On the Relationship between Thermal Stability and Catalytic Power of Enzymes[†]

Maite Roca*, Hanbin Liu, Benjamin Messer and Arieh Warshel*

RUNNING TITLE: Flexibility Does Not Increase Enzyme Catalysis

Department of Chemistry, University of Southern California, 418 SGM Building, 3620 McClintock Avenue, Los Angeles, CA, 90089-1062, USA

* Corresponding authors. E-mail: mtrm@usc.edu, warshel@usc.edu Phone: +1 213 740 4114; Fax +1 213 740 2701

[1]Abbreviations: Tm, thermophile; Ms, mesophile; DHFR, dihydrofolate reductase; FEP/US, free energy perturbation umbrella sampling; EVB, empirical valence bond; RC, reaction coordinate; FC, folding coordinate; RS, reactant state; TS, transition state; PS, product state; DP, dispersed polaron.

Abstract

The possible relationship between the thermal stability and the catalytic power of enzymes is of great current interest. In particular, it has been suggested that thermophilic or hyperthermophilic (Tm) enzymes have lower catalytic power at a given temperature than the corresponding mesophilic (Ms) enzymes, because the thermophilic enzymes are less flexible (assuming that flexibility and catalysis are directly correlated). These suggestions presume that the reduced dynamics of the thermophilic enzymes is the reason for their reduced catalytic power. The present paper takes the specific case of dihydrofolate reductase (DHFR) and explores the validity of the above argument by simulation approaches. It is found that the Tm enzymes have restricted motions in the direction of the folding coordinate, but this is not relevant to the chemical process, since the motions along the reaction coordinate are perpendicular to the folding motions. Moreover, it is shown that the rate of the chemical reaction is determined by the activation barrier and the corresponding reorganization energy, rather than by dynamics or flexibility in the ground state. In fact, as far as flexibility is concerned, we conclude that the displacement along the reaction coordinate is larger in the Tm enzyme than in the Ms enzyme and that the general trend in enzyme catalysis is that the best catalyst involves less motion during the reaction than the less optimal catalyst. The relationship between thermal stability and catalysis appears to reflect the fact that in order to obtain small electrostatic reorganization energy it is necessary to invest some folding energy in the overall preorganization process. Thus, the optimized catalysts are less stable. This trend is clearly observed in the DHFR case.

Enzymes play fundamental roles in almost all life processes. They accelerate a great variety of metabolic reactions and control processes such as signaling, energy transduction and the translation of genetic information. The ability of enzymes to catalyze reactions by many orders of magnitude allows cells to carry out reactions that otherwise would not occur on biologically useful time scales. There is, therefore, a broad interest in understanding the origin of this catalytic power on a molecular level. Although it is clear that electrostatic effects play a major role in stabilizing the transition states of enzymatic reactions (for reviews, see refs (*1, 2*)), it is important to explore the importance of other factors. In particular, there is great current interest in dynamical contributions to catalysis (*3-11*). That is, although in our view there is no compelling evidence that can be used to support the dynamical proposal (see discussion in refs (*2, 12*)) it is possible to use studies of thermal adaptation of enzymes to support this proposal (*8-11*). More specifically, thermophilic (Tm) enzymes that evolved to function at high temperatures should be more stable at most temperatures that the corresponding mesophilic (Ms) enzymes that evolved to function at room temperature. Therefore, Tm enzymes should have smaller fluctuations than the corresponding Ms enzymes at most temperatures, and thus involve less dynamical motions. It is also known that Tm enzymes have lower catalytic power than the Ms enzymes at the same low temperature (*8, 13, 14*), and thus one may argue that the low catalytic power is due to the presumed decrease in dynamical motions. This argument may be used to support the idea that dynamics is the key to catalysis and that flexibility and enzyme catalysis are directly correlated. Although this is not the only argument in favor of the idea that the low catalytic power of Tm enzymes is associated with rigidity and lack of dynamical effects, it is presented here as the

clearest argument in favor of the relationship between dynamics and the catalytic power of the Tm.

This work takes the specific case of dihydrofolate reductase (DHFR) and explores the validity of the argument presented above as well as the general issue of the relationship between the catalytic power of enzymes and their thermal stability. DHFR catalyzes the reduction of 7,8-dihydrofolate (DHF) to 5,6,7,8-tetrahidrofolate (THF) through the oxidation of the coenzyme nicotidamide adenine dinucleotide phosphate (NADPH) (see Figure 1 for the reaction mechanism). DHFR is necessary for cellular metabolism because it maintains intracellular pools of THF, which is a coenzyme in a number of one-carbon metabolic processes and is essential for the biosynthesis of purines, thymidylate and several aminoacids. This property has made DHFR a clinical target for anticancer and antibacterial drugs (*15*).

It is found that the Tm enzymes less able to move in the direction of the folding coordinate, but this is not relevant to the chemical process since the motions along the reaction coordinate are perpendicular to the folding motion. Furthermore, it is shown that the rate of the chemical reaction is entirely determined by the activation barrier and the corresponding reorganization energy of the polar groups in the active site. The relationship between thermal stability and catalysis appears to be due to the fact that in order to obtain low reorganization energy and a preorganized active site, it is necessary for the system to invest some folding energy and thus to become less stable (*16, 17*).

Methods

Exploring the changes in the folding landscape. To explore the landscape effects and the probability of being at different configurations, it is important to be able to sample the protein configurational space in an efficient way. At present, it is hard to accomplish this task by all atom models and the optimal option involves the use of a simplified protein model of the type used in simulations of protein folding (e.g.(*18-23*)). In this study, we are not focused on the best simplified model, since we are mainly interested in the ability to reproduce the general feature of the difference between the Tm and Ms landscapes (namely the restricted motions of the Tm in the direction of the unfolding coordinate). Thus, we used here a simplified folding motion similar to the one used in ref (*24*) noting that this version is still under refinement.

The simplified model presented here is created by representing the explicit side chain of each residue by an effective unified "atom" and an additional dummy atom. The unified atoms are placed at the center of mass of the corresponding side chains (with a residue dependent charge and van der Waals radius), and the dummy atoms are placed along the corresponding C_α - C_β vectors and serve as tools for rotational transformations in the process of moving between the simplified and explicit models (which is not being used here). The dummy atoms do not have any charge or van der Waals interaction with the rest of the system. The backbone atoms of each residue are treated explicitly, and the interactions between main chain atoms are identical to those used in the explicit model.

The potential energy surface of the simplified model can therefore be written as,

$$U_{simplified} = U_{main} + U_{main-side} + U_{side-side} + U_{solvation}^{self}$$ (1)

U_{main} is basically the force field used in our explicit simulation, which is a standard part of the MOLARIS software package (*25, 26*). However, the torsional potential and hydrogen bonding interaction terms are modified to account for the missing solvent (*27*). $U_{side-side}$ is described by using an "8-6" potential (as was done in ref. (*24*) with a residue-dependent well depth (ε) and equilibrium separation radius (r^0).

$$U_{side-side} = \sqrt{\varepsilon_i \varepsilon_j} \left[3 \left(\frac{\sqrt{r_i^0 r_j^0}}{r_{ij}} \right)^8 - 4 \left(\frac{\sqrt{r_i^0 r_j^0}}{r_{ij}} \right)^6 \right] + 332 \frac{Q_i Q_j}{r_{ij} \varepsilon_{eff}}$$ (2)

where U is given in kilocalories per mole, r in angstroms and Q in atomic charge units. The parameters ε and r^0 are given in Table 1 (and they reflect both the average interaction between the side chains and the effect of the hydrophobic forces on these interactions). In addition to this standard treatment, we include a special treatment of the electrostatic interaction between ionized residues (the charges in Eq. (2), Q, are zero for the neutral groups and +1 or -1 for ionized groups). In this treatment we describe the interaction between side chains by using a large charge-charge dielectric constant, (ε_{eff} = 40), following the insight from our studies of electrostatic energies in proteins (*28*) and electrostatic contributions to protein stability (*29*). The solvation of each ionized group is treated by $U_{solvation}^{self}$ that will be considered below.

The $U_{main-side}$ term describes the interaction between the effective side chains and the main chain atoms. This interaction includes van der Waals interactions and are described by the first term of Eq. (2) where ε and r^0 of the main chains are taken from the corresponding van der

Waals potential of the main chain atoms. In addition we consider the electrostatic interactions between the main chain residual charges and ionized side chains (this is done using a dielectric constant $\varepsilon'_{eff} = 10$)

Since our treatment describes specifically ionized groups, it is crucial to account for the change in the solvation energy of each of these groups upon moving from water to its protein site. This is done by,

$$U_{solvation}^{self} = \sum_i \left[U_{np}^i(N_{np}^i) + U_{polar}^i(N_{polar}^i) \right] \tag{3}$$

where i runs over all ionized residues while N_{np} and N_{polar} are the number of nonpolar and polar neighboring residues, respectively. The functions U_{np} and U_{polar} are given by,

$$\begin{cases} U_{np}(N_{np}) = 4\exp\left(-0.2(N_{np} - 6)^2\right) & N_{np} \leq 6 \\ U_{np}(N_{np}) = 4 & N_{np} > 6 \end{cases} \tag{4}$$

$$\begin{cases} U_{polar}(N_{polar}) = -2\exp\left(-0.2(N_p - 4)^2\right) & N_p \leq 4 \\ U_{polar}(N_{polar}) = -2 & N_p > 4 \end{cases} \tag{5}$$

The number of polar and nonpolar neighboring residues is expressed by an analytical function of the form,

$$N_{np}^{(i)} = \sum_{j(np)} G(r_{ij}) \tag{6}$$

$$\begin{aligned} G(r_{ij}) &= 1 & r_{ij} \leq r_{np} \\ G(r_{ij}) &= \exp\left(-6(r_{ij} - r_{np})^2\right) & r_{ij} > r_{np} \end{aligned} \tag{7}$$

where r_{np} is the nonpolar radius that defines the cutoff range for nonpolar neighboring residues. The same treatment is used for polar residues (N_{polar}). This treatment is aimed at capturing the fact that an ionized group has to pay very large amount of energy for moving from water to a nonpolar environment (*30, 31*) and is usually surrounded by polar residues or water molecules (*28, 31*). The energy in Eq. (3) is expressed in a way that if the ionized group is surrounded by nonpolar groups (i.e. $N_{np} \approx 6$) the desolvation penalty is ~4 kcal/mol reflecting the very large compensation of desolvation energies found in modeling protein stability (*29*). On the other hand, when the ionized residue is surrounded by polar residues, it is more stable than in water by ~2 kcal/mol.

The parameters in eqs (3-7) were determined by comparing the results obtained with these equations to the results of the actual solvation energy evaluated by the semimacroscopic version of the Protein Dipole Langevin Dipole (PDLD/S) in its linear response approximation (LRA) version (PDLD/S-LRA). The overall simplified folding model was explored by preliminary validation studies which will be described elsewhere together with a more detailed description of our treatment (*27*). It is important to note that our conclusions do not depend drastically on the parameters in the model (see below).

Our simplified folding model can be used to explore the free energy landscape of a protein as a function of any given set of generalized coordinates (e.g., contact order, native contacts or native hydrogen bonds (*32-34*)) . In our case, we evaluated the free energy landscape in terms of two parameters; the radius of gyration (Rg) and the root-mean-square deviation (rmsd) from the folded structure (which was taken as the most stable structure of the simplified model).

The free energy surface was evaluated by using the free energy perturbation umbrella sampling (FEP/US) method in the same way that was used in our previous studies of electron transfer reactions (e.g. (*35, 36*)) and in EVB studies (*37*) while following the same treatment used in studies of ion channels (*38*) applying a mapping potential of the form,

$$\varepsilon_m = (1 - \lambda_m)\varepsilon_1 + \lambda_m\varepsilon_2 \tag{8}$$

$$\varepsilon_1(X_1, X_2) = U_{simplified} + K(Rg - Rg^{(1)})^2$$

$$\varepsilon_2(X_1, X_2) = U_{simplified} + K(Rg - Rg^{(2)})^2 \tag{9}$$

where Rg is the radius of gyration. Using ε_m and the FEP/US approach, we sort the system in two dimensions (X_1=Rg, X_2=rmsd) rather than in one dimension as in ref (*38*).

The starting points for the free energy landscapes were taken as the most stable structure of the simplified model after 200 ps of equilibration. After that, we obtained the free energy surfaces, following the above FEP/US method and applying a force constant of 100 kcal/mol·Å^2, by unfolding the systems increasing their Rg along 21 frames of 60 ps each at 300 K and with 1 fs time step.

In addition to the FEP/US mapping we also explored regions near the folded state by running 4 ns molecular dynamics (MD) simulations without any constraints. The probability of being at different points on the surface was evaluated by following the corresponding Boltzmann distribution.

Exploring the energetics and dynamics along the chemical reaction coordinate. To analyze the relationship between thermal stability and catalysis we need to evaluate the activation barrier for the chemical step in addition to the study of the folding landscape reported above. This requires very extensive sampling, which is hard to accomplish reliably by current molecular orbital QM/MM methods, despite significant advances in this direction in studies of different enzymes in general (for a review see ref (*39*)) and DHFR in particular (e.g. refs (*40-43*)). We believe that at present, the most effective strategy is to use the empirical valence bond (EVB) method. This method has been described extensively elsewhere (e.g. ref (*2, 17*)) and only some key details are given here. This is particularly important in view of the need to evaluate the reorganization which appears to be a major factor in determining the catalytic effect in DHFR (see ref. (*44*)). Thus, we use here the EVB as our main simulation tool.

The reacting system is described schematically in Figure 1 and the location of the substrate and the coenzyme in DHFR, in the mesophilic (EcDHFR) and thermophilic (TmDHFR) systems, is depicted in Figure 2. The EVB describes the reacting system by two zero order (diabatic states),

$$\Psi_1 = \text{D-H} \ \ \text{A}^+$$

$$\Psi_2 = \text{D}^+ \ \ \text{H-A} \tag{10}$$

where DH, D^+, A^+ and AH are the NADPH, NADP^+, DHF-H^+ and THF systems, respectively, described in Figure 1. The energy of each of these states is described by a force-field-like empirical potential function and the diabatic functions are mixed with an off diagonal term to give the actual potential surface for the reaction. The effect of the environment is included in the energy of each diabatic state. The EVB is parametrized to reproduce the ab-initio surface (and relevant experimental information about the reaction in aqueous solution and then used without

11

changing any parameter (only the environment changes) to generate the energy surface in the enzyme active site. Since this approach focuses on the change of moving between different environments rather than on the absolute energy of the reacting substrate, it provides a very effective tool for studies of enzyme catalysis.

The EVB treatment was identical to that used recently in our study of DHFR and its mutants (*44*) by using the MOLARIS simulation program (*25*) and applying the ENZYMIX force field. The EVB activation barriers were evaluated using the same FEP/US approach (*17, 45*) used in all of our EVB studies. This approach gives, in addition to the activation free energies, the reorganization energies (see ref. (*44*) and the discussion below). The simulation systems were solvated by the surface-constrained all atom solvent (SCAAS) model (*46*) using a radius for the explicit region of 20 Å, and the long-range electrostatic effects were treated by the local reaction field (LRF) method (*47*). The free energy profile for the reaction was mapped over different initial structures to determine an average free energy barrier. The FEP mapping of each step was typically evaluated after an equilibration time of 300 ps, followed by 31 frames of 30 ps each at 300 K with a time step of 1 fs for moving along the reaction coordinate with our all atom surface constrained spherical model.

Results and discussion

The folding landscape is indeed more confined in the Tm than in the Ms. In the first step of our study we explored the folding landscape for the hyperthermophilic (Tm) dihydrofolate reductase from *Thermotoga maritime* (TmDHFR) (*48*) and the mesophile (Ms) from *Escherichia coli*

(EcDHFR) (*49*). The coordinates for the TmDHFR and EcDHFR structures were obtained from the Protein Data Bank (entries 1D1G and 1RX2, respectively). The proteins are described in Figure 2. The calculations were carried out with the simplified folding model described in Methods. Note that in the Ms we considered the entire system in generating the folding energy surface, but in the Tm, we allow only one subunit to unfold. The calculated results are summarized in Figure 3 and, as seen from the figure, the folding free energy surface is stiffer for the Tm than for the Ms, which is what would be expected from common knowledge (*8, 48, 50*). These results indicate that the motion on the Tm surface is more confined to a smaller space than the corresponding motion in the Ms case. We also explored the folding free energy of these two systems in ref (*29*) and reproduce a folding energy of -34.6 kcal/mol for the Tm and -8.1 kcal/mol for the Ms (the corresponding experimental values (*50, 51*) are -34 kcal/mol and -6 kcal/mol for the Tm and Ms respectively). Interestingly, we obtained better results in terms of the shape of the surface (we obtained more stable Tm) using the electrostatic term of eq 2, where we ionized all the residues which are ionized at pH=7. It is important to note that the results of this study did not change significantly with a change in the parameters of the simplified model. In other words, this part of our study mainly confirms the common view that the Tm landscape is stiffer than that of the Ms landscape.

The activation barrier for the chemical step are determined by the reorganization energy. In the next step, we evaluated the activation barriers of the two enzymes by the EVB-FEP/US procedure, applying extensive averaging over different initial conditions. The calculated results are summarized in Figure 4. The calculated activation barriers reasonably reproduced the trend in

the observed rate constant (*13, 49, 52*) as is evident from Table 2, where the barrier in the Tm is higher than in the Ms.

The EVB calculations provided, in addition to the adiabatic activation barriers, the diabatic free energy functions and its decomposition into solvent and solute contributions and the corresponding total reorganization energies (see Figure 5). These energies (λ in Fig. 5) represent the reduction in the energy of the protein (or the solvent) when the reacting system is placed on the potential surface of the product state, at the equilibrium coordinate of the reactant state, and thus allowed to relax to the product equilibrium coordinate. In general enzymes reduce their activation barriers by reducing the reorganization energy (*2, 44*). The calculations, which are summarized in Figure 5, establish that the reorganization energy for the chemical step is significantly larger for the Tm than the Ms, which is in full agreement both with the corresponding change in activation barriers and with our previous study (*44*).

The curvature of the energy surface along the chemical coordinate is similar in the Tm and Ms. The reorganization energy evaluated above was used to construct a reduced two-dimensional free energy surfaces for the chemical step (Figures 6 and 7). In Figure 6, we decomposed the effective free energy surface of the system to its solute and solvent components (see ref. *38* for details). Since the displacement between the minima of the reactant and product state is determined by the reorganization energy (*44*), we see in the figure a larger displacement for the Tm than for the Ms (as seen from Figure 6, the minimum changes from approximately -16 (unitless) in EcDHFR to approximately -20 (unitless) in TmDHFR). Thus, the overall reaction coordinate is longer in the Tm than in the Ms.

Another useful representation that shows that the reaction pathway is slightly longer in the Tm and presents a higher activation barrier is obtained by plotting the donor-acceptor (D-A) distance versus the donor-hydrogen (D-H) distance. This representation is shown schematically in Figure 7. The surfaces were generated by fitting a simple EVB potential to the behavior of the actual surface at selected points. As seen from the figure, the donor-acceptor distance at the region of the RS is slightly larger for the Tm (~4.4 Å for the Ms and 4.7 Å for the Tm) and there is a significant difference along the pathway between both surfaces, where the TmDHFR presents a higher activation free energy. That is, in the Ms system, the D-A distance roughly reaches 3.4 Å before the energy starts to increase and the TS occurs with a donor-acceptor distance of ~2.6 Å. On the other hand, in the TmDHFR system, the free energy starts to rapidly increase compared to that of the Ms system, presenting a larger D-A distance, at ~4.0 Å and the TS also occurs at ~2.6 Å (see also ref. *38*).

Apparently, as seen from Figures 6 and 7, the reaction pathway is significantly longer in the Tm. However, the length of the chemical pathway cannot be correlated with the free energy curvature at the reactant state. That is, as seen from Figure 5, the curvature at the RS in the direction of the reaction coordinate, is similar for the Tm and Ms. This is fully consistent with the view that the free energy curves are harmonic functions, with equal curvature and with shifted origins. The fact that the curvature of the reaction coordinate is similar for the Tm and Ms can be verified on a more quantitative level by dispersed polaron (DP) (Spin Boson) calculations (*53*) that provide the projection of the normal modes of the system onto the chemical coordinate (Figure 8). That is, as seen from the figure, the frequencies of the motions that have a projection on the reaction coordinate are similar for the Tm and Ms and the main difference is that the

amplitudes of these motions are larger for the Tm, since the reaction in this system involves a larger reorganization energy. Note that the overall reorganization energy (λ) is given by (*12, 53*),

$$\lambda = \sum_i \frac{1}{2} \hbar \omega_i \Delta_i^2$$

(11)

where Δ_i is the origin shift of the ith mode. Thus the Δ_i are larger for the Tm, however, the ω_i values are similar for the Tm and Ms enzymes. This means that the curvature of the surfaces is similar in the chemically relevant direction. Note that if the frequencies of the most effective reactive modes were significantly smaller in the Ms than in the Tm, we would have a peak at a high frequency in the Tm case and at a lower frequency in the Ms case.

Reexaming the relationship between the flexibility and catalytic power. With the above findings in mind, we may return to the issue of the relationship between flexibility and catalytic power. The folding calculations, which are summarized in Figure 3, confirm the idea that the free energy profile for the folding process is steeper for the Tm. Thus, the fluctuations in the direction of the folding coordinate at a given temperature are smaller than the corresponding fluctuations for the Ms. On the other hand, the above calculation of the chemical step indicate that the fluctuations in the direction of the reaction coordinate are similar in the RS of the Tm and Ms and that thus the relationship between thermal stability to catalysis is not related to the corresponding increase in rigidity.

To explore the seemingly conflicting results discussed above, we evaluated the vectors that represent the multidimensional reaction coordinate (calculated as the difference between product state and reactant state) and the vector that represents the folding coordinate in the Ms (EcDHFR). The folding vector was evaluated in two ways. In the first approach, we evaluated a

structure of a particular unfolded protein (along the minimum energy path in Figure 3a). In the

second approach, we considered the change between the folded (namely closed, see ref. (*49*))

structure to the occluded structure (*49*) (PDB entry 1RX7) as a representative of the folding

coordinate. During the catalytic cycle, DHFR undergoes conformational changes between the

closed and occluded states which, respectively, describe whether the active site is closed or

occluded, repectively, by the Met20 loop (*49*). As shown in Figure 9, the two sets of coordinates

are nearly perpendicular in both cases. Thus, there is no reason to have a correlation between the

two sets of motions even in the extremely unlikely case of coherent motions (see below).

To clarify the nature of the chemical process, we should focus on the relationship between the

motions in the direction of the reaction coordinate in the Ms and Tm enzymes. Here, as seen in

Figure 5, we have similar available landscapes for the RS complex in the Tm and Ms and thus

similar average thermal motions at room temperature (*17*). The only difference is that the barrier

is higher in the case of the Tm due to the larger reorganization energy and the larger work term

(the free energy of bringing the donor and acceptor to the interaction distance at the reactant

state) (*44*). This means that in the case of the Tm, the system has to wait longer time for the rare

fluctuations that reaches the transition state (TS), because the activation barrier is larger.

To further validate the above argument, we calculated the productive trajectories for the Tm

and Ms by running downhill trajectories from the TS (note that the time reversal of the downhill

trajectory provides the reactive trajectory that moves from the RS to the TS, (see ref (*6, 17*)). As

illustrated in the Figure 10 and as found in any related calculation (*6*), the system starts to move

downhill with at most two excursions at the TS and then moves to the RS and stays at this state

for a very long time, (only a short time is depicted in the figure). This means that the time

reversal of the downhill trajectory would correspond to a very long random motion on the RS and then eventually to a very rare event of a generation of a reactive trajectory, with a probability that is determined by the activation free energy. Comparing both simulations, we find that in the Tm case the system oscillates along a larger range of donor-acceptor distances and acceptor-hydrogen distances in the RS region than the Ms system. This finding is in total agreement with the results observed in Figures 6 and 7 where the reaction pathway is slightly longer in the Tm system.

To further explore the dynamical behavior, we evaluated the projection of the reactive trajectory onto the chemical reaction coordinate and the folding coordinate in the Ms (EcDHFR). This was done by considering the scalar products,

$$\Delta \boldsymbol{r}_{RC} \cdot \delta \boldsymbol{r}(t)_{RC} = (\Delta \boldsymbol{r}_{RC}^{I}, \Delta \boldsymbol{r}_{RC}^{II}) \cdot (\delta \boldsymbol{r}^{I}(t), \delta \boldsymbol{r}^{II}(t)) / |\Delta \boldsymbol{r}_{RC}| =$$

$$\left[(\Delta \boldsymbol{r}_{RC}^{I} \cdot \delta \boldsymbol{r}^{I}(t)) + (\Delta \boldsymbol{r}_{RC}^{II} \cdot \delta \boldsymbol{r}^{II}(t)) \right] / |\Delta \boldsymbol{r}_{RC}| \tag{12}$$

$$\Delta \boldsymbol{r}_{FC} \cdot \delta \boldsymbol{r}(t)_{RC} = \left[(\Delta \boldsymbol{r}_{FC}^{I} \cdot \delta \boldsymbol{r}^{I}(t)) + (\Delta \boldsymbol{r}_{FC}^{II} \cdot \delta \boldsymbol{r}^{II}(t)) \right] / |\Delta \boldsymbol{r}_{FC}| \tag{13}$$

where $\Delta \boldsymbol{r}_{RC}$, $\Delta \boldsymbol{r}_{FC}$ and $\delta \boldsymbol{r}(t)_{RC}$ are, respectively, the vectors of displacements between the reactant and product, the displacement along the folding coordinate and the time dependent downhill trajectory. Although the figure was generated using the transition from the folded structure to the occluded one, we obtain similar results considering the transition to a partially unfolded structure generated using the free energy landscape (Figure 3a). The vectors are divided into the components in the EVB region (substrate, DHF-H^{+}, and coenzyme NADPH) and the residues within 5 Å of the EVB atoms ($\Delta \boldsymbol{r}_{RC}^{I}$) and the surrounding protein ($\Delta \boldsymbol{r}_{RC}^{II}$). Here, we found that it is most informative to include only the first part of Eq. (12 and 13), namely the

18

projections that include only the reacting region and its neighboring surroundings. This reflects the fact that the changes in the distant residues of the $\Delta \boldsymbol{r}_{RC}^{II}$ and $\Delta \boldsymbol{r}_{FC}^{II}$ are small and the corresponding changes in $\delta r(t)_{RC}^{II}$ during the downhill trajectory are modulated by random thermal motions that lead to different results in each trajectory. At any rate, as seen from Figure 11, the projections are much larger in the case of the reaction coordinate than in the folding coordinate. This result again indicates that the reaction coordinate is not directly related to the folding coordinate, since the folding coordinate does not change significantly during the reactive trajectory.

Perhaps more instructive analysis is obtained from the autocorrelation of $\Delta \boldsymbol{r}_{RC} \cdot \delta \boldsymbol{r}(t)_{RC}$ and $\Delta \boldsymbol{r}_{FC} \cdot \delta \boldsymbol{r}(t)_{RC}$ (Figure 12). As seen from the figure, the system loses its "memory" about the reactive motion in a very short timespan of only a few picoseconds. This shows that the information from the milisecond folding motions cannot be transferred in a ballistic (non thermal) way to the reactive trajectory. This point is further clarified in Figure 13 where we consider the nature of the fluctuations of the system. In the case of the Ms, we have a shallow folding potential but even in this case a fluctuation from a partially unfolded structure ((a) in Figure 13a) to the RS will be completely randomized before it can be coupled to the reaction coordinate. The situation is even more obvious in the case of the Tm where we do not have significant fluctuations in the folding direction so that the system is trapped in the RS region before generating the reactive trajectory. More significantly, the protein spends most of the time at the RS in both systems and the motions are randomized by the thermal energy with similar average amplitudes in the Tm and Ms. In both cases, the thermal motions in the RS lead to reactive trajectories only according to the Boltzmann probability of being in the TS. Thus,

although a "frictionless" fluctuation along the folding coordinate (a)→(b) can in principle lead

from a partially unfolded structure to the RS and TS, the thermal energy of such a fluctuation will

be completely randomized before it can be transferred to the "chemical" coordinate. The fast

randomization of the thermal energy can be realized by considering Figure 12, and also by

considering studies of much faster processes such as the photoisomerization process of

bacteriorhodopsins when the thermal energy is almost fully dissipated in a few picoseconds (*54*).

Nevertheless, it will be interesting to consider longer time simulation to fully establish the view

that emerges from the present study.

Concluding Remarks

The idea that thermophilic enzymes have slow reaction rate due to their restricted dynamical

ability (*8-11*) has been frequently raised as support to the general hypothesis that enzymes

catalyze their reactions by dynamical effects. Before addressing this issue, it is important to

clarify a point that is sometimes not emphasized by the proponents of the above idea;

thermophilic enzymes do not need to evolve to have a very large catalytic power but they simply

need to remain folded at high temperature. Such enzymes exploit the fact that the rate constants

of reactions in condensed phases, with a barrier of more than a few kilocalories per mole, follow

transition state theory (TST) (see e.g. (*17*)) with a minor modification and can be written as,

$$k = \kappa \left(\frac{1}{2} \langle | \dot{x} | \rangle_{TS} \Big/ \Delta x^{\ddagger} \right) \exp\left[- \Delta G^{\ddagger} \beta \right] = \kappa \, k_{TST} \tag{14}$$

where κ is the transmission factor, $\langle |\dot{x}| \rangle_{TS}$ is the average of the absolute value of the velocity along the reaction coordinate at the transition state ($\Delta x^{\ddagger}$), and $\beta = 1/k_B T$ (where k_B is the Boltzmann constant and T the absolute temperature). The term $\Delta G^{\ddagger}$ designates the multidimensional activation free energy that reflects the probability for the system to be in the TS region. Because of the exponential 1/T dependence of the rate constant, the Tm enzymes can function with a much larger activation barrier than the corresponding Ms enzymes, at its intended temperature range. Thus, there is no need to postulate any exotic dynamical reason for the fact that a Tm enzyme has a slower rate than the corresponding Ms at the temperature where the Ms works. It is quite logical to assume that the Ms evolved to reduce the activation barrier as much as possible (or needed) while the Tm evolved to stay folded at a high temperature and thus was unable to preorganize in perfect way and to provide a low activation barrier. This is of course consistent with the inverted relationship between stability and activation barrier (*2, 16*), where the Tm has a larger folding energy (larger stability) and a higher activation barrier and thus a slower reaction rate. Note that this relationship is in one direction. That is, enzymes with larger k_{cat} are usually less stable than mutants with much smaller k_{cat} (less good catalysis), sometimes less stable enzymes can still be good catalysts if the instability is not coupled to the active site preorganization.

Despite these considerations, it is important to explore the proposal that the Ms enzyme is a better catalyst than the Tm enzyme (at the same temperature) because of dynamical effects and increased flexibility. The present examination of this proposal led to the finding that the difference in rate constants between the Tm and the Ms is due to the difference in activation free energy for the chemical step (the calculated activation barriers and Eq. (14) fully accounted for the difference in rate). It was also found that the change in the activation barrier is due to the

21

change in reorganization energy and work function (for a more detailed analysis, see also ref (*44*)). Furthermore, it was concluded, in agreement with previous proposals, that the motions of the Tm are more restricted in the folding direction than those of the Ms. However, these folding motions are orthogonal to the reaction coordinate and are unlikely to be coupled to the motions involved in the catalytic process. Even in the case when the motions in the folding direction have some projection on the direction of the reaction coordinate, the autocorrelation of this projection decays in a few picoseconds so that the information from motions along the folding coordinate is unlikely to be transferred to the reactive trajectories.

As stated above, the folding surface is stiffer in the Tm case than in the Ms case, and thus can be considered to have "less motion". However, the surface along the chemical coordinate does not have a larger curvature in the Tm than in the Ms case. This can be seen from inspection of the reactant and product regions in Figure 5. It can also be observed from the DP analysis of Figure 8. In other words, a careful examination of the nature of the reactive motions indicated that these motions have similar frequencies in the Tm and Ms, but in the case of the Tm, the displacement along these modes on the way to the TS is larger. Thus the system involves larger reorganization energy and has to wait a longer time until a reactive fluctuation occurs because the barrier is higher. This well known relationship between barrier height and rate (see e.g. (*17*)) has very little to do with dynamical effects, but rather with the fact that the chance for a reactive trajectory is proportional to the Boltzmann factor of the corresponding barrier (i.e. to Eq. (14)).

At this point it might be instructive to consider the specific difference between the Tm and Ms of DHFR. This issue was explored here mainly in an "operational" way by evaluating the relevant activation barriers and folding landscapes. However, it appears from our study of the stability of the Tm and Ms (*29*) that the dimer uses the interaction between its ionized groups to

gain a remarkable amount of extra stability. We also know from our study of the chemical process in the Tm that this protein pulls the donor and acceptor further apart and also is unable to provide small reorganization energy (*44*). The exact way by which the extra stabilization of the Tm leads to an increase in the reorganization energy has not been explored here. However, this effect is fully consistent with the inversed relationship between folding stability and catalysis.

The entropic features of psychrophilic enzymes have also been raised as a support for the idea that flexibility is related to catalysis (*55-57*). That is, such enzymes reduce the activation enthalpy ($\Delta H^{\ddagger}$) (relative to the corresponding thermophilic enzymes) while increasing the contribution of the activation entropy of the system ($-T\Delta S^{\ddagger}$). This feature may be rationalized by suggesting that the motion in the RS of the psychrophilic enzymes is larger than in the TS. Since this effect is larger in the psychrophilic than in Ms, one may assume that there are more motions in the psychrophilic RS than in the Ms RS. However, such an assumption faces the following serious problems. (i) The actual effect of ground state (GS) motions (namely the effect of $-T\Delta S^{\ddagger}$) is, in fact, anticatalytic, since it increases the activation barrier rather than decreases it. Thus one can argue that motions reduce catalysis rather than increase it. (ii), the catalytic effect of the psychrophilic at low temperature is due the reduction of $\Delta H^{\ddagger}$ which is most probably associated with the reduction of the reorganization energy and not with changes in dynamical effects. In fact, $-T\Delta S^{\ddagger}$ only tells us about the difference between the space available for thermal motions in the RS and TS, but not about the difference between the motions in psychrophilic and Ms enzymes. Thus, even if the local flexibility is larger in the RS of psychrophilic than in the RS of Ms, elucidating of the actual nature of the catalytic effect should involve calculations of the activation entropy and enthalpy (as done in ref (*58*)). It is most likely that such a study will reproduce the observed fact but will not establish a direct relationship between flexibility and

catalysis. Such a study, which is left to subsequent works, would probably find that a reduction in flexibility leads to lower $\Delta H^{\ddagger}$, but that some other requirements (not related to catalysis) lead to the apparent change in flexibility.

Overall it appears that enzymes catalyze their reactions by reducing the reorganization energies of the polar and ionized groups that stabilize the TS, and rather than by changing their flexibility. The relationship between catalysis and flexibility (if any) can be either coincidental or a result of the need to reduce the reorganization energy. For example (see (*6*)) if the charge distribution is more localized in the TS than in the RS it is simpler to reduce the barrier and reorganization energy at low temperature by fixing the dipoles at the TS (this will be an enthalpic effect) and by entropic effect at high temperature. However, due to entropy/enthalpy compensation, the main issue here is the reduction of the reorganization free energy and not the changes in its entropic and enthalpic components. In other words, the enzyme tries to have a preorganized active site where the contribution of the enzyme structural change to the free energy of moving from the RS structure to the TS structure is minimal. This task can be accomplished by using either enthalpy or entropy depending on the specific case (e.g. the change in charge during the reaction) and perhaps on inherent stability constraints. Thus, postulating that catalysis is related to flexibility is not useful since it does not tell us how and why. Exploring what determines the reorganization energy (e.g. the relationship between folding energy and reorganization energy) is very useful and potentially predictive. Asking how the reorganization energy depends on temperature and what the constraints are on such a dependence is much more complex and requires further simulation studies.

In concluding this work, we would like to emphasize a key argument against the idea that an increase in catalysis is associated with an increase in flexibility. Apparently, catalysis is related to the magnitude of the reorganization energy and thus to the displacement along the reaction coordinate between the reactant and product states (*12*). A better catalyst has a smaller displacement rather than a larger one (see Figure 6). Thus, we in fact need less motion rather than more motion in a good catalyst.

With regard to the relationship between the thermal stability and the catalytic power of enzymes, it seems that the present work provide further support to the idea (*16*) that some folding energy and stability have to be sacrificed in order to increase catalytic power. In other words, in order to obtain small reorganization energy and small activation free energy with a preorganized active site, it is necessary for the system to create local instability in the folded structure and this is reflected in the reduction of the overall folding energy.

ACKNOWLEDGMENT

REFERENCES

1. Marti, S., Roca, M., Andres, J., Moliner, V., Silla, E., Tunon, I., and Bertran, J. (2004) Theoretical insights in enzyme catalysis. *Chemical Society Reviews 33*, 98-107.

2. Warshel, A., Sharma, P. K., Kato, M., Xiang, Y., Liu, H. B., and Olsson, M. H. M. (2006) Electrostatic basis for enzyme catalysis. *Chemical Reviews 106*, 3210-3235.

3. Epstein, D. M., Benkovic, S. J., and Wright, P. E. (1995) Dynamics of the Dihydrofolate-Reductase Folate Complex - Catalytic Sites and Regions Known to Undergo Conformational Change Exhibit Diverse Dynamical Features. *Biochemistry 34*, 11037-11048.

4. Schnell, J. R., Dyson, H. J., and Wright, P. E. (2004) Structure, dynamics, and catalytic function of dihydrofolate reductase. *Annual Review of Biophysics and Biomolecular Structure 33*, 119-140.

5. Hammes-Schiffer, S. (2004) Quantum-classical simulation methods for hydrogen transfer in enzymes: a case study of dihydrofolate reductase. *Current Opinion in Structural Biology 14*, 192-201.

6. Olsson, M. H. M., Parson, W. W., and Warshel, A. (2006) Dynamical contributions to enzyme catalysis: Critical tests of a popular hypothesis. *Chemical Reviews 106*, 1737-1756.

7. Wang, L., Tharp, S., Selzer, T., Benkovic, S. J., and Kohen, A. (2006) Effects of a distal mutation on active site chemistry. *Biochemistry 45*, 1383-1392.

8. Zavodszky, P., Kardos, J., Svingor, A., and Petsko, G. A. (1998) Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins. *Proceedings of the National Academy of Sciences of the United States of America 95*, 7406-7411.

9. Ringe, D., and Petsko, G. A. (1999) Quantum enzymology - Tunnel vision. *Nature 399*, 417-418.

10. Kohen, A., Cannio, R., Bartolucci, S., and Klinman, J. P. (1999) Enzyme dynamics and hydrogen tunnelling in a thermophilic alcohol dehydrogenase. *Nature 399*, 496-499.

11. Kohen, A., and Klinman, J. P. (2000) Protein flexibility correlates with degree of hydrogen tunneling in thermophilic and mesophilic alcohol dehydrogenases. *Journal of the American Chemical Society 122*, 10738-10739.

12. Warshel, A., and Parson, W. W. (2001) Dynamics of biochemical and biophysical reactions: insight from computer simulations. *Quaterly Reviews of Biophysics 34*, 563-679.

13. Maglia, G., Javed, M. H., and Allemann, R. K. (2003) Hydride transfer during catalysis by dihydrofolate reductase from Thermotoga maritima. *Biochemical Journal 374*, 529-535.

14. Wrba, A., Schweiger, A., Schultes, V., Jaenicke, R., and Zavodszky, P. (1990) Extremely Thermostable D-Glyceraldehyde-3-Phosphate Dehydrogenase from the Eubacterium Thermotoga-Maritima. *Biochemistry 29*, 7584-7592.

15. Blakley, R. L. (1984) *Folates and Pterins*, Wiley, New York.

16. Warshel, A. (1978) Energetics of Enzyme Catalysis. *Proceedings of the National Academy of Sciences of the United States of America 75*, 5250-5254.

17. Warshel, A. (1991) *Computer Modeling of Chemical Reactions in Enzymes and Solutions*, Wiley Interscience, New York.

18. Levitt, M., and Warshel, A. (1975) Computer-Simulation of Protein Folding. *Nature 253*, 694-698.

19. Bryngelson, J. D., and Wolynes, P. G. (1987) Spin-Glasses and the Statistical-Mechanics of Protein Folding. *Proceedings of the National Academy of Sciences of the United States of America 84*, 7524-7528.

20. Hinds, D. A., and Levitt, M. (1992) A Lattice Model for Protein-Structure Prediction at Low Resolution. *Proceedings of the National Academy of Sciences of the United States of America 89*, 2536-2540.

21. Shakhnovich, E., Abkevich, V., and Ptitsyn, O. (1996) Conserved residues and the mechanism of protein folding. *Nature 379*, 96-98.

22. Dill, K. A. (1990) Dominant Forces in Protein Folding. *Biochemistry 29*, 7133-7155.

23. Olszewski, K. A., Kolinski, A., and Skolnick, J. (1996) Folding simulations and computer redesign of protein A three-helix bundle motifs. *Proteins-Structure Function and Genetics 25*, 286-299.

24. Fan, Z. Z., Hwang, J. K., and Warshel, A. (1999) Using simplified protein representation as a reference potential for all-atom calculations of folding free energy. *Theoretical Chemistry Accounts 103*, 77-80.

25. Lee, F. S., Chu, Z. T., and Warshel, A. (1993) Microscopic and Semimicroscopic Calculations of Electrostatic Energies in Proteins by the Polaris and Enzymix Programs. *Journal of Computational Chemistry 14*, 161-185.

26. MOLARIS version alpha9.06, Chu, Z. T., Villa, J., Strajbl, M., Schutz, C. N., Shurki, A., and Warshel, A., University of Southern California, Los Angeles.

27. Messer, B., Roca, M., and Warshel, A. *(in preparation)*.

28. Warshel, A., Sharma, P. K., Kato, M., and Parson, W. W. (2006) Modeling electrostatic effects in proteins. *Biochimica Et Biophysica Acta-Proteins and Proteomics 1764*, 1647-1676.

29. Roca, M., Messer, B., and Warshel, A. (2007) Electrostatic contributions to protein stability and folding energy. *FEBS Letters 581*, 2065-2071.

30. Warshel, A., Russell, S. T., and Churg, A. K. (1984) Macroscopic Models for Studies of Electrostatic Interactions in Proteins - Limitations and Applicability. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences 81*, 4785-4789.

31. Warshel, A., and Russell, S. T. (1984) Calculations of Electrostatic Interactions in Biological-Systems and in Solutions. *Quarterly Reviews of Biophysics 17*, 283-422.

32. Fersht, A. R. (2000) Transition-state structure as a unifying basis in protein-folding mechanisms: Contact order, chain topology, stability, and the extended nucleus mechanism. *Proceedings of the National Academy of Sciences of the United States of America 97*, 1525-1529.

33. Robinson-Rechavi, M., and Godzik, A. (2005) Structural Genomics of Thermotoga maritima proteins shows that contact order is a major determinant of protein thermostability. *Structure 13*, 857-860.

34. Shea, J. E., Onuchic, J. N., and Brooks, C. L. (2002) Probing the folding free energy landscape of the src-SH3 protein domain. *Proceedings of the National Academy of Sciences of the United States of America 99*, 16064-16068.

35. Warshel, A. (1982) Dynamics of Reactions in Polar-Solvents - Semi-Classical Trajectory Studies of Electron-Transfer and Proton-Transfer Reactions. *Journal of Physical Chemistry 86*, 2218-2224.

36. Hwang, J. K., and Warshel, A. (1987) Microscopic Examination of Free-Energy Relationships for Electron-Transfer in Polar-Solvents. *Journal of the American Chemical Society 109*, 715-720.

37. Aqvist, J., and Warshel, A. (1993) Simulation of Enzyme-Reactions Using Valence-Bond Force-Fields and Other Hybrid Quantum-Classical Approaches. *Chemical Reviews 93*, 2523-2544.

38. Kato, M., and Warshel, A. (2005) Through the channel and around the channel: Validating and comparing microscopic approaches for the evaluation of free energy profiles for ion penetration through ion channels. *Journal of Physical Chemistry B 109*, 19516-19522.

39. Shurki, A., and Warshel, A. (2003) Structure/function correlations of proteins using MM, QM/MM, and related approaches: Methods, concepts, pitfalls, and current progress. *Adv. Protein Chem. 66*, 249-313.

40. Garcia-Viloca, M., Truhlar, D. G., and Gao, J. L. (2003) Reaction-path energetics and kinetics of the hydride transfer reaction catalyzed by dihydrofolate reductase. *Biochemistry 42*, 13558-13575.

41. Thorpe, I. F., and Brooks, C. L. (2005) Conformational substates modulate hydride transfer in dihydrofolate reductase. *Journal of the American Chemical Society 127*, 12997-13006.

42. Castillo, R., Andres, J., and Moliner, V. (1999) Catalytic mechanism of dihydrofolate reductase enzyme. A combined quantum-mechanical/molecular-mechanical characterization of transition state structure for the hydride transfer step. *Journal of the American Chemical Society 121*, 12140-12147.

43. Cummins, P. L., and Gready, J. E. (2003) Computational methods for the study of enzymic reaction mechanisms. II. An overlapping mechanically embedded method for hybrid semi-empirical-QM/MM calculations. *Journal of Molecular Structure: THEOCHEM 632*, 247-257.

44. Liu, H., and Warshel, A. (2007) The Catalytic Effect of Dihydrofolate Reductase and Its Mutants Is Determined by Reorganization Energies. *Biochemistry 46*, 6011-6025.

45. Hwang, J. K., King, G., Creighton, S., and Warshel, A. (1988) Simulation of Free-Energy Relationships and Dynamics of Sn2 Reactions in Aqueous-Solution. *Journal of the American Chemical Society 110*, 5297-5311.

46. King, G., and Warshel, A. (1989) A Surface Constrained All-Atom Solvent Model for Effective Simulations of Polar Solutions. *Journal of Chemical Physics 91*, 3647-3661.

47. Lee, F. S., and Warshel, A. (1992) A Local Reaction Field Method for Fast Evaluation of Long-Range Electrostatic Interactions in Molecular Simulations. *Journal of Chemical Physics 97*, 3100-3107.

48. Dams, T., Auerbach, G., Bader, G., Jacob, U., Ploom, T., Huber, R., and Jaenicke, R. (2000) The crystal structure of dihydrofolate reductase from Thermotoga maritima: Molecular features of thermostability. *Journal of Molecular Biology 297*, 659-672.

49. Sawaya, M. R., and Kraut, J. (1997) Loop and subdomain movements in the mechanism of Escherichia coli dihydrofolate reductase: Crystallographic evidence. *Biochemistry 36*, 586-603.

50. Dams, T., and Jaenicke, R. (1999) Stability and folding of dihydrofolate reductase from the hyperthermophilic bacterium Thermotoga maritima. *Biochemistry 38*, 9169-9178.

51. Ionescu, R. M., Smith, V. F., O'Neill, J. C., and Matthews, C. R. (2000) Multistate equilibrium unfolding of Escherichia coli dihydrofolate reductase: Thermodynamic and spectroscopic description of the native, intermediate, and unfolded ensembles. *Biochemistry 39*, 9540-9550.

52. Miller, G. P., and Benkovic, S. J. (1998) Stretching exercises - flexibility in dihydrofolate reductase catalysis. *Chemistry & Biology 5*, R105-R113.

53. Warshel, A., Chu, Z. T., and Parson, W. W. (1989) Dispersed Polaron Simulations of Electron-Transfer in Photosynthetic Reaction Centers. *Science 246*, 112-116.

54. Warshel, A., Chu, Z. T., and Hwang, J. K. (1991) The Dynamics of the Primary Event in Rhodopsins Revisited. *Chemical Physics 158*, 303-314.

55. Georlette, D., Blaise, V., Collins, T., D'Amico, S., Gratia, E., Hoyoux, A., Marx, J. C., Sonan, G., Feller, G., and Gerday, C. (2004) Some like it cold: biocatalysis at low temperatures. *Fems Microbiology Reviews 28*, 25-42.

56. Olufsen, M., Smalas, A. O., Moe, E., and Brandsdal, B. O. (2005) Increased flexibility as a strategy for cold adaptation - A comparative molecular dynamics study of cold- and warm-active uracil DNA glycosylase. *Journal of Biological Chemistry 280*, 18042-18048.

57. Papaleo, E., Riccardi, L., Villa, C., Fantucci, P., and De Gioia, L. (2006) Flexibility and enzymatic cold-adaptation: A comparative molecular dynamics investigation of the elastase family. *Biochimica Et Biophysica Acta-Proteins and Proteomics 1764*, 1397-1406.

58. Villa, J., Strajbl, M., Glennon, T. M., Sham, Y. Y., Chu, Z. T., and Warshel, A. (2000) How important are entropic contributions to enzyme catalysis? *Proceedings of the National Academy of Sciences of the United States of America 97*, 11899-11904.

Table 1. Force field parameters for simplified side chain model[a]

Residue	r^0	ε_1^0	ε_2^0	polarity
A	2.8	0.05	0.05	nonpolar
C	3.1	1.50	0.10	polar
D	3.4	1.50	0.21	polar
E	4.4	1.50	0.27	polar
F	4.1	0.16	0.16	nonpolar
H	3.8	1.50	0.33	polar
I	3.8	0.21	0.21	nonpolar
K	3.8	1.50	0.27	polar
L	3.5	0.21	0.21	nonpolar
M	3.8	0.21	0.21	nonpolar
N	3.3	1.50	0.21	polar
P	3.4	0.39	0.39	nonpolar
Q	3.7	1.50	0.27	polar

R	4.1	1.50	0.39	polar
S	2.9	1.50	0.10	polar
T	3.4	1.50	0.16	polar
V	3.5	0.16	0.16	nonpolar
W	4.4	0.45	0.45	nonpolar
Y	4.2	1.50	0.45	polar

[a] The energies are given in kilocalories per mole and distances in angstroms. ε_1^0 is used for interactions involving residues of the same polarity as well as interactions with ionized residues, and ε_2^0 is used for interactions involving residues of different polarity.

Table 2. Activation Energies at different runs in the EVB simulations[a]

	run 1	run 2	run 3	run 4	run 5	run 6	total map	ave run	obs
EcDHFR	12.8	14.1	14.6	16.1	15.9	15.9	14.7	14.9	13.5
TmDHFR	22.7	23.7	24.3	23.3	23.6	23.8	23.9	23.6	18.7

[a] The energies are given in kilocaries per mole. The FEP/US mapping was done in each individual run and in the total simulations (total map). "Ave run" designates the average activation energy of all the simulations and "observed" designates the experimental activation energy. The first 10% of data collected in each simulation was excluded from the FEP/US calculations.

Figure Legends

Figure 1. A schematic representation of the reaction catalyzed by DHFR. This reaction involves a hydride transfer from the NADPH coenzyme to the 7,8-dihydrofolate substrate.

Figure 2. X-ray structures of the mesophilic Dihydrofolate Reductase (monomer) from *Escherichia coli* (EcDHFR) and the hyperthermophilic Dihydrofolate Reductase (dimer) from *Thermotoga maritime* (TmDHFR). The 7,8-dihydrofolate substrate (in orange) and NADPH coenzyme (in blue) are represented in sphere model.

Figure 3. (a) Free energy landscapes for the EcDHFR and TmDHFR. The free energy surfaces were obtained by the FEP/US method and are represented in terms of the radius of gyration (Rg) and the root mean square deviation (rmsd) from the most stable structure. Energies are expressed in kcal/mol and distances in Å. (b) Free energy surfaces for the EcDHFR and TmDHFR obtained by long molecular dynamics simulations without any constraint.

Figure 4. EVB free energy profiles for the hydride transfer reaction of the EcDHFR (in blue) and the TmDHFR (in red). The profiles were obtained by averaging several EVB simulations as described elsewhere (*44*) and are given in terms of the EVB reaction coordinate ($\Delta\varepsilon = \varepsilon_2 - \varepsilon_1$).

Figure 5. Diabatic free energy functionals for EcDHFR (black) and TmDHFR (blue): (a) The total free energy functionals, (b) the solvent contribution and (c) the solute contribution. The figures display the functional after the minima have been shifted to the same heights. The intersection region is magnified in each case. As discussed in ref (*44*) the larger the reorganization (λ) the higher the activation barrier.

Figure 6. Free energy surface for EcDHFR (a) and TmDHFR (b). The surfaces are given in terms of the solute and solvent coordinates, which reflect the solute and solvent reorganization energies. The activation energies of the reaction are represented in EcDHFR and TmDHFR which are 15 kcal/mol and 23 kcal/mol, respectively. RS, TS and PS designate reactant state, transition state and product state, respectively.

Figure 7. Free energy surface as a function of the donor-acceptor distance (R(D-A)) and donor-hydrogen (R(D-H)) for the reaction catalyzed by the EcDHFR (a) and the TmDHFR (b). The contour lines are given in kcal/mol. RS, TS and PS designate reactant state, transition state and product state, respectively.

Figure 8. Dispersed polaron (DP) spectral distribution analysis, taking into account only the electrostatic contribution of the solvent, for the EcDHFR (in blue) and TmDHFR (in red). This analysis gives the projection of the modes of the protein system along the reaction coordinate

where the projection of each mode is represented by the corresponding reorganization energy

$$\lambda_i = \frac{1}{2}\hbar\omega_i\Delta_i^2 .$$

Figure 9. Relationship between the folding motions and the reactive modes in the reactant state of EcDHFR. The red arrows represent the vector of the motion along the reaction coordinate for the residues near or at the active site while with black arrows present the direction of the folding coordinate. The figure considers the direction of the folding coordinate starting from (a) a partially unfolded structure and (b) starting from the occluded structure (*49*). The angles between the two vectors are 96º and 93º in case (a) and (b), respectively.

Figure 10. Representation of the donor-acceptor distance versus the acceptor-hydrogen distance along the reactive downhill trajectories in EcDHFR (a) in TmDHFR (b). The figures show that both systems oscillate in the RS for a very long time until it generates the rare reactive trajectory that goes to the TS.

Figure 11. Average projection of the downhill trajectories in EcDHFR on the folding coordinate (in blue) and on the reaction coordinate (in red). The negative time is taken relative to the start of the downhill trajectory and its time reversal corresponds to a reactive trajectory from the RS. The projection has been evaluated while considering the EVB reacting region and its surroundings up to 5 Å from the EVB atoms (see text).

Figure 12. Normalized autocorrelation of $\Delta r_{RC} \cdot \delta r(t)_{RC}$ (in red) and $\Delta r_{FC} \cdot \delta r(t)_{RC}$ (in blue) for the EcDHFR system. The fact that the autocorrelation decays in a very short time indicates that it is very unlikely that the fluctuation along the folding coordinate would transfer energy to the reaction coordinate.

Figure 13. Schematic representation of the motion along the folding coordinate and the reaction processes in the mesophile and the thermophile. The motion starts from a partially unfolded straucture (a) moves to the folded structure (b) and then continues to the TS (c). The rates in the figure correspond to the overall time until a given process occurs rather than to the time of a reactive trajectory. The rates for the motion along the folding coordinate are somewhat arbitrary. Note also that the actual time for a reactive trajectory is ~1ps. The figure illustrates schematically the fact that the enzyme spends most of the time at the reactant state in both systems and that the motions are randomized by the thermal energy and lead to reactive trajectories only according to the Boltzmann probability of being in the TS.

Figure 1

DHF-H⁺ ... NADPH ... EcDHFR ... TmDHFR
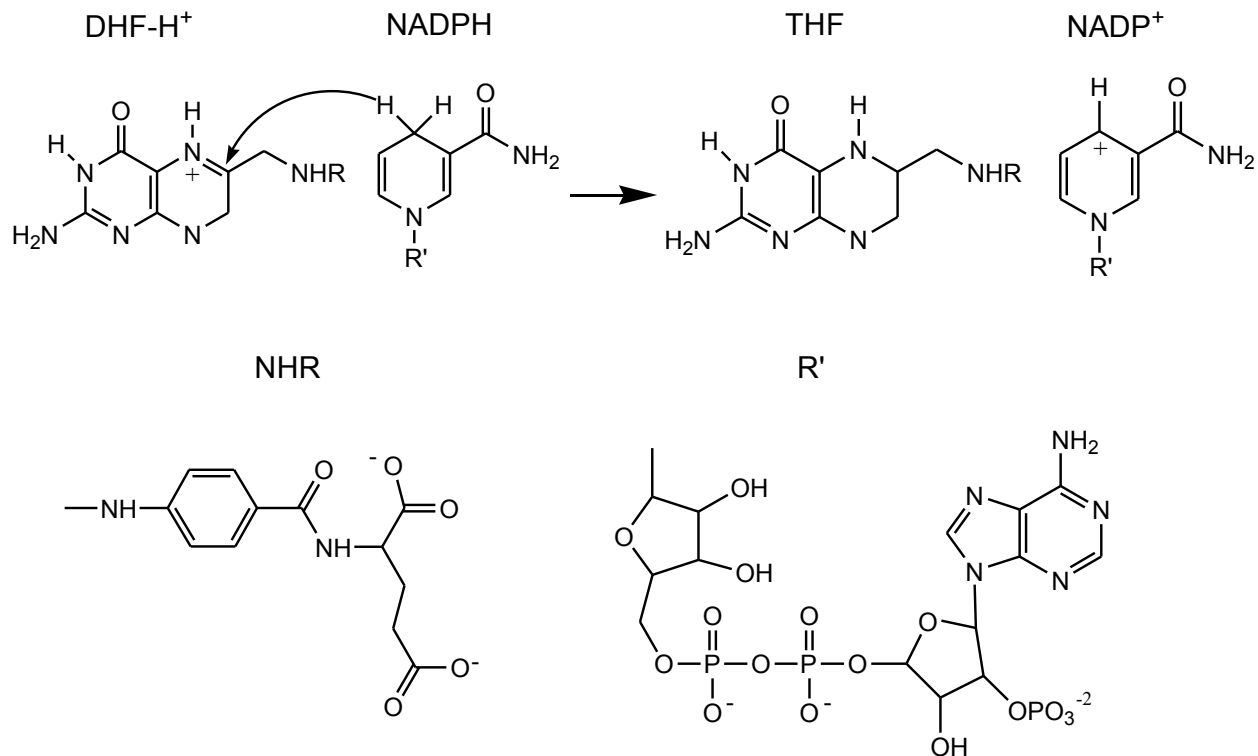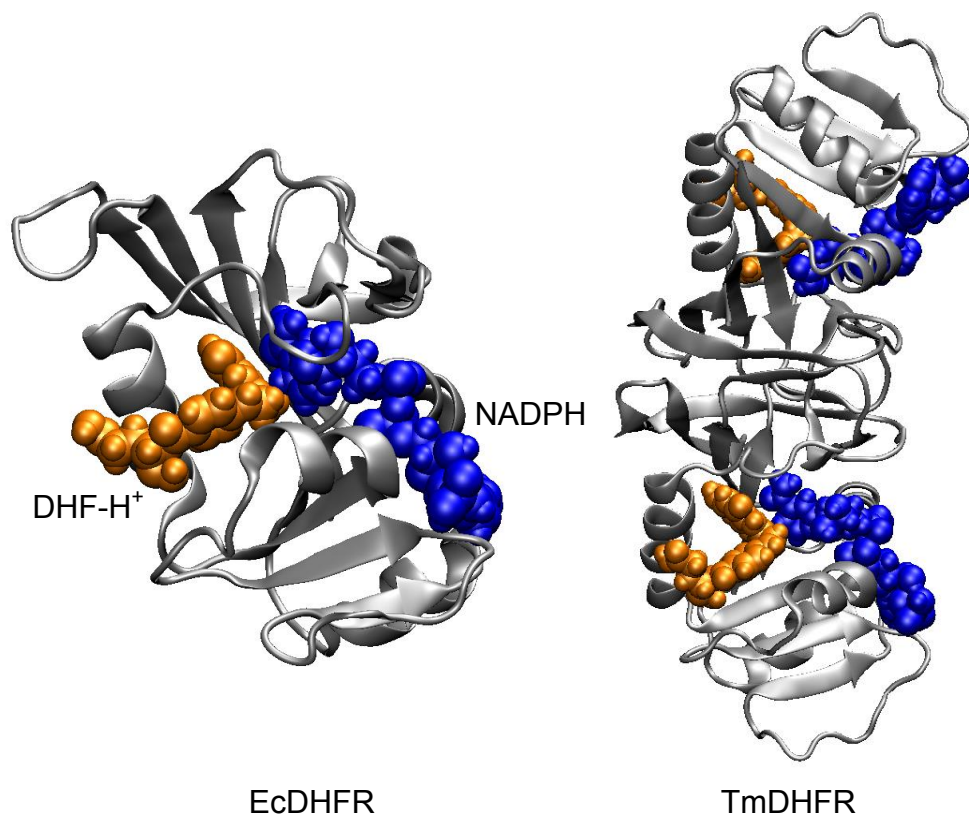
Figure 2

(a)
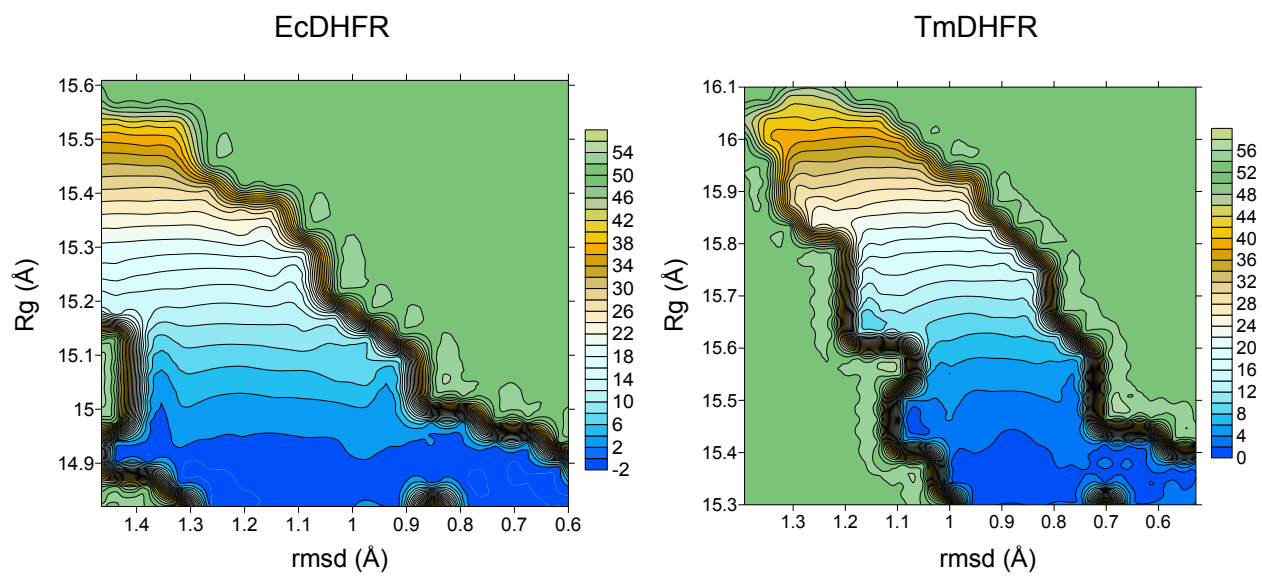
(b)
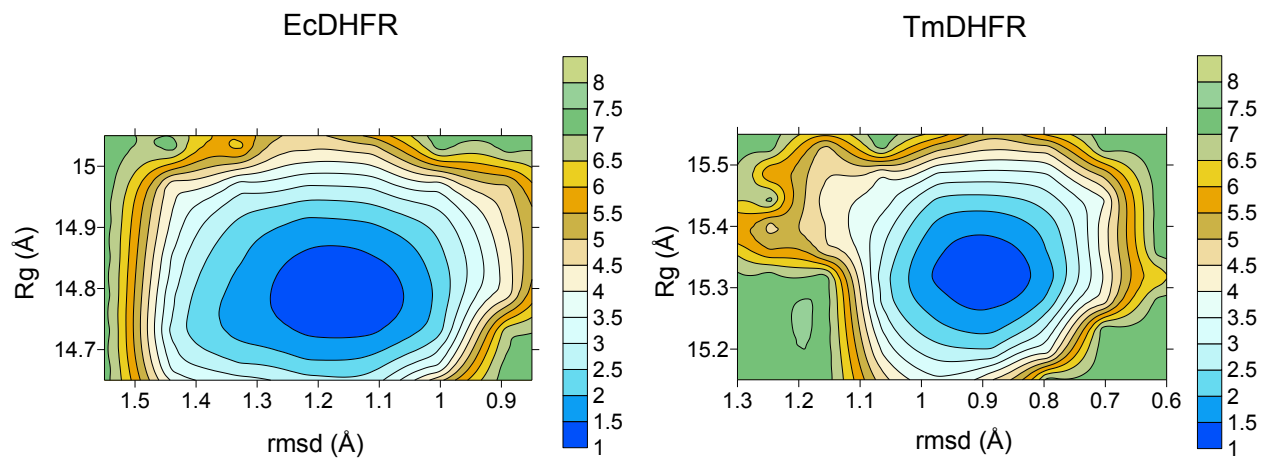
Figure 3

Figure 4

(a)

(b)

(c)

Figure 5

(a)

EcDHFR

(b)

TmDHFR

Figure 6

46

Figure 7

Figure 8

(a)

....▶ Reaction coordinate

....▶ Folding coordinate

49

(b)

····▶ Reaction coordinate

·····▶ Folding coordinate

Figure 9

(a)

EcDHFR

(b)

TmDHFR
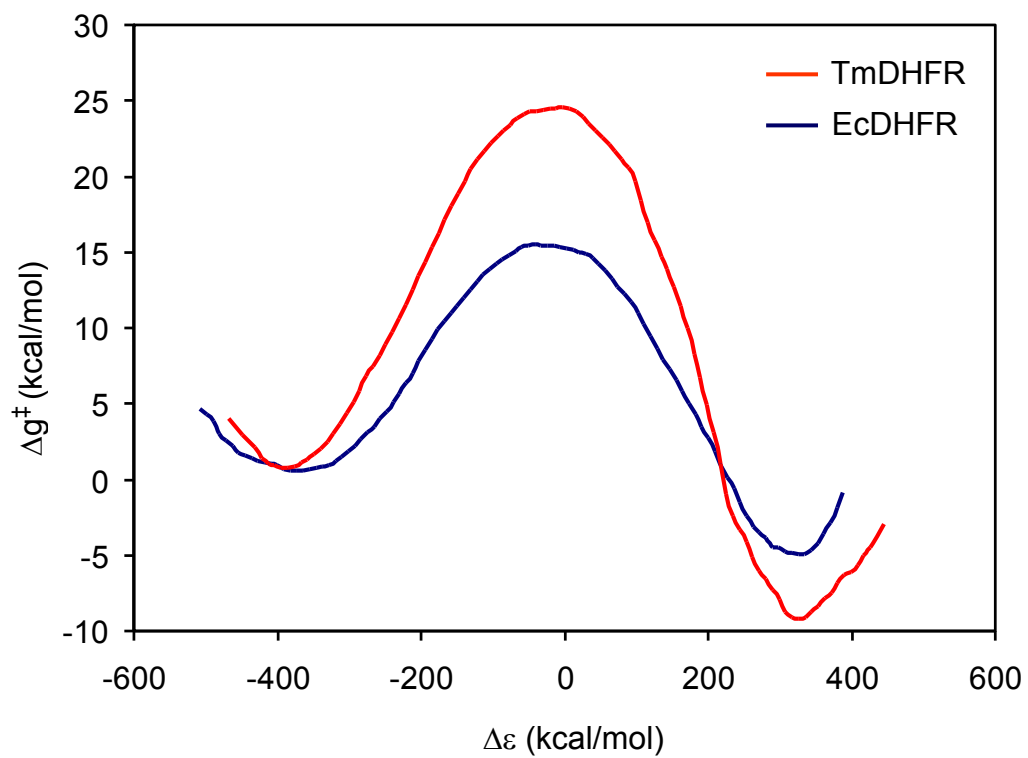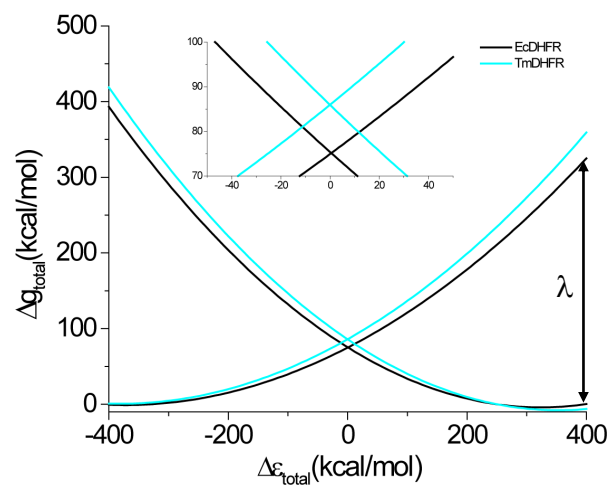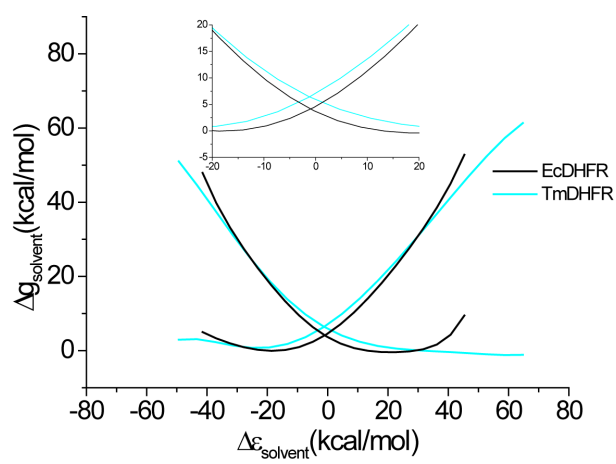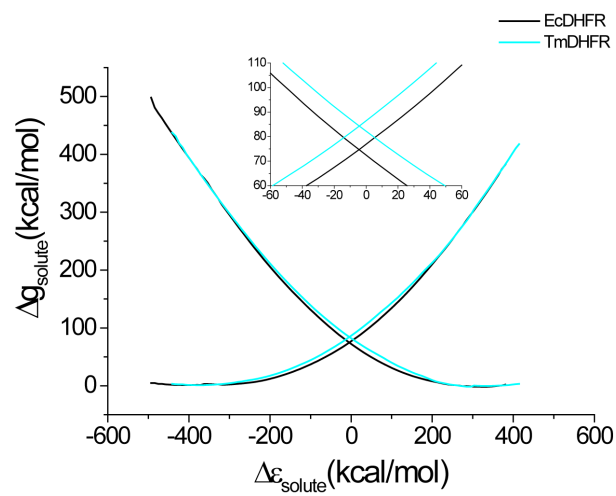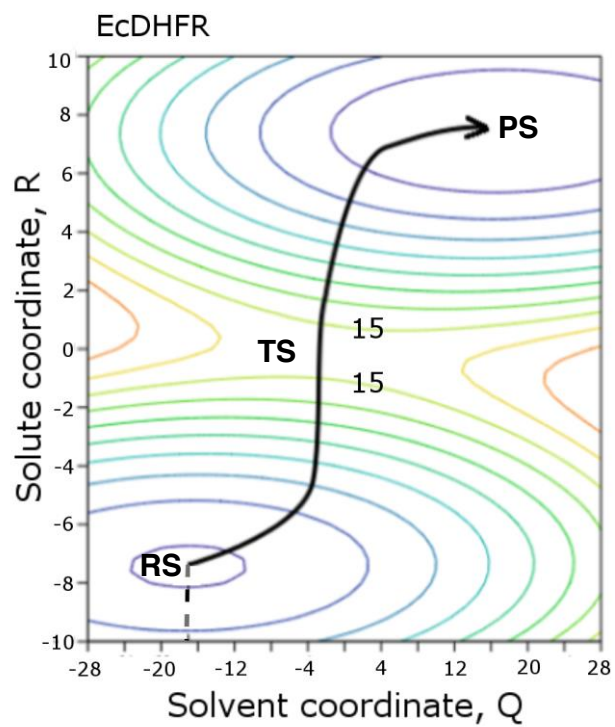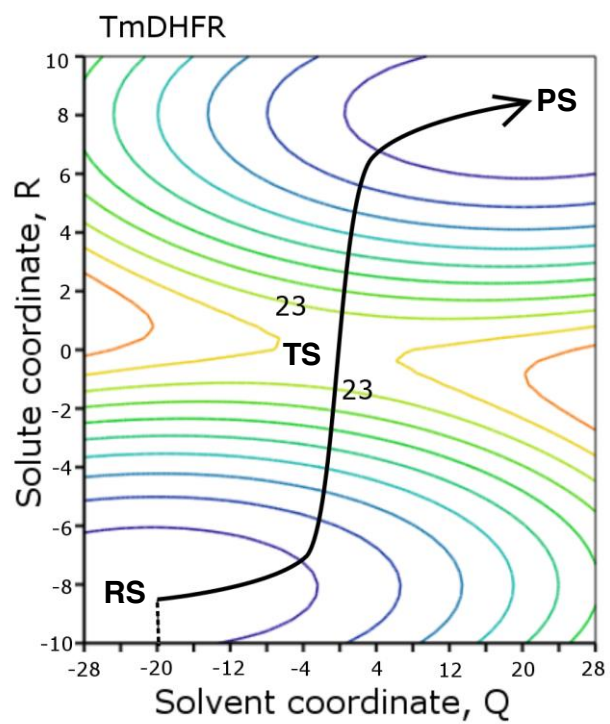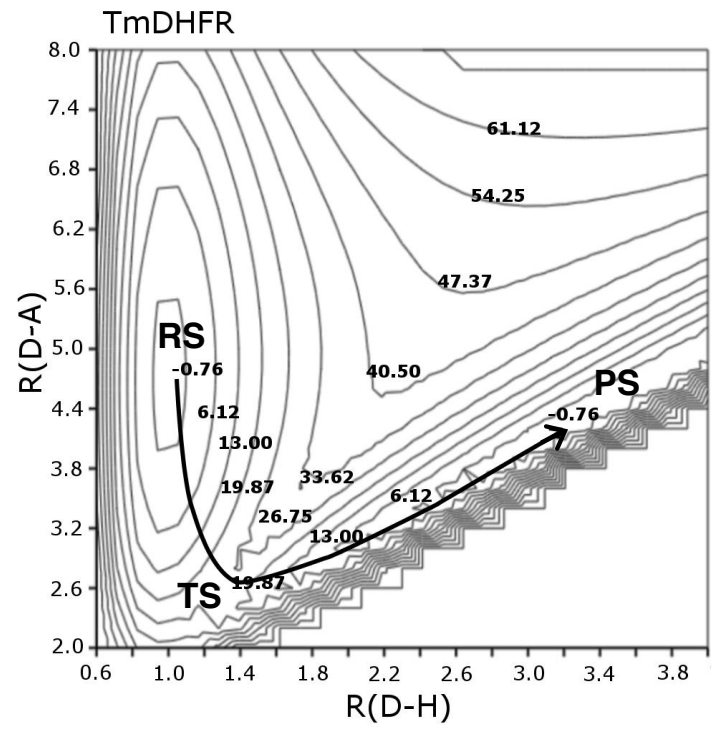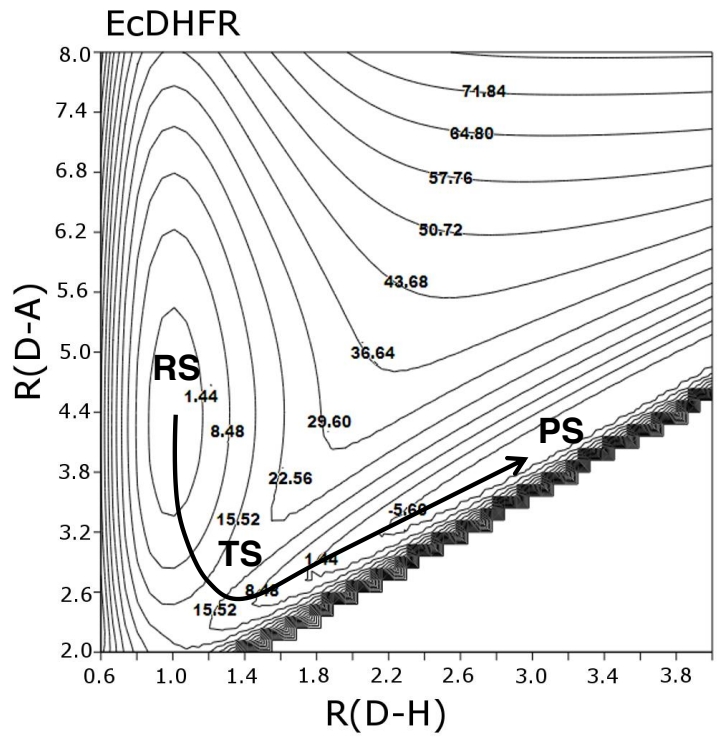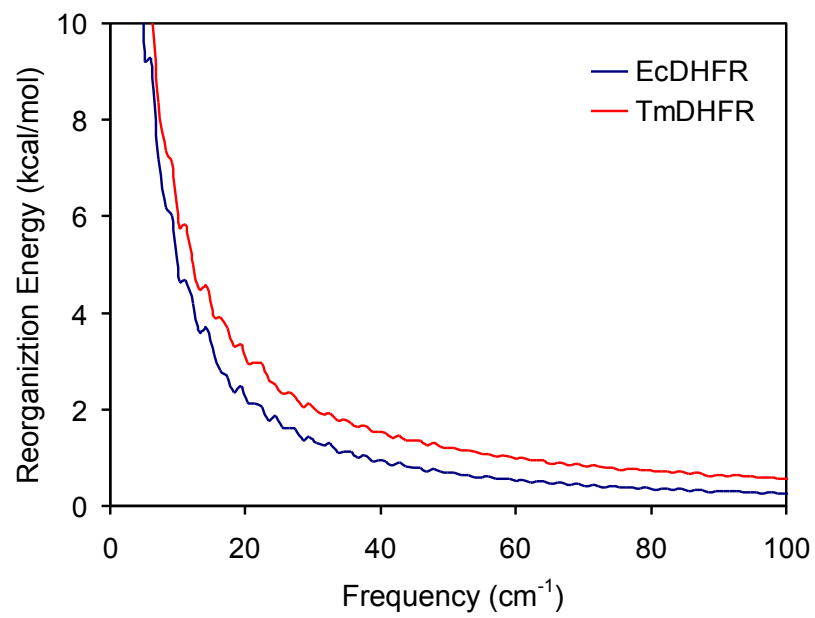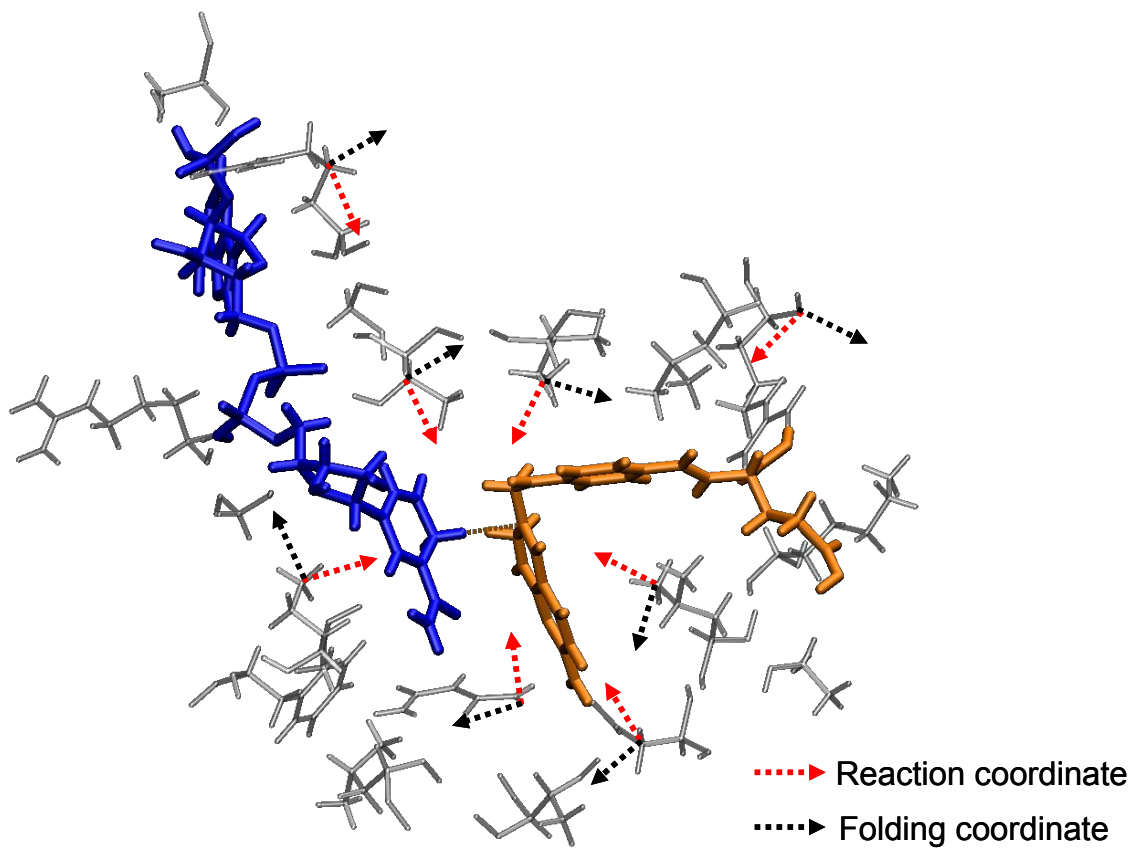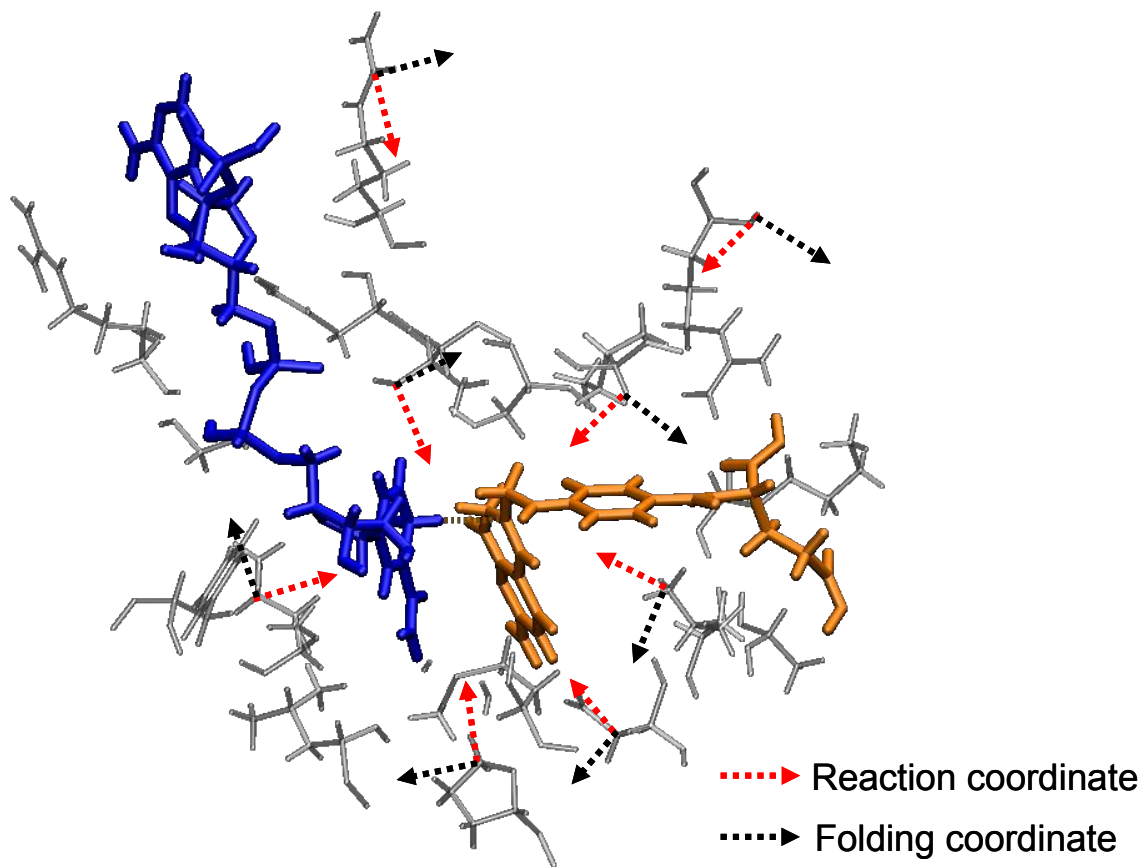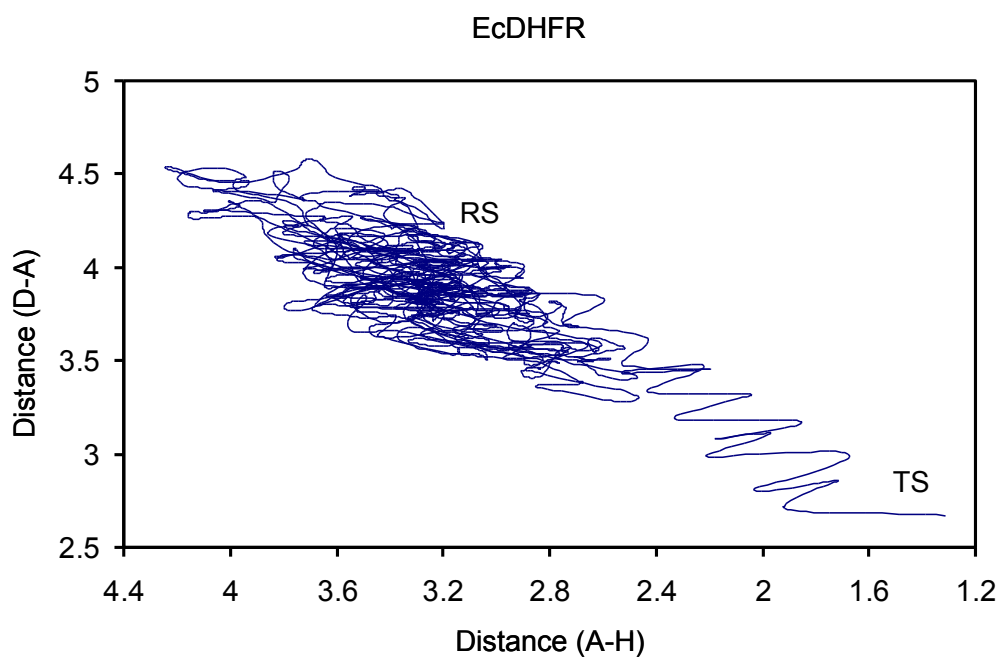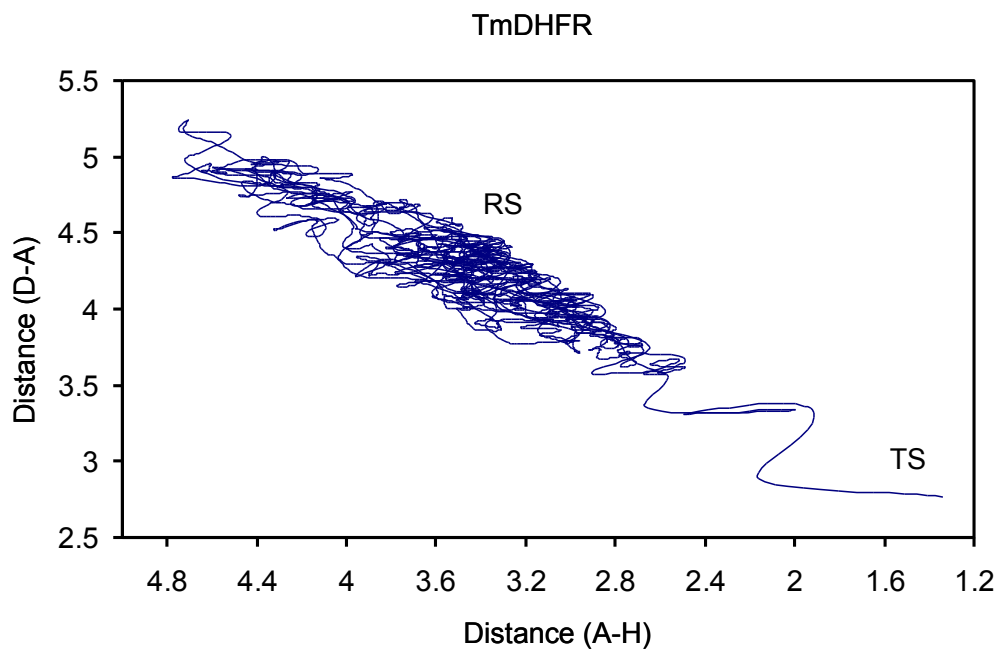
Figure 10

Figure 11

Figure 12

Mesophile

Thermophile

$k_{bc}^{Ms} \approx 10^2 \text{ s}^{-1}$

(c)

$k_{ab}^{Ms} \approx 10^9 \text{ s}^{-1}$

(a)

Folding coordinate

(b) Reaction coordinate

$k_{bc}^{Tm} \approx 1 \text{ s}^{-1}$

(c)

$k_{ab}^{Tm} \approx 10^{11} \text{ s}^{-1}$
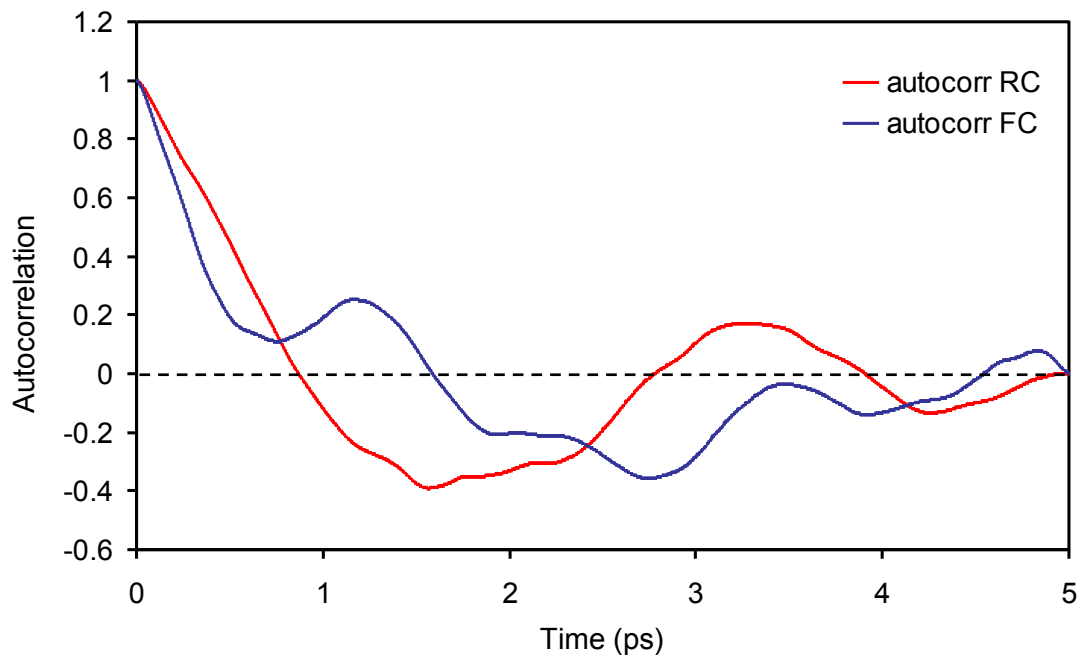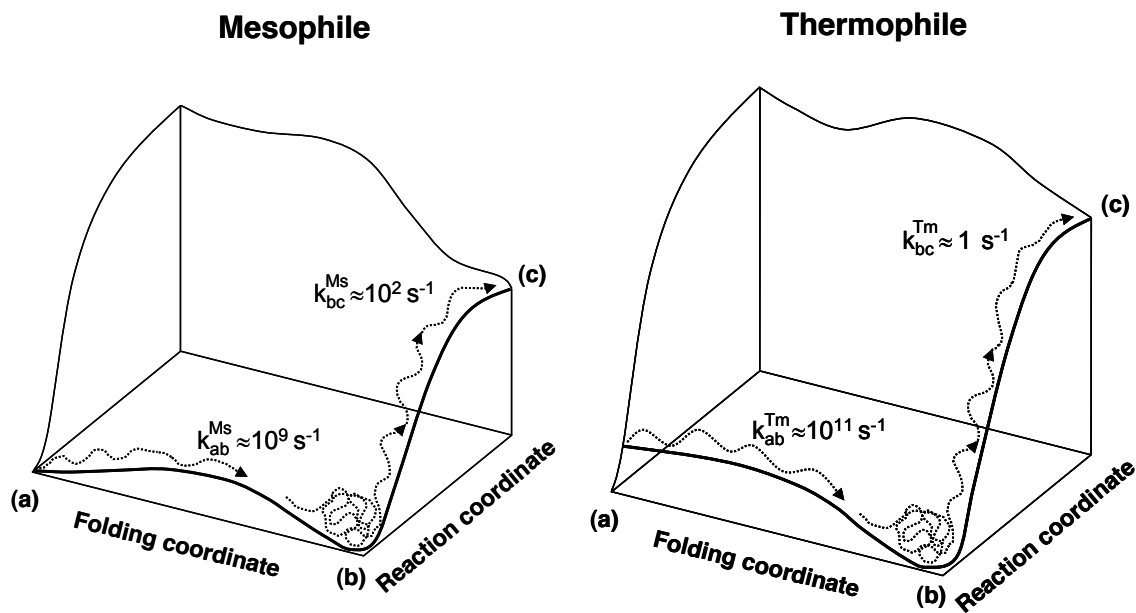
(a)

Folding coordinate

(b) Reaction coordinate
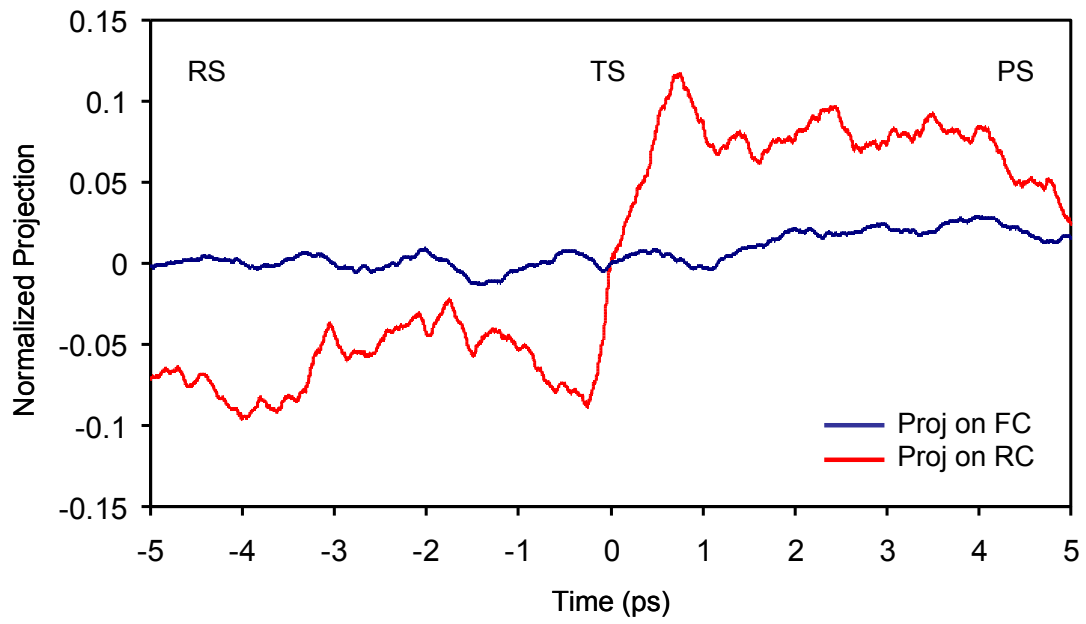
Figure 13

For Table of Contents Use Only

On the relationship between thermal stability and catalytic power of enzymes[†]

Maite Roca*, Hanbin Liu, Benjamin Messer and Arieh Warshel*

$$\Delta r_{FC}^{Tm} < \Delta r_{FC}^{Ms}$$

$$\Delta r_{RC}^{Tm} \approx \Delta r_{RC}^{Ms}$$

RC
FC