Enzymatic conformational fluctuations along the reaction coordinate of cytidine deaminase

RYAN C. NOONAN,[1] CHARLES W. CARTER, JR.,[2] AND CAREY K. BAGDASSARIAN[1]

[1]Department of Chemistry, The College of William and Mary, Williamsburg, Virginia 23187-8795, USA
[2]Department of Biochemistry and Biophysics, School of Medicine, University of North Carolina, Chapel Hill, North Carolina 27599-7260, USA

(RECEIVED January 17, 2002; FINAL REVISION February 28, 2002; ACCEPTED March 5, 2002)

Abstract

Analysis of the crystal structures for cytidine deaminase complexed with substrate analog 3-deazacytidine, transition-state analog zebularine 3,4-hydrate, and product uridine establishes significant changes in the magnitude of atomic-scale fluctuations along the (approximate) reaction coordinate of this enzyme. Differences in fluctuations between the substrate analog complex, transition-state analog complex, and product complex are monitored via changes in corresponding crystallographic temperature factors. Previously, we reported that active-site conformational disorder is substantially reduced in the transition-state complex relative to the two ground-state complexes. Here, this result is statistically corroborated by crystallographic data for fluorinated zebularine 3,4-hydrate, a second transition-state analog, and by multiple regression analysis. Multiple regression explains 70% of the total temperature factor variation through a predictive model for the average B-value of an amino acid as a function of the catalytic state of the enzyme (substrate, transition state, product) and five other physical and structural descriptors. Furthermore, correlations of atomic fluctuation magnitudes throughout the body of each complex are quantified through an autocorrelation function. The transition-state analog complex shows the greatest correlations between temperature factor magnitudes for spatially separated atoms, underscoring the strong ability of this reaction-coordinate species to "organize" enzymatic fluctuations. The catalytic significance for decreased atomic-scale motions in the transition state is discussed. A thermodynamic argument indicates that the significant decreases in local enzymatic conformational entropy at the transition state result in enhanced energetic stabilization there.

Keywords: Cytidine deaminase; conformational fluctuations; temperature factors; catalysis

Transition-state stabilization has long been recognized as the hallmark for the tremendous rate accelerations achieved during enzymatically catalyzed chemical transformations (Radzicka and Wofenden 1995; Mader and Bartlett 1997; Kraut 1988). Enzyme molecules are exquisitely designed to bind strongly to the transition states of chemical reactions, whereas ground-state interactions with substrates or products are energetically weaker. The fact that chemically stable transition-state analogs are bound much more tightly to enzymes than are substrates lends strong support to this idea. For *Escherichia coli* cytidine deaminase (CDA), the focus of this present work, the transition-state analog zebularine 3–4 hydrate dissociates from the enzyme with an equilibrium dissociation constant of $K_i = 1.2 \times 10^{-12}$ M, whereas $K_M = 5.0 \times 10^{-5}$ M for the substrate cytidine (Frick et al. 1989; Radzicka and Wolfenden 1995). Via a well-known thermodynamic cycle, the hypothetical value for the dissociation constant of the actual transition-state–enzyme complex can be calculated (Wolfenden 1972). In terms of the enzymatic affinity for substrate (K_M) and the ratio of rate constants for noncatalyzed and catalyzed reactions, the affinity for transition state is found to be

Reprint requests to: Carey K. Bagdassarian, Department of Chemistry, The College of William and Mary, P.O. Box 8795, Williamsburg, Virginia 23187-8795, USA; e-mail: ckbagd@wm.edu; fax: (757) 221-2715.

Article and publication are at http://www.proteinscience.org/cgi/doi/10.1110/ps.0202102.

$K_{TX} = (k_{non}/k_{cat})K_M$. For CDA, $k_{non}/k_{cat} = 2.5 \times 10^{-12}$ (Frick et al. 1987), reflecting the dramatic preferential binding of enzyme to transition state.

These ideas comprise a static picture of catalysis: The barrier to chemical reaction is reduced through enzyme–transition-state binding interactions. However, a protein molecule is dynamic, exploring numerous conformational states through its thermal fluctuations (McCammon and Harvey 1987; Brooks et al. 1988). These "breathing" modes present a fascinating temperature dependence. Below a transition temperature, the atomic fluctuations of the protein are harmonic, with anharmonic motions setting in above this temperature (Loncharich and Brooks 1990; Rasmussen et al. 1992; Hagen et al. 1995; More et al. 1995; Ringe and Petsko 1999). Molecular dynamics simulations of enzymatic fluctuations, in addition to a host of experimental techniques, highlight features of these thermal breathing modes (Mulholland et al. 1993; Eurenius et al. 1996; Karplus and Ichiye 1996; Verma et al. 1997; Melchionna et al. 1998; Radkiewicz and Brooks 2000). For example, work on the fluctuation dynamics of the homodimeric enzyme superoxide dismutase has revealed an asymmetry between the two monomers (Melchionna et al. 1998). On the time-scale of the simulation (up to 900 ps), one of the subunits has larger average fluctuations than the other, prompting speculation as to whether an approaching substrate molecule encounters a symmetric or asymmetric dimer.

As stressed by Cannon et al. (1996), catalysis itself is a dynamical transformation that involves molecular rearrangements from substrate to transition state to product. These investigators emphasize that the strong geometrical and electrostatic complementarity between active-site amino acids and the transition state eliminates otherwise slow solvent reorganizational events and thereby accelerates the chemistry. In fact, they suggest that dynamical events within the active site of the enzyme-reaction-coordinate complex play a crucial role in catalysis, and recently, even the idea of transition-state stabilization has been challenged (Bruice and Benkovic 2000). Klinman's group has found that enzymatic fluctuations are crucial in promoting the chemistry in a thermophilic alcohol dehydrogenase (Kohen et al. 1999). Specifically, as the temperature is decreased below 30°C, the increased rigidity of the thermophile hinders a hydrogen tunneling event. Balabin and Onuchic (2000) have shown that conformational fluctuations are crucial for electron transfer in a bacterial photosynthetic reaction center. Recent quantum dynamical calculations on liver alcohol dehydrogenase show that the activation barrier to reaction is strongly sensitive to the atomic-scale conformation of the enzyme (Alhambra et al. 2001; Billeter et al. 2001).

An important question arises: Is the three-dimensional molecular architecture of the enzyme-substrate complex engineered to favor catalytically productive atomic-scale motions while restricting "useless" or "stray" fluctuations? Furthermore, how do magnitudes of atomic-scale motions and their correlations change as the enzyme-ligand system proceeds along the reaction coordinate?

Here we address fluctuation patterns along the chemical reaction coordinate once the substrate molecule is sequestered in the active site—and not large-scale conformational changes as the enzyme adjusts by induced fit to accommodate substrate (Wierenga et al. 1992; Ilyin et al. 2000). The latter conformational rearrangements are known to be correlated with enzyme function (Rasmussen et al. 1992), but less evidence exists to support the role of conformational dynamics within the active site itself in bond making and breaking events (Kohen et al. 1999; Ringe and Petsko 1999).

In the present paper, we examine correlations between the differential affinity of an enzyme for substrate analog, transition-state analog, and product and the magnitude of atomic-scale fluctuations in the respective enzyme-ligand complexes. Previously, we showed that thermal agitation in the active site of *E. coli* CDA substantially decreases when the enzyme is bound with transition-state analog (Alper et al. 2001). A detailed analysis of this conclusion, as well as statistical justification for its validity, is offered here.

E. coli CDA is a homodimer with two active sites that catalyzes the hydrolytic deamination of cytidine into uridine (for review, see Schramm and Bagdassarian 1999). Several crystal structures will be studied: that of the enzyme complexed with (1) substrate analog 3-deazacytidine, (2) transition-state analog hydrated pyrimidin-2-one ribonucleoside (also called zebularine 3-4 hydrate), (3) product uridine, and (4) a fourth inhibitor, 3,4-dihydrozebularine (Betts et al. 1994; Xiang et al. 1995, 1996, 1997; Carter 1995). The structures of these ligands, along with the chemical reaction coordinate, are shown in Figure 1. The enzyme complexed with 5-flourozebularine 3,4-hydrate (Betts et al. 1994) provides a second transition-state structure, closely related to that of zebularine hydrate (see Fig. 1). The transition-state analog zebularine 3,4-hydrate (see Fig. 1) binds to CDA with an inhibition constant $K_i = 1.2 \times 10^{-12}$M. Replacement of of the -OH group of this analog with a hydrogen atom to give 3,4-dihydrozebularine decreases the affinity of the enzyme for dihydrozebularine to $K_i = 3.0 \times 10^{-5}$M (Frick et al. 1989). The fluorine derivative of zebularine 3,4-hydrate has $K_i = 3.9 \times 10^{-11}$M (Frick et al. 1989). For the product uridine, $K_i = 2.5 \times 10^{-3}$ M (Radzicka and Wolfenden 1995). Substrate cytidine is reported with $K_M = 5 \times 10^{-5}$M (Radzicka and Wolfenden 1995). Finally, the substrate analog 3-deazacytidine is not expected to bind as tightly as does cytidine. The sequence of enzyme-inhibitor complexes 3-deazacytidine, zebularine hydrate, and uridine approximates the reaction coordinate from substrate to transition state to product, and provides "snapshots" for the

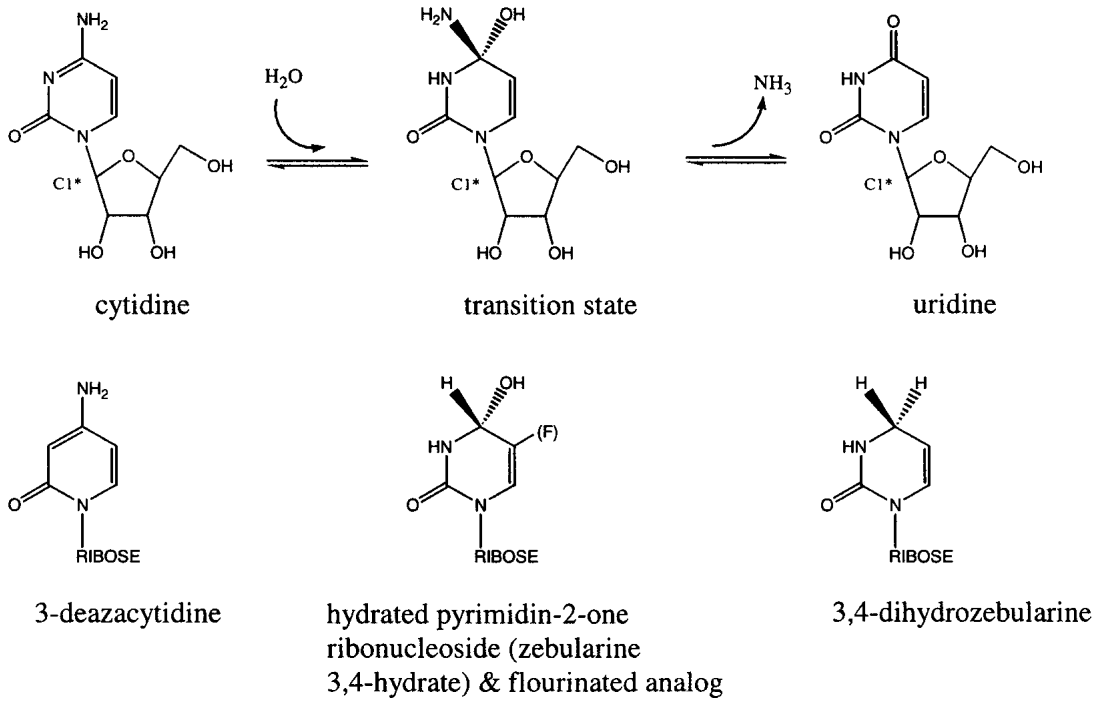

Fig. 1. Reaction coordinate for cytidine deaminase and reaction-coordinate analogs.

changing atomic-scale conformation of the enzyme as the reaction proceeds.

In the present work, we compare sets of crystallographic temperature factors for the enzyme-inhibitor complexes that approximate the chemistry under the assumption that they contain significant information about structural fluctuations. Surprisingly, we find that CDA temperature-factor values differ according to several physical features derived from the structures of the complexes themselves, and that in the context of these sources of variation, the state along the reaction coordinate for deamination (substrate, transition state, or product) is among the most significant predictor of the overall variation.

A related goal of this work is calculation of fluctuation correlations within each enzyme-inhibitor complex. Explicitly, how do correlations in fluctuation magnitude decay with increasing spatial separation between atomic loci under consideration? Fluctuation correlations are found to be greatest in the transition state, underscoring the ability of a transition-state analog to "organize" active-site thermal disorder.

Results and Discussion

Theoretical approach

The crystallographically derived atomic coordinates and temperature factors for CDA complexed with 3-deazacytidine (1ALN), zebularine 3-4 hydrate (1CTU), uridine (1AF2), and 3,4-dihydrozebularine (1CTT; Carter 1995; Xiang et al. 1995, 1996, 1997) are available from the Protein Data Bank. For simplicity, the first three species are referred to as substrate, transition state, and product, with the understanding that substrate means the enzyme-substrate-analog complex (that of the enzyme with 3-deazacytidine) and transition state refers to the enzyme–transition-state analog complex (enzyme complexed with zebularine 3–4 hydrate). The fourth inhibitor complex will be called dihydrozebularine, and the fluorinated-zebularine hydrate complex (Betts et al. 1994) will be referred to as F-transition state. Fluctuation magnitudes for substrate, transition state, and product are compared through analysis of the temperature factors of the enzyme-analog complexes, as justified shortly. The average temperature factor of all atoms comprising each amino acid residue is used to simplify the data for study of fluctuations along the reaction coordinate. Because our conclusions are unchanged by averaging over all atoms of an amino acid or just over those of the side-chain, the all-atom results are presented.

The five CDA-inhibitor complexes constitute an unusually coherent structural reaction profile. Moreover, they were crystallized, and their structures determined under identical or near identical conditions (Betts et al. 1994; Carter 1995; Xiang et al. 1995, 1996, 1997). The structure of the 5-flourozebularine hydrate complex (Betts et al. 1994) was used to initiate refinements of the 3-deazacytidine, zebularine hydrate, and 3,4-dihydrozebularine complexes. The uridine complex was initially refined against the

zebularine hydrate structure. Finally, the CDA crystal structures are all highly isomorphous, so differences in mean thermal parameters are more likely to be intrinsically related to enzyme-ligand interactions and not to unrelated consequences of, for example, different crystal packing arrangements. Therefore, with a good deal of confidence, we can quantify the magnitudes of thermal fluctuations along the reaction coordinate by studying atomic temperature factors derived from the crystallographic analysis.

The isotropic Debye-Waller factor, B, represents the total coordinate variance, $\langle \Delta r^2 \rangle = B/8\pi^2$, for atoms in a crystal, with a model based on the assumption that all variance is independent and spherically symmetric. These assumptions are patently false, because B represents both dynamic and static disorder and because atomic motions are strongly correlated within side-chains and along secondary structural elements (McCammon and Harvey 1987; Karplus and Petsko 1990). Temperature factors have nonetheless proved adequate to support numerous valid conclusions regarding atomic motion. Variation of B with temperature was, for example, used to show heterogeneity in side-chain orientation (Frauenfelder et al. 1979). It is widely appreciated that ligand binding is coupled to protein stability and, in particular, to the magnitude of observed coordinate variances, and that binding leads in general to lower refined B-values. Structural biologists routinely adduce variation in B-values, in a qualitative sense, as evidence for functional interpretations of atomic motion.

It, therefore, seemed reasonable to evaluate variations in B quantitatively over the series of refined CDA structures, which constitute a series of complexes closely related to the structural reaction profile, together with a significant amount of redundancy. We compare the effect of the presumed reaction coordinate to other sources of variation in the observed B-values of proteins. Thus, our comparison scheme offers the opportunity to investigate the variation of B-values quantitatively, by using multiple regression models, and by using Student *t*-testing to evaluate the significance of different effects. Additionally, the refined structure of the fluorinated zebularine complex, compared with that of the unfluorinated transition-state analog zebularine hydrate, provides an important estimate of the error in crystallographic determination of the thermal parameters.

Because all the CDA-ligand complexes are structurally very similar, those 29 amino acids spatially closest to the bound substrate analog in the enzyme-substrate complex are used to define the canonical set of amino acids used for multiple regression analysis. The inhibitor-to-inhibitor distance in all of the homodimeric CDA complexes is ~22 Å. Restricting the analysis to these 29 amino acids, the most distant of which is within 10 Å of one of the bound inhibitors, ensures an unambiguous study of the associated active site. Various factors, derived from the overall physical structure of the enzyme itself, may cause variation between amino acid B-values within a single complex. Possible such factors include the degree of solvent exposure and the distance between a residue and various zones within the protein that might limit atomic displacements, such as nonpolar core regions or the inhibitor. Each of these factors is expected to influence the temperature factor of a given residue: Amino acids closer to the inhibitor, those buried away from the surface, and those closer to the nonpolar core regions of the protein will likely have lower B-values. Additionally, the catalytic state of the enzyme (substrate, transition state, or product) contributes to the magnitude of the average temperature factor of a residue (Alper et al. 2001).

We used multiple regression to formulate a new model for amino acid average B-values. The proportion of temperature factor variation that is correlated with the catalytic state (State: 1 = substrate; 2 = transition state; 3 = product) and its statistical significance, relative to other possible determinants, is thereby evaluated. The best model encodes determinants including which of the two subunits donates a given residue (Subunit: 1; 2), the fraction of the amino acid surface area that is exposed to solvent (Surf), and distances from the bound ligand (Dist) and from the centroids of nonpolar cores (D_Core1 and D_Core2) previously identified for the two domains in the CDA monomer (Navaratnam et al. 1998). In short, the model predicts the average temperature factor $\langle B \rangle$ of an amino acid residue as a function of State, Subunit, Surf, Dist, D_Core1, and D_Core2. Details concerning the generation and evaluation of models are included in Materials and Methods.

For auto-correlation calculations, all atomic coordinates and corresponding temperature factors are used: The fluctuation pattern of an enzyme-inhibitor complex is approximated by the spatial distribution of temperature factors for each atom in each residue. For any two atoms in a given enzyme-ligand complex, fluctuations are correlated if the temperature factors of both atoms are either high or low (compared with the mean temperature factor over all atoms). As elaborated more fully in Materials and Methods, an auto-correlation function is applied to each enzyme-inhibitor complex to quantify fluctuation-magnitude correlations within the body of each complex. The auto-correlation function stems from a combination of those used by Moran (Cliff and Ord 1973) and by Wagener et al. (1995):

$$I(a) = \frac{n}{L} \frac{\sum_{i,j} (f_i - \bar{f})(f_j - \bar{f})}{\sum_{i} (f_i - \bar{f})^2}. \quad (1)$$

I(a) is the fluctuation correlation calculated from all atom-pairs in an enzyme-ligand complex separated by distance a.

f_i is the B-factor of atom i in the enzyme-inhibitor complex of interest; f_j is the value at atom j, which is separated from i by distance a; and $\bar{f}$ is the overall average temperature factor for the complex.

Fluctuation magnitudes along the reaction coordinate and multiple regression analysis

Figure 2a–c shows the distributions of some relevant residues around the three reaction-coordinate species, along with their average B-factors. This figure is based on those of Xiang et al. (1995, 1996, 1997) and features residues H-bonded or in hydrophobic contact with ligand and those coordinated to zinc.

A striking feature is immediately apparent: The average temperature factor for the substrate analog is 27 $Å^2$ and decreases for the transition-state analog to 15 $Å^2$, only to increase again in the product to 28 $Å^2$. This decrease in temperature factor at the transition state is corroborated by the 17 $Å^2$ value for the F-transition–state analog. As the transition state is accessed along the reaction coordinate, thermal fluctuations in the bound reaction-coordinate species decrease. Furthermore, examination of active site amino acids reveals smaller thermal fluctuations in the transition-state complex compared with either ground state, whereas the substrate and product active sites are similar in fluctuation magnitude. A study of hypoxanthine-quanine phosphoribosyltransferase is in accord with our results (Wang et al. 2001). Transition-state analog binding significantly tightens the catalytic site, as assessed by hydrogen/deuterium exchange experiments.

Of the set of 29 amino acids (which includes the residues shown in Fig. 2) closest to the bound inhibitors, 23 are less thermally agitated in the transition state compared with substrate. Furthermore, for 19 of these, thermal motion in the transition state is decreased by >14%. This set of 19 amino acids is particularly interesting because, for it, substrate and product fluctuations on average are remarkably similar, as are thermal motions in the two transition states. For these 19 amino acids, Figure 3 shows the strong fluctuation similarities between the two transition states, between the two ground states, and the greater differences between transition and ground states. It is tempting to speculate that these amino acids—significantly constrained in the transition state but equally noisy in the two ground states—undergo correlated motions in the ground-state complexes that are linked to locking the transition-state species within the active site.

In the transition-state complex, the average temperature factor for these amino acids is 18 $Å^2$; values for substrate and product amino acids are essentially identical at 23.5 $Å^2$ and 23.3 $Å^2$, respectively. On average, then, the ground-state complexes are 27% more noisy than transition state. The average fluctuation magnitude for F-transition state (17 $Å^2$) is close to that of unfluorinated transition state.

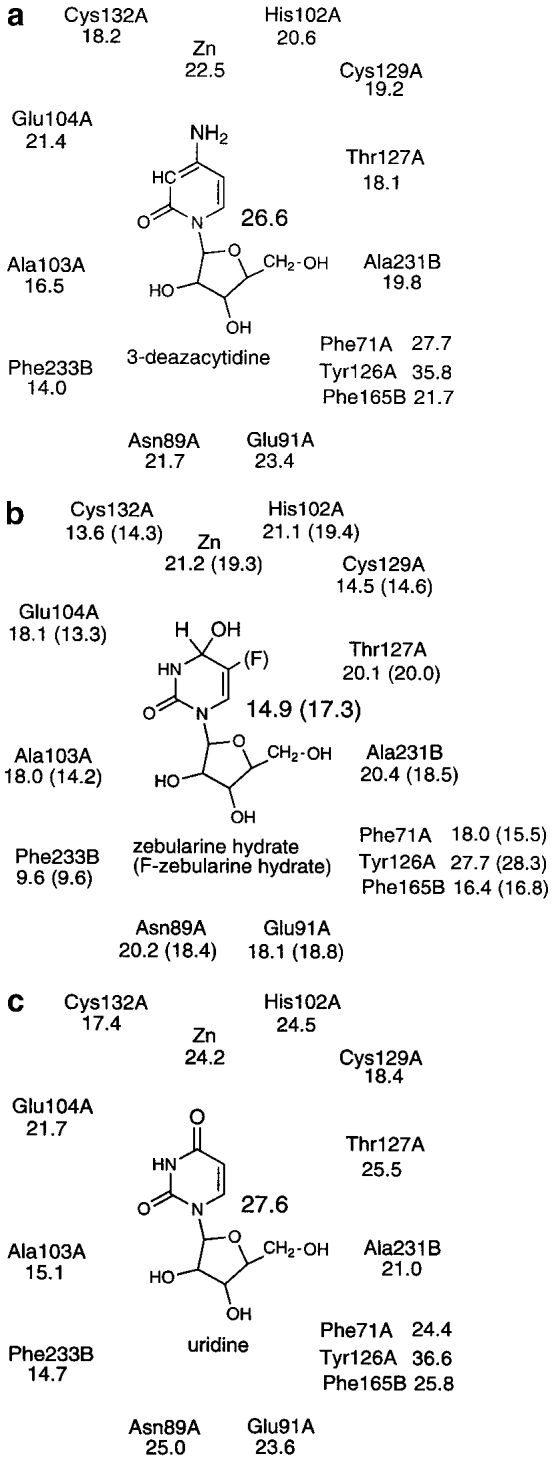

Fig. 2. (*a*) Average temperature factors for active-site amino acids and bound substrate analog. (*b*) Temperature factors for active-site amino acids and bound transition-state analog. Data for F-transition state in parentheses. (*c*) Average temperature factors for active-site amino acids and bound product.

We estimate inherent errors in crystallographic temperature factors from the root mean square difference between the two different but structurally similar transition-state ana-

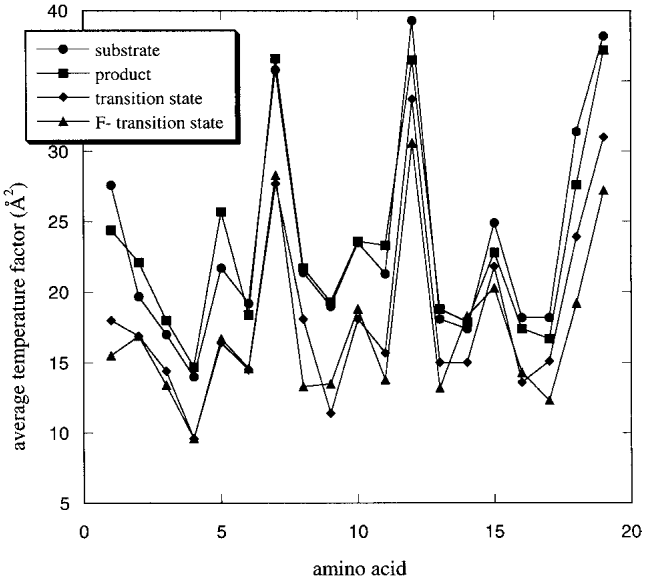

Fig. 3. Average temperature factor per amino acid for substrate, product, transition-state, and F-transition-state enzyme complexes. Data shown for 19 amino acids that are >14% noisier in the substrate complex than in transition state. As the amino-acid index increases (from 1 to 19), the distance of that residue to bound inhibitor increases.

log complexes. This value, $\sigma = 2.4$ Å^2, for those 19 amino acids markedly less noisy in the transition state is less than half that of the root mean square difference (5.4 Å$^2 = 2.25\sigma$) between transition state and either product or substrate. For comparison of product and substrate, the root mean square difference is 1.9 Å^2 and lies within the estimate of error from the two transition states. Note that because there are probably real differences between the two transition-state complexes, the value for σ may be an overestimate.

Figure 2 reveals an important feature: Temperature factors for both substrate analog and product uridine are generally greater than those of the surrounding amino acids, whereas thermal motion in the transition-state analogs is either less than or comparable to that of most surrounding residues. This indicates that relative to the ground states, the "fit" of the transition state to the active site of the enzyme improves significantly and in subtle ways not readily apparent to the observer. Furthermore, although the substrate and product complexes on average are equally thermally agitated, several residues in Figure 2 show considerably increased thermal fluctuations in the latter, with a possible correlation to product release. For example, the average B-factor of Thr127A is 41% greater in the product than in the substrate complex. Thr127A, in the actual chemical reaction coordinate, is presumptively H-bonded to the leaving amino group of the transition state (see Fig. 1). We suggest that the global enzymatic active site responds to structural changes in the bound reaction-coordinate species in a way that results in transiently increased mobility of Thr127A immediately after the transition state, thus facilitating release of the leaving group. Note that for Phe71A, Phe233B, Cys129A, Cys132A, Glu91A, Tyr126A, and Phe165B—all markedly less thermally agitated in the transition state—the B-factors, with the exception for those for Phe71A, in both transition states are very similar. These side-chains show reduced atomic fluctuations as the fit with the ligand reaches its optimal configuration.

The enzyme molecule responds to the demand of nestling around successive reaction-coordinate species by using correlated motions that have not as yet been completely understood. The average behavior of this extraordinarily complex "many-body" problem is highlighted by the few representative amino acids focused on.

Table 1 summarizes the analysis of variance for the multivariate model that best predicts amino acid average temperature factors in terms of the physical descriptors described above and in Materials and Methods. This model

$$\begin{aligned}\langle B\rangle_{calc} = {} & \beta_0 + \beta_{State} * State + \beta_{State-State} * State * State \\ & + \beta_{Subunit} * Subunit + \beta_{Subunit-Surf} * Subunit * Surf \\ & + \beta_{Dist} * Dist + \beta_{D\_Core1} * D\_Core1 \\ & + \beta_{Subunit-D\_Core2} * Subunit * D\_Core2 \\ & + \beta_{D\_Core1-D\_Core2} * D\_Core1 * D\_Core2.\end{aligned}$$

explains ~70% of the total variation in overall temperature factors for the group of 29 amino acids closest to the bound inhibitors. The remaining 30% of the variation not explained by the model is caused in unknown part by the inherent errors in the crystallographic temperature factor values, so this "explained" fraction of the variance is substantial. Coefficients in the model have *t* test probabilities between 10^{-3} and 10^{-11} and are therefore highly significant, as is the F-ratio test for the model as a whole. Thus, although each of the factors might be expected to contribute to the variance of $\langle B\rangle$, Table 1 provides quantitative evidence regarding their relative proportions.

There are three distinct kinds of effects separated in Table 1 by the line spacing. The first group shows that the catalytic state is a very significant variable and that $\langle B\rangle_{calc}$ is approximately quadratic in the reaction coordinate variable, consistent with the minimum value observed in the transition state. This effect accounts for ~10% of the overall variation in $\langle B\rangle$. The second group involves the relative solvent exposure of the side-chains. The third involves the significant effects of distances between the amino acids and three locations that have low $\langle B\rangle$, namely, the inhibitor and the centroids of the nonpolar clusters in each of the two domains in the monomer. The latter two effects account for roughly equal portions of the remaining variation in $\langle B\rangle$.

The first two categories of predictors are illustrated graphically in Figure 4. The minimum value of the overall

Table 1. *Analysis of variance for multivariate temperature-factor model*

Variable	Coefficient	STD Error	T	P (2-Tail)
Constant	33.931	2.870	11.822	0.10E-14
State	−6.310	0.965	−6.536	0.99E-09
State*State	1.836	0.314	5.850	0.31E-07
Subunit	−31.292	4.733	−6.612	0.67E-09
Subunit*Surf	13.800	1.785	7.730	0.16E-11
Dist	−0.821	0.197	−4.155	0.55E-04
D_Core1	1.501	0.266	5.632	0.89E-07
Subunit*D_Core2	1.047	0.180	5.803	0.39E-07
D_Core1*D_Core2	−0.059	0.015	−3.965	0.11E-03

Source	Sum-of-squares	DF	Mean-square	F-Ratio	P
Regression	4609.062	8	576.133	39.554	0.999201E-15
Residual	2126.581	146	14.566		

Dependent variable is mean B. n = 155. Squared multiple R = 0.684. Standard error of estimate is 3.81.

average temperature factor occurs at a reaction coordinate of 1.95, close to the value assigned to the transition state compounds (Fig. 4a). The two-way interaction between subunit and exposed surface area shows, as expected, that exposed residues have higher $\langle B \rangle$ (Fig. 4b). Moreover, the effect is much more substantial in subunit 2 than it is in subunit 1. As noted in the previous paragraph, Student *t* tests are highly significant for all coefficients, establishing support for the conclusion that the quadratic dependence of $\langle B \rangle_{calc}$ on the catalytic state is a relatively important effect compared with the other structure-based predictors.

We turn now to the divergent behavior of the last inhibitor. It is found that of the five complexes, the dihydrozebularine-enzyme assembly shows the largest fluctuations in both the inhibitor molecule itself and in surrounding amino acids. The crystallographic data reveal a trapped water molecule juxtaposed against the tetrahedral carbon of the pyrimidine ring of the inhibitor (see Fig. 1; Xiang et al. 1995). Because this water molecule introduces steric strain through excessively close contacts with the carbon, the surrounding amino acid residues may possibly settle into an energetic (local) minimum characterized by thermal flexibility.

Auto-correlation functions

Results from application of the auto-correlation function (equation 1) to the four enzyme-inhibitor complexes—transition state, substrate, product, and dihydrozebularine—are shown in Figure 5a. The result for F-transition state is nearly indistinguishable from that of the unfluorinated analog and is not shown. The auto-correlation function I(a) calculates the average correlation in atomic temperature factors over all atom-pairs separated by distance a Å, and a is chosen here to range from 2 to 34. To avoid scoring correlations between two identical atoms, that is, one from each monomer, the calculation was performed on a single monomer. For each complex, and with similar functional dependencies, fluctuation auto-correlations between atom-pairs decay with increasing distance a (Fig. 5a).

However, a "differential" and different application of the auto-correlation function reveals differences between the various complexes in a most telling way. For a given complex, we chose all atoms between 1.0 and 3.0 Å from the C1* carbon of the ribose ring of the inhibitor (see Fig. 1), thus forming a 2-Å thick spherical shell of atoms centered around C1*. Because the active site requires both enzymatic subunits, atoms from both monomers are collected, forming a complete shell. Similarly, four other shells are created by collecting all atoms between 3.0 and 5.0 Å, 5.0 and 7.0 Å, 7.0 and 9.0 Å, and 9.0 and 11.0 Å from the central atom. The auto-correlation function is now calculated to reveal atomic fluctuation correlations between first and second shell atoms, second and third shell atoms, third and fourth shells, and fourth and fifth shells, for a total of four values for the auto-correlation I. Specifically, for the first and second shell comparison, every atom belonging to the first shell is compared with every atom belonging to the second. Then every second shell atom is compared to every third shell atom, and so on. Because the successive shells radiate outward from the bound inhibitor, a natural spatial relationship with the ligand is imposed in calculating I—a feature lacking in the first application of Equation 1. In this way, effects of the tight binding transition-state analog on "organizing" fluctuations with increasing distance from the analog can be studied.

Results of this application in Figure 5b show that fluctuation correlations between successive shells of atoms are substantially increased when the two transition-state analogs occupy the active site. Correlations between the two

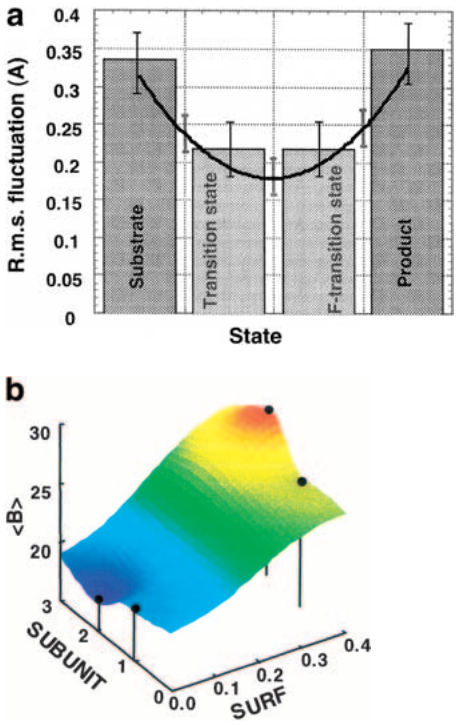

Fig. 4. Graphical representation of the most significant predictors in the temperature factor model in Table 1. (*a*) Quadratic dependence of overall average temperature factor, expressed as $\langle \Delta r^2 \rangle^{1/2} = (\langle B \rangle / 8\pi^2)^{1/2}$, on the catalytic state. The bar graph represents root mean square fluctuations calculated from the crystallographic data; the curve is from the model described in Table 1. Error bars for crystallographic $\langle \Delta r^2 \rangle^{1/2}$ values and for the model are given on the histograms (thin black lines) and on the fitted curve (thick gray lines), respectively. (*b*) The two-way interaction between the subunit to which an amino acid residue belongs and its solvent accessible surface area. Data shown is for the 29 amino acids used for multiple regression analysis, representing one of the two functional active sites of cytidine deaminase . Solvent exposed residues belonging to the second subunit tend to have significantly higher and buried residues, a significantly lower $\langle B \rangle$, than do comparably exposed residues belonging to the first subunit, which contains the Zn atom of this active site and other catalytic residues.

shells most distant from the central atom are found to be very similar for the five complexes: Structural differences among the various inhibitors affect atomic fluctuations to ~9 Å from C1*.

With the exception of the particularly noisy dihydrozebularine complex, an immediate relationship is found between the binding free energies of transition state, F-transition state, substrate, product, and the auto-correlation profiles of the corresponding complexes. As the free energy becomes more favorable for binding, that is, more negative, fluctuation correlations between atoms of successive shells become stronger.

Implications for catalysis

Conformational fluctuations are smallest at the (approximate) transition state of the CDA reaction. This is advantageous: If the transition-state structure is most rigidly held within the active site, those amino acid residues responsible for the chemistry have an increasingly fixed molecular target with which to interact for maximal catalytic efficacy as the reaction proceeds away from substrate.

The reduced atomic excursions characteristic of the transition-state complex have three potential implications for catalysis. First, we expect—because of the increasing stiff-

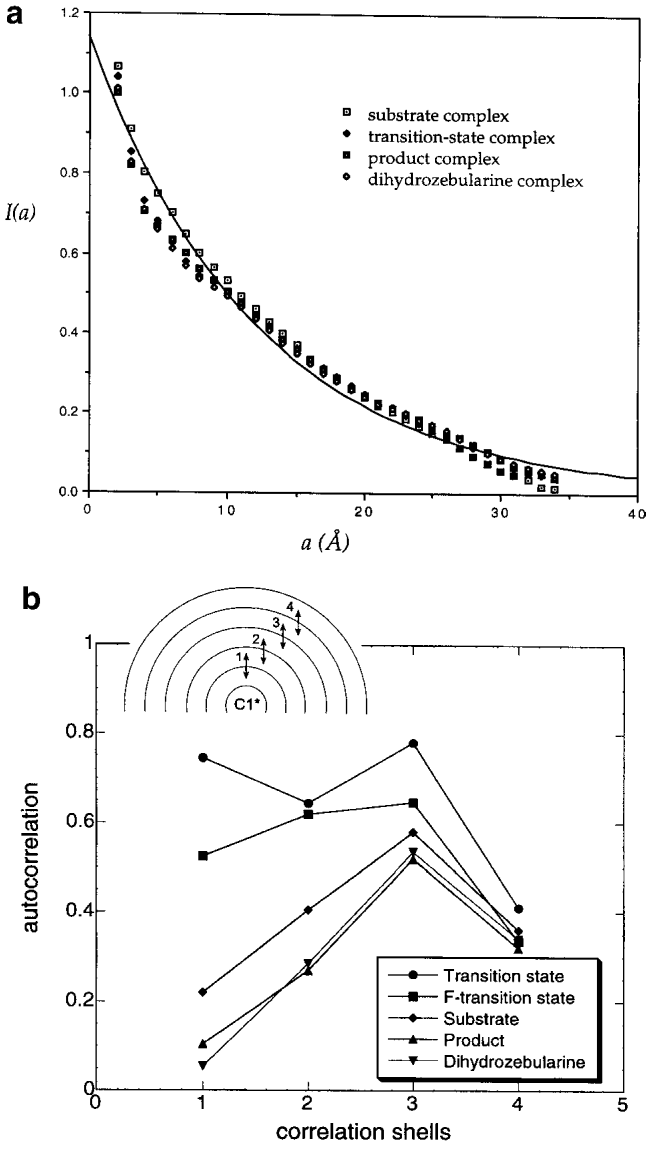

Fig. 5. (*a*) Decay of fluctuation correlations I(a) with increasing distance a between atom-pairs. (*b*) Fluctuation correlations between shells of enzymatic atoms surrounding bound inhibitors. On the x-axis, 1 indicates the correlation between the first and second shells of atoms around a given inhibitor; 2, the calculation for the second and third shells; etc. The pictorial insert sketches the five spherical shells of atoms (shown as hemispheres for brevity) as they radiate outward from C1*, and the numerical labeling within highlights that 1 indicates the auto-correlation calculation between the first and second shells, and so on.

ness of the enzyme matrix—a diminished frequency of "stray" fluctuations during the actual chemical transformation from substrate to transition state; that is, those fluctuations remaining at the transition state are conjectured to be crucial to the chemistry. The importance of an increased proportion of catalytically productive fluctuations at the transition state is further accentuated by the fact that this species is short-lived compared with the substrate: A tightly secured transition-state structure again ensures that these remaining fluctuations can interact favorably with it. These ideas dovetail with the work of Young and Post (1996) who showed that the energetically allowed conformational fluctuations in the (lactate dehydrogenase)-NADH-pyruvate complex are those that distort the cofactor to look like transition state.

Second, atomic-scale enzymatic conformational fluctuations are necessarily coupled to fluctuations in the potential energy of the enzyme-ligand complex. Focus on the local active site potential energy. Greater thermal agitation in the substrate and product complexes, compared with the enzyme-transition-state complex, leads to larger potential energy fluctuations for these ground-state species. Therefore, the enzymatic reaction progresses along a fluctuating potential energy surface, where energetic deviations from the mean potential energy are diminished as the transition state is accessed. A theory for enzymatic catalysis along such a fluctuating potential energy surface is presented elsewhere (Alper et al. 2001).

Finally, we can restate with increased precision the general notion that the entropic barrier to catalysis at the transition state must be compensated by enhanced enthalpic binding interactions between the transition-state species and the active site of the enzyme. Diminished active-site thermal motions characteristic of the transition state imply a loss in local conformational entropy at precisely the point in catalysis at which the equilibrium binding affinity of the active site reaches a maximum. During a spontaneous elementary chemical process, the active-site free energy, $\Delta G = \Delta H - T\Delta S$, must be negative for conversion of enzyme-substrate complex (ES) to enzyme-transition-state complex (ES*). The active-site entropy change ΔS is also negative. Thus, the enthalpy change ΔH for converting ES to ES* must be even more negative for the actual enzymatic process than for a comparable process in which structural fluctuations remained comparable in both states, ensuring $\Delta G < 0$ for spontaneous reaction.

Snider, Wolfenden and coworkers (Snider et al. 2000; Snider and Wolfenden 2001) have reported that the entropic change in CDA in going from the substrate complex to the transition-state complex is slightly positive, with $T\Delta S = 0.9$ kcal/mole. Consequently, the formation of bound transition state from bound substrate is favored from an entropic and enthalpic standpoint. These measurements target the entire solvated enzyme and may potentially mask the restrictions in active-site conformational freedom that we report: Even though the active site itself appears to become increasingly ordered in the transition state, the overall entropy change leading to transition state formation is slightly positive. A deeper resolution to this conundrum may lie either in the precise details of the normal modes of enzyme motion in the transition state, which may broaden along the productive direction, or in otherwise indirect effects, such as the release by the transition-state complex of solvent water molecules that are bound to the ground-state complexes. This question may be addressed by comparing the bound water structures of the various complexes, and this is in progress.

Experimental studies are addressing increasingly profound questions in protein dynamics (Zaccai 2000). Furthermore, large-scale quantum mechanical calculations have recently been performed on CDA to geometrically optimize 1330 active-site atoms (Lewis et al. 1998). As computational studies and continuing experimental work address the dynamical features of cytidine deaminase, correlations between catalytic efficacy and fluctuation dynamics will be confirmed in atomic-scale detail. Not surprisingly, fast atomic motions recorded in molecular dynamics simulations can be used to calculate temperature factors for comparison to those derived from experiment (Karplus and Petsko 1990; Loncharich and Brooks 1990; Verma et al. 1997; Melchionna et al. 1998). These quantities, from experimental data, simulations, or theory, are useful in characterizing the thermal mobilities of enzymatic domains (Loll and Lattman 1989; Hynes and Fox 1991; Bahar and Jernigan 1999). Because analysis of molecular dynamic trajectories can reveal fluctuation correlations (Lins et al. 1999; Radkiewicz and Brooks 2000), it would be intriguing to model CDA with bound reaction-coordinate species.

Materials and methods

Multiple regression and analysis of variance

Data for the set of 30 amino acids closest to the bound inhibitors were reformatted into an experimental matrix with one independent variable, the average temperature-factor value $\langle B \rangle$ for an amino acid, per line. This matrix was completed by encoding possible predictors in additional columns. Potential predictors of $\langle B \rangle$ were encoded numerically for each amino acid. They included the catalytic state or position along the reaction coordinate (State: 1, substrate; 2, transition state; 3, product), the subunit to which an amino acid belongs (Subunit: 1, subunit containing the zinc ligands and Glu104; 2, the second subunit), the solvent accessible surface area of the side-chain in the folded protein as a percentage of that for an extended chain in solution (Surf), the distance of an amino acid centroid from C1* (see Fig. 1) of the bound pyrimidine ligand (Dist, Å), and the distances from each of two nonpolar core packing regions, one in each domain (D_Core1 and D_Core2, Å). Identification of these core regions was performed by Delaunay tessellation and likelihood scoring, as described previously (Na-

varatnam et al. 1998). They consist of clusters of tetrahedra formed by nearest neighbors composed of the amino acids valine, leucine, isoleucine, tyrosine, phenylalanine, tryptophan, or methionine that occur in the database of known structures more than four times as often as expected by chance. The search for linear models was performed with the SYSTAT statistics program (Wilkinson 1987), using stepwise multiple regression to compare different possible models. The model in Table 1 was obtained stepping either forward or backward and was subsequently reevaluated by full least squares. Estimation by multiple regression and analysis of variance is robust with respect to contributions to the observed variation from unknown predictors. Thus, the analysis is capable of estimating jointly the statistical significance of several different factors simultaneously without undue confounding by unknown factors.

Auto-correlation functions

The auto-correlation function (Equation 1) for an enzyme-inhibitor complex

$$I(a) = \frac{n}{L} \frac{\sum_{i,j} (f_i - \bar{f})(f_j - \bar{f})}{\sum_{i} (f_i - \bar{f})^2}$$

(with f_i indicating the B-value at atom i; f_j, the value at atom j which is separated from i by distance a; and $\bar{f}$, the average temperature factor) is used as follows. Operationally, the summation in the numerator extends over all pairs of atoms i, j that are separated by a distance of $a \pm \delta$ (in units of angstroms), and the denominator sums over all individual atoms contained in the enzyme-inhibitor complex. δ is typically chosen to be 1.0 Å; n is the total number of atoms in the complex; and L is the number of atom-pairs separated by the appropriate distance $a \pm \delta$. I(a), therefore, calculates the average correlation in fluctuation magnitudes over all atoms separated by $a \pm \delta$. It is clear that if atoms i, j are both noisy with f_i and f_j larger than the mean fluctuation $\bar{f}$, a positive contribution to I(a) will result for this correlated atom-pair. Conversely, if atom i is less noisy than the mean, the two atoms are not correlated in their fluctuations, and a negative contribution to I(a) results. A larger positive value for I(a) indicates more strongly correlated fluctuations. Auto-correlation between atom-pairs of increasing separation a in the enzyme-inhibitor complex can be studied by computing I(a) as a function of a.

Acknowledgments

C.K.B thanks Vern Schramm and Ben Braunheim for illuminating discussions. This work was supported by grants from the Jeffress Memorial Trust (J-432; C.K.B.) and Research Corporation (CC4937; C.K.B).

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

References

Alhambra, C., Corchado, J., Sanchez, M.L., Garcia-Viloca, M., Gao, J., and Truhlar, D.G. 2001. Canonical variational theory for enzyme kinetics with the protein mean force and multidimensional quantum mechanical tunneling dynamics: Theory and application to liver alcohol dehydrogenase. *J. Phys. Chem. B* **105:** 11326–11340.

Alper, K.O., Singla, M., Stone, J.L., and Bagdassarian, C.K. 2001. Correlated conformational fluctuations during enzymatic catalysis: Implications for catalytic rate enhancement. *Protein Sci.* **10:** 1319–1330.

Bahar, I. and Jernignan, R.L. 1999. Cooperative fluctuations and subunit communication in tryptophan synthase. *Biochemistry* **38:** 3478–3490.

Balabin, I.A. and Onuchic, J.N. 2000. Dynamically controlled protein tunneling paths in photosynthetic reaction centers. *Science* **290:** 114–117.

Betts, L., Xiang, S., Short, S.A., Wolfenden, R., and Carter, Jr., C.W. 1994. Cytidine deaminase: The 2.3 Å crystal structure of an enzyme:transition-state analog complex. *J. Mol. Biol.* **235:** 635–656.

Billeter, S.R., Webb, S.P., Agarwal, P.K., Iordanov, T., and Hammes-Schiffer, S. 2001. Hydride transfer in liver alcohol dehydrogenase: Quantum dynamics, kinetic isotope effects, and role of enzyme motion. *J. Am. Chem. Soc.* **123:** 11262–11272.

Brooks, C.L., Karplus, M., and Pettitt, B.M. 1988. *Proteins: A theoretical perspective of dynamics, structure and thermodynamics*. Wiley, New York, NY.

Bruice, T.C. and Benkovic, S.J. 2000. Chemical basis for enzyme catalysis. *Biochemistry* **39:** 6267–6274.

Cannon, W.R., Singleton, S.F., and Benkovic, S.J. 1996. A perspective on biological catalysis. *Nat. Struct. Biol.* **3:** 821–833.

Carter, Jr., C.W. 1995. The nucleoside deaminases for cytidine and adenosine: Structure, transition state stabilization, mechanism, and evolution. *Biochimie* **77:** 92–98.

Cliff, A.D. and Ord, J.K. 1973. *Spatial autocorrelation*. Pion, London, England.

Eurenius, K.P., Chatfield, D.C., Brooks, B.R., and Hodoscek, M. 1996. Enzyme mechanisms with hybrid quantum and molecular mechanical potentials, I: Theoretical considerations. *Int. J. Quant. Chem.* **60:** 1189–1200.

Frauenfelder, H., Petsko, G.A., and Tsernoglou, D. 1979. Temperature-dependent X-ray diffraction as a probe of protein structural dynamics. *Nature* **280:** 558–563.

Frick, L., MacNeela, J.P., and Wolfenden, R. 1987. Transition state stabilization by deaminases: Rates of nonenzymatic hydrolysis of adenosine and cytidine. *Bioorganic Chem.* **15:** 100–108.

Frick, L., Yang, C., Marquez, V.E., and Wolfenden, R. 1989. Binding of pyrimidin-2-one ribonucleoside by cytidine deaminase as the transition-state analogue 3,4-dihydrouridine and the contribution of the 4-hydroxyl group to its biding affinity. *Biochemistry* **28:** 9423–9430.

Hagen, S.J., Hofrichter, J., and Eaton, W.A. 1995. Protein reaction dynamics in a room-temperature glass. *Science* **296:** 959–962.

Hynes, T.R. and Fox, R.O. 1991. The crystal structure of staphylococcal nuclease refined at 1.7 Å resolution. *Proteins* **10:** 92–105.

Ilyin, V.A., Temple, B., Hu, M., Li, G., Yin, Y., Vachette, P., and Carter, Jr., C.W. 2000. 2.9 Å crystal structure of ligand-free tryptophanyl-tRNA synthetase: Domain movements fragment the adenine nucleotide binding site. *Protein Sci.* **9:** 218–231.

Karplus, M. and Ichiye, T. 1996. Comment on a "fluctuation and cross correlation analysis of protein motions observed in nanosecond molecular dynamics simulations." *J. Mol. Biol.* **263:** 120–122.

Karplus, M. and Petsko, G.A. 1990. Molecular dynamics simulations in biology. *Nature* **347:** 631–639.

Kohen, A., Cannio, R., Bartolucci, S., and Klinman, J.P. 1999. Enzymatic dynamics and hydrogen tunneling in a thermophilic alcohol dehydrogenase. *Nature* **399:** 496–499.

Kraut, J. 1988. How do enzymes work? *Science* **242:** 533–540.

Lewis, J.P., Carter, Jr., C.W., Hermans, J., Pan, W., Lee, T.-S., and Yang, W. 1998. Active species for the ground-state complex of cytidine deaminase: A linear-scaling quantum mechanical investigation. *J. Am. Chem. Soc.* **120:** 5407–5410.

Lins, R.D., Briggs, J.M., Straatsma, T.P., Carlson, H.A., Greenwald, J., Choe, S., and McCammon, J.A. 1999. Molecular dynamics studies on the HIV-1 integrase catalytic domain. *Biophys. J.* **76:** 2999–3011.

Loll, P.J. and Lattman, E.E. 1989. The crystal structure of the ternary complex of staphylococcal nuclease, Ca^{2+}, and the inhibitor pdTp, refined at 1.65 Å.*Proteins* **5:** 183–201.

Loncharich, R.J. and Brooks, B.R. 1990. Temperature dependence of dynamics of hydrated myoglobin. *J. Mol. Biol.* **215:** 439–455.

Mader, M.M. and Bartlett, P.A. 1997. Binding energy and catalysis: The implications for transition-state analogs and catalytic antibodies. *Chem. Rev.* **97:** 1281–1301.

McCammon, J.A. and Harvey, S.C. 1987. *Dynamics of proteins and nucleic acids*. Cambridge University Press, Cambridge, UK.

Melchionna, S., Falconi, M., and Desideri, A. 1998. Effect of temperature and hydration on protein fluctuations: molecular dynamics simulation of Cu, Zn superoxide dismutase at six different temperatures: Comparison with neutron scattering data. *J. Chem. Phys.* **108:** 6033–6041.

More, N., Daniel, R.M., and Petach, H.H. 1995. The effect of low temperatures on enzyme activity. *Biochem. J.* **305:** 17–20.

Mulholland, A.J., Grant, G.H., and Richards, W.G. 1993. Computer modelling of enzyme catalyzed reaction mechanisms. *Protein Eng.* **6:** 133–147.

Navaratnam, N., Fujino, T., Bayliss, J., Jarmuz, A., How, A., Richardson, N., Somasekaram, A., Battacharya, S., Carter, Jr., C.W., and Scott, J. 1998. *Escherichia coli* cytidine deaminase provides a molecular model for ApoB RNA editing and a mechanism for RNA substrate recognition. *J. Mol. Biol.* **275:** 695–714.

Radkiewicz, J.L. and Brooks, C.L. 2000. Protein dynamics in enzymatic catalysis: Exploration of dihydrofolate reductase. *J. Am. Chem. Soc.* **122:** 225–231.

Radzicka, A. and Wolfenden, R. 1995. Transition state and multistate analog inhibitors. *Meth. Enzymol.* **246:** 284–312.

Rasmussen, B.F., Stock, A.M., Ringe, D., and Petsko, G.A. 1992. Crystalline ribonuclease A loses function below the dynamical transition at 220 K. *Nature* **357:** 423–424.

Ringe, D. and Petsko, G.A. 1999. Quantum enzymology: Tunnel vision. *Nature* **399:** 417–418.

Schramm, V.L. and Bagdassarian, C.K. 1999. Deamination of nucleosides and nucleotides and related reactions. In *Comprehensive natural products chemistry,* vol. 5 (ed. C.D. Poulter), pp. 71–100. Elsevier, Amsterdam, The Netherlands.

Snider, M.J. and Wolfenden, R. 2001. Site-bound water and the shortcomings of a less than perfect transition state analogue. *Biochemistry* **40:** 11364–11371.

Snider, M.J., Gaunitz, S., Ridgway, C., Short, S.A., and Wolfenden, R. 2000. Temperature effects on the catalytic efficiency, rate enhancement, and transition state affinity of cytidine deaminase, and the thermodynamic consequences for catalysis of removing a substrate "anchor." *Biochemistry* **39:** 9746–9753.

Verma, C.S., Caves, L.S.D., Hubbard, R.E., and Roberts, G.C.K. 1997. Domain motions in dihydrofolate reductase: A molecular dynamics study. *J. Mol. Biol.* **266:** 776–796.

Wagener, M., Sadowski, J., and Gasteiger, J. 1995. Autocorrelation of molecular surface properties for modeling *coricosteroid binding globulin* and cytosolic *Ah* receptor activity by neural networks. *J. Am. Chem. Soc.* **117:** 7769–7775.

Wang, F., Shi, W., Nieves, E., Angeletti, R.H., Schramm, V.L., and Grubmeyer, C. 2001. A transition-state analogue reduces protein dynamics in hypoxanthine-guanine phosphoribosyltransferase. *Biochemistry* **40:** 8043–8054.

Wierenga, R.K., Borchert, T.V., and Noble, M.E.M. 1992. Crystallographic binding studies with triosephosphate isomerases: Conformational changes induced by substrate and substrate-analogues. *FEBS* **307:** 34–39.

Wilkinson L. 1987. *SYSTAT: The system for statistics*, version 5.2.1. SYSTAT, Inc., Chicago, IL.

Wolfenden, R. 1972. Analog approaches to the structure of the transition state in enzymatic reactions. *Acc. Chem. Res.* **5:** 10–17.

Xiang, S., Short, S.A., Wolfenden, R., and Carter, Jr., C.W. 1995. Transition-state selectivity for a single hydroxyl group during catalysis by cytidine deaminase. *Biochemistry* **34:** 4516–4523.

———. 1996. Cytidine deaminase complexed to 3-deazacytidine: A "valence buffer" in zinc enzyme catalysis. *Biochemistry* **35:** 1335–1341.

———. 1997. The structure of the cytidine deaminase-product complex provides evidence for efficient proton transfer and ground-state destabilization. *Biochemistry* **36:** 4768–4774.

Young, L. and Post, C.B. 1996. Catalysis by entropic guidance from enzymes. *Biochemistry* **35:** 15129–15133.

Zaccai, G. 2000. How soft is a protein?: A protein dynamics force constant measured by neutron scattering. *Science* **288:** 1604–1607.