4. QM/MM studies of enzyme-catalysed reactions

4.1. *New insights into 'textbook' enzyme mechanisms*

4.1.1. *Hen egg white lysozyme*

Lysozymes belong to a large class of enzymes known as glycosidases. They catalyse the hydrolysis of the β(1–4) glycosidic bond between *N*-acetyl-muramic acid (NAM) and *N*-acetyl-D-glucosamine (NAG) in peptidoglycans, a component of some bacterial cell walls, particularly Gram-positive bacteria. Hen egg white lysozyme (HEWL) was the first enzyme to have its three-dimensional structure solved by X-ray crystallography [191]. It was also the first enzyme to have a mechanism proposed based on structural data, and

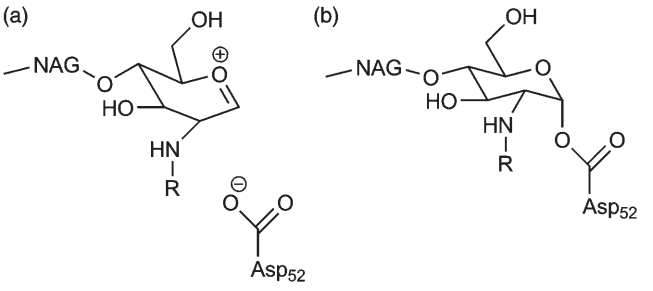

Figure 4. (a) The oxocarbenium ion intermediate in the mechanism for HEWL proposed by Phillips [191] and (b) the covalent intermediate in the mechanism for glycosidases proposed by Koshland [194]. HEWL was the subject of the first QM/MM investigation of an enzyme [25]. Recent QM/MM simulations support the Koshland mechanism, i.e. they indicate that the reaction of (wild-type HEWL with its natural substrate) proceeds via a covalent intermediate [195].

the mechanism proposed by Phillips [191] is commonly found in biochemistry textbooks. In this mechanism, the D (–1) site NAM was proposed to be distorted from the chair conformation when bound to the enzyme. Glu35 donates a proton to the glycosidic oxygen between rings D and E, leading to the breakage of the glycosidic bond and formation of an oxocarbenium ion intermediate. Experimental evidence from crystallography and electrospray ionisation mass spectroscopy [192,193] indicates that instead, HEWL follows a mechanism of the type proposed by Koshland, i.e. via a covalent intermediate rather than the oxocarbenium ion [194] (Figure 4). However, the use of mutant enzymes and/or unnatural substrates in these experiments does not categorically preclude the orignal mechanism proposed by Phillips in the biologically relevant reaction of the natural substrate and the wild type enzyme.

Lysozyme was the subject of the first QM/MM study of an enzyme, which was published in 1976 by Warshel and Levitt [25]. They studied the stability of the oxocarbenium ion and concluded that electrostatic stabilisation is the most important factor in forming this intermediate. Recent calculations performed by Bowman *et al.* [195] provide evidence that supports reaction in HEWL proceeding via a covalent intermediate, as proposed by Koshland (and subsequently found for other β-retaining glycosidases [193]), rather than the original mechanism proposed by Phillips [191]. PM3-CHARMM22 umbrella sampling simulations were used to calculate the free energy profile for the reaction in HEWL with high-level energy corrections calculated at the B3LYP/6-31 + G(2d) and MP2/6-311 + G(2d) levels. Spontaneous formation of the covalent intermediate was observed following the protonation of the glycosidic bond by Glu35 and cleavage of the glycosidic bond (Figure 5). The total barrier to the formation of the covalent intermediate was calculated to be $\sim$18 kcal mol^{-1} (for the reverse reaction), which is in good agreement with the value of 17 kcal mol^{-1} derived from the experimental rate constant (for the forward reaction). The TS to breakage of the glycosidic bond was found to have oxocarbenium ion character. The stability of the oxocarbenium ion compared to the covalent intermediate was investigated by simulating their interconversion. The free energy profile showed the covalent intermediate to be significantly more stable than the ionic form, by around 30 kcal mol^{-1} [195].

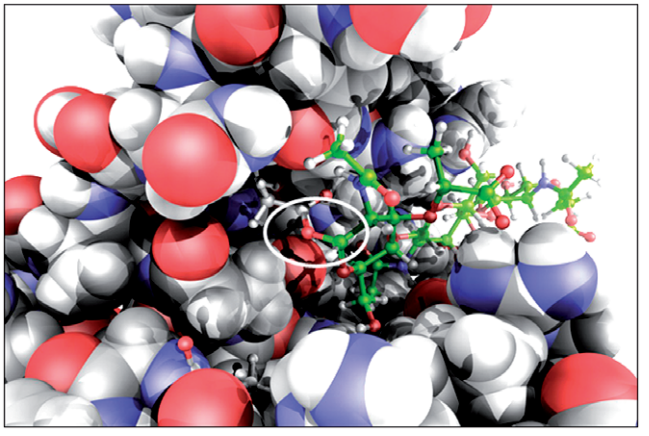

Figure 5. The covalent intermediate formed in lysozyme from PM3-CHARMM22 MD simulations in the work of Bowman *et al.* [195]. The covalent bond formed between the D(−1) site NAM sugar and Asp52 is circled.

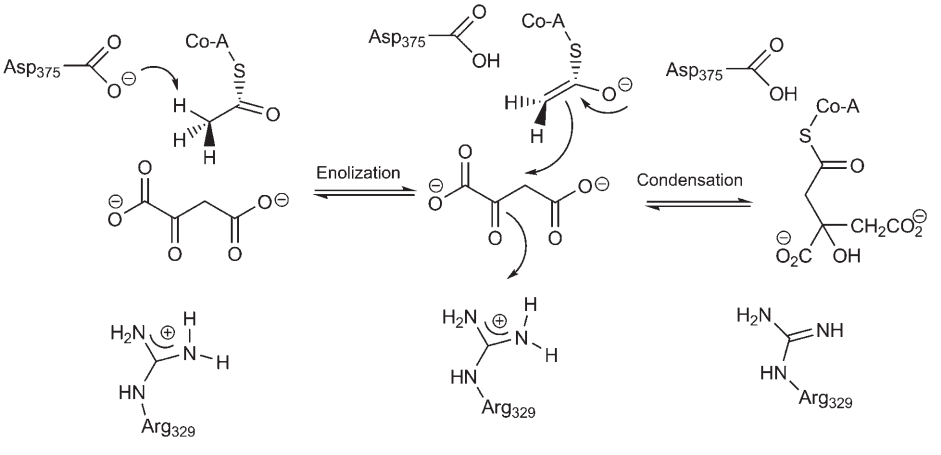

Figure 6. The mechanism for the formation of citryl-CoA in citrate synthase (involving the side chain of an arginine residue acting as the general acid) as suggested by the QM/MM calculations performed by van der Kamp *et al.* [203].

4.1.2. *Citrate synthase*

Citrate synthase catalyses the first reaction of the citric acid cycle, the conversion of oxaloacetate to citrate, using acetyl-coenzyme A (acetyl-CoA), an important model for mechanisms of carbon–carbon bond formation in biological catalysis. There are several steps in the reaction: (i) proton abstraction from acetyl-CoA; (ii) condensation of the deprotonated acetyl-CoA with oxaloacetate to form a citryl-CoA intermediate and (iii) hydrolysis of this intermediate to form citrate and CoA (Figure 6). The nature of the intermediate formed by proton abstraction from acetyl-CoA was not certain from

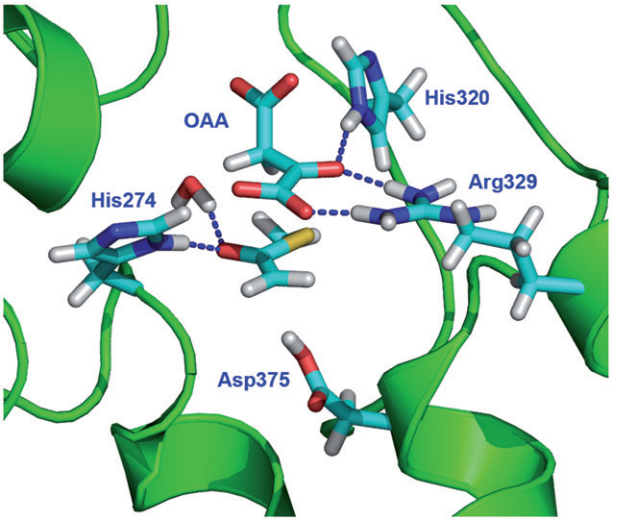

Figure 7. [Colour online] The enolate formed during the formation of citryl-CoA from the QM/MM calculations performed by van der Kamp *et al.* [203]. Carbon atoms are shown in cyan, oxygens in red and nitrogens in blue.

experiments [196]. The intermediate had variously been proposed to be either the enolate or enol of the acetyl-CoA thioester, or an 'enolic' form with a so-called 'low-barrier' (e.g. shared proton) hydrogen bond between His274 and the acetyl-CoA oxygen [197–201]. Early studies at the AM1/CHARMM level [76] and subsequently with *ab initio* (MP2/6-31 + G(d)//6-31G(d)/CHARMM) QM/MM methods [11,202] indicated that the enolate was energetically favoured. Hydrogen bonds donated by His274 and a crystallographically observed water molecule provide significant stabilisation of the enolate intermediate, but they are conventional rather than 'low-barrier' hydrogen bonds [11,202]. Figure 7 shows the hydrogen bonds present between the enolate intermediate and the active site of citrate synthase.

In the condensation step, the former carbonyl oxygen of oxaloacetate must be protonated to form citryl-CoA, but the nature of the proton donor has been the subject of debate. Recent high-level QM/MM modelling (at the MP2/aug-cc-pVDZ-CHARMM//B3LYP/6-31G(d)-CHARMM level) has suggested that Arg329 acts as the acid to donate the proton [203]. This is a relatively unexpected role for an arginine side chain, which is quite basic. Several mechanisms for condensation were modelled, with the arginine pathway giving the most reasonable results. The use of a variety of reaction coordinates for the condensation step indicates that carbon–carbon bond formation is concerted with the proton transfer from Arg329 to the former carbonyl carbon. The overall calculated reaction barrier of 14.2 kcal mol^{-1} is in excellent agreement with the experimental barrier of 14.7 kcal mol^{-1} [204]. This proposed mechanism (i.e. with Arg329 as the general acid for citryl-CoA formation) avoids overstabilisation of the citryl-CoA intermediate, and could help to trigger the large-scale conformational change necessary for the hydrolysis reaction or product release.

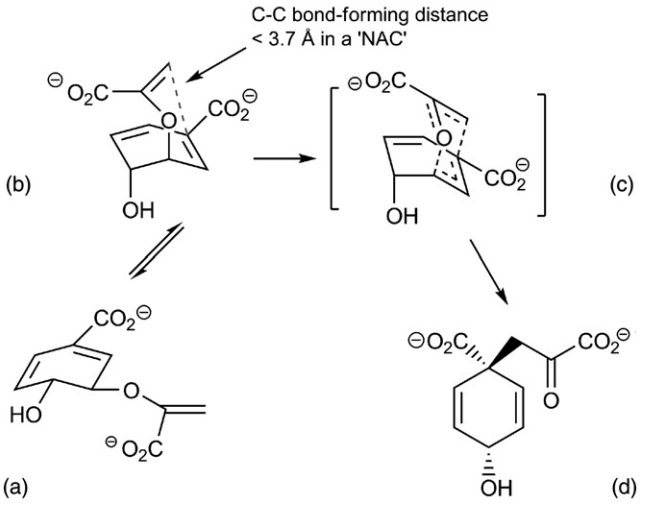

Figure 8. Mechanism for the conversion of chorismate to prephenate catalysed by chorismate mutase showing (a) the non-reactive diequatorial form of chorismate, (b) the diaxial form of chorismate indicating one definition of a 'NAC' [13–17], (c) the chair-like TS and (d) prephenate. The rearrangement of chorismate to prephenate has been the subject of many QM/MM studies to provide insight into the fundamental principles of catalysis [133,208,210,211,215] and also to aid the development and testing of new QM/MM methods [72,142,157,190,227–231].

4.2. *Chorismate mutase: analysing the origins of catalysis*

Chorismate mutase has played a central role in the development of theories of enzyme catalysis [9,205]. It catalyses the Claisen arrangement of chorismate to prephenate (Figure 8), a key step in the shikimate pathway for the biosynthesis of aromatic amino acids in plants, fungi and bacteria. As a consequence, the enzyme is a potential target for new herbicides, fungicides and antibiotics. The mechanism of the reaction is the same in solution and in the enzyme, and the enzymatic reaction does not involve any covalent interaction between the enzyme and the substrate, so it is also an ideal model system to investigate catalysis. At 25°C an apparent activation free energy of $\Delta^{\ddagger}G = 15.4\,\mathrm{kcal\,mol^{-1}}$ has been found for *Bacillus subtilis* chorismate mutase (BsCM), compared to $\Delta^{\ddagger}G = 24.5\,\mathrm{kcal\,mol^{-1}}$ for the uncatalysed reaction in aqueous solution (the reaction is believed to be a unimolecular pericyclic reaction in both cases), a rate acceleration by the enzyme of 10^6 [206].

The first QM/MM modelling of this enzyme at the AM1-CHARMM22 level indicated that stabilisation of the TS in the enzyme is important and suggested that substrate destabilisation or strain may also play a role [207]. Subsequent studies at the semiempirical [208–210] and *ab initio* [211,212] levels of QM/MM theory also indicate that TS stabilisation is important in this enzyme. The active site of chorismate mutase at the TS from HF/6-31(d)-CHARMM calculations [212] is shown in Figure 9. Detailed comparison of the reaction in solution and in the enzyme, with semiempirical QM/MM methods, indicates that both substrate preorganisation and TS stabilisation contribute to catalysis in the enzyme [166,208,210,213,214]. Claeyssens *et al.* [133] calculated 16 different potential

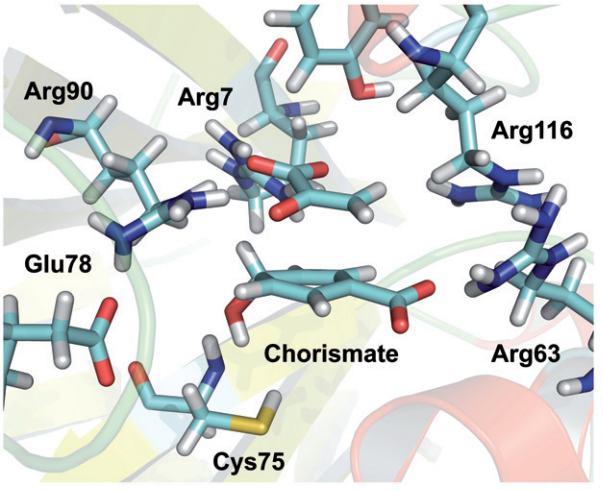

Figure 9. [Colour online] The active site of chorismate mutase at the TS from *ab initio* QM/MM calculations (HF/6-31G(d)-CHARMM22) [212]. Carbon atoms are shown in cyan, oxygens in red and nitrogens in blue.

energy profiles of the reaction in the enzyme using DFT QM/MM methods (B3LYP/6-31G(d)/CHARMM27). These multiple adiabatic mapping calculations (using multiple starting structures from lower level QM/MM MD simulations) gave an average potential energy barrier of 12.0 ($\pm$1.7) kcal mol^{-1}, in excellent agreement with the experimental activation enthalpy of 12.7 kcal mol^{-1}. By comparing the energy profiles in the enzyme environment and in the gas-phase (which for this enzyme has shown to give similar barriers as the reaction in solution [211,215,216]), it was found that the TS is stabilised significantly more than the reactant. This stabilisation varies depending on the local structure of the active site, but on average, the enzyme was found to stabilise the TS 4.2 kcal mol^{-1} more than it stabilises the reactant. The barrier height of the individual profiles also correlated strongly with the amount of TS stabilisation, indicating that the latter is the determining factor for reactivity in the enzyme.

Claeyssens *et al.* [30] have used high-level QM/MM calculations to provide activation enthalpies for the conversion of chorismate to prephenate to within chemical accuracy. Geometries from the 16 reaction paths generated at the B3LYP/6-31G(d)-CHARMM level, as mentioned above [133], were used to perform single point QM/MM energy calculations using high-level QM/MM methods, including the effects of thermal and zero-point energy corrections (at 300 K). Activation enthalpies of 28.3 ($\pm$2.1), 10.2 ($\pm$1.8), 9.5 ($\pm$1.0) and 13.1 ($\pm$1.1) kcal mol^{-1} were calculated at the HF, B3LYP, LMP2 and LCCSD(T0) QM/MM levels, respectively. The numbers in parentheses indicate the standard deviation of the 16 calculated paths. The L in these acronyms indicates that local approximations were used, T0 is an approximate triples correction and CCSD denotes coupled-cluster theory with single and double excitations. The effect of the size of the QM region was tested: an extended QM region (including Arg7, Arg63, Glu78 and Arg90 side chains and three water molecules in addition to the substrate) only changed the B3LYP/MM barrier by a very small amount, around 0.7 kcal mol^{-1}. DFT and LMP2

methods overestimated the effects of electron correlation quite significantly, giving barriers that are 3–5 kcal mol^{-1} lower than the experimental value. Only the LCCSD(T0) results were found to be in close agreement with experiment. The accuracy provided by CCSD(T0) methods, which were previously only feasible for small molecules in the gas-phase, is now achievable for enzyme systems within a QM/MM scheme. Reaction rates of course depend on activation free energies rather than on reaction enthalpies. For chorismate mutase, the entropic contribution at 300 K is computed (by low-level QM/MM MD simulations) to be 2.5 kcal mol^{-1}, in good agreement with the experimental value of 2.7 kcal mol^{-1} [206]. This advance in accuracy means that, in the best cases, calculations can predict enzyme mechanisms with quantitative accuracy. Results for the *para*-hydroxybenzoate hydrolase system, also included in this study, were in excellent agreement with the experiment [30].

Bruice and co-workers [14,15,17,217–221] have proposed that catalysis in chorismate mutase (and many other enzymes) is due to the ability of the enzyme to bind what they termed 'near attack conformers' of the substrate (NACs), and that almost no stabilisation of the TS is necessary for catalysis. Several different definitions of an NAC have been suggested, such as conformations in which the reacting atoms are at van der Waals distance (e.g. a carbon–carbon distance $\leq$3.7 Å in chorismate mutase) or the distance criterion with additional terms relating to the angle of attack. The 'NAC effect' was originally proposed based on MM MD simulations only, where NAC populations were counted to estimate the mole fraction in the enzyme and in solution [14,218,219]. The calculation of potential catalytic contributions through unrestrained MD simulations in this way has been criticised as being unreliable: it is known that populations of unfavourable (high energy) states will be underestimated generally in typical unbiased MD simulations [222].

In contrast to the 'NAC proposal', QM/MM modelling of the enzyme reaction and mutagenesis studies indicate significant TS stabilisation in chorismate mutase [207,209,210,223,224]. Repasky *et al*. [225] have studied solvent effects on the reaction, and their AM1 QM/MM Monte Carlo/Free Energy Perturbation (MC/FEP) study suggested that the conversion of non-NACs to NACs provides no free energy contribution to catalysis [226]. Chorismate does adopt different conformations in the enzyme and solution [16,209,213,226]. It is more generally useful to ask what the catalytic benefit is of forming the substrate conformation bound to the enzyme rather than trying to define 'NACs' through the use of several parameters. This question has been investigated using semiempirical AM1-CHARMM22 QM/MM simulations involving extensive FEP MD methods. It was found that forcing the conformation of chorismate in solution into the significantly different conformation found in the enzyme 'costs' 3.8–4.6 kcal mol^{-1} [16]. A similar estimate of 5 kcal mol^{-1} was calculated using EVB methods [13]. The agreement between these results, obtained with different theoretical methods, is encouraging and suggests that they are reliable. The difference between the experimental activation barriers for the reaction in the enzyme and in solution is 9.1 kcal mol^{-1} [206], as mentioned above. The catalytic benefit of the binding of chorismate to the enzyme in a more reactive conformation thus cannot be the sole factor in catalysis. Further stabilisation of the TS relative to the bound substrate (or reactant) must therefore be involved, as found earlier [133,208,210,211,215]. When the results for differential stabilisation of the TS (4.2 kcal mol^{-1} [133]) are combined with the calculated cost of forming the enzyme-bound

conformation of chorismate (3.8–5 kcal mol^{-1} [16]), a value very close to the experimentally observed catalytic rate acceleration by the enzyme (9.1 kcal mol^{-1}) is obtained. It can therefore be concluded that these two effects together account for catalysis by the enzyme. The affinity of the enzyme is higher for the TS than for the reactant (or product), and so both catalytic contributions can be understood as arising from the complementarity of the enzyme to the TS.

Chorismate mutase has served as a 'benchmark' system for testing new developments in QM/MM methods [72,142,157,190,227–231], and will continue to do so in future, being an ideal test system. It is recommended as a test case for methodological development and testing. The replica-path method for pathway generation has been applied to the chorismate mutase system [232] and used as a test for pathway optimisation techniques at the SCC-DFTB-CHARMM QM/MM level [233]. The effect of including MM polarisation has also been investigated for chorismate mutase [234]. Transition path sampling methods [235,236] and multiple steering free energy simulations using Jarzynski's equality [175] have also been tested on this enzyme. Crespo *et al.* [175] used multiple steering MD techniques to generate the free energy profile for the conversion of chorismate to prephenate at the DFT-MM level using SIESTA [237] for the QM calculation and the MM force field of Wang *et al.* [238] to include the protein environment. QM-MM calculations were performed using the PBE density functional with a DZVP basis sets and a pseudoatomic orbital energy shift of 50 meV with a grid cutoff of 150 Ry. The enzyme system containing chorismate or prephenate was first equilibrated by purely 2.5 ns MM MD simulations and 20 structures from each simulation were chosen as starting structures for the multiple steering dynamics simulations. Each structure was equilibrated at the DFT-MM level for 0.5 ps at 300 K before the multiple steering dynamics simulations were carried out along a defined reaction coordinate. They also calculated the free energy profile using umbrella sampling techniques for comparison with a total simulation time of 60 ps for both types of simulation. With both methods a free energy barrier of ~8 kcal mol^{-1} was obtained, lower than the experimental value due to the limitations of the PBE DFT method for this reaction.

Crehuet and Field [235] used transition path sampling techniques to generate a transition path ensemble for the conversion of chorismate to prephenate in chorismate mutase and in solution in order to gain insight into the geometry of the ensembles, rather than for the calculation of rate constants. The simulations were carried out at the AM1-OPLS-AA QM/MM method with 970 trajectories of 110 structures analysed for the reaction in the enzyme and 467 trajectories for the reaction in water. Reactants were defined as having C–O distances shorter than 1.7 Å and C–C distances longer than 2.9 Å and products had corresponding distances of 2.9 and 1.8 Å. Different methods of identifying TSs were tested: one from the calculation of committor probabilities (which is computationally expensive) and a less intensive method involving the calculation of the steepest descents path to find structures which lead to reactant and product basins. Principal component analysis was also carried out to identify characteristics of the ensemble. The committor probability is in principle the exact reaction coordinate for a reaction as it indicates the likelihood that a structure will go to reactants and products but it is difficult to determine and its value is only significant in regions close to the TS. Crehuet and Field [235] attempted to correlate the committor probabilities with various geometrical coordinates finding that the commonly used combination of distances did

show some correlation but others based on dihedral angles showed little correlation with the committor probabilities. Analysis of structures from the transition path ensemble showed that whilst all enzyme structures met the definition of a 'NAC' in terms of C–C distance, most of the structures do not meet the angle criteria and would therefore not be considered 'NACs' if this definition is used [220]. Some conformational compression was also identified. The distances between the carbon atom of the ring carboxylate of chorismate and C8 of Leu115 and the carbon of the C–O bond broken during reaction and CD1 of Phe57 were at a minimum ~150 fs before and 50 fs after the TS, suggesting that conformational compression of the substrate may play a role in catalysis. Electrostatic interactions were found to vary along the path, with the TS stabilised more significantly than the reactant in both environments, but the reorganisation energy needed to generate reactive conformations was not calculated and is likely to be greater in water than in the enzyme. B-factors were also calculated for individual residues and hydrogen bonding networks analysed. An extension to this study in which the trajectories were extended to 1.9 ps has also been reported [236]. Leu115 showed interesting behaviour in the simulations: the calculated B-factor for this residue was significantly lower for the transition path ensemble than for simulations of chorismate, prephenate or the TSA, despite it being unable to form hydrogen bonds, and analysis of the variation of the temperature of different groups during the simulations shows that this residue is far hotter than any other residue around the substrate as most of the energy released by the substrate upon reaction flows into this residue. Experimentally, mutations have been made in the C-terminus region of chorismate mutase to determine their effect on activity and Leu115 was shown to be the most conserved residue in this region [239].

4.3. *Quantum tunnelling in enzyme-catalysed reactions*

KIEs are a useful probe of enzyme mechanisms. In some cases, the primary KIE found on the substitution of a transferring hydrogen by deuterium or tritium is very large and can also display unusual temperature dependence. These high primary KIEs (high meaning above 7 for H/D substitution in a rate-limiting step) can be explained by a quantum mechanical tunnelling mechanism (a measured value below 7 does not necessarily rule out tunnelling; kinetic complexity may significantly reduce the observed value).

Atomically detailed understanding of enzyme reactions is needed to interpret experimental data, and to test hypotheses regarding the role of quantum tunnelling in catalysis. Computational molecular simulation methods, many based on QM/MM potential energy functions, are proving to be central in this area. One of the most widely used methods for calculating KIEs for enzymatic reactions is VTST with multidimensional tunnelling (MT) corrections, developed by Truhlar and co-workers [138,240,241]. QM/MM methods have formed the basis for applications of these methods to enzyme-catalysed reactions, to calculate energy profiles (and Hessian matrices). These calculations require the calculation of multiple Hessians, and variational TS optimisation (and can involve QM/MM classical MD simulations to calculate (classical) free energy profiles). As such, they are computationally demanding and have mostly used semiempirical QM methods such as AM1. The semiempirical methods can be parameterised for the reaction (e.g. to fit the reaction energy and barrier to results of higher level calculations on a small model). The development of specific reaction parameters (SRPs or system specific parameters, SSPs) for semiempirical methods, such as AM1 and PM3, is a useful way to improve

the potential energy surface [79,164], but developing new parameters for a reaction is a time-consuming process.

Several different protocols have been adopted for VTST/MT calculations within the QM/MM framework to model enzyme-catalysed reactions, which differ, for example, in the number of protein configurations used and whether or not the classical free energy profile is also calculated, as well as the type of method used to calculate tunnelling corrections [241]. The incorporation of quantum effects involves the determination of the MEP for a reaction; typically this can involve the optimisation of positions of the atoms in the QM region in a 'bath' of frozen MM atoms, though the optimised and frozen regions need not correspond exactly to the QM and MM regions. The vibrational partition functions and generalised free energies of activation are then calculated using the harmonic oscillator approximation. The transmission coefficient including tunnelling can be calculated using the small curvature tunnelling (SCT) approximation [242]. This procedure is then repeated with the isotopic substitutions in order to calculate the KIE. Variants of the VTST/MT methodology has been applied to many enzyme systems such as liver alcohol dehydrogenase [243,244], dihydrofolate reductase [245,246], xylose isomerase [247], triesophosphate isomerase [248], soybean lipoxygenase [249], methylamine dehydrogenase (MADH) [244,250–254] and aromatic amine dehydrogenase (AADH) [20,255], giving results for primary kinetic isotopes in good agreement with experimental data. Some attempts have been made recently to model the temperature dependence of the KIEs in dihydrofolate reductase [241,246], but this is a challenging task.

AADH and MADH belong to a family of dehydrogenases that catalyse the conversion of primary amines to the corresponding aldehyde and ammonia [256]. Reaction occurs at the TTQ co-factor, a prosthetic group formed by post-translational modifications of two tryptophan residues (Figure 10). Experimental data are available for both MADH and AADH with a range of substrates, and one of the largest (primary) H/D KIEs (55 ± 6 T = 5–20°C) observed for an enzymic proton transfer has been obtained for AADH with tryptamine as a substrate [20]. The availability of kinetic data and crystal structures has made MADH an important model for tunnelling in enzymes of this class. Several groups [244,250–252,254] have used the VTST/MT approach (within slightly different methodologies) to model the proton transfer step in the reaction of methylamine in MADH, and also of ethanolamine as the substrate for the enzyme [253]. High-resolution crystal structures have been solved for multiple intermediates in the reaction of AADH with tryptamine [20] formed the basis of QM/MM modelling of the proton transfer step in the AADH/tryptamine system [20,255].

Alhambra *et al.* [250] performed umbrella sampling simulations along a reaction coordinate to calculate the classical potential of mean force for the proton transfer reaction in the MADH/methylamine system, using SRPs for the PM3 method. In the second stage of the calculations, multiple MEPs from different protein configurations were calculated. The quantum-corrected potential of mean force was calculated by accounting for the difference in the quantal and classical vibrational free energies along the path. The transmission coefficient was then calculated to include the effects of tunnelling, using the SCT approximation [242]. The classical activation barrier of 20.3 kcal mol^{-1} was reduced by 3.2 kcal mol^{-1} due to the quantisation of the vibrational energies and the inclusion of tunnelling reduced the barrier by a further 2.5 kcal mol^{-1}. The activation

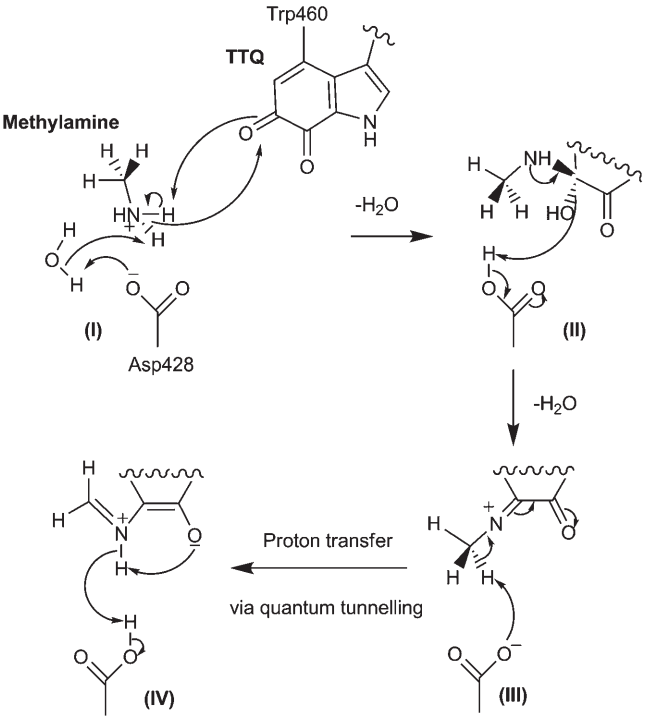

Figure 10. The proposed mechanism of the reductive half reaction in the oxidative deamination of methylamine by MADH [257]. The conversion of (III) to (IV) is believed to be the rate-limiting step for the reaction with a primary hydrogen/deuterium KIE of ~17 [257]. This reaction in MADH has been the subject of a number of QM/MM studies [244,250–252,254]. The mechanism is the same for the AADH/tryptamine system, which exhibits a primary hydrogen/deuterium KIE of ~55 [20].

barrier of 14.6 kcal mol^{-1} when all quantum corrections had been included was in very good agreement with the experimental activation barrier of ~14 kcal mol^{-1}. The predicted KIE at 298 K of 18.3 of the MADH/methylamine system was also in very good agreement with the experimental result of ~17 [257].

Faulder *et al.* [251] initially calculated only a single path for the proton-tunnelling step in the reaction of methylamine with MADH. They obtained a KIE of ~11 using PM3-AMBER QM/MM methods, suggesting that the semiempirical PM3 method may not model the energetics of the proton transfer sufficiently well to produce KIEs in agreement with the experiment. In subsequent studies [253,258], MM AMBER simulations generated an ensemble of structures for QM/MM VTST/MT calculations of the reaction of methylamine and ethanolamine in MADH. A wide range of barriers and KIEs were calculated for the MADH/methylamine system depending on the protein configuration and which oxygen of the catalytic aspartate the proton was transferred to [258]. *Ab initio* QM/MM methods (HF/3-21G-MM) were also used to generate MEPs for the proton transfer, and generated different degrees of reaction path curvature from the semiempirical methods [258]. Results were also improved by the use of SRPs developed for the PM3 method to describe the reaction [253]. Modelling of ethanolamine as a substrate

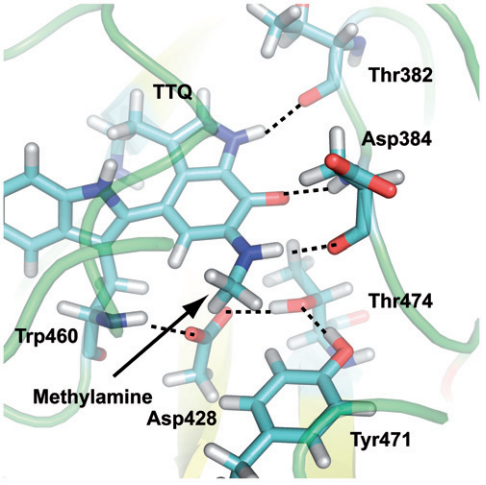

Figure 11. [Colour online] The active site of the MADH/methylamine system from PM3-CHARMM22 MD simulations described in the work of Ranaghan *et al.* [254] (carbon atoms are shown in cyan, oxygens in red and nitrogens in blue).

highlighted two different binding modes for the ethanolamine substrate, which gave rise to different kinetic behaviour [253]. The authors suggested that both conformers could be contributing to the experimentally observed temperature dependence of the KIE, with the less stable structure, playing a more important role at higher temperatures.

Ranaghan *et al.* [254] have also investigated the rate-limiting proton transfer step in the MADH/methylamine system and compared the results with the related AADH/tryptamine system [20,255]. The two carboxylate oxygens of the aspartate base (Asp 128β in AADH and Asp428β in MADH) are in different hydrogen bonding environments: OD1 is hydrogen bonded to a threonine residue (Thr172β in AADH and Thr474β in MADH) and OD2 is hydrogen bonded to a tryptophan (the unmodified tryptophan of TTQ: Trp160β in AADH and Trp460β in MADH). QM/MM MD simulations provided a key insight into the substrate complex for the proton transfer as it is poised for reaction in both systems. The active site of the MADH/methylamine system is shown in Figure 11. Both oxygen atoms were found to be within sufficient proximity to abstract the proton. PM3-CHARMM and AM1-CHARMM QM/MM MD simulations of the AADH/tryptamine complex (without any restraints imposed on the reaction coordinate) show that the proton to be abstracted interacts strongly with OD2 the closer of the two oxygens in the substrate complex, but also with OD1 for a significant part of the simulation (50% PM3-CHARMM, 80% AM1-CHARMM) [255]. In the MADH/methylamine system this is not the case, interactions between OD1 and the proton are quite rare (3–15% PM3-CHARMM). In AADH/tryptamine [255] transfer to OD1 of Asp128β, the oxygen more distant from the proton in simulations of the substrate complex, was both kinetically (by $\sim$4 kcal mol^{-1}) and thermodynamically more favourable than transfer to OD2 (PM3-SRP//PM3-CHARMM). In MADH/methylamine, and at the same level of theory, both proton transfers presented similar barrier heights and tunnelling contributions leading to a phenomenological activation free energy of 16.5 ± 0.9 kcal mol^{-1} but transfer

to OD1 is much less endergonic [254]. One possible cause for this observation is that MADH contains a non-conserved tyrosine residue (Tyr471β in MADH, which is equivalent to Phe169β in AADH). As a result, Thr474β forms hydrogen bonds with both the catalytic aspartate (Asp428β) and Tyr471β, reducing the mobility of Asp428β and Thr474β. These interactions are likely to impact on the ability of OD1 to reach the proton for hydrogen bonding and reaction. These environmental effects and the different preferences for the two carboxylate atoms of the catalytic base are possible causes of the differences in the observed KIEs for these two related enzyme systems.

4.4. *β-Lactamases*

Resistance of pathogenic bacteria to antibiotics is a serious and an ever increasing problem, posing a serious threat to human health. β-Lactamase enzymes are responsible for much bacterial resistance against β-lactam antibiotics such as penicillin, cephalosphorins and carbapenems. These enzymes are divided into four classes (A–D). In Classes A, C and D an active site residue such as serine is responsible for the acylation reaction, while Class B enzymes require one or two zinc ions and are known as metallo-β-lactamases. Improved understanding of their mechanisms could assist the design of inhibitors for these enzymes and new antibiotics. The reaction mechanism of a Class A β-lactamase with benzylpenicillin (Figure 12a) has been investigated by QM/MM methods. Potential energy surfaces for the acylation reaction of benzylpenicillin in the TEM1 β-lactamase were calculated using adiabatic mapping techniques at the AM1-CHARMM level of theory [259] and subsequently with hybrid DFT corrections [134]. The deacylation step has also been modelled following a similar B3LYP/6-31G(d)//AM1-CHARMM22 protocol [260], identifying Glu166 as the base in both acylation and deacylation reactions. The role of Lys73 in catalysis has also been debated; in these calculations, Hermann *et al.* found that Lys73 stabilises the TS of the acylation step and has a vital role as a proton shuttle in subsequent steps. *Ab initio* QM/MM (MP2/6-31 + G(d)-AMBER//HF/3-21G-AMBER) calculations have been used to study the TEM1 β-lactamase with penicillic acid (Figure 12b) as a substrate [261]: these calculations studied two possible pathways for the acylation reaction either with Glu166 or Lys73 acting as the general base, involving only a 4 kcal mol^{-1} difference in barrier height. These enzymes react with a wide range of substrates and it might be possible that they may perhaps utilise different pathways for different substrates. The question of the mechanism of class A β-lactamases is certainly worthy of further investigation.

Hermann *et al.* [262] have recently extended calculations on the first step of benzylpenicillin in the TEM-1 Class A β-lactamase reaction. Structures along the reaction path for tetrahedral intermediate formation were optimised at the B3LYP/6-31 + G(d)/CHARMM27 level, with energies for key points calculated up to the *ab initio* SCS-MP2/aug-cc-pVTZ/CHARMM27 level. The structure of the tetrahedral intermediate from B3LYP/6-31 + G(d)/CHARMM27 calculations is shown in Figure 13. The results support the mechanism involving Glu166 acting to remove a proton (via an intervening water molecule) from Ser70, which in turn attacks the β-lactam of the antibiotic. The calculated barriers range from 3 to 12 kcal mol^{-1} for this step (depending on the level of QM/MM theory), reasonably consistent with the experimental data. Modelling of this reaction step in a model of the K73A mutant gave a slightly higher barrier to reaction, indicating that

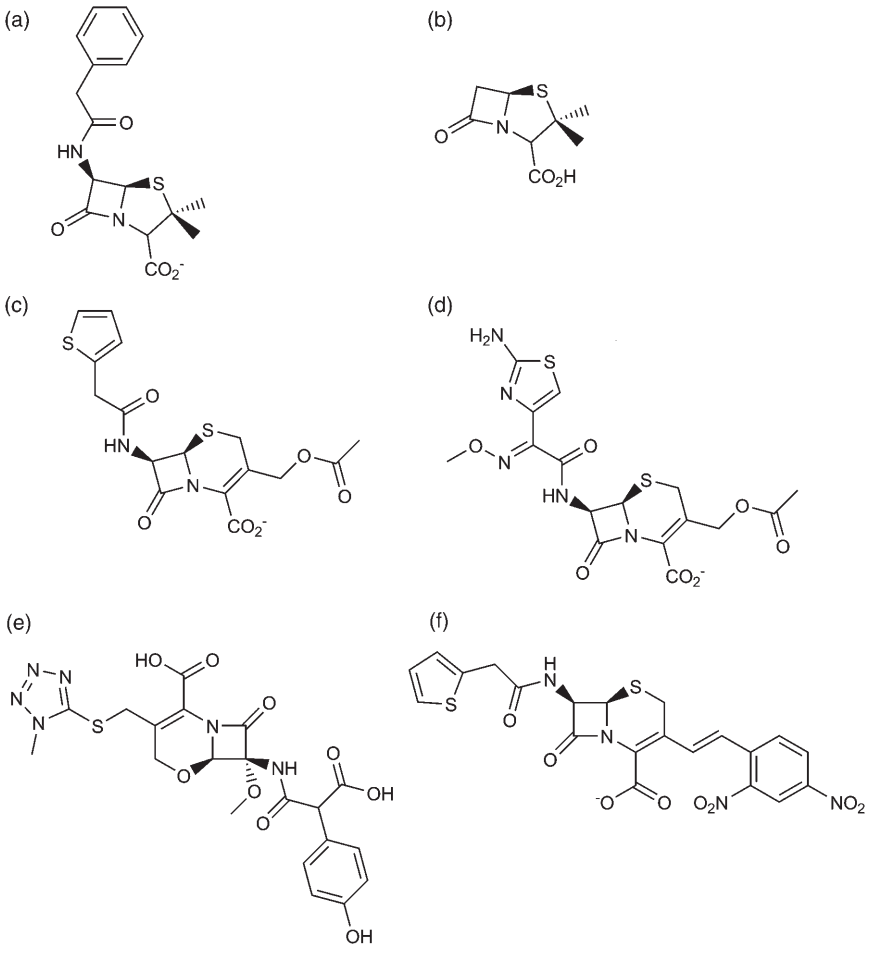

Figure 12. β-Lactamase enzymes are responsible for much bacterial resistance against β-lactam antibiotics. Examples of β-lactam antibiotics are (a) benzylpenicillin: the reaction of benzylpenicillin with a TEM-1 β-lactamase has been the subject of several QM/MM studies [134,259,260,262], (b) penicillic acid: a TEM-1 β-lactamase with penicillic acid as a substrate has been modelled by Meroueh *et al.* [261], (c) cephalothin: the substrate for a class C β-lactamase (P99) and also a penicillin-binding protein (PBP) in the work of Gherman *et al.* [263], (d) cefotaxime: the substrate for zinc-dependent β-lactamase in the QM/MM study by Dal Peraro *et al.* [171], (e) moxalactam: modelled as a substrate for a β-lactamase in the work of Xu *et al.* [264] and (f) nitrocefin: modelled as a substrate for a β-lactamase enzyme in the work of Park *et al.* [265].

Lys73 stabilises the TS, in particular deprotonated Ser70, lowering the barrier by about 1.7 kcal mol^{-1}. This finding may help to explain the conservation of Lys73, in addition to the role previously found for it in the later stages of the reaction [260].

Gherman *et al.* [263] have modelled the deacylation reaction of cephalothin (Figure 12c) with a class C β-lactamase (P99) and also a penicillin binding protein (PBP). They used B3LYP/6-31G(d)-OPLS-AA QM/MM methods to calculate free energy

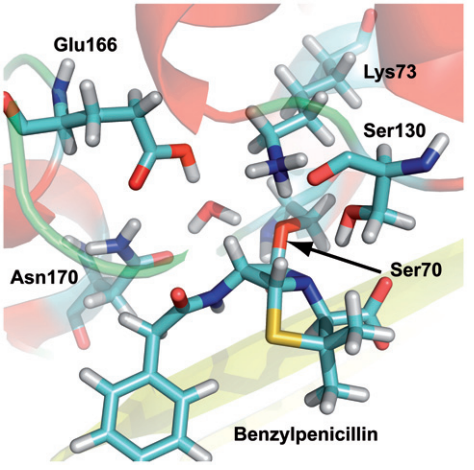

Figure 13. [Colour online] The tetrahedral intermediate formed in the acylation reaction of benzylpenicillin in TEM-1 β-lactamase from calculations at the B3LYP/6-31 + G(d)-CHARMM27 level of QM/MM theory in the work of Hermann *et al.* [262].

profiles (via static calculations with vibrational analysis) for the hydrolysis reaction including 150 atoms in the QM region. They considered two pathways: one involving a tyrosine as the general base and the other involving a concerted Tyr/Lys pathway. In the PBP the Tyr159 general base pathway was favoured, whereas in P99 the concerted Tyr159/Lys65 was favoured. They found the differences in the activity of P99 and PBP are due to differences in hydrogen bonds in the active site (β-lactamases such as P99 are deacylated efficiently, in contrast to PBPs, which are the targets for β-lactam antibiotics).

The metallo-β-lactamases have also been investigated using QM/MM methods; various such enzymes exist that contain either one or two zinc ions. The first step of the mechanism is believed to involve the nucleophilic attack of the zinc-bound hydroxide on the β-lactam carbonyl. In dizinc β-lactamases, the nucleophilic hydroxide is in a bridging position between two zinc ions. The second step is the cleavage of the C–N bond assisted by the protonation of the nitrogen or stabilisation by the second zinc ion in dizinc species. The proton donor to the nitrogen species has been the subject of much debate. Dal Peraro *et al.* [171] used hybrid CPMD methods at the BLYP-AMBER level of theory to study the mechanism of cefotaxime (Figure 12d) hydrolysis by a monozinc β-lactamase. The initial structure for the free energy simulations was obtained after 2.2 ns of MD. The system was first equilibrated by QM/MM CPMD for 0.1 ps with the QM region frozen, followed by 3 ps CPMD on the entire system. Constrained dynamics was then carried out along the first reaction coordinate involving the distance between the hydroxide and C2 of the substrate. The reaction coordinate was followed between values of 3.6 and 1.4 Å in 12 steps with 0.5 ps equilibration and 1 ps of production CPMD carried out at each point on the reaction coordinate. The intermediate formed by following this reaction coordinate was then simulated for a further 2 ps. Two more reaction coordinates (for protonation of the substrate nitrogen and protonation of the newly formed carboxylate upon ring opening) were modelled in the same way to complete the reaction. They find that during the first

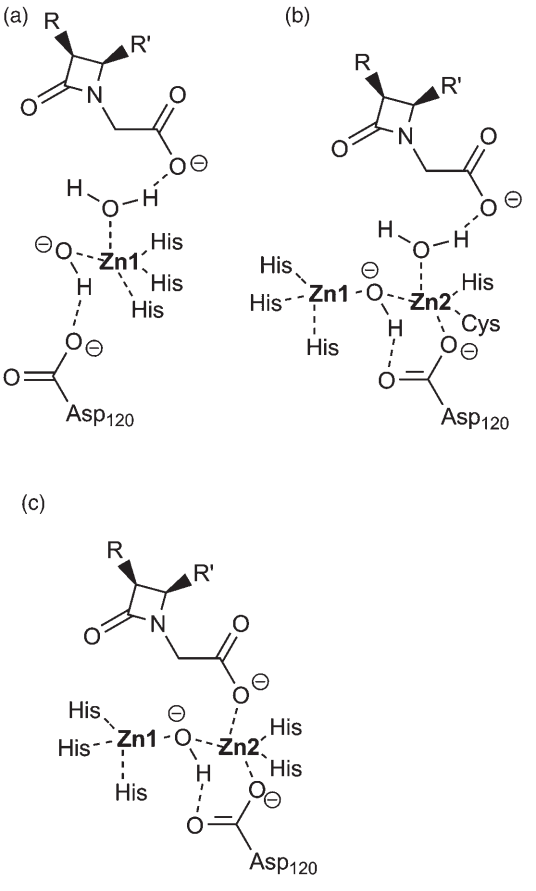

Figure 14. A schematic diagram showing the coordination of the zinc ions in the reactant state of class B β-lactamases: (a) a monozinc β-lactamase in the work of Dal Peraro *et al.* [171], (b) a dizinc β-lactamase in the work of Dal Peraro *et al.* [173] and (c) a dizinc β-lactamase in the work of Xu *et al.* [264].

step a nearby water molecule enters the solvation sphere of the zinc changing its geometry from tetrahedral to trigonal bipyramidal, altering the pattern of hydrogen bonds in the active site. Once the hydroxide has been transferred, the zinc returns to a distorted tetrahedral geometry as found in the initial structure (after 2.2 ns MM MD). This zinc-bound water molecule is then well positioned to act as a proton donor in the second step of the mechanism. The hydrolysis of cefotaxime (and other β-lactams) by a dizinc β-lactamase has also been investigated using similar Car-Parrinello QM/MM methods, suggesting that the reaction occurs in a single step [173]. Xu *et al.* [264] calculated free energy profiles for the ring opening reaction of moxalactam in a dizinc L1 β-lactam at the SCC-DFTB-CHARMM level of theory, which were then corrected using hybrid DFT-CHARMM energies. In contrast to the work of Dal Peraro *et al.* [173], Xu *et al.* [264] find that the β-lactam binds in a different manner and that the reaction occurs in a stepwise manner (Figure 14). They found that the reaction of moxalactam (Figure 12e)

is dominated by the nucleophilic addition of the hydroxide to the carbonyl carbon of the lactam with a B3LYP-CHARMM//SCC-DFTB-CHARMM free energy barrier of $\sim$24 kcal mol^{-1} in quite good agreement with the apparent experimental value of 18.5 kcal mol^{-1}. Cleavage of the C–N bond is concerted with the proton transfer of the hydroxyl proton to Asp120. One of the zinc ions stabilises the nitrogen-leaving group while the other stabilises the newly formed carboxylate. Park *et al*. [265] have carried out 200 ps PM3/AMBER QM/MM MD simulations of nitrocefin (Figure 12f) bound to a dizinc β-lactamase, before carrying out pathway calculations on smaller enzyme models using DFT. The β-lactam binds through its aminocarbonyl oxygen to the active site zinc ion with a tetrahedral coordination (Zn1) and the bridging hydroxyl was found to move away from Zn2 during the dynamics simulation to a position suitable for nucleophilic attack on the carbonyl carbon. The substrate was also observed to form a hydrogen bond with Zn2 through the carboxylate oxygen within the first 30 ps of the simulation, which remains constant for the remainder of the simulation. QM/MM dynamics simulations at the SCC-DFTB-CHARMM level (600 ps: 200 ps equilibration + 400 ps production) have been used to study inhibitor complexes of the B1 metallo-β-lactamase IMP-1 [266]. They studied three enzyme-inhibitor complexes, which function by displacing the hydroxide nucleophile bridging the two zinc ions, giving results in good agreement with the experimentally determined structures.

4.5. *Fatty acid amide hydrolase*

Fatty acid amide hydrolase (FAAH) is an enzyme involved in endocannabinoid metabolism, a promising target for the treatment of central and peripheral nervous

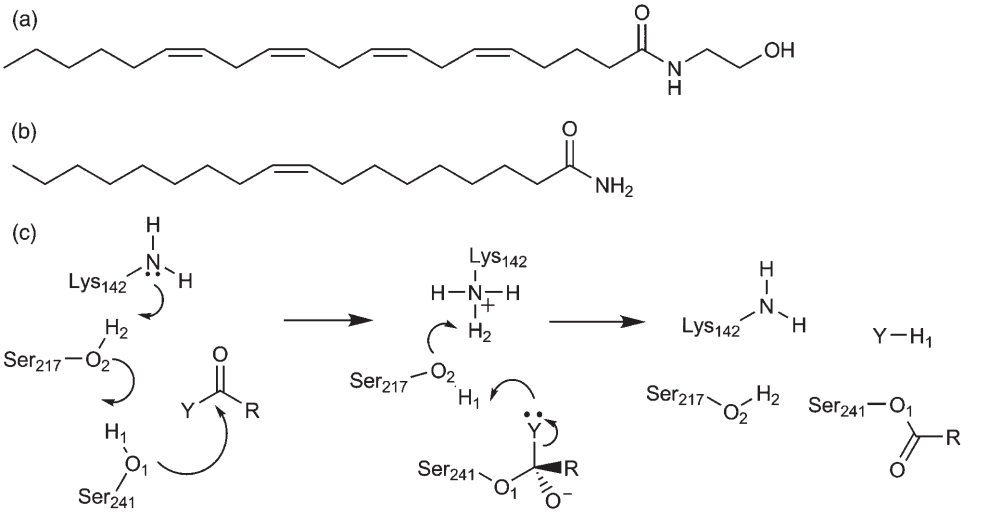

Figure 15. FAAH is an important enzyme in endocannabinoid metabolism, and is a promising target for treatment of nervous system disorders such as anxiety, pain and depression. It has been the subject of several investigations with different QM/MM techniques. (a) Structure of anandamide (b) structure of oleamide and (c) acylation mechanism showing the role of the Ser-Ser-Lys catalytic triad as studied by QM/MM methods [135,267,268].

system disorders such as anxiety, pain and hypertension. It is responsible for hydrolysing bioactive lipid amides including the endocannabinoids anandamide and oleamide (Figure 15a and b). Unusually it can also hydrolyse esters at a simlar rate as amides; both reactions occur via a mechanism where acylation is rate limiting. Unlike classic serine hydrolases, which contain a Ser–His–Asp triad, FAAH contains a Ser–Ser–Lys catalytic triad, the function of which is supported by QM/MM calculations [135]. The B3LYP/6-31 + G(d)//PM3-CHARMM QM/MM potential energy surfaces calculated by Lodola *et al.* [135] (Figure 16), which give a barrier consistent with experiment, indicate that neutral Lys142 deprotonates the nucleophilic residue Ser241 via Ser217, which acts as a proton shuttle during the reaction (Figure 15c). This mechanism is supported by independent QM/MM FEP calculations and Monte Carlo simulations at the PDDG/PM3 QM level [267]. This study also describes similar simulations of the reaction in the Lys142Ala mutant, which impressively reproduce the observation that this mutation decreases the rate of hydrolysis for oleamide significantly more than for methyl oleate. It is encouraging that independent studies using different QM/MM techniques arrive at the same qualitative mechanistic conclusions and strengthen these conclusions. Simulations

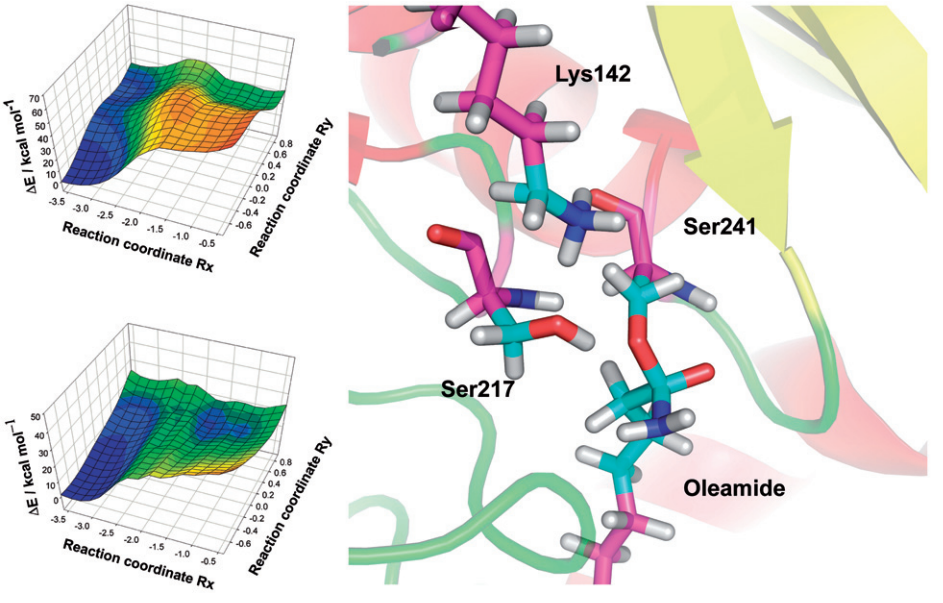

Figure 16. QM/MM potential energy surfaces for the formation of the tetrahedral intermediate for acylation of oleamide in FAAH from the work of Lodola *et al.* [135]. Top left, the potential energy surface calculated at the PM3/CHARMM22 level and bottom left, the higher level B3LYP//PM3/CHARMM22 surface. The surfaces are shown in a colour gradient from blue to red indicating increasing energy. The TS structure (as identified from the B3LYP QM/MM surface) is also shown. Carbon atoms in the QM region are shown in cyan and in the MM region are shown in pink, oxygen atoms are shown in red and nitrogens in blue.

have also highlighted interesting conformational effects in (wild type) FAAH: QM/MM and MD simulations of FAAH show that reaction (acylation with oleamide) occurs via a high-energy conformation of the enzyme–substrate complex [268].

Carbamic acid aryl esters, such as URB524 and its derivatives, are interesting pharmaceutical lead compounds: they inhibit FAAH irreversibly, and in rodents show analgesic, antidepressant-like and anxiolytic-like effects [269]. There has been uncertainty over the mode of binding of these inhibitors: they can be docked in two possible orientations in the FAAH active site; both modes place the carbamic group close to the nucleophilic Ser241 and both are consistent with QSAR data. QM/MM modelling of covalent adduct formation for URB524 in FAAH at the B3LYP/6-31 + G(d)// PM3-CHARMM22 level of theory [270] shows that only one binding orientation is consistent with the experimentally observed irreversible carbamoylation of the nucleophilic serine. This is a potentially crucial insight for designing new covalent inhibitors of this promising drug target, and understanding structure–activity relationships for these inhibitors. This application is an example of how QM/MM modelling can give useful insight for drug design (e.g. for covalent inhibitors), when traditional docking approaches alone may fail.

4.6. *Cytochrome P450 enzymes*

Cytochrome P450 enzymes (P450s) are an intriguing class of enzymes capable of catalysing a great variety of stereo- and regioselective reactions (such as hydroxylation, epoxidation and heteroatom oxidation) for a wide range of compounds. Several P450 isoenzymes are of particular pharmaceutical interest due to their involvement in drug metabolism. Apart from their influence on the bioavailability of drugs, P450 enzymes are often the cause of unwanted drug–drug interactions and the variability of drug metabolism between individuals (related to common polymorphisms in P450 genes). Understanding selectivity and reactivity in P450 enzymes is therefore pharmaceutically relevant. QM/MM modelling is an attractive tool for this purpose as the complex electronic structure of the haem group can be treated using reliable QM methods (e.g. hybrid DFT), while the effects of the protein environment included with MM modelling can potentially reveal differences between isoenzymes (both of different class and through polymorphisms) [6].

Because of the availability of crystal structures (and other experimental data), many QM/MM studies of P450 enzymes have focused on P450cam [271–280], a bacterial P450 that catalyses the hydroxylation of camphor. Calculations have found that hydrogen bonding in the enzyme environment can affect the spin density in the haem porphyrin and cysteine sulfur [274]. The protonation states of residues around the propionate side chains of the prophyrin ring also influenced the results significantly [276,279], showing that the protonation states of active site residues must be carefully considered in the preparation of an enzyme model. The active site of cytochrome P450cam is depicted in Figure 17, showing the reactive Compound I species.

Crystal structures of human P450 enzymes have also been determined, more recently [281–285]. Bathelt *et al.* [286] have used QM/MM calculations to study the reactant complexes of substrates in three human P450 isoenzymes (2C9, 2D6 and 3A4) with the drugs ibruprofen and diclofenac. The electronic structure of Compound I does not vary significantly between the three isoenzymes and the results were insensitive to a number

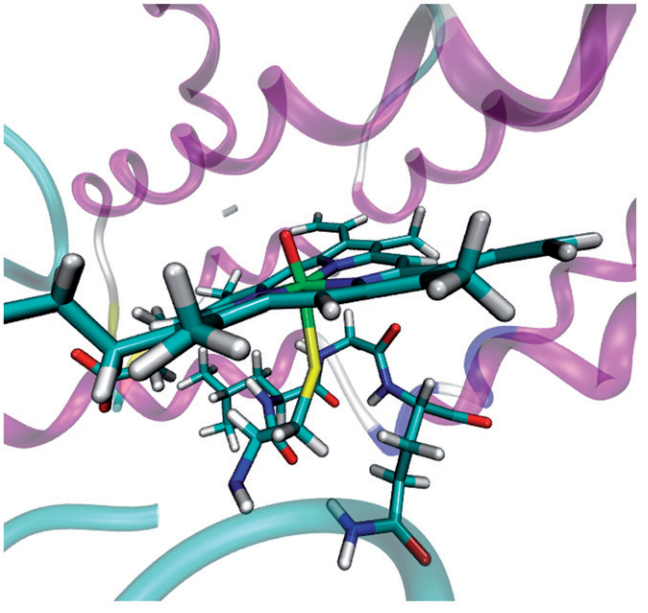

Figure 17. Many QM/MM studies of P450 enzymes have focused on P450cam [271–280], a bacterial P450 that catalyses the hydroxylation of camphor. The active site of cytochrome P450cam illustrating the hydrogen bonding environment of the haem-bound cysteine residue. Calculations have found that hydrogen bonding in the enzyme environment can affect the spin density on the haem porphyrin and cysteine sulfur [274].

of tests such as the size of the QM region, choice of starting structure (e.g. structural 'snapshots' taken from MD simulations), the DFT functional or the presence/absence of a substrate in the active site. This indicates that the experimentally observed large differences in substrate selectivity are not due to differences in the intrinsic reactivity between their respective Compound I species, but rather to differences in the interactions between substrates and the active site (e.g. different preferred binding orientations).

Cytochrome P450 3A4 can accommodate up to three substrates in its active site, giving rise to atypical kinetic behaviour [287,288]. A recent B3LYP/CHARMM27 QM/MM study examined features of Compound I in 3A4 isoenzyme with no, one or two molecules of diazepam bound [289]. In the absence of a substrate, the Fe–S bond tends to elongate and hydrogen bonds are formed between Compound I and active site residues, which may lead to its destruction. It was found that when two molecules of diazepam bind to 3A4, the second does not interact with Compound I directly, but helps to position the other diazepam molecule close to Compound I so that it can undergo reaction. The second diazepam helps to displace the hydrogen bonds between Compound I and active site residues, preventing its destruction.

The mechanism of benzene hydroxylation in cytochrome P450 2C9 has recently been studied using QM/MM methods at the B3LYP/CHARMM27 level. Bathelt *et al.* [290] found that aromatic oxidation may occur from both side-on and face-on geometries (with respect to the orientation of the benzene molecule and the porphyrin plane) in the

enzyme, in contrast to previous gas-phase results where a side-on orientation was preferred to a face-on orientation of the benzene substrate. However, substituents on the benzene substrate altered this preference. Three different pathways for the rearrangement of the initially formed σ-adduct were examined, leading to the formation of epoxide, ketone or N-protonated porphyrin species. From a face-on geometry, the barrier to the formation of epoxide and ketone products is similar at 5.1 and 6.3 kcal mol^{-1}, respectively. There was more variation in energy barriers from a side-on geometry (1.2–8.9 kcal mol^{-1}), with epoxide formation being favoured. The rearrangement to a N-protonated porphyrin is competitive with epoxide formation from a side-on geometry. The results suggest that several energetically feasible pathways will compete during aromatic hydroxylation in P450 2C9.

Cytochrome P450 StaP (CYP 245A1) converts chromopyrrolic acid to an antitumor drug, which proceeds by an unusual C–C bond coupling reaction [291]. In the X-ray crystal structures of P450 StaP [292], the substrate is situated in a position quite distant from the haem group of Compund I, suggesting that direct reaction of the substrate may not be possible without some movement of the substrate. QM/MM calculations at the UB3LYP/CHARMM22 level of theory have reavealed an important role for a crystallographic water molecule (Wat664) and a water molecule liberated during the formation of Compound I (Wat789). The water molecule analagous to Wat789 plays an important role in camphor dehydroxylation in P450cam [275,277]. The proposed reaction mechanism for the reaction in P450 StaP consists of 10 steps. The initial step was modelled as a proton-coupled electron transfer (PCET) reaction. The hydrogen-bonded triad of Wat664-His250-Wat789 acts as proton shuttles to transfer a proton from the substrate to Compound I. Subsequent steps involve carbon–carbon bond formation, several proton rearrangements and another electron transfer. QM/MM calculations of potential energy profiles for this reaction highlight the role of water molecules in the mechanism. These calculations also show that in this particular P450 enzyme Compound I behaves more like Compound II in cytochrome c peroxidase [293].

Recently, QM/MM calculations have been used to model the reaction path for hydrogen abstraction from N-palmitoylglycine in P450 BM3 [294]. In the crystal structure, the ligand was positioned in a 'distal state' and was unable to undergo reaction, so an induced-fit docking procedure was used to generate a model of the reactive complex. The reaction was then modelled at the UB3LYP/LACVP*-OPLS 1999 level of theory, giving free energy barriers (calculated from a static profile with vibrational analysis) including the effects of zero-point energy of 13.3 and 15.6 kcal mol^{-1} for quartet and doublet states, respectively. Tian and Friesner also identify a crystal water molecule that plays an important role in the reaction. The authors suggest that the similarity of the results for the P450 BM3 and P450cam systems [277] is possibly due to limitations in the performance of the DFT method for these systems.

4.7. *Glutathione S-transferase*

Glutathione *S*-transferases (GSTs) are a family of enzymes that catalyse the conjugation of glutathione, a γ-Glu-Cys-Gly tripeptide, to a wide range of compounds, a critical process in the detoxification of many xenobiotic (including carcinogenic) compounds and drugs. Through this process, xenobiotic compounds become more hydrophilic, allowing their

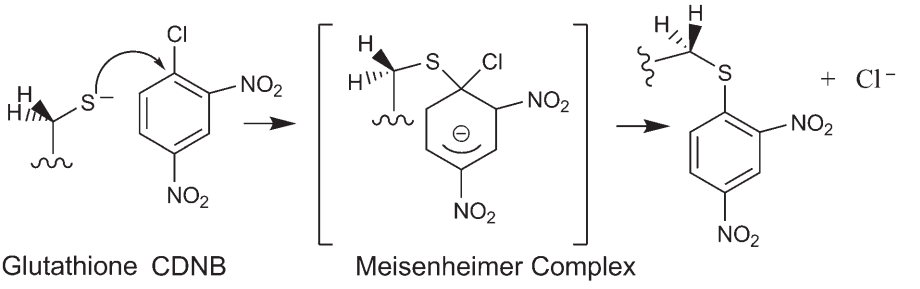

Figure 18. The conjugation of the glutathione anion with CDNB via a Meisenheimer complex as catalysed by GST modelled by Bowman *et al*. [295]. This is a critical process in the detoxification of many xenobiotic (including carcinogenic) compounds and drugs.

excretion from the body. Recently, Bowman *et al*. [295] have modelled reactions of 1-chloro-2,4-dinitrobenzene (CDNB) and glutathione in the wild type and a series of active site mutants of M1-1 GST. The mechanism for the reaction of glutathione with CDNB involves the formation of a glutathione thiolate anion, which can attack CDNB to form the so-called Meisenheimer complex (Figure 18). After the elimination of Cl^-, the glutathione-DNB product is formed. It is particularly important to include conformational sampling in the study of this reaction because it involves significant changes in solvation, which simple 'adiabatic mapping' approaches will not model correctly. QM/MM umbrella sampling simulations were used to calculate the free energy profile for the reaction using SRPs for the AM1 method (AM1-SRP) developed to overcome the limitations of the standard AM1 for this reaction. The reaction was modelled in solution, in the wild-type enzyme and three mutants, Tyr6Phe, Tyr115Phe and Met108Ala. Overall, the simulations of the reaction between glutathione and CDNB in the different environments compared very well to experiment, providing atomic-level insight into the effect of the mutations. For the wild-type enzyme, the release of the CDNB product is believed to be rate limiting, and the simulation results are consistent with this, as they show only a $\sim$1 kcal mol^{-1} barrier to form the Meisenheimer complex. The Tyr6Phe mutant shows a higher barrier to the formation of the Meisenheimer complex at $\sim$4 kcal mol^{-1}, consistent with the reduced activity of this mutant. After the formation of the Meisenheimer complex, the energy profile is the same as the wild type, implying that Tyr6 is not essential to catalyse the attack of the thiolate to CDNB, but does play a catalytic role through lowering the pK_a of the enzyme-bound thiolate. The role of water molecules in stabilising the glutathione thiolate was also confirmed by the observation that up to four water molecules could form hydrogen bonds to the glutathione thiolate. The simulation of Tyr115Phe mutant resulted in a similar free energy profile compared to the wild-type enzyme. Experimentally, this mutant shows increased activity compared to the wild-type enzyme, but the experimental data suggest that the increased activity is due to a difference in the rate of product release. Similar techniques have been used to model the reaction of this mutant enzyme with phenanthrene 9,10-oxide, which shows reduced activity compared to the wild-type enzyme for this substrate [125]. The free energy profile obtained for the Met108Ala mutant was also similar to that obtained for the wild-type

enzyme, consistent with the proposal that Met108 is important for substrate binding but not for catalysis.

Recently Dourado *et al.* [296] have used ONIOM QM/MM techniques to study the mechanism of the deprotonation of the thiol group in GST. It has been proposed that a water molecule is involved in the proton transfer from the thiol to the glutamyl α-carboxylate group. Calculations were carried out at the B3LYP/6-31G(d)-AMBER level giving a barrier of $\sim$13 kcal mol^{-1} in good agreement with experiment, whereas direct transfer between the glutamyl α-carboxylate and thiol groups gives a barrier of $\sim$35 kcal mol^{-1}.

4.8. *Butyrylcholinesterase*

Butyrylcholinesterase (BChE) is the primary enzyme involved in the metabolism of cocaine, leading to biologically inactive metabolites. A small proportion of cocaine is deactivated by cytochrome P450 enzymes, but metabolism by that route leads to the production of the hepatotoxic compound norcocaine. However, BChE shows significantly lower catalytic activity for the naturally occurring (–)-cocaine than for the biologically inactive (+)-cocaine (Figure 19). Design of a BChE mutant with increased activity for (–)-cocaine is an important target for therapeutic use in the treatment of cocaine addiction. The mechanism of cocaine metabolism in BChE is similar to that of ester hydrolysis by serine hydrolases [297]. BChE contains a catalytic triad of Ser198–His438–Glu325, like the Ser–His–Glu triad of acetylcholinesterase (AChE) (Section 4.10). Gao and Zhan [298,299] have calculated the reaction pathway for (+)-cocaine and (–)-cocaine using HF/3-21G: AMBER ONIOM methods with single point energy corrections at the MP2/6-31+G(d):AMBER level. The potential energy barriers for the four steps of the reaction of (–)-cocaine were 13.0, 0.1, 14.2 and 7.2 kcal mol^{-1}. For (+)-cocaine, similar potential energy barriers of 12.1, 0.4, 14.2 and 7.2 kcal mol^{-1} were obtained for the four reaction steps. The reaction barriers for the first and third steps are quite similar so that it could be difficult to assess experimentally which is the rate-determining step, if indeed a single step

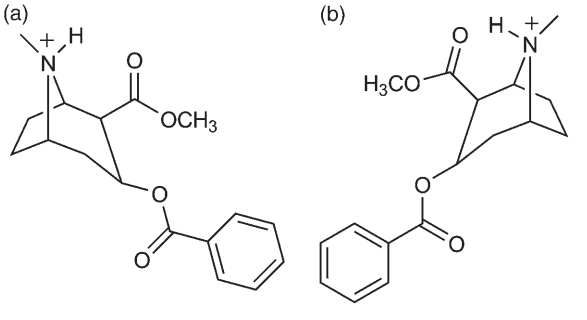

Figure 19. The structure of (a) (−)-cocaine and (b) (+)-cocaine, an illegal drug primarily metabolised in the body by the enzyme BChE. BChE shows significantly lower catalytic activity for the naturally occurring (–)-cocaine than for the biologically inactive (+)-cocaine. Design of a BChE mutant with increased activity for (–)-cocaine is an important target for therapeutic use in the treatment of cocaine addiction and has been investigated with ONIOM methods by Gao and Zhan [298,299].

does determine the overall rate. However, kinetic data for the A328W/Y332A mutant show that the first step is rate limiting with a much higher barrier than for the third step, which may indicate limitations of the calculations. The ONIOM results of Gao and Zhan show interesting interactions in the oxyanion hole consisting of Glu116, Glu117 and Ala119, different from those observed in AChE. Glu116 is not available for hydrogen bond formation with the substrate as it forms a hydrogen bond with the side chain of Asp197. These calculations also predicted that the mutation A199S/A328W/Y332G could significantly decrease the energy barrier for the third step of the reaction.

4.9. *Human immunodeficiency virus type 1 enzymes: reverse transcriptase, protease and integrase*

Acquired immunodeficiency syndrome (AIDS) caused by human immunodeficiency virus type 1 (HIV-1) is currently one of the most serious world health problems with more than 30 million people infected worldwide. There are three enzymes that are essential for the replication of HIV-1: reverse transcriptase, protease and integrase. These enzymes are important targets in the development of anti-AIDS drugs and QM/MM methods can help understand the mechanism, specificity and genetic variation of these enzymes. Common treatments involve a combination of drugs to inhibit HIV-1 reverse transcriptase and HIV-1 protease. A significant problem in anti-AIDS drug therapy is drug resistance caused by the mutation of the reverse transcriptase enzyme. Recently, attention has also been focused on HIV-1 integrase due to its role in the life cycle of the virus and the fact that it has no human analogues, so unwanted side effects may be minimised.

HIV-1 protease is a homodimeric enzyme that belongs to the class of aspartic proteases characterised by a dyad formed by two aspartic acids. Note that notation Asp25/25′ is used to indicate residues from different monomer subunits in this enzyme. These Asp residues are each part of a highly conserved Asp–Thr–Gly motif, located at the base of the substrate binding site. Above the catalytic aspartic acids are two flaps that control ligand entry into the active site. HIV-1 protease cleaves peptide bonds between Phe and Pro or Tyr and Pro. X-ray crystal structures (see, e.g. [300]) have identified a water molecule that bridges the substrate and flap regions, which is believed to be important in catalysis. Liu *et al.* [301] investigated the mechanism of HIV-1 protease using QM/MM umbrella sampling simulations combining the PM3 semiempirical QM method with the GROMOS87 force field. Several possible reaction pathways were considered, concluding that the nucleophilic attack and proton transfer processes were concerted. In another early study of HIV-1 protease, Chatfield *et al.* [302] used *ab initio* QM/MM techniques at the HF/STO-3G–CHARMM level of theory to compare possible mechanisms such as general acid–general base with a neutral intermediate or direct nucleophilic attack by an aspartate residue. At the time these calculations were very challenging and both mechanisms were shown to be stepwise with nucleophilic attach preceding protonation of the carbonyl oxygen, in contrast to the work of Liu *et al.* [301] at the semiempirical level. Piana *et al.* [170] used a QM/MM *ab initio* CPMD approach to study the reaction mechanism of HIV-1 protease. They found two distinct energy barriers: one for nucleophilic attack of the water molecule and a second for the protonation of the peptide nitrogen, with barriers of 18 and 21 kcal mol^{-1}, respectively. They tested the accuracy of their method using gas-phase models and found that the barrier heights are represented well by DFT methods,

but the stability of the intermediate is underestimated compared to MP2 and coupled-cluster doubles (CCD) methods [170]. They also tested the effect of the protein environment by switching off the electrostatic field generated by the protein and found this to increase the reaction free energy by 3 kcal mol^{-1}.

It has been proposed that the aspartic acid dyad in HIV-1 protease forms a 'low-barrier' hydrogen bond. Low-barrier hydrogen bonds (in which the proton is effectively shared between the hydrogen bonded partners in a charged system) were suggested to play a role in catalysis in many enzymes. QM/MM studies of, for example, citrate synthase [11] do not support this hypothesis in the majority of enzymes that have been investigated, and simulations indicate that low-barrier hydrogen bonds would not generally offer any energetic catalytic advantage [201]. QM/MM methods at the B3LYP/6-31 + G(d,p)-(effective fragment potential (EFP)) level have been used to investigate the possibility of a proposed low-barrier hydrogen bond in unbound HIV-1 protease [12]. Porter and Molina obtained a double well potential with a ground state vibrational energy that is 0.87 kcal mol^{-1} below the barrier for proton transfer. They also calculated nuclear magnetic resonance (NMR) chemical shifts for the proton involved in the low-barrier hydrogen bond in unbound HIV-1 protease and obtained a value of 0.72 ppm, which is consistent with experimental data for other enzyme complexes.

Suresh *et al*. [303] have investigated the role of the conserved water molecule that bridges the active site and flap regions of HIV-1 protease/inhibitor complexes. Six different structures were selected and calculations were carried out within the ONIOM method at the B3LYP/6-31G(d):AMBER level to obtain the binding energy of the structural water molecule. The structural water molecule forms four hydrogen bonds – two with the carbonyl groups of the ligand and two with the amide nitrogen of the flap region. Saen-oon *et al*. [304] have also used an ONIOM approach to study the interaction between HIV-1 protease and saquinavir (Figure 20d), the first inhibitor licensed for clinical use. The effect of protein mutations associated with saquinavir resistance was also investigated. They conclude that the mutation of residue 48 causes disruption of the direct interaction between the drug and the pocket lowering the binding affinity, whereas the mutation of the more distant residue 90 causes repositioning of residues throughout the protein structure. The mutation weakens inhibitor binding and increases the stability of the enzyme.

Molecular docking is a technique widely used to aid drug discovery [305–307]. QM/MM calculations of protein–ligand interactions can offer advantages over traditional docking methods and aid the development of more accurate scoring functions for the prediction of binding free energies [308]. Protein–ligand interactions have been examined for complexes of HIV-1 protease with the high-affinity inhibitors nelfinavir, mozenavir and tipranavir (Figure 20a–c) using AM1-CHARMM MD simulations [309]. In the energy decomposition analysis, 10,000 snapshots were taken from a 10 ps production simulation. The total electrostatic interaction was of similar magnitude for nelfinavir and tipranavir but that of mozenavir was significantly smaller. They found electronic polarisation of the inhibitors by the enzyme environment to be a significant contribution to the total electrostatic interaction energy at 32–39%. The contributions of individual amino acids were also investigated showing that the increased potency of tipranavir can be attributed to interactions with Asp25/25′ and Ile50/50′. They predict that placing a hydrogen bond donor in the position of the sulfonamide of tipranavir would optimise the interaction with Asp30 and make a more potent inhibitor. Recently, Fong *et al*. [310] have also investigated

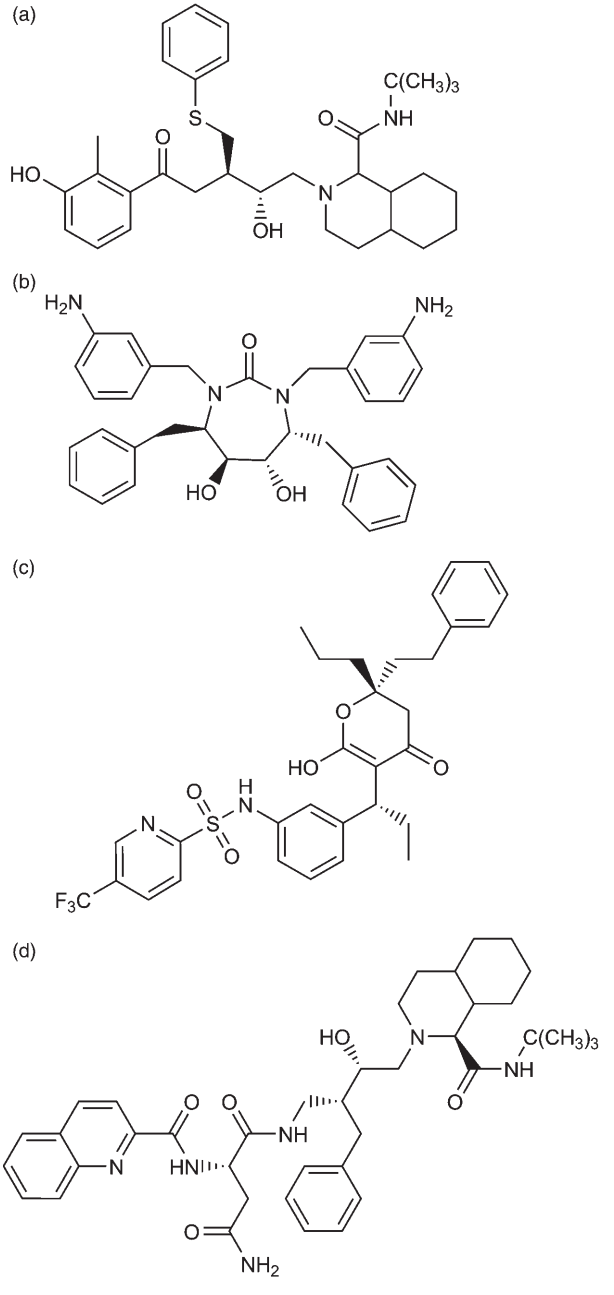

Figure 20. Common treatments for HIV involve a combination of drugs to inhibit HIV-1 reverse transcriptase and HIV-1 protease. The structure of HIV protease inhibitors (a) nelfinavir, (b) mozenavir, (c) tipranavir and (d) saquinavir used in the treatment of HIV. QM/MM methods have been used to provide insight into enzyme–drug interactions in HIV-1 protease/inhibitor complexes. Interactions in HIV-1 protease complexes with nelfinavir, mozenavir and tipranavir have been studied with QM/MM MD simulations at the AM1-CHARMM22 level of theory [309]. ONIOM methods have been used to investigate the interactions between HIV-1 protease and saquinavir [304].

HIV-1 protease inhibitors to assess the use of QM/MM scoring functions in docking procedures. They calculated protein–ligand energetics for six different HIV-1 protease inhibitors using an ONIOM method (with electrostatic embedding) at the HF/6-31G(d), AM1/d and PM3 and also a larger QM treatment at the PM3/d level with the AMBER force field for the MM part. Solvent effects were also included using a continuum model. They also compared results with the ChemScore [311] and GoldScore [312] scoring functions. The best QM/MM scoring function found was the QM/MM solvent corrected HF/6-31G(d)-AMBER model, which exceeded the performance of the ChemScore and GoldScore scoring functions. Semiempirical methods exhibited a lower discriminative power. The protonation state of the catalytic aspartic acid residues also affected the ability of the scoring functions to discriminate the native binding pose.

HIV-1 reverse transcriptase catalyses the multi-step conversion of the single-stranded viral RNA genome into a double-stranded DNA copy, which can be integrated into the host genome. Rungrotmongkol *et al.* [313] have investigated the mechanisms of the deprotonation and DNA polymerisation reactions using AM1-CHARMM22 methods. They proposed that Asp185 is the proton acceptor in the reaction, abstracting the 3′-OH hydrogen atom from the primer terminus prior to the nucleophilic attack. The free energy profiles showed the product to be unstable compared to the substrate and intermediate complexes leading to the suggestion that alternative mechanisms needed to be explored. In a subsequent study, they investigated the different possible protonation states of the deoxythymidine triphosphate (dTTP) substrate and two active site aspartates in HIV-1 reverse transcriptase using AM1-CHARMM22 and PM3-CHARMM22 methods [314]. Both QM/MM methods gave realistic structures of the dTTP triphosphate in good agreement with the crystal structure.

HIV-1 integrase catalyses two main reactions: (i) cleavage of the last two nucleotides from each 3′-end of viral DNA and (3′-processing) and (ii) attack of the 3′-OH end of the viral DNA on the phosphodiester bond in the host DNA human chromosome (strand transfer). Diketo acids are potent and selective inhibitors of HIV-1 integrase and the inhibitor 5CITEP has been crystallised in the catalytic domain of HIV-1 integrase, which has formed the basis of many classical MD and docking studies. Alves *et al.* have reported a QM/MM study of the protein–ligand interactions of 11 inhibitors of HIV-1 integrase. One nanosecond of AM1-CHARMM22 dynamics was carried out for each protein-inhibitor system with averaged interaction energies calculated from 100,000 configurations from the last 100 ps and decomposed into electrostatic and polarisation terms. The strength of the interaction energy correlated well with the experimental activity of the inhibitor. Active site residues Asp64, Thr66, Gln148, Glu152, Asn155, Lys156 and Lys159 and the Mg^{2+} ion were found to determine the anti-HIV-1 activity. They observed that the compounds showing highest anti-HIV-1 activity all show a high interaction between the positive charge of the Mg^{2+} ion and the quadrupole moment of the substrate aromatic ring, predicting that changing the position of (or removing) the pyrrole ring substituent will improve the activity of the inhibitor.

Nunthaboot *et al.* [315] have also used QM/MM MD techniques to study HIV-1 integrase diketo acid inhibitor complexes. They performed MM and PM3-CHARMM22 simulations for 1.2 ns for the HIV-1 integrase/5CITEP and HIV-1 integrase/DKA complexes. DKA and 5CITEP are similar in structure but in 5CITEP there is a tetrazole ring instead of a carboxylic acid group (Figure 21). DKA is 16-fold and 50-fold more

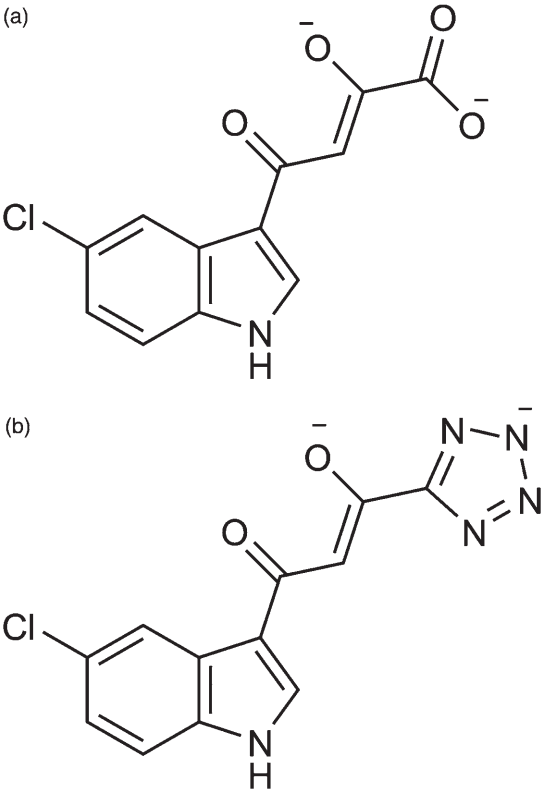

Figure 21. HIV-1 inteagrase is one of three enzymes essential for HIV replication and is an important target for the development of anti-AIDS drugs. HIV-1 integrase inhibitors (a) DKA and (b) 5CITEP modelled in complex with HIV-1 integrase by PM3-CHARMM22 MD simulations [315].

active than 5CITEP in the presence of Mg^{2+} for the 3′-processing and strand transfer reactions in HIV-1 integrase, respectively. They found quite significant differences in the structures obtained from MM and QM/MM simulations. In QM/MM simulations, Mg^{2+} is found in a distorted octahedral coordination whereas it is octahedral in the MM simulations. The mode of binding for 5CITEP is also different with the ligand forming a salt bridge with active site residue Lys159, not possible in the MM simulations. The two inhibitor complexes also show different dynamics. In the HIV-1 integrase/DKA complex, Asp64 interacts with the Mg^{2+} ion through both carboxylate oxygens whereas in the HIV-1 integrase/5CITEP only OD1 of Asp64 is bound to the Mg^{2+} ion. The interaction between the inhibitor and Lys159 was stronger in the DKA complex and may be the reason for its higher potency compared to 5CITEP.

Ruiz-Pernia *et al.* [316] have recently investigated the mechanism for the 3′-processing step in the reaction catalysed by HIV-1 integrase. In the reaction studied, a water molecule is activated by an active site aspartate and then attacks the phosphodiester substrate causing one ester group to leave, abstracting a proton from another water molecule. A PM3-OPLS-AA potential energy surface for the reaction was generated by scanning

along two reaction coordinates, showing that the reaction occurs by a stepwise mechanism. Proton transfer from Wat 1 to Asp64 involves a barrier of ~33 kcal mol^{-1} and the second proton transfer from Wat 2 to the methoxide leaving group involves a barrier of ~32 kcal mol^{-1}. Umbrella sampling simulations were performed to generate free energy profiles for the reaction. The first reaction coordinate was simulated in 63 windows with 10 ps production QM/MM MD at each point and the second reaction coordinate was followed in 73 simulation windows. High-level MP2/6-31G(d,p) energy corrections and zero-point energy effects were also calculated giving a free energy barrier of 17.6 kcal mol^{-1} for the first step and 22.6 kcal mol^{-1} for the second.

4.10. *Acetylcholinesterase*

AChE is an important enzyme in neurotransmission. It hydrolyses the neurotransmitter acetylcholine (Figure 22a), whereby the impulse transmissions are terminated. Malfunctioning of this enzyme and other components of cholinergic synapses are involved in several important human diseases, such as glaucoma, Alzheimer's and Parkinson's. Interestingly, the action of chemical warfare agents such as sarin and tabun involves the inhibition of AChE, which can also prove fatal. The enzyme is therefore an important target for drug design. AChE belongs to the serine hydrolase family that relies on the well-known Ser–His–Glu (or Asp) catalytic triad for their action. In the acylation reaction, serine serves as a nucleophile that attacks the carbonyl oxygen of the acetylcholine ester, while its hydroxyl proton is taken up by histidine. This results in a tetrahedral intermediate. After taking up a proton from the histidine and elimination of a primary alcohol, the acyl-enzyme is formed. Zhang *et al*. [317] calculated a potential energy barrier of 10.5 kcal mol^{-1} for the acylation at the MP2/6-31 + G(d)-MM//HF/3-21G-MM level

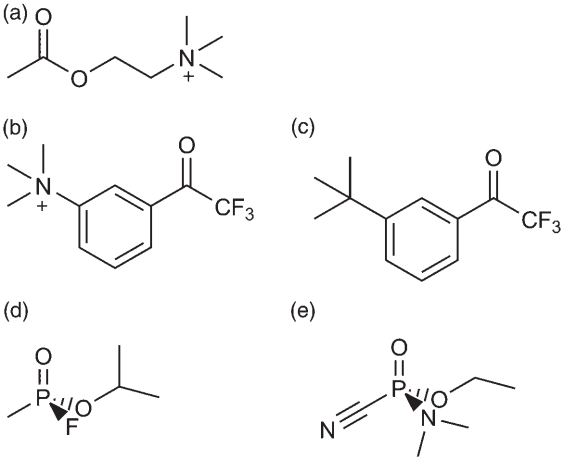

Figure 22. AChE hydrolyses the neurotransmitter acetylcholine. Malfunctioning of this enzyme and other components of cholinergic synapses are involved in several important human diseases, such as glaucoma, Alzheimer's and Parkinson's. The structure of (a) acetylcholine, (b) TFK$^+$, (c) TFK^0, (d) sarin and (e) tabun. TFK$^+$ and TFK^0 are inhibitors of AChE modelled by Cheng *et al*. [319] using QM/MM-FE techniques. The action of chemical warfare agents sarin and tabun also involves inhibition of AChE, recently modelled by Kwasnieski *et al*. [321].

in good agreement with the experimental value of 12 kcal mol^{-1}. The exact role of Glu334 (from the catalytic triad) in the mechanism has been debated in the past; mechanisms such as 'low-barrier' hydrogen bond formation and 'charge-relay' where it acts as a general acid–base catalyst have been suggested. Zhang *et al.* [317] found that this residue plays an important role stabilising the TS, and in the acyl-enzyme intermediate, through electrostatic interactions in agreement with the EVB study of Fuxreiter and Warshel [318].

TFK$^+$ and TFK^0 (Figure 22b and c) are inhibitors of AChE, with the charged species showing a slightly higher binding affinity than the neutral analogue with the wild-type enzyme. Interesting experimental studies involving the mutation of the catalytic His447 to an isoleucine showed that TFK$^+$ was still observed to react with the mutant enzyme, whereas the neutral species did not. Cheng *et al.* [319] used a combination of MD simulation and B3LYP/6-31G(d)/AMBER95 QM/MM-FE techniques to model the reaction of TFK$^+$ and TFK^0 with the mutant enzyme. They observed that a water molecule adopted a similar position to that of the imidazole ring of His447, forming strong interactions with Ser203 and Glu334 of the catalytic triad. Such behaviour was never observed in equivalent simulations with the TFK^0 inhibitor. To see whether the water molecule could perform a similar function to His447 in the original catalytic triad, three QM/MM-FE simulations were performed. A reaction coordinate was used to force the water molecule to accept one proton from Ser203 and donate another proton to Glu334. The height of the barrier to reaction varies depending on the protein configuration and level of QM theory, but overall the energy barrier is about 8 kcal mol^{-1}, indicating that the proposed reaction is feasible. These QM/MM-FE and MD studies therefore explain how TFK$^+$ is still able to demonstrate a high binding affinity to a mutant AChE enzyme in which the histidine residue of the catalytic triad is replaced by isoleucine.

Sarin and tabun (Figure 22d and e) are potent nerve agents that inhibit AChE irreversibly as a covalent bond is formed between the oxygen of Ser203 and the phosphorus of the nerve agent. Understanding the mechanism of inhibition will be important for the design of effective drugs to treat organophosphorus poisoning. AChE phosphonylation by sarin has been studied using QM methods on a small model system, showing that the mechanism proceeds via an addition–elimination pathway [320]. Kwasnieski *et al.* [321] have used QM/MM methods to study the inhibition of AChE by tabun and investigate why one enantiomer of tabun shows a higher potency than the other. Density functional theory was used for the QM part (BP86/TZVP) with the CHARMM27 force field used for the MM region. As in the mechanism of sarin, the inhibition of the reaction proceeds via an addition–elimination pathway, with cyano group departure as the rate-limiting step as tabun fixation on Ser203 is almost barrierless. The *S*-Syn enantiomer was found to be the most active enantiomer. Another interesting observation was that the cyano group does not leave *anti* to the oxygen of Ser203 due to repulsive interactions between cyanide and several active site phenylalanine residues.

4.11. *Avian influenza virus H5N1 enzymes*

The avian influenza virus subtype H5N1 (HPH5) is highly pathogenic and poses a serious threat to human health. QM/MM methods have been used to investigate the first step of the cleavage of the HPH5 virus by furin [322]. The catalytic triad of Asp102–His194–Ser368 in furin is the same as in other serine proteases, and it has been proposed that

the reaction mechanism will be similar. Thus His194 would deprotonate Ser368, which would then attack the carbonyl carbon of the S1–Arg on the HPH5's cleavage site leading to the formation of a tetrahedral intermediate. Pathways were generated at the PM3-CHARMM22 level, indicating that the reaction is concerted. The energy barrier of 33.5 kcal mol^{-1} at the PM3-CHARMM22 level was reduced to 16.2 kcal mol^{-1} with gas-phase B3LYP/6-31+G(d) corrections. They found that Asp153 provides the most significant electrostatic stabilisation to the TS and that the oxyanion hole formed by Asn295 and Ser368 stabilises the tetrahedral intermediate.

4.12. *Computational design of enzymes and catalytic antibodies*

Computer-aided design of enzymes and catalytic antibodies is a tantalising goal in enzymology [323–326], with the potential for industrial and medical application. Understanding of factors affecting catalysis will aid the design or prediction of mutations that may improve the catalytic activity of an enzyme or antibody, and therefore QM/MM methods will be likely to play an important role in the future of enzyme design. Martí *et al.* [327] have used QM/MM techniques to propose active site mutations to a catalytic antibody (1F7) generated to catalyse the conversion of chorismate to prephenate (Section 4.2). These workers calculated free energy profiles for the chorismate to prephenate rearrangement in aqueous solution, in chorismate mutase (from *B. subtilis* and *Escherichia coli*), and in the catalytic antibody 1F7, in good agreement with the observed activities of the different species. Comparing interactions in the active sites of the different enzymes and the antibody they predicted that the mutation of AsnH33 to serine will improve the efficiency of the catalytic antibody. The free energy profile for the mutant antibody is 4.5 kcal mol^{-1} lower than that predicted for the wild-type antibody, and only 2.4 kcal mol^{-1} higher than the free energy barrier for the wild-type BsCM enzyme.

The enzyme isochorismate pyruvate lyase (IPL) primarily catalyses the conversion of isochorismate into pyruvate and salicylate, but also exhibits secondary activity for the conversion of chorismate into prephenate. IPL reduces the barrier to the rearrangement of chorismate by 2.3 kcal mol^{-1} compared to the barrier in solution. Using a protocol similar to that described in their study of antibody 1F7, Martí *et al.* [328] predicted that mutation of Ala38 to valine would improve activity for the chorismate to prephenate rearrangement in IPL. The free energy barrier in the mutant IPL is 4.4 kcal mol^{-1} lower than that in wild-type IPL, and 2 kcal mol^{-1} higher than that for BsCM.

A series of artificial enzymes and catalytic antibodies have been designed to catalyse the Kemp elimination reaction of 5-nitrobenzioxazole [329] (Figure 23). In solution, the reaction is thought to follow an E2 mechanism with concerted proton transfer and N–O bond cleavage. The enzymes use Asp, Glu or His residues as the catalytic base, with some enzymes also containing an acidic residue (or residues) near the N–O bond. QM/MM Monte Carlo FEP calculations at the PDDG/PM3-OPLS-AA QM/MM level have been used to calculate free energy surfaces for the reaction in solution and the artificial enzymes KE07, KE10(V131N), KE15 and KE16 [330]. Two reaction coordinates were perturbed during the simulations: one to describe the proton transfer and one to describe the N–O bond cleavage. Each FEP window was sampled for 5×10^6 configurations of solvent equilibration, 10×10^6 configurations of equilibration of the entire system and 25×10^6 configurations. The enzymes KE07, KE10(V131N) and KE15 were predicted to be

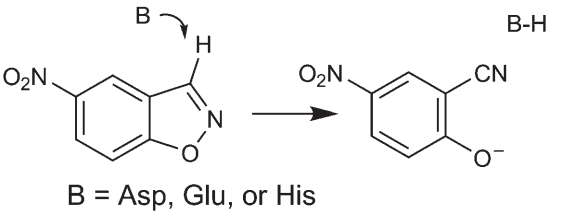

Figure 23. A series of artificial enzymes and catalytic antibodies have been designed to catalyse the Kemp elimination reaction of 5-nitrobenzioxazole [329]. QM/MM Monte Carlo FEP calculations at the PDDG/PM3-OPLS-AA QM/MM level have been used to calculate FE surfaces for the reaction in solution and the artificial enzymes KE07, KE10(V131N), KE15 and KE16 [330] and also the catalytic antibody 34E4 and the E50D variant [331].

catalytic, but the results for KE16 were inconclusive. The free energy surfaces agree that the reaction is a concerted process, but it is asymmetric with the proton transfer more advanced in the TS than the cleavage of the N–O bond. The authors suggest that the activity of the artificial enzymes could be further optimised by increasing the basicity of the catalytic base and better positioning of the substrate to stabilise the breaking N–O bond at the TS.

The catalytic antibody 34E4 and the E50D variant that catalyse the same reaction have also been studied using QM/MM Monte Carlo methods [331] using a similar protocol to that described above. Antibody 34E4 shows significant catalytic activity, with Glu50 acting as the catalytic base; mutation to Asp leads to a 30-fold reduction in catalytic activity. Free energy surfaces were calculated for the reaction in the two species predicting a 2.4 kcal mol^{-1} difference in activation energy. Key structural differences were identified that give rise to the different activities of the two antibodies. The relative positions of the substrate and the base were found to be less optimal in the mutant antibody. Residue Trp91 is unable to stabilise the developing negative charge on the substrate in the E50D mutant. Asn58 and some water molecules form hydrogen bonds with the substrate and catalytic base, not observed in the native species.

4.13. *QM/MM methods in the interpretation of the vibrational spectra of biomolecules: bacteriorhodopsin*

Bacteriorhodopsin in a transmembrane protein that functions as a light-driven proton pump [332–334]. Photoexcitation of the Schiff-base *trans*-retinal initiates a photocyle, which involves several intermediate states (J, K, L, M, N and O) before return to the resting state. During the photocycle, bacteriorhodopsin transports protons from the inside of the membrane to the outside in a unidirectional manner. Light absorption causes the isomerisation of *trans*-retinal to 13-*cis*-retinal giving rise to the K intermediate. The K state undergoes relaxation processes to the L intermediate and a proton is released to the extracellular bulk during the L to M transition. Crystal structures of several intermediates in the pathway have been solved [335–338], but they do not offer much insight into proton storage sites in this enzyme. Several Fourier transform infrared (FTIR) studies have been carried out on bacteriorhodopsin [339–350], but the interpretation of the spectra is not trivial. QM/MM calculations can be used to calculate FTIR spectra to provide a way to link the experimental data with what is happening on the atomic level.

X-ray crystal structures of the ground state of bacteriorhodopsin [335,336] have shown that three water molecules form a hydrogen-bonded network in the active site. Structures of other intermediates in the photocycle [337,338] show the dislocation of one or two of these water molecules from the active site, which was proposed to play an important role in the proton transfer network. The strength of hydrogen bonds formed by these water molecules during the photocycle has been the subject of FTIR experiments [342,344,351]. The experimental spectrum for the resting state of bacteriorhodopsin was obtained using substitution with D_2O, with bonds sensitive to substitution by D_2O displaying O–D or N–D stretching vibrations widely distributed in the 2100–2700 cm^{-1} region, suggesting that several different hydrogen bonding environments are present [344]. No N–D signal for the Schiff base could be assigned for the experimental spectrum, due to the angular dependence of the IR intensities. In an early study, Hayashi and Ohmine [352] carried out an *ab initio* (HF/DVZ-AMBER) calculation of the proton transfer mechanism in bacteriorhodopsin, including a normal mode analysis. Six O–D stretching modes were obtained for the three water molecules (WA, WB and WC; Figure 24) in the region 2257–2721 cm^{-1}. WA forms a weak hydrogen bond with Asp212 and a stronger hydrogen bond with Asp85 with O–D stretching frequencies assigned at $\nu_2 = 2618\,cm^{-1}$ and $\nu_6 = 2257\,cm^{-1}$, respectively. WB forms a hydrogen bond to Asp85 giving rise to the $O–D_2$ stretching frequency of $\nu_5 = 2320\,cm^{-1}$ and the other $O–D_1$ frequency is higher ($\nu_1 = 2721\,cm^{-1}$) as it does not participate in any hydrogen bond. WC has symmetric and antisymmetric stretching frequencies of $\nu_4 = 2468\,cm^{-1}$ and $\nu_3 = 2601\,cm^{-1}$, respectively. One limitation of the model is that Arg82, which forms a hydrogen bond with WC, is treated by MM and hence the modes associated with WC may not be well described.

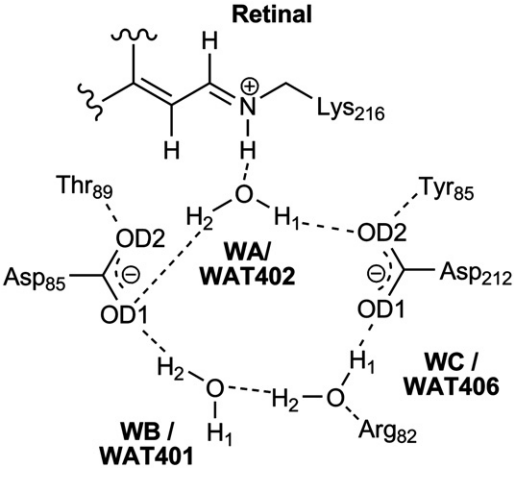

Figure 24. Bacteriorhodopsin is a light-driven proton pump, converting the light absorbed by retinal into a transmembrane proton gradient. Proton storage sites can be difficult to identify by experimental techniques alone. It has been proposed that water clusters could play an important role in the transport of the protons in bacteriorhodopsin. *Ab initio* QM/MM methods have been used to calculate the IR spectra for water clusters in the active site of bacteriorhodopsin in the resting state [352] and the K intermediate [356] of the photocycle. The hydrogen bonding environment of the water molecules is indicated by dotted lines.

The angles of IR intensities are quite different from the angles of the O–D bonds due to the sensitivity of the IR intensity to vibrational mixing and to electron-vibration interaction. A value of $\nu_{ND} = 2113\,cm^{-1}$ was obtained for the N–D stretching mode of the Schiff base. This mode is of low frequency due to the strong hydrogen bond with WA. The intensity of this band was large (46.8) and the angle of the IR intensity of 520° is very different from the angle of the bond (190°), due to the fact that the dipole derivatives with respect to the local N–D stretch mode is directed along the total dipole moment of the QM system instead of the N–D bond itself. Further experiments confirmed the assignment of the vibrational modes in the resting state of bacteriorhodopsin, including the assignment of the N–D stretching mode of the Schiff base. Subsequently, crystal structures [353,354] and a theoretical model [355] of the K intermediate in the photocyle were solved. QM/MM modelling, including normal mode analysis, was carried out [356] following the same protocol as used for the resting state to link protein structure to the FTIR difference spectrum of the K and resting states [357]. Upon the formation of intermediate K, the frequency of the N–D bonds upshifts by $420\,cm^{-1}$ to $\nu_{ND} = 2533\,cm^{-1}$, with a significant reduction in the IR intensity from 46.8 to 5.8. An upshift in the signal for the stretching frequencies of many of the O–D bonds of the water molecules is also observed suggesting a decrease in the strength of the hydrogen bonds. For example, the stretching frequency of WA/WAT401 O–D_2 was $\nu_6 = 2257\,cm^{-1}$ in the resting state and $\nu_4 = 2423\,cm^{-1}$ in the K intermediate, a shift of $166\,cm^{-1}$. The hydrogen bond between Wat402 and OD1 of Asp85 is 0.05 Å longer in the K intermediate than in the resting state. Overall, the spectra indicate that hydrogen bonds between the Schiff base and Wat402 and Wat402 and Asp85 (Figure 24) are weaker in the K intermediate than in the resting state, which has been proposed as a method for storage of the light energy absorbed by retinal [355,356,358,359].

QM/MM CPMD simulations have also been used to calculate spectral features of a pentagonal water cluster in the active site of bacteriorhodopsin in the resting state [360]. The model system consisted of bacteriorhodopsin embedded in a POPC (1-palmitoyl-2-oleyl-sn-glycero-3-phosphatidylcholine) membrane and fully solvated to form a $80 \times 80 \times 79\,Å^3$ box. The model system was equilibrated for 8 ns using classical MD employing the OPLS-AA force field. Three QM regions were tested: model A contained only the three water molecules, model B contained the three water molecules and the side chains of Asp85 and Asp212 and model C also includes the protonated retinal Schiff base and the side chains of Lys216 and Arg82. The QM region was treated using the BLYP functional with a dual basis set method combining Gaussian and plane wave basis functions. QM/MM CPMD simulations were carried out for 25 ps (5 ps equilibration + 20 ps production) for each QM/MM model. The IR spectra were calculated as the Fourier transforms of the dipole autocorrelation functions with harmonic quantum corrections. For models A and B, the pentagonal water cluster is stable throughout the simulations, whereas changes in hydrogen bonding networks are observed in simulations of model C. In simulations of model C after 5 ps, Arg82 forms a hydrogen bond with Asp212 causing W406 to move towards Asp85 and its hydrogen bond with W401 to break. The hydrogen bond between W406 and Asp212 switches its acceptor oxygen. A hydrogen bond is also formed between W401 and W406. Despite changes in hydrogen bonding, the water cluster still forms a pentagonal structure. The hydrogen bond formed betweeen Asp212 and Arg82 may be a consequence of the choice of QM region but the temperature factors of Arg82 are higher than other residues in this region [335], indicating that this conformation

may not be unreasonable. All models have absorptions in the range 2800–3650 cm^{-1}, typical for water molecules depending on their hydrogen bonding strength [361]. Only model C showed significant intensity below 3000 cm^{-1} on accordance with the experimentally observed broad abssorbance in this region [340]. In model C, where the Schiff base is also treated by QM methods, there is a very broad peak around 2800 cm^{-1} with a shoulder ranging down to 2500 cm^{-1}, which cannot be attributed to O–H strectching alone. It arises as a result of the mutual polarisation between the Schiff-base proton and W402 in the field of the two aspartic acids. When the spectral features due to Arg82 are considered, both conformations could give rise to the experimentally observed peaks. The results show that quite a large QM region is necessary to include intermolecular couplings and correctly predict the structural features of the IR spectrum of bacteriorhodopsin [360].

The origin of the proton released during the L–M transition remains a matter of debate, as it does not come from Asp85, which was previously protonated by a proton transfer from the Schiff base [341]. X-ray crystal structures and IR-studies have provided key data but cannot alone resolve this debate. It was proposed that Glu204 and/or Glu194 were the proton release group [362,363], but this was refuted because the expected shifts for the deprotonation of carboxylate groups of these residues were not observed in the IR spectra of the L to M transition in wild-type bacteriorhodopsin and a series of mutants. The decay of a continuum band around 2000 cm^{-1} (also observed in spectra of protonated water clusters [364]) during the rise of the M state led to the proposal that the proton is instead stored on a water cluster in the active site [339,340].

Rousseau *et al.* [365] have used QM/MM CPMD techniques to model protonated water clusters in bacteriorhodopsin in the L intermediate state embedded in a POPC (1-palmitoyl-2-oleyl-sn-glycero-3-phosphatidylcholine) membrane. They modelled $H_5O_2^+ \cdot (H_2O)_4$, a symmetrically solvated Zundel cation (which contains a shared proton between two water molecules) and $H_3O^+ \cdot (H_2O)_3$, a solvated Eigen cation (a localised proton at one water molecule) [366] embedded in the active site with a deprotonated Schiff base and a 13-*cis*,15-*anti*-retinal conformation. The system was equilibrated classically before 10 ps QM/MM CPMD simulations were carried out at the BLYP-MM level (GROMOS 43a2 force field). The IR spectrum was then obtained from the dipole autocorrelation function. Restraints were imposed on the Zundel complex to keep it as symmetrical as possible. For both models, hydrogen bonds were observed with Glu194, Glu204, Arg82 and Tyr57. There was only one prominent absorption for the Eigen complex extending between 2000 and 3000 cm^{-1}, assigned to linear combinations of symmetric and antisymmetric stretching modes of the central H_3O^+. The Zundel model did not have much spectral density in this region but instead had a very broad band covering the 500–2000 cm^{-1} region associated with a large amplitude motion of the symmetric hydrogen bond and the bending motion of the two terminal water molecules.

Mathias and Marx [367] have carried out QM/MM CPMD (BLYP-GROMOS 43a2 force field) simulations of protonated water clusters in bacteriorhodopsin in the resting state. Clusters of the form $H^+ \cdot (H_2O)_4$ and $H^+ \cdot (H_2O)_3$ were modelled as the QM region in simulations of 32 and 51 ps (after 5 ps equilibration), respectively. The clusters adopted structures more like a protonated water chain than the Zundel and Eigen forms modelled by Rousseau *et al.* [368]. The excess proton was found to mainly reside on W405 with some movement to W403 and W409* in the large cluster regardless of where it was

initially placed. The computed IR spectra for the two models are similar in the region up to 1700 cm^{-1}. In the region above 1800 cm^{-1}, both models exhibit a broad band with a flat maximum at ~1940 cm^{-1}. The peak of the $H^+ \cdot (H_2O)_4$ cluster is broader and more intense than that obtained for the $H^+ \cdot (H_2O)_3$ cluster. The band at 1900 cm^{-1} is due to the antisymmetric stretch of the water on which the proton is located and the band at 2500 cm^{-1} is due to the symmetric mode. Both spectra have intense peaks at 2940 and 2950 cm^{-1} and a weaker peak at 3350 cm^{-1} due to the stretching vibrations of O–H bonds that are hydrogen bonded to Glu194 and Glu204 and a free O–H bond. It is notable that the use of a fictitious mass in the CPMD method may red-shift these frequencies by ~100 cm^{-1}, this effect is less significant for frequencies below 3000 cm^{-1}. The peaks at 1940 cm^{-1} are in good agreement with the experimental spectrum [364], although slightly blue-shifted with respect to the experimental maximum. The authors propose that the spectrum for the $H^+ \cdot (H_2O)_4$ cluster is in better agreement with the experimental results as it carries more overall intensity than the $H^+ \cdot (H_2O)_3$ cluster [367].

QM/MM MD calculations of bacteriorhodopsin in the L intermediate state at the B3LYP/6-31+G(d,p)-CHARMM and SCC-DFTB-CHARMM levels have been used to identify the origin of the proton released into the extracellular bulk during the L to M transition [369]. QM/MM MD simulations were carried out with both the protonated water cluster and the side chains of Glu194 and Glu204 included in the QM region. At the B3LYP/6-31 + G(d,p)-CHARMM level, 5 ps simulations were carried out and at the SCC-DFTB-CHARMM level 5 separate simulations were carried out for 2 ns per trajectory. The IR spectrum was computed from the Fourier transform of the dipole autocorrelation function collected from the SCC-DFTB-CHARMM MD simulations. The nuclear quantum effect on the IR spectrum was approximated with a harmonic quantum correction factor. During the QM/MM MD simulations at both the B3LYP/6-31 + G(d,p)-CHARMM and SCC-DFTB-CHARMM levels, the excess proton on the water cluster was unstable and the proton moved to become delocalised between Glu194 and Glu204. Analysis of distances during the MD simulations showed that the excess proton is shared by the pair of glutamates with a slight preference for Glu204. The average distance of the two oxygens of Glu204 and Glu194 was ~2.5 Å at both levels of QM/MM theory, in good agreement with distances found in crystal structures, e.g. [370,371] (see Table 1 in [369]). As an independent test, QM/MM MD simulations were also carried out treating only the protonated water cluster in the QM region as used in QM/MM CPMD simulations [367]. With this QM/MM model, the oxygen–oxygen distances for the two glutamates increased to distances significantly longer than those observed in high-resolution crystal structures (see Table 1 in [369]).

The calculated IR spectrum showed a continuum band in the 1800–2000 cm^{-1} region in agreement with the experimental observation [339]. The continuum band was found to be absent if one of the glutamate residues is treated by MM rather than QM, suggesting that the continuum band in this region arises due to the delocalisation of the proton between the two glutamates. The ν_{COOH} band of Glu194/204 is significantly red-shifted by this strong hydrogen bond to $< 1700\ cm^{-1}$, a region normally associated with amide bands. This is also a possible explanation of why this band could not be identified in experimental spectra [347,350]. Several mutant forms of bacteriorhodopsin have also been studied as part of the investigation into the nature of the proton release group [339,347]. These spectra of these mutant enzymes have also been calculated following the same

QM/MM protocol. In simulations of a E194D mutant enzyme, the excess proton becomes largely localised on Glu204 and the continuum band in the IR spectrum is no longer present in agreement with the experimental spectrum [339]. In simulations of a E204D mutant the proton remains localised mostly on Asp204 but there is some sharing of the proron with Glu194 as a weak continuum band is observed in the 1800–2000 cm^{-1} region also in agreement with the experimentally obtained spectrum [339]. Mutant enzymes where the mutation is in the region of the two gluatmates have also been studied: S193A, Y83F and R82Q [339]. In these systems, the proton remains delocalised between the two glutamates as in the wild-type enzyme. The band in the 1800–2000 cm^{-1} region is observed for all three mutants but with different intensities as is observed in the experimental studies [339]. The authors suggest that using this form of 'low-barrier hydrogen bond' (Section 4.9) to stabilise two repulsive groups might be advantageous in terms of balancing binding affinity and rapid proton release in proton pumps of this type [369].

4.14. *X-ray crystal structure and NMR structure refinement using QM/MM methods*

The determination of an X-ray crystal structure typically involves several steps: crystallisation, data collection, phase determination, model building, refinement and validation of the model [372,373]. To produce an accurate model, several iterations of refinement must usually be carried out, in which the coordinates, perhaps occupancies and B-factors are changed to improve the fit to the data [372]. Minimisation or simulated annealing calculations using an MM force field are often part of the process. The accuracy of the model could be improved by including a QM/MM calculation in the refinement process rather than an MM calculation, particularly for substrate or inhibitor atoms and metal centres, which may not be treated very accurately by the MM force field [374–377]. Ryde *et al*. [374] have incorporated a widely used crystallographic refinement program (CNS: crystallography and NMR system [378,379]) in their QM/MM program ComQum-X. In their protocol, the QM/MM calculation takes the form of a subtractive scheme (Equation (2)) as in type A models described in Section 2.1. This method, which they call quantum refinement, has been tested on enzymes such as ferrochelatase [374], cytochrome c553 [375,376] and alcohol dehydrogenase [380]. For cytochrome c553, crystal structures obtained under the same conditions are available at different resolutions: 1.7 Å and 0.97 Å [381]. This enzyme contains a haem group where the Fe^{3+} moiety is liganded by a histidine and a methionine residue and the two crystal structures show quite significant differences in the iron–ligand bond lengths: in the lower resolution structure the $Fe–N_{HIS}$ distance is 2.31 Å and in the high-resolution structure this distance is 1.99 Å, likewise the $Fe–S_{MET}$ distance is 2.21 Å in the low-resolution structure and 2.33 Å in the high-resolution structure. The structure of the haem group from the low-resolution structure was optimised using ComQum-X at the BP86/6-31G(d)-MM level [375,376]. The positions of the haem group in the ComQum-X structure are much closer to those in the high-resolution data, showing that it is possible to improve lower resolution X-ray structures by including QM/MM calculations in the refinement process.

QM/MM calculations can also be useful in the interpretation of crystal structures, e.g. in determining the protonation states of metal-bound solvent molecules. Metal–ligand bond lengths are affected by the nature of other ligands making it difficult to distinguish between a metal–OH bond and a metal–H_2O bond. In the case of alcohol dehydrogenase,

the pK_a of the solvent molecule bound to the zinc ion is known from kinetic measurements [382]. Nilsson and Ryde [380] have studied a complex between alcohol dehydrogenase, NAD^+ and trifluoroethanol at 2 Å resolution, in which the alcohol should be present as an alkoxide ion. They calculated structures with ComQum at the BP89/6-31G(d)-MM level with both the alkoxide and alcohol forms and found that the alkoxide form does give a much better fit to the experimental data in terms of several parameters. In another structure of alcohol dehydrogenase where experimental data indicate that a water molecule is bound to the zinc ion [383], the results of QM/MM calculations also show a better fit to the experimental data for a water molecule rather than a hydroxide ion.

Yu *et al*. [384] have also developed a QM/MM method for the refinement of X-ray crystal structures in which the gradients from the QM/MM energy function are combined with those from the X-ray target to refine the crystal structure. This method also uses the X-ray refinement program CNS [378,379], but it is coupled to AMBER [49] to perform QM/MM calculations using the linear scaling semiempirical electronic structure program DivCon [385–388] and the SANDER module. The QM/MM X-ray refinement method is used to generate a set of structures that satisfy the X-ray data equally well and their stability is scored using a more accurate energy function with a continuum representation of the bulk solvent. This method has been applied to β-secretase bound to an inhibitor, an aspartyl protease that is a key target in the pathogenesis of Alzheimer's disease, to assign the protonation states of the important aspartate residues [384]. Eight different protonation patterns of the aspartyl Asp315/Asp32 dyad were subjected to the QM/MM refinement process. The model consisted of 13,195 atoms with 68–70 treated by the AM1 QM method in the QM/MM refinement process. The models were refined by 1000–1500 steps with the MM force field under the restraints derived from the X-ray data before being refined by 700–1000 steps of QM/MM refinement. The relative stability of the structures, compared to AspH/Asp in solution, was evaluated through the use of a thermodynamic cycle calculated using QM calculations in the presence of a self-consistent reaction field. It was found that the monoprotonated form with Asp32 protonated on the inner oxygen was the most favourable configuration.

Protein–ligand complexes obtained by X-ray crystallography play an important role in structure-based drug design [307,389,390], so it is important that the structures of the ligands are accurately represented during the crystallographic refinement process, something that can be difficult with traditional (e.g. MM based) refinement techniques [391,392]. Distorted geometries of ligands or protein–ligand clashes are not uncommon especially in structures refined against low to medium resolution crystallographic data [393]. Li *et al*. [394] have used their QM/MM refinement method with *ab initio* QM methods to study conformational variability of benzamidinium-based inhibitors in protein–ligand complexes. Benzamidinium compounds are serine protease inhibitors and they can be divided into three classes depending on their substitution: unsubstituted amidines, singly substituted at the *para* or *meta* positions or disubstituted amidines with different substitution patterns. In the Protein Data Bank (PDB: www.rscb.org/pdb) [395,396] there are 87 structures of the first type, 60 of the second type and 6 of the third class of benzamidinium inhibitors. These structures were analysed to find the dihedral angle distribution describing the orientation of the phenyl ring with respect to the amidinium moiety, finding one group corresponding to planar conformations and another

corresponding to twisted conformations [394]. Data for the amidines in the Cambridge Structural Database (CSD) [397,398] give only non-planar structures, showing that there may be significant differences in the bound geometries of these ligands or errors in the description of these molecules during the refinement process. Three crystal structures were selected from the PDB as the starting geometries for the QM/MM refinement process: 1RKW a quaternary ammonium compound repressor QacR complexed with pentadiamidine, 1RPW a QacR with hexadiamidine bound [399] and 1Y3X a trypsin–ligand complex [400]. The inhibitors in these complexes are all predicted to adopt planar conformations in the original crystal structures. Different refinement methods predict significantly different values for this torsion angle. For example, in the complex of QacR with pentadiamine the crystal structure predicts a value of 1.60° for the torsion of one of the benzamidinium groups located in binding site 1, standard refinement with CNS predicts a torsion of 0.23° (both planar geometries) and QM/MM refinement predicts a value of 34.06° for this torsion, a twisted geometry [394]. This may have important implications for structure-based drug design because incorrect conclusions could be drawn if the wrong geometry is used as a lead.

NMR spectroscopy is the other most important technique for protein structure determination. Data from NMR experiments provide information such as distances between pairs of atoms, constraints in dihedral angles and hydrogen bonds. These are converted to a three-dimensional structure by the use of distance geometry methods and restrained MD [401]. As in X-ray crystallography, the experimental data are supplemented with MM force field data to ensure that the structures are chemically reasonable, e.g. in terms of bond lengths or the planarity of aromatic rings. Similar problems exist for the treatment of small molecules such as ligands and inhibitors. Hsiao *et al.* [402] have developed a method for NMR structure refinement, similar to that applied for X-ray crystal structure refinement, using their QM/MM program QomQum-N. NMR refinement involves a compromise between the energy calculated by the MM force field and the sum of all the NMR restraint energies. The NMR restraints must be given a suitable weighting to make them comparable to the MM energy. A QM/MM energy term can be used to replace the standard MM term and improve the accuracy of the description of the ligand or metal centre. The procedure for NMR structure refinement also includes the calculation of electrostatic interactions; hydrogen atoms are typically included in the NMR model. During the refinement procedure, the geometry of the QM region is optimised keeping the MM region fixed and then the geometry of the MM part is optimised by extensive energy minimisation under the NMR constraints (note no restrained MD is carried out, unlike standard NMR refinement procedures). Hsiao *et al.* [402] have applied this methodology to the two calcium binding sites in the epidermal growth factor-like domains 3 and 4 in vitamin K-dependent protein S (EGF34). The calcium ion is known to have flexible geometric preferences, binding six to eight ligands with different metal–ligand bond lengths [403], making EFG34 an interesting test case for the method. The calcium binding sites are forms by Asp160, Val161, Glu163, Asn178 and Ile179 for one site and Asp202, Ile203, Glu205, Asn217 and Tyr 218 for the other. The second and fifth residues of each site bind through their backbone amide oxygen atom, whereas the other residues bind via their side chains. One water molecule is also bound to the calcium ions in both sites. Crystal structures of the EFG-like domain in human clotting factor IX predict six coordinate calcium ions arranged in a pentagonal

bipyramidal arrangement with one carboxylate group binding in a bidentate manner [404]. QomQum-N refinement was applied using 10 structures of each of the mono and bidentate binding modes of the Asp and Glu residues. QM calculations were performed using the BP86 density functional with a 6-31G(d) basis set. The coordination of the calcium ions was found to change in many of the simulations. In five calculations starting from the monodentate ligand structures the final refined structure contains a bidentate ligand (four with an Asp residue and one with a Glu). In calculations starting from structures with bidentate ligand binding, the coordination of the calcium also changes becoming monodentate in some structures and some showing an increase of the calcium coordination to seven with both Glu and Asp binding in a bidentate manner. Analysis of the energies showed that bidentate binding of Glu is unfavourable, but it was not possible to determine the binding of the Asp residue with any certainty. It is possible that the two modes of binding may interchange on the NMR timescale due to the small barrier associated with the process [405].

4.15. *Other recent examples of QM/MM studies*

The number of QM/MM modelling studies of enzymes and their reactions continues to increase and only a few studies can be mentioned here. The hydroxylation step in the mechanism of the 3-hydroxylation of *para*-hydroxybenzoate in the enzyme *para*-hydroxybenzoate hydrolase has been studied with QM/MM techniques at the density functional and *ab initio* levels of theory [30,406,407], giving results in excellent agreement with experiment. QM/MM modelling of *para*-hydroxybenzoate hydrolase [30,406,407] and the related phenol hydrolase [408], supports the proposal that aromatic substitution is rate determining in these enzymes and provides further evidence that TST is a good basis for understanding their reactivity. The electronic structure of Compound II in chloroperoxidase has been studied using QM/MM methods to interpret different iron species indicated by Mössbauer parameters [409]. The formation of a covalent link between haem and Trp51 in cytochrome c peroxidase supports the proposal of radical formation on Trp51 [410]. Racemisation reactions in proline racemase [411] and glutamate racemase [412] have been studied using QM/MM methods. Phosphoryl transfer reactions have been studied using QM/MM methods, for example applying a new version of the AM1 semiempirical method, which takes into account the *d* orbitals of the phosphorus atom [413] and also new parameters for the SCC-DFTB method [414]. Conformational transitions have also been found to play a role in ATP hydrolysis in myosin [415]. Phosphate transfer in human uridine-cytidine kinase 2 has been studied by QM/MM techniques [416], as it has the mechanism for the phosphorylation of ligands in mitogen-activated kinases [417]. Multiple reaction pathways for the reductive half reaction of aldehyde oxidoreductase in the formation of acetaldehyde have been calculated using DFT-MM methods [418], showing the catalytic effect of Glu869 in lowering the barrier to nucleophilic attack so that hydride transfer becomes rate limiting. A new method for umbrella sampling simulations has recently been introduced where important degrees of freedom are identified to optimise the reaction coordinate [157]. A QM/MM metadynamics study has been performed to investigate the mechansim of cellulose hydrolysis by inverting GH8 endoglucanases, showing a boat-like TS and the presence of an oxocarbenium ion [188]. Car-Parrinello QM/MM MD simulations have been used to study the reaction mechanism

of farnesyltransferase [419], a target for anticancer drug design. QM/MM methods have also been used to describe long-range protein–protein electron transfer pathways in enzymes such as cytochrome P450cam, ascorbate peroxidase and cytochrome c peroxidase [420].

Some enzymes have become 'guinea pigs', model systems that allow detailed comparison not only with experimental data, but also between different modelling approaches. Such enzymes include include chorismate mutase [16,133,156,208,210,212,226,235], citrate synthase [11,76,129,202,203,421], *para*-hydroxybenzoate hydroxylase [30,109,407,422,423], cytochrome P450cam [276,278–280] and FAAH [135,267,270,424]. Agreement between mechanistic findings for several enzymes based on quite different QM/MM techniques (e.g. FAAH, discussed in Section 4.5) provides validation of the results and indicates that reliable conclusions can be drawn from careful QM/MM modelling. Given the complexity of enzymes and of QM/MM reaction simulations, it is important to test how the setup of the model (e.g. choice of structure, solvation, protonation states, parameters) and details of the calculation (e.g. level of QM theory, MM force field, method for reaction modelling) influence the calculated reaction energetics. This can analyse uncertainties in calculations, and provide error estimates, helping to solidify mechanistic conclusions. When testing the appropriate level of QM theory to use in QM/MM simulations of enzymic reactions, high-level QM/MM methods [30,407] can provide a benchmark for testing and validating lower-level computational approaches.

5. Summary and conclusions

Modelling enzymes, and in particular the mechanisms of the reactions they catalyse, presents many challenges. QM/MM techniques have become powerful tools that can provide detailed information about many aspects of enzyme reactions. QM/MM modelling has provided insight into enzyme catalysis, enzyme mechanisms, the effect of mutations and the basis of differential reactivity and selectivity highlighted by some of the studies of enzymes reviewed here. QM/MM modelling has been used to provide crucial evidence to support or dispute proposed enzymatic reaction mechanisms, and is becoming a key part of assessing the validity of mechanistic proposals. Advances in computer hardware and software mean that an unprecedented level of accuracy is now possible for the QM part of the calculation. Many challenges remain for the future development of QM/MM methods. The extensive configurational sampling required to obtain converged free energies means that most QM/MM MD or Monte Carlo simulations are still limited to semiempirical QM/MM techniques. Approaches based on high-level QM corrections provide a practical way to overcome this problem. Improvements in the MM treatment such as polarisation and the better treatment of electrostatics could also improve accuracy in QM/MM calculations. Careful consideration of aspects of model preparation such as protonation states is also a key factor in the reliability of a QM/MM model.

Computational enzymology has matured rapidly over the past two decades [6,205], and is likely to have a significant and growing impact in many areas. For example, the ability to model mutations accurately will allow mutational studies initially to be performed using computational methods, which will greatly assist the analysis of genetic differences associated with disease or drug metabolism. 'Designer' enzymes for specific synthetic applications remain an important practical goal [329] and will be an area of

significant growth in the future. The potential accuracy and flexibility of QM/MM methods will mean that their use will become more widespread in fields such as structure-based drug design [6,308]. Overall, it is clear that QM/MM techniques have the ability to provide useful information about fundamental aspects of enzyme-catalysed reactions. The importance and impact of these methods in enzymology and biochemistry more widely looks certain to increase in future [425].

Acknowledgements

A.J. Mulholland is an EPSRC Leadership Fellow, and (with KER) thanks the EPSRC for support. A.J. Mulholland thanks his co-workers in the work described here. Some of the works reviewed here were carried out using the computational facilities of the Advanced Computing Research Centre, University of Bristol (http://www.acrc.bris.ac.uk/acrc/index.htm). We would also like to thank Dr Richard Lonsdale, Dr Alessio Lodola, Dr Marc van der Kamp, Dr Johannes Hermann and Ian Grant for help with the preparation of the figures.

References

[1] A. Warshel and W. W. Parson, Q. Rev. Biophys. **34**, 563 (2001).
[2] M. J. Field, J. Comput. Chem. **23**, 48 (2002).
[3] A. Warshel, Annu. Rev. Biophys. Biomol. Struct. **32**, 425 (2003).
[4] M. Garcia-Viloca, J. Gao, M. Karplus, and D. G. Truhlar, Science **303**, 186 (2004).
[5] A. Warshel, P. K. Sharma, M. Kato, Y. Xiang, H. B. Liu, and M. H. M. Olsson, Chem. Rev. **106**, 3210 (2006).
[6] A. J. Mulholland, Drug Discov. Today **10**, 1393 (2005).
[7] H. Lin and D. G. Truhlar, Theor. Chem. Acc. **117**, 185 (2007).
[8] P. Sherwood, B. R. Brooks, and M. S. P. Sansom, Curr. Opin. Struct. Biol. **18**, 630 (2008).
[9] H. M. Senn and W. Thiel, Angew. Chem. Int. Ed. **48**, 1198 (2009).
[10] W. W. Cleland, P. A. Frey, and J. A. Gerlt, J. Biol. Chem. **273**, 25529 (1998).
[11] A. J. Mulholland, P. D. Lyne, and M. Karplus, J. Am. Chem. Soc. **122**, 534 (2000).
[12] M. A. Porter and P. A. Molina, J. Chem. Theor. Comput. **2**, 1675 (2006).
[13] M. Štrajbl, A. Shurki, M. Kato, and A. Warshel, J. Am. Chem. Soc. **125**, 10228 (2003).
[14] S. Hur and T. C. Bruice, Proc. Natl Acad. Sci. U.S.A. **99**, 1176 (2002).
[15] S. Hur, Z. E. R. Newby, and T. C. Bruice, Proc. Natl Acad. Sci. U.S.A. **101**, 2730 (2004).
[16] K. E. Ranaghan and A. J. Mulholland, Chem. Commun. 1238 (2004).
[17] X. D. Zhang, X. H. Zhang, and T. C. Bruice, Biochemistry **44**, 10443 (2005).
[18] M. H. M. Olsson, W. W. Parson, and A. Warshel, Chem. Rev. **106**, 1737 (2006).
[19] S. Hammes-Schiffer, Acc. Chem. Res. **39**, 93 (2006).
[20] L. Masgrau, A. Roujeinikova, L. O. Johannissen, P. Hothi, J. Basran, K. E. Ranaghan, A. J. Mulholland, M. J. Sutcliffe, N. S. Scrutton, and D. Leys, Science **312**, 237 (2006).
[21] H.-H. Limbach, J. Miguel Lopez, and A. Kohen, Philos. Trans. R. Soc. B Biol. Sci. **361**, 1399 (2006).
[22] E. Hatcher, A. V. Soudackov, and S. Hammes-Schiffer, J. Am. Chem. Soc. **129**, 187 (2007).
[23] J. P. Klinman, Chem. Phys. Lett. **471**, 179 (2009).
[24] J. Villa, M. Štrajbl, T. M. Glennon, Y. Y. Sham, Z. T. Chu, and A. Warshel, Proc. Natl Acad. Sci. U.S.A. **97**, 11899 (2000).
[25] A. Warshel and M. Levitt, J. Mol. Biol. **103**, 227 (1976).
[26] U. C. Singh and P. A. Kollman, J. Comput. Chem. **7**, 718 (1986).
[27] M. J. Field, P. A. Bash, and M. Karplus, J. Comp. Chem. **11**, 700 (1990).

[28] M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, and J. J. P. Stewart, J. Am. Chem. Soc. **107**, 3902 (1985).
[29] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, J. Comput. Chem. **4**, 187 (1983).
[30] F. Claeyssens, J. N. Harvey, F. R. Manby, R. A. Mata, A. J. Mulholland, K. E. Ranaghan, M. Schutz, S. Thiel, W. Thiel, and H. J. Werner, Angew. Chem. Int. Ed. **45**, 6856 (2006).
[31] F. Himo and P. E. M. Siegbahn, Chem. Rev. **103**, 2421 (2003).
[32] P. E. M. Siegbahn and T. Borowski, Acc. Chem. Res. **39**, 729 (2006).
[33] P. E. M. Siegbahn and F. Himo, J. Biol. Inorg. Chem. **14**, 643 (2009).
[34] A. Warshel, *Computer Modeling of Chemical Reactions in Enzymes and Solutions* (John Wiley & Sons, New York, 1997).
[35] J. Bentzien, R. P. Muller, J. Florian, and A. Warshel, J. Phys. Chem. B **102**, 2293 (1998).
[36] J. Gao, Acc. Chem. Res. **29**, 298 (1996).
[37] P. D. Lyne, M. Hodoscek, and M. Karplus, J. Phys. Chem. A **103**, 3462 (1999).
[38] J. L. Gao and C. Alhambra, J. Chem. Phys. **107**, 1212 (1997).
[39] A. T. Hadfield and A. J. Mulholland, Int. J. Quant. Chem. (Biophys. Q.) **73**, 137 (1999).
[40] C. L. Brooks III, M. Karplus, and B. M. Pettitt, *Proteins: A Theoretical Perspective of Dynamics, Structure and Thermodynamics* (Wiley, New York, 1988).
[41] W. Im, S. Berneche, and B. Roux, J. Chem. Phys. **114**, 2924 (2001).
[42] P. Schaefer, D. Riccardi, and Q. Cui, J. Chem. Phys. **123**, 014905 (2005).
[43] D. Bakowies and W. Thiel, J. Phys. Chem. **100**, 10580 (1996).
[44] F. Maseras and K. Morokuma, J. Comp. Chem. **16**, 1170 (1995).
[45] T. Matsubara, S. Sieber, and K. Morokuma, Int. J. Quant. Chem. **60**, 1101 (1996).
[46] T. Vreven and K. Morokuma, Theor. Chem. Acc. **109**, 125 (2003).
[47] T. Vreven, K. S. Byun, I. Komaromi, S. Dapprich, J. A. Montgomery, K. Morokuma, and M. J. Frisch, J. Chem. Theory Comput. **2**, 815 (2006).
[48] W. D. Cornell, P. Cieplak, C. I. Bayly, I. Gould, R. Merz, K. M. Freguson, D. M. Spellmeyer, D. C. Fox, T. Caldwell, J. W. M. Kollman, and P. A., J. Am. Chem. Soc. **117**, 5179 (1995).
[49] D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, and R. J. Woods, J. Comput. Chem. **26**, 1668 (2005).
[50] A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus, J. Phys. Chem. B **102**, 3586 (1998).
[51] W. L. Jorgensen, D. S. Maxwell, and J. TiradoRives, J. Am. Chem. Soc. **118**, 11225 (1996).
[52] G. A. Kaminski, R. A. Friesner, J. Tirado-Rives, and W. L. Jorgensen, J. Phys. Chem. B **105**, 6474 (2001).
[53] G. A. Kaminski, H. A. Stern, B. J. Berne, R. A. Friesner, Y. X. X. Cao, R. B. Murphy, R. H. Zhou, and T. A. Halgren, J. Comput. Chem. **23**, 1515 (2002).
[54] P. Y. Ren and J. W. Ponder, J. Phys. Chem. B **107**, 5933 (2003).
[55] G. A. Kaminski, H. A. Stern, B. J. Berne, and R. A. Friesner, J. Phys. Chem. A **108**, 621 (2004).
[56] S. Patel, A. D. Mackerell, and C. L. Brooks, J. Comput. Chem. **25**, 1504 (2004).
[57] V. W. Anisimov, G. Lamoureux, I. V. Vorobyov, N. Huang, B. Roux, and A. D. MacKerell, J. Chem. Theory Comput. **1**, 153 (2005).
[58] N. Gresh, J. P. Piquemal, and M. Krauss, J. Comput. Chem. **26**, 1113 (2005).
[59] E. Harder, B. C. Kim, R. A. Friesner, and B. J. Berne, J. Chem. Theory Comput. **1**, 169 (2005).
[60] Z. X. Wang, W. Zhang, C. Wu, H. X. Lei, P. Cieplak, and Y. Duan, J. Comput. Chem. **27**, 781 (2006).
[61] I. V. Vorobyov, V. M. Anisimov, and A. D. MacKerell, J. Phys. Chem. B **109**, 18988 (2005).
[62] A. Warshel, M. Kato, and A. V. Pisliakov, J. Chem. Theory Comput. **3**, 2034 (2007).

[63] T. D. Rasmussen, P. Y. Ren, J. W. Ponder, and F. Jensen, Int. J. Quant. Chem. **107**, 1390 (2007).
[64] N. Gresh, G. A. Cisneros, T. A. Darden, and J. P. Piquemal, J. Chem. Theory Comput. **3**, 1960 (2007).
[65] M. J. S. Dewar and W. Thiel, J. Am. Chem. Soc. **99**, 4899 (1977).
[66] U. C. Singh and P. A. Kollman, J. Comp. Chem. **7**, 718 (1986).
[67] J. L. Gao, P. Amara, C. Alhambra, and M. J. Field, J. Phys. Chem. A **102**, 4714 (1998).
[68] V. Thery, D. Rinaldi, J. L. Rivail, B. Maigret, and G. G. Ferenczy, J. Comput. Chem. **15**, 269 (1994).
[69] X. Assfeld and J. L. Rivail, Chem. Phys. Lett. **263**, 100 (1996).
[70] G. Monard, M. Loos, V. Thery, K. Baka, and J. L. Rivail, Int. J. Quant. Chem. **58**, 153 (1996).
[71] Hypercube, Inc, *HyperChem*(Hypercube, Inc, Gainesville, Florida, 1985–2007).
[72] R. J. Hall, S. A. Hindle, N. A. Burton, and I. H. Hillier, J. Comp. Chem. **21**, 1433 (2000).
[73] N. Reuter, A. Dejaegere, B. Maigret, and M. Karplus, J. Phys. Chem. A **104**, 1720 (2000).
[74] K. P. Eurenius, D. C. Chatfield, B. R. Brooks, and M. Hodoscek, Int. J. Quant. Chem. **60**, 1189 (1996).
[75] I. Antes and W. Thiel, J. Phys. Chem. A **103**, 9290 (1999).
[76] A. J. Mulholland and W. G. Richards, Prot.: Struct. Funct. Genet. **27**, 9 (1997).
[77] P. D. Lyne, M. Hodoscek, and M. Karplus, J. Phys. Chem. A **103**, 3462 (1999).
[78] A. J. Mulholland, P. D. Lyne, and M. J. Karplus, J. Am. Chem. Soc. **122**, 534 (2000).
[79] L. Ridder, I. M. C. M. Rietjens, J. Vervoort, and A. J. Mulholland, J. Am. Chem. Soc. **124**, 9926 (2002).
[80] D. Das, K. P. Eurenius, E. M. Billings, P. Sherwood, D. C. Chatfield, M. Hodoscek, and B. R. Brooks, J. Chem. Phys. **117**, 10534 (2002).
[81] P. H. Konig, M. Hoffmann, T. Frauenheim, and Q. Cui, J. Phys. Chem. B **109**, 9082 (2005).
[82] Y. K. Zhang, T. S. Lee, and W. T. Yang, J. Chem. Phys. **110**, 46 (1999).
[83] M. Garcia-Viloca and J. L. Gao, Theor. Chem. Acc. **111**, 280 (2004).
[84] J. Z. Pu, J. L. Gao, and D. G. Truhlar, J. Phys. Chem. A **108**, 5454 (2004).
[85] J. Z. Pu, J. L. Gao, and D. G. Truhlar, ChemPhysChem **6**, 1853 (2005).
[86] J. Z. Pu, J. L. Gao, and D. G. Truhlar, J. Phys. Chem. A **108**, 632 (2004).
[87] A. Rodriguez, C. Oliva, M. Gonzalez, M. van der Kamp, and A. J. Mulholland, J. Phys. Chem. B **111**, 12909 (2007).
[88] R. B. Murphy, D. M. Philipp, and R. A. Friesner, Chem. Phys. Lett. **321**, 113 (2000).
[89] U. C. Singh and P. A. Kollman, QCPE Bull. **2**, 117 (1982).
[90] P. K. Weiner and P. A. Kollman, J. Comput. Chem. **2**, 287 (1981).
[91] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, J. Comput. Chem. **14**, 1347 (1993).
[92] P. Sherwood, J. Mol. Graphics Modell. **16**, 275 (1998).
[93] R. D. Amos, I. Alberts, L. Andrews, J. S. Colwell, S. M. Handy, N. C. Jayatilaka, D. Knowles, P. J. Kobayashi, R. Koga, N. Laidig, K. E. Maslen, P. E. Murray, C. W. Rice, J. E. Sanz, J. Simandrias, E. D. Stone, A. J. Su, and M.-D., *CADPAC: The Cambridge Analytic Derivatives Package* (Cambridge University Press, Cambridge, 1995).
[94] Q. Cui, M. Elstner, E. Kaxiras, T. Frauenheim, and M. Karplus, J. Phys. Chem. B **105**, 569 (2001).
[95] *QSite, A QM/MM software*. 2001, Schroedinger, Inc: Portland, OR.
[96] *Jaguar 4.1*. 2000, Schrodinger, Inc.: Portland, Oregon.
[97] *ChemShell, a computational chemistry shell,* http://www.chemshell.org/
[98] P. Sherwood, A. H. de Vries, M. F. Guest, G. Schreckenbach, C. R. A. Catlow, S. A. French, A. A. Sokol, S. T. Bromley, W. Thiel, A. J. Turner, S. Billeter, F. Terstegen, S. Thiel, J. Kendrick, S. C. Rogers, J. Casci, M. Watson, F. King, E. Karlsen, M. Sjovoll, A. Fahmi, A. Schafer, and C. Lennartz, J. Mol. Struct. Theochem. **632**, 1 (2003).

[99] J. N. Harvey, Faraday Discuss. **127**, 165 (2004).
[100] H.-J. Werner, P. J. Knowles, R. Lindh, F. R. Manby, M. Schütz, P. Celani, T. Korona, A. Mitrushenkov, G. Rauhut, T. B. Adler, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, E. Goll, C. Hampel, G. Hetzer, T. Hrenar, G. Knizia, C. Köppl, Y. Liu, A. W. Lloyd, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklaß, P. Palmieri, K. Pflüger, R. Pitzer, M. Reiher, U. Schumann, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, M. Wang, and A. Wolf, *Molpro*. 2006, www.molpro.net
[101] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, and J. A. Pople, *Gaussian* 94. 1994, Gaussian, Inc.: Pittsburgh PA.
[102] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, S. Adamo Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, A. G. Baboul, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, J. L. Andres, C. Gonzalez, M. Head-Gordon, E. S. Replogle, and J. A. Pople, *Gaussian 98* (Gaussian, Inc., Pittsburgh PA, 1998).
[103] J. W. Ponder, *TINKER: Software Tools for Molecular Design* (Washington University School of Medicine, Saint Louis, MO, 2003).
[104] S. Bjelic and J. Aqvist, Biochemistry **43**, 14521 (2004).
[105] G. J. Davies, L. Mackenzie, A. Varrot, M. Dauter, A. M. Brzozowski, M. Schülein, and S. G. Withers, Biochemistry **37**, 11707 (1998).
[106] J. Antosiewicz, J. M. Briggs, A. H. Elcock, M. K. Gilson, and J. A. McCammon, J. Comp. Chem. **17**, 1633 (1996).
[107] D. C. Bas, D. M. Rogers, and J. H. Jensen, Prot.: Struct. Funct. Bioinform. **73**, 765 (2008).
[108] H. Li, A. D. Robertson, and J. H. Jensen, Prot.: Struct. Funct. Bioinform. **61**, 704 (2005).
[109] L. Ridder, A. J. Mulholland, I. Rietjens, M. C. M. Vervoort, and J., J. Mol. Graphics Mod. **17**, 163 (1999).
[110] A. J. Mulholland and W. G. Richards, Proteins **27**, 9 (1997).
[111] B. Guillot, J. Mol. Liq. **101**, 219 (2002).
[112] W. L. Jorgensen, J. Am. Chem. Soc. **103**, 335 (1981).
[113] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).
[114] M. W. Mahoney and W. L. Jorgensen, J. Chem. Phys. **112**, 8910 (2000).
[115] W. L. Jorgensen and J. Tirado-Rives, Proc. Natl Acad. Sci. USA **102**, 6665 (2005).
[116] Y. A. Mantz, B. Chen, and G. J. Martyna, J. Phys. Chem. B **110**, 3540 (2006).
[117] E. Neria, S. Fischer, and M. Karplus, J. Chem. Phys. **105**, 1902 (1996).
[118] Shaw, K.E., Woods, C.J., and Mulholland, A.J., J. Phys. Chem. Lett. **1**, 219–223 (2010).
[119] L. Ridder and A. J. Mulholland, Curr. Top. Med. Chem. **3**, 1241 (2003).
[120] S. Fischer, S. Michnick, and M. Karplus, Biochemistry **32**, 13830 (1993).

[121] A. J. Mulholland and M. Karplus, Biochem. Soc. Trans. **24**, 247 (1996).
[122] J. J. P. Stewart, J. Comp. Aid. Mol. Design **4**, 1 (1990).
[123] K. Senthilkumar, J. I. Mujika, K. E. Ranaghan, F. R. Manby, A. J. Mulholland, and J. N. Harvey, J. R. Soc. Interf. **5**, S207 (2008).
[124] A. J. Mulholland, in *Theoretical Biochemistry – Processes and Properties of Biological Systems*, edited by L. A. Eriksson (Amsterdam, Elsevier, 2001), pp. 597–653.
[125] L. Ridder, I. Rietjens, J. Vervoort, and A. J. Mulholland, J. Am. Chem. Soc. **124**, 9926 (2002).
[126] M. J. Field, P. A. Bash, and M. Karplus, J. Comput. Chem. **11**, 700 (1990).
[127] M. Freindorf and J. L. Gao, J. Comput. Chem. **17**, 386 (1996).
[128] U. Pentikäinen, K. E. Shaw, K. Senthilkumar, C. J. Woods, and A. J. Mulholland, J. Chem. Theory Comput. **5**, 396 (2009).
[129] M. W. van der Kamp, F. Perruccio, and A. J. Mulholland, Prot Struct. Funct. Bioinform. **69**, 521 (2007).
[130] D. Riccardi, G. H. Li, and Q. Cui, J. Phys. Chem. B **108**, 6467 (2004).
[131] J. A. McCammon and S. C. Harvey, *Dynamics of Proteins and Nucleic Acids* (Cambridge University Press, Cambridge, 1987).
[132] L. Ridder, A. J. Mulholland, J. Vervoort, and I. M. C. M. Rietjens, J. Am. Chem. Soc. **120**, 7641 (1998).
[133] F. Claeyssens, K. Ranaghan, F. Manby, J. Harvey, and A. Mulholland, Chem. Commun. 5068 (2005).
[134] J. C. Hermann, C. Hensen, L. Ridder, A. J. Mulholland, and H. D. Holtje, J. Am. Chem. Soc. **127**, 4454 (2005).
[135] A. Lodola, M. Mor, J. C. Hermann, G. Tarzia, D. Piomelli, and A. J. Mulholland, Chem. Commun. 4399 (2005).
[136] K. Fukui, Acc. Chem. Res. **14**, 363 (1981).
[137] G. Monard, X. Prat-Resina, A. Gonzalez-Lafont, and J. M. Lluch, Int. J. Quant. Chem. **93**, 229 (2003).
[138] D. G. Truhlar, J. Gao, C. Alhambra, M. Garcia-Viloca, J. Corchado, M. Sanchez, and J. Villa, Acc. Chem. Res. **35**, 341 (2002).
[139] R. B. Murphy, D. M. Philipp, and R. A. Friesner, J. Comput. Chem. **21**, 1442 (2000).
[140] J. Kastner, S. Thiel, H. M. Senn, P. Sherwood, and W. Thiel, J. Chem. Theory Comput. **3**, 1064 (2007).
[141] S. Martí and V. Moliner, J. Chem. Theory Comput. **1**, 1008 (2005).
[142] S. Martí, V. Moliner, I. Tuñón, and I. H. Williams, J. Phys. Chem. B **109**, 3707 (2005).
[143] R. Elber and M. Karplus, Chem. Phys. Lett. **139**, 375 (1987).
[144] R. Czerminski and R. Elber, Int. J. Quant. Chem. Quant. Chem. Symp. **24**, 167 (1990).
[145] G. Mills, H. Jonsson, and G. K. Schenter, Surf. Sci. **324**, 305 (1995).
[146] G. Henkelman and H. Jonsson, J. Chem. Phys. **113**, 9978 (2000).
[147] L. Xie, H. Liu, and W. Yang, J. Chem. Phys. **120**, 8039 (2004).
[148] G. Henkelman, B. P. Uberuaga, and H. Jonsson, J. Chem. Phys. **113**, 9901 (2000).
[149] H. Liu, Z. Lu, G. A. Cisneros, and W. Yang, J. Chem. Phys. **121**, 697 (2004).
[150] P. Y. Ayala and H. B. Schlegel, J. Chem. Phys. **107**, 375 (1997).
[151] O. S. Smart, Chem. Phys. Lett. **222**, 503 (1994).
[152] S. Fischer and M. Karplus, Chem. Phys. Lett. **194**, 252 (1992).
[153] H. A. Carlson and W. L. Jorgensen, J. Am. Chem. Soc. **118**, 8475 (1996).
[154] Y. K. Zhang, H. Y. Liu, and W. T. Yang, J. Chem. Phys. **112**, 3483 (2000).
[155] T. H. Rod and U. Ryde, J. Chem. Theory Comput. **1**, 1240 (2005).
[156] C. R. W. Guimaraes, M. Udier-Blagovic, I. Tubert-Brohman, and W. L. Jorgensen, J. Chem. Theory Comput. **1**, 617 (2005).
[157] E. Rosta, M. Klahn, and A. Warshel, J. Phys. Chem. B **110**, 2934 (2006).

[158] P. Hu and Y. K. Zhang, J. Am. Chem. Soc. **128**, 1272 (2006).
[159] Z. Lu and W. Yang, J. Chem. Phys. **121**, 89 (2004).
[160] H. Hu, Z. Y. Lu, and W. T. Yang, J. Chem. Theory Comput. **3**, 390 (2007).
[161] C. Bartels and M. Karplus, J. Comput. Chem. **18**, 1450 (1997).
[162] G. M. Torrie and J. P. Valleau, J. Comput. Phys. **23**, 187 (1977).
[163] J. Kastner and W. Thiel, J. Chem. Phys. 123 (2005).
[164] A. González-Lafont, T. N. Truong, and D. G. Truhlar, J. Phys. Chem. **95**, 4618 (1991).
[165] P. A. Bash, L. L. Ho, A. D. Mackerell, D. Levine, and P. Hallstrom, Proc. Natl. Acad. Sci. USA **93**, 3698 (1996).
[166] S. Martí, J. Andrés, V. Moliner, E. Silla, I. Tuñón, J. Bertrán, and M. Field, J. Am. Chem. Soc. **123**, 1709 (2001).
[167] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).
[168] P. Carloni, U. Röthlisberger, and M. Parrinello, Acc. Chem. Res. **35**, 455 (2002).
[169] M. Eichinger, P. Tavan, J. Hutter, and M. Parrinello, J. Chem. Phys. **110**, 10452 (1999).
[170] S. Piana, D. Bucher, P. Carloni, and U. Röthlisberger, J. Phys. Chem. B **108**, 11139 (2004).
[171] M. Dal Peraro, L. I. Llarrull, U. Röthlisberger, A. J. Vila, and P. Carloni, J. Am. Chem. Soc. **126**, 12661 (2004).
[172] A. V. Vargiu, P. Ruggerone, A. Magistrato, and P. Carloni, J. Phys. Chem. B **110**, 24687 (2006).
[173] M. Dal Peraro, A. J. Vila, P. Carloni, and M. L. Klein, J. Am. Chem. Soc. **129**, 2808 (2007).
[174] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).
[175] A. Crespo, M. A. Marti, D. A. Estrin, and A. E. Roitberg, J. Am. Chem. Soc. **127**, 6940 (2005).
[176] M. A. Mart, M. C. G. Lebrero, A. E. Roitberg, and D. A. Estrin, J. Am. Chem. Soc. **130**, 1611 (2008).
[177] C. Dellago, P. G. Bolhuis, F. S. Csajka, and D. Chandler, J. Chem. Phys. **108**, 1964 (1998).
[178] G. E. Crooks and D. Chandler, Phys. Rev. E **64**, 026109 (2001).
[179] P. G. Bolhuis, D. Chandler, C. Dellago, and P. L. Geissler, Annu. Rev. Phys. Chem. **53**, 291 (2002).
[180] J. E. Basner and S. D. Schwartz, J. Am. Chem. Soc. **127**, 13822 (2005).
[181] S. L. Quaytman and S. D. Schwartz, J. Phys. Chem. A **113**, 1892 (2009).
[182] S. Saen-oon, V. L. Schramm, and S. D. Schwartz, Z. Phys. Chem. Int. J. Res. Phys. Chem. Chem. Phys. **222**, 1359 (2008).
[183] M. Iannuzzi, A. Laio, and M. Parrinello, Phys. Rev. Lett. **90**, 238302 (2003).
[184] A. Laio, A. Rodriguez-Fortea, F. L. Gervasio, M. Ceccarelli, and M. Parrinello, J. Phys. Chem. B **109**, 6714 (2005).
[185] B. Ensing, M. De Vivo, Z. W. Liu, P. Moore, and M. L. Klein, Acc. Chem. Res. **39**, 73 (2006).
[186] D. Branduardi, F. L. Gervasio, and M. Parrinello, J. Chem. Phys. **126**, 054103 (2007).
[187] C. L. Stanton, I. F. W. Kuo, C. J. Mundy, T. Laino, and K. N. Houk, J. Phys. Chem. B **111**, 12573 (2007).
[188] L. Petersen, A. Ardevol, C. Rovira, and P. J. Reilly, J. Phys. Chem. B **113**, 7331 (2009).
[189] A. Fersht, *Structure and Mechanism in Protein Science. A Guide to Enzyme Catalysis and Protein Folding*, 2nd ed. (Freeman, New York, 1999).
[190] S. Martí, V. Moliner, I. Tuñón, and I. H. Williams, Org. Biomol. Chem. **1**, 483 (2003).
[191] D. C. Phillips, Proc. Natl Acad. Sci. USA **57**, 484 (1967).
[192] A. J. Kirby, Nat. Struct. Biol. **8**, 737 (2001).
[193] D. J. Vocadlo, G. J. Davies, R. Laine, and S. G. Withers, Nature **412**, 835 (2001).
[194] D. E. Koshland, Biol. Rev. Cambr. Philos. Soc. **28**, 416 (1953).
[195] A. L. Bowman, I. M. Grant, and A. J. Mulholland, Chem. Commun. 4425 (2008).
[196] C. T. Evans, L. C. Kurz, S. J. Remington, and P. A. Srere, Biochemistry **35**, 10661 (1996).
[197] J. A. Gerlt and P. G. Gassman, J. Am. Chem. Soc. **115**, 11552 (1993).
[198] J. A. Gerlt, M. M. Kreevoy, W. W. Cleland, and P. A. Frey, Chem. Biol. **4**, 259 (1997).

[199] G. Alagona, C. Ghio, and P. A. Kollman, J. Am. Chem. Soc. **117**, 9855 (1995).
[200] W. W. Cleland and M. M. Kreevoy, Science **264**, 1887 (1994).
[201] A. Warshel and A. Papazyan, Proc. Natl Acad. Sci. USA **93**, 13665 (1996).
[202] M. W. van der Kamp, F. Perruccio, and A. J. Mulholland, J. Mol. Graphics Modell. **26**, 676 (2007).
[203] M. W. van der Kamp, F. Perruccio, and A. J. Mulholland, Chem. Commun. 1874 (2008).
[204] G. M. Alter, J. P. Casazza, Z. Wang, P. Nemeth, P. A. Srere, and C. T. Evans, Biochemistry **29**, 7557 (1990).
[205] M. W. van der Kamp and A. J. Mulholland, Nat. Prod. Rep. **25**, 1001 (2008).
[206] P. Kast, M. Asif-Ullah, and D. Hilvert, Tetrahedron Lett. **37**, 2691 (1996).
[207] P. D. Lyne, A. J. Mulholland, and W. G. Richards, J. Am. Chem. Soc. **117**, 11345 (1995).
[208] S. Martí, J. Andrés, V. Moliner, E. Silla, I. Tuñón, and J. Bertrán, Theor. Chem. Acc. **105**, 207 (2001).
[209] S. Martí, J. Andrés, V. Moliner, E. Silla, I. Tuñón, and J. Bertrán, Chem. Eur. J. **9**, 984 (2003).
[210] K. E. Ranaghan, L. Ridder, B. Szefczyk, W. A. Sokalski, J. C. Hermann, and A. J. Mulholland, Mol. Phys. **101**, 2695 (2003).
[211] B. Szefczyk, A. J. Mulholland, K. E. Ranaghan, and W. A. Sokalski, J. Am. Chem. Soc. **126**, 16148 (2004).
[212] K. E. Ranaghan, L. Ridder, B. Szefczyk, W. A. Sokalski, J. C. Hermann, and A. J. Mulholland, Org. Biomol. Chem. **2**, 968 (2004).
[213] S. Martí, J. Andrés, V. Moliner, E. Silla, I. Tuñón, and J. Bertrán, J. Phys. Chem. B **104**, 11308 (2000).
[214] H. Guo, Q. Cui, W. N. Lipscomb, and M. Karplus, P. Nat. Acad. Sci. **98**, 9032 (2001).
[215] K. E. Ranaghan, L. Ridder, B. Szefczyk, W. A. Sokalski, J. C. Hermann, and A. J. Mulholland, Org. Biomol. Chem. **2**, 968 (2004).
[216] M. Štrajbl, A. Shurki, M. Kato, and A. Warshel, J. Am. Chem. Soc. **125**, 10228 (2003).
[217] T. C. Bruice, Acc. Chem. Res. **35**, 139 (2002).
[218] S. Hur and T. C. Bruice, J. Am. Chem. Soc. **125**, 5964 (2003).
[219] S. Hur and T. C. Bruice, J. Am. Chem. Soc. **125**, 1472 (2003).
[220] S. Hur and T. C. Bruice, Proc. Natl Acad. Sci. USA **100**, 12015 (2003).
[221] X. H. Zhang and T. C. Bruice, Proc. Natl Acad. Sci. USA **102**, 18356 (2005).
[222] A. Shurki, M. Štrajbl, J. Villà, and A. Warshel, J. Am. Chem. Soc. **124**, 4097 (2002).
[223] S. T. Cload, D. R. Liu, R. M. Pastor, and P. G. Schultz, J. Am. Chem. Soc. **118**, 1787 (1996).
[224] A. Kienhöfer, P. Kast, and D. Hilvert, J. Am. Chem. Soc. **125**, 3206 (2003).
[225] M. P. Repasky, C. Ruch Werneck Guimaraes, J. Chandrasekhar, J. Tirado-Rives, and W. L. Jorgensen, J. Am. Chem. Soc. **125**, 6663 (2003).
[226] C. Ruch Werneck Guimaraes, M. P. Repasky, J. Chandrasekhar, J. Tirado-Rives, and W. L. Jorgensen, J. Am. Chem. Soc. **125**, 6892 (2003).
[227] A. Crespo, D. A. Scherlis, M. A. Martí, P. Ordejon, A. E. Roitberg, and D. A. Estrin, J. Phys. Chem. B **107**, 13728 (2003).
[228] S. Martí and V. Moliner, J. Chem. Theory Comput. **1**, 1008 (2005).
[229] J. J. Ruiz-Pernia, E. Silla, I. Tunon, and S. Marti, J. Phys. Chem. B **110**, 17663 (2006).
[230] H. L. Woodcock, M. Hodoscek, A. T. B. Gilbert, P. M. W. Gill, H. F. Schaefer, and B. R. Brooks, J. Comput. Chem. **28**, 1485 (2007).
[231] T. Ishida, J. Chem. Phys. **129**, 125105 (2008).
[232] H. L. Woodcock, M. Hodoscek, P. Sherwood, Y. S. Lee, H. F. I. Shaefer, and B. R. Brooks, Theor. Chem. Acc. **109**, 140 (2003).
[233] H. L. Woodcock, M. Hodoscek, and B. R. Brooks, J. Phys. Chem. A **111**, 5720 (2007).
[234] C. J. R. Illingworth, K. E. Parkes, C. R. Snell, S. Martí, V. Moliner, and C. A. Reynolds, Mol. Phys. **106**, 1511 (2008).
[235] R. Crehuet and M. J. Field, J. Phys. Chem. B **111**, 5708 (2007).

[236] A. Jimenez, P. Clapes, and R. Crehuet, J. Mol. Model. **14**, 735 (2008).
[237] J. M. Soler, E. Artacho, J. D. Gale, A. Garcia, J. Junquera, P. Ordejon, and D. Sanchez-Portal, J. Phys. Conden. Matter **14**, 2745 (2002).
[238] J. Wang, P. Cieplak, and P. A. Kollman, J. Comput. Chem. **21**, 1049 (2000).
[239] M. Gamper, D. Hilvert, and P. Kast, Biochemistry **39**, 14087 (2000).
[240] Truhlar, D. G., in *Isotope Effects in Chemistry and Biology*, H. A. K. Limbach, ed. 2005, Marcel Dekker, Inc.: New York.
[241] J. Z. Pu, J. Gao, and D. Truhlar, Chem. Rev. **106**, 3140 (2006).
[242] Y. P. Liu, G. C. Lynch, T. N. Truong, D. H. Lu, D. G. Truhlar, and B. C. Garrett, J. Am. Chem. Soc. **115**, 2408 (1993).
[243] C. Alhambra, J. Corchado, M. L. Sanchez, M. Garcia-Viloca, J. Gao, and D. G. Truhlar, J. Phys. Chem. B **105**, 11326 (2001).
[244] G. Tresadern, J. P. McNamara, M. Mohr, H. Wang, N. A. Burton, and I. H. Hillier, Chem. Phys. Lett. **358**, 489 (2002).
[245] M. Garcia-Viloca, D. G. Truhlar, and J. Gao, Biochemistry **42**, 13558 (2003).
[246] J. Z. Pu, S. H. Ma, J. L. Gao, and D. G. Truhlar, J. Phys. Chem. B **109**, 8551 (2005).
[247] M. Garcia-Viloca, C. Alhambra, D. G. Truhlar, and J. L. Gao, J. Comput. Chem. **24**, 177 (2003).
[248] Q. Cui and M. Karplus, Abstr. Pap. Am. Chem. Soc. **220**, 501 (2000).
[249] I. Tejero, M. Garcia-Viloca, A. Gonzalez-Lafont, J. M. Lluch, and D. M. York, J. Phys. Chem. B **110**, 24708 (2006).
[250] C. Alhambra, M. L. Sanchez, J. Corchado, J. Gao, and D. G. Truhlar, Chem. Phys. Lett., **347**, 512; erratum: **355**, 388, 2002 (2001).
[251] P. F. Faulder, G. Tresadern, K. K. Chohan, N. S. Scrutton, M. J. Sutcliffe, I. H. Hillier, and N. A. Burton, J. Am. Chem. Soc. **123**, 8604 (2001).
[252] G. Tresadern, S. Nunez, P. F. Faulder, H. Wang, I. H. Hillier, and N. A. Burton, Faraday Discuss. **122**, 223 (2002).
[253] G. Tresadern, H. Wang, P. F. Faulder, N. A. Burton, and I. H. Hillier, Mol. Phys. **101**, 2775 (2003).
[254] K. E. Ranaghan, L. Masgrau, N. S. Scrutton, M. J. Sutcliffe, and A. J. Mulholland, ChemPhysChem **8**, 1816 (2007).
[255] L. Masgrau, K. E. Ranaghan, N. S. Scrutton, A. J. Mulholland, and M. J. Sutcliffe, J. Phys. Chem. B **111**, 3032 (2007).
[256] V. L. Davidson (ed.) in *Principles and Applications of Quinoproteins* (Marcel Dekker, New York, 1993) pp. 73–85.
[257] J. Basran, M. J. Sutcliffe, and N. S. Scrutton, Biochemistry **38**, 3218 (1999).
[258] G. Tresadern, S. Nunez, P. F. Faulder, H. Wang, I. H. Hillier, and N. A. Burton, Faraday Discuss. **122**, 223 (2003).
[259] J. C. Hermann, L. Ridder, A. J. Mulholland, and H. D. Holtje, J. Am. Chem. Soc. **125**, 9590 (2003).
[260] J. C. Hermann, L. Ridder, H. D. Hotje, and A. J. Mulholland, Org. Biomol. Chem. **4**, 206 (2006).
[261] S. O. Meroueh, J. F. Fisher, H. B. Schlegel, and S. Mobashery, J. Am. Chem. Soc. **127**, 15397 (2005).
[262] J. C. Hermann, J. Pradon, J. N. Harvey, and A. J. Mulholland, J. Phys. Chem. A **113**, 11984 (2009).
[263] B. F. Gherman, S. D. Goldberg, V. W. Cornish, and R. A. Friesner, J. Am. Chem. Soc. **126**, 7652 (2004).
[264] D. Xu, H. Guo, and G. Cui, J. Am. Chem. Soc. **129**, 10814 (2007).
[265] H. Park, E. N. Brothers, and K. M. Merz, J. Am. Chem. Soc. **127**, 4232 (2005).
[266] C. Wang and H. Guo, J. Phys. Chem. B **111**, 9986 (2007).

[267] I. Tubert-Brohman, O. Acevedo, and W. L. Jorgensen, J. Am. Chem. Soc. **128**, 16904 (2006).
[268] A. Lodola, M. Mor, J. Zurek, G. Tarzia, D. Piomelli, J. N. Harvey, and A. J. Mulholland, Biophys. J. **92**, L20 (2007).
[269] D. Piomelli, G. Tarzia, A. Duranti, A. Tontini, M. Mor, T. R. Compton, O. Dasse, E. P. Monaghan, J. A. Parrott, and D. Putman, CNS Drug Rev. **12**, 21 (2006).
[270] A. Lodola, M. Mor, S. Rivara, C. Christov, G. Tarzia, D. Piomelli, and A. J. Mulholland, Chem. Commun. 214 (2008).
[271] J. C. Schoneboom and W. Thiel, J. Phys. Chem. B **108**, 7468 (2004).
[272] J. C. Schoneboom, S. Cohen, H. Lin, S. Shaik, and W. Thiel, J. Am. Chem. Soc. **126**, 4017 (2004).
[273] H. Lin, J. C. Schoneboom, S. Cohen, S. Shaik, and W. Thiel, J. Phys. Chem. B **108**, 10083 (2004).
[274] S. Shaik, D. Kumar, S. P. de Visser, A. Altun, and W. Thiel, Chem. Rev. **105**, 2279 (2005).
[275] A. Altun, S. Shaik, and W. Thiel, J. Comput. Chem. **27**, 1324 (2006).
[276] J. Zurek, N. Foloppe, J. N. Harvey, and A. J. Mulholland, Org. Biomol. Chem. **4**, 3931 (2006).
[277] A. Altun, V. Guallar, R. A. Friesner, S. Shaik, and W. Thiel, J. Am. Chem. Soc. **128**, 3924 (2006).
[278] J. J. Zheng, A. Altun, and W. Thiel, J. Comput. Chem. **28**, 2147 (2007).
[279] A. Altun, S. Shaik, and W. Thiel, J. Am. Chem. Soc. **129**, 8978 (2007).
[280] M. Altarsha, T. Benighaus, D. Kumar, and W. Thiel, J. Am. Chem. Soc. **131**, 4755 (2009).
[281] P. A. Williams, J. Cosme, A. Ward, H. C. Angova, D. M. Vinkovic, and H. Jhoti, Nature **424**, 464 (2003).
[282] P. A. Williams, J. Cosme, D. M. Vinkovic, A. Ward, H. C. Angove, P. J. Day, C. Vonrhein, I. J. Tickle, and H. Jhoti, Science **305**, 683 (2004).
[283] E. E. Scott, M. A. White, Y. A. He, E. F. Johnson, C. D. Stout, and J. R. Halpert, J. Biol. Chem. **279**, 27294 (2004).
[284] M. R. Wester, J. K. Yano, G. A. Schoch, C. Yang, K. J. Griffin, C. D. Stout, and E. F. Johnson, J. Biol. Chem. **279**, 35630 (2004).
[285] J. K. Yano, M. R. Wester, G. A. Schoch, K. J. Griffin, C. D. Stout, and E. F. Johnson, J. Biol. Chem. **279**, 38091 (2004).
[286] C. M. Bathelt, J. Zurek, A. J. Mulholland, and J. N. Harvey, J. Am. Chem. Soc. **127**, 12900 (2005).
[287] M. G. Shou, Q. Mei, M. W. Ettore, R. K. Dai, T. A. Baillie, and T. H. Rushmore, Biochem. J. **340**, 845 (1999).
[288] M. Shou, R. Dai, D. Cui, K. R. Korzekwa, T. A. Baillie, and T. H. Rushmore, J. Biol. Chem. **276**, 2256 (2001).
[289] D. Fishelovitch, C. Hazan, H. Hirao, H. J. Wolfson, R. Nussinov, and S. Shaik, J. Phys. Chem. B **111**, 13822 (2007).
[290] C. M. Bathelt, A. J. Mulholland, and J. N. Harvey, J. Phys. Chem. A **112**, 13149 (2008).
[291] Y. Wang, H. Chen, M. Makino, Y. Shiro, S. Nagano, S. Asamizu, H. Onaka, and S. Shaik, J. Am. Chem. Soc. **131**, 6748 (2009).
[292] M. Makino, H. Sugimoto, Y. Shiro, S. Asamizu, H. Onaka, and S. Nagano, Proc. Natl Acad. Sci. USA **104**, 11591 (2007).
[293] T. P. Barrows, B. Bhaskar, and T. L. Poulos, Biochemistry **43**, 8826 (2004).
[294] L. Tian and R. A. Friesner, J. Chem. Theory Comput. **5**, 1421 (2009).
[295] A. L. Bowman, L. Ridder, I. Rietjens, J. Vervoort, and A. J. Mulholland, Biochemistry **46**, 6353 (2007).
[296] D. Dourado, P. A. Fernandes, B. Mannervik, and M. J. Ramos, Chem. Eur. J. **14**, 9591 (2008).
[297] C. G. Zhan, F. Zheng, and D. W. Landry, J. Am. Chem. Soc. **125**, 2462 (2003).
[298] C. G. Zhan and D. Q. Gao, Biophys. J. **89**, 3863 (2005).
[299] D. Q. Gao and C. G. Zhan, Proteins Struct. Funct. Bioinform. **62**, 99 (2006).

[300] A. Wlodawer and J. W. Erickson, Annu. Rev. Biochem. **62**, 543 (1993).
[301] H. Liu, F. Müller-Plathe, and W. F. van Gunsteren, J. Mol. Biol. **261**, 454 (1996).
[302] D. C. Chatfield, K. P. Eurenius, and B. R. Brooks, J. Mol. Struct. Theochem. **423**, 79 (1998).
[303] C. H. Suresh, A. M. Vargheese, K. P. Vijayalakshmi, N. Mohan, and N. Koga, J. Comput. Chem. **29**, 1840 (2008).
[304] S. Saen-oon, O. Aruksakunwong, K. Wittayanarakul, P. Sompornpisut, and S. Hannongbua, J. Mol. Graphics Modell. **26**, 720 (2007).
[305] T. Tuccinardi, Comb. Chem. High Throughput Screen. **12**, 303 (2009).
[306] M. H. J. Seifert, Drug Discov. Today **14**, 562 (2009).
[307] O. Guvench and A. D. MacKerell, Curr. Opin. Struct. Biol. **19**, 56 (2009).
[308] M. P. Gleeson and D. Gleeson, J. Chem. Inform. Model. **49**, 670 (2009).
[309] C. Hensen, J. C. Hermann, K. H. Nam, S. H. Ma, J. L. Gao, and H. D. Holtje, J. Med. Chem. **47**, 6673 (2004).
[310] P. Fong, J. P. McNamara, I. H. Hillier, and R. A. Bryce, J. Chem. Inform. Model. **49**, 913 (2009).
[311] M. D. Eldridge, C. W. Murray, T. R. Auton, G. V. Paolini, and R. P. Mee, J. Comp. Aid. Mol. Design **11**, 425 (1997).
[312] M. L. Verdonk, J. C. Cole, M. J. Hartshorn, C. W. Murray, and R. D. Taylor, Proteins Struct. Funct. Genet. **52**, 609 (2003).
[313] T. Rungrotmongkol, S. Hannongbua, and A. Mulholland, J. Theor. Comput. Chem. **3**, 491 (2004).
[314] T. Rungrotmongkol, A. J. Mulholland, and S. Hannongbua, J. Mol. Graphics Modell. **26**, 1 (2007).
[315] N. Nunthaboot, S. Pianwanit, V. Parasuk, J. O. Ebalunode, J. M. Briggs, and S. Kokpol, Biophys. J. **93**, 3613 (2007).
[316] J. J. Ruiz-Pernia, C. N. Alves, V. Moliner, E. Silla, and I. Tuñón, J. Mol. Struct. Theochem. **898**, 115 (2009).
[317] Y. K. Zhang, J. Kua, and J. A. McCammon, J. Am. Chem. Soc. **124**, 10572 (2002).
[318] M. Fuxreiter and A. Warshel, J. Am. Chem. Soc. **120**, 183 (1998).
[319] Y. H. Cheng, X. L. Cheng, Z. Radic, and J. A. McCammon, J. Am. Chem. Soc. **129**, 6562 (2007).
[320] J. Wang, J. D. Gu, and J. Leszczynski, J. Phys. Chem. B **110**, 7567 (2006).
[321] O. Kwasnieski, L. Verdier, M. Malacria, and E. Derat, J. Phys. Chem. B **113**, 10001 (2009).
[322] T. Rungrotmongkol, P. Decha, P. Sompornpisut, M. Malaisree, P. Intharathep, N. Nunthaboot, T. Udommaneethanakit, O. Aruksakunwong, and S. Hannongbua, Proteins Struct. Funct. Bioinform. **76**, 62 (2009).
[323] V. Nanda, Nature Chem. Biol. **4**, 273 (2008).
[324] L. Jiang, E. A. Althoff, F. R. Clemente, L. Doyle, D. Rothlisberger, A. Zanghellini, J. L. Gallaher, J. L. Betker, F. Tanaka, C. F. Barbas, D. Hilvert, K. N. Houk, B. L. Stoddard, and D. Baker, Science **319**, 1387 (2008).
[325] K. Kikuchi, R. B. Hannak, M. J. Guo, A. J. Kirby, and D. Hilvert, Bioorg. Med. Chem. **14**, 6189 (2006).
[326] R. Muller, E. W. Debler, M. Steinmann, F. P. Seebeck, I. A. Wilson, and D. Hilvert, J. Am. Chem. Soc. **129**, 460 (2007).
[327] S. Martí, J. Andrés, E. Silla, V. Moliner, I. Tuñón, and J. Bertrán, Angew. Chemie Int. Edn. **46**, 286 (2007).
[328] S. Martí, J. Andrés, V. Moliner, E. Silla, I. Tuñón, and J. Bertrán, J. Am. Chem. Soc. **130**, 2894 (2008).
[329] D. Rothlisberger, O. Khersonsky, A. M. Wollacott, L. Jiang, J. DeChancie, J. Betker, J. L. Gallaher, E. A. Althoff, A. Zanghellini, O. Dym, S. Albeck, K. N. Houk, D. S. Tawfik, and D. Baker, Nature **453**, 190 (2008).

[330] A. N. Alexandrova, D. Rothlisberger, D. Baker, and W. L. Jorgensen, J. Am. Chem. Soc. **130**, 15907 (2008).
[331] A. N. Alexandrova and W. L. Jorgensen, J. Phys. Chem. B **113**, 497 (2009).
[332] U. Haupts, J. Tittor, and D. Oesterhelt, Annu. Rev. Biophys. Biomol. Struct. **28**, 367 (1999).
[333] W. Kuhlbrandt, Nature **406**, 569 (2000).
[334] M. A. Van der Horst and K. J. Hellingwerf, Acc. Chem. Res. **37**, 13 (2004).
[335] H. Luecke, B. Schobert, H. T. Richter, J. P. Cartailler, and J. K. Lanyi, J. Mol. Biol. **291**, 899 (1999).
[336] H. Belrhali, P. Nollert, A. Royant, C. Menzel, J. P. Rosenbusch, E. M. Landau, and E. Pebay-Peyroula, Structure **7**, 909 (1999).
[337] K. Edman, P. Nollert, A. Royant, H. Belrhali, E. Pebay-Peyroula, J. Hajdu, R. Neutze, and E. M. Landau, Nature **401**, 822 (1999).
[338] H. Luecke, B. Schobert, H. T. Richter, J. P. Cartailler, and J. K. Lanyi, Science **286**, 255 (1999).
[339] F. Garczarek, L. S. Brown, J. K. Lanyi, and K. Gerwert, Proc. Natl Acad. Sci. USA **102**, 3633 (2005).
[340] F. Garczarek and K. Gerwert, Nature **439**, 109 (2006).
[341] K. Gerwert, G. Souvignier, and B. Hess, Proc. Natl Acad. Sci. USA **87**, 9774 (1990).
[342] M. Hatanaka, H. Kandori, and A. Maeda, Biophys. J. **73**, 1001 (1997).
[343] H. Kandori, Y. Yamazaki, J. Sasaki, R. Needleman, J. K. Lanyi, and A. Maeda, J. Am. Chem. Soc. **117**, 2118 (1995).
[344] H. Kandori, N. Kinoshita, Y. Shichida, and A. Maeda, J. Phys. Chem. B **102**, 7899 (1998).
[345] H. Kandori, M. Belenky, and J. Herzfeld, Biochemistry **41**, 6026 (2002).
[346] V. A. Lorenz-Fonfria and H. Kandori, J. Am. Chem. Soc. **131**, 5891 (2009).
[347] R. Rammelsberg, G. Huhn, M. Lubben, and K. Gerwert, Biochemistry **37**, 5001 (1998).
[348] M. Shibata, T. Tanimoto, and H. Kandori, J. Am. Chem. Soc. **125**, 13312 (2003).
[349] V. Spassov, Z. Luecke, H. Gerwert, K. Bashford, and D., J. Mol. Biol. **312**, 203 (2001).
[350] C. Zscherp, R. Schlesinger, J. Tittor, D. Oesterhelt, and J. Heberle, Proc. Natl Acad. Sci. USA **96**, 5498 (1999).
[351] A. Maeda, J. Sasaki, Y. Shichida, T. Yoshizawa, M. Chang, B. F. Ni, R. Needleman, and J. K. Lanyi, Biochemistry **31**, 4684 (1992).
[352] S. Hayashi and I. Ohmine, J. Phys. Chem. B **104**, 10678 (2000).
[353] Y. Matsui, K. Sakai, M. Murakami, Y. Shiro, S. Adachi, H. Okumura, and T. Kouyama, J. Mol. Biol. **324**, 469 (2002).
[354] B. Schobert, J. Cupp-Vickery, V. Hornak, S. O. Smith, and J. K. Lanyi, J. Mol. Biol. **321**, 715 (2002).
[355] S. Hayashi, E. Tajkhorshid, and K. Schulten, Biophys. J. **83**, 1281 (2002).
[356] S. Hayashi, E. Tajkhorshid, H. Kandori, and K. Schulten, J. Am. Chem. Soc. **126**, 10516 (2004).
[357] T. Tanimoto, Y. Furutani, and H. Kandori, Biochemistry **42**, 2300 (2003).
[358] R. R. Birge, T. M. Cooper, A. F. Lawrence, M. B. Masthay, C. Vasilakis, C. F. Zhang, and R. Zidovetzki, J. Am. Chem. Soc. **111**, 4063 (1989).
[359] R. R. Birge and T. M. Cooper, Biophys. J. **42**, 61 (1983).
[360] M. Baer, G. Mathias, I. F. W. Kuo, D. J. Tobias, C. J. Mundy, and D. Marx, ChemPhysChem **9**, 2703 (2008).
[361] H. Kandori and Y. Shichida, J. Am. Chem. Soc. **122**, 11745 (2000).
[362] L. S. Brown, J. Sasaki, H. Kandori, A. Maeda, R. Needleman, and J. K. Lanyi, J. Biol. Chem. **270**, 27122 (1995).
[363] S. P. Balashov, E. S. Imasheva, T. G. Ebrey, N. Chen, D. R. Menick, and R. K. Crouch, Biochemistry **36**, 8671 (1997).
[364] J. M. Headrick, E. G. Diken, R. S. Walters, N. I. Hammer, R. A. Christie, J. Cui, E. M. Myshakin, M. A. Duncan, M. A. Johnson, and K. D. Jordan, Science **308**, 1765 (2005).

[365] R. Rousseau, V. Kleinschmidt, U. W. Schmitt, and D. Marx, Angew. Chemie Int. Edn. **43**, 4804 (2004).
[366] G. Zundel, Adv. Chem. Phys. **111**, 111, 1 (2000).
[367] G. Mathias and D. Marx, Proc. Natl Acad. Sci. USA **104**, 6980 (2007).
[368] R. Rousseau, V. Kleinschmidt, U. W. Schmitt, and D. Marx, Phys. Chem. Chem. Phys. **6**, 1848 (2004).
[369] P. Phatak, J. S. Frahmcke, M. Wanko, M. Hoffmann, P. Strodel, J. C. Smith, S. Suhai, A. N. Bondar, and M. Elstner, J. Am. Chem. Soc. **131**, 7064 (2009).
[370] T. Kouyama, T. Nishikawa, T. Tokuhisa, and H. Okumura, J. Mol. Biol. **335**, 531 (2004).
[371] K. Edman, A. Royant, G. Larsson, F. Jacobson, T. Taylor, D. van der Spoel, E. M. Landau, E. Pebay-Peyroula, and R. Neutze, J. Biol. Chem. **279**, 2147 (2004).
[372] G. J. Kleywegt and T. A. Jones, Macromol. Crystallography, Pt B **277**, 208 (1997).
[373] A. Wlodawer, W. Minor, Z. Dauter, and M. Jaskolski, FEBS J. **275**, 1 (2008).
[374] U. Ryde, L. Olsen, and K. Nilsson, J. Comput. Chem. **23**, 1058 (2002).
[375] U. Ryde and K. Nilsson, J. Mol. Struct. Theochem. **632**, 259 (2003).
[376] U. Ryde and K. Nilsson, J. Am. Chem. Soc. **125**, 14232 (2003).
[377] U. Ryde, Dalton Trans. **6**, 607 (2007).
[378] A. T. Brunger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J. S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, R. J. Read, L. M. Rice, T. Simonson, and G. L. Warren, Acta Crystallographica Sect. D Biol. Crystallography **54**, 905 (1998).
[379] A. T. Brunger, Nature Protocols **2**, 2728 (2007).
[380] K. Nilsson and U. Ryde, J. Inorg. Biochem. **98**, 1539 (2004).
[381] S. Benini, A. Gonzalez, W. R. Rypniewski, K. S. Wilson, J. J. Van Beeumen, and S. Ciurli, Biochemistry **39**, 13115 (2000).
[382] G. Pettersson, CRC Crit. Rev. Biochem. **21**, 349 (1987).
[383] J. K. Rubach, S. Ramaswamy, and B. V. Plapp, Biochemistry **40**, 12686 (2001).
[384] N. Yu, S. A. Hayik, B. Wang, N. Liao, C. H. Reynolds, and K. M. Merz, J. Chem. Theory Comput. **2**, 1057 (2006).
[385] T.-S. Lee, D. M. York, and W. Yang, J. Chem. Phys. **105**, 2744 (1996).
[386] A. van der Vaart, D. Suarez, and K. M. Merz, J. Chem. Phys. **113**, 10512 (2000).
[387] S. L. Dixon and K. M. Merz, J. Chem. Phys. **104**, 6643 (1996).
[388] S. L. Dixon and K. M. Merz, J. Chem. Phys. **107**, 879 (1997).
[389] A. D. Andricopulo, L. B. Salum, and D. J. Abraham, Curr. Top. Med. Chem. **9**, 771 (2009).
[390] A. M. Davis, S. A. St-Gallay, and G. J. Kleywegt, Drug Discov. Today **13**, 831 (2008).
[391] D. E. Tronrud, L. F. Teneyck, and B. W. Matthews, Acta Crystallographica Sect. A **43**, 489 (1987).
[392] A. Jack and M. Levitt, Acta Crystallographica Sect. A **34**, 931 (1978).
[393] J. W. M. Nissink, C. Murray, M. Hartshorn, M. L. Verdonk, J. C. Cole, and R. Taylor, Prot. Struct. Funct. Genet. **49**, 457 (2002).
[394] X. Li, X. He, B. Wang, and K. Merz, J. Am. Chem. Soc. **131**, 7742 (2009).
[395] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. Shindyalov, N. Bourne, and P. E., Nucleic Acids Res. **28**, 235 (2000).
[396] F. C. Bernstein, T. F. Koetzle, G. J. B. Williams, E. F. Meyer, M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, and M. Tasumi, J. Mol. Biol. **112**, 535 (1977).
[397] F. H. Allen, Acta Crystallograph. Sect. B Struct. Sci. **58**, 380 (2002).
[398] F. H. Allen and W. D. S. Motherwell, Acta Crystallograph. Sect. B Struct. Sci. **58**, 407 (2002).
[399] D. S. Murray, M. A. Schumacher, and R. G. Brennan, J. Biol. Chem. **279**, 14365 (2004).
[400] J. Fokkens and G. Klebe, Angew. Chemie Int. Edn. **45**, 985 (2006).
[401] J. Cavanagh, W. J. Fairbrother, A. G. Palmer III, N. J. Skelton, Protein NMR Spectroscopy: Principles and Practice, (Academic Press, London, 1996).
[402] Y. W. Hsiao, T. Drakenberg, and U. Ryde, J. Biomol. NMR **31**, 97 (2005).

[403] J. J. R. Frausto Da Silva and R. J. P. Williams. The Biological Chemistry of the Elements: The Inorganic Chemistry of Life, Oxford University Press: Oxford, England, UK; New York, New York, USA. Illus, xxi+561pp (1991).
[404] Z. Rao, P. Handford, M. Mayhew, V. Knott, G. G. Brownlee, and D. Stuart, Cell **82**, 131 (1995).
[405] U. Ryde, Biophys. J. **77**, 2777 (1999).
[406] L. Ridder, J. Harvey, I. Rietjens, J. Vervoort, and A. J. Mulholland, J. Phys. Chem. B **107**, 2118 (2003).
[407] R. A. Mata, H. J. Werner, S. Thiel, and W. Thiel, J. Chem. Phys. **128**, 025104 (2008).
[408] L. Ridder, B. A. Palfey, J. Vervoort, and I. Rietjens, FEBS Lett. **478**, 197 (2000).
[409] W. Z. Lai, H. Chen, and S. Shaik, J. Phys. Chem. B **113**, 7912 (2009).
[410] Z. Pipirou, V. Guallar, J. Basran, C. L. Metcalfe, E. J. Murphy, A. R. Bottrill, S. C. Mistry, and E. L. Raven, Biochemistry **48**, 3593 (2009).
[411] A. Rubinstein and D. T. Major, J. Am. Chem. Soc. **131**, 8513 (2009).
[412] E. Puig, E. Mixcoha, M. Garcia-Viloca, A. Gonzalez-Lafont, and J. M. Lluch, J. Am. Chem. Soc. **131**, 3509 (2009).
[413] K. Nam, Q. Cui, J. L. Gao, and D. M. York, J. Chem. Theory Comput. **3**, 486 (2007).
[414] Y. Yang, H. B. Yu, D. York, M. Elstner, and Q. Cui, J. Chem. Theory Comput. **4**, 2067 (2008).
[415] Y. Yang, H. B. Yu, and Q. Cui, J. Mol. Biol. **381**, 1407 (2008).
[416] A. J. T. Smith, Y. Li, and K. N. Houk, Org. Biomol. Chem. **7**, 2716 (2009).
[417] A. Gustavo Turjanski, G. Hummer, and J. S. Gutkind, J. Am. Chem. Soc. **131**, 6141 (2009).
[418] S. Metz, D. Q. Wang, and W. Thiel, J. Am. Chem. Soc. **131**, 4628 (2009).
[419] M. H. Ho, M. De Vivo, M. Dal Peraro, and M. L. Klein, J. Chem. Theory Comput. **5**, 1657 (2009).
[420] F. Wallrapp, D. Masone, and V. Guallar, J. Phys. Chem. A **112**, 12989 (2008).
[421] A. J. Mulholland and W. G. Richards, J. Mol. Struct. (Theochem.) **427**, 175 (1998).
[422] L. Ridder, A. J. Mulholland, I. Rietjens, and J. Vervoort, J. Mol. Graphics Modell. **17**, 163 (1999).
[423] H. M. Senn, S. Thiel, and W. Thiel, J. Chem. Theory Comput. **1**, 494 (2005).
[424] A. Lodola, M. Mor, J. Sirirak, and A. J. Mulholland, Biochem. Soc. Trans. **37**, 363 (2009).
[425] M. W. van der Kamp, K. E. Shaw, C. J. Woods, and A. J. Mulholland, J. R. Soc. Interface **5**, S173 (2008).