Effect of enzyme dynamics on catalytic activity

DIMITRI ANTONIOU, JODI BASNER, SARA NÚÑEZ and STEVEN D. SCHWARTZ

Department of Biophysics, Albert Einstein College of Medicine, Bronx, New York, USA

1 Introduction

It is widely accepted that the catalytic properties of enzymes are a consequence of binding energy differences between reactants and transition state (TS), arising from the arrangement of residues in the active site. There are two versions of this concept, both of which have experimental support. In the TS binding picture,[1] the configuration that binds most strongly to the enzyme is assumed to be at the top of the uncatalyzed barrier to reaction. This binding releases energy that stabilizes the TS, lowering the energetic barrier to reaction. In the ground state destabilization picture[2] the role of the enzyme is to make the reactants less stable, leading again to a lower barrier to reaction.

In the last few years it has been suggested by us and other groups that enzyme dynamics may play a role in catalysis. We do not claim that these dynamical effects contribute more to catalysis than the standard binding energy effects, but that they should be taken into account in the interpretation of, for example, kinetic isotope effects (KIE) measurements[3] and that they may provide insight to some puzzling data. In particular, our work on the relation between catalysis and enzyme dynamics originated in the effort to understand some unusual properties of the

ADVANCES IN PHYSICAL ORGANIC CHEMISTRY
VOLUME 41 ISSN 0065-3160 DOI: 10.1016/S0065-3160(06)41006-6

following three systems:

1. It is now widely accepted that for some enzymes (e.g., liver alcohol dehydrogenase, thermophilic alcohol dehydrogenase, etc.) proton transfer proceeds through quantum tunneling. The high activation barriers in these systems were consistent with tunneling. However, the KIE were modest, when tunneling would seem to imply high KIE.
2. The enzyme lactate dehydrogenase (LDH) catalyzes the interconversion of lactate to pyruvate. There are two isoforms in the body to accommodate different substrates. Despite the fact that the active site is identical in these two isoforms, one favors the production of lactate and the other production of pyruvate.
3. Crystal structures of human purine nucleoside phosphorylase with several TS analogs showed an unusual geometric arrangement of three oxygens, lying in a close stack. One may question whether this geometry serves a catalytic purpose.

We will show that the answers to all the three of these puzzles involve the dynamics of the enzyme. There has recently been a disagreement among some authors regarding the meaning of the term ''dynamical'', with some suggesting that the term should be reserved for non-equilibrium motions, while others would use it for equilibrium motions. For clarity, we will define the meaning we give to the term ''dynamical'' in this review.

Let us assume that a variable $A(t)$ is coupled to the reaction coordinate and that (A) is its mean value. If a measurement of some property P depends on (A), but not on the particular details of the time dependence of $A(t)$, then we will call it a ''statistical'' dependence. If the property P depends on particular details of the dynamics of $A(t)$ we will call it a ''dynamical'' dependence. Note that in this definition it is not the mode $A(t)$ alone that causes dynamical effects, but it also depends on the timescale of the measured property P. Promoting vibrations (to be discussed in Sections 2–4) are a ''dynamic'' effect in this sense, since their dynamics is coupled to the reaction coordinate and have similar timescales. Conformation fluctuations that enhance tunneling (to be discussed in Section 5) are a ''statistical'' effect: the reaction rate is the sum of transition state theory (TST) rates for barriers corresponding to some configuration, weighted by the probability that the system reaches that configuration. This distinction between dynamic and statistical phenomena in proteins was first made in the classic paper of Agmon and Hopfield.[4]

We will discuss three kinds of motions:

1. ''Rate-promoting'' quasi-harmonic motions, a fast sub-ps effect we and other have proposed (Section 3).
2. Other kinds of sub-ps motions that involve correlated motions of several residues (Section 4).
3. Conformation fluctuations (Section 5).

The structure of this review is as follows. In Section 2 we will review the concept of ''rate-promoting'' vibrations. We will first need to review briefly the theory of

quantum hydrogen/hydride transfer, because it is the large proton mass (relative to an electron) that makes the reaction rate very sensitive to motions that modulate its transfer distance. We will then identify the experimental and computational signatures of these promoting vibrations. We will close with investigating some objections to the possibility of existence of such promoting vibrations. In Section 3 we will apply the theory of Section 2 to the three enzymatic systems that we mentioned earlier in the Introduction.

In Section 4 we will use two theoretical techniques (transition path sampling (TPS) and essential dynamics (ED)) to analyze molecular dynamics trajectories. We will explain how we were able to identify in atomic detail collective motions that affect catalysis.

Finally, in Section 5 we will briefly discuss recent work by Truhlar, Brooks, and Hammes-Schiffer on the relation of conformation fluctuations and catalysis in dihydrofolate reductase (DHFR) and we will propose a new method for studying much slower motions (such as conformation fluctuations) that may affect catalysis.

This review is not meant to be comprehensive of all work on enzyme dynamics and catalysis. The emphasis will be on the work done by our group, mainly on fast sub-ps enzyme motions, while other groups have studied mostly conformational fluctuations. When necessary, we will provide brief descriptions and references to the current work by other groups.

2 Proton transfer and rate-promoting vibrations

We first identified rate-promoting vibrations in enzymatic systems where proton transfer proceeds via quantum tunneling (in this theoretical section, we will use the terms hydrogen, hydride, proton as if they were equivalent). In order to understand why systems with proton tunneling are good candidates for identifying promoting vibrations, we must review the modern theory of quantum charge transfer in condensed phase, which will be the subject of this section.

Excellent recent reviews of the experimental work in tunneling in enzymes have been written by Scrutton,[5] Romesberg and Schowen,[6] and Kohen.[7]

QUANTUM THEORY OF PROTON TRANSFER

The reaction rate of proton transfer in condensed phases depends on several parameters: temperature, potential barrier height, transfer distance, reactant frequency, strength of coupling to the environment. For different values of these parameters, different physical mechanisms dominate which have been described by different theoretical models, in the chronological order they were studied, the parameter regions were:

Region I: The dynamics is over the barrier (as described by TST) or just below the barrier (small quantum corrections).

Region II: The dynamics takes place by tunneling from excited energy states in the reactant well (moderate to large quantum effects).

Region III: The dynamics takes place by tunneling from the ground state in the reactant well (large quantum effects).

Special care must be taken if one attempts to draw conclusions from the temperature dependence of the rate, because the rate has Arrhenius form in all the three regions, but the activation energy has different meaning in each regime.

Region I: Small quantum corrections

This region is often studied with the methods described in the book by Bell,[8] even though it is not really correct to use gas-phase approaches in condensed phase reactions. An assumption is made that there are several energy levels below the top of the barrier and that over the barrier transfer is described by classical dynamics. The TST result for the transfer rate is:

$$k = \frac{k_B T}{2\pi} \frac{1}{Z_0} e^{-\beta V} \tag{1}$$

where $\beta = 1/k_B T$ is the inverse temperature and Z_0 is the partition function for an oscillator of frequency ω_H. If we describe the motion at the reactant well quantum mechanically (QM), then

$$\frac{1}{Z_0} = 2\sinh\left(\frac{\beta\omega_H}{2}\right) \tag{2}$$

If $\beta\omega_H/2 \ll 1$, one arrives at the familiar textbook TST result for the transfer rate, $k = (\omega_H/2\pi)\ \exp(-\beta V)$. However, a typical frequency for a proton–carbon bond vibration is $3000\,cm^{-1}$, so in the present case the opposite limit $\beta\omega_H/2 \gg 1$ is relevant. In this limit $1/Z_0 \simeq \exp(\beta\omega_H/2)$ and the semiclassical result for the rate is obtained by:

$$k = A_H e^{-\beta(V-\omega_H/2)} \tag{3}$$

Equation (3) predicts a KIE (assuming $A_H = A_D$):

$$\ln\frac{k_H}{k_D} = \frac{\beta}{2}(\omega_H - \omega_D) = \frac{\beta}{2}\left(1 - \sqrt{\frac{m_H}{m_D}}\right)\omega_H \tag{4}$$

For C–H bond cleavage, Equation (4) predicts a KIE equal to $k_H/k_D \sim 7$ at room temperature. In the limit where the semiclassical theory is valid, experimentalists measure the Schaad–Swain exponent, $\ln(k_H/k_T)/\ln(k_D/k_T)$. In the special case that the pre-Arrhenius factor A_L is the same for all isotopes (which is not true in most cases) then semiclassical theory predicts for this exponent a value 3.26. Deviations from this value are often interpreted as signs of increased tunneling, but in our opinion this line of argument is based on an oversimplified model of quantum transfer in condensed phases. Note that in tunneling reactions where the ratio $A_H/A_D \neq 1$, the semiclassical theory predicts an exponent that is *not* equal to 3.26 and is temperature dependent.

Region III: Large quantum effects

When the energetic barrier is very high, tunneling takes place from the ground state. In this limit, the Marcus–Levich–Dogonadze theory[9] has been used in the

study of electron transfer in solution and biomolecules. It is assumed that the environment can be described by a 1-dimensional (1-D) coordinate (a questionable assumption when details of environment dynamics are important) and one exploits the fact that in the deep tunneling limit the tunneling matrix element Δ can be used as a small parameter in a perturbative approach and find a transfer rate equal to:

$$k \sim \Delta^2 e^{-\beta(E_r+\varepsilon)^2/4E_r} \tag{5}$$

where E_r is the reorganization energy of the *environment*, ε is the exothermicity and Δ is the tunneling matrix element. Note that the quantum result Equation (5) predicts an Arrhenius form for the rate, similarly to the TST result. In this theory, the KIE is equal to $k_H/k_D \sim \Delta_H^2/\Delta_D^2$, which in the deep tunneling limit has a value of the order 10^3–10^4, much larger than the measured KIE in biological systems. However, we will see later in this section that the Marcus theory approach, while valid for electron transfer, has to be modified for proton transfer.

Region II: Moderate quantum effects

This is the case when tunneling takes place from excited states, but not close to the barrier top, and had eluded solution for decades because there is no small parameter available (as in the previous two cases). Finally, this problems was numerically solved in the mid 1990s[10] and the solution was analytically verified with a different method by our group.[11]

It is instructive to have a feeling for when and how the semiclassical theory fails. In Table 1 we compare some exact results[10] with the predictions of the semiclassical model Equation (3). There is a range of values for the exact result, because the rate depends on friction (an effect that cannot be captured in the semiclassical model). We note that when the reactant frequency becomes large, the semiclassical theory of the rate fails (we should note that despite the failure in predicting the rate, the semiclassical model gives reasonably good prediction for the KIE).

We will briefly describe how it became possible in the 1990s to describe QM this region of moderate quantum effects. One *assumes* that the classical charge transfer problem is adequately described by the Langevin equation:[12]

$$m\ddot{s} = -\frac{\mathrm{d}V(s)}{\mathrm{d}s} - \int_0^t \mathrm{d}t'\gamma(t-t')\dot{s} + F_{\mathrm{env}}(t) \tag{6}$$

where s is the reaction coordinate and $V(s)$ is the potential energy surface (PES). The influence of the enzymatic environment in Equation (6) is represented by the random force $F_{\mathrm{env}}(t)$, which is related to the friction $\gamma(t)$ through the fluctuation–dissipation theorem.[13]

Table 1 The ratio of rates $k_{\mathrm{exact}}/k_{\mathrm{semiclassical}}$

$\omega/_R/k_BT$	Exact	Semiclassical
3.2	0.8–1.8	1.5
4.8	2.0–3.8	2.3
9.6	30–1000	12.9

A most important result in rate theory was the proof[14] that the classical dynamics of s governed by the Hamiltonian equation:

$$H = \frac{p_s^2}{2m_s} + V(s) + \sum_k \frac{P_k^2}{2m_k} + \sum_k \frac{1}{2} m_k \omega_k^2 \left(q_k - \frac{c_k s}{m_k \omega_k^2} \right)^2 \tag{7}$$

is described by Equation (6). The harmonic oscillators ω_k constitute a fictitious effective environment that is constructed to generate the correct friction kernel $\gamma(t)$. We must emphasize that there is no approximation in using harmonic oscillators to describe an anharmonic environment, since these oscillators are an effective medium that is really a Fourier decomposition of the friction kernel $\gamma(t)$, and are only indirectly related to protein dynamics. The approximation lies in assuming that the Langevin equation is a good approximation for describing the classical charge transfer. The advantage of casting the problem in terms of the Hamiltonian Equation (7) is that there are many tools for solving quantum Hamiltonians that have harmonic terms, while there are no methods for solving the quantum version of the Langevin equation.

The beautiful point of this formulation is that the Marcus–Levich–Dogonadze result Equation (5) is the solution of the Hamiltonian Equation (7) in the deep tunneling limit. In addition, the solution of the Hamiltonian Equation (7) in the classical limit reproduces the TST result, corrected for recrossings of the barrier and for memory effects.[12] These results mean that the Zwanzig Hamiltonian provides a unified description of proton transfer reactions in all the three parameter regions defined earlier in this section.

Finally, we should mention that the approach described above is not the only possible for studying quantum proton transfer in condensed phases. Truhlar and coworkers have followed a different methodology, based on variational TST, summarized in detail recently by Truhlar.[15] An advantage of that methodology is that its inputs are related naturally to quantities that are obtained in quantum chemistry calculations. On the other hand, because of its phenomenological character, one cannot easily understand when its approximations are justified.

RATE-PROMOTING VIBRATIONS

The theory outlined in the previous subsection is appropriate for electron transfer, but has to be modified for proton transfer. In this section we will describe the physical justification and mathematical formalism that incorporates these effects.

Hynes' theory of promoting vibrations

In the early 1990s, in their studies of proton transfer in solution using Marcus' rate theory Equation (5), Hynes and coworkers[16,17] noticed the following limitation. If Q is the tunneling distance, it can be shown that the tunneling matrix element that appears in Equation (5) has the form $\Delta \sim e^{-\alpha Q}$. For typical electron transfer reactions

$\alpha \simeq 1\,\text{Å}^{-1}$, while for typical proton transfer reactions $\alpha \simeq 30\,\text{Å}^{-1}$. This means that while electron transfer rates are insensitive to variations of the tunneling distance Q, proton transfer rates, because of the large value of α, depend strongly on motions that possibly reduce the transfer distance Q.

Hynes assumed that the deviation δQ of the transfer distance from its equilibrium value has a harmonic time dependence, $\delta Q = \delta Q_0 \cos(\Omega_{pv} t)$, and calculated the rate using Fermi's golden rule (i.e. the same level of approximation as Marcus' theory). He found for the rate:

$$k = \Delta_0^2 \; e^{-\beta E_M} \; e^{2\alpha^2/\beta M_{pv}\Omega_{pv}^2} \tag{8}$$

where Δ_0 is the tunneling splitting that corresponds to the equilibrium transfer distance, E_M is the activation energy of the ordinary Marcus theory and M_{pv}, Ω_{pv} are, respectively, the mass and frequency of the promoting vibration.

In the previous section we mentioned that Marcus' theory is not a plausible model for describing proton transfer because it predicts a KIE $k_H/k_D = \Delta_H^2/\Delta_D^2$, that is very large, since the tunneling splittings decrease exponentially with the square root of the mass. This problem is remedied by Hynes' work, since the KIE in the presence of a promoting vibration becomes:

$$\frac{k_H}{k_D} = \frac{\Delta_H^2}{\Delta_D^2} \exp\left\{ -\frac{\alpha_D^2 - \alpha_H^2}{\beta M_{pv}\Omega_{pv}^2} \right\} \tag{9}$$

Since α scales like $\sqrt{m}$, the exponential in Equation (9) reduces significantly the KIE. Recently, Hynes reviewed[18] his approach, with emphasis on applications to enzymes.

Benderskii's theory of promoting vibrations

Hynes' formulation is intuitively very appealing, but there are some drawbacks. The environment is described by a 1-D coordinate and the promoting vibration is an artifact that was introduced to modulate the tunneling splitting. In a series of papers on gas-phase proton transfer,[19–21] Benderskii had examined the same effect, tunneling rate modulated by fluctuations of transfer distance, using a Hamiltonian formalism. We will briefly summarize his approach.

Let us assume a symmetric double well PES, $V(s) = as^4 - bs^2$. Its barrier height is $b^2/4a$ and the transfer distance $\sqrt{2b/a}$. Let us assume that a harmonic mode $Q(t)$ is coupled to s through a term cs^2Q. Effectively, the parameter b of the original PES is replaced by $b - cQ(t)$. As $Q(t)$ oscillates, the transfer distance also oscillates around its equilibrium value. In addition, when the transfer distance decreases, the barrier height is lowered. In summary, the simple, symmetric with respect to s, functional form cs^2Q of the coupling reproduces the desired behavior of a PES whose barrier is lowered as the transfer distance decreases.

Benderskii managed to solve this problem in the deep tunneling limit using the instanton method. Roughly speaking, an instanton is the most probable among the

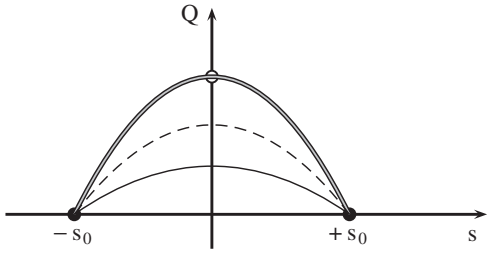

Fig. 1 The PES surface in Benderskii's model. The dark circles are the reactant/product minima and the open circle is the saddle point. The double line is the MEP. The solid line is the instanton tunneling path for H, and the dashed line for D. Because D is heavier, it tunnels closer to the saddle point where the barrier is lower. As a result, the KIE is relatively low, even though the reaction proceeds through tunneling.

possible classical trajectories when one solves the problem QM. We summarize Benderskii's results in Fig. 1.

The double line represents the minimum energy path (MEP), which is the reaction path assumed by TST. The single line represents the instanton trajectory for proton tunneling and the dashed line the instanton trajectory for deuteron tunneling. The heavier deuteron tunnels closer to the MEP, where the barrier is lower. These distinct instanton paths are the reason for the lowering of the KIE by the promoting vibration that we mentioned earlier.

There are other important features in this seemingly simple diagram. The instanton trajectories cross the TS parallel to the *s*-axis, which means that tunneling is happening instantaneously in the timescale of the promoting vibration. But this does not mean that the frequency of the promoting vibration does not play a role! In fact, the result depends on the ratio of the promoting vibration frequency over the barrier frequency Ω_{pv}/ω_b. We have to distinguish between the following limits:[22]

- When $2\Omega_{pv}/\omega_b \ll 1$, we are in the "fast-flip", or "sudden approximation", or "corner-cutting", or "large curvature" limit, where the reaction coordinate follows theMEP, but before it reaches the saddle point it tunnels along the *s* coordinate in a time that is short compared to the timescale of the *Q* vibration.
- When $2\Omega_{pv}/\omega_b \gg 1$, we are in the "slow-flip", or "adiabatic", or "small curvature" limit, where the *Q* vibration adiabatically follows the *s* coordinate and transfer takes place along the MEP path (i.e. at the saddle point).

Therefore very fast promoting vibrations do not affect the rate, and we should expect to see an effect when $2\Omega_{pv}/\omega_b \ll 1$, i.e. when Ω_{pv} is smaller than roughly $500\,cm^{-1}$, which explains why we mentioned earlier that we studied sub-ps motions.

Theory of promoting vibrations in condensed phase

The formulation of promoting vibrations by Benderskii is very satisfactory because it formulated the problem in a Hamiltonian language. On the other hand, the Hynes

formulation, even though it uses an ensatz for the promoting vibration, has the advantage that includes interaction with the environment (with the limitation of the Marcus model, i.e. that the environment is represented by a single degree of freedom). We can incorporate these two theories into the Zwanzig Hamiltonian equation (7) and obtain a theoretically satisfying framework for the description of proton transfer in condensed phases, coupled to a promoting vibration.[22]

We add a term $cs^2Q + \frac{1}{2}M_Q\,\Omega_{\mathrm{pv}}^2$ in Equation (7), to obtain the Hamiltonian:

$$\frac{p_s^2}{2m_s} + V(s) + cs^2Q + \frac{1}{2}M_Q\Omega_{\mathrm{pv}}^2 + \sum_k \frac{P_k^2}{2m_k} + \sum_k \frac{1}{2}m_k\omega_k^2\left(q_k - \frac{c_k s}{m_k\omega_k^2}\right)^2 \tag{10}$$

This incorporates the advantages of Benderskii's and Hynes' ideas, and in addition contains a more realistic description of the environment. One important difference is that Hynes and Benderskii studied systems in which the oscillator Q was a bond vibration, while in Equation (10) Q(t) can be any variable that modulates the PES, for example, it can be a distance between the donor and a nearby residue that changes in time because of equilibrium fluctuations of the enzyme.

To establish a relationship between the Hamiltonian equation (10) and the actual enzymatic system one performs a molecular dynamics simulation to obtain the force $F(t)$ acted upon the reaction coordinate. Then the force autocorrelation function $\langle \mathrm{F(t)F(0)} \rangle$, which is proportional to the friction kernel $\gamma(t)$, is related to the parameters of the fictitious medium of Equation (10) through

$$\gamma(t) = \frac{1}{k_{\mathrm{B}}T}\langle F(t)F(0)\rangle \sim \sum_{k=1}^{N}\frac{c_k^2}{m_k\omega_k^2}\cos(\omega_k t) \tag{11}$$

This equation permits the mapping of the computed force $F(t)$ to the parameters of our Hamiltonian.

To obtain some insight into the behavior of the solutions of the Hamiltonian equation (10), we performed a numerical simulation of a model system:[23] we assumed that $V(s)$ is a symmetric double well, we coupled s to 1000 harmonic oscillators ω_k with frequencies ranging from 10 to 1000 cm^{-1}, and symmetrically to one oscillator Ω_{pv}. Even though the simulation is completely classical, we obtained instructive results that illustrate several of the points we have mentioned in this section.

The next two figures show results for proton transfer in a symmetric double well potential that has barrier height equal to 6 kcal mol^{-1} and transfer distance 1 Å; the proton is coupled symmetrically to an oscillator of frequency 300 cm^{-1}. In Fig. 2 we show the progress of the reaction coordinate from reactants to products. The very fast oscillations are bond oscillations in the reactant/product wells. The slower oscillation that envelopes the bond oscillations is the promoting vibration. Note that the promoting vibration is fast enough that the time the barrier is lowered is not long enough for a reactive event to occur, i.e. it requires several oscillations of the promoting vibration for the charge transfer to occur. Once the crossing over the barrier happens though, it is practically instantaneous in the timescale of the promoting vibration.

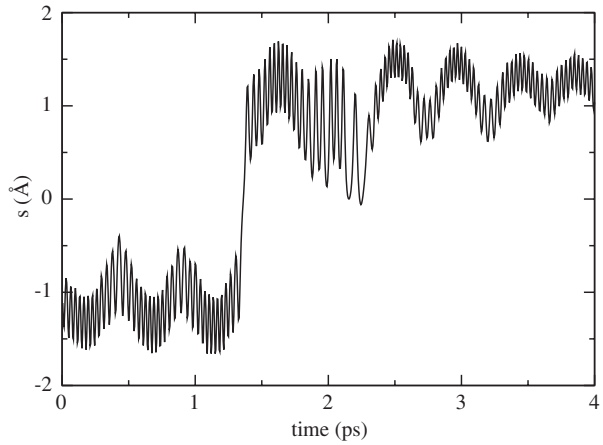

Fig. 2 An example of a reactive trajectory. The TS is at the $s = 0$ line and the reactant/product wells at the $s = \mp 1$ lines.

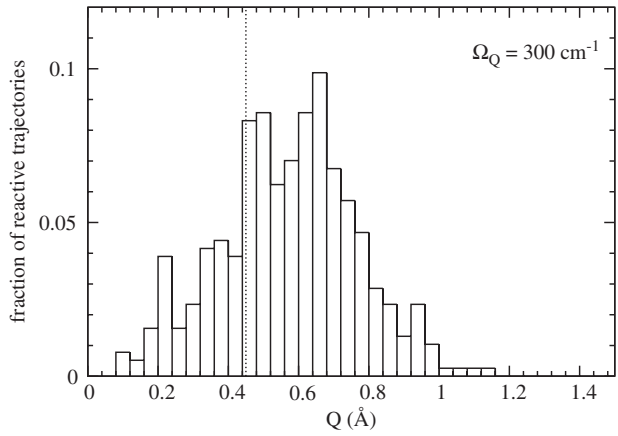

Fig. 3 A histogram of the values of the promoting vibration coordinate Q, as the reactive trajectories cross the TS. The dashed line corresponds to the location of the saddle point.

In Fig. 3 we plot a statistics of the values of the promoting vibration coordinate as the reactive trajectories cross the TS. The saddle point in Fig. 1 corresponds to the value of Q shown with a dashed line in Fig. 3. We note that the reactive trajectories do not pass through the saddle point, but rather through a broad region centered at the saddle point. This picture shows that the dynamics cannot be described by a single TS. The deeper reason for this is that the assumption of separability fails:[24] the promoting vibration is strongly coupled to the reaction coordinate in the TS region and the frequency of the promoting vibration is comparable to the inverted barrier frequency, so there is no separation of timescales.

Finally, in our opinion, there is an open question in quantum theories of proton transfer. The theories we have been discussing use a mean field potential as the PES. This is undoubtedly correct when the environment relaxes quickly in the timescale of the reaction. However, the transfer step in quantum tunneling is very fast, and it is not certain that all the configurations that enter the mean field potential are sampled. This point was first raised by Hynes[25] and implemented in studying proton transfer in solution by Hynes and Wilson.[26]

COMPUTATIONAL SIGNATURE OF PROMOTING VIBRATIONS

In a realistic system there are many other motions present, so there is no guarantee that the effect of the promoting vibration will not be masked by other interactions that are present in an enzyme. We need a diagnostic that when we perform a computer simulation of the dynamics of an enzyme, will allow us to identify whether a promoting vibration is present.

We have found[27] such a computational signature in the framework of the Langevin equation (6). Let us recall that if we compute from a simulation the force $F(t)$ on the reaction coordinate, then the friction kernel $\gamma(t)$ is proportional to the autocorrelation of that force. We have shown that if we add a term $cs^2Q + \frac{1}{2}M_Q\Omega_{\rm pv}^2$ to the Langevin equation, and allowing the promoting vibration to be coupled to the environment with coupling strength ζ, the friction kernel becomes

$$\gamma_{\rm pv}(t) = \gamma(t) + \frac{4c^2}{M\Omega_{\rm pv}^2} s(t)s(0)\left[\cos(\tilde{\Omega}_{\rm pv}t) + \frac{\zeta}{2\tilde{\Omega}_{\rm pv}}\sin(\tilde{\Omega}_{\rm pv}t)\right]e^{-\zeta t/2} \tag{12}$$

Here $\gamma(t)$ is the friction of the original Langevin equation and $\tilde{\Omega}_{\rm pv} = \sqrt{\Omega_{\rm pv}^2 - \zeta^2/4}$ is the effective frequency of the promoting vibration, modified due to its coupling to the environment.

Equation (12) provides the diagnostic we have been looking for. Note that the correction to the friction kernel due to the promoting vibration is proportional to $s(t)s(0)$. Suppose we perform a simulation where we have imposed constraints to keep the transferred proton fixed, so that the correction term is proportional to $s(0)^2$. If we keep the proton fixed near the TS, $s = 0$, the correction term will be very small. If we keep it fixed away from the TS (most simply, at the reactant or product configuration), the correction term will be nonzero. In addition, if we take the Fourier transform of Equation (12), the presence of the trigonometric terms in the correction term will produce large peaks at the frequency of the promoting vibration.

In conclusion, if we perform simulations with the transferred proton fixed near and away the TS, and then take a Fourier transform of the calculated friction kernel, if we see sharp peaks for the latter simulation that are absent when the proton is held fixed near the TS, then we have evidence that a promoting vibration is present, and the position of the peak is an indication of its frequency. In the next section we will discuss examples of enzyme simulations where this diagnostic was successful.

EXPERIMENTAL SIGNATURE OF PROMOTING VIBRATIONS

The discussion in this section has suggested a clear experimental signature for the presence of promoting vibrations. Hynes' formalism, Benderskii's formalism and our quantum calculations using the Hamiltonian equation (10) all predict a low KIE, i.e. much lower than what one would expect for transfer through tunneling. In fact, the initial resistance to accepting that tunneling can occur in enzymes was exactly the low values of KIE. As mentioned earlier, Hynes has recently reviewed[18] the topic of influence of promoting vibrations on KIE, but with a perspective different than ours.

Furthermore, the tunneling rate depends exponentially on the reorganization energy of the environment, which for an enzymatic system depends strongly on the rigidity of the enzyme. This raises some intriguing possibilities for the interpretation of certain experiments on thermophilic enzymes. Thermophilic enzymes show low enzymatic activity at mesophilic temperatures, and the conservation of the active site structure and chemical mechanisms suggests that the same chemical mechanism is present at both thermophilic and mesophilic temperatures. One possible interpretation focuses on the temperature-dependent motions of proteins, because experiments have shown that the rigidity of thermophilic proteins is less at thermophilic rather than at mesophilic temperatures.[28]

The work that is of interest with respect to promoting vibrations is a study[29] of hydrogen tunneling in alcohol dehydrogenase from *Bacillus stearothermophilus*. The unusual features of this system are (i) the activation energy is smaller in the thermophilic regime, which in a naive interpretation would imply that tunneling is enhanced with increasing temperature; (ii) the primary KIE is small and temperature-independent in the thermophilic regime, but larger and temperature-dependent in the mesophilic regime. We have shown[30] that by assuming the presence of a promoting vibration in the thermophilic regime, and assuming that it freezes out in the mesophilic regime, one is able to reproduce all the trends mentioned earlier in this paragraph. Reproducing trends is not definite proof, but at the very least it is a reminder that the presence of a minimal dynamical element makes the problem sufficiently complex, that conclusions derived by studying Arrhenius plots should not be trusted.

Anomalous values of the secondary KIE have been interpreted as indications of tunneling. The reader should consult the arguments regarding secondary KIEs in reviews of experimental literature.[6,7]

FOUR OBJECTIONS TO PROMOTING VIBRATIONS

The real protein dynamics is not harmonic

"Proteins are very anharmonic systems, so what is the justification of modeling the protein environment as a set of harmonic oscillators in Equation (10)?" The real approximation we have made is modeling the proton transfer by the Langevin

equation (6). One expects this to be a good approximation for the short timescales relevant for barrier crossing. Within the Langevin equation, the effect of the environment on the proton is captured through the friction kernel, which is the force autocorrelation function. No approximation is made regarding this friction kernel, it is what is obtained from a simulation. This time-dependent friction kernel can be decomposed into Fourier components, which are the harmonic oscillators appearing in Equation (10). That is, these oscillators are fictitious quantities introduced to reproduce the exact form of the friction kernel, therefore no simplification regarding the protein anharmonic dynamics is made.

The promoting vibration is quickly dephased

"*If we embed a harmonic oscillator in a medium, it will be very quickly dephased. Therefore, the identification in an enzyme of a promoting vibration with a definite frequency is not plausible*." Here, the misunderstanding is that the promoting vibration is not a harmonic oscillator embedded in the enzymatic environment.

The motion of a protein on its PES can be described as anharmonic motions near local minima (i.e. conformations), with rare hops between conformations. While the system executes this motion, we can record, for example, the distance $Q(t)$ between two residues. If the Fourier transform of $Q(t)$ is relatively peaked, then the distance between these residues varied in time like a damped harmonic motion. The quantity $Q(t)$ is not an oscillator with energy levels, that is embedded in the enzyme, rather it is an internal distance of, for example, residues that participate in the equilibrium fluctuations of the enzyme.

In simple model calculations we can mimic this effect by writing a Hamiltonian like Equation (10) in which $Q(t)$ appears as an independent oscillator, but it must be understood that this Hamiltonian is a simplified model designed to produce a fluctuating PES, and $Q(t)$ is a quantity that parametrizes this fluctuation. However, it is true that one cannot assume beforehand that the distance $Q(t)$ is harmonic, one has to calculate it, Fourier transform it and check whether it is peaked at some frequency, as we will do in the examples in the next section.

Parenthetically, we would like to use this opportunity to correct a misunderstanding that is common in enzymatic literature. Highly anharmonic potentials do not necessarily exclude harmonic dynamics! An example is water: the interatomic potential is extremely anharmonic (hard spheres), but water supports harmonic waves (sound). The resolution of the paradox is that the variable that describes sound waves (density) is not the variable that enters the anharmonic interatomic potentials, so it is possible for equilibrium fluctuations, like sound, to have harmonic dynamics.

The promoting vibration is much faster than the turnover rate

"*The turnover rate for enzymes has timescales μs–ns, but the promoting vibration has ps-timescale, therefore they can't be related*." This argument is correct for some gas-phase reactions, but it is not valid for condensed phase reactions.

Let us look at a simple condensed phase reaction, electron transfer in liquids, which is described by Marcus' theory equation (5). As is well known, electron tunneling takes place only when the interaction with the solvent symmetrizes the PES. The reorganization energy required for the solvent to reach its configuration that will symmetrize the PES of the electron, is the activation energy appearing in Equation (5). Once the PES is symmetrized, the probability that the electron will transfer is proportional to the square Δ^2 of the tunneling matrix element. Therefore, even for simple transfer reactions in condensed phases, several events of widely varying timescales contribute to reaction and the turnover rate is not simply related to the timescale of a single event.

The method was designed to identify ps-timescale motions

"*The reason this method identified promoting vibrations for some systems is that by construction it searches only for fast ps motions, and therefore it cannot identify e.g., slower conformational motions that possibly affect the reaction rate.*" This criticism is absolutely correct, if there are conformations that bring donor and acceptor close together thus enhancing tunneling, they are outside the grasp of a method that is based on Langevin equation. In Section 5 we will outline some ideas for searching the conformation space for configurations that affect the reaction rate.

3 Examples of rate-promoting motions in enzymatic systems

We will now present examples of enzymatic systems where we applied the ideas and formalism of the previous section, and we were able to identify rate-promoting vibrations.

HORSE LIVER ALCOHOL DEHYDROGENASE

The first system in which we identified a promoting vibration was horse liver alcohol dehydrogenase (HLADH).[31] The active site and surrounding residues are shown in Fig. 4. The suspicion that dynamics may play a role existed because two specific mutations had been identified, Val203→Ala and Phe93→Trp, which significantly affect enzyme kinetics. Both residues are located at the active site. Val203 impinges directly on the face of the nicotinamide ring in the nicotinamide adenine dinucleotide (NAD) cofactor distal to the alcohol substrate. Additionally, there is evidence from molecular dynamics simulations,[32] that Val203 forces the nicotinamide ring of NAD^+ into closer proximity to the substrate, thus facilitating the hydride transfer to produce the corresponding aldehyde. These facts suggest that motions of these residues may play an important role in catalysis.

In the previous section we found a computational signature for the existence of promoting vibrations: fix the transferred proton at the TS and away from it, and compare the magnitude of the Fourier transform of the forces on the proton, as it is

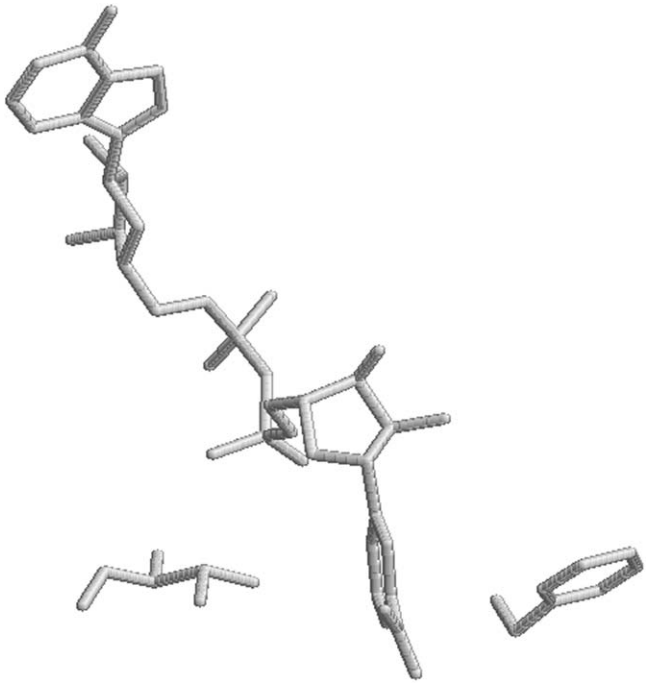

Fig. 4 Active site of HLADH: Val203, the NAD cofactor and the alcohol substrate.

held fixed at different positions. Since the location of the TS is unknown, we tried several intermediate positions and selected the one for which the spectral density is smallest, i.e. a position of minimum coupling.

We shall present results from simulations for three configurations – reactants (R), products (P), and minimal coupling (MC):

- The R configuration consists of NAD^+ and the deprotonated benzyl alcohol ($PhCH_2O^-$).
- The P consists of NADH and benzaldehyde.
- In the third configuration MC, we have NAD^+ and $PhCH_2O^-$ with the pro-R hydrogen restrained so that it is equidistant from the hydroxyl α-carbon (hydride donor) and the C4 carbon (hydride acceptor) in the nicotinamide ring.

We performed a 30-ps simulation and saved the time series for several distances between atoms. In Fig. 5 we plot, for all three configurations, the Fourier transform of the donor–acceptor distance. We see a peak at ca. $100\,cm^{-1}$, common to all three configurations, signifying that the two motions appear to be in resonance.

The Fourier transform of the force on the reaction coordinate is shown in the right panel of Fig. 5. Again, the peak at ca. $100\,cm^{-1}$ is common to both configurations. This is a strong indication that the motion of the transferred proton is

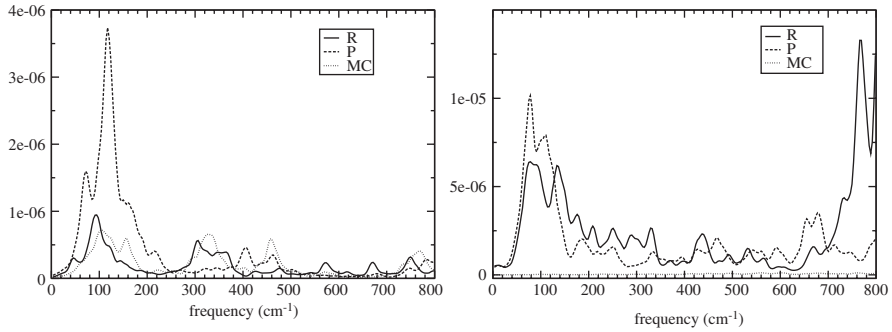

Fig. 5 Left: the spectral density for the donor–acceptor relative motion in the wild type; it monitors the donor–acceptor distance. Right: the spectral density for the reaction coordinate, in the wild type. The three lines represent the reactants configuration (R), the products configuration (P), and the minimal coupling configuration (MC).

coupled to the relative motion between the alcohol and the nicotinamide ring. On the scale of the graph, the spectral density for MC appears as the flat line close to the horizontal axis, which shows that the spectral density is position-dependent, as predicted by Equation (12). Furthermore, if we had shown the MC line in magnification we would see that although the high frequency peaks in the P and C are still present in MC, the peak at ca. $100\,\mathrm{cm}^{-1}$ is almost absent in MC. These results strongly suggest that the reaction coordinate is coupled to an oscillation of frequency ca. $100\,\mathrm{cm}^{-1}$.

In view of our earlier analysis, our results indicate that the reaction coordinate is coupled to the alcohol-NAD motion, which in turn is coupled to Val203, whose side chain impinges directly on the face of the nicotinamide ring. It is thus shown that in HLADH this dynamic coupling is central to the catalytic process. Finally, we should point out that Cui and Karplus[33] used our concept of symmetrically coupled vibrations and performed a simulation on HLADH following a different method, and found results that are in agreement with the results presented here.

LACTATE DEHYDROGENASE

The next system we studied was the isoforms of human LDH. The first part of the calculation was similar to HLADH. However, we encountered interesting complications.

LDH catalyzes the interconversion of lactate and pyruvate with the coenzyme NAD (see Fig. 6). This enzyme plays a fundamental role in respiration, and multiple isozymes have evolved to enable efficient production of substrate in different microenvironments. Two main subunits, referred to as heart and muscle, are combined in the functional enzyme as a tetramer, and subunit combinations range from pure heart to pure muscle. The kinetic properties of the heart and muscle isoforms are

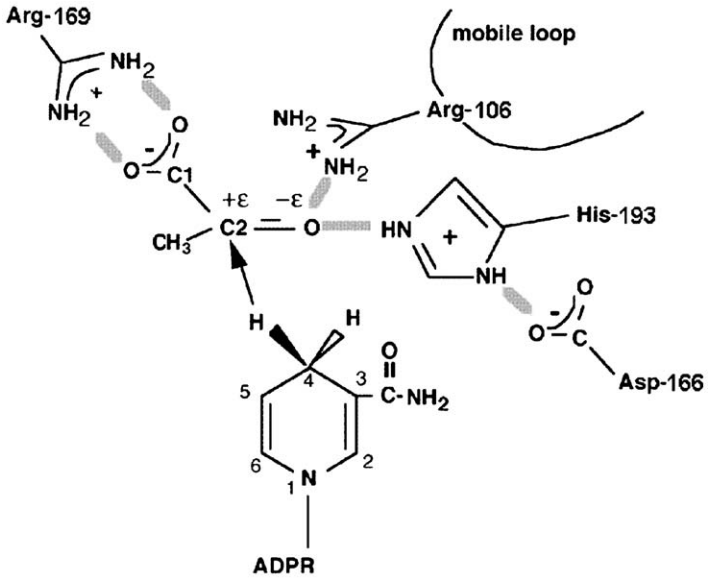

Fig. 6 Diagram of the binding site of LDH with bound NADH and pyruvate showing hydrogen bonds between the substrate and key catalytically important residues of the protein. The catalytic event involves the hydride transfer of the C4 hydrogen of NADH from the pro-R side of the reduced nicotinamide ring to the C2 carbon of pyruvate and protein transfer from the imidazole group of His-193 to pyruvate's keto oxygen.

distinct: the heart favors production of lactate and the muscle of pyruvate. Despite this difference, the domain structure, subunit association, and amino acid content of the active sites of the two isoforms are almost identical, leading to the puzzle of what is the cause of their difference in activity. We propose that placement of the TS, coupled with a promoting vibration, can influence kinetic control of hydride transfer.

Results will be presented for the following configurations: the heart and muscle isoform each with lactate and NAD^+ bound; the heart and muscle isoform each with pyruvate and NADH bound; two simulations of the heart isoform with lactate and NAD^+ bound where the hydride distance was restrained at a point between donor and acceptor carbons. In the following figures, in the language of the previous subsection, the reactant configuration for the heart isoform is the line "lactate" and the product configuration is the line "pyruvate." And vice versa, the reactant configuration for the muscle isoform is the line "pyruvate" and the product configuration is the line "lactate."

In Fig. 7 we show results for the heart LDH isoform: (1) the Fourier transform of the force on the transferred hydride (left); (2) the Fourier transform of the relative motion between the substrate C2 carbon and carbon C4N of the nicotinamide ring of the cofactor NAD^+/NADH (right). In Fig. 8 we show the corresponding figures

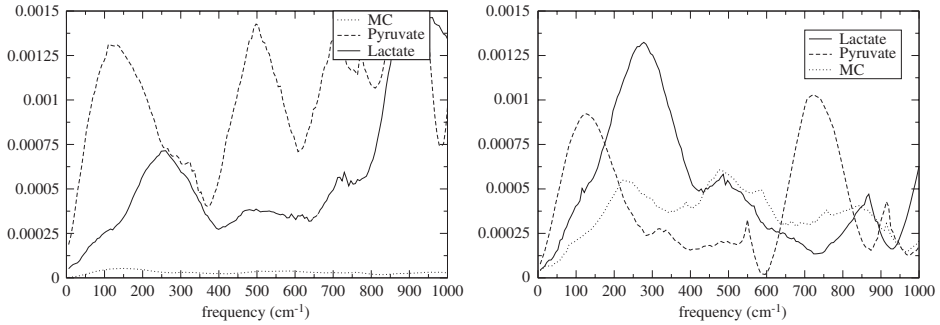

Fig. 7 Wild-type heart LDH isoform: Fourier transform of the force on the reaction coordinate (left) and Fourier transform of relative donor–acceptance distance (right).

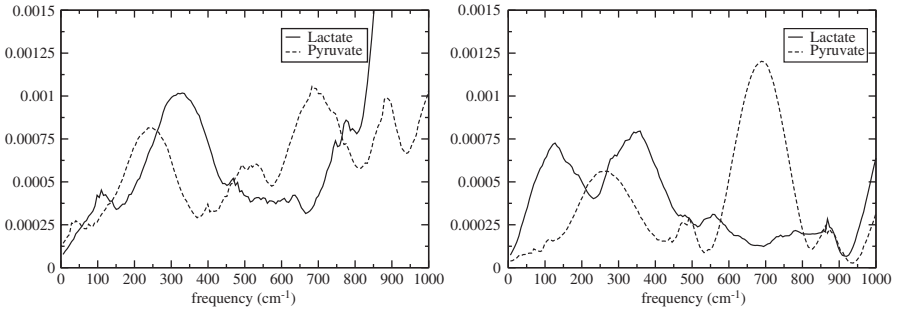

Fig. 8 Wild-type muscle LDH isoform: Fourier transform of the force on the reaction coordinate (left) and Fourier transform of relative donor–acceptance distance (right).

for the muscle LDH isoform. Similarly, to the HLADH simulation we found clear evidence of the presence of a promoting vibration:

(1) The spectral density is much lower when the hydride is near the TS, as predicted by Equation (12) (line MC in the Figure).
(2) The peaks of the Fourier transform of the force on the hydride, are in resonance with the peaks of the Fourier transform of the relative donor–acceptor motion.

These results show strong evidence for the presence of a promoting vibration, but they are symmetric for both isoforms, so they don't resolve the paradox why one isoform favors the production of pyruvate and the other of lactate. However, there is a hint: in the left panels of Figs 7 and 8, the pyruvate peak is higher for the heart isoform and the lactate peak higher for the muscle isoform. Recall that the height of the spectral density is proportional to the force on the reaction coordinate. According to Equation (12), what may lead to bigger force, is bigger $s(t)$, i.e. bigger deviation from the TS. Since the lines "lactate" and "pyruvate" in Figs 7 and 8 correspond to the hydride bound to the donor/acceptor (and vice versa for the other

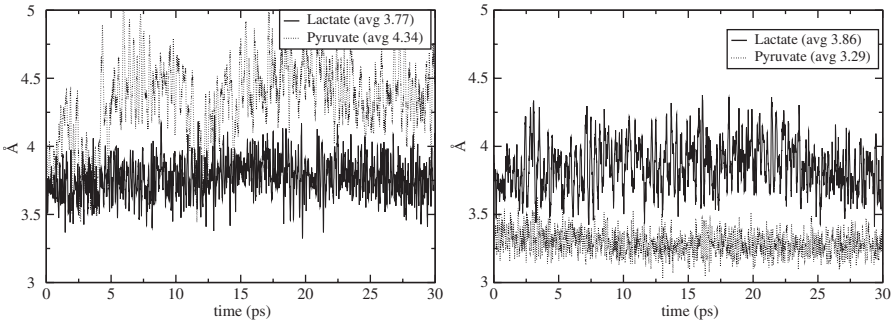

Fig. 9 Donor–acceptor distance for the wild-type human heart LDH isoform (left) and the muscle isoform (right).

isoform), it would be interesting to see what is the average distance of donor–acceptor in the two isoforms.

In Fig. 9 we show the results of a 30-ps simulation for the donor–acceptor distance, i.e. the distance between the C2 carbon of substrate and carbon C4N of the nicotinamide ring of the cofactor NAD^+/NADH. Fig. 9 shows that the average donor–acceptor distance is shorter for the heart isoform when lactate and NAD^+ are bound, and for the muscle isoform when pyruvate and NADH are bound. We propose that the different kinetic activity of the two isoforms is due to the reduced donor–acceptor distance when lactate is bound to the heart isoform and pyruvate is bound to the muscle isoform.

The next question is to identify residues near the active site that may modulate the donor–acceptor distance. In Fig. 10 we show the active site and some nearby residues. In the spirit of the previous results, in order to predict the degree that the motion of these residues is correlated with the donor–acceptor motion, we can calculate the Fourier transform of the autocorrelation of the residue motion, and then order the residues according to the height of the peak of the spectral density.[34] In Fig. 11 we show one result, the spectral densities for the motion, projected first along the residue–donor axis and then along the donor–acceptor direction, of three residues, two of them strongly correlated and one not correlated.

Finally, to bring our argument to its logical conclusion, we can test for the consistency of our interpretation: we substituted residue 31, Valine, with a less bulky one, Alanine. If our interpretation is correct, we expect that the Fourier transform of the force on the hydride would have a higher peak in the wild type than in the mutant, since the bulkier Valine is more efficient in pushing the nicotinamide ring of the NAD^+, and as a result the average donor–acceptor distance would be smaller for the wild type than in the mutant. In Fig. 12, we show the results of the simulation, which are consistent with our prediction. As we mentioned earlier, the peak of the spectral density can be lowered either because the coupling to the reaction coordinate is weaker, or because the reaction coordinate is fixed closer to the TS. The right panel of Fig. 12 shows that the average donor–acceptor distance is larger

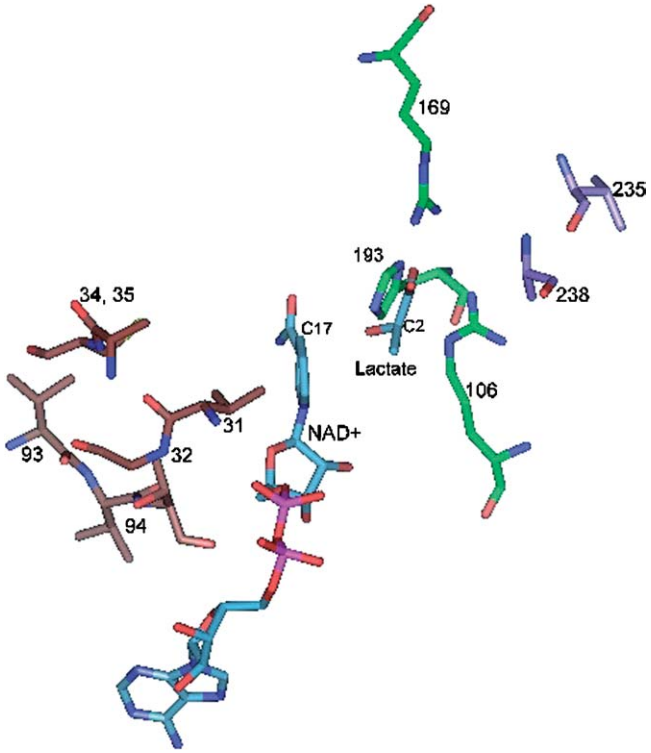

Fig. 10 The structure of the heart isoform with lactate and NAD^+ bound.

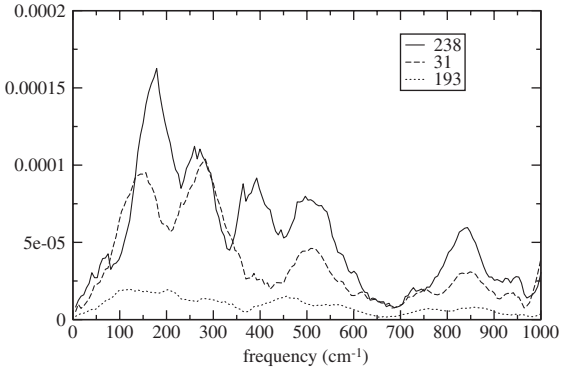

Fig. 11 Spectral density of motion of three residues in the wild-type human heart LDH isoform with lactate and NAD^+ bound: residue 238 is the most strongly correlated residue, 31 is strongly correlated and substituted in the mutagenesis simulation, and residue 193, the essential active site histidine, is poorly correlated.

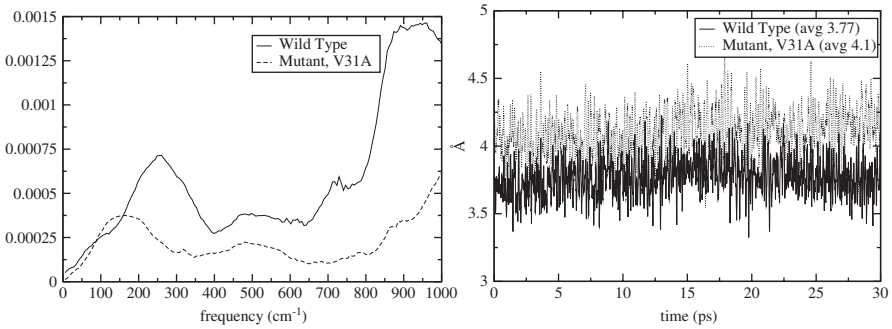

Fig. 12 Left, spectral density for the reaction coordinate in the wild type and mutant human heart LDH isoform with lactate and NAD^+ bound: the solid line is the wild-type configuration where residue 31 is Valine, and the dashed line is the mutant configuration where residue 31 is Alanine. Right, donor–acceptor distance for the wild type and mutant human heart LDH isoform with lactate and NAD^+ bound: the solid line represents the wild-type configuration where residue 31 is Valine and the dashed line represents the mutant configuration where residue 31 is Alanine.

for the mutant result, therefore the lower peak of the spectral density in the left panel necessarily means that the coupling is weaker for the less bulky Alanine.

HUMAN PURINE NUCLEOSIDE PHOSPHORYLASE

Up to this point, all the examples we discussed, concerned protein motions that affect reactions that involve a light particle transfer. However, rate-promoting enzymatic motions can also exist in other types of reactions. For example, we have shown[35] that such motions can influence catalysis by acting through electron density fluctuations caused by geometrical changes. The system we investigated is human purine nucleoside phosphorylase (hPNP), which catalyzes reversible bond cleavage of 6-oxopurine nucleosides to form phosphorylated α-D-ribose products in the presence of phosphate, as seen in Fig. 13.

The cleavage of the C-1′–N-9 ribosidic bond (for atom terminology see Fig. 14) occurs in a dissociative mechanism that forms a TS with a substantial oxycarbenium ion character. The phosphate provides an electrostatic stabilization of this oxycarbenium ion, encouraging TS formation.[36,37] As the N-ribosidic bond is cleaved, electron density is expelled by the oxygen-stack compression towards the purine ring, and improves electrostatic interactions with nearby residues and facilitates the abstraction of a proton from a close-by proton donor, making the purine a better leaving group and accelerating catalysis. In summary, oxycarbenium stabilization, increased phosphate ionization and purine ring activation, contribute in concert to catalytic acceleration. Crystallographic data of several hPNP complexes with TS analogs, showed an unusual geometric arrangement of the atoms O-5′, O-4′, and O_P, lying in a close threeoxygen stack (Fig. 14), which was later corroborated by

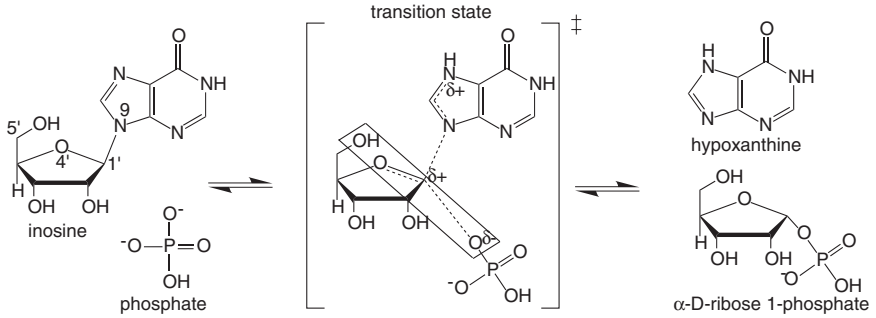

Fig. 13 hPNP-catalyzed phosphorolysis of the purine nucleoside. The guanine leaving group and phosphate nucleophile are well-separated from the oxycarbenium ion, defining a highly dissociative TS. Note the oxygen stacking, indicated by the rectangle, of O5′, O4′, and the phosphate nucleophile O_P.

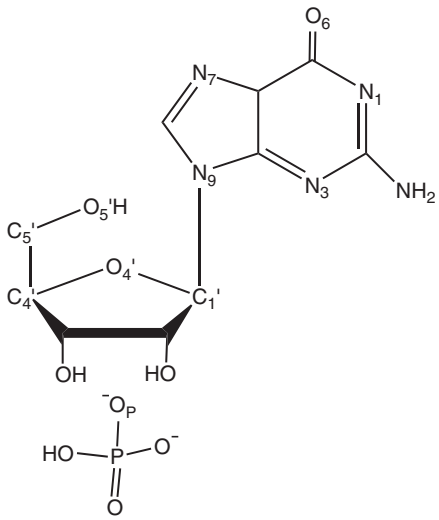

Fig. 14 Atomic nomenclature of the purine nucleoside and phosphate nucleophile.

extensive experimental KIE analysis.[38] We have shown[35] that protein motions in hPNP and its substrates cause the O-5′, O-4′, and O_P oxygens to squeeze together and push electrons towards the purine ring, stabilizing the oxycarbenium character of the TS.

The starting point of our computations was the 2.5 Å resolution structure of hPNP complexed with the TS analog ImmH and phosphate. Following standard computational procedures for enzymes we performed both classical molecular dynamics simulations and hybrid quantum/molecular simulations.[35]

The enzymatic system was divided into two concentric zones. These consisted of the QM region, containing the atoms being treated QM, and the molecular mechanics

(MM) region, in which the atoms were treated with a MM potential. In the case of the simulation of the E···S complex, the QM region contained the substrates guanosine and phosphate, and the conserved catalytic water. For the E···I complex, the QM region contained ImmG, phosphate and a catalytic water. The AM1 semi-empirical Hamiltonian was employed as the QM potential, and the all-atom force field of CHARMM22 was used to describe the atoms in the MM region.

For example, for the E···S complex, we chose the QM/MM model shown schematically in Fig. 15. The residues N243 and E201 are actively involved in catalysis by stabilization of the TS, but since they are not directly involved in the acid catalysis, they were not included in the QM region due to computational limitations of QM/MM methods. However, their contribution to the stabilization of the TS structure is introduced by the QM/MM electrostatic term of the Hamiltonian. We were interested in the last step of the reaction, the cleavage of the N-ribosidic bond and protonation of the purine ring to yield the protonated guanine and phosphorylated α-D-ribose. It is essential that a good reaction coordinate be used in order to represent the enzymatic reaction properly to obtain meaningful PESs. After investigating a variety of reaction coordinates, we found that a suitable reaction coordinate to describe the phosphorolysis reaction was the interatomic C-1′–N-9 and C-1′–O_P distances (Fig. 15).

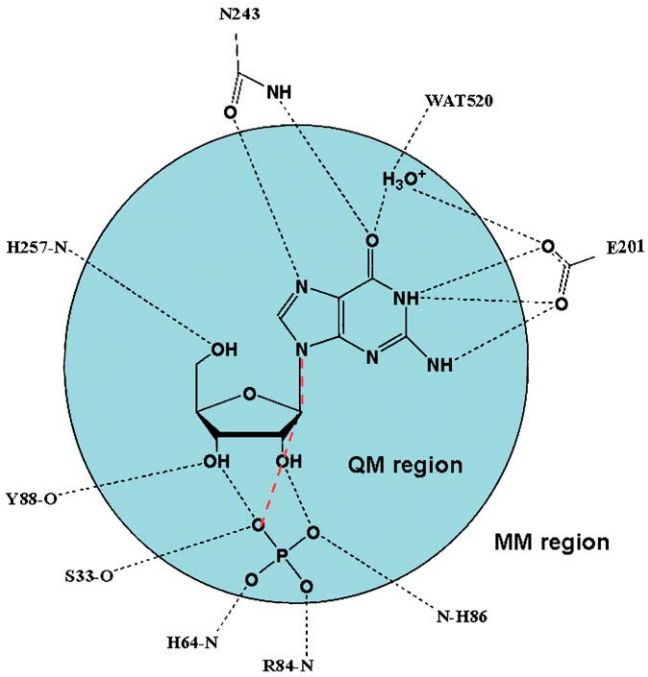

Fig. 15 QM/MM model used to obtain the PES. The reaction coordinate used to produce these surfaces is shown with dashed lines in the figure.

Results: dynamics

During the classical MD simulation we saved the values of the O-5′–O-4′ and O-4′–O_P distances. We found that these distances are quite stable, deviating only up to 0.3 Å from their average value. This implies that the O-5′–C-5′–C-4′–O-4′ dihedral is not rotating freely as it does in aqueous solution, but is restricted by the enzymatic environment. We show the time series for that dihedral in both the enzymatic environment and in aqueous solution in Fig. 16. We note that the mean value in each case is different, and that there is a higher deviation from the mean value in solution, verifying that this dihedral's motion is restricted in the enzyme.

In Table 2 we show the average O-5′–O-4′ and O-4′–O_P distances and standard deviations, for the WT and for several mutants of hPNP complexed with guanosine and phosphate (the maximum fluctuation of the O-5′–O-4′ distance is 0.6 Å, therefore a difference of, for example, 0.2 Å in the average distance is substantial). For each mutation the changes in compression of the oxygen stack were different. For example, we found that for the H257G mutant, the average O-5′–O-4′ distance was considerably higher than that of the WT, meaning that the residue H257 plays a role in keeping the two oxygens compressed.

We now move to a dynamical analysis of the time series for these distances. In the left panel of Fig. 17, we show the Fourier transforms of the O-5′–O-4′ distance autocorrelation function, and of the O-5′–O-4′, O-4′–O_P distance–distance correlation function. The spectra are very similar, indicating that O-5′−O-4′ and O-4′–O_P

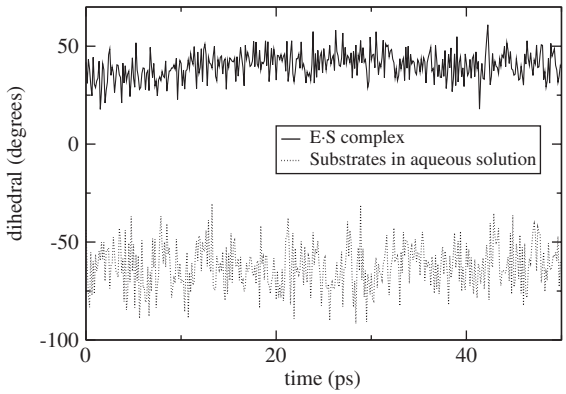

Fig. 16 O-5′−C-5′−C-4′−O-4′ dihedral (degrees) of the ribose ring during the WT classical MD simulation in aqueous solution and in the E···S complex.

Table 2 Average O-5′−O-4′ and O-4′−O_P distances (Å) for the WT and mutated hPNPs

Distance	WT	E201G	V260G	H257G
O-5′−O-4′	2.81 ± 0.08	2.88 ± 0.11	2.78 ± 0.08	3.04 ± 0.08
O-4′−O_P	4.34 ± 0.10	6.58 ± 0.11	3.90 ± 0.11	4.69 ± 0.10

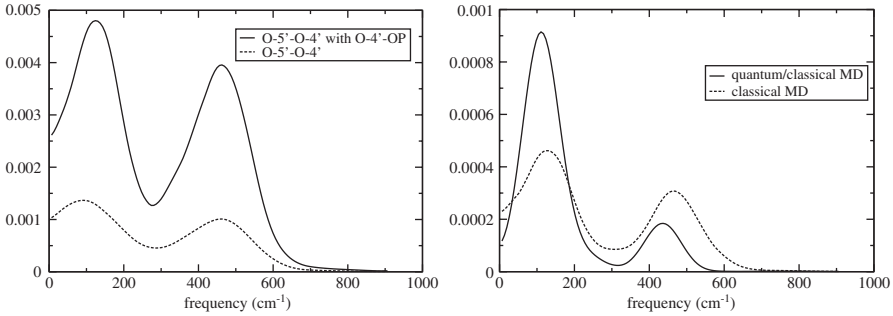

Fig. 17 Left: Comparison of the spectra of the O-5′−O-4′ distance autocorrelation function and the O-5′−O-4′ and O-4′−O_P distance–distance correlation function. Right: The similarity of the spectra of the O-5′−O-4′ distance autocorrelation function for the E···S complex based on the classical and quantum/classical MD simulation, shows an agreement to the classical simulations.

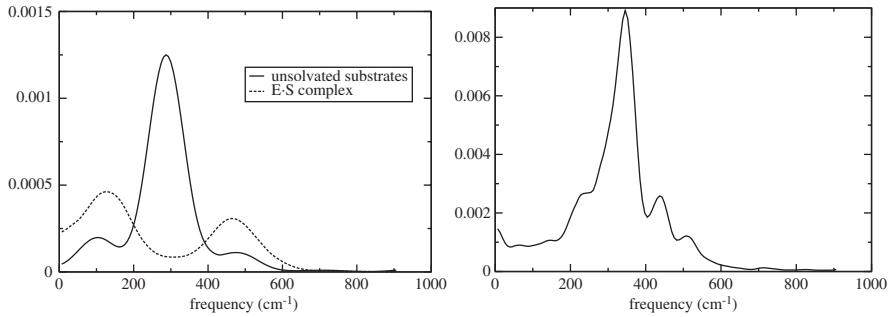

Fig. 18 Left: spectra of the O-5′−O-4′ distance autocorrelation function for hPNP and unsolvated substrates. Note that the natural vibration of the oxygen centers, i.e. 285 cm^{-1}, is altered in the presence of the enzyme. Right: the power spectrum of the H257G mutant of the E···S complex shows a distinct peak at 333 cm^{-1}, very similar to the result for the solvated substrate.

oscillate at the same frequencies: 125 and 465 cm^{-1} in the enzyme environment. In the right panel of Fig. 17, we compare the Fourier transforms of the classical and mixed quantum/classical MD simulations, and find that they are similar.

Next we examine whether these vibrations are unique in the enzymatic environment or they are inherent in the substrates. In the left panel of Fig. 18 we compare the calculation in the enzyme with a simulation of the substrates in aqueous solution, in the absence of hPNP. The spectrum of the O-5′−O-4′ distance autocorrelation function of the classical MD of solvated substrates showed a peak at 330 cm^{-1}, and of the unsolvated substrates at 285 cm^{-1}, i.e. distinct from the peaks in the presence of the enzyme, revealing that hPNP is directly affecting the way in which these oxygens naturally vibrate.

In order to assess the effect of the mutation of nearby residues on the three oxygen stack electronic interaction, we performed classical simulations of mutated hPNPs.

We studied whether changes in protein structure have an influence on the pattern of spectra discussed above for WT hPNP. In particular, we mutated F200G, E201G, H257G, H257A, V260G, and L261G. The spectrum of the O-5′−O-4′ motion for the F200G, E201G, V260G, and L261G mutated hPNPs was very similar to that of the WT. However, as can be seen in the left panel of Fig. 18, the H257G was very similar to that of the unsolvated substrates, giving further evidence that H257 is responsible for modifying the power spectrum in the E···S complex.

Results: energetic barrier

We turn our attention to the coupling of the rate-promoting motion we have been describing, to the reaction coordinate. We used the hybrid QM/MM method described earlier to obtain the PES for the phosphorolysis reaction for various E···S complexes, for a range of O-5′−O-4′ and O-4′–O_P interatomic distances. We have presented elsewhere[35] detailed results for the activation energies (kcal mol^{-1}) as a function of oxygen interatomic distances of the E···S complex and the E···TS complex. This is collection of structures covered a wide range of geometries, with O-5′−O-4′ ranging from 2.78 to 3.64 Å and O-4′−O_P ranging from 3.35 to 4.30 Å. Thus, these structures were sufficient to represent the different E···S geometries that can be found in the PES.

To quantify this correlation between lower activation energy and compression of the oxygen distances, we fitted the results for the activation energies to a function that had up to quadratic terms, as seen in Fig. 19 (a parabolic fitting had smaller χ^2 than the also plausible linear fitting). The parabolic dependence of the activation on energy on distance may have a simple interpretation: as the distances between the oxygen increases the energy rises as carbocation stabilization is lost, but at shorter distances, as we approach distances equal to twice the van der Waals radius, electron–electron repulsion will compete with the stabilization and there is a point where there are no further reduction of the activation energy. From the values of the

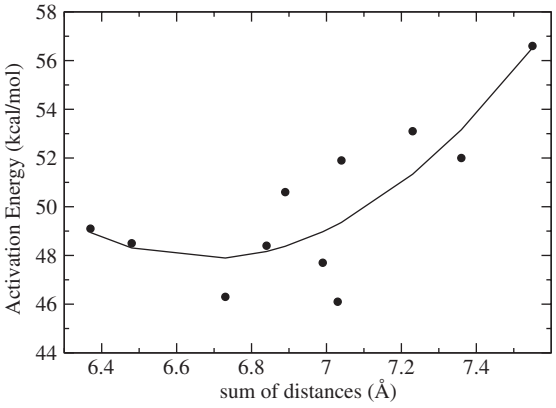

Fig. 19 Activation energies (kcal mol^{-1}) versus the sum of O-5′−O-4′ and O-4′−O_P distances (Å) for the E···S complex.

curvature of the parabolic fittings we can obtain a spring constant, and by using a mass equal to the reduced mass of O−O, we can obtain an effective harmonic frequency associated with the parabolic fittings shown in Fig. 19. This effective frequency was equal to 180 cm^{-1}, in surprising agreement with the position of the peaks of the spectra of the oxygen motions.

In conclusion, the protein motion that compresses the oxygen stack, is one of the factors which makes the reaction possible, leading up to a 20% decrease in barrier height.

Results: charge fluctuations

In addition, we have studied[39] charge fluctuations in the ribose and purine groups of the enzyme-bound ribonucleoside. If these fluctuations are resonant with the oxygen-stack promoting motion, it would further validate the hypothesis that the promoting motion pushes electrons across the ribosidic bond.

Partial charge calculations were employed to follow changes in electron density in the substrate guanosine, using Gaussian98 at the B3LYP level of theory with the 6-31G(d,p) basis set. Partial charges for each atom were calculated using the CHELPG algorithm. Charges for the ribose and purine ring components of the ribonucleoside substrate were calculated by adding up the partial charges of the constituent atoms of each.

We generated time series for the partial charges, and as usual took the Fourier transforms, shown in the left panel of Fig. 20. The two spectra nearly overlap, demonstrating that the charge fluctuations in the ribose are resonant with the charge fluctuations in the purine ring, with a dominant peak at 450 cm^{-1} and a smaller peak at 160 cm^{-1}. Recalling that the spectrum of the oxygen-stack compression motion had peaks 125 and 465 cm^{-1}, we note that the partial charge spectra are clearly resonant with those of the oxygen motion.

In addition, we probed fluctuations across the N9-C1′ ribosidic bond, since this is the bond that is broken in this reaction. The spectra are shown in the right panel of Fig. 20. They are resonant with each other and with the spectra for ribose and purine

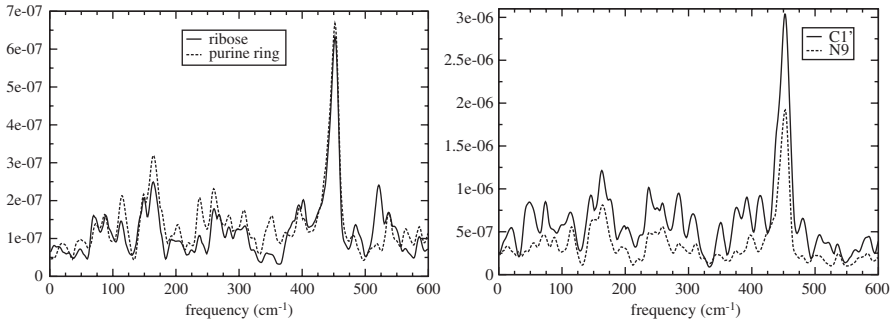

Fig. 20 Left: Power spectra for ribose and purine ring partial charge time series. Right: Power spectra for atoms C1′ and N9 partial charge time series.

ring, and also feature peaks at 160 and 450 cm^{-1}, i.e. they contain the signature peaks of the oxygen-stack compression spectra. These results confirm that the oxygen-stack compression is the cause of these partial charge fluctuations.

4 Description in atomic detail of correlated protein motions

In this section we will examine other kinds of correlated protein motions (with ps- or ns-timescales) and methods that can identify them. The rate-promoting vibrations we examined in the previous section are just one example of correlated protein motions. Because promoting vibrations involve residues in the immediate vicinity of donor and acceptor, it was relatively easy to identify them. In the more general case of extended correlated motions, it would be a challenge to identify residues that take part in them. In this section we describe two methods that have been successfully used for identifying atomic motions of interest, the TPS and the ED method. We will apply them to two enzymes we already studied in the previous section, LDH and PNP.

TRANSITION PATH SAMPLING

TPS was originally developed for studying rare reactive events. The most difficult problem in studying reactive events, which is hidden by the wide acceptance of TST, is to define an appropriate reaction coordinate and find the location of the TS. In addition, if one tries to simulate rare reaction events using a molecular dynamics simulation, most of the trajectories that start from the reactants will not cross to the products, and as a consequence the calculation becomes computationally not feasible, because the time step that must be used in the MD simulation is much smaller than the timescale of interest.

TPS[40,41] addresses these problems by performing a Monte Carlo search in the trajectory space. It can simulate rare events without the knowledge of a reaction coordinate or the TS. TPS needs another algorithm (molecular dynamics) to generate trajectories, while TPS itself is an algorithm for searching the trajectory space. The essence of TPS is that the chaotic nature of classical multi-dimensional systems guarantees a fast Monte Carlo search of the trajectory space.

A brief description of the algorithm follows. Let us assume a transition between R and P (i.e. reactants and products). Since R and P are long-lived states, they can accommodate equilibrium fluctuations, and can be characterized by a variable, called ''order parameter'', which can be used as a criterion for deciding whether the system is localized in R or P. Let us further assume that we somehow know one reactive trajectory that starts from R and ends in P. In the TPS algorithm we randomly select a time slice along this reactive trajectory, we perturb slightly all momenta at that time slice, and starting from that time slice and using the new momenta, we propagate (''shoot'') forward and backward in time, examining whether the new trajectory is reactive or not. In the usual Monte Carlo fashion, the

new trajectory is accepted or not, according to some probability distribution. Because of the ergodicity of classical dynamics, new trajectories are guaranteed to quickly deviate from old ones, leading to a fast sampling of the trajectory space.

An ensemble of reactive trajectories in LDH

We have applied[42] the TPS algorithm to LDH, i.e. to the enzyme in which we identified a rate-promoting vibration, as discussed in the previous section. This work, along with a paper by Schlick[43] were the first that applied the TPS algorithm to a realistic protein. By finding common features in all the harvested trajectories, one can get insights for defining an appropriate reaction coordinate, and for identifying experimental targets for future studies. Recall that LDH catalyzes the interconversion of the hydroxy-acid lactate and the keto-acid pyruvate with the coenzyme NAD. The reaction involves a double transfer: a proton transfer between the active site histidine and the C2 substrate oxygen and a hydride transfer between NC4 of the coenzyme and C2 of the substrate (see Fig. 6). In fact, there was a controversy regarding whether the transfer steps are concerted or sequential, that our study was able to resolve. Since the reaction involves bond cleavage we must use a quantum potential for describing the reactive potential surface. The details of our choice for the quantum description are explained elsewhere.[42] Below we briefly define the variables and concepts that our TPS simulation used.

Definition of the order parameter. The first step in the TPS algorithm is to define an "order parameter," i.e. a variable that describes whether the system is in the reactants, products or in an intermediate region, as shown schematically in Fig. 21. The pyruvate and lactate regions were defined by the values of the appropriate bond lengths.[42]

Decorrelation of trajectories. The goal of TPS is to generate reactive trajectories that span the whole trajectory space. We must ensure that the ensemble of trajectories

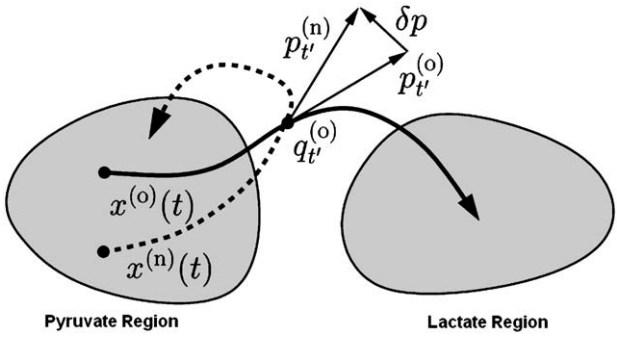

Fig. 21 Schematic representation of the TPS algorithm. The shaded regions are identified by the order parameter as reactants and products. The solid line is a reactive trajectory. A "shooting" move is shown: a time slice was chosen along the reactive trajectory, momenta were perturbed, and then the system was propagated forward and backward in time, resulting in a non-reactive trajectory shown with the dashed line.

we generated do not lie "near each other" in the trajectory space. The variable that monitors this is an autocorrelation function of the appropriate variable. This autocorrelation function is shown in Fig. 22. As can be seen, the trajectories became uncorrelated after about 30 successful iterations of the algorithm.

Results: atomic description of relevant catalytic motions

A compression of residues facilitates catalysis. We have identified a compression–relaxation sequence of residue motions that facilitates catalysis. These are the residues 31 and 65, located behind the cofactor and transferring hydride, and 106 and 195, located on the acceptor side behind the substrate, as seen in Fig. 23.

In Fig. 24 we plot various distances that reveal a compression of several residues that occurs close to the reaction event. These are taken from a trajectory with a concerted hydride and proton transfer. All graphs are plotted in the pyruvate to lactate reaction direction.

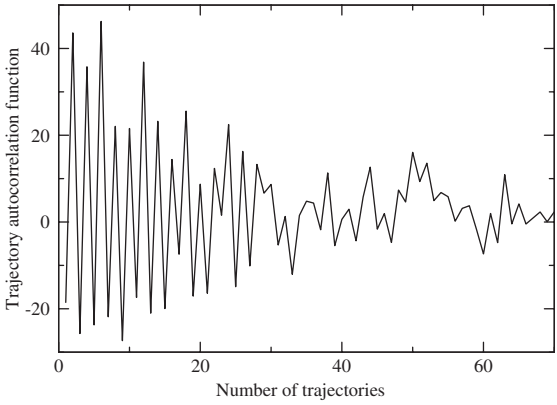

Fig. 22 Autocorrelation function for trajectories, plotted versus the number of reactive trajectories that have been generated after a given trajectory. The figure shows a decorrelation of reactive trajectories generated by TPS.

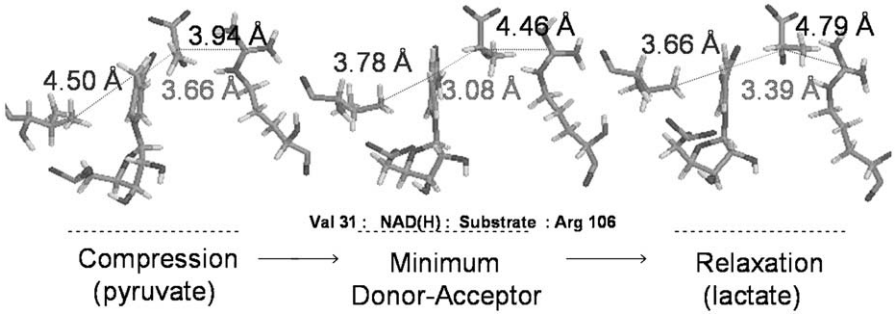

Fig. 23 Three snapshots of the compression of residues that leads to a reactive event.

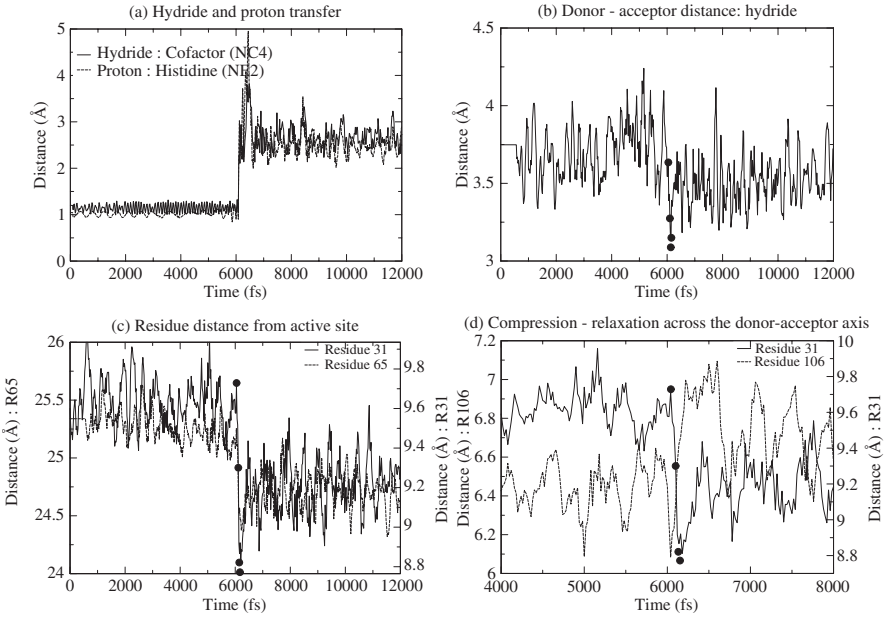

Fig. 24 Comparison of motions of donor–acceptor and surrounding residues reveals a compression-relaxation motion. See text for description of the panels.

The panels of Fig. 24 describe the following:

(a) Plots the distance of the hydride from the cofactor reactive carbon and the distance of the proton from the histidine reactive nitrogen. At time = 6100 fs the atoms begin to transfer to the substrate.
(b) Plots the donor–acceptor distance for the hydride. The minimum distance of the donor–acceptor distance, and the distance of residues Valine 31 and Arginine 106 from the active site are marked by solid circles. The hydride donor–acceptor distance reaches its minimum at time = 6132 fs. A similar plot can be drawn for the proton transfer.
(c) Plots the distance of residues 31 and 65 from the active site. They are located behind the cofactor and transferring hydride.
(d) Plots the distance of residues 31 and 106 from 4 to 8 ps. Residue 106, responsible for polarization of the substrate carbonyl bond via hydrogen bonding, initially compresses towards the active site reaching a minimum distance at 6043 fs (marked by solid circle). By 6153 fs (marked by solid circle) residue 31 has reached its minimal compression towards the active site.

The compression towards the active site revealed in Fig. 24 is what causes the donor–acceptor distances for the hydride and proton transfers to reach their minimum. When they reach their minimum, interactions across the donor, transferring atom, and acceptor are initiated. The events that occur next are critical for

completion of the reaction. The continued compression of the donor side residues towards the active site, are not only involved with bringing the donor–acceptor distance closer together, but also with shifting the entire enzyme. While the donor–acceptor distances increase again, it will be the motion of the surrounding residues that ultimately determine whether the atoms transfer. In particular, the compression of the donor side residues cause the acceptor side residues to relax away and the reaction to complete.

Concerted vs. *stepwise transfers.* We now address the question of whether the hydride and proton transfers are concerted, or whether the hydride transfer precedes the proton in the pyruvate to lactate reaction direction. Our TPS study showed that both mechanisms are possible. In Fig. 25 we show the distribution of the time lag between the hydride and proton transfer for all reactive trajectories. We note that both concerted and sequential transfers are possible, and that 74% of the trajectories have a time lag greater than 10 fs, indicating that the majority of reactive trajectories have sequential transfer steps.

Figure 26 compares two trajectories: one for concerted hydride and proton transfers with one for sequential transfers, as seen in panel (a). In panel (b) we can see that the donor–acceptor distance for the stepwise trajectory reaches its minimum first, due to the earlier combined compression of residues 31 and 106. Why does the proton take longer to transfer? If we look at panel (c) at the distance of residue 31 after it reaches its minimum, it jumps back away from the active site causing a delay in the relaxation of residue 106, as seen in panel (d), and a delay in the increase of the proton donor–acceptor distance. Soon after, panel (d) shows that residue 106 relaxes away from the active site while the donor–acceptor distance increases, completing the reaction.

Perturbation of the donor–acceptor axis and compression reaction coordinate. We will now demonstrate the effects of disrupting the compression–relaxation of the donor–acceptor axis residues close to the reactive event. For the perturbation we used the coordinates and velocities of a time slice 160 fs away from the reactive event

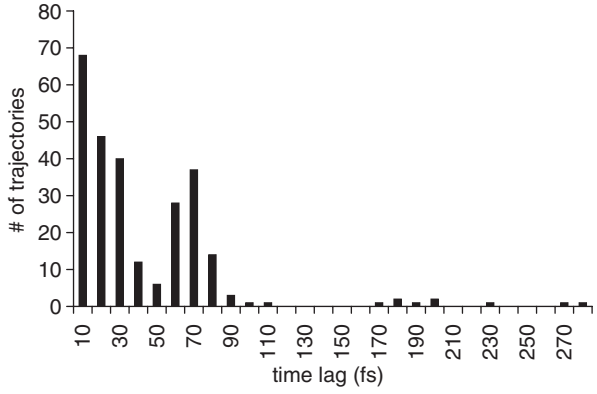

Fig. 25 Distribution of the time lag between the hydride and proton transfer.

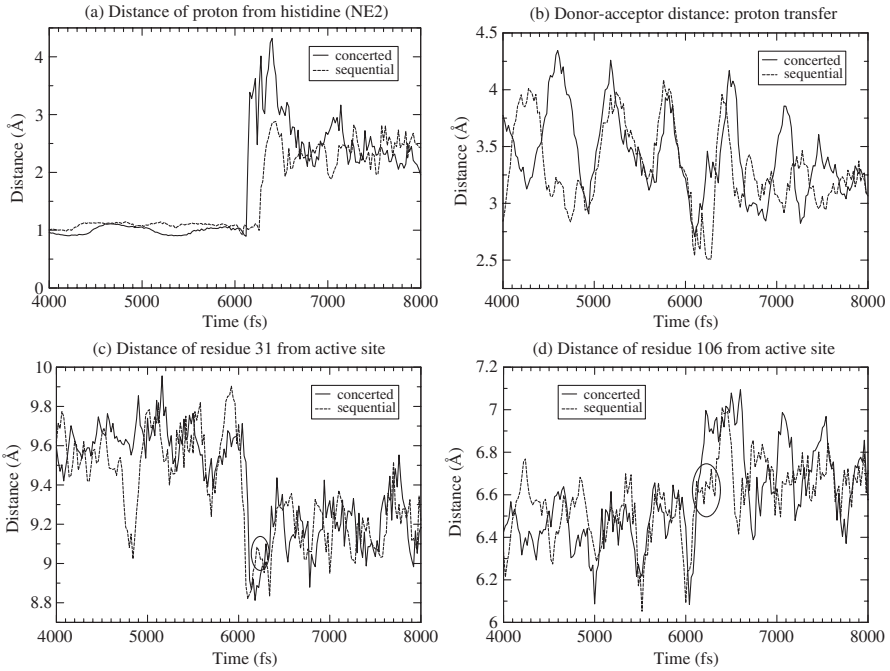

Fig. 26 Comparison of motions of donor–acceptor and surrounding residues reveals when the double transfer is sequential and when it is concerted. See text for description of the panels. The circles in the residue trajectories in the lower panels, highlight the moment the transfer event happens.

of a concerted transfer trajectory. The momenta of residue 31 were perturbed along the residue-active site axis, 1 Å away from the active site.

Figure 27 contrasts the original reactive trajectory to the perturbed one, shown in panel (a). Panel (b) shows that after a delay, the hydride donor–acceptor distance begins to deviate from the original reactive trajectory, unable to reach its minimum without the full compression of the donor side residues, i.e. 65 shown in panel (c) and the perturbed 31 (not shown). Additionally, the absence of the compression prevents the relaxation of the acceptor side residues, for example, of 106 shown in panel (d). The donor–acceptor distance comes closer, since at that time residue 106 is still compressing and the perturbed residue 31 has a weaker, delayed compression. Due to this, the hydride starts to transfer. However, since the compression–relaxation transition does not occur, the reaction is not completed.

ESSENTIAL DYNAMICS

ED (or principal component analysis), is a method commonly used for identifying large scale motions in proteins, e.g., protein folding or substrate binding. A summary of the method follows. Let $\vec{R} = (R_{1,x}, R_{1,y}, \ldots, R_{N,z})$ be the positions of the

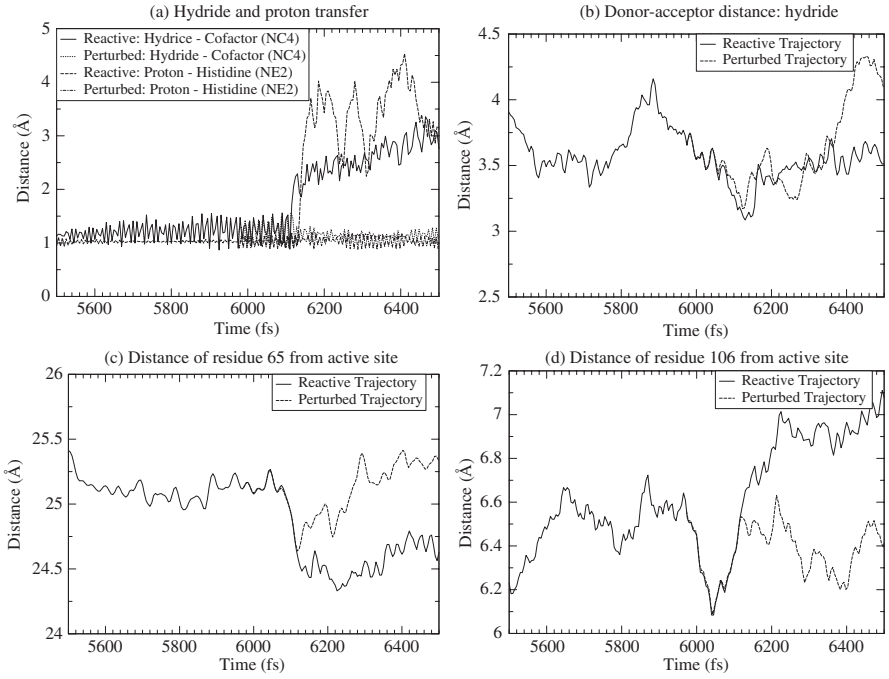

Fig. 27 How perturbation of a residue makes a trajectory nonreactive.

protein atoms. One assumes that the solution of Newton's equation of motion can be written in the following approximate form:

$$\vec{R}(t) \simeq \sum_{m=1}^{3N} a_m(t)\vec{\eta}_m \tag{13}$$

One then tries to find the coefficients and basis functions a_m, $\vec{\eta}_m$ that provide the best approximation (in a least-squares sense) of $\vec{R}(t)$ to the exact solution of Newton's equation of motion. This condition is satisfied by choosing the spatial basis set $\vec{\eta}_m$ to be the eigenfunctions of the covariance matrix

$$C_{ij} = \langle (R_i - \langle R_i \rangle)\ (R_j - \langle R_j \rangle) \rangle \tag{14}$$

where $\langle \cdots \rangle$ denotes time average over the entire MD trajectory. When a few eigenvalues of the covariance matrix are much bigger than the rest, the corresponding eigenvectors dominate the dynamics. When sorted by their eigenvalues, the ED modes are sorted according to their contribution to the total mean-square fluctuation. These few dominant modes are designed to provide a good fitting to the trajectory. In a protein, there may be concerted motions of groups of atoms (e.g., loop motions) which provide a significant number of covariant matrix elements C_{ij}, and therefore dominate the principal eigenvectors of the covariant matrix. In this case, ED separates the conformational space into an *essential subspace* containing

only a few dominant collective modes, and a remaining space which contains random atomic fluctuations.[44,45]

In Section 3, we found that protein motions in hPNP accelerate the chemical step. Now we will present an application of the ED method, used to identify protein motions that increase turnover by creating substrate binding affinity. In particular, we have studied[46] the conformational change in the 241–265 loop, and identify variations in its orientation, which is crucial in determining the substrate accessibility to the active site.

Substrate binding in PNP

As a reference for the discussion below, in Fig. 28 show a plot of the active site of PNP. We will present the experimental reasons that led us to focus our attention on the 241–265 loop. If one assumes that each X-ray crystal structure represents a possible conformation in solution, considering several crystals and their relative atomic B-factors presents an experimental view of the conformational flexibility of the protein at an atomic resolution. The relative B-factors of the C_α atoms of apo hPNP, and hPNP complexed with the TS analog Immucillin-H show that the loop residue E250 is the centroid of highly mobile region, while residues G63 and E183 are centroids of more localized mobile regions. Additionally, from the structural differences between the crystal pairs of apo hPNP–hPNP · guanine (formation of the Michaelis complex), and apo hPNP–hPNP · ImmH (formation of the TS analog

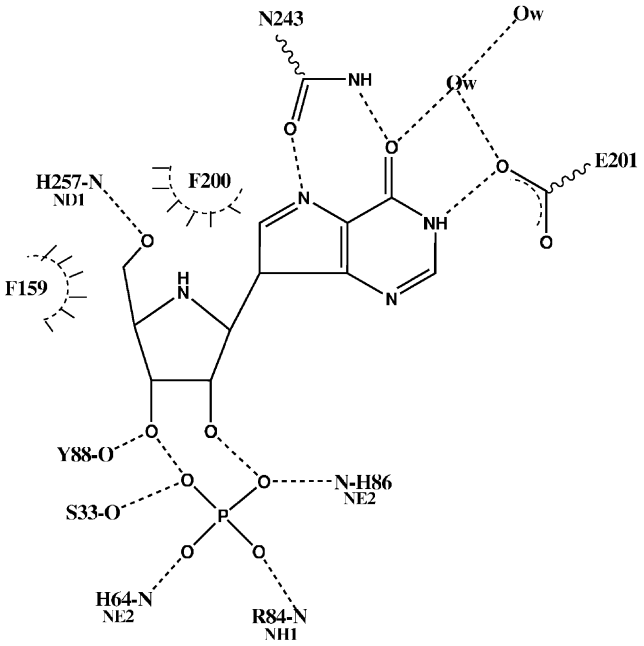

Fig. 28 Active site of hPNP with the TS analog ImmH and the phosphate nucleophile.

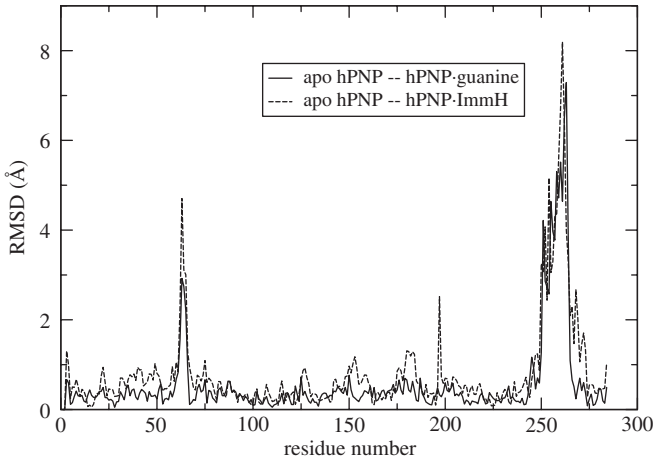

Fig. 29 Detailed root-mean-square displacement (RMSD) Å, per residue, of the geometric difference between the apo hPNP–hPNP···guanine (black), and apo hPNP–hPNP···ImmH (red) crystal structure pairs. The different RMSDs of both pairs of crystals, provide evidence that the TS analog is bound to hPNP much tighter than the substrate analog.

complex), it has been questioned whether correlated motions exist to allow the necessary conformational change of the 241–265 loop for substrate binding and turnover. It is accepted that reorganization of this flexible loop, which makes up part of the active site, is essential for allowing different substrates/inhibitors to enter the active site.[47] The loop displacement is conjectured to push the substrates towards the active center favoring more reactive configurations. In Fig. 29 we show the root-mean-square deviation of the C_α geometric difference between the apo hPNP–hPNP · guanine, and the apo hPNP–hPNP · ImmH pairs. The same geometrical difference trend appears for any combination of two apo hPNP structures with six substrate/TS analogs, confirming the substantial loop rearrangement upon substrate analog/TS analog binding.

We used principal component analysis to identify correlated motions in different forms of hPNP, namely, its apo and complexed forms, and assess whether they facilitate the 241–265 loop rearrangement prior to the subsequent phosphorolysis reaction. We compared the principal components for the apo and complexed hPNP simulations, and examined the different correlated motions for each form of the enzyme, comparing directly to the crystallographic B-factors. Finally, via experimental site-directed mutagenesis, several residues implicated in the correlated motion were mutated, and the kinetic constants k_{cat} and K_M (fingerprints of catalytic efficiency), were measured to weigh the impact of these residues in the phosphorolytic efficiency.

Results: mobile residues in the active site

We performed molecular dynamics simulations for the APO, ES, and EI-APO models. The objective of the EI-APO simulation is to capture putative concerted

motions in hPNP during the loop conformational change taking place upon going from an "E···I geometry" towards an "apo hPNP geometry", which reversibly translates into existing motions in hPNP during substrate binding and TS formation.

For the APO simulations, from the first few eigenvectors, we observed prominent correlated motions in the neighboring residues F159 and P150, adjacent to the purine substrate. Moreover, it is not surprising that G63 and E183 also show strong motions, granted their large B-factors. Additionally, to a lesser extent, the active site residues S33, H84, H86, and F200 display high mobilities. These concerted motions were observed both in a 1-ns vacuum simulation and in a solvated simulation. Overall, the most significant feature of the APO simulations is the correlated motions found in the spatially neighboring F159 and P150 residues.

In the case of the ES simulations, we found that there exist correlated motions around the F200 and E250 residues, the latter being the epicenter of the loop dynamics. To a lesser extent, residues S33 and P150 also showed concerted behavior. It is not surprising that residues H183 and N121 are involved in dynamics granted their high mobilities implied by its high B-factors. We found differences in the concerted motions between the APO and ES models of hPNP: the APO model presents concerted dynamics around the F159 and P150 residues, while for the simulation of the Michaelis complex (ES), concerted motions are mostly found around the active site residue F200 and loop residue E250. This in turn suggested that dynamics in the apo enzyme is concentrated around the residues F159 and P150, though once the substrate enters the active site, these become dormant, and active site residue F200 and loop residue E250 (embracing the purine substrate), come into action.

For the EI-APO simulations, the correlated motions were again found around the neighboring residues F159 and P150, as well as around F200, E250, and S33. This behavior resembles a combination of the APO and ES models. These results suggest these residues are coordinated in the enzyme so that it can successfully accommodate the substrate in the active site and achieve TS formation.

We observe that the three computational models (APO, ES, and EI-APO) present distinct characteristics in their correlated motions. Some similarities are found, for example, in the high mobilities of G63 and E183, in agreement with their high crystallographic B-factors. However, other residues which had not been seen as highly mobile based on their B-factors (namely F159, P150, H230, and F200), have been found to have correlated dynamics. This implies that observation of crystallographic structures alone may not be enough to infer important dynamic behavior in proteins, and that additional MD studies are needed to identify correlated dynamical modes. Finally, we have been able to successfully recognize correlated motions on the 241–265 loop, particularly present in the simulations of the Michaelis complex, and in a lesser extent in the EI-APO simulations. This suggests that the loop may have a favorable role in capturing the substrate *as well as* in the chemical step. Fig. 30 highlights the residues that were identified by the principal component analysis: note the presence of residues F159 and F200, whose involvement in binding was not obvious from B-factors alone.

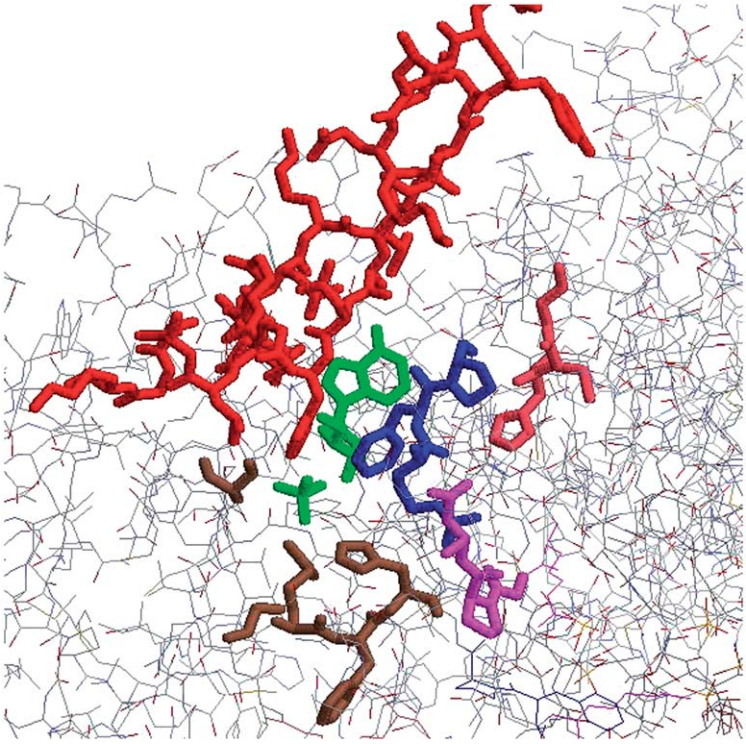

Fig. 30 hPNP with marked residues that show correlated motions: S33 (brown), H64 (brown), H86 (brown), P150 (magenta), F159 (blue), H230 (pink) and 241–265 loop (red).

Table 3 Experimental kinetic parameters for different mutants which show correlated motions in hPNP

Mutant	WT	F200G	F159G	H230G	N121G
K_M (μM)	76.8	1580	2260	145	64.9
k_{cat} (s^{-1})	32.4	3.8	2.3	2.6	33.0
k_{cat}/K_M ($10^5 s^{-1} M^{-1}$)	4.2	0.024	0.01	0.1	5.0

Experimental site-directed mutagenesis

The X-ray crystal structures indicate that all direct contacts between hPNP and the substrates guanosine and phosphate are through amino acid side chains. Our site-directed mutagenesis strategy replaced each residue with glycine in order to assess the contribution of the residue's side chain to substrate binding and/or catalysis. Results from the kinetic studies (Table 3) indicate that substrate binding affinity was most sensitive to the F159G and F200G mutations. In the phosphorolytic direction, an increase of ca. 400-fold in K_M (and modest change in k_{cat}) is observed for the

F159G mutant. Loss of the herringbone-type interaction between the β-face of the ribose and the hydrophobic surface created by this residue is important for substrate binding, thus essential for hPNP activity. Additionally, the H230G mutant showed modest changes in both K_M and k_{cat}, unexpectedly given its long distance from the active site. Finally, the N121G mutant, did not show any relevant change in catalytic efficiency, regardless of proving substantial concerted dynamic behavior during the Michaelis complex simulation.

5 Conformational fluctuations

The motions we examined in the previous sections had ps- or ns-timescales, therefore they were accessible with ordinary molecular dynamic simulations. However, it is well known that conformation fluctuations are sometimes coupled to the reaction coordinate. From a computational point of view, such effects are practically impossible to study directly. To understand the reason, let us recall the shape of the energy landscape of a protein (a description of the modern energy landscape view of biomolecules can be found in the recent textbook of Wales[48]). It consists of a very large number of local minima (conformations), separated by barriers whose height ranges from moderate to high (left panel of Fig. 32). The time evolution of the system in this landscape consists of relatively long periods of oscillations in the local minima, separated by hops between conformations (Fig. 31).

The rate-promoting vibrations we examined earlier are fluctuations within a single conformation. The problem we want to address is, whether there are some conformations which favor catalysis, for example, because in them the average

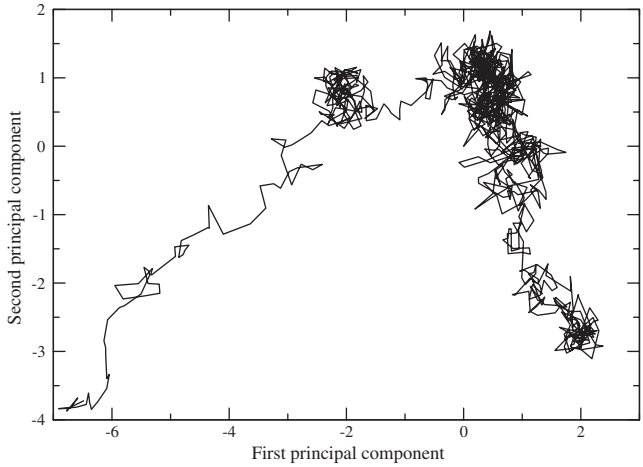

Fig. 31 A 100 ps MD trajectory of PNP, plotted in the plane of the two largest principal components. The two darker regions, where the trajectory fluctuates locally, are two conformations. If we are interested in sampling many conformations, the long time the system spends in local fluctuations in a given conformation, is a waste of resources.

donor–acceptor distance is shorter. To study this problem, we would like to have our molecular dynamics trajectory visit as many conformations as possible. However, the system spends most of its time in fluctuations within a conformation local minimum, and only rarely hops to a different conformation minimum. Therefore, MD simulations may be useful for studying other problems, but they are not a practical tool for searching the conformation space. However, there are other techniques that may be useful. In the rest of this section, we will first briefly review recent work on conformation fluctuations that enhance tunneling in DHFR; we will close with a computational scheme we propose for studying this class of problems.

DIHYDROFOLATE REDUCTASE

DHFR catalyzes the reduction of 7,8-dihydrofolate (DHF) to 5,6,7,8-tetrahydrofolate (THF) through the oxidation of the coenzyme NADPH. DHFR has been studied extensively due to its role in maintaining intracellular pools of THF, which is an important component of several metabolic pathways. The reaction catalyzed by DHFR is thought to involve transfer of a proton followed by a hydride to DHF.

It has been shown through nuclear magnetic resonance (NMR) studies[49,50] that the catalytic cycle involves conformational changes of the M20 loop. This mobile loop is close to the active site and is assumed to play an important role in binding of the cofactor and substrate to DHFR. Also, mutational studies have identified distal residues[51] that affect catalysis. These studies strongly suggest that conformational changes are related to catalysis.

In addition, classical MD simulations[52] have identified correlated domain motions in the reactant DHF complex but not in product complexes, indicating they are related to catalysis. These domains are in the same regions highlighted by the NMR studies.

We will only present a very brief review of theoretical studies on this system and refer to reader to publications of the groups that have studied it.[53–55] Since the catalytic step involves a hydride transfer, a major difficulty is how to treat quantum effects. These works followed different paths.

Truhlar, Gao, and Garcia-Viloca[53] use a geometric reaction coordinate, the difference between acceptor-hydrogen and donor-hydrogen distances. The system is divided into a small primary zone at the active site, and a secondary zone. The results are averaged over several secondary zone configurations. Electronic quantum effects are included with a semi-empirical QM/MM potential that is augmented with a valence bond term, parametrized to fit the experimental free energies. The free energy profile is calculated with an umbrella sampling along the above reaction coordinate. Brooks and Thorpe,[55] also used a semi-empirical QM/MM potential. Only structures from the reactant side were used, and those with donor–acceptor distance smaller than 2.5 Å were used in the QM/MM calculations. For each of these configurations, a QM/MM optimization was performed for a series of fixed hydrogen-acceptor distances. Only atoms within 10 Å from the transferring hydrogen, were allowed to move. Hammes-Schiffer[54] used as a reaction coordinate the

energy difference between reactant and product diabatic states, averaged over the ground state of the vibrational wavefunction of the transferred hydrogen. Electronic quantum effects are included through an EVB potential, parametrized to fit experimental free energies of activation. The free energy profile is generated by sampling over the entire range of that collective reaction coordinate. Once the energy profile is found, they used TST to calculate the reaction rate. Note that in real time the protein dynamics may need e.g., 1 ms to move along the reaction pathway of the reactive coordinate used in these works, so the generation of the energy profile allowed them to describe an event that requires time longer than any feasible molecular dynamics trajectory.

There is a long-standing debate whether approaches that calculate the free energy barrier using a geometric reaction coordinate can sufficiently sample protein conformations and whether methods that use the energy gap as the reaction coordinate are superior. In the present system, it has been shown[53] that the two methods generate equivalent results. In fact both methods produced similar free energy profiles.

Coming back to the work of Truhlar and coworkers,[53] and Hammes-Schiffer,[54] after they generated a free energy profile along their choice of reaction coordinate, the next step was to examine the ensemble of conformations used to generate a point on the energy profile, and generate statistics and comparisons for various geometric characteristics along the energy profile. They both found a correlation of structural changes as the system approaches the TS, which show a tightening of some hydrogen bonds and of the donor–acceptor distance.

We should emphasize that the structural changes found in those works were not motions in real time. Even if it is true that conformational changes affect catalysis, it is not clear if the enhancement of the rate is due to a coupling to the motion of the hydride, or due to different thermodynamic averages, for example, because the transfer distance became shorter. The fact that in the calculations of Thorpe and Brooks,[55] most of the enzyme was held fixed, is an indication that the effect is structural rather than dynamical. In this last work, a wide range of barrier heights was found, which led to questioning whether approaches that are based on mean field potentials can capture the range of protein conformations that led to that range of energy profiles. Various opinions on these questions can be found in the original papers.

To avoid semantic confusion we should clarify that Hammes-Schiffer uses the term ''promoting motion'' to describe an ensemble of conformations that, for example, have a shorter donor–acceptor distance. The ''motion'' in her case is the implied intra-conversion among these conformations. The ''promoting vibration'' we described earlier refers to fluctuations within a single conformation. The two effects are different and occur in separate timescales.

The identification of the way conformation fluctuations in DHFR affect particular distances at the TS, is a significant achievement. There are some details that maybe improved, but we should emphasize that the point we will make this is not a criticism of those works, but rather of the limitations of simulations that are currently possible. These simulations identified distances that change in the TS, but they can't distinguish if they must change in a particular order (e.g., as found with TPS in

LDH). Also, they generated an ensemble of conformations near the TS, but since they have no notion of time, they cannot give details of how the system samples the conformations. In the next section we will propose a scheme that may provide a starting point for a discussion that will address these questions.

A PROPOSED SCHEME FOR SEARCHING THE CONFORMATION SPACE

It would be desirable to have a more systematic method for searching for conformation fluctuations that are coupled to the reaction coordinate. In the rest of this section, we will propose a computational scheme that tries to address this problem. This scheme is unpublished work. It consists of three elements:

(1) The topological structure of the conformation space.
(2) A computational scheme (kinetic Monte Carlo) for describing dynamics in the conformation space.
(3) An algorithm (string method) for calculating energetic barriers between conformations.

The topological structure of conformation space

It is well known that conformations that have similar geometries have similar energies, but the opposite is not true, conformations that are unrelated may accidentally have similar energies. In addition, long MD runs that visit several conformations, sample geometrically similar conformations. The above observations suggest that conformations populate their configuration space in some structured way. The structure of the conformation space was studied by Berry and Kunz,[56] Becker and Karplus,[57] and Becker.[58] They termed it "topological" structure because the key in the description is neither geometrical similarity nor energy, but rather a connectivity concept, i.e. given a conformation, which other conformations are accessible to it, and how many saddle points are crossed by the pathway that connects them.

Their main conclusions are summarized in Fig. 32. In the left panel, we see a free energy profile. At the energy level marked as "–1," the conformations that are

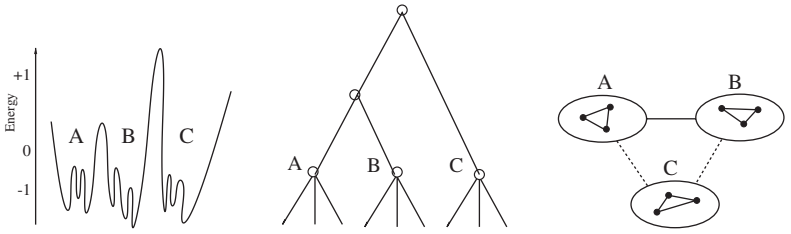

Fig. 32 The topological structure of conformation space. Left: the energetic profile. Middle: the connectivity graph. Right: the clustering of conformations into basins.

clustered under A are accessible to each other by crossing a single saddle point. The same is true for the conformations under B and under C. At the energy level marked as "0," one can go from A to B by crossing one saddle point. A group of conformations like the ones clustered A or B, is called a "basin." There are three basins visible in the diagram: A, B, and C. At the energy level marked as "+1," one can go from A to B, but to go from basin B to basin C, one has to cross another saddle point. We say that A and B belong to the same "superbasin," but C belongs to a different superbasin. The above structure is depicted in the "connectivity graph" shown in the middle panel, that shows how the various conformations and saddle points are connected. An easier to understand schematic picture is shown in the right panel: the solid dots are conformations and they are grouped in their basins. Solid lines represent transitions through a single saddle point, possible either between conformations in the same basin, or between basin A and B. Basins A and B are connected to basin C by a dashed line, in order to signify that they belong to a different superbasin and one has to cross 2 saddle points instead of 1, for a transition to C.

Becker and Karplus[57] classified in this manner all the minima and saddle points of a small peptide, met-enkephalin, and found the topological structure shown in Fig. 32. It should be noted that Frauenfelder and Wolynes,[59] from their analysis of experimental data, had predicted earlier that the energy landscape of proteins would have the topological structure shown in Fig. 32. In addition, they found that conformations that belong to the same basin are separated by low barriers of a few $k_{\mathrm{B}}T$. Basins are separated from each other by a high barrier. That is, a basin consists of geometrically similar conformations that are separated by low barriers, and is separated from another basin by a high barrier. This structure explains the observations mentioned at the beginning of this section. A constellation of basins can have various structures,[48] one of which corresponds to the well-known "funnel" landscape of proteins.[60]

From the dynamic point of view, the system frequently hops among conformations in the same basin, and it rarely hops to a different basin.

Basin hopping in the conformation space

In the beginning of this section we mentioned that even an MD trajectory that is several ns long, can sample only a few conformations. With the help of the basin picture, we now realize that the conformations sampled by MD run, all lie in the same basin. The reason is that, by definition, only these conformations are separated by low barriers, which can be overcome in the timescale accessible to MD runs. A simple corollary is that the MD run would be able to sample more conformations, if it could somehow be able to hop to a neighboring basin.

Now, imagine that we have two conformations that belong to different basins, and that we know the height of the barrier that separates these two conformations. Things become more interesting if we recall that the conformations that belong to the same basin are geometrically and energetically similar to each other, therefore we can pick any conformation as representative of that basin. We can make the

following assumption, first suggested by Berry[56] (he recently applied it to a study of small peptide[61]): each basin is represented by any conformation in it, and hopping among basins can be approximated by transitions among their representative conformations. In our case, one can use TST to describe these transitions.

This is a summary of the proposed scheme for studying dynamics in conformation space:

(1) Generate many protein conformations.
(2) Classify them into basins, and keep one representative conformation for each basin.
(3) Find the energetic barriers between these representative conformations (see next subsection).
(4) Once a set of conformations and the rates of conversions among them have been determined, we can use a standard algorithm (e.g., kinetic Monte Carlo[62]) to describe the dynamics of hopping in the conformation space.

What we gained compared to standard MD simulations, is that the extremely long trajectories required until the system hops to a different basin, have been replaced by a transition described by TST and kinetic Monte Carlo. In this way, one is able to sample many more conformations than with an ordinary MD simulation.

To connect these abstract concepts with the subject of this review paper, let us imagine that in an enzyme where proton is transferred through tunneling, there are some conformations that have donor–acceptor distance that is on average closer than other conformations. Then, conformation fluctuations that bring the protein towards these favorable for reaction conformations, will strongly enhance the rate. Unlike the rate-promoting vibrations we discussed earlier, whose timescale were so fast that several oscillations were needed before the proton is transferred, once the protein reaches a conformation in which donor and acceptor are close, it stays in that conformation for long enough time that tunneling takes place in that configuration of donor–acceptor.

The string method for finding reaction pathways

In order to implement the algorithm suggested in the previous subsection, we need a method for identifying the energetic barrier between two conformations. This is a specific case of a more general problem: given two states that lie in local minima of the energy landscape, find the MEP that connects them. This is a difficult problem that has attracted a lot of activity in the last decade.

The most successful method for identifying the MEP and energetic barrier between two stable states, is the ''nudged elastic band'' method.[63] This elastic band is a collection of system images that connect in configuration space the initial and final stable states. Adjacent system images are connected with fictitious harmonic strings. If this elastic band lies along the MEP, the force on it is equal to zero. If it is displaced from the MEP, then the forces acting on it (arising from the protein potential) will tend to return it towards the MEP. If there were no fictitious springs

that connected the adjacent system images, the images would relax towards either the initial or final stable states, i.e. the function of the springs is to keep the system images well separated. The presence of the fictitious springs guarantees that the elastic band will slide towards the MEP. The computational procedure is conceptually simple (but difficult to implement in practice): interpolate a few images between initial and final states; connect them with fictitious springs; use the protein potential to calculate the forces on the system images; Newton's equations of motion will make the elastic band slide towards the MEP.

The elastic band method has been applied with remarkable success.[63,64] However, it cannot be applied to problems that have rugged energy landscapes, like proteins. Recently, a similar method has been developed, that was designed to be applicable to rugged landscape systems. It is called the "finite-temperature string method".[65]

The difference with respect to the nudged elastic band method is the following: again, a chain of system images that connect the initial and final stable states is formed, but instead of connecting them with fictitious springs, the system images are assumed to be elements of a string of uniform density. Similarly to the elastic band, Newtonian dynamics using the protein potential makes the string slide towards the MEP, and during its descent it stretches, while keeping its property of uniform density.

The advantage of setting up the chain of images as a string rather as an elastic band, is that while for the elastic band the dynamics of the fictitious springs have to be solved explicitly (causing difficulties in rugged landscapes where the springs get stuck), the string dynamics that keeps its density uniform can be implemented with the standard SHAKE algorithm of molecular dynamics.

Recently, the string algorithm was used for finding the transition path in Alanine dipeptide. If one can apply it successfully to a large protein, then this will be the final missing link in the scheme outlined above for studying dynamics in conformation space.

References

1. Pauling, L. (1948). The nature of forces between large molecules of biological interest. *Nature* **161**, 707–709
2. Jencks, W. (1975). Binding energy, specificity, and enzymic catalysis: the Circe effect. *Adv. Enzym.* **43**, 219–310
3. Schramm, V. (2005). Enzymatic transition states: thermodynamics, dynamics and analogue design. *Arch. Biochem. Biophys.* **433**, 13–26
4. Agmon, N. and Hopfield, J. (1983). CO binding to heme proteins: a model for barrier height distributions and slow conformational changes. *J. Chem. Phys.* **79**, 2042–2053
5. Masgrau, L., Basran, J., Hothi, P., Sutclifie, M. and Scrutton, N. (2004). Hydrogen tunneling in quinoproteins. *Arch. Biochem. Biophys.* **428**, 41–51
6. Romesberg, F. and Schowen, R. (2004). Isotope effects and quantum tunneling in enzyme-catalyzed hydrogen transfer. Part I. The experimental basis. *Adv. Phys. Org. Chem.* **39**, 27–77
7. Kohen, A. (2005). Probes for hydrogen tunneling and coupled motion in enzymatic systems. In: Schowen, R., Klinman, J. and Hynes, J. (eds), *Handbook of Hydrogen, Vol. 2: Biological Aspects of Hydrogen Transfer*. Wiley, Weinheim

8. Bell, R.P. (1980). *The Tunnel Effect in Chemistry*. Chapman and Hall, New York
9. Marcus, R.A. and Sutin, N. (1985). Electron transfers in chemistry and biology. *Biochim. Biophys. Acta.* **811**, 265–322
10. Topaler, M. and Makri, N. (1994). Quantum rates for a double well coupled to a dissipative bath: accurate path integral results and comparison with approximate theories. *J. Chem. Phys.* **101**, 7500–7519
11. Schwartz, S.D. (1996). Quantum activated rates: an evolution operator approach. *J. Chem. Phys.* **105**, 6871–6879
12. Hänggi, P., Talkner, P. and Borkovec, M. (1990). Reaction-rate theory: fifty years after Kramers. *Rev. Mod. Phys.* **62**, 251–341
13. Straub, J.E., Borkovec, M. and Berne, B.J. (1987). Calculation of dynamic friction on intramolecular degrees of freedom. *J. Phys. Chem.* **91**, 4995–4998
14. Zwanzig, R. (1973). Nonlinear generalized Langevin equations. *J. Stat. Phys.* **9**, 215–220
15. Truhlar, D. (2005). Variational transition state theory and multidimensional tunneling for simple and complex reactions in the gas phase, solids, liquids, and enzymes. In: Kohen, A. and Limbach, H. (eds), *Isotope Effects in Chemistry and Biology*. Marcel Dekker, New York
16. Borgis, D. and Hynes, J.T. (1989). Proton transfer reactions. In: Cooper, A., Houben, J. and Chien, L. (eds), *The Enzyme Catalysis Process*. Plenum, New York p. 293
17. Borgis, D. and Hynes, J.T. (1991). Molecular dynamics simulation for a model nonadiabatic proton transfer reactions in solution. *J. Chem. Phys.* **94**, 3619–3628
18. Kiefer, P. and Hynes, J. (2004). Kinetic isotope effects for nonadiabatic proton transfer reactions in a polar environment: 1 Interpretation of tunneling kinetic isotopic effects. *J. Phys. Chem. A.* **108**, 11793–11808
19. Benderskii, V., Goldanskii, V. and Makarov, D. (1990). The theory of cryochemical reaction rates in the Leggett formalism. *Chem. Phys. Lett.* **171**, 91–96
20. Benderskii, V., Goldanskii, V. and Makarov, D. (1991). Low-temperature chemical reactions. Effect of symmetrically coupled vibrations in collinear exchange reactions. *Chem. Phys.* **154**, 407–424
21. Benderskii, V., Makarov, D. and Wight, C. (1994). Chemical dynamics at low temperatures. *Adv. Chem. Phys.* **88**, 1–385
22. Antoniou, D. and Schwartz, S.D. (1998). Activated chemistry in the presence of a strongly symmetrically coupled vibration. *J. Chem. Phys.* **108**, 3620–3625
23. Antoniou, D., Abolfath, M.R. and Schwartz, S.D. (2004). Transition path sampling study of classical rate-promoting vibrations. *J. Chem. Phys.* **121**, 6442–6447
24. Miller, W.H. (1976). Importance of nonseparability in quantum mechanical transition state theory. *Acc. Chem. Res.* **9**, 306–312
25. van der Zwan, G. and Hynes, J.T. (1983). Nonequilibrium solvation dynamics in solution reactions. *J. Chem. Phys.* **78**, 4174–4185
26. Bergsma, J.P., Gertner, B.J., Wilson, K.R. and Hynes, J.T. (1987). Molecular dynamics of a model S_N2 reaction in water. *J. Chem. Phys.* **86**, 1356–1376
27. Caratzoulas, S. and Schwartz, S.D. (2001). A computational method to discover the existence of promoting vibrations for chemical reactions in condensed phases. *J. Chem. Phys.* **114**, 2910–2918
28. Zavodszky, P., Kardos, J., Svingor, A. and Petsko, G. (1998). Adjustment of conformational flexibility is a key event in the thermal adaptation of protein. *Proc. Natl. Acad. Sci. USA* **95**, 7406–7411
29. Kohen, A., Cannio, R., Bartolucci, S. and Klinman, J. (1999). Enzyme dynamics and hydrogen tunneling in a thermophilic alcohol dehydrogenase. *Nature* **399**, 496–499
30. Antoniou, D. and Schwartz, S.D. (2001). Internal enzyme motions as a source of catalytic activity: rate promoting vibrations and hydrogen tunneling. *J. Phys. Chem. B* **105**, 5553–5558

31. Caratzoulas, S., Mincer, J.S. and Schwartz, S.D. (2002). Identification of a protein promoting vibration in the reaction catalyzed by horse liver alcohol dehydrogenase. *J. Am. Chem. Soc.* **124**, 3270–3276
32. Luo, J., Kahn, K. and Bruice, T.C. (1999). The linear dependence of $\log(k_{cat}/k_m)$ for reduction of NAD^+ by $PhCH_2OH$ on the distance between reactants when catalyzed by horse liver alcohol dehydrogenase and 203 single point mutants. *Bioorg. Chem.* **27**, 289–296
33. Cui, Q. and Karplus, M. (2002). Promoting modes and demoting modes in enzyme-catalyzed proton transfer reactions: a study of models and realistic systems. *J. Phys. Chem. B* **106**, 7927–7947
34. Basner, J.E. and Schwartz, S.D. (2004). Donor–acceptor distance and protein promoting vibration coupling to hydride transfer: a possible mechanism for kinetic control in isozymes of human lactate dehydrogenase. *J. Phys. Chem. B* **108**, 444–451
35. Núñez, S., Antoniou, D., Schramm, V.L. and Schwartz, S.D. (2004). Promoting vibrations in human PNP: a molecular dynamics and hybrid quantum mechanical/molecular mechanical study. *J. Am. Chem. Soc.* **126**, 15720–15729
36. Lewandowicz, A. and Schramm, V.L. (2004). Transition state analysis for human and plasmodium falciparum purine nucleoside phosphorylases. *Biochemistry* **43**, 1458–1468
37. Schramm, V.L. (2003). Enzymatic transition state poise and transition state analogues. *Acc. Chem. Res.* **36**, 588–596
38. Fedorov, A., Shi, W., Kicska, G., Fedorov, E., Tyler, P.C., Furneaux, R.H., Hanson, J.C., Gainsford, G.J., Larese, J.Z., Schramm, V.L. and Almo, S.C. (2001). Transition state structure of PNP and principles of atomic motion in enzymatic catalysis. *Biochemistry* **40**, 853–860
39. Mincer, J.S., Núñez, S. and Schwartz, S.D. (2004). Coupling protein dynamics to reaction center electron density in enzymes: an electronic protein promoting vibration in human purine nucleoside phosphorylase. *J. Theor. Comp. Chem.* **3**, 501–509
40. Dellago, C., Bolhuis, P., Csajka, F. and Chandler, D. (1998). Transition path sampling and the calculation of rate constants. *J. Chem. Phys.* **108**, 1964–1977
41. Dellago, C. and Chandler, D. (2003). Bridging the time scale gap with transition path sampling. In: Nielaba, P., Mareschal, M. and Ciccotti, G. (eds), *Bridging the Time Scales: Molecular Simulations for the Next Decade* (*Vol. 605 of Lecture Notes in Physics*). Springer Verlag, New York
42. Basner, J.E. and Schwartz, S.D. (2005). How enzyme dynamics helps catalyze a reaction, in atomic detail: a transition path sampling study, J. *Am. Chem. Soc.* **127**, 13822–13831
43. Radhakrishnan, R. and Schlick, T. (2004). Orchestration of cooperative events in DNA synthesis and repair mechanism unraveled by transition path sampling of DNA polymerase β's closing. *Proc. Natl. Acad. Sci. USA* **101**, 5970–5975
44. Amadei, A., Linssen, A.B.M. and Berendsen, H.J.C. (1993). Essential dynamics of proteins. *Proteins* **17**, 412–425
45. de Groot, B.L., Amadei, A., Scheek, R.M., van Nuland, N.A.J. and Berendsen, H.J.C. (1996). An extended sampling of the configurational space of HPr from E. coli. *Proteins* **26**, 314–322
46. Núñez, S., Wing, C., Antoniou, D., Schramm, V.L. and Schwartz, S.D. (2006). Insight into catalytically relevant correlated motions in human purine nucleoside phosphorylase, J. *Phys. Chem. A* **110**, 463–472
47. Erion, M.D., Takabayashi, K., Smith, H., Kessi, J., Wagner, S., Honger, S., Shames, S. and Ealick, S.E. (1997). Purine nucleoside phosphorylase. 1. Structure-function studies. *Biochemistry* **36**, 11725–11734
48. Wales, D. (ed.), (2003). *Energy Landscapes: Applications to Clusters, Biomolecules and Glasses*. Cambridge University Press, Cambridge, UK

49. Osborne, M., Schnell, J., Benkovic, S., Dyson, H. and Wright, P. (2001). Backbone dynamics in dihydrofolate reductase complexes: role of loop flexibility in the catalytic mechanism. *Biochemistry* **40**, 9846–9859
50. Sawaya, M. and Kraut, J. (1997). Loop and subdomain movements in the mechanism of *E. coli* dihydrofolate reductase: crystallographic evidence. *Biochemistry* **36**, 586–603
51. Rajagopalan, P., Lutz, S. and Benkovic, S. (2002). Coupling interactions of distal residues enhance dihydrofolate reductase catalysis: mutational effects on hydride ransfer rates. *Biochemistry* **41**, 12618–12628
52. Radkiewicz, J. and Brooks, C.L. (2000). Protein dynamics in enzymatic catalysis: exploration of dihydrofolate reductase. *J. Am. Chem. Soc.* **122**, 225–231
53. Garcia-Viloca, M., Truhlar, D. and Gao, J. (2003). Reaction-path energetics and kinetics of the hydride transfer reaction catalyzed by dihydrofolate reductase. *Biochemistry* **42**, 13558–13575
54. Hammes-Schiffer, S. (2004). Quantum-classical simulation methods for hydrogen transfer in enzymes: a case study of dihydrofolate reductase. *Curr. Opin. Struct. Biol.* **14**, 192–201
55. Thorpe, I. and Brooks, C. (2004). The coupling of structural fluctuations to hydride transfer in dihydrofolate reductase. *Proteins: Struct. Funct. Bioinf.* **57**, 444–457
56. Berry, R.S. and Kunz, R. (1995). Topography and dynamics of multidimensional interatomic potential surfaces. *Phys. Rev. Lett.* **20**, 3951–3954
57. Becker, O. and Karplus, M. (1997). The topology of multidimensional potential energy surfaces: theory and application to peptide structure and kinetics. *J. Chem. Phys.* **106**, 1495–1517
58. Becker, O. (1997). Geometrical versus topological clustering: insight into conformation mapping. *Proteins: Struct. Funct. Genet.* **27**, 213–226
59. Frauenfelder, H., Sligar, S. and Wolynes, P. (1991). The energy landscapes and motions of proteins. *Science* **254**, 1598–1603
60. Wolynes, P. (2005). Energy landscapes and solved protein-folding problems. *Phil. Trans. Roy. Soc. Lond. A* **363**, 453–464
61. Despa, F., Fernandez, A., Berry, R.S., Levy, Y. and Jortner, J. (2003). Interbasin motion approach to dynamics of conformationally constrained peptides. *J. Chem. Phys.* **118**, 5673–5682
62. Fichthorn, K. and Weinberg, W. (1991). Theoretical foundations of kinetic Monte Carlo simulations. *J. Chem. Phys.* **95**, 1090–1096
63. Henkelman, G., Johannesson, G. and Jonsson, H. (2000). Methods for finding saddle points and minimum energy paths. In: Schwartz, S.D. (ed.), *Theoretical Methods in Condensed Phase Chemistry*. Kluwer Academic Publishers, Dordrecht, The Netherlands
64. Evans, D. and Wales, D. (2003). The free energy landscape and dynamics of metenkephalin. *J. Chem. Phys.* **119**, 9947–9955
65. Weinan, E., Ren, W. and Vanden-Eijnden, E. (2005). Finite-temperature string method for the study of rare events. *J. Phys. Chem. B* **109**, 6688–6693