J Mol Model (2008) 14:735–746
DOI 10.1007/s00894-008-0283-2

ORIGINAL PAPER

A dynamic view of enzyme catalysis

Aurora Jiménez · Pere Clapés · Ramon Crehuet

Received: 13 November 2007 / Accepted: 1 February 2008 / Published online: 6 March 2008

Abstract Recent experimental advances have shown that enzymes are flexible molecules, and point to a direct link between dynamics and catalysis. Movements span a wide time range, from nano- to milli-seconds. In this paper we introduce two aspects of enzyme flexibility that are treated with two appropriate techniques. First, transition path sampling is used to obtain an unbiased picture of the transition state ensemble in chorismate mutase, as well as its local flexibility and the energy flow during the chemical step. Second, we consider the binding and release of substrates in L-rhamnulose-1-phosphate aldolase. We have calculated the normal modes of the enzyme with the elastic network model. The lowest frequency modes generate active site deformations that change the coordination number of the catalytic zinc ion. The coordination lability of zinc allows the binding and release of substrates. Substitution of zinc by magnesium blocks the exchange of ligands.

Keywords Enzyme catalysis · Enzyme dynamics · Transition path sampling · Energy relaxation · QM/MM · Chorismate mutase · Aldolase

Electronic supplementary material The online version of this article (doi:10.1007/s00894-008-0283-2) contains supplementary material, which is available to authorized users.

A. Jiménez · P. Clapés · R. Crehuet (✉)
Departament de Química de Pèptids i Proteïnes,
Institut d´Investigacions Químiques i Ambientals
de Barcelona (IIQAB–CSIC),
Jordi Girona 18,
08034 Barcelona, Catalonia, Spain
e-mail: rcsqtc@iiqab.csic.es

Introduction

That proteins are flexible is now well established [8, 32, 58, 69]. Enzymes, being proteins, are also expected to be flexible, but it was not until recently that the influence of motion upon catalysis was realised. [3, 6] The traditional, and still most important, source of structural information, i.e. X-ray crystallography, has recently been expanded with new techniques that focus on dynamics. Advances in spectroscopic fluorescence single-molecule experiments [28, 70, 82] and relaxation-dispersion NMR [25, 26] have shown that enzymes have a static and dynamic disorder that is coupled to their catalytic cycle. It has even been suggested that nature has evolved to a point where "the chemical reactions in the active site are so fast that the limitation of the enzyme catalytic power is its ability to move" [26].

Dynamics can span a wide range [65, 75, 79]. The catalytic step in the enzyme active site corresponds to the cleavage or formation of a few bonds. This takes place at the typical bond stretching frequencies, below picoseconds. The rattling of residues jumping among different rotamers involves barriers of the magnitude of the thermal energy, and can take place as rapidly as tens of picoseconds [65, 75]. At the other end of the spectrum we have large scale protein motions that can be as slow as hundreds of microseconds, or even protein folding and unfolding events in the scale of seconds or minutes [9, 79]. The spatial scale is correlated to the temporal scale. Local events, involving a few atoms moving short distances, are fast, while large scale movements, involving whole protein domains are slower. Obviously, no single computational, or experimental, technique can span all these time and spatial scales.

When dynamics involve a well-defined set of states, as opposed to diffusive motion, then kinetic theory can be

applied [43]. In this case, it is worth remembering that fast events are not related to large rate constants. A large rate constant connecting two states involves a short life-time or residence time in each state. This time will be orders of magnitude larger than the transition time—the time it takes to travel between the states, otherwise, kinetic theory is not applicable [14, 43]. In fact, larger energy barriers involve lower rate constants but shorter transition times [18, 94].

Local flexibility of side chain and substrate can make an important contribution to the free energy. This is accounted for in the calculation of potentials of mean force (PMF), i.e. the free energy along a defined reaction coordinate. PMFs are the usual method used to describe energy profiles in enzymes, rather than the minimum energy paths used in small molecule calculations. For chemical reactions, a driving potential is needed to overcome the energy barrier, and this is usually achieved by using a harmonic constraining potential. This technique is termed umbrella sampling [87]. After sampling, several approaches to remove the effect of the artificial potential and obtain an unbiased PMF are available, including the weighted histogram analysis method (WHAM) [62], or the apparently more promising "umbrella integration" [57]. There are many alternative methods based, for example, on non-equilibrium procedures such as Bennett's acceptance ratios [7, 81], or on Jarzynski's equation [53].

The essence of these methods is the use of a constraining potential along a reaction coordinate. The need to define a reaction coordinate poses two problems. First, when several bonds take part in the reaction, the definition of the reaction coordinate can be complex. For instance, if a nucleophilic substitution is accompanied by proton abstraction by the leaving group, four bonds are involved in this reaction, in a (a priori unknown) synchronous or asynchronous way. Second, the precise choice of the reaction coordinate forces a certain mechanism and can exclude other possibilities. Continuing with the previous example, the choice of which proton is being abstracted, or the choice of a reaction coordinate that forces a step-wise mechanism, can bias the results. Although this can be corrected by trying different reaction coordinates, it is always possible to miss a relevant mechanism. Besides, description based on driving potential does not allow a dynamic picture of the reaction to be obtained, but only the (equilibrium) free energy profile. Transition path sampling (TPS) was developed to overcome these limitations [11, 12, 21, 22, 37]. The essence of TPS is to generate an equilibrium ensemble of reactive trajectories. This can be used, after an expensive calculation, to obtain the rate constant for the process under study. It can also be used, as we have done here, to locate an ensemble of transition structures and to study the average dynamic evolution. This ensemble is not biased by any predefined reaction coordinate.

TPS generates trajectories that become locally decorrelated after some steps, but the time scale spanned is too short to sample large scale movements. When we consider protein domain displacements, coarse-graining becomes a good approach in many cases [88]. One of the key discoveries in modelling protein dynamics was that sequence determines dynamics only in an indirect way. Indeed, sequence determines the 3-dimensional (3D) folding, and it is this 3D structure that directly determines collective dynamics [64, 71, 85]. That means that if a 3D structure is known (from NMR or crystallography) one can get a good description of its dynamics without a detailed description of each residue. These coarse-grained models have been successfully applied to describe the mobility of several proteins and to improve the resolution of X-ray and cryo-microscopy structures [20]. It has also been shown that catalytically active residues lie on mechanical elements of enzyme, such as hinges [93]. Thorough comparisons with atomic-scale models have confirmed the validity of these approaches [80]. However, because coarse-grained methods cannot describe atomic detail, they pose a significant limitation to the study of enzyme catalysis.

In this paper we present two aspects of the integration of dynamics into the catalytic step of enzymes. The first is the application of TPS to chorismate mutase (CM). The second, a study of the influence of large scale motions of L-rhamnulose-1-phosphate aldolase (RhuA) to the active site and the catalytic cycle. CM catalyses the Claisen rearrangement of chorismate into prephenate, and has been investigated by many groups (see, e.g. [27, 39, 40, 41, 44, 48–52, 63, 66–68, 76–78, 84, 90, 95]). A scheme of the reaction

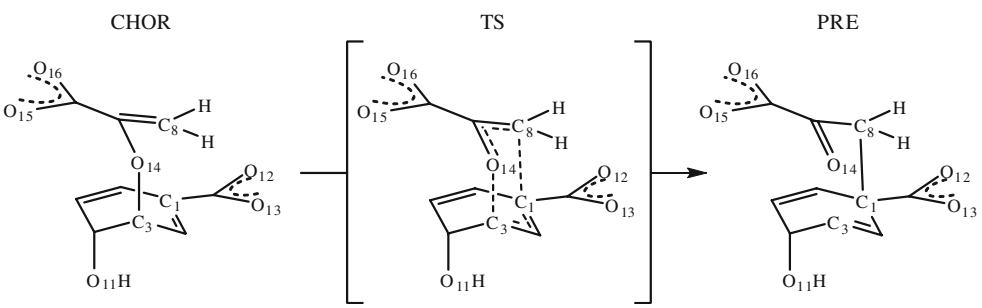

Fig. 1 Scheme of the reaction catalysed by chorismate mutase (CM) with the atom numbering employed in the text

along with the atom numbering is given in Fig. 1. Despite the large number of experimental and theoretical studies, the role of some residues is still unknown and the origin of catalysis is debated. RhuA is a tetrameric enzyme that catalyses the stereoselective addition of dihydroxyacetone phosphate (DHAP) on L-lactaldehyde to give L-rhamnulose-1-phosphate. Crystallographic results have suggested that large scale domain motions are involved in the catalytic event. We explore the effect of these motions on a model of the active site and conjecture a role for its metal ion. This sheds light on how enzyme dynamics can be coupled to the active site to modulate the chemical properties of its metal and, vice-versa, how evolution has picked up a metal that can bind or unbind ligands depending on the conformation of the protein. Thus, chemical properties become intimately connected with conformational changes.

Methods

Chorismate mutase

We refer the reader to reference [68] for a full description of the preparation of the system, and to reference [17] for the implementation and application of TPS to CM, and the location of transition structures in the TPS ensemble. The general procedure for implementing a TPS algorithm can be found in [22]. A brief description of the calculation steps follows. We employed one homotrimer from the PDB 1COM structure [15, 16], solvated it, and generated the substrate structure from the transition state analogue in the crystal. The simulation system contained 17,183 atoms and we fixed 13,532 atoms—equivalent to those that were more than 12 Å away from the substrate. All QM/MM calculations were performed with the DYNAMO library [29, 30]. The QM region comprised the substrate and was treated with the AM1 Hamiltonian. The remaining atoms of the protein and solvent were in the MM region and were described with the OPLS-AA force field [55].

The TPS calculations were performed in a canonical (NVT) ensemble at 300 K. TPS is a Monte Carlo procedure that, starting from an arbitrary reactive trajectory, generates an equilibrium ensemble of reactive trajectories. The length of each trajectory is fixed (160 fs). We chose to describe all reactive events, so than all trajectories started from the reactant (chorismate) basin and ended in the product (prephenate) basin. These basins are defined according to the bond distances characteristic of chorismate and prephenate [17]. A trajectory may have one or more transition structures. A transition structure is defined as a structure that has the same probability of reaching the reactant or the product basin: $p_{\text{reactant}} = 1 - p_{\text{product}} = 0.5$ [22]. To calculate the probability of ending in the reactant or product regions, several trajectories with random velocities were initiated from each frame of each trajectory. The number of trajectories was chosen to obtain an error in p_i of ±0.03, evaluated from the standard deviation of a binomial distribution [17, 22]. This leads to the calculation of about 200 trajectories if the probability of reaching the reactant basin is $p_{reactant}$=0.5. The results shown in Fig. 6 were calculated in this way. All trajectory frames having $p_{\text{reactant}} = 0.5 \pm 0.03$ define the ensemble of transition structures. An alternative method of locating only transition structures, based on an intrinsic reaction coordinate developed in [17] was used to generate the results shown in Fig. 4.

The ensemble of transition structures located from the TPS trajectories was compared with an ensemble of structures where the reaction coordinate $s = d_{C1-C8} - d_{C3-O14}$ (see Fig. 1) was constrained to 0.275 Å with a harmonic potential. This value was obtained from the transition state described in [68], and corresponds to the maximum of the PMF. The position of this maximum depends on the calculation method, but in this work the setup of the system is exactly the same as ours. These structures have been obtained from 15 ps sampling steps, after 5 ps of equilibration; the time-step was 1 fs. This follows the standard procedure used when calculating PMFs—transition structures generated this way are termed TS-PMF to differentiate them from the transition state ensemble (TSE) derived from the unconstrained TPS.

Because TPS introduces no constraint on the dynamics, we can calculate the average dispersion of transition state structures in the TSE. This is easily expressed with the temperature, or B-factor, as is commonly done in crystallography [38]. The B-factor is defined as:

$$B_i = \frac{8\pi^2}{3M} \sum_{i=1}^{M} (\mathbf{r}_i - \overline{\mathbf{r}}_i)^2 \qquad (1)$$

where M is the number of structures, r_i is the position of the ith atom and $\overline{\mathbf{r}_i}$ the average position, previously calculated as $\overline{\mathbf{r}_i} = \frac{\sum_{i=1}^{M} \mathbf{r}_i}{M}$. We also calculated the temperature factors for the reactant, the product, and a transition state analogue (TSA, 8-hydroxy-2-oxa-bicyclo[3.3.1]non-6-ene-3,5-dicarboxylic acid) that had been used in previous crystallisation studies [16]. These were obtained from a 5.5 ns simulation. To avoid correlated structures, structures to be used in Eq. 1 were separated by 600 fs.

A representation of the active site is shown in Fig. 2. We consider the following residues to belong to the active site and refer to them as "active site residues": Phe57, Ala59, Lys60, Arg63, Val73, Thr74, Cys75, Arg7*, Glu78*, Arg90*, Tyr108*, Leu115*, Arg116* (where the * indicates that these residues belong to a different chain from the others).

To study the energy dissipation, trajectories longer than 160 fs are needed. We have chosen 300 reactive trajectories

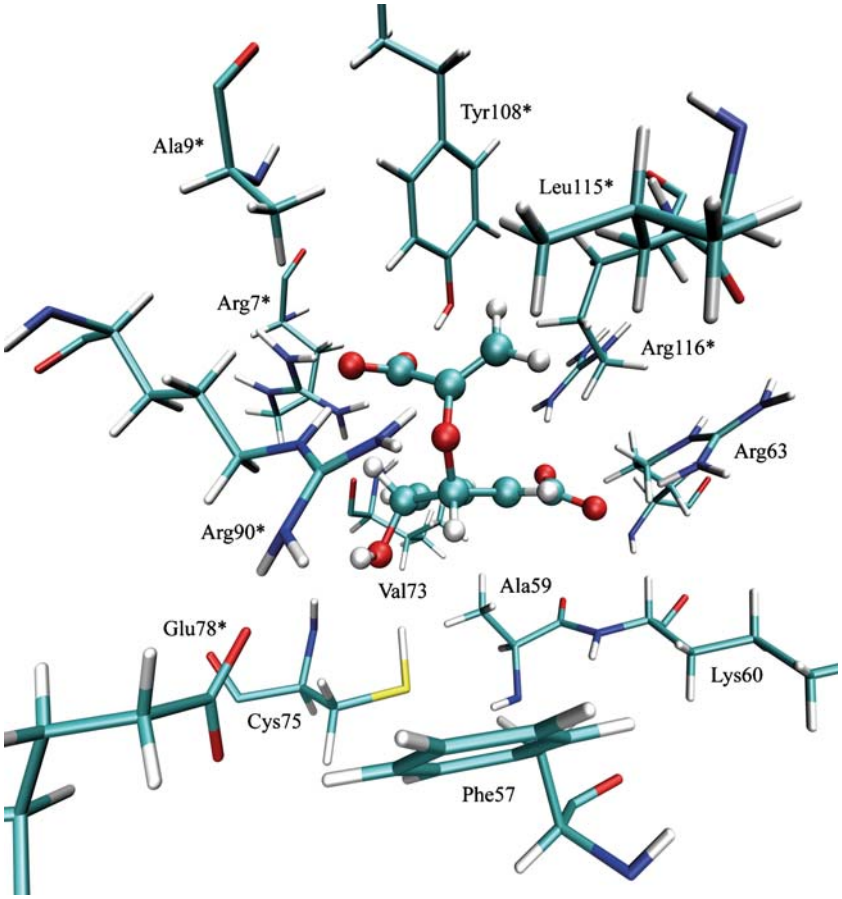

Fig. 2 The active site of CM. The chorismate substrate is drawn with *balls* and *sticks* and the protein residues with *sticks* only

from the TPS calculation and elongated them 0.9 ps from each side to obtain trajectories of a total of 1.9 ps. Because kinetic energy is an extensive property, and we have compared fragments of the enzyme–ligand system with different sizes, we compared the results in terms of an intensive property, i.e. the temperature. The temperature is just an average kinetic energy, defined as:

$$T = \frac{1}{3Nk_B}\sum_{i=1}^{N} m_i \mathbf{v}_i \mathbf{v}_i \qquad (2)$$

Where m and v_i are the mass and velocity of each atom, respectively, k_B the Boltzmann constant, and N the number of atoms for the fragment considered.

L-rhamnulose-1-phosphate aldolase

Our aim was to investigate whether domain motions and the coordination of the metal ion in the active site are coupled. This involves the challenge of coupling a large amplitude motion with a local chemical step. We devised an approximate procedure, described below, which may give a first insight into this issue.

We calculated the normal modes of the RhuA tetramer with the approximate C-alpha model [46] using MMTK software [45]. The C-alpha is a type of elastic network model, where a coarse-grained description of the protein is used. Only the C_α of each amino acid is kept in the simplified representation of the protein, and a harmonic potential is applied between a C_α and its neighbours. This allows very fast calculation of normal modes. Several studies have shown that results for these approximate normal modes describe the same conformational space as more accurate normal modes or essential dynamics studies [60, 80]. We generated a model of the active site with the first coordination sphere of the zinc ion, and generated structures along the six lowest energy modes. The phosphoglycolohydroxamic acid (PGH) molecule was simplified as shown in Fig. 3 (and supplementary Fig. S3, see electronic supplementary material). Because normal modes describe only the displacements of the amino acid residues, the PGH position had to be extrapolated from the movements of the closest residues (see electronic supplementary material). In order to check that the arbitrary nature of this choice does not create an artifact, two different choices of fixed atoms and driving residues were designed and the calculations were repeated for these two sets of model systems. Here, we discuss only one set of results, but the other set is consistent with it and is presented in the electronic supplementary material. All the calculations were

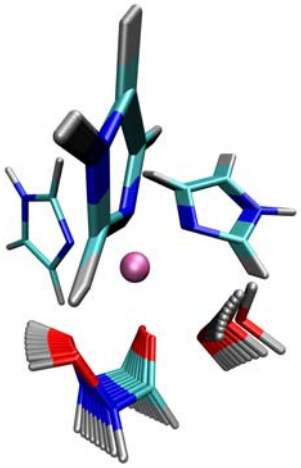

Fig. 3 Model of the L-rhamnulose-1-phosphate aldolase (RhuA) active site depicting the structures generated along normal mode 3. Note how the water molecule becomes looser with the reorientation of phosphoglycolohydroxamic acid (PGH). Colour coding: *red* O, *cyan* C, *blue* N, *grey* H, *purple* Zn

performed with the B3LYP functional and the LANL2DZ basis set, a method that performs correctly for zinc coordination [83]. The calculations were performed under vacuum. Although a higher dielectric constant for the medium does influence the binding properties of the Zn [24], our focus is on trends along different normal modes, not on absolute binding energies. Gaussian 03 [33] was used for DFT calculations.

The proposed model has several simplifications that do not allow precise quantitative results to be deduced. However, our aim here was only to test the validity of our hypothesis to see if further investigations were relevant. The use of a molecular dynamics approach to describing the whole active site requires the use of semi-empirical methods, and the results obtained so far are not satisfactory. Neither AM1, PM3 nor the modified zinc parameters for PM3 [13] can correctly reproduce the zinc lability obtained with DFT. Therefore, although our model is simple, it can give better results than the traditional QM/MM approach. We are working towards a more complete description of the active site and its mobility without losing the accuracy of DFT.

Results

Chorismate mutase

Does the transition state have a stronger binding?

The idea that enzymes preferentially stabilise transition state (TS) structures is well known to most researchers. Many studies have analysed the energetic contributions of different residues, and of different energy components. However, this has been an experimentally elusive issue. Recently, Noonan et al. [73] showed that, in the enzyme cytidine deaminase, TS analogues had lower B-factors than the reactant or product analogues. This was an indirect proof of tight TS binding. But did the inevitable use of analogues introduce a caveat?

Figure 4 compares the B-factor for the TSA and the actual TSE. It also shows the results for the reactant and product ensemble. Although the TSA has lower B-factors than the reactant or the product, the values for the TSE are the lowest. From that we can conclude that the crystallographic results are correct but that they underestimate the strongest binding of the TS. These results can be explained by considering the network of H-bonds around the ligand. Stronger interactions reduce the mobility of the atoms and result in lower B-factors, as can be seen in Fig. 5 and in reference [17]. Leu115* cannot be explained by H-bond interactions because it cannot establish them. However, the results for the TSE are markedly lower than for any other of the species, suggesting close contact between the ligand and this residue. The importance of this residue in energy dissipation will be discussed below in the section on dynamically relevant residues.

How good is the TS from a simple reaction coordinate?

In our previous study [17] we showed that the use of a single reaction coordinate defined as $s = d_{C1-C8} - d_{C3-O14}$ was not sufficient to discriminate structures committed to the reactant basin, to the transition state region, or to the product basin. Figure 6 shows that the value usually assigned to the TS [68] is s=0.275 Å, there are structures that have a probability of staying in the reactant basin of $P_{reactant}$=0.7, and others that have the same probability of staying on product basin, $P_{reactant}$=0.3. Therefore, in terms of dynamics, this is not the best reaction coordinate.

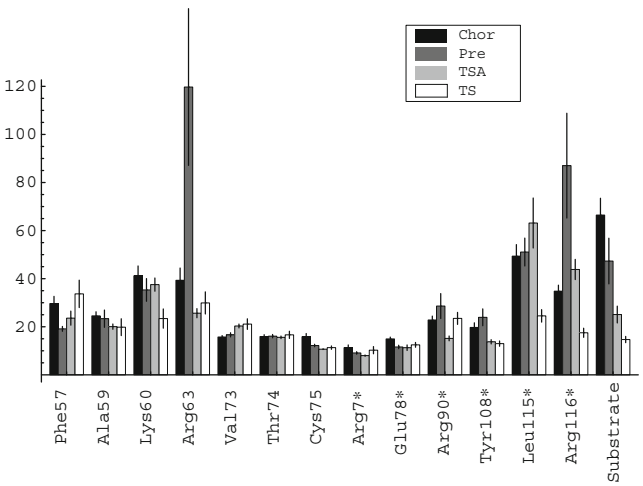

Fig. 4 Calculated B-factors for the substrate and active site residues. *Chor* Chorismate, *Pre* prephenate, *TSA* transition state analogue, *TS* transition state ensemble from transition path sampling. The *error bars* were determined as the variance of the block-average mean using data block sizes of 2^{10} points

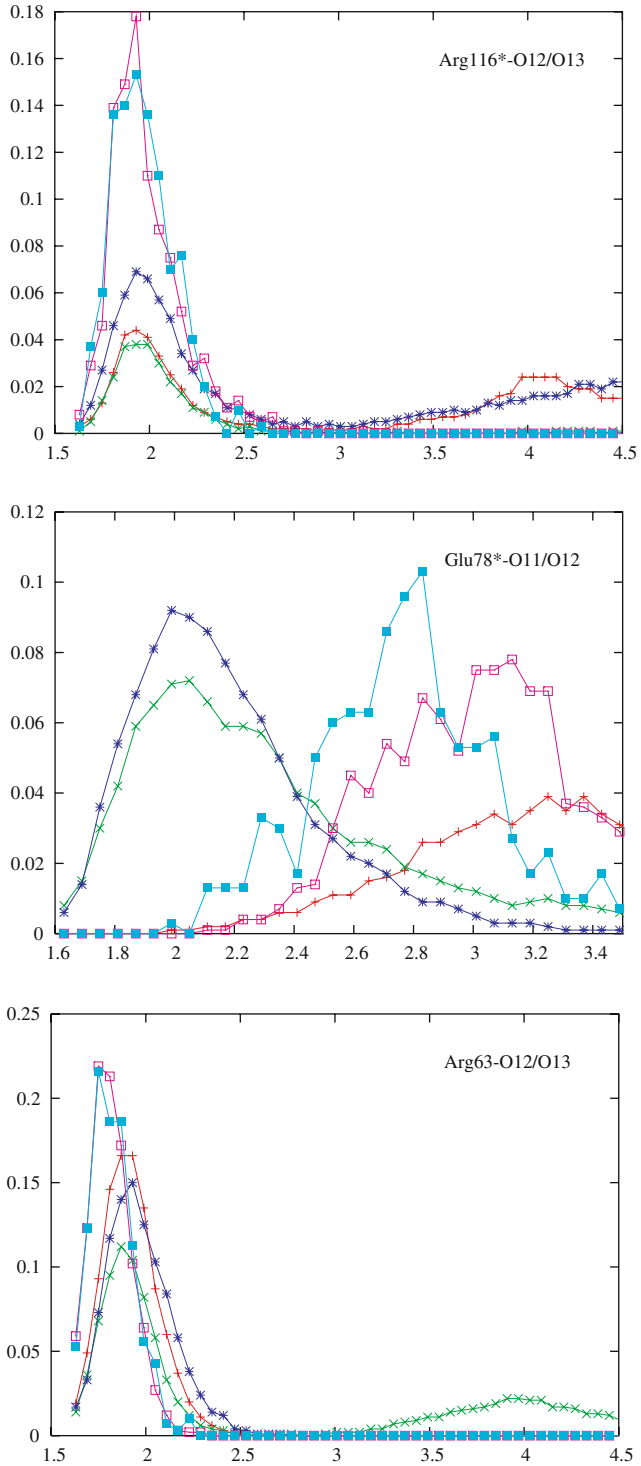

Fig. 5 Hydrogen-bond length distributions for a representative set of ligand-active site residue interactions: *red* chorismate, *green* prephenate, *blue* transition state analogue, *magenta* transition state ensemble, *cyan* TS-PMF (potentials of mean force)

In the previous section, we showed that the interactions in the TSE are specific to this ensemble, and are different from that of reactant, product or even a TS analogue. Some H-bonds with critical residues are present only in the TSE, thus conferring more rigidity on this ensemble. This can be seen in the temperature factors of the substrate and active site residues. The TS analogue behaviour is closer to that of the TSE, but the structures that give rise to the lowest temperature factors are the TSEs, both for the substrate and for most of the residues, as can be seen in Fig. 4.

When free energies are calculated from a PMF, the TS is assumed to be the ensemble of structures (TS-PMF) that lie on the top of the profile. We have studied the H-bond distributions between the substrate and relevant residues in the active site. We have found that, in all but one case, the TS-PMF behaves as the true TSE. This is good news, because it means traditional TS-PMF can be used to analyse the interactions in the TS region. The harmonic constraints do not affect the interactions of this ensemble. Figure 5 shows three representative cases. Interactions characteristic of the TSE, either because they are stronger (as with Arg63–O12/O13) or because they are present mainly in the TSE (as with Arg116*–O12/O13), have the same pattern as the TS-PMF. The only small inconsistency occurs with Glu78*–O12/O13. In this case the distance for the TSE seems to be slightly longer than for the TS-PMF. In both cases, however, the H-bond is not present. In contrast, the TSA does show the presence of the H-bond. Thus, the TS-PMF and the TSE agree and point to a differential behaviour with respect to the TSA. The TS-PMF can be obtained at a much lower computational cost than the TSE, which requires a long TPS calculation.

Dynamically relevant residues

Not only do enzymes catalyse a chemical reaction, but they also behave as a bath for the substrate, because they are

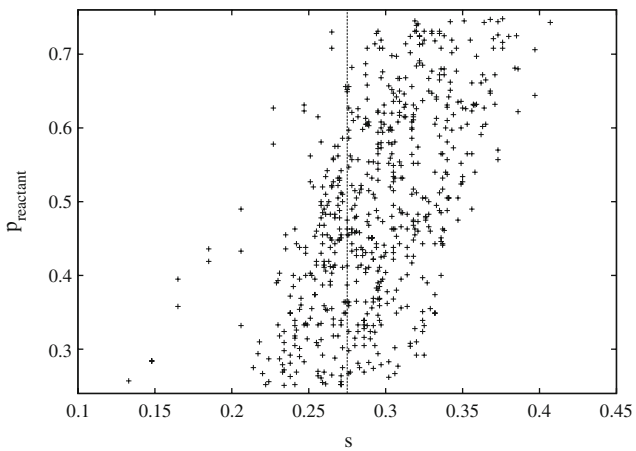

Fig. 6 Probability of reaching the reactant region with respect to the reaction coordinate $s = d_{C1-C8} - d_{C3-O14}$. The *vertical line* indicates the position of the TS-PMF (see text), s=0.275 Å. Each *cross* represents a frame from a transition path sampling (TPS) trajectory. Its probability is calculated by fleeting multiple trajectories from that frame, as explained in Methods

much bulkier than most synthetic catalysts. This allows them to capture the energy released in a chemical reaction and convert it to other forms, such as that needed for motor proteins and ion pumps. Indeed, the catalytic site seems to be located so as to be in contact with residues relevant to large-scale movements. [93] For enzymes that have no mechanical activity, dynamics can be relevant in two aspects. First, dynamics may contribute to energy dissipation to avoid overheating of the enzyme [4, 28]. Second, they may be actively coupled to the catalytic step, promoting the reaction [1–3, 5, 6, 42]. To our knowledge, the first hypothesis has never been tested, either computationally or experimentally. The second proposition is not completely accepted [89, 91] and has been discussed mainly in theoretical works.

Here, we consider the flow of energy from the enzyme to the substrate, and within the substrate, to the reacting atoms. Figure 7 plots the temperature of all the mobile atoms of the system. The effect of the finite size of the system can be seen. The peak in the t=0 region corresponds to the TS. Because these conformations have more potential energy and less kinetic energy, the equilibrium temperature reached by the system outside the TS region is higher by 1 degree for the reactant and 2 degrees for the product. This difference in heating agrees with the exothermicity of the reaction, which is approximately the same as the energy barrier (with the AM1 Hamiltonian) [67]. A change of 1–2 degrees is not relevant for the overall results but could be important if a smaller system is considered when applying TPS.

In the substrate, the four reacting atoms acquire the initial energy release (see Fig. 8). These atoms remain hot for the entire trajectory, but receive a burst in energy about 50 fs after the TS. This timing is similar to the reaction

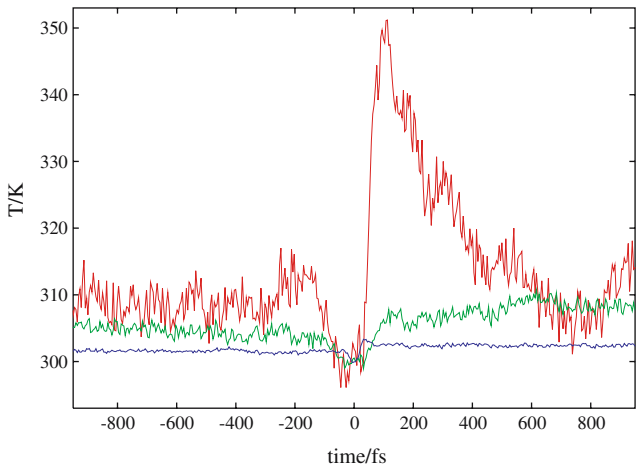

Fig. 7 Evolution of the temperature of selected sets of atoms in CM: *red* Leu115*, *green* active site residues, as defined in Fig. 4 and the text, *blue* all 3,651 free atoms

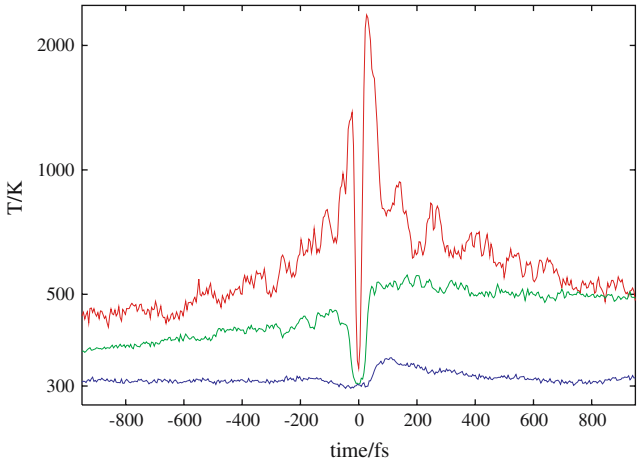

Fig. 8 Evolution of the temperature (in log scale) of selected sets of atoms in CM: *red* the four reacting atoms, *green* the remaining substrate atoms, *blue* Leu115*. Note the change of scale with respect to Fig. 7

event, which lasts less than 125 fs [17]. During the chemical step, the kinetic energy is transfered to potential energy to overcome this barrier. The remaining substrate atoms are also depicted in Fig. 8. Compare their variation with the rest of the residues in Fig. 7. As expected, the energy is redistributed fastest within the substrate, because there are covalent bonds joining the atoms. It is interesting to note that, even after 900 fs, the temperature of the product is still 200 degrees above the rest of the enzyme.

The most surprising result in this section is the behaviour of Leu115*. As can be seen in Fig. 7, the temperature of this residue rises far above the average of all the other residues around the substrate. Furthermore, it also responds faster than the average residue. Leu115* lies just above the substrate, in close contact with O14 and C8. Our calculations show that most of the energy is released from the substrate towards this residue. Gamper et al. [36] studied which residues could be mutated in the C-terminal tail without drastically diminishing the activity of the enzyme. It was found that Leu115* was the most conserved residue. Here, we suggest that its dynamic role could be the main reason underlying its conservation. The need for close contact with the ligand in the TS region could also explain why we found very low B-factors for Leu115* in the TSE with respect to reactants, products or the TSA (see Fig. 4). This also agrees with the computational findings of Kitaura and co-workers [52], who reported a movement of the C-terminal helix that places Leu115* closest to the substrate in the TS region.

Just on the other side of the substrate lies Ala59. We studied the temperature evolution of this residue to see if it parallelled the behaviour of Leu115*. Our results (see electronic supplementary material) show a much smaller effect for this residue. Interestingly, the modest amount of

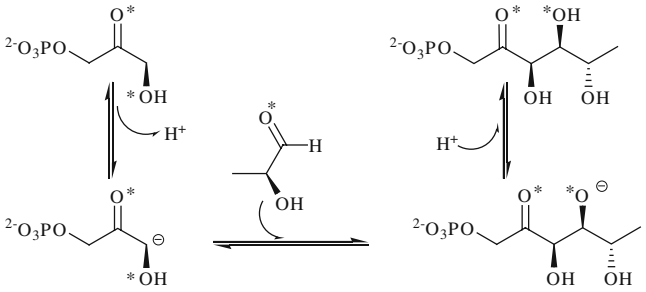

Fig. 9 Scheme of the reaction catalysed by L-rhamnulose-1-phosphate aldolase (RhuA). *Asterisks* indicate atoms that are likely to bind the Zn ion in some of the substrate conformations

heating that takes place is in the reactant region, and not in the product. We have no explanation for this at present. All these results highlight the importance of the time evolution study of the energy flow and demonstrate that an analysis of geometrical contacts in the TS is insufficient to understand the role of active site residues.

In this section we have completely neglected quantum effects on nuclei. To further understand the energy redistribution and quantum effects, it would be interesting to compare our results with results from solution and to apply the techniques developed for this field [56, 72], or to consider the analytical quantum approaches taken by others [34, 35].

L-rhamnulose-1-phosphate aldolase

The biological unit of RhuA is a homotetramer of 274 amino acid residues. The enzyme was recently crystallized [61] with a dihydroxyacetone phosphate analogue, phosphoglycolohydroxamic acid (PGH), which is a bidentate ligand that coordinates the zinc ion in the active site. This zinc ion is hexacoordinated; three positions are occupied by three histidine residues and the remaining one by a water molecule. In the free enzyme, the Zn ion is probably coordinated to more than one water or an active site residue, which have to be substituted by DHAP to give a structure analogous to the crystallographic structure. During the catalytic cycle, the remaining water molecule has to be substituted by the aldehyde. After the aldolic condensation takes place, the negative charge that was located mainly in the initial DHAP oxygen is transferred to the initial aldehyde oxygen. This structure is better stabilised if this oxygen is coordinated to the positively charged metal [19]. Finally the substrate must be released. This cycle is depicted schematically in Fig. 9. All this implies that the coordination sphere of the zinc ion changes significantly during the reaction cycle. Until now, it was assumed that the only function of the zinc was to bind the substrate and reduce the pKa of the alpha proton [54, 59].

A surprising result that emerged during our modelling studies with this aldolase was the fact that full optimisation of a model system containing the active site residues resulted in a structure where Zn was pentacoordinated, with a monodentate PGH, whereas the crystallographic structure shows an hexacoordinated Zn with a bidentate PGH. This immediately points to an entactic geometry for the metal in the active site, where the coordination of the metal is in an energised conformation [92]. Entactic active sites are common in metalloproteins and modulate the metal chemical properties—in this case, a switch from four to six coordination (and possibly also a change in Lewis acid strength). Entactic sites, because they are not relaxed, need to be in rigid areas of the protein, usually buried in β-sheets or cross-linked α-helices [92]. The active site of RhuA, however, corresponds to the contact section of two different chains. Such regions are usually dynamic, playing a mechanical role as hinges. Indeed, the three histidines that form the three upper ligands of zinc belong to the core of the protein formed by β-sheets. However, not only do the residues that bind the PGH belong to a different chain, they also correspond to loops or α-helices. An analysis based on crystallographic temperature factors showed that the

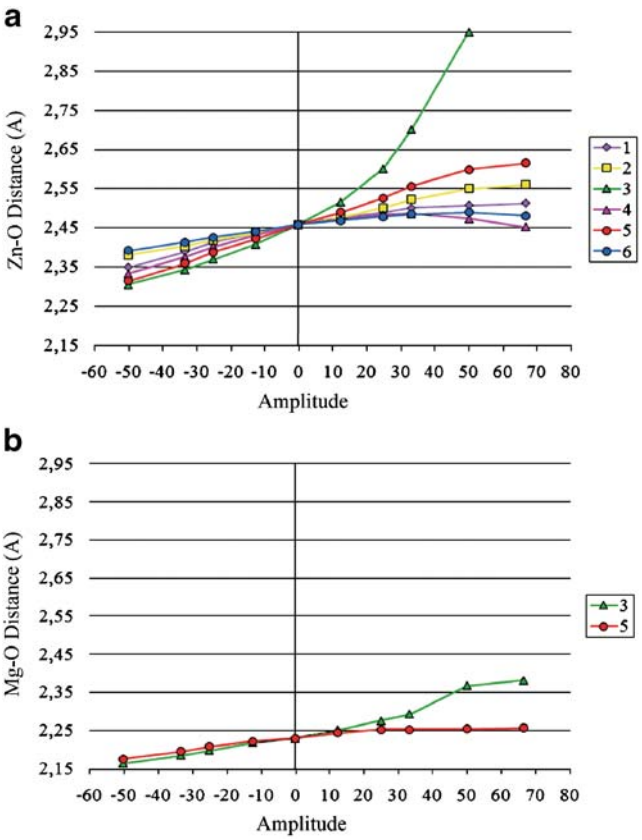

Fig. 10 Variation of the Zn–O (in H_2O) and Mg–O distances for the optimised structures along the selected normal modes of RhuA

N-terminal domain had a higher mobility than the C-domain, and that this motion pushed the N-domain towards the Zn ion and the L-lactaldehyde, favouring the reaction [61].

Normal modes show that the three histidines have little relative movement, which is as expected because they belong to the protein core. However, some of the low frequency modes show a considerable displacement between the histidine residues relative to those residues binding PGH. This suggests that these modes could generate entactic and relaxed zinc states with different properties, in particular with different coordination number preference. Figure 10 shows the distance of the water molecule to the zinc ion for different structures along each normal mode. Modes 3 and 5 produce a large change in this distance, to the point that, for mode 3, the water molecule unbinds the zinc ion.

If nature has indeed used the lability of the zinc atom, coupled to active site mobility, then other metals should show this effect much more weakly. For the modes that change most (3 and 5) we carried out calculations substituting the Zn ion by a Mg ion, as both these ions are of similar size but the latter prefers six coordination environments [10]. Figure 10 shows that the Mg ion is considerably more insensitive to large amplitude motions of the active site, and maintains its rigid hexacoordination sphere throughout. Additional calculations carried out in our group have shown that both metals increase alpha proton acidity to a similar extent [19]. Therefore, we suggest that nature's choice of Zn is strongly related to the latter's coordination flexibility, although other factors are surely also involved, such as the side reactions that have been found with cobalt [47]. The fact that L-fuculose-1-phosphate aldolase, with a similar active site, lacks this water molecule [23], could also be explained by a slightly perturbed active site geometry in the crystal form.

Conclusions

In this TPS study of CM, we have shown how to account for the local dynamics of the residues around the active site during a catalytic event. We have shown that, due to specific TS interactions, the TSE has lower mobility than reactants, products, or TS analogues. These interactions can also be correctly described using a constrained potential, which validates the approaches based on umbrella sampling. These approaches have a much lower computational cost but do not allow evaluation of the time evolution properties along a catalytic event.

This analysis of the kinetic energy of different regions of the enzyme–ligand compound expands the results of our previous work [17]. We have shown how energy flows from the reactant atoms to the bulk of the enzyme in very different time scales, all of which are much larger than that required for the chemical step. One of the residues, Leu115*, displays a particular behaviour, as this residue captures most of the kinetic energy of the reactant. This correlates with its being one of the most relevant residues in the C-terminal region [36].

The results presented in the RhuA study suggest that the type of movements depicted for the approximate modes and the chosen ligands take place throughout the whole catalytic cycle and can contribute to ligand binding and release, further explaining why zinc is the element of choice in the wild type active site of RhuA. This enzyme provides a striking example of how large amplitude motions can be coupled to a chemical property. This leads us to recommend that these motions be included when studying enzyme mechanisms computationally, as is starting to be done in the field of drug design [31, 86] and substrate binding [3, 74].

Acknowledgements R.C. thanks the Spanish Ramón y Cajal Program for financial support, and A.J. the CSIC and the European Social Fund for her I3P fellowship. We acknowledge the financial support of the MEC (Grant CTQ2006–01345/BQU) and the Generalitat de Catalunya (Grant 2005SGR00111). This research has been partly performed using CESCA resources.

References

1. Agarwal P (2005) Role of protein dynamics in reaction rate enhancement by enzymes. J Am Chem Soc 127(43):15248–15256
2. Agarwal PK, Billeter SR, Rajagopalan PR, Benkovic SJ, Hammes-Schiffer S (2002) Network of coupled promoting motions in enzyme catalysis. Proc Natl Acad Sci USA 301:2794–2799
3. Antoniou D, Basner J, Núñez S, Schwartz SD (2006) Computational and theoretical methods to explore the relation between enzyme dynamics and catalysis. Chem Rev 106(8):3170–3187
4. Astumian D (2002) Protein conformational fluctuations and free-energy transduction. Appl Phys A 75(2):193–206
5. Basner JE, Schwartz SD (2005) How enzyme dynamics helps catalyze a reaction in atomic detail: a transition path sampling study. J Am Chem Soc 127:13822–13831
6. Benkovic SJ, Hammes-Schiffer S (2003) A perspective on enzyme catalysis. Science 301:1196–1202
7. Bennett CH (1976) Efficient estimation of free energy differences from monte carlo data. J Comput Phys 22(2):245–268
8. Berendsen HJ, Hayward S (2000) Collective protein dynamics in relation to function. Curr Opin Struct Biol 10:165–169
9. Bernado P, Blackledge M (2004) Local dynamic amplitudes on the protein backbone from dipolar couplings: toward the elucidation of slower motions in biomolecules. J Am Chem Soc 125 (25):7760–7761
10. Bock C, Katz A, Markham G, Glusker J (1999) Manganese as a replacement for magnesium and zinc: functional comparison of the divalent ions. J Am Chem Soc 121(32):7360–7372
11. Bolhuis PG, Chandler D, Dellago C, Geissler PL (2002) Transition path sampling: Throwing ropes over rough mountain passes, in the dark. Annu Rev Phys Chem 53:291–318

12. Bolhuis PG, Dellago C, Chandler D (2000) Reaction coordinates of biomolecular isomerization. Proc Natl Acad Sci USA 97 (11):5877–5882
13. Brothers EN, Suarez D, Deerfield DW II, Merz KM Jr (2004) PM3-compatible zinc parameters optimized for metalloenzyme active sites. J Comput Chem 25(14):1677–1692
14. Chandler D (1998) Barrier crossings: classical theory of rare but important events. In: Berne B, Cicotti G, Coker DF (eds) Classical and quantum dynamics in condensed phase simulations. World Scientific, Singapore, pp 3–23
15. Chook YM, Gray JV, Ke H, Lipscomb WN (1994) The monofunctional chorismate mutase from bacillus subtilis. Structure determination of chorismate mutase and its complexes with a transition state analog and prephenate, and implications for the mechanism of the enzymatic reaction. J Mol Biol 240:476–500
16. Chook YM, Ke H, Lipscomb WN (1993) Crystal structures of the monofunctional Chorismate Mutase from *Bacillus subtilis* and its complex with a transition state analog. Proc Natl Acad Sci USA 90:8600–8603
17. Crehuet R, Field M (2007) A transition path sampling study of the reaction catalyzed by the enzyme chorismate mutase. J Phys Chem B 111(20):5708–5718
18. Crehuet R, Field MJ, Pellegrini E (2004) Transition events in one dimension. Phys Rev E 69:012,101
19. Crehuet R, Jimenez A (2008) Work in preparation
20. Cui Q, Bahar I (2006) Normal mode analysis. Theory and applications to biological and chemical systems. Mathematical and computational biology series. Chapman & Hall/CRC, Boca Raton
21. Dellago C, Bolhuis PG, Csajka FS, Chandler D (1998) Transition path sampling and the calculation of rate constants. J Chem Phys 108(5):1964–1977
22. Dellago C, Bolhuis PG, Geissler PL (2002) Transition path sampling. Adv Chem Phys 123:1–86
23. Dreyer MK, Schulz GE (1996) Refined high-resolution structure of the metal-ion dependent L-fuculose-1-phosphate aldolase (class II) from *Escherichia coli*. Acta Crystallogr, D Biol Crystallogr 52 (6):1082–1091
24. Dudev T, Lim C (2000) Metal binding in proteins: The effect of the dielectric medium. J Phys Chem B 104(15):3692–3694
25. Eisenmesser EZ, Bosco DA, Akke M, Kern D (2002) Enzyme dynamics during catalysis. Science 295(5559):1520–1523
26. Eisenmesser EZ, Millet O, Labeikovsky W, Korzhnev DM, Wolf-Watz M, Bosco DA, Skalicky JJ, Kay LE, Kern D (2005) Intrinsic dynamics of an enzyme underlies catalysis. Nature 438 (7064):117–121
27. Elstner M, Cui Q, Munih P, Kaxiras E, Frauenheim T, Karplus M (2003) Modeling zinc in biomolecules with the self consistent charge-density functional tight binding (scc-dftb) method: applications to structural and energetic analysis. J Comput Chem 24 (5):565–581
28. Engelkamp H, Hatzakis NS, Hofkens J, Schryver FCD, Nolte RJM, Rowan AE (2006) Do enzymes sleep and work? Chem Commun 9:935–940
29. Field MJ (1999) A practical introducction to the simulation of molecular systems. Cambridge University Press, Cambridge, UK
30. Field MJ, Albe M, Bret C, Proust-de Martin F, Thomas A (2000) The DYNAMO library for molecular simulations using hybrid quantum mechanical and molecular mechanical potentials. J Comput Chem 21(12):1088–1100
31. Floquet N, Marechal JD, Badet-Denisot MA, Robert CH, Dauchez M, Perahia D (2006) Normal mode analysis as a prerequisite for drug design: application to matrix metalloproteinases inhibitors. FEBS Lett 580(22):5130–5136
32. Frauenfelder H, Sligar S, Wolynes PG (1991) The energy landscapes and motions of proteins. Science 254:1598–1603
33. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA Jr, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA (2004) Gaussian 03, Revision C.02. Gaussian, Wallingford, CT
34. Fujisaki H, Bu L, Straub JE (2006) Probing vibrational energy relaxation in proteins using normal modes. In: Cui Q, Bahar I (eds) Normal mode analysis. Theory and applications to biological and chemical systems. Mathematical and computational biology series. Chapman & Hall/CRC, Boca Raton, pp 301–323
35. Fujisaki H, Straub JE (2005) Vibrational energy relaxation in proteins. Proc Natl Acad Sci USA 102(19):6726–6731
36. Gamper M, Hilvert D, Kast P (2000) Probing the role of the c-terminus of *Bacillus subtilis* chorismate mutase by a novel random protein-termination strategy. Biochemistry 39:14087–14094
37. Geissler PL, Dellago C, Chandler D (1999) Kinetic pathways of ion pair dissociation in water. J Phys Chem B 103:3706–3710
38. Giacovazzo C, Monaco H, Artioli G, Viterbo D, Ferraris G, Gilli G, Zanotti G, Catti M (2002) Fundamentals of crystallography, 2nd edn. International Union of Crystallography Texts on Crystallography. Oxford University Press
39. Guimarães CRW, Repasky MP, Chandrasekhar J, Tirado-Rives J, Jorgensen WL (2003) Contributions of conformational compression and preferential transition state stabilisation to the rate enhancement by chorismate mutase. J Am Chem Soc 125:6892–6899
40. Guimarães CRW, Udier-Blagović M, Tubert-Brohman I, Jorgensen WL (2005) Effects of Arg90 neutralization on the enzyme-catalyzed rearrangement of chorismate to prephanate. J Chem Theory Comput 1:617–625
41. Guo H, Cui Q, Lipscomb WN, Karplus M (2001) Substrate conformational transitions in the active site of chorismate mutase: Their role in the catalytic mechanism. Proc Natl Acad Sci USA 98 (16):9032–9037
42. Hammes-Schiffer S (2002) Impact of enzyme motion on activity. Biochemistry 41(45):13335–13343
43. Hänggi P (1990) Reaction–rate theory: fifty years after Kramers. Rev Mod Phys 62(2):251–341
44. Higashi M, Hayashi S, Kato S (2007) Transition state determination of enzyme reaction on free energy surface: application to chorismate mutase. Chem Phys Lett 437:293–297
45. Hinsen K (2000) The molecular modeling toolkit: a new approach to molecular simulations. J Comput Chem 21(2):79–85
46. Hinsen K, Petrescu AJ, Dellerue S, Bellissent-Funel MC, Kneller GR (2000) Harmonicity in slow protein dynamics. Chem Phys 261(1–2):25–37
47. Hixon M, Sinerius G, Schneider A, Walter C, Fessner WD, Schloss JV (1996) Quo vadis photorespiration: atale of two aldolases. FEBS Lett 392(3):281–284
48. Hur S, Bruice TC (2002) The mechanism of catalysis of the chorismate to prephenate reaction by the *Escherichia coli* mutase enzyme. Proc Natl Acad Sci USA 99(3):1176–1181

49. Hur S, Bruice TC (2003) Enzymes do what is expected (chalcone isomerase versus chorismate mutase). J Am Chem Soc 125:1472–1473
50. Hur S, Bruice TC (2003) Just a near attack conformer for catalysis (chorismate to prephenate rearrangements in water, antibody, enzymes, and their mutants). J Am Chem Soc 125:10540–10542
51. Hur S, Bruice TC (2003) The near attack conformation approach to the study of the chorismate to prephenate reaction. Proc Natl Acad Sci USA 100(21):12015–12020
52. Ishida T, Fedorov D, Kitaura K (2006) All electron quantum chemical calculation of the entire enzyme system confirms a collective catalytic device in the chorismate mutase reaction. J Phys Chem B 110(3):1457–1463
53. Jarzynski C (1997) Nonequilibrium equality for free energy differences. Phys Rev Lett 78(14):2690–2693
54. Joerger AC, Gosse C, Fessner WD, Schulz GE (2000) Catalytic action of fuculose 1-phosphate aldolase (class II) as derived from structure-directed mutagenesis. Biochemistry 39(20):6033–6041
55. Jorgensen WL, Maxwell DS, Tirado-Rives J (1996) Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. J Am Chem Soc 118:11225–11236
56. Käb G, Schröder C, Schwarzer D (2002) Intramolecular vibrational redistribution and energy relaxation in solution: a molecular dynamics approach. Phys Chem Chem Phys 4:271
57. Kästner J, Thiel W (2005) Bridging the gap between thermodynamic integration and umbrella sampling provides a novel analysis method: "umbrella integration". J Chem Phys 123 (14):144104. DOI 10.1063/1.2052648
58. Kay LE (1998) Protein dynamics from NMR. Nat Struct Biol 5:513–517
59. Kimura E, Gotoh T, Koike T, Shiro M (1999) Dynamic enolate recognition in aqueous solution by zinc(II) in a phenacyl-pendant cyclen complex: Implications for the role of zinc(II) in class II aldolases. J Am Chem Soc 121(6):1267–1274
60. Kondrashov DA, Van Wynsberghe AW, Bannen RM, Cui Q, Phillips GN Jr (2007) Protein structural variation in computational models and crystallographic data. Structure 15(2):169–177
61. Kroemer M, Merkel I, Schulz GE (2003) Structure and catalytic mechanism of L-rhamnulose-1-phosphate aldolase. Biochemistry 42(36):10560–10568
62. Kumar S, Rosenberg JM, Bouzida D, Swendsen RH, Kollman PA (1992) The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. J Comput Chem 13 (8):1011–1021
63. Lee YS, Worthington SE, Krauss M, Brooks BR (2002) Reaction mechanism of chorismate mutase studied by the combined potentials of quantum mechanics and molecular mechanics. J Phys Chem B 106:12059–12065
64. Ma J (2006) Applications of normal mode analysis in structural refinement of supramolecular complexes. In: Cui Q, Bahar I (eds) Normal mode analysis. Theory and applications to biological and chemical systems. Mathematical and computational biology series. Chapman & Hall/CRC, Boca Raton, pp 137–154
65. Markwick P, Bouvignies G, Blackledge M (2007) Exploring multiple timescale motions in protein gb3 using accelerated molecular dynamics and nmr spectroscopy. J Am Chem Soc 129 (15):4724–4730
66. Martí S, Andrés J, Moliner V, Silla E, Tuñón I, Bertrán J (2003) Conformational equilibrium of chorismate. A QM/MM theoretical study combining statistical simulations and geometry optimisations in gas phase and in aqueous solution. Theochem 632:197–206
67. Martí S, Andrés J, Moliner V, Silla E, Tuñón I, Bertrán J (2004) A comparative study of Claisen and Cope rearrangements catalyzed by chorismate mutase. An insight into enzymatic efficiency: transition state stabilization or substrate preorganization? J Am Chem Soc 126:311–319
68. Martí S, Andrés J, Moliner V, Silla E, Tuñón I, Bertrán J, Field MJ (2001) A hybrid potential reaction path and free energy study of the chorismate mutase reaction. J Am Chem Soc 123:1709–1712
69. McCammon JA, Gelin BR, Karplus M (1977) Dynamics of folded proteins. Nature 267:585–590
70. Min W, English BP, Luo G, Cherayil BJ, Kou SC, Xie XS (2005) Fluctuating enzymes: lessons from single-molecule studies. Acc Chem Res 38(12):923–931
71. Ming D, Kong Y, Lambert MA, Huang Z, Ma J (2002) How to describe protein motion without amino acid sequence and atomic coordinates. Proc Natl Acad Sci USA 99(13):8620–8625
72. Neufeld AA, Schwarzer D, Schröeder J, Troe J (2003) Molecular dynamics approach to vibrational energy relaxation. Quantum-classical versus purely classical non-equilibrium simulations. J Chem Phys 119:2502–2512
73. Noonan RC, Carter CW, Bagdassarian CK (2002) Enzymatic conformational fluctuations along the reaction coordinate of cytidine deaminase. Protein Sci 11:1424–1434
74. Núñez S, Wing C, Antoniou D, Schramm V, Schwartz S (2006) Insight into catalytically relevant correlated motions in human purine nucleoside phosphorylase. J Phys Chem A 110(2):463–472
75. Parak FG (2003) Physical aspects of protein dynamics. Rep Prog Phys 66:103–129
76. Ranaghan KA, Ridder L, Szefczyk B, Sokalski WA, Hermann JC, Mulholland AJ (2003) Insights into enzyme catalysis from QM/MM modelling: transition state stabilisation in chorismate mutase. Mol Phys 101(17):2695–2714
77. Ranaghan KE, Ridder L, Szefczyk B, Sokalski WA, Hermann JC, Mulholland AJ (2004) Transition state stabilisation and substrate strain in enzyme catalysis: ab initio QM/MM modelling of the chorismate mutase reaction. Org Biomol Chem 2(7):968–980
78. Repasky MP, Guimarães CRW, Chandrasekhar J, Tirado-Rives J, Jorgensen WL (2003) Investigation of solvent effects for the Claisen rearrangement of chorismate to prephenate: mechanistic interpretation via near attack conformations. J Am Chem Soc 125:6663–6672
79. Daniel RM, Dunn RV, Finney JL, Smith JC (2003) The role of dynamics in enzyme activity. Annu Rev Biophys Biomol Struct 32:69–92
80. Rueda M, Chacon P, Orozco M (2007) Thorough validation of protein normal mode analysis: a comparative study with essential dynamics. Structure 15(5):565–575
81. Shirts MR, Pande VS (2005) Comparison of efficiency and bias of free energies computed by exponential averaging, the bennett acceptance ratio, and thermodynamic integration. J Chem Phys 122(14):144,107
82. Smiley RD, Hammes GG (2006) Single molecule studies of enzyme mechanisms. Chem Rev 106(8):3080–3094
83. Sousa SF, Fernandes PA, Ramos MJ (2007) Comparative assessment of theoretical methods for the determination of geometrical properties in biological zinc complexes. J Phys Chem B 111(30):9146–9152
84. Szefczyk B, Claeyssens F, Mulholland AJ, Sokalski WA (2007) Quantum chemical analysis of reaction paths in chorismate mutase: conformational effects and electrostatic stabilization. Int J Quantum Chem 107(12):2274–2285
85. Tama F, Wriggers W, Brooks CL III (2002) Exploring global distortions of biological macromolecules and assemblies from

low-resolution structural information and elastic network theory. J Mol Biol 321(2):297–305
86. Teague SJ (2003) Implications of protein flexibility for drug discovery. Nat Rev Drug Discov 2(7):527–541
87. Torrie G, Valleau J (1977) Nonphysical sampling distributions in monte carlo free-energy estimation:umbrella sampling. J Comput Phys 23(2):187–199
88. Tozzini V (2005) Coarse-grained models for proteins. Curr Opin Struct Biol 15(2):144–150
89. Villà J, Warshel A (2001) Energetics and dynamics of enzymatic reactions. J Phys Chem B 105:7887–7907
90. Štrajbl M, Shurki A, Kato M, Warshel A (2003) Apparent NAC effect in chorismate mutase reflects electrostatic transition state stabilisation. J Am Chem Soc 125:10228–10237
91. Warshel A, Sharma PK, Kato M, Xiang Y, Liu H, Olsson MHM (2006) Electrostatic basis for enzyme catalysis. Chem Rev 106 (8):3210–3235
92. Williams RJ (1995) Energised (entatic) states of groups and of secondary structures in proteins and metalloproteins. Eur J Biochem 234(2):363–381
93. Yang LW, Bahar I (2005) Coupling between catalytic site and collective dynamics: a requirement for mechanochemical activity of enzymes. Structure 13(6):893–904
94. Zhang BW, Jasnow D, Zuckerman DM (2007) Transition-event durations in one-dimensional activated processes. J Chem Phys 126:074504
95. Zhang X, Zhang X, Bruice TC (2005) A definitive mechanism for chorismate mutase. Biochemistry 44(31):10443–10448