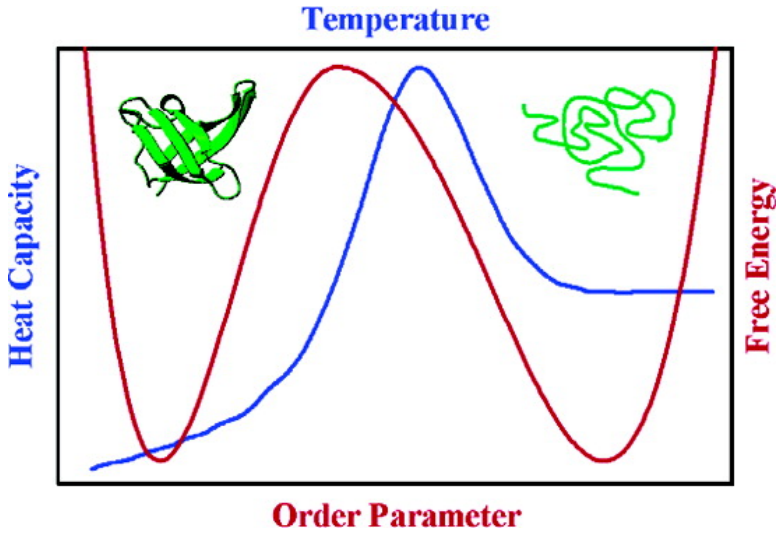
Communication

Direct Measurement of Barrier Heights in Protein Folding

Athi N. Naganathan, Jose M. Sanchez-Ruiz, and Victor Muoz

J. Am. Chem. Soc., **2005**, 127 (51), 17970-17971• DOI: 10.1021/ja055996y • Publication Date (Web): 01 December 2005



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 8 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
COMMUNICATIONS

Published on Web 12/01/2005

Direct Measurement of Barrier Heights in Protein Folding

Athi N. Naganathan,[†] Jose M. Sanchez-Ruiz,[#] and Victor Muñoz*,[†]

Department of Chemistry and Biochemistry, and Center for Biomolecular Structure and Organization, University of Maryland, College Park, Maryland 20742, and Departamento de Química Física, Facultad de Ciencias, Universidad de Granada, 18071 Granada, Spain

Received August 31, 2005; E-mail: vmunoz@umd.edu

In traditional chemical kinetics, reaction mechanisms are resolved by measuring the reaction free energy surface as a function of relevant coordinates. However, protein folding reactions require the formation of myriads of weak noncovalent interactions and the synchronization of large numbers of degrees of freedom,[1] thus making the mapping of detailed free energy surfaces a nearly impossible task. A practical solution involves using a projection onto a low-dimensional free energy surface that reproduces the average molecular behavior.[2] In this case, the dynamic term in the rate equation (i.e., the pre-exponential in $k = k_0 \exp(-\Delta G/RT)$) becomes a complex function of the projected surface and cannot be derived from first principles. The inherent limitation is, of course, that kinetic measurements of folding only provide relative barrier heights.

Much effort has been devoted in the past decade to obtain empirical folding speed limits, which could be used to estimate barrier heights from kinetic experiments[3] and to scale theoretical predictions of folding barriers.[4] Recently, we proposed a different approach in which folding barrier heights are extracted from differential scanning calorimetry (DSC) experiments.[5] The approach readily distinguishes between folding transitions with high free energy barriers and globally downhill.[5] Theoretically, our method could even provide absolute folding barriers, thus allowing discrimination between energetic and dynamic contributions to folding relaxation rates. There are, however, intrinsic limitations that need evaluation to assess its real impact on the study of protein folding. The method approximates the free energy surface to a Landau polynomial,[5] instead of attempting the notoriously ill-defined model-free inversion of the Laplace transform.[6] Therefore, proteins must have free energy barriers within the sensitivity of the approximation and energy landscapes smooth enough to produce simple thermodynamic and kinetic behavior.[7] Here we explore these issues.

As a first step, we investigated method sensitivity with a simple test based on retrieving the barrier heights from simulated DSC thermograms of theoretically calculated 1-D free energy surfaces. The 1-D free energy surfaces were produced with a variant of Zwanzig's statistical mechanical model for protein folding,[8] in which the stabilization energy is assumed to decrease exponentially as the degree of nativeness of the protein increases. In addition to its extreme simplicity, the model reproduces the scaling of thermodynamic and kinetic properties found for natural proteins. Figure 1 shows a plot comparing the barrier height extracted from the analysis of simulated DSC thermograms (i.e,. the parameter β) and the theoretical barrier height measured directly on the free energy surface. In the figure, it is apparent that the extracted and theoretical barriers are in very close agreement when the barrier is small and progressively diverge as the barrier increases. This is not surprising, as the method should be maximally sensitive when the population at the top of the barrier is maximal. The data tend asymptotically to ~15 kJ/mol (~6 *RT*), signaling the sensitivity limit of the procedure that corresponds to ~0.25% population at the top of the barrier. More practically, Figure 1 indicates that barrier heights can be measured quite precisely up to ~10 kJ/mol (~4 *RT* or 1.8%). From 10 up to 15 kJ/mol, there still is sensitivity to detect relative differences, but the extracted barrier heights are increasingly underestimated. In this analysis, we also found little dependence of the extracted barrier height on the native heat capacity baselines and specific shape of the (always smooth) theoretical free energy surfaces.

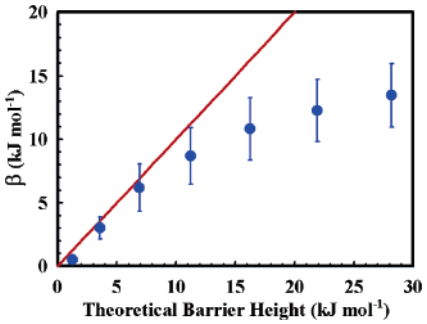


Figure 1. Comparison between theoretical barrier heights produced by statistical mechanical model and barrier heights extracted from analysis of simulated DSC thermograms (β).

At this point, the important question was whether this sensitivity range is useful to measure folding free energy barriers of natural proteins. This issue cannot be addressed directly because there are no alternative methods available to measure absolute folding barriers. Therefore, we used an indirect approach in which we compared the free energy barrier extracted from DSC experiments with folding rates in water obtained from kinetic experiments. For this purpose, we built a database of 13(+2) proteins for which reasonable DSC (no signs of irreversibility) and kinetic folding data are available and which do not show signs of populated intermediates (see Supporting Information). The database is by no means comprehensive and is quite uniform in protein size (i.e., 64 ± 15 residues). However, it does include proteins from the three main structural classes and spans a wide range of folding times (i.e., 4 orders of magnitude).

To deal with the heterogeneity in the available DSC data, we devised a general analysis scheme. In two of the proteins, we used our own DSC data, analyzed exactly as explained before.[5] For all other proteins, we digitized the original published data. We then analyzed the DSC thermograms using the lowest temperature point to set a native baseline with temperature dependence as predicted by Freire's empirical correlation.[9] This procedure rendered good fits for a large fraction of the proteins. Figure 2A shows data and fit for CspB as an example within that group. The inset shows the

[†] University of Maryland.
[#] Universidad de Granada.

10.1021/ja055996y CCC: $30.25

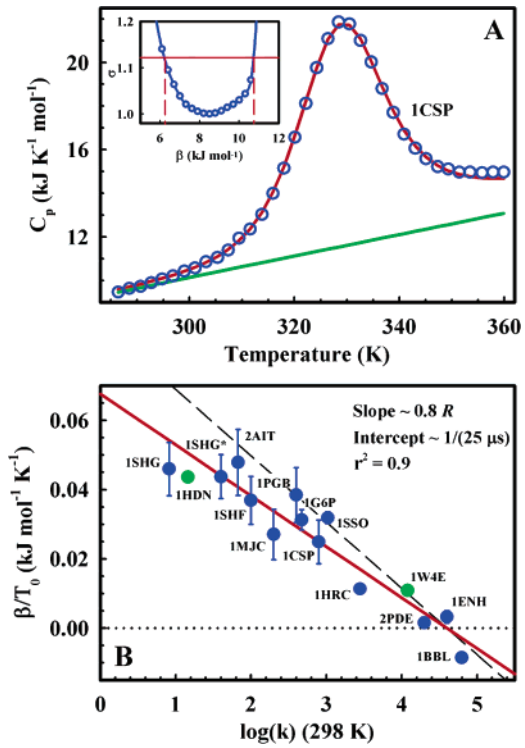


Figure 2. (A) DSC thermogram of CspB (1CSP) (blue circles) and fit to variable-barrier model (red line), together with native baseline (green line). Inset: sum of least squares as a function of parameter β in a grid search. The red line signals a 95% confidence interval. (B) Correlation between folding rates at 298 K and the ratio between barrier height (β) and characteristic temperature (T_0). The dashed line shows the expectation for a slope of R. For 1W4E, the folding rate at 298 K was obtained by scaling the available rate at 325 K for the changes in water viscosity (i.e., a factor of ~2 decrease). For HDN, slight aggregation was originally reported, but we could still obtain a reasonable fit. Neither of these two proteins was included in the correlation.

quality of the fit as a function of the barrier height (β), from which we derive the error in the measurement. For the remaining proteins, such baseline either crossed the thermogram (three cases) or could not reproduce the low-temperature data (two cases). For the first group, we allowed the baseline to downshift during the fitting, while for the second group, we fixed the native baseline to the lowest temperature data points (see Supporting Information).

The obtained free energy barrier heights range from negative (i.e., −3 kJ/mol) for the global downhill folder BBL to ~18 kJ/mol for Tendamistat (see Supporting Information for all data). These barriers correspond to the characteristic temperature (T_0) of the protein, which spans a range of ~60 K. For the comparison with the folding rates at 298 K, we use the ratio between the barrier height (β) and T_0. This ratio corrects for differences in stability and facilitates comparison of energy scales. Figure 2B shows the correlation between β/T_0 and the logarithm of the folding rates. The correlation coefficient (r^2) is 0.9, and the slope is ~0.8 R, indicating that the two parameters do indeed have similar energy scales. A slope slightly below R and the skewing on the high barrier limit agree with expectations from the theoretical analysis (Figure 1). The level of agreement is remarkable, given that there is no significant correlation between folding rates and protein size ($r^2 < 0.2$) or unfolding enthalpy ($r^2 = 0.25$) for this dataset. Furthermore, the correlation with β values directly is of similar quality ($r^2 = 0.86$, slope ~0.9 RT), possibly because the net temperature dependence of folding free energy barriers is weak as result of enthalpy−entropy compensations.[10]

These results confirm the suitability of our approach to measure free energy barriers in protein folding. Overall, the range of barrier heights measured from DSC experiments agrees with estimates based either on speed limits[3] or on size-scaling[11] arguments. For cold shock protein from *T. maritima* (1G6P), for example, we obtain a barrier of 11.4 ± 1.1 kJ/mol, well within the limits obtained by single molecule spectroscopy.[12] The procedure also detects differences between homologous proteins (e.g., 1G6P and 1CSP). The sensitivity reaches the level of detecting changes induced by single mutations. An example is E3BD, a structural homologue of the downhill protein BBL. The wild-type sequence has no net barrier (2PDE in Figure 2B), while a single (although evolutionarily nonconservative) F → W mutation results in a small hump of ~4 kJ/mol (1W4E in Figure 2B). From our results, it appears that single domain proteins can be classified in three distinct groups: with marginal or no barriers (<2 *RT*, or 0.017 in Figure 2B), two-state-like (>4 *RT*, or 0.033 in Figure 2B), and twilight zone proteins (<4 *RT* and >2 *RT*). The latter group should comply with most criteria for two-stateness, but show significant sensitivity to perturbations.

In Figure 2B, the intercept for 0 barrier is found at ~1/(25 μs), corresponding to an average pre-exponential of ~40 000 s^{-1} at 298 K. This value is ~3 times slower than a recent upper limit estimated from internal friction measurements on cytochrome *c*.[13] As a pre-exponential, this average value applies only to proteins with significant barriers and should not be confused with a folding "speed limit".[14] In proteins with no barriers, a favorable free energy gradient will speed up folding (e.g., 1BBL in Figure 2B). Furthermore, other factors, such as sequence, structure, stability, and temperature (e.g., our pre-exponential scales to >100 000 s^{-1} at temperatures more typical of T-jump experiments just from the decrease in water viscosity), also affect the folding speed of proteins with marginal barriers. For these proteins, the combination of our calorimetric analysis with kinetic measurements opens the opportunity of measuring individual diffusion coefficients for folding.

Acknowledgment. This work was supported by NSF Grant MCB-0317294 and NIH Grant GM066800-1 (V.M.) and Spanish Ministry of Education and Science Grant BIO2003-0229 (J.M.S.R.).

Supporting Information Available: List of the (13+2) proteins used, with references to the original DSC and kinetic data, and description of the fits performed and fitting parameters obtained. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Bryngelson, J. D.; Onuchic, J. N.; Socci, N. D.; Wolynes, P. G. *Proteins: Struct., Funct., Genet.* **1995**, *21*, 167−195.
(2) Onuchic, J. N.; LutheySchulten, Z.; Wolynes, P. G. *Annu. Rev. Phys. Chem.* **1997**, *48*, 545−600.
(3) Kubelka, J.; Hofrichter, J.; Eaton, W. A. *Curr. Opin. Struct. Biol.* **2004**, *14*, 76−88.
(4) Henry, E. R.; Eaton, W. A. *Chem. Phys.* **2004**, *307*, 163−185.
(5) Muñoz, V.; Sanchez-Ruiz, J. M. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17646−17651.
(6) Chan, H. S. *Proteins: Struct., Funct., Genet.* **2000**, *40*, 543−571.
(7) Chan, H. S.; Shimizu, S.; Kaya, H. *Methods Enzymol.* **2004**, *380*, 350−379.
(8) Zwanzig, R. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 9801−9804.
(9) Freire, E. *Protein Stability and Folding*; Humana Press: Totowa, NJ, 1995.
(10) Akmal, A.; Muñoz, V. *Proteins: Struct., Funct., Bioinf.* **2004**, *57*, 142−152.
(11) Naganathan, A. N.; Muñoz, V. *J. Am. Chem. Soc.* **2005**, *127*, 480−481.
(12) Schuler, B.; Lipman, E. A.; Eaton, W. A. *Nature* **2002**, *419*, 743−747.
(13) Hagen, S. J.; Qiu, L. L.; Pabit, S. A. *J. Phys.: Condens. Matter* **2005**, *17*, S1503−S1514.
(14) Yang, W. Y.; Gruebele, M. *Nature* **2003**, *423*, 193−197.

JA055996Y