

Temperature dependence of reactions with multiple pathways

Muhammad H. Zaman,^{ac} Tobin R. Sosnick^{bc} and R. Stephen Berry^{*ad}

^a Department of Chemistry, The University of Chicago, Chicago, IL 60637, USA.
E-mail: berry@uchicago.edu

^b Department of Biochemistry and Molecular Biology, The University of Chicago, Chicago, IL 60637, USA

^c Institute of Biophysical Dynamics, The University of Chicago, Chicago, IL 60637, USA

^d The James Franck Institute, The University of Chicago, Chicago, IL 60637, USA

Received 3rd January 2003, Accepted 29th April 2003

First published as an Advance Article on the web 13th May 2003

The rate coefficient $k(T)$ of a barrier-limited, two-state reaction proceeding through multiple pathways is determined by using transition state theory. The entropic and enthalpic contributions to the overall rate are treated separately. The rate coefficient of reaction is determined as a function of the variance of the distribution of pathways and activation energies. The relative densities of state in the initial and transition state lead to curvature in the dependence of $\ln(k)$ on $1/T$. The sign of the curvature depends on the relative increase in densities of state of the initial and the transition (saddle) state. The analysis has implications for the interpretation of macromolecular reactions, including the protein-folding process. The inclusion of the density of states in the initial and the saddle states yields rate coefficients whose temperature dependence reproduces the curvatures in experimental Arrhenius plots for protein-folding and unfolding processes.

Introduction

Recent advances in experimental and theoretical methods to study macromolecules have stirred a rebirth of interest in the temperature-dependent kinetics of complex reactions.^{1,2} These theoretical and experimental methods³ have provided extremely useful information about the potential energy surfaces and the pathways that lead from one state to another. However, the picture is still incomplete.

One uncertainty is the effect of multiple pathways on the temperature dependence of reactions. An analysis of this issue can provide vital information about the shape of the potential energy surface. The problem has been addressed by both experimental^{4,5} and theoretical² research in recent years. One example is the protein-folding problem where the energy landscape picture has been quite useful. Determining the presence or absence of multiple pathways can provide critical information about the landscape and the overall folding behavior. Therefore, a clear experimental signature of multiple pathways can be a very useful tool for the interpretation of complex reactions. In this study, we show that the temperature dependence of rate coefficients $k(T)$, specifically the deviation from the traditional Arrhenius linear dependence of $\ln(k)$ on $1/T$, can contain this information.

Kinetics through multiple pathways can be divided into three categories; parallel pathways, sequential pathways and combinations thereof. In this paper, we address the first case, and apply transition state theory to calculate the rate of a two-state reaction with parallel pathways having different activation energies. Our method takes into account the implicit temperature dependence of the reaction rates, and addresses the consequences of the distribution of activation energies of pathways. We find that the presence of multiple pathways results in curvature in Arrhenius plots. The sign of the curvature reveals the relative rates of increase, with energy, of the densities of states of the saddle and the initial state.

Methods

According to transition state theory (TST),

$$k = \frac{k_B T}{h} \exp\left(-\frac{\Delta G^\ddagger}{RT}\right) = \frac{k_B T}{h} \exp\left(-\frac{\Delta H^\ddagger}{RT}\right) \exp\left(\frac{\Delta S^\ddagger}{R}\right) \quad (1)$$

where the superscript \ddagger refers to the activation parameters.

The activation enthalpy of many macromolecular systems (e.g. protein-folding process) is of the order of a few kilocalories per mole.⁶ Therefore in our model, we use the Arrhenius approximation of $\Delta H^\ddagger = E_a - RT$. The entropy of activation of the system is calculated by considering the densities of states, $\rho(E)$, for the initial and the saddle states in the rate-determining step (see eqns. (2) and (3) below). The contribution to the entropy of activation from initial and transition (saddle) state densities depends on temperature. For our model, we assume that the population in the density of states either increases linearly or quadratically with temperature. More complex behavior is plausible, but in the absence of specific information about any particular system, and in the present situation of our examination of general behavior, the linear and quadratic models are sufficient to demonstrate the behavior we seek to interpret.

We also assume that if the density of states of the reactant increases linearly with temperature then the density of states for the saddle state increases quadratically and *vice versa*. Once again, this is to ensure that we are using the simplest possible model. For a process such as folding or unfolding, the densities of states of the initial and the saddle states do not increase at the same rate. Therefore we use this assumption to model such processes in which the densities of states of the initial and the saddle state increase at different rates. If the densities of states of the initial and the transition (or the final) state increase at the same rate, there is no curvature in Arrhenius plots.

The change of entropy of between two given states (initial and transition or transition or final) at a given temperature

is then calculated by using the densities of states:

$$\Delta S(T) = R(\ln \rho^{\text{TS}}(T)) - (\ln \rho^{\text{I}}(T)) \quad (2)$$

where the superscripts TS and I refer to the densities of states at the saddle state and the initial state respectively. The overall rate coefficient is thus given by:

$$k(T) = \frac{k_{\text{B}}T}{h} \exp\left(1 - \frac{E_{\text{a}}}{RT}\right) \exp(\Delta S(T)) \quad (3)$$

The density of states at the saddle is a measure of the multiplicity of pathways, but no reference need be made to the extent to which there is mode-coupling among them. In the system with multiple pathways, the pathways differ from each other due to different E_{a} .

For an ensemble of pathways, each with different activation energies $E_{\text{a},i}$, the observed rate coefficient is the sum of rate coefficients through each individual pathway i is given by:

$$k_{\text{observed}} = \sum_i k_i \quad (4)$$

Assuming that the activation energies have a Gaussian distribution, we can determine $P(E_{\text{a}})$, by using the standard equation of a normalized Gaussian distribution with mean activation energy μ and variance $\sigma_{E_{\text{a}}}$:

$$P(E_{\text{a}}) = \frac{1}{\sigma_{E_{\text{a}}} \sqrt{2\pi}} \exp\left(-\frac{(E_{\text{a}} - \mu)^2}{2\sigma_{E_{\text{a}}}^2}\right) \quad (5)$$

This value of $P(E_{\text{a}})$ is multiplied to its corresponding E_{a} in eqn. (3) to ensure the Gaussian weighting of activation energy. Hence the overall rate depends on the nature of the distribution of activation energies, the mean activation energy and the variance.

Model systems and results

Model 1

Two different model systems are examined in this study. In the first model, there are two routes, a singular low energy pathway with activation energy E_{a} , and an ensemble of g -fold degenerate higher energy pathways having a higher activation energy $E_{\text{b}} = E_{\text{a}} + \delta E_{\text{a}}$ (Fig. 1a). The purpose of this model is to examine how the presence of an ensemble of high energy pathways affects the behavior of k_{obs} ; explicitly, identifying the temperature range where the flux switches from the lower energy route to the higher energy routes. The net rate of the reaction depends, in addition to the temperature and the activation energy barriers of the two routes, on the degeneracy of the higher energy pathways.

If the reaction rate through each pathway obeys the Arrhenius approximation, the net rate coefficient of the reaction can be written as the products of the exponentials of entropic and enthalpic terms:

$$\begin{aligned} k_{\text{obs}} &= A \exp(\ln[1]) \exp\left(-\frac{E_{\text{a}}}{RT}\right) + B \sum \exp(\ln[g]) \exp\left(-\frac{E_{\text{b}}}{RT}\right) \\ &= A \exp\left(-\frac{E_{\text{a}}}{RT}\right) + B \sum \exp(\ln[g]) \exp\left(-\frac{E_{\text{a}} + \delta E_{\text{a}}}{RT}\right) \end{aligned} \quad (6)$$

where A and B are constants which are obtained from the initial conditions of the reaction. The summation is over all the possible pathways. The degeneracy of the low energy pathway is 1 and hence its entropic contribution is zero. The degeneracy of the higher energy pathway is g , and hence its entropic contribution is $\ln[g]$.

The rate coefficient of the reaction for a system depicted in Fig. 1a as a function of temperature is plotted in Fig. 1b with $g = 4000$, $E_{\text{a}} = 10 \text{ kcal mol}^{-1}$ and $\delta = 0.51$.

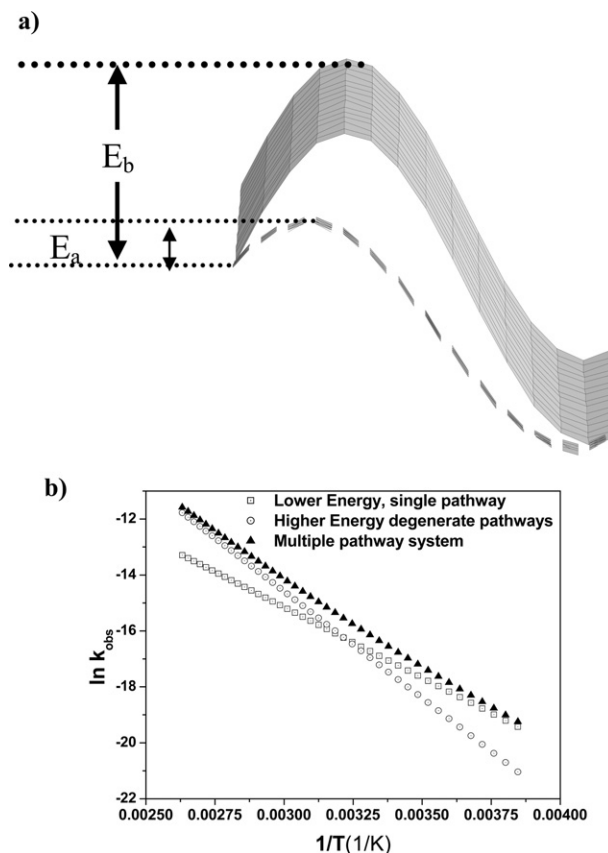


Fig. 1 Model #1: The two-pathway reaction. (a) The reaction diagram of a two-pathway system. The single lower energy pathway, with a single activation energy E_{a} , and a higher energy ensemble of pathways with activation energy $E_{\text{b}} = E_{\text{a}} + \delta E_{\text{a}}$ and degeneracy g , where $g = 4 \times 10^3$. (b) The Arrhenius plot for the system shown in (a), with $\delta = 0.51 E_{\text{a}}$. The observed rate coefficient through only the lower energy pathway, and through the ensemble is plotted to show the “temperature switch” that changes the flux from one pathway to another pathway.

Model 2

This model is an extension of the first model, as this model spreads the sharp distribution of energies in Model 1 over a range of energies. The reactant molecules can now go through a wider distribution of pathways, *i.e.* pathways with a Gaussian distribution of activation energies and a variance $\sigma_{E_{\text{a}}}$. The overall rate through each pathway then depends upon the density of states in the reactant and the saddle state, whose populations vary with temperature as discussed above in the methods section. The rate coefficient of reaction through each individual pathway is calculated according to eqn. (3) and the overall observed rate coefficient is given by eqn. (4).

The plots of $\ln(k)$ versus $1/T$ for different mean activation energies are shown in Fig. 2(a and c). These plots of rate coefficients are obtained by calculating the rates for N ($N = 1500$) pathways in 5 K temperature intervals. Fig. 2(a) shows the behavior of a two-state system for which the density of states for the reactant increases faster (increases quadratically with T) than the density of states of the in the saddle state (increases linearly as T), *i.e.* $\rho^{\text{I}} > \rho^{\text{TS}}$. Fig. 2(c) depicts the opposite behavior ($\rho^{\text{I}} < \rho^{\text{TS}}$).

The plots in Fig. 2(a and c) show increased curvature at lower and higher temperatures. The curvatures increase several fold at even higher ($> 373 \text{ K}$) and lower ($< 270 \text{ K}$) temperatures; however such a temperature range is unlikely to be observed experimentally for biological systems. The results are plotted for various mean activation energies (μ); the variance of the distribution (σ_E) is kept the same for these

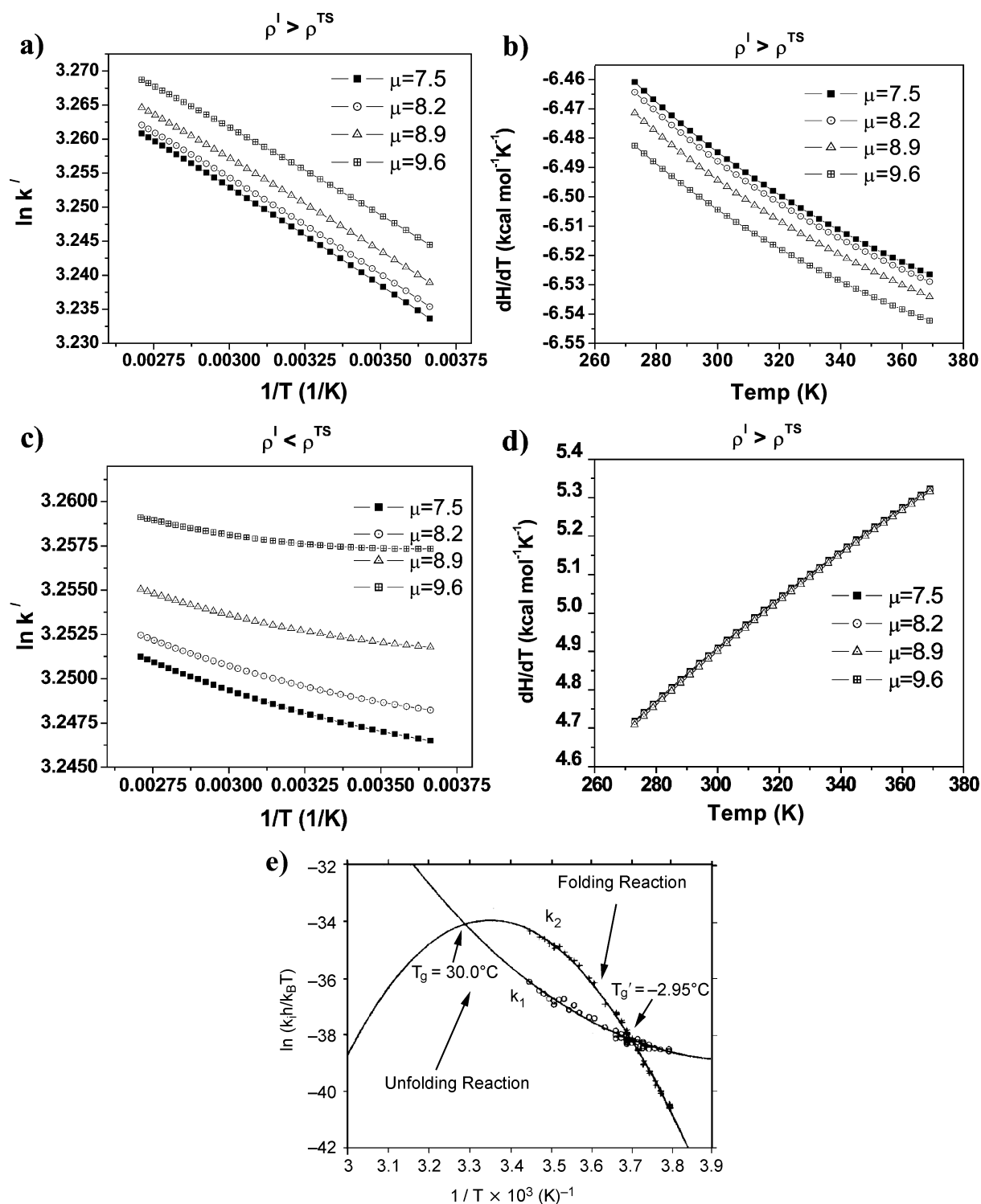


Fig. 2 Model #2: Log rate coefficient vs. 1/Temperature for different mean activation energies. (a) The log of the observed rate coefficient of the reaction versus $1/T$, also known as an Arrhenius plot. The y-axis is $k' = (kh/k_B T)$ as often used in macromolecular kinetics plots.^{9,10} The origin of curvature is described in detail in the text. The plot represents a system for which the density of states of the initial state increases at a faster rate than that of the saddle state. (b) $dH/dT (= C_p)$ versus temperature for a system for which the density of states of the initial state increases at a faster rate than that of the saddle state. The curvature in the plot suggests that the slope is non-zero ($dC_p/dT < 0$). (c) The y-axis is the same as in Fig. 2a. The plot represents a system for which the density of states of the initial state increases at a slower rate than that of the saddle state. The curvature in the plot suggests that the slope is non-zero ($dC_p/dT > 0$). (d) dH/dT versus temperature for a system for which the density of states of the initial state increases at a slower rate than that of the saddle state. The curvature in the plot suggests that the slope is non-zero ($dC_p/dT > 0$). (e) Rate-coefficient plot for folding and unfolding of mutant of phage T4 lysozyme (from ref. 9, reproduced with author's permission). The plot shows opposite curvatures for folding and unfolding reactions, as predicted in Figs. 2(a) and 2(c). The plots are based on different mean activation energies of the Gaussian distribution, however the variances of the distributions are kept the same.

plots. If the density of states of either the reactant or the saddle state were higher and varied more rapidly, the curvature would of course appear within a narrower range of temperature. For this system, our calculations suggest that the effective dH/dT is

not a constant with temperature, so we have a non-zero value of dC_p/dT . For a process whose the density of states increases faster in the initial state than in the saddle state (*i.e.* $\rho^I > \rho^{TS}$), dC_p/dT is negative. The opposite effect is observed if the

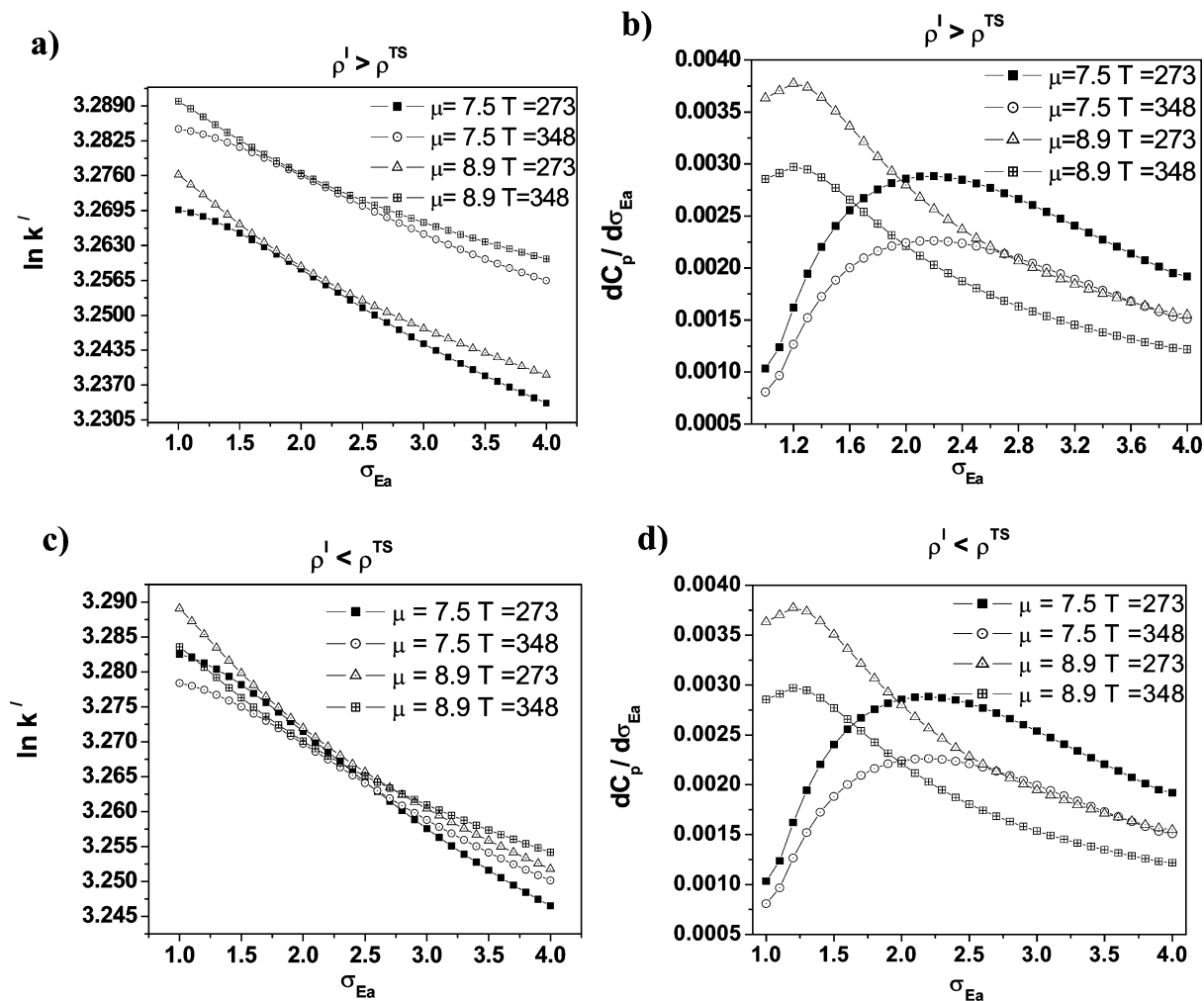


Fig. 3 Model #2: Log Rate Coefficient vs. Variance of Gaussian distribution plots for different mean activation energies and different temperatures. (a) The log of the observed rate coefficient of the reaction as a function of the variance of the Gaussian distribution. The y-axis shown is $k' = (kh/k_B T)$. The plot represents a system for which the density of states of the initial state increases at a faster rate than that of the saddle form. (b) The plot of change of C_p with variance magnifies the fact that these variance plots are very sensitive to changes in activation energy and insensitive to changes in temperature. (c) The plot is similar to (a) and represents a system for which the density of states of the initial state increases at a slower rate than of the saddle state. (d) The plot is similar to plot (b), except the system plotted is the one for which the density of states of the initial state increases at a slower rate than that of the saddle state. The plots have two different mean activation energies and are plotted at two different temperatures.

density of states of the saddle state increases faster (Figs. 2b and d respectively).

Fig. 3 shows the behavior of the natural log of observed rate coefficient *versus* the variance of the distribution at different temperatures. The plots show the behavior of the system at different mean activation energies (μ). Fig. 3(a) show the behavior of the model two-state system whose density of states of the initial state increases at a faster rate (increases quadratically with temperature) than that of the saddle state (increases linearly with temperature), whereas the opposite effect is depicted in Fig. 3(c).

Discussion

The first model involving pathways at just two energies, with one lower-energy pathway and an ensemble of degenerate higher energy pathways, though very simple, is quite useful. It shows that the presence of high energy paths can influence the rate of the reaction significantly at physiological temperatures (Fig. 1b), under certain conditions, *e.g.* when the degeneracy of higher energy paths is high ($g > 10^3$) and the activation energy gap between the lower and higher energy paths is small ($E_b - E_a \leq 0.5E_a$). This model also shows the presence of a temperature “switch”, at which the dominant

flux switches from one pathway to the other. Such information is going to be very useful to design, conduct and interpret macromolecular kinetic experiments at different temperatures.

The results using model #2 also show curvature in the Arrhenius plots. As stated earlier, the plot shows more curvature at higher (> 373 K) temperatures than in the intermediate temperature range ($273 < T < 373$). The curvature in our model is due to the presence of multiple pathways and the individual rates of reaction through those pathways. Since we are working with a Gaussian distribution, at lower temperatures pathways with very low activation energies are populated, whereas at $T > 310$ K the flux goes through all the possible pathways.

The origin of curvature in Arrhenius plots in macromolecular processes such as protein folding has been a subject of debate.^{1–3,6–8} According to some theories, the origin is due to diffusion on the landscape^{1,2,6} whereas the experimentally observed curvature in Arrhenius plots is usually attributed to hydrophobic effect and other temperature dependent interactions that stabilize the folding molecule.³

First-principle theoretical methods have failed to reproduce the experimentally observed curvature in Arrhenius plots for protein-folding reactions. In fact, these theoretical results predict curvature in the Arrhenius plots with a sign opposite to

that of experiments. This is because the entropic contribution is modeled as a term completely independent of temperature, and hence there is an increase in the observed rate at higher temperatures (also shown in model #1; Fig. 1b). Based on our model (model #2, in which the densities of states of initial and the saddle state, and their implicit dependence on temperature are taken into account) we can correctly account for the sign of the curvature for protein-folding and unfolding reactions. Though there might be a difference in magnitude of the rate coefficient (because of complexities that our model does not take into account), our results agree fairly well with the experimental results^{9,10} in predicting the sign of the curvature in the Arrhenius plots.

For a protein-folding reaction, the density of states of the initial unfolded state increases faster than that of a saddle state, whereas the opposite effect is to be expected for the unfolding reaction. This is because the random coil is less restrictive than the transition (saddle) state and hence is expected to have a density of states, up to some rather high energy, that increases rapidly. The transition state is presumably more constrained, hence its density of state increases more slowly, and the native state is the most restricted of all, and hence has the slowest-growing density of states. Interpreting our model with this assumption gives the observed curvature for both the unfolding and the folding reactions. Our results agree well with experimentally determined folding and unfolding rate-coefficient plot (Fig. 2e), where the folding and unfolding curvatures have opposite signs, as predicted by our theoretical model (Fig. 2a and 2c).

Our results also suggest that for a multiple path process the effective dH/dT is nonzero. Though the change in magnitude is relatively small, the slope of the plot of $C_p (= dH/dT)$ versus temperature is non-zero for both cases discussed above (Fig. 2b and d).

From a theoretical perspective, reaction along a single pathway will result in a constant C_p if no other interactions (*e.g.* hydrophobic effect) are taken into account. Ours is a phenomenological model in which the densities of states of the initial and saddle states are populated as functions of temperature, and result in a temperature-dependent heat capacity, *i.e.* $dC_p/dT \neq 0$.

It is sometimes said that the hydrophobic effect results in the curvature in the Arrhenius plots.³ The origins of hydrophobic effect (difference in C_p in the unfolded and the folded state) lie in the difference in organization of water around the folded and the unfolded molecule and the difference in accessible energy states. Therefore the hydrophobic effect can be interpreted as a microscopic model that carries with it a difference in densities of conformational states between the unfolded and transition (or folded) states. This, in turn, results in curvature in the Arrhenius plots. Therefore our analysis is not contradictory, but complimentary to other work showing the curvature in Arrhenius plots due to hydrophobic effect.³ Rather, our model shows that *any* effect associated with the differing patterns of densities of state in initial and transition states will yield such curvature.

Another interesting feature of the presence of multiple pathways is the compensation between enthalpic and entropic factors. The enthalpic term in the free energy for a folding reaction is fairly small, and is the only term depending explicitly on the activation energy. However, this might not be the case for other macromolecular reactions, where the enthalpy plays a more significant role. Eqn. (1) indicates that enthalpy and entropy act in opposite directions. In our model, for a system whose density of states of the reactant increases faster than the saddle state in the rate-determining step (*i.e.* $\rho^I > \rho^{TS}$), the entropic contribution is positive, so entropy and enthalpy work in opposite directions (see eqns. (1), (2) and (3)). Thus for a system with a small enthalpic contribution, or small E_a (*e.g.* folding reaction) the curvature is primarily

due to the entropic contribution. On the other hand, the enthalpy and entropy terms in the rate of the reaction act in the same direction for a system whose density of states of the saddle state increases faster than that of the initial state (*i.e.* $\rho^I < \rho^{TS}$). This is because the entropic contribution in eqns. (1) and (3) is negative, and thus both entropy and enthalpy have the same sign. This is observed from Fig. 2c ($\mu = 9.6$), which has the maximum curvature, due to higher activation energy and hence greater enthalpy. Because the enthalpic component of the overall reaction rate depends only upon the activation energy and is independent of any variation in the density of states, the enthalpic component shows the same behavior for the forward and the backward reaction. Consequently, for macromolecular processes which are governed overwhelmingly by enthalpic forces, we predict that the shape of the $\ln(k)$ vs. $1/T$ plot would be the same for forward and backward reactions.

We can also predict the rates of reactions through parallel pathways as a function of the variance of the distribution of activation energies (Fig. 3). Though it is not yet possible to observe the rate of the reaction experimentally as a function of this parameter, nonetheless these plots can give us valuable information about the behavior of the system as a function of the nature of the distribution of activation energies. One of the most striking features of these plots is the change in the shape of the curve as a function of the mean activation energy μ (Fig. 3a). At the same temperatures, the two curves (Fig. 3a ($\mu = 7.5$) and ($\mu = 8.9$)) have different shapes, whereas the curves with the same activation energy (Fig. 3a) have similar shapes even though the systems have different temperatures.

A similar observation can be made about Fig. 3c for which the density of states of the initial state is higher than that of the saddle state in the rate-determining step (*i.e.* $\rho^I > \rho^{TS}$). The only source of difference in these curves is the difference between the activation energies. Similarly the variations in temperature do not seem to have any significant effect on the shape of the curves. These results show the effect of the change in E_a (*i.e.* enthalpic contribution) on the overall rate coefficient which is not obvious from the Arrhenius ($\ln k$ vs. $1/T$) plots. It is interesting to note that whereas variations in the entropic term in the rate equation result in changes in the shape of Arrhenius plots, the variance plots show the dependence of the rate coefficient on the changes in activation energy and are insensitive to temperature changes. This observation is magnified in the corresponding plots in which the derivative of C_p with variance is plotted against variance (Fig. 3b and d).

From a theoretical perspective, the Arrhenius plots represent only part of the overall picture. We believe that these rate-variance plots, in combination with rate-temperature plots will give an improved picture of the contribution of various thermodynamic terms to the overall rate.

Conclusion

We have presented a simple method to address the issue of reactions across a single barrier per pathway in a system of multiple pathways. The presence or absence of multiple pathways can play a very important role in visualizing the potential energy surface of a reaction. In this paper, we have addressed the issue of the effect of multiple parallel pathways on the overall kinetics of complex reactions. Though initial theoretical studies of macromolecular kinetics failed to agree with experiments, our study shows that the inclusion of the density of states in the initial and the saddle state corrects for the previous discrepancy between theoretical and experimental results. Our study indicates that presence of multiple parallel pathways and the density of states assignment to the initial and the saddle state plays a key role in providing this curvature observed in the Arrhenius plots in protein folding and unfolding reactions.

We also show the sensitivity of the Arrhenius plots to the density of states, as well as to the variance of the distribution of energies. The rates at which actual densities of states vary with temperature are presumably more complicated than the model discussed here; however the good agreement with experiments suggests that more sophisticated modifications of the dependence of densities of states on temperature will only improve the quality of the results. The success of our model in predicting the correct sign of curvature in Arrhenius plots is intended to stimulate experiments designed and conducted to observe the presence of multiple pathways, since initial comparisons to experimental results are very encouraging.^{3,9,10} More sophisticated calculations taking into account the densities of states of the initial and saddle states, as well as different rates at which these two densities of states increase will provide a more rigorous test of our theory.

Abbreviations

k_{obs} , rate coefficient of observed reaction; C_p , heat capacity at constant pressure; E_a , activation energy; ΔH , change in heat capacity, ΔH^\ddagger , activation enthalpy; ΔS , change in entropy; ΔS^\ddagger , activation entropy; $\rho(E)$, density of states.

Acknowledgements

We thank Drs R. Marcus, W. Eaton and F. Despa for helpful comments and useful suggestions. We also thank members of

our group for many enlightening discussions. We wish to thank Prof. John Schellman and his coworkers B. Chen and W. A. Baase for permission to reproduce their Fig. 3 from ref. 9 as our Fig. 2e.

This work was supported by Packard Foundation grant to RSB and TRS. MHZ is a Burroughs Wellcome interdisciplinary research Fellow. RSB wishes to acknowledge the support of a grant from the National Science Foundation.

References

- 1 J. D. Bryngelson, J. N. Onuchic, N. D. Socci and P. G. Wolynes, *Proteins: Struct. Func. Genet.*, 1995, **21**, 167.
- 2 H. S. Chan and K. A. Dill, *Proteins Struct. Func. Genet.*, 1998, **30**, 2.
- 3 L. M. Scalley and D. Baker, *Proc. Natl. Acad. Sci. USA*, 1997, **94**, 10636.
- 4 B. A. Krantz and T. R. Sosnick, *Nat. Struct. Biol.*, 2001, **8**, 1042.
- 5 L. B. Moran, J. P. Schneider, A. Kentsis, G. Reddy and T. R. Sosnick, *Proc. Natl. Acad. Sci. USA*, 1999, **96**, 10699.
- 6 P. G. Wolynes, Z. Luthey-Schulten and J. N. Onuchic, *Chem. Biol.*, 1996, **3**, 425.
- 7 J. D. Bryngelson and P. G. Wolynes, *J. Phys. Chem.*, 1989, **93**, 6902.
- 8 N. D. Socci, J. N. Onuchic and P. G. Wolynes, *J. Chem. Phys.*, 1996, **104**, 5860.
- 9 B. Chen, W. A. Baase and J. A. Schellman, *Biochemistry*, 1989, **28**, 691.
- 10 S. Jackson and A. R. Fersht, *Biochemistry*, 1991, **30**, 10436.