TUTORIAL REVIEW

CSR
www.rsc.org/csr

Theoretical insights in enzyme catalysis

Sergio Martí,[a] Maite Roca,[a] Juan Andrés,[a] Vicent Moliner,*[a] Estanislao Silla,[b] Iñaki Tuñón*[b] and Juan Bertrán[c]

[a] *Departament de Ciències Experimentals, Universitat Jaume I, Box 224, Castellón, Spain. E-mail: moliner@exp.uji.es; Fax: +34964728066; Tel: +34964728084*

[b] *Departament de Química Física/IcMol, Universidad de Valencia, 46100 Burjasot, Spain. E-mail: Ignacio.Tunon@uv.es; Fax: +34963544564; Tel: +34963544880*

[c] *Departament de Química, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain*

Received 2nd July 2003
First published as an Advance Article on the web 9th December 2003

In this *tutorial review* we show how the methods and techniques of computational chemistry have been applied to the understanding of the physical basis of the rate enhancement of chemical reactions by enzymes. This is to answer the question: Why is the activation free energy in enzyme catalysed reactions smaller than the activation free energy observed in solution? Two important points of view are presented: Transition State (TS) theories and Michaelis Complex (MC) theories. After reviewing some of the most popular computational methods employed, we analyse two particular enzymatic reactions: the conversion of chorismate to prephenate catalysed by *Bacillus subtilis* chorismate mutase, and a methyl transfer from *S*-adenosylmethionine to catecholate catalysed by catechol *O*-methyltransferase. The results and conclusions obtained by different authors on these two systems, supporting either TS stabilisation or substrate preorganization, are presented and compared. Finally we try to give a unified view, where a preorganized enzyme active site, prepared to stabilise the TS, also favours those reactive conformations geometrically closer to the TS.

1 Introduction

Enzymes are biological catalysts that allow organisms to carry out biological reactions with time scales compatible with life. For example, an ancient ship can persist for centuries in the bottom of the sea due to the fact that the hydrolysis of glycosidic bonds of cellulose would require several million years to reach its half-time in the absence of the appropriate catalyst.[1] Enzymes are able to speed up chemical reactions by an order of magnitude of 10^6 to 10^{20}, which represent an amazing enhancement of chemical kinetics with respect to the counterpart reaction in solution. These catalysts are not only very efficient but they are also specific. The origin of these features of enzymes is the question to be answered in this review. In order to reach this goal, molecular simulations in solution and in the enzyme are presented and analysed.

The simplest kinetic scheme used to understand enzymatic processes is that proposed by Michaelis and Menten, which proceeds with the formation of a substrate–catalyst complex prior to the reaction step during which the catalyst is recovered:

$$\mathrm{E + S \leftrightarrow ES \rightarrow P + E} \qquad (1)$$

Fig. 1, together with the uncatalysed process in solution. The formation of the enzyme–substrate complex (the Michaelis complex, MC) has a free energy of binding ($\Delta G^{\mathrm{MC}}_{\mathrm{Bind}}$) related to the apparent dissociation constant ($K_{\mathrm{m}} = RT \ln \Delta G^{\mathrm{MC}}_{\mathrm{Bind}}$). Many enzymes have evolved so efficiently that the rate constant for the second step (k_{cat}) approaches the diffusion limit. Moreover, in some cases this chemical step is even faster, the rate limiting step then being diffusion of reactants (or products). If this is not the case, and following Transition State Theory, this rate constant can be formally related to an activation free energy ($k_{\mathrm{cat}} = \frac{k_{\mathrm{B}}T}{h}\exp\left(\Delta G^{\ddagger}_{\mathrm{cat}}\right)$). At low substrate concentration regime, the initial reaction rate, v_0, can be expressed as:

$$v_0 = \frac{k_{\mathrm{cat}}}{K_{\mathrm{m}}}\left[E\right]_0\left[S\right]_0 \qquad (2)$$

At high substrate concentration regime, the initial reaction rate, v_0, becomes:

$$v_0 = k_{\mathrm{cat}}\left[E\right]_0 \qquad (3)$$

From eqn. (2) the global rate constant is given by $k_{\mathrm{cat}}/k_{\mathrm{m}}$ and the global activation free energy is then $\Delta G^{\ddagger}_{E} = \Delta G^{\mathrm{MC}}_{\mathrm{Bind}} + \Delta G^{\ddagger}_{\mathrm{cat}}$. From eqn. (3) the observed activation free energy is $\Delta G^{\ddagger}_{\mathrm{cat}}$.

The counterpart reaction in solution can usually be related to a unimodal-shaped free energy profile. This starts with the separated, fully solvated, reactants which transform into products with a single rate constant k_{uncat}, associated to the activation free energy of the uncatalysed reaction in solution $\Delta G^{\ddagger}_{\mathrm{uncat}}$. Rate enhancement is usually defined as the ratio $k_{\mathrm{cat}}/k_{\mathrm{uncat}}$,[1] this is comparing $\Delta G^{\ddagger}_{\mathrm{cat}}$ and $\Delta G^{\ddagger}_{\mathrm{uncat}}$. In this way, the analysis avoids the substrate concentration dependency.

In this paper we will show how the methods and techniques provided by Computational Chemistry can be applied to understand the physical basis of the rate enhancement of the chemical reactions by the enzymes. This is to answer the question: Why is the activation free energy in enzyme-catalysed reactions smaller than the activation free energy observed in solution?

2 Theories

According to the thermodynamic quantities presented in Fig. 1, the reduction of the activation free energy can be expressed in terms of the binding energy of both the reactant and transition states as deduced from Scheme 1:

$$\Delta G^{\ddagger}_{\mathrm{uncat}} - \Delta G^{\ddagger}_{\mathrm{cat}} = \Delta G^{\mathrm{MC}}_{\mathrm{Bind}} - \Delta G^{\mathrm{TS}}_{\mathrm{Bind}} \qquad (4)$$

If the left-hand side of eqn. (4) is positive, the $\Delta G^{\mathrm{TS}}_{\mathrm{Bind}}$ magnitude has to be larger, in absolute value, than $\Delta G^{\mathrm{MC}}_{\mathrm{Bind}}$. This is, the enzyme presents a larger affinity for the transition state than for the reactant state.

In a coarse-grained sense there are two groups of theories that can be used to understand the ability of enzymes to speed up

DOI: 10.1039/b301875j

This journal is © The Royal Society of Chemistry 2004

Sergio Marti obtained his European PhD (2002) degree in chemistry from Universitat Jaume I, under the supervision of Vicent Moliner and Iñaki Tuñón. During this time he spent several months in the research groups of Martin J. Field (Grenoble, France), Manolo Ruiz-López (Nancy, France) and Ian H. Williams (Bath, United Kingdom). Nowadays he focuses his attention on computational and applied aspects of statistical hybrid quantum mechanical/molecular mechanical methods.

Maite Roca obtained her BSc (2000) from Universitat Jaume I and her MSc (2002), under the supervision of Vicent Moliner and Iñaki Tuñón. She studied PhD courses on theoretical and computational chemistry at the University of Valencia (2001) and spent two months with Ian H. Williams at the University of Bath (2002). Her research interests are in the use of hybrid quantum mechanical/molecular mechanical (QM/MM) and molecular dynamics methods for the study of reaction mechanisms in active enzyme sites.

Juan Andrés obtained his BSc (1978) and his PhD (1982) from the University of Valencia, under the supervision of Estanislao Silla. He carried out postdoctoral research with O. Tapia (University of Uppsala, 1983–1997). He was invited Professor at the Pierre and Marie Curie University of Paris (1998) and the University of Sao Carlos, Brazil (1998, 2003). Since 1994 he has been Full Professor of Physical Chemistry at Universitat Jaume I. His research interests are focused on theoretical studies of chemical bonds, solid state chemistry, structure and chemical reactivity, in both gas and condensed (solvent and enzyme) phases.

Vicent Moliner obtained his BSc (1988) from the University of Valencia and his PhD (1993) from Universitat Jaume I, under the supervision of Juan Andrés. He carried out postdoctoral research with Ian H. Williams (University of Bath, 1996–1997), and also with O. Tapia (Uppsala University) and J. Krechl (Prague University). He was appointed as Assistant Professor at Universitat Jaume I in 1994 and has been Lecturer (Titular de Universidad) in Physical Chemistry since 1998. His research interest is mainly focused on theoretical studies of chemical reactivity in enzyme catalysis.

Estanislao Silla obtained his PhD (1974) from Universitat Autónoma de Madrid working with Professors Fernández-Alonso and Bertrán. After a postdoctoral stage at LCQEM (Pisa, Italy) with Professors Scrocco and Tomasi he moved to CUC (Castellón, Spain). He was appointed as Lecturer (Titular de Universidad) at the University of Valencia in 1981 and has been Full Professor of Physical Chemistry there since 1992. His research interests involve the application of theoretical methodologies to the study of organic and biochemical processes: reaction mechanisms, solvent effects and enzyme catalysis.

Iñaki Tuñón obtained his PhD (1993) from the University of Valencia working under the supervision of Professor Silla on continuum models. He conducted postdoctoral research in 1994 at the University of Nancy with Professor Rivail. During this time he worked on DFT/MM models. After one year he took up a postdoctoral position at the University of Valencia and has been a Lecturer (Titular de Universidad) in Physical Chemistry there since 1998. His research is focused on theoretical studies of chemical reactivity in enzymes and in solution.

Juan Bertrán obtained his BSc (1961) and his PhD (1969) from the University of Barcelona. His PhD was under the supervision of Raymond Daudel, working at the "Centre de Mécanique Ondulatoire Appliquée" in Paris. He has been Professor of Physical Chemistry at different Spanish Universities. He has also been invited Professor at the Henry Poincaré University of Nancy and at Pisa University. At present, he is Emeritus Professor at the "Universitat Autònoma de Barcelona". His research interest is focused on theoretical studies of chemical reactivity, particularly in relation to reaction mechanisms, solvent effects and enzyme catalysis.

Sergio Marti

Maite Roca

Juan Andrés

Vicent Moliner

Estanislao Silla

Iñaki Tuñón

Juan Bertrán

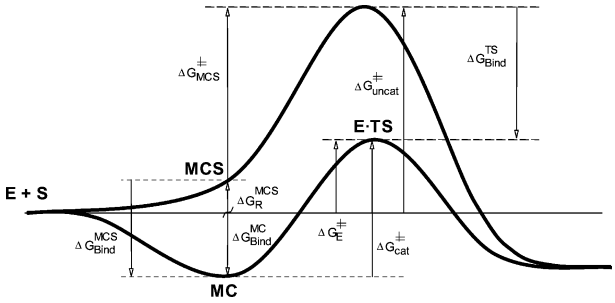


Fig. 1 Schematic free energy diagram of the enzyme catalysis compared to the reaction in solution.

$$\begin{array}{ccc} E + S & \xrightarrow{\Delta G^{\ddagger}_{uncat}} & E + S^{\ddagger} \\ \Delta G^{MC}_{Bind} \downarrow & & \downarrow \Delta G^{TS}_{Bind} \\ ES^{MC} & \xrightarrow{\Delta G^{\ddagger}_{cat}} & ES^{\ddagger} \end{array}$$

Scheme 1

chemical reactions. A first interpretation focuses on the larger binding of the enzyme with the TS, which appears as a consequence of a better interaction of the enzyme with the TS relative to the MC. This idea was first formulated by Pauling[2] assuming the complementarity between the enzyme's active site and the transition structure. Warshel and co-workers[3–5] have explored different enzymatic mechanisms showing that the transition state stabilisation is basically due to the electrostatic environment provided by the active site of the enzyme. According to these authors, the active site displays an electric field prepared to accommodate the charge distribution of the transition state. This provides a strong stabilisation of the transition state relative to the reactants without changing the enzymatic environment too much. In contrast, in aqueous solution, water molecules adapt to the reaction charge flow, which has an associated energetic price. Consequently, the catalytic effect is attributed to the preorganization of the enzyme. These energetic contributions are shown in Fig. 2.

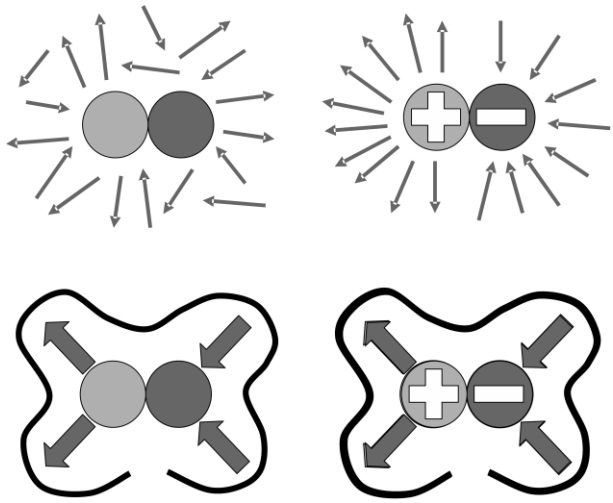


Fig. 2 Contribution of the reorganization energy to the transition state stabilisation. The figure displays a particular ideal reaction process where apolar reactants are transformed in a strongly polar transition state. In aqueous solution, water dipoles must be reoriented during the reaction, spending a part of the interaction energy. In continuum solvent models this energy price amounts to one-half of the solute–solvent interaction energy. In the enzyme the electrostatic environment is already prepared to accommodate the transition state charge distribution.

While the previous authors have emphasized the TS stabilisation as the main role of the enzyme catalysis, others focus their attention in the formation of the Michaelis complex (ES^{MC} in Scheme 1), *i.e.* the left hand side of the energy profile of Fig. 1. This second line of thinking is based on the fact that the binding energies can include non-negligible contributions due to the geometrical changes of the substrate. In the case of the transition state, the most important geometrical parameters are assumed to remain unchanged when passing from the solution to the enzyme. Thus, ΔG^{TS}_{Bind} essentially contains the variation of the interaction energy when passing from water to the active site. However, ΔG^{MC}_{Bind} may contain an important energy contribution due to the change in the substrate geometry as far as one usually goes from separated and fully solvated reacting groups to a spatial rearrangement where these are in close contact and in a proper orientation to react. The comparison between the free energy profiles for the reaction in the enzyme active site and in solution can be helped if we consider an imaginary intermediate state in solution that corresponds with a similar structure to the MC in the enzyme, although it was not a real free energy minimum. This MC-like structure in solution (MCS) would have the same value of the reaction coordinate as the MC and then it would present a reactive conformation ready to progress up to the transition state. In this way, the binding energy of the reactants (ΔG^{MC}_{Bind}) can be split into two terms, the free energy required in going from the solvent-separated reactants to the MCS (ΔG^{MCS}_{R}) and the binding energy of this structure (ΔG^{MCS}_{Bind})

$$\Delta G^{MC}_{Bind} = \Delta G^{MCS}_{R} + \Delta G^{MCS}_{Bind} \quad (5)$$

Now, the binding energies of the TS and the MCS can be more directly compared as far as the main contribution in both is the variation in the interaction energies. Substituting eqn. (5) into (4):

$$\Delta G^{\ddagger}_{uncat} - \Delta G^{\ddagger}_{cat} = \Delta G^{MCS}_{R} + \Delta G^{MCS}_{Bind} - \Delta G^{TS}_{Bind} \quad (6)$$

Since ΔG^{MCS}_{R} is a positive value, a reduction in the activation free energy ($\Delta G^{\ddagger}_{uncat} - \Delta G^{\ddagger}_{cat} > 0$) can be obtained even if the enzyme presents a better affinity for the Michaelis complex (ΔG^{MCS}_{Bind}) than for the transition state (ΔG^{TS}_{Bind}), provided that the most important contribution due to the geometrical changes of the reactants (ΔG^{MCS}_{R}) has been withdrawn.

Thus, following this model, the key of the catalysis would be focused on the preorganization of the substrate that takes place favourably in the enzyme active site and has an energetic cost in solution. This basic idea has been expressed in different ways depending on the components chosen to follow the process. Page and Jencks[6] emphasize the concept of the entropic trap. The MC formation implies the loss of translational and rotational degrees of freedom rendering a loss of entropy that is, *a priori*, always a non-favourable term. This implies that in solution the ΔG^{MCS}_{R} term is essentially of entropic nature. In the enzyme active site this entropy loss is compensated by favourable interactions, resulting in a negative binding energy. This explanation has been named as the entropic trap. Kollman *et al.*[7] centred the attention on the free energy needed to preorganize the reactants in solution, ΔG^{MCS}_{R}, which is not only due to entropic terms but also to enthalpy contributions. The latter term mainly comes from the new inter and intramolecular interactions appearing when the reacting groups are approached. This MCS free energy of formation in solution is called as "cratic" energy by these authors. Koshland *et al.*[8] stressed on the orbital steering, arguing that an increased overlap between the molecular orbitals involved in the chemical process plays a major role in enzyme catalysis. Geometrical factors have been also explored to understand the preorganization of the substrate induced by the enzyme. Bruice *et al.*[9] introduced the Near Attack Conformation (NAC) concept as those ground state conformers that closely resemble the TS. Bruice considers NACs as turnstiles through which the ground state must pass to enter the TS. According to this idea the enzyme active site would decrease the activation free energy by increasing the probability of finding NAC-like structures. Menger *et al.*[10] in their spatial-temporal hypothesis have previously used closely related arguments.

In conclusion, although other effects, based for instance on dynamic factors, have been also evoked to explain enzyme

catalysis, (for references to such proposals see refs. 3 and 4) we can classify the most popular theories about enzyme catalysis into two groups: *TS-theories* and *MC-theories*, this is, transition state stabilisation and substrate preorganization, respectively. A note of caution here about terminology: substrate preorganization is not to be confused with enzyme preorganization, mentioned previously.

3 Methodology

Potential energy function

Historically, first insights into enzymatic reaction mechanisms were obtained through gas phase calculations, using standard programs of quantum chemistry, assuming the invariance of the transition structures. However, the properties of stationary structures of enzymatic processes can be different from those obtained in the gas phase and, obviously, the interaction energy with the environment is not considered in these calculations. A better understanding was obtained by including a small part of the active centre into the calculations. The problem is that in this case the optimised structures did not always fit into the enzyme active site. An approximate solution was obtained anchoring some key atoms of the enzyme to their crystallographic positions and optimising the rest of the coordinates of the model. The computational cost of these calculations rapidly increases as more atoms of the environment are explicitly included. Linear-scaling quantum methods can be used to partially avoid this bottleneck.

In the aforementioned strategies the enzyme flexibility and long-range effects on the nuclear and electronic polarisation of the chemical system by the environment are not easily incorporated. Shortly after those first gas phase calculations, there was a major breakthrough that included the enzyme in the calculations.[11] Although this approach was ignored for long time, now it is the most popular computational methodology for studying enzymatic reactions and is based on a combination of quantum mechanics and molecular mechanics. In these methods the quantum mechanical (QM) description is reserved to a small portion of the system, the region where the most important chemical changes (*i.e.* bond breaking and forming processes) take place. The rest of the system can be described by means of molecular mechanics (MM) potentials. In this way a very large number of atoms can be explicitly considered in the calculations. The division of the full system into two sub-systems may require cutting a covalent bond. Different techniques[12] have been developed to treat this problem, fulfilling the valence of the quantum atom placed in the boundary with hydrogen atoms or frozen orbitals (see Fig. 3). The

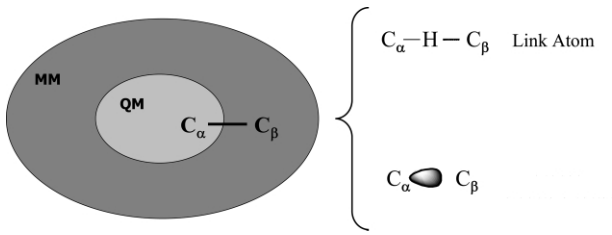


Fig. 3 In QM/MM methods the subsystem is divided into a QM part and a MM part. The QM subsystem must include all those atoms directly involved in the bond breaking and forming process. This partition may require cutting a covalent bond (between a QM atom α and a MM atom β) in the boundary and thus special techniques should be applied to fulfil the valence of the QM atom (α).

combination of these two levels, by means of the appropriated coupling terms, is generically known as hybrid QM/MM methods.[11,13] In this way the wave function of the quantum subsystem, and thus any related property, can be obtained under the influence of the environment.

Exploration of the potential energy surface

The existence of this QM/MM potential energy function, which combines reliability and computational efficiency, is not the only requirement to deal with chemical reactions. To describe such a process we should be able to locate and characterize the set of stationary structures (reactants, products, transition state and possible intermediates) that defines a particular reaction mechanism. This is complicated by the large dimensionality of the Potential Energy Surface (PES) when the environment is explicitly included. Most of the programs using hybrid calculations contain algorithms to locate energy minima. Several approximate strategies have been proposed to locate these structures. In the simplest one only the positions of the QM atoms are varied while the MM part remains frozen at some particular positions obtained from X-ray data or previous optimisations. A more accurate way of optimising the relevant structures is to allow also the relaxation of the environment coupled with the chemical system. However, this simulation is computationally taxing.

In a first approximation to follow a chemical process a distinguished geometrical coordinate can be chosen and energy minimisations carried out for different values of this coordinate. Obviously, this procedure can not be always convenient in that the true reaction coordinate can be different from the selected one. The direct location of transition structures is even more difficult and is not included in any commercial programs. The high dimensionality of the surface prevents the calculation of the full Hessian matrix and thus the proper location and characterization of saddle points of index one. Recently some algorithms have been proposed to calculate approximate Hessians for chemical processes in very large systems.[14] For example, a reduced Hessian matrix can be defined containing only the relevant coordinates for the chemical process under study. The difference is that the Hessian matrix now guides the geometrical search and thus second derivative based optimisation methods can be employed to explore the selected coordinates while the rest is minimised at each step of this search. The procedure is schematically shown in Fig. 4. The stationary

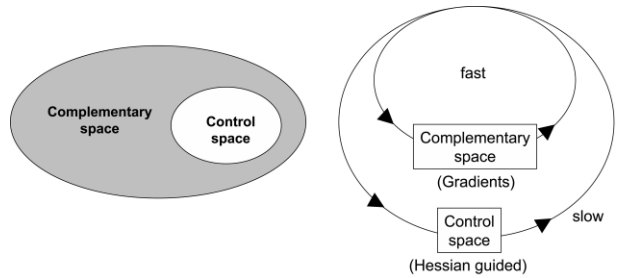


Fig. 4 Stationary structures in very large systems can be located and characterized using a partition of the coordinates space into the control space (all the geometrical variables with a significant contribution to the reaction coordinate) and the complementary space (the rest of the coordinates). A reduced Hessian matrix is calculated for the control space and used to guide the search of the stationary structure. At each step of the search the complementary space is minimised using only gradients. The obtained structure is thus a stationary point on the PES (all gradients are zero) and it contains the correct number of imaginary frequencies) in the control space.

structure obtained in this way can be characterized as a true stationary structure (all the first derivatives are equal to zero) and having the correct number of imaginary frequencies in the approximate Hessian matrix (*i.e.* zero for minima and only one for transition state). From this structure the transition vector and reaction paths can be defined and used to get relevant knowledge about the reaction mechanism.

Statistical simulations

One substantial difference between gas phase and condensed media reactions is that in the former, the reactant, the product and the transition state usually correspond to well-defined single structures. Thus thermodynamic properties can be obtained applying standard formulae to the different energy levels of these structures. In solution or enzymatic environments exploration of the PES is not enough to get magnitudes directly comparable to experiment. The

PES contains a myriad of stationary structures mainly due to the great number of possible conformations accessible to the enzyme and the solvent molecules. Thus, a statistical ensemble of minima and transition structures must be explored to properly define the reactant, product and transition states. Simulations can be carried out using Monte Carlo or Molecular Dynamics methods obtaining detailed information from which averaged and thermodynamic properties can be derived.[15] In particular, free energies associated with the transformation from the reactant state to the transition state (the activation free energy) and to the product state (the reaction free energy) can be extracted using different techniques applied to molecular simulations.[16]

One of the most popular techniques to get reaction free energy profiles for enzymatic reactions is the quantum mechanics-free energy perturbation (QM-FEP) method developed by Kollman and co-workers.[7] In this approach the reaction path is obtained for a gas phase model of the active site. Coordinates and charges of the chemical system are afterwards used in purely classical simulations where only the changes in the environment are sampled. Free Energy Perturbation (FEP) is then used to get the reaction and activation free energies as the sum of the gas phase reaction energy and environment free energy contributions. The main advantage of this approach is that high-level quantum methods can be used for the gas phase calculations. The main drawbacks are: i) the lack of the chemical system flexibility contribution to the free energy and ii) during the simulation of the reaction path, the environment is not incorporated in the calculation. A similar but more coupled scheme is obtained if the reaction path is determined in the active site by means of QM/MM iterative optimisations along a distinguished coordinate.[17]

Free energies profiles can be also obtained as a Potential of Mean Force (PMF) appearing along a particular reaction coordinate.[15] Selection of this reaction coordinate should be based on the exploration of the PES including the environment, or even better, on IRCs traced down to the corresponding products and reactants valleys from transition structures located and characterized in the enzyme active site. The umbrella sampling method is used to place the chemical system at different values of the reaction coordinate that cannot be sampled frequently enough by thermal fluctuations. This is done by adding an adequate parabolic energy function centred at the value of the reaction coordinate that is to be explored. Simulations are then carried out sampling all the degrees of freedom of the system except for the reaction coordinate (see Fig. 5). Once the reaction coordinate has been fully explored from

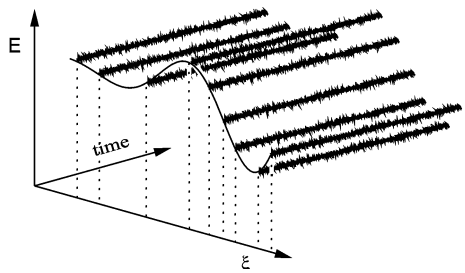
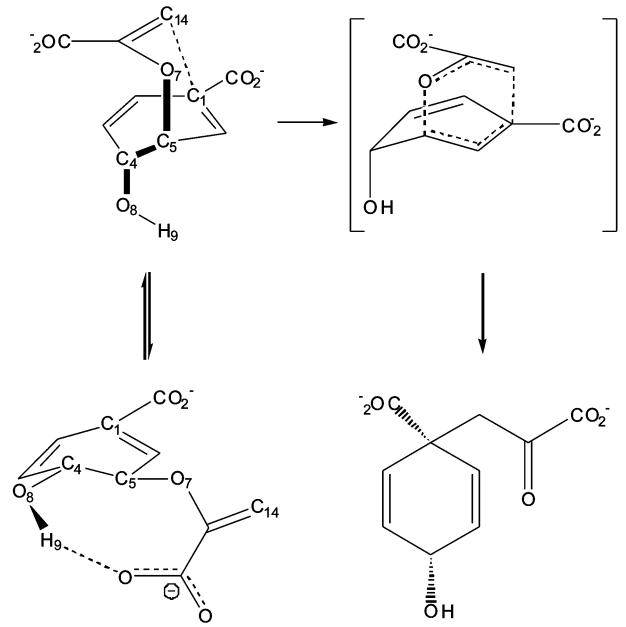


Fig. 5 The PMF is obtained by means of a series of molecular dynamic simulations where all the degrees of freedom and a specified reaction coordinate are sampled. Umbrella sampling is used to place the system at an adequate value of the reaction coordinate (ζ_0) and then Monte Carlo or Molecular Dynamics simulations are run. The fluctuations of the reaction coordinate are finally pieced together obtaining the full distribution function and thus the free energy profile.

reactants to products, the total probability distribution function is obtained and thus the free energy profile is calculated. Furthermore, averaged properties can be derived from the dynamics obtained on the maximum and the minima regions of the PMF profile, characterizing in this way the transition and reactant and product states, respectively. The main advantage of this technique is then the inclusion of all contributions to the free energy, but this require a very large number of energy evaluations. These calculation are expensive and nowadays restricted to semi-empirical Hamiltonians[13] or empirical valence bond (EVB)[3,4] methods. This latter can be fitted to gas phase *ab initio* surfaces and charges and then the information transferred in a consistent coupled way to FEP calculations in solution or in enzymatic media.

4 Chorismate mutase: an example of a unimolecular enzymatic reaction

After presenting the different computational methodologies used to simulate enzymatic processes, we have selected an example of an enzyme reaction that has been used by most of the previously cited authors to explain the origin of the enzyme catalysis: the chorismate to prephenate rearrangement catalysed by chorismate mutase. This system has been so popular probably due to the fact that it presents several advantages: i) The rearrangement of chorismate to prephenate catalysed by the enzyme has its counterpart in solution, and experimental[18–21] and theoretical,[9,22–38] studies have demonstrated that the reaction takes place following the same molecular mechanism. This is a very important feature as it allows direct comparison of the results obtained in both media and gives an insight into the role of the enzyme. We have to keep in mind that it is quite frequent that catalysts accelerate the chemical rate by changing the mechanism and in such a case the comparisons would not give information of the generic aspects of enzyme catalysis. ii) There are data available in the literature[18,21] that offer the opportunity to compare theoretical and experimental results. Furthermore, although some debate appeared in the literature as to which step was rate limiting in chorismate mutase, more recent studies based on kinetic isotope effects have demonstrated that the chemical reaction is preponderantly rate limiting in this enzyme.[19] iii) No covalent bonds are formed between the substrate and the protein, avoiding technical problems of frontier treatments between QM and MM regions, as explained previously in the methodology section. iv) Since the rearrangement of chorismate to prephenate is a unimolecular reaction, the first step of the energy profile depicted in Fig. 1, the contribution of bringing two separated reactant species together to form the MC in a bimolecular process, is simplified to a conformational problem: the work of changing a non-reactive chorismate conformer structure into a new one which is ready to proceed the rearrangement to prephenate (see Scheme 2).



Scheme 2 (Adapted in part with permission from ref. 26. Copyright 2003 Wiley-VCH.)

The conversion of (−)-chorismate to prephenate constitutes a key step in the shikimate pathway for biosynthesis of the phenylalanine and tyrosine aromatic amino acids in bacteria, fungi

and higher plants. This reaction is formally a Claisen rearrangement, and a rare example of an enzyme-catalysed pericyclic process that takes place through a diaxial-chairlike transition state. As depicted in Scheme 2, a conformational equilibrium prior to the chemical reaction between reactive and non-reactive conformers can take place. We will label as reactive reactants those chorismate conformers closer to the TS conformation: they present a pseudo-diaxial character and a significant close distance between the two carbon atoms to be bonded in the rearrangement (C1–C14). The non-reactive reactants present a conformation with a long C1–C14 distance and/or a pseudo-equatorial conformation; far from the TS. Menger has emphasized the importance of the C1–C14 distance.[36] Classification of the possible chorismate conformers is shown in Fig. 6.

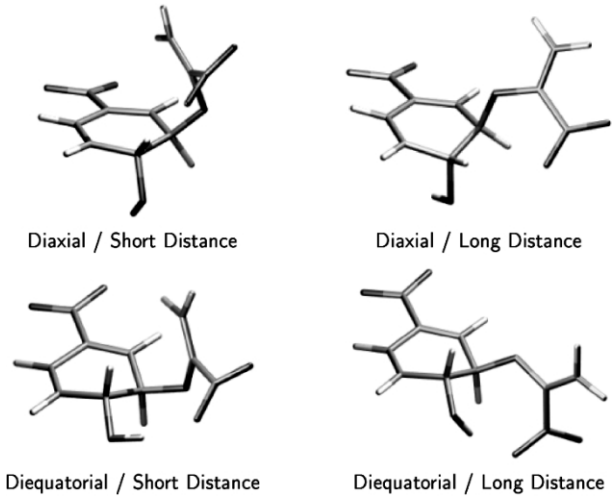


Fig. 6 Possible conformations available to chorismate. The classification criterium is based on the C1–C14 distance (short and long distance conformers) and the relative position of the ring substituents (diaxial and diequatorial). (Adapted in part with permission from ref. 26. Copyright 2003 Wiley-VCH).

MC theories

It is obvious that in the gas phase the pseudo-diequatorial chorismate conformers are the most stable due to the presence of an intramolecular hydrogen bond. In solution, the formation of solute–solvent hydrogen bond interactions can dramatically change this behaviour. Direct NMR studies of this conformation equilibrium in solution carried out by Knowles and co-workers indicated that the pseudo-diequatorial forms were only 0.9–1.4 kcal mol^{-1} lower in energy than the pseudo-diaxial ones.[21] In this sense, while preliminary results of Jorgensen[22] based on Monte Carlo free energy perturbation simulations reproduced this experimental data, very recently these authors realized that the diaxial character of the chorismate conformer is not enough to classify the reactants into the reactive reactants group.[23] As mentioned above, the distance between the two atoms that are bonded during the chemical reaction (C1–C14) has to be considered to classify the species as reactive reactants.

The pioneering QM/MM work of Richards on *Bacillus subtilis* chorismate mutase (BsCM)[24] demonstrated that the inclusion of the enzyme environment in the calculations facilitates the stabilisation of a minimum energy structure which presents a distorted geometry compared to the ground state structure of chorismate in the gas phase. This chorismate–enzyme complex is displaced towards the structure of the TS. In this way, these results provided an example of the concept of the preorganization of the substrate by the enzyme. A direct QM/MM potential energy minimisation of six different starting point chorismate–enzyme complexes carried out by Martí *et al.*[25] clearly render the diaxial and short C1–C14 distance conformer as the most stable structure. The analysis of the different energy terms showed that this conformer was the most stable one due to the lowest deformation of the enzyme. This study points out that the compression effect of the enzyme preorganizing the substrate is caused by the inertia of the enzyme structure to be deformed. In a subsequent paper[26] PMF profiles were obtained for the diequatorial–diaxial and for the short–long C1–C14 distance equilibria, based on QM/MM dynamics in aqueous and protein environments. The analysis of this comparative study showed that while only pseudo-diaxial and short distance chorismate conformers were stable in the enzyme active site, the aqueous solvated solute calculations render an almost degenerated equilibrium between short and long C1–C14 distance with the pseudo-diequatorial conformer slightly more stable than the diaxial one. Thus the enzyme favours the population of the reactive reactants, if compared with the solvent behaviour. A similar result was obtained by Karplus *et al.*[27] using a chorismate mutase enzyme of another organism (*yeast* CM). According to them, the results obtained from their QM/MM molecular dynamics simulations, starting from reactive and non-reactive chorismate conformations, suggest that one contribution of the enzyme is to bind the more prevalent non-reactive conformer and transform it into an active form as a first step, before the following chorismate to prephenate rearrangement.

In an attempt to quantify, to some extent, the importance of the substrate preorganization energy term in the overall free energy barrier, Bruice *et al.*[28] have evaluated it from Boltzman distribution analysis derived from the probabilities of population of NACs coming from classical molecular dynamics calculations. The NAC was defined in this work as those ground state conformations showing a C1–C14 distance $\leq$ 3.7 Å and an attack angle $\leq$ 30 degrees. This definition, based in an arbitrary criterion, is useful to distinguish between non-reactive and reactive chorismate conformers and, as a consequence, to evaluate the substrate pre-organization. The mole fraction of NAC that they obtained was 0.00007% in water and 34% in the enzyme. These results correspond to standard free energies for NAC formation of 8.4 and 0.6 kcal mol^{-1} in water and in the *E. coli* CM enzyme, respectively. Considering that the experimental value of the difference in the activation free energy between aqueous and enzyme is around. 9.1 kcal mol^{-1}, the authors concluded that the origin of the enzyme catalysis is related to this different energy cost required to form the NAC in solution and in the enzyme active site. Nevertheless, it is well known that it is very hard to obtain the proper probability of being at a high-energy region by a direct MD simulation.[35] As a matter of fact, a much lower estimate of free energies for NAC formation was obtained by FEP and PMF calculations of Warshel,[35] Martí[26] and Jorgensen.[23]

Menger *et al.*[36] have carried out a study to quantify the compression of the substrate by the enzyme following a different strategy than the one used by Bruice. Investigating experimentally the chorismate to prephenate rearrangement on model compounds and by quantum chemistry calculations at density functional theory level, a direct relationship between activation barrier lowering and the distance between the reactive termini (C1–C14) was deduced. Assuming the geometrical results of Lyne *et al.*[24] that showed that this carbon–carbon distance in reactants goes from *ca.* 3.30 Å in gas phase to 2.85 Å in the enzyme active site, Menger deduced that this reduction of *ca.* 0.45 Å would imply a reduction of around 10 kcal mol^{-1} in the activation barrier. This value, based on gas phase calculations, is overestimated if compared to the conclusions obtained in the work of Lyne *et al.* including the effect of the environment.[24] Lyne *et al.* proposed that TS stabilisation is also important. In contrast, Hillier *et al.*,[29] using QM/MM optimisations of the ground state and the TS in BsCM suggested that the compression of the reactants does not contribute to the decreasing of the barrier. They suggested that the increase in the enzyme–substrate electrostatic interactions when going from reactants to TS was the real origin of the rate enhancement for the reaction step.

TS theories

After having presented those theories based on the substrate preorganization as the main responsible factor of the enzyme catalysis, following the same CM enzymatic system as the conductor wire, we are presenting now those studies that emphasized the enzyme role on stabilising the TS.

Warshel and co-workers,[30] based on the idea that the largest catalytic effect of enzymes is related to the electrostatic complementarity of the active site, obtained the most favourable dipole environment to stabilise the charge distribution of the TS for the chorismate to prephenate rearrangement. From this simulated enzyme electric field, they could improve the binding of a TS analogue. A similar strategy of searching an electrostatic complementarity between the TS and the environment has been carried out by Kangas and Tidor[31] in order to predict the optimal TS analogue (TSA) for this reaction. Using an endo-oxabicyclic TSA as the starting geometry, they tried to improve the electrostatic affinity for the enzyme using a charge optimisation method. Their calculations predict that by replacing the carboxylate group of the ring by a neutral nitro group the binding energy would be improved. This approach has been recently supported by Hilvert *et al*.[32] that experimentally confirmed the prediction.

A test of this hypothesis of preferential stabilisation of the TS for this particular chemical reaction has been carried out by Martí *et al*.[33] Using the antisymmetric combination of the forming and breaking bond distances as the distinguished reaction coordinate, the PMF profiles obtained in aqueous solution and in the enzyme, as described in previous section, are plotted in Fig. 7. From the

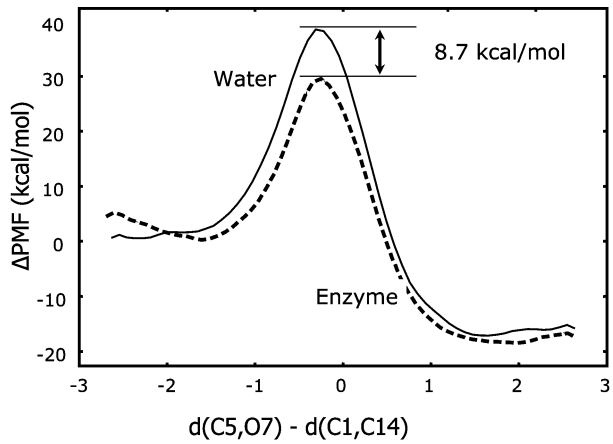


Fig. 7 Chorismate to prephenate PMF profiles obtained in aqueous solution and in the BsCM obtained from the data of ref 33.

analysis of this figure it is possible to conclude that the barriers appear overestimated, basically due to the use of a semi-empirical Hamiltonian to describe the QM region of the system. In an attempt to correct this source of error a correction of the description of the QM region has been carried at DFT level, obtaining results in good agreement with the experimental data. Nevertheless, the difference between the solvent and enzyme theoretical free energy barriers obtained at AM1/MM level is in very good agreement with the experimental data (8.7 *vs.* 9.1 kcal mol^{-1}, respectively). As mentioned above, this enzymatic system presents, as one of the advantages, the fact that no covalent bonds exist between the QM and the MM regions. This feature will allow the carrying out of a very interesting decomposition analysis of the potential energy barrier in aqueous solution and in the enzymatic environment. The total QM/MM potential energy is decomposed into three different contributions:[26]

$$E_T = E_{solute/substrate} + E_{int} + E_{env} \quad (7)$$

where the first term ($E_{solute/substrate}$) is the energy of the substrate or the solute, the second term (E_{int}) is the interaction energy between the two sub-systems and the last term (E_{env}) is the energy of the MM environment (water or enzyme). According to this decomposition, the energy barrier of the reaction is the sum of three contributions, which are given in Table 1 for the reaction in water and in BsCM.

Table 1 Averaged values for the free energy and potential energy barrier and its components (see eqn. (7)) for the chorismate rearrangement in water solution and in BsCM (from ref. 33). All values in kcal/mol

	Water	BsCM
$\Delta G^{\ddagger}$	38.0	29.3
$\Delta E^{\ddagger}$	39.0	27.1
$\Delta E_{solute/substrate}$	40.4	42.1
ΔE_{int}	−2.7	−16.2
ΔE_{env}	1.4	1.2

It must be kept in mind that the reported values in Table 1 come from the average of the structures appearing in QM/MM MD generated at the maximum (TS) and the minimum (Reactants) of the free energy profiles in Fig. 7. The MD simulations have been calculated until convergence of the differences in energies. Nevertheless, as previously mentioned, these results have to be treated with caution and the conclusions have to be considered as qualitative, because they can be affected by large statistical errors, especially the change in the MM environment energy. As can be seen in the table, from the comparison between the free energy barriers and the potential energy barriers, the low influence of entropic factors can be deduced. The main contribution to the potential energy barrier lowering does not come from the solute or the substrate energy, which in fact is in the opposite trend, but from the preferential interaction of the enzyme with the TS. It must be emphasized that the nature of this contribution is essentially electrostatic. It can be surprising that while the interaction contribution is much more important (in the absolute sense) in the enzyme than in water, the energetic change of the environment is very similar in both media. In water solution this last term is very close to the one half value of the interaction energy as predicted by linear response solvent models while in the enzyme this represents a much lower relative contribution. Inside the enzyme, there is a large electrostatic effect associated with a very small reorganization. This analysis is in agreement with the preferential stabilisation of the TS by electrostatic interactions in the enzyme rather than in water as previously pointed out by Warshel *et al*.[3,4] More recently, they have evaluated the binding free energy of the ground state and the transition state in CM, demonstrating that the enzyme works by transition state stabilisation (TSS).[35] The evaluation, using linear response methods, of the different contributions to the reduction of the activation energy established that TSS resulted from electrostatic effects.

It is interesting to compare these later results with the ones published recently by Jorgensen *et al*.[38] using gas phase optimised structures along the reaction path. The energy values derived from these optimisations were added to the classical Monte Carlo simulations of the protein in order to obtain the final free energy profile of the process (QM-FEP described in the methodology section). Although their free energy profiles are very close to the ones obtained by Martí *et al*.,[33] the conclusions are different. Their thermodynamic analysis would suggest that preferential TS stabilisation in the enzyme environment relative to water plays a secondary role in the rate enhancement, the main contribution being the gas phase conformational compression of reactants. Nevertheless, as has been emphasized,[36] the energy of deforming the substrate in the gas phase cannot be used in analysing enzyme catalysis, unless one considers the energy of solvating the substrate as well, in aqueous solution and in the enzyme environment. Finally, the central role of TS stabilisation in this enzyme has been emphasized by Mulholland *et al*.[37] According to these authors, the active site is exquisitely complementary to the TS, stabilising it more than the substrate, so reducing the barrier to reaction.

Combining MC theories with TS theories: an integrated view

In Fig. 8 a detail of the substrate and the amino acids that form the active site of the enzyme is depicted for the averaged TS structures,

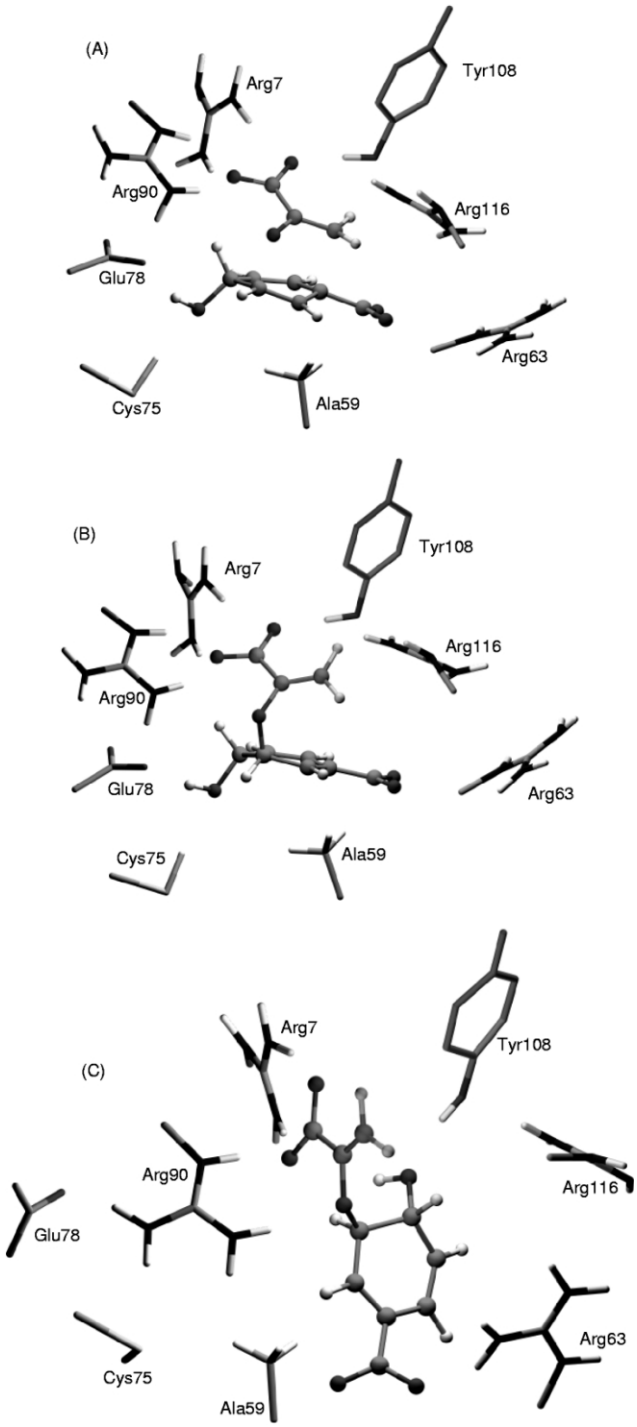


Fig. 8 Detail of the averaged structures of a) TS, b) reactive reactants and c) non-reactive reactants of the chorismate to prefenate rearrangement in the active site of the BsCM. (Adapted in part with permission from ref. 26. Copyright 2003 Wiley-VCH.)

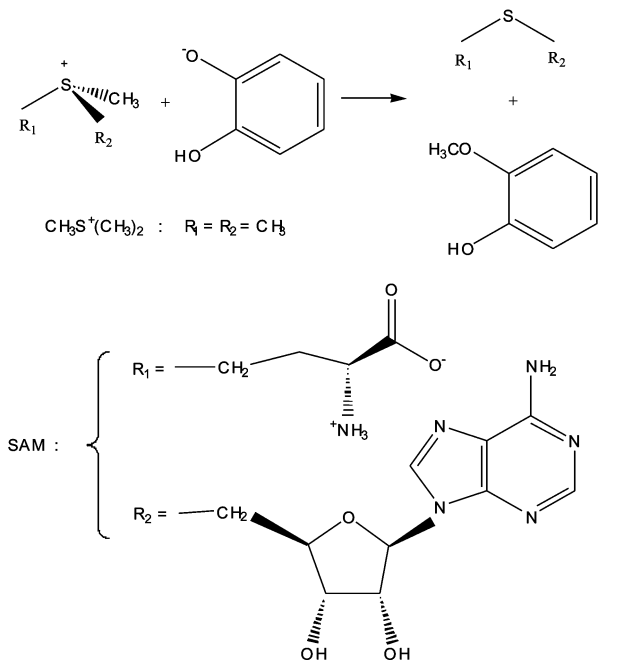
reactive reactants and one of the non-reactive reactants (8a, 8b and 8c, respectively).[26] It is clear that while the patterns of interactions in the TS and the reactive reactants are very similar; this is not the case when the non-reactive reactant complex is formed. This implies that while the enzyme structure remains essentially unaltered during the reaction progress, it must be considerably deformed to accommodate reactants conformations very different from the TS structure. As the enzyme deformation has an associated energy penalty, this suggests that an enzyme site prepared to accommodate the TS should also favour those reactants more similar to the TS structure. The fact that the enzyme does not require to be deformed in going from reactants to transition state has been also observed in the studies of Worthinton *et al.*[34] generated by *ab initio*/MM calculations, even when the quantum subsystem is enlarged by including some of the residues in the surroundings of the active site. The importance of Glu78 residue, the Arg90 that activates the ether bond and stabilises the TS, or the role of Arg7, Tyr108 and Arg115 that present direct ionic interactions to the substrate being catalytically significant, in addition to their obvious role in binding, has been emphasized by these authors.

Arriving at this point, MC theories and TS theories that appeared to be opposite lines of thinking perhaps are not so different. If the ideas of Pauling and Warshel are accepted and we also take into account that the enzyme deformation requires an energy penalty, then, and this is the important point here, it is clear that this same enzyme structure has a considerable effect on the reactants. In the global energy balance the equilibrium among reactants' substrate conformers is displaced towards those reactive conformations geometrically closer to the TS, thus avoiding the energetic penalty associated with the deformation of the full enzyme–substrate system. That is, substrate preorganization and the preferential stabilisation of the TS by favourable electrostatic interactions with the protein have a common origin in the enzyme structure, which is in turn preorganized to favour the reaction progress. Both effects are two faces of the same coin. Consequently, the final conclusion for this system is that the origin of the catalysis has to be searched for in the enzyme structure. Another integrated vision of TS and MC theories has been proposed recently by Warshel *et al.*[35] According to these authors, the apparent NAC effect is not the reason for the catalytic effect but the result of the TS stabilisation; the key catalytic effect is electrostatic in nature. However, since the charge distribution of the TS and the reactive reactants is similar, the stabilisation of TS leads to reduction in the distance between the reacting atoms in the reactant state.

5 Catechol *O*-methyltransferase: an example of a bimolecular enzymatic reaction

In order to check the previous conclusions derived from a unimolecular reaction, we now present a study of a bimolecular reaction, namely methyl transfer from *S*-adenosylmethionine (SAM) to the hydroxylate oxygen of a substituted catechol catalysed by catechol *O*-methyltransferase (COMT). COMT is important in the central nervous system where it metabolises dopamine, adrenaline, noradrenaline and various xenobiotic catechols. This bimolecular S_N2 process can be formally considered as an inverse Menchutkin reaction where ionic reactants proceed towards neutral products (Scheme 3). The enzymatic process also requires the presence of a magnesium cation (Mg^{2+}) in the active site. As in the preceding example, no covalent bonds are formed between the substrate, the cofactor and the enzyme. The counterpart reaction in solution is the transfer of a methyl group from trimethylsulfonium ion to cathecolate, see also Scheme 3, a reaction which proceeds with the same reaction mechanism and for which experimental data are also available.

Using the QM-FEP approach, Kollman *et al.*[39] traced the free energy profile from the TS to the MC in the enzyme and the corresponding MCS in solution. These structures were obtained using some constraints on a few selected internal coordinates. The profiles accounted only for around 1/3 of the total activation free energy lowering provided by the enzyme. Using a combination of gas phase and continuum model calculations these authors estimated the free energy change associated with the transformation of separated reactants to MCS, named as 'cratic' free energy in their work, which amounted up to ~10 kcal mol^{-1}; this is 2/3 of the total catalytic effect. Roca *et al.*[40] obtained the AM1/MM PMF for the



Scheme 3 (Adapted in part with permission from ref. 40. Copyright 2003 American Chemical Society.)

reaction in solution from the TS to the fully solvated separated reactants passing through the MCS and a solvent separated ion pair complex. From this profile it is clear that the MCS, which is not a local minimum, is an advanced point of the reaction path in solution and obviously the free energy barrier measured from this point is largely reduced (by about 7 kcal mol^{-1} in their estimation).

From the electrostatic point of view this reaction presents an interesting feature: it proceeds from oppositely charged reactants (presenting then a large dipole moment, μ) to neutral products (with a small dipole moment) through a TS that will be clearly less polar than the reactants. Thus, solvation stabilises the reactants more than the TS and the interaction energy increases the energy barrier. From long MD simulations carried out at different points of the PMF Roca *et al.*[40] obtained an interaction energy contribution to the energy barrier of 45.7 kcal mol^{-1} in solution while in the enzyme this accounted only for 32.6 kcal mol^{-1}. The reason for this difference was found in the very different nature of the electric fields in the solution and in the enzyme. In solution, the TS is destabilised with respect to the reactants for two reasons: (i) the polarity of the solute diminishes and (ii) consequently the reaction field is also decreased. The interaction energy roughly depends on the product of the solute dipole moment and the magnitude of the reaction field ($\vec{R}$). This last, in turn, is a function of the solute dipole moment. Thus, the TS destabilisation, relative to the reactants, is a function of the square of the change in solute polarity.

$$\Delta E \propto \Delta(\vec{R} \cdot \vec{\mu}) \propto \Delta\mu^2 \tag{8}$$

In the enzyme, the electric field ($\vec{R}$) was found to be a permanent field and then, the variation of the interaction energy would be only a function of the variation of the substrate polarity:

$$\Delta E \propto \vec{E} \cdot \Delta\vec{\mu} \tag{9}$$

Therefore, destabilisation of the TS is expected to be lower than in solution and the activation barrier will be smaller. The important point is that the permanent electric field created in the enzyme active site is able to stabilise the TS with respect the solution phase; while the relative stability of reactants remains unaltered. These facts assure an energetically favourable binding process along the enzyme reaction pathway.

Roca *et al.*[40] also compared the electrostatic environment provided by the enzyme for the TS and the MC, concluding that the enzyme presents an electrostatic preorganization, just as found in BsCM. This result can be used to give a unified view of TS stabilisation and substrate preorganization. If the enzyme environment is prepared to assist the charge transfer (relative to the in solution process) and this environment does not change a lot in arriving at the MC, it must also favour reactants rearrangements where this charge transfer was more advanced than in solution. In this latter medium, the most stable reactant conformation would be that structure with the largest dipole moment; *i.e.* fully separated solvated species. Effectively, comparison of reactant trajectories in solution and in the enzyme active site clearly confirms the preorganization of the substrate, correctly positioning the methyl group for the subsequent transfer. This effect was related to the electrostatic potential profile provided by the enzyme.

6 Future perspectives

In previous sections we have presented different theories that have been proposed to explain the origin of enzyme catalysis applied to a unimolecular and to a bimolecular reaction catalysed by enzymes; the CM and the COMT, respectively. The common conclusion is that, for both systems, the structure of the enzyme is the origin of the catalysis. Although this conclusion has been derived from two particular cases where the mechanism in solution and in the enzyme remains unchanged, probably this can be considered as a general conclusion. The enzyme structure is the result of a long evolution that has taken place over ages to optimise the kinetics of the reaction. The rate enhancement is achieved because the active site is exquisitely complementary to the TS from electrostatic and geometrical points of view, thus stabilising it more than the reactant state, so reducing the barrier to reaction. The fact that the enzyme is flexible does not directly imply that its deformation is without an energy cost. Thus, a force that compresses the chemical system in the direction of the TS appears as a natural consequence of the enzyme structure. The relationship between chemical and mechanical energy has been recently the topic of some studies to interpret some chemical processes. This point of view renders a lot of possibilities to understand the enzymatic reactions.

The conclusions presented in this paper are, obviously, the result of experimental and theoretical studies that are constrained by the limitations on the current state of the methods and techniques. In regard to Computational Chemistry, as consequence of the synergic development of computers and methodologies, accurate simulations have arrived at the level where the full enzyme and explicit solvent molecules can be included. Present QM/MM studies are oversimplified in several aspects: i) the level of theory that describes the QM region; ii) the size of the QM region; iii) the lack of a complete sampling of the MM region; and iv) the lack of a complete sampling of the QM region. To solve the first limitation the correlation energy has to be included in the calculations to properly describe the process of bond breaking and formation. For the second limitation, the main difficulty is the description of charge transfer between sub-systems, a problem that can be reduced by increasing the size of the QM region. Nevertheless, the question of cutting a covalent bond between two zones that are not described with the same level of theory is inherent to QM/MM methods. The problem of sampling the MM region will be solved by increasing the total length of the simulations, which is associated with the development of computer capabilities. A similar answer will be given to the last limitation, although more complicated due to the specific features of the QM methods that require a lot of CPU time at each evaluation of the wave function and its derivatives. Methods based on DFT will probably be an efficient alternative in the future.

From the studies commented above, it looks that a static perspective of enzyme structures is incomplete. Although some methods have already appeared to study the dynamics of the chemical processes in enzymes, the coupling between the dynamics of the protein and the chemical reaction is still in an early stage of

its development. There are also some open questions related to the quantitative estimation of the conformational changes of the protein in going from the solvated state to the enzyme–substrate solvated complex. This is especially important in those cases where the chemical reaction is not the rate-limiting step. Anyway, the calculation of this first step of the full process, the formation of the Michaelis complex, seems to be essential for modelling the full catalytic process.

In this paper we have presented two enzymatic reactions where the mechanisms in solution and in the enzyme are the same. This is not the most common situation due to the fact that enzyme catalysis, as any other kind of catalysis, usually works by changing the molecular mechanism of the reaction to a multi-step process. The enzyme flexibility may then play an important role preparing the active site to the requirements of each chemical step. The conclusions derived from this review could be still useful if we focus on the rate-limiting step of the full process. The knowledge of the origin of enzyme catalysis and the full mechanism that governs the process will render very important consequences. These could include rationalized drug design and the more efficient synthesis of catalytic antibodies. In spite of the progress carried out in this field, as shown in this review, a long path is still waiting to be walked.

References

1 R. Wolfenden and M. J. Snider, *Acc. Chem. Res.*, 2001, **34**, 938.
2 L. Pauling, *Chem. Eng. News*, 1946, **24**, 1375.
3 J. Villà and A. Warshel, *J. Phys. Chem. B*, 2001, **105**, 7887.
4 A. Warshel, *J. Biol. Chem.*, 1998, **273**, 27035.
5 A. Warshel, *Proc. Natl. Acad. Sci. USA*, 1978, **75**, 5250.
6 M. I. Page and W. P. Jenks, *Proc. Natl. Acad. Sci.*, 1971, **68**, 1678.
7 P. A. Kollman, B. Kuhn, O. Donini, M. Peräkylä, R. Stanton and D. Bakowies, *Acc. Chem. Res.*, 2001, **34**, 72.
8 A. D. Mesecar, B. L. Stoddard and D. E. J. Koshland, *Science*, 1997, **277**, 202.
9 T. C. Bruice, *Acc. Chem. Res.*, 2002, **35**, 139.
10 F. M. Menger, *Acc. Chem. Res.*, 1993, **26**, 206.
11 A. Warshel and M. Levitt, *J. Mol. Biol.*, 1976, **103**, 227.
12 N. Reuter, A. Dejaegere, B. Maigret and M. Karplus, *J. Phys. Chem. A*, 2000, **104**, 1720.
13 J. Gao and D. G. Truhlar, *Annu. Rev. Phys. Chem.*, 2002, **53**, 467.
14 A. J. Turner, V. Moliner and I. H. Williams, *Phys. Chem. Chem. Phys.*, 1999, **1**, 1323.
15 M. J. Field, *A practical Introduction to the Simulation of Molecular Systems*, Cambridge University Press, Cambridge, UK, 1999.
16 P. Kollman, *Chem. Rev.*, 1993, **93**, 2395.
17 Y. Zhang, H. Liu and W. Yang, *J. Chem. Phys.*, 2000, **112**, 3483.
18 P. R. Andrews, G. D. Smith and I. G. Young, *Biochemistry*, 1973, **12**, 3492.
19 D. J. Gustin, P. Mattei, P. Kast, O. Wiest, L. Lee, W. W. Cleland and D. Hilvert, *J. Am. Chem. Soc.*, 1999, **121**, 1756.
20 P. Kast, M. Asif-Ullah and D. Hilvert, *Tetrahedron Lett.*, 1996, **37**, 2691.
21 S. D. Copley and J. R. Knowles, *J. Am. Chem. Soc.*, 1987, **109**, 5008.
22 H. A. Carlson and W. L. Jorgensen, *J. Am. Chem. Soc.*, 1996, **118**, 8475.
23 M. P. Repasky, C. R. Werneck Guimarães, J. Chandrasekhar, J. Tirado-Rives and W. L. Jorgensen, *J. Am. Chem. Soc.*, 2003, **125**, 6663.
24 P. D. Lyne, A. J. Mulholland and W. G. Richards, *J. Am. Chem. Soc.*, 1995, **45**, 11345.
25 S. Martí, J. Andrés, V. Moliner, S. Silla, I. Tuñón and J. Bertrán, *J. Phys. Chem. B*, 2000, **104**, 11308.
26 S. Martí, J. Andrés, V. Moliner, E. Silla, I. Tuñón and J. Bertrán, *Chem. Eur. J.*, 2003, **9**, 984.
27 H. Guo, Q. Cui, W. N. Lipscomb and M. Karplus, *Proc. Natl. Acad. Sci.*, 2001, **98**.
28 S. Hur and T. H. Bruice, *J. Am. Chem. Soc.*, 2003, **125**, 5964.
29 R. J. Hall, S. A. Hindle, N. A. Burton and I. H. Hillier, *J. Comput. Chem.*, 2000, **21** , 1433.
30 M. Barbany, H. Gutiérrez-de-Terán, F. Sanz, J. Villà-Freixa and A. Warshel, *ChemBioChem*, 2003, **4**, 277.
31 E. Kangas and B. Tidor, *J. Phys. Chem. B*, 2001, **105**, 880.
32 A. Mandal and D. Hilvert, *J. Am. Chem. Soc.*, 2003, **125**, 5598.
33 S. Martí, J. Andrés, V. Moliner, E. Silla, I. Tuñón, J. Bertrán and M. J. Field, *J. Am. Chem. Soc.*, 2001, **123**, 1709.
34 Y. L. Lee, E. Worthington, M. Krauss and B. R. Brooks, *J. Phys. Chem. B*, 2002, **106**, 12059.
35 M. Strajbl, A. Shurki, M. Kato and A. Warshel, *J. Am. Chem. Soc.*, 2003, **125**, 10228.
36 N. A. Khanjin, J. P. Snyder and F. M. Menger, *J. Am. Chem. Soc.*, 1999, **121**, 11831.
37 K. E. Ranaghan, L. Ridder, B. Szefczyk, W. A. Sokalski, J. C. Hermann and A. J. Mulholland, *Mol. Phys.*, 2003, **101**, 2695.
38 C. R. Werneck Guimarães, M. P. Repasky, J. Chandrasekhar, J. Tirado-Rives and W. L. Jorgensen, *J. Am. Chem. Soc.*, 2003, **125**, 6892.
39 B. Kuhn and P. A. Kollman, *J. Am. Chem. Soc.*, 2000, **122**, 2586.
40 M. Roca, S. Martí, J. Andrés, V. Moliner, I. Tuñón, J. Bertrán and I. H. Williams, *J. Am. Chem. Soc.*, 2003, **125**, 7726.