



# The Folding Pathway of T4 Lysozyme: An On-pathway Hidden Folding Intermediate

Hidenori Kato<sup>†</sup>, Ngoc-Diep Vu<sup>†</sup>, Hanqiao Feng, Zheng Zhou and Yawen Bai<sup>\*</sup>

Laboratory of Biochemistry  
Center for Cancer Research  
National Cancer Institute  
Building 37, Room 6114E  
Bethesda, MD 20892, USA

T4 lysozyme has two easily distinguishable but energetically coupled domains: the N and C-terminal domains. In earlier studies, an amide hydrogen/deuterium exchange pulse-labeling experiment detected a stable submillisecond intermediate that accumulates before the rate-limiting transition state. It involves the formation of structures in both the N and C-terminal regions. However, a native-state hydrogen exchange experiment subsequently detected an equilibrium intermediate that only involves the formation of the C-terminal domain. Here, using stopped-flow circular dichroism and fluorescence, amide hydrogen exchange-folding competition, and protein engineering methods, we re-examined the folding pathway of T4-lysozyme. We found no evidence for the existence of a stable folding intermediate before the rate-limiting transition state at neutral pH. In addition, using native-state hydrogen exchange-directed protein engineering, we created a mimic of the equilibrium intermediate. We found that the intermediate mimic folds with the same rate as the wild-type protein, suggesting that the equilibrium intermediate is an on-pathway intermediate that exists after the rate-limiting transition state.

Published by Elsevier Ltd.

<sup>\*</sup>Corresponding author

**Keywords:** T4 lysozyme; hydrogen exchange; protein folding; hidden intermediate; protein engineering

## Introduction

To understand the mechanism of protein folding, it is necessary to characterize the energy landscape of protein folding in detail, which includes partially unfolded intermediates, transition states, and their orders in the folding process. A number of methods have been developed to do so. Among them, amide hydrogen exchange (HX) has been particularly useful for obtaining the backbone structural information of folding intermediates at the level of residues and has been applied to study protein folding in both kinetic and equilibrium modes.<sup>1–4</sup> For example, in the kinetic mode, an amide hydrogen/deuterium exchange pulse-labeling method has identified kinetic folding intermediates that transiently accumulate before the rate-limiting

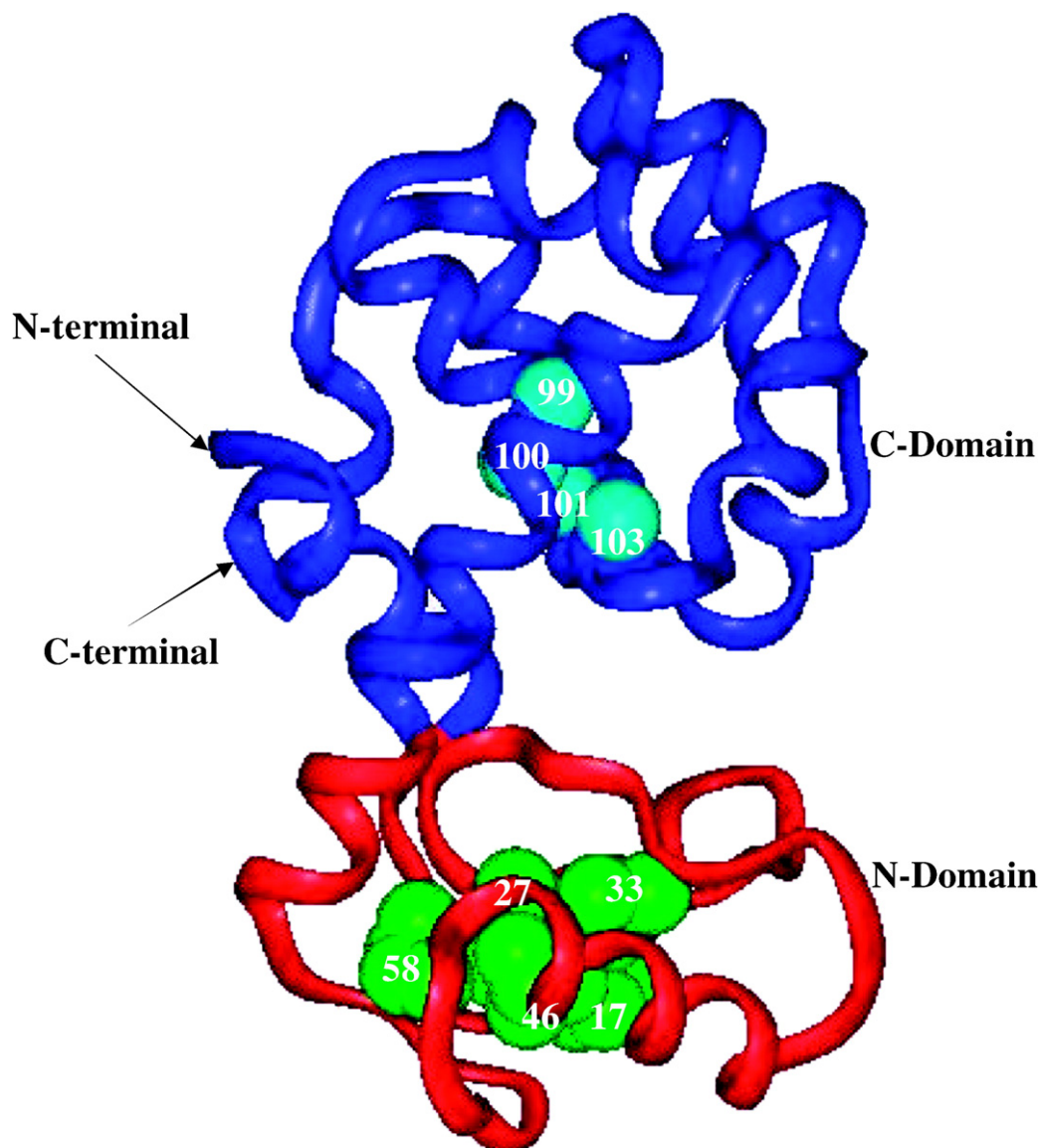
transition state.<sup>5–9</sup> In the equilibrium mode, a native-state hydrogen exchange method has detected equilibrium folding intermediates that populate only in miniscule amount under native conditions.<sup>10–14</sup> Both methods have been applied to T4 lysozyme.<sup>12,15</sup>

T4 lysozyme has two easily distinguishable but energetically coupled domains: the N and C-terminal domains (see Figure 1). The C-terminal domain has an  $\alpha$ -helical structure, whereas the N-terminal domain includes both  $\alpha$ -helices and  $\beta$ -strands. The earlier amide hydrogen/deuterium exchange pulse-labeling experiment on T4 lysozyme suggests that a stable submillisecond folding intermediate transiently accumulates before the rate-limiting transition state. It involves the partial formation of structures in both the N and C-terminal domains<sup>15</sup> and is  $\approx 3$  kcal/mol more stable than the unfolded state in the presence of 1.5 M urea at pH 6 and 25 °C. In contrast, the native-state hydrogen exchange experiment has identified an intermediate with the N-terminal domain (from residues 14 to 65) unfolded and C-terminal domain (from residues 1 to 13 and 66 to 164) folded.<sup>12</sup>

<sup>†</sup> H.K. and N.D.V. contributed equally to this work.

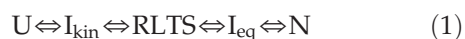
Abbreviations used: HX, hydrogen exchange; HSQC, heteronuclear single quantum coherence.

E-mail address of the corresponding author:  
yawen@helix.nih.gov



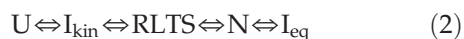
**Figure 1.** Structure of T4 lysozyme. The region with blue ribbon represents the N-terminal domain. The region with red ribbon represents the C-terminal domain. Small balls in cyan represent the four slowest exchanging amide protons, including residues 99, 100, 101, and 103. The hydrophobic residues I17, I27, L33, L46, and L58 in the N-terminal domain are shown in CPK models.

The simplest folding pathway for T4 lysozyme would be a sequential reaction with both intermediates on the folding pathway, as described in scheme (1):



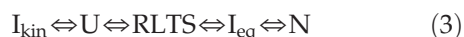
Here,  $I_{\text{kin}}$  and  $I_{\text{eq}}$  represent the intermediates that are detected in the pulse-labeling and the native-state hydrogen exchange experiments, respectively. RLTS represents the rate-limiting transition state.  $I_{\text{eq}}$  is placed after the rate-limiting transition state because it is not observed in the kinetic folding experiment. However, the difference in the structures of the kinetic and equilibrium intermediates argues that it is unlikely that the two intermediates are on a sequential folding pathway. This is because

it would require the unfolding of the N-terminal region in  $I_{\text{kin}}$  in order to form  $I_{\text{eq}}$ . Therefore, other possibilities need to be considered to define the folding pathway of T4 lysozyme. For example, the intermediate identified in the pulse-labeling experiment may be an on-pathway intermediate, whereas the intermediate detected by the native-state HX method is an off-pathway intermediate, as described in scheme (2):

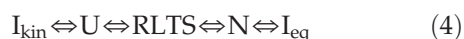


However, the observation of an intermediate in a kinetic folding experiment alone does not necessarily mean that the intermediate is on the kinetic folding pathway.<sup>16–18</sup> An off-pathway intermediate can also produce the same observable folding kinetics as an

on-pathway intermediate when the folding rate from the unfolded state to the intermediate is too fast to be measurable by the conventional stopped-flow machine. In other words, the available experimental results can also be interpreted by assuming that the  $I_{\text{kin}}$  is an off-pathway intermediate, whereas the  $I_{\text{eq}}$  is an on-pathway intermediate that exists after the rate-limiting transition state, as shown in the folding reaction scheme (3):



Other possibilities include that both intermediates are off the folding pathway, as shown in scheme (4):



To investigate the kinetic role of the two intermediates, we re-examined the kinetic folding of T4 lysozyme by using a number of methods, including stopped-flow circular dichroism (CD) and fluorescence, hydrogen exchange-folding competition, and the native-state hydrogen exchange-directed protein engineering.<sup>19</sup> We found no evidence for the population of a stable submillisecond folding intermediate under neutral pH. In contrast, we were able to create a mimic of the equilibrium intermediate identified in the native-state HX experiment. Importantly, the mimic of the equilibrium intermediate folded with the same rate as the native state, suggesting that the equilibrium intermediate is an on-pathway intermediate that exists after the rate-limiting transition state.

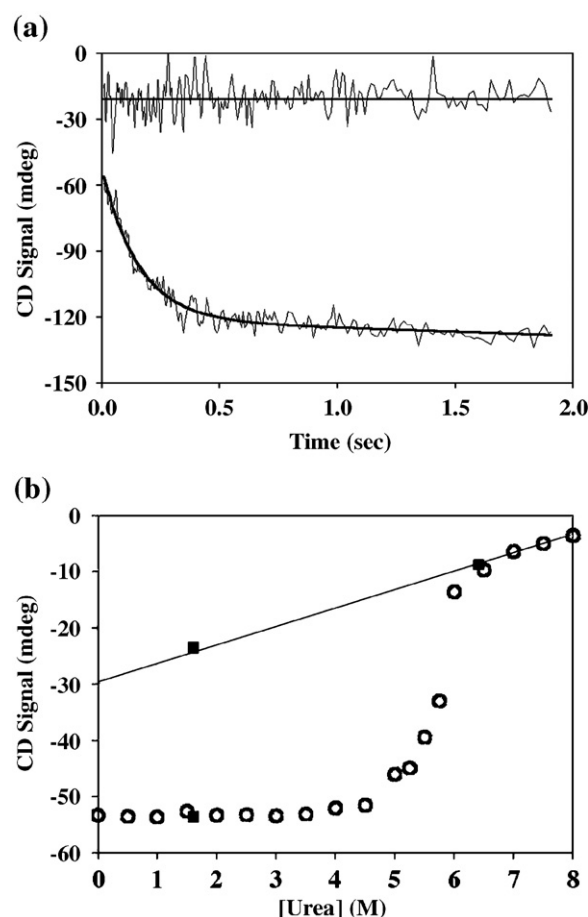
## Results and Discussion

### The submillisecond folding events

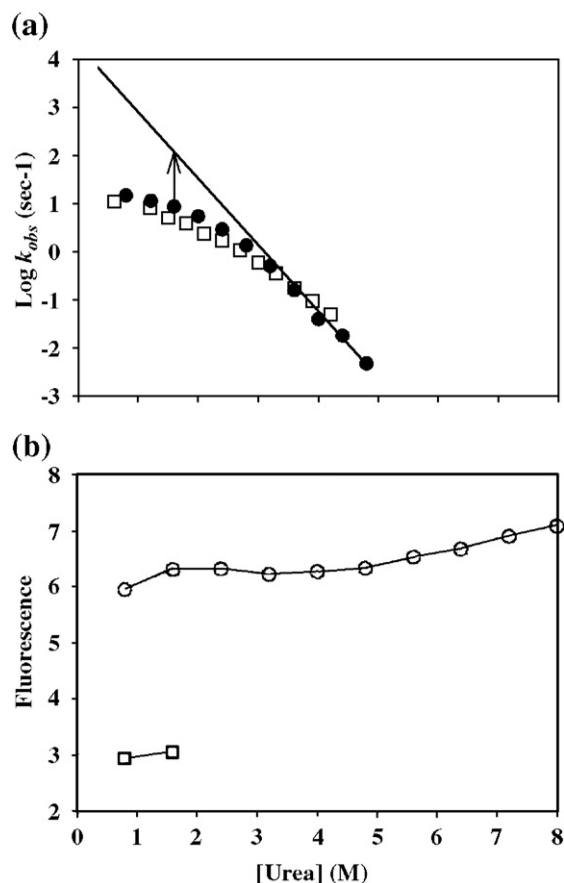
For the past several years, the views and experimental results regarding the sub-millisecond folding events have been controversial. Several proteins that were considered to have stable submillisecond intermediates based on earlier experiments have been called into question by more recent experiments.<sup>20</sup> One of the problems is that transient aggregations have occurred in the folding of several proteins.<sup>21–23</sup> Since the earlier pulse-labeling experiment on T4 lysozyme was done at extremely high protein concentrations (7 mg/ml during folding) and its kinetic folding intermediate is different from the equilibrium intermediate, a re-examination of the kinetic folding pathway appears to be necessary for clarifying this issue. Toward this goal, we first examined the CD and fluorescence signals at 3 ms after the initiation of the folding of T4 lysozyme. Folding of T4 lysozyme has two observable kinetic phases as monitored by both CD and fluorescence.<sup>15,24</sup> The fast phase has a rate constant  $\sim 17 \text{ s}^{-1}$  and occupies  $\approx 80\%$  of the observed amplitude. The slow phase has  $\approx 20\%$  of the total amplitude, which is attributed to the process of proline isomerization. Studies on the folding of T4 lysozyme have focused on the fast

phase. Unfolding of the protein at high concentrations of denaturant shows a single exponential kinetic phase. Figure 2(a) shows the kinetic traces of CD signals measured at 1.6 M urea and 6.0 M urea with the starting protein sample at 8 M urea (pH 6.0 and 25 °C). Figure 2(b) compares the results from the kinetic folding experiments with those from equilibrium unfolding experiments. No missing amplitude of the CD signal at 1.6 M urea was observed when the signals of the unfolded state at high concentrations of urea were linearly extrapolated to low concentrations of urea. It should be noted that in these studies T4 lysozyme is referred to a mutant WT\* C54T/C97A.

We further measured the folding rate by the stopped-flow fluorescence at pH 6.0 and 9.6, corresponding to the folding and pulse conditions in the earlier pulse-labeling experiment. Figure 3(a) shows the logarithms of the folding rates as a function of denaturant concentrations. A non-linear behavior occurred. If one assumes that this non-linear behavior was caused by the existence of an early folding intermediate, it can be estimated that at



**Figure 2.** Kinetic folding and equilibrium unfolding measured by CD. (a) Kinetic traces monitored by CD (222 nm) at 6.4 M (up curve) and 1.6 M (bottom curve) urea, respectively. (b) Equilibrium unfolding results (open circles) and the normalized values of the CD signals obtained at 3 ms (filled squares). The values have been normalized to the same protein concentration.



**Figure 3.** Folding rate constants and fluorescence signals for the early folding events determined in kinetic folding experiments. (a) Logarithms of the folding rate constants as a function of urea concentrations at pH 6 (filled circles) and pH 9.6 (open squares) determined in stopped-flow fluorescence experiments. At 1.5 M urea, the measured folding rate is smaller than the predicted value from the linear extrapolation by about a factor of 10, suggesting that a putative intermediate, if it exists, would be more stable than the unfolded state by  $\sim 1.4$  kcal/mol. (b) Fluorescence signals (open circles) observed at 3 ms after the initiation of folding as a function of urea concentrations in stopped-flow fluorescence experiments. The non-cooperative unfolding suggests the absence of a stable intermediate at 1.5 M urea. The data (open squares) indicate fluorescence signals for the native state.

1.5 M urea, at which the earlier pulse-labeling experiment was performed, the putative intermediate would have a stability of  $\sim 1.4$  kcal/mol (see legends in Figure 3(a)). This value, however, is significantly smaller than the value ( $\sim 2.8$  kcal/mol) measured in the earlier pulse-labeling experiment.<sup>15</sup> In addition, we also measured the fluorescence signal at 3 ms of the folding as a function of denaturant concentrations. At 3 ms, the putative intermediate should have formed without converting to the native state. Therefore, the fluorescence signal should fully come from the proposed intermediate. Since the signal shows no cooperative transition as a function of denaturant, it does not support the existence of a stable intermediate (see

Figure 3(b)). Therefore, these results do not support the existence of a stable submillisecond folding intermediate.

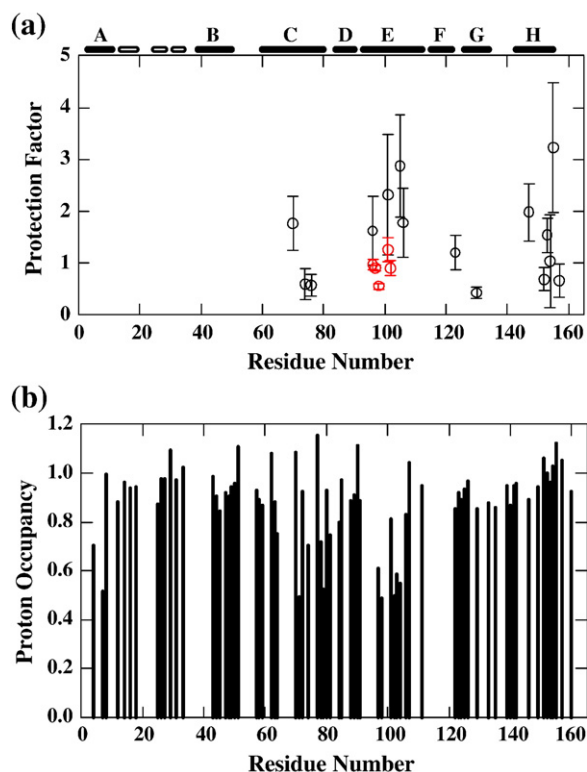
### Hydrogen exchange-folding competition and dead-time pulse-labeling

A hydrogen exchange pulse-labeling experiment involves two important steps: folding at neutral pH and pulse at high pH. If one of the two steps involves transient aggregations, it would lead to the protection of amide protons and the conclusion for the existence of early folding intermediates. To examine the two steps separately, we performed the hydrogen exchange-folding competition experiment at pH 6.8 and 25 °C.<sup>25,26</sup> Under this condition, the folding rate is close to the intrinsic exchange rates of many amide protons of T4 lysozyme, which allows more accurate measurement of the protection factors ( $P_f = k_{\text{int}}/k_{\text{ex}}$ ) for amide protons in the submillisecond state (see Materials and Methods). Here,  $k_{\text{int}}$  and  $k_{\text{ex}}$  are the intrinsic exchange rate constant<sup>27</sup> and the measured exchange rate constant, respectively. Because most amide protons can exchange in the native state through local structure fluctuations, only 15 amide protons that are heavily protected in the native state could be used to characterize the structure formation of the submillisecond state in the competition experiment. We found that the protection factors for these amide protons ranged from 0.5 to 3.5, a typical range for amide protons in an unfolded polypeptide chain (Figure 4(a)). These include four amide protons from residues R96, N101, Q105 and M106, which are shown to be heavily protected with  $P_f \sim 100$  in the earlier pulse-labeling experiment.<sup>15</sup> Thus, these results suggest that the observation of the protections of amide protons in the earlier pulse-labeling experiment is not due to the formation of a submillisecond intermediate at neutral pH.

To examine the folding at pulse condition, we performed a dead-time pulse-labeling experiment, in which the unfolded protein sample in  $^2\text{H}_2\text{O}$  and 8.5 M D-urea was mixed with a  $\text{H}_2\text{O}$  solution at pH 10.2 by a ratio of 1:9. After 13 ms, the sample was quenched at pH 3.0. At pH 10.2 and 25 °C, unprotected amide deuterons can exchange with solvent protons in less than 1 ms.<sup>27</sup> Figure 4(b) shows the proton occupancy measured at pH 3.0 and 20 °C, the same condition used in the earlier pulse-labeling experiment. We found that some amide protons in the A, C, D, and E helices were strongly protected within the dead time ( $\sim 1$  ms) of our quenched-flow instrument. However, no significant protections for amide protons in the N-terminal domain occurred. Thus, these results suggest that the protections observed in the earlier pulse-labeling experiment occurred during the pulse at high pH.

Although amide proton protections were observed in the dead-time pulse labeling experiment, the pattern of the protection is different from that in the earlier pulse-labeling experiment. For example, some amide protons in the N-terminal domain are also strongly protected in the earlier pulse-labeling





**Figure 4.** Hydrogen exchange-folding competition and dead-time pulse-labeling experiments for WT\* and its mutants. (a) Competition experiment for wild-type at pH 6.8 and 25 °C (black) and the T4\_IM at pH 7.1 and 25 °C (red). (b) Dead-time pulse-labeling experiment for WT\* with labeling time of 13 ms at pH 10.2. The error bars in (a) and (b) were estimated from three independent experiments. The location of the secondary structures is shown on top of the Figure. The open bar indicates the position of  $\beta$ -strand. The folding rates ( $k_f$ ) used to calculate exchange rates ( $k_{ex}$ ) by equation (9) are 14 and 34  $s^{-1}$  for wild-type and the T4\_IM, respectively. The larger folding rate of the T4\_IM is likely due to the shorter linker.

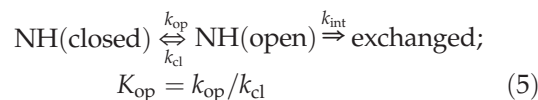
experiment, whereas they are not protected in the current experiment (Figure 4(b)), suggesting that pulse-labeling results are sensitive to subtle differences in experimental conditions. Since we observed no concentration-dependent folding rates down to micromolar protein concentrations (unpublished results), it is intriguing that protections of amide protons occurred at high pH but the folding rate remains the same as that at neutral pH (see Figure 3(a)). One possible interpretation is that the interactions formed in the early stage of the folding remain formed in the rate-limiting transition state.

### Measurement of unfolding rate by hydrogen exchange at high pH

A non-linear logarithm of the folding rate constant as a function of denaturant concentration is commonly taken as evidence for the existence of an early folding intermediate. However, it has been shown recently that the non-linear behavior could also be caused by the change of rate-limiting transition

states<sup>28</sup> or a funnel-like energy landscape.<sup>29,30</sup> For example, in the case of barnase, the curvature in the folding arm of the logarithm of folding/unfolding rate constants *versus* denaturant concentration (chevron plot) is caused by the switch of the rate-limiting transition states rather than population of stable submillisecond intermediates. The evidence is that the logarithm of the unfolding rate constant is curved down equally as the logarithm of the folding rate constant at low concentrations of denaturant (for details, please read Takei *et al.*<sup>26</sup>). In contrast, if the curvature of the logarithm of the folding rate constant is caused completely by the formation of a stable early folding intermediate, the logarithm of the unfolding rate constant should be a linear function of denaturant concentrations.

To distinguish the above two possibilities, we performed amide hydrogen exchange experiment at high pH at zero denaturant, which allows the measurement of unfolding rate constant under native conditions. Exchangeable amide hydrogen atoms (NH) involved in the hydrogen-bonded structure can exchange with solvent hydrogen atoms only when they are transiently exposed to solvent in some kind of closed-to-open reactions,<sup>31</sup> as indicated in equation (3). Here,  $k_{op}$  is the kinetic opening rate constant and  $k_{cl}$  is the kinetic closing rate constant.  $k_{int}$  is the intrinsic exchange rate constants in the open form, which is normally derived from unfolded peptide models:<sup>27</sup>



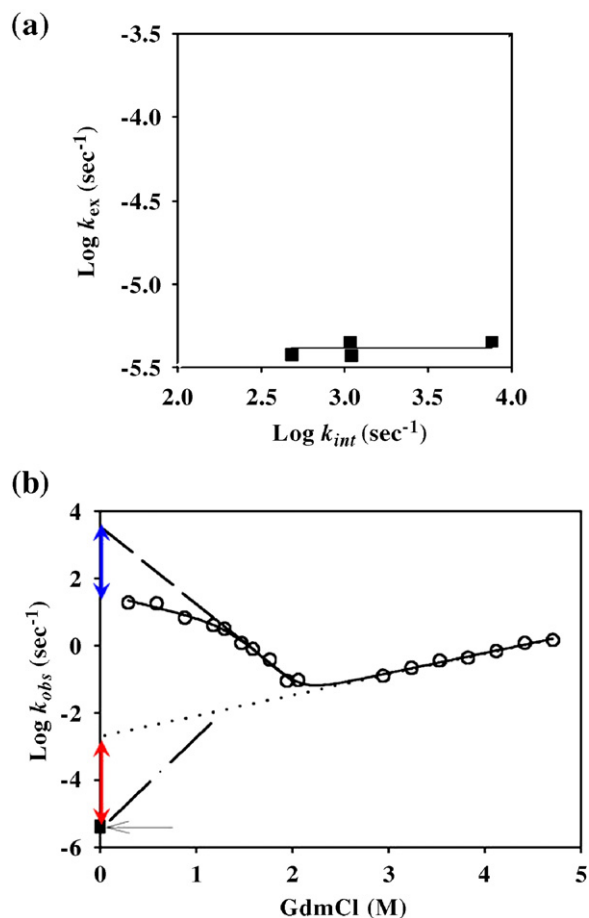
When the closing reaction is faster than  $k_{int}$  (EX2 condition), the exchange rate of any hydrogen,  $k_{ex}$ , is determined by its chemical exchange rate in the open form multiplied by the equilibrium opening constant,  $K_{op}$ :

$$k_{ex} = K_{op} \times k_{int} \quad (6)$$

This leads to free energy for the dominant opening reaction:

$$\Delta G_{HX} = -RT \ln K_{op} = -RT \ln (k_{ex}/k_{int}) \quad (7)$$

where  $R$  is the gas constant and  $T$  is the temperature. When the closing reaction is slower than  $k_{int}$  (EX1 condition), the exchange rate,  $k_{ex}$ , is determined by the opening rate constant  $k_{op}$ . For the slowly exchanging amide protons that can only exchange through global unfolding, the  $k_{op}$  will be the global unfolding rate constants. Thus, under EX1 conditions, the measurement of the exchange rates of slowly exchanging amide protons provides a way to measure the global unfolding rates of proteins. Figure 5(a) plots the logarithms of the measured exchange rates for the four slowest exchanging amide protons (99, 100, 101, and 103) *versus* their intrinsic exchange rates at pH 9.6 and 25 °C. The measured values are very similar among the four protons even though their intrinsic exchange rates spread over an order of magnitude, suggesting that



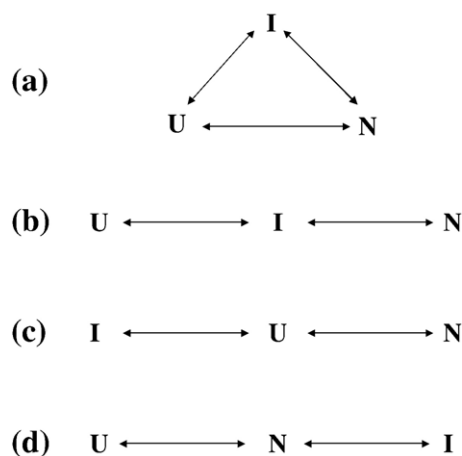
**Figure 5.** Folding and unfolding rates at pH 9.6 and 25 °C. (a) Test of hydrogen exchange mechanism. The independence of the measured hydrogen exchange rates for the slowest exchanging amide protons on the intrinsic exchange rates suggests that the exchange occurs in EX1 mechanism and the measured exchange rate represents the unfolding rate. (b) Folding and unfolding rate constants measured by stopped-flow and hydrogen exchange. The continuous curve is a fit to a three-state model with equations of (8)–(10). GdmCl was used to obtain the broad unfolding arm of the chevron plot. The broken line in the folding arm of the chevron plot represents the linear extrapolation of the data near the bottom region of the chevron plot. The linear dotted line represents the extrapolation of the unfolding arm of the chevron plot. The dashed-dotted line connected to the measured unfolding rate (filled square indicated by the single arrow) by the HX method represents a possible deviation of  $\log k_u$  from the linear region. At zero denaturant, the measured unfolding rate by HX under EX1 condition is much smaller than that extrapolated from the unfolding arm of the chevron plot. The difference in the unfolding rate (red double arrow) is similar to that between the measured folding rate and the linearly extrapolated folding rate (blue double arrow), suggesting that the curvature in the folding arm of the chevron plot is caused by the movement or switch of the rate-limiting transition state rather than the existence of a stable early folding intermediate.

the exchange indeed occurs through an EX1 mechanism. Therefore, the measured exchange rates are the unfolding rate of the native state.

We also performed stopped-flow folding/unfolding experiments at pH 9.6 and 25 °C. In this experiment, we used GdmCl, since it allows the unfolding arm of the chevron plot to be measured in a broader range. Figure 5(b) compares the unfolding rate measured by the hydrogen exchange method with those measured in the kinetic unfolding experiments at high denaturant concentrations. The unfolding rate constant measured by the hydrogen exchange method at zero denaturant is  $\approx$ two orders of magnitude smaller than the value extrapolated linearly from the logarithm of the unfolding rates at high concentrations of denaturant. This difference is very close to that between the measured folding rate at zero denaturant and the value extrapolated from the linear region near the bottom of the chevron plot. In other words, the directly measured  $k_u/k_f$  at zero denaturant is close to that linearly extrapolated from the linear region near the bottom of the chevron plot. Thus, the change of the rate-limiting transition state rather than the population of a stable early folding intermediate caused the curvature of the chevron plot.

### The order of a folding reaction with an intermediate

A general description of a protein folding reaction with an intermediate is the triangle scheme as shown in Figure 6(a). Depending on the relative conversion rate of each folding step, three special cases can occur. The first case is that the direct folding from U to N is the slowest and I is considered as on-pathway (Figure 6(b)). The second case is that the direct folding from I to N is the slowest and I is considered as off-pathway (Figure 6(c)). The third case is that the folding from U to I is the slowest and I is considered as kinetically isolated from U or I is a conformation state that fluctuates in the ensemble of the native state (Figure 6(d)).



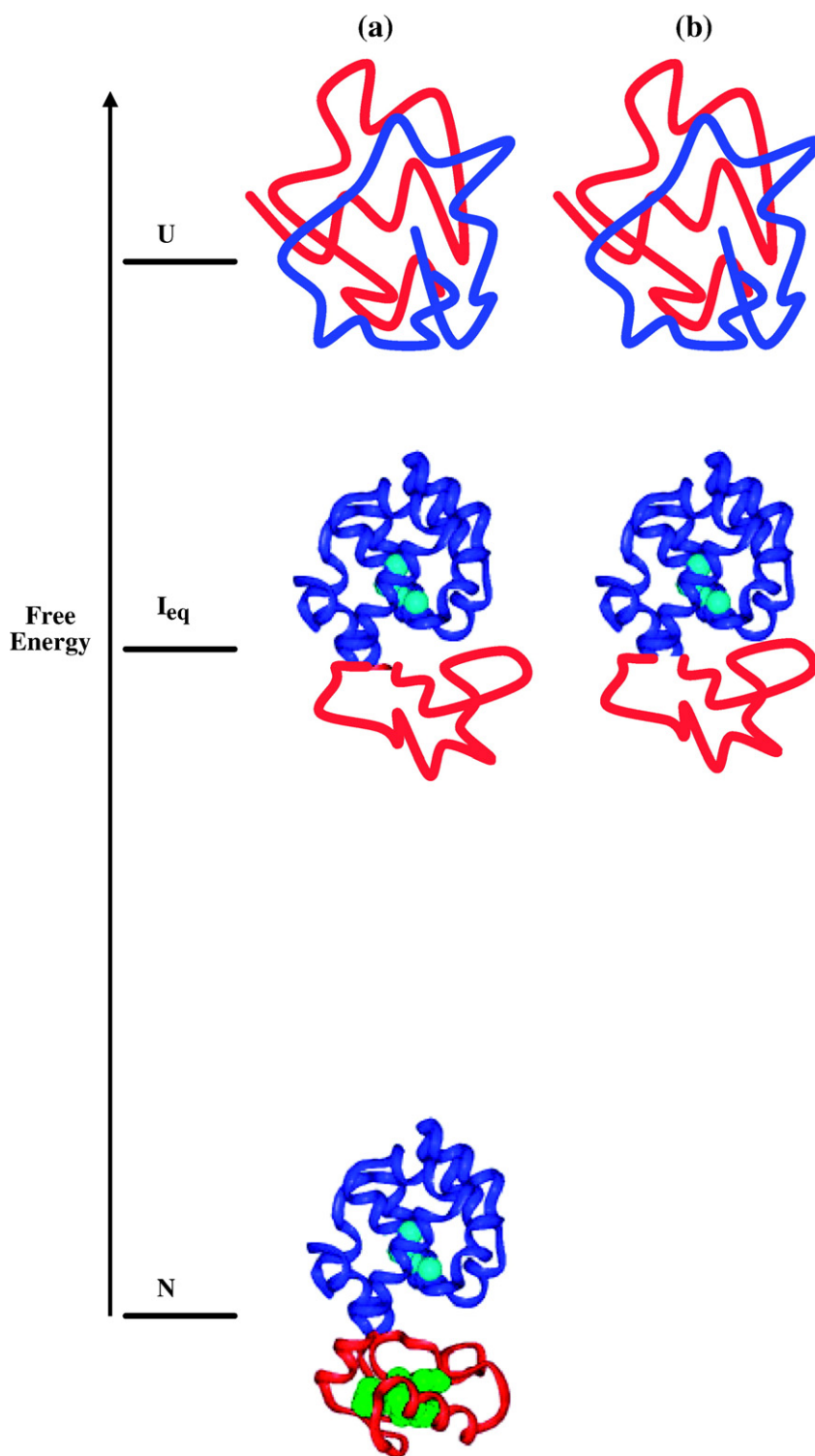
**Figure 6.** Folding reaction schemes with an intermediate. (a) A general triangle scheme. (b) On-pathway intermediate. Direct folding from U to N is the slowest. (c) Off-pathway intermediate. Folding from I to N is the slowest. (d) The intermediate is kinetically isolated. Folding from U to I is the slowest.

Accordingly, if the apparent folding from U to I can be shown to have the same rate for the apparent folding from U to N, the third case, i.e. a kinetically isolated I, can be excluded. Moreover, if I can be shown to be undetectable in the kinetic folding experiment, then the second case can be further disproved, which would leave the first case as the reaction scheme. In the following, we use protein engineering to obtain a mimic of the  $I_{eq}$  of T4 lysozyme and demonstrate that it folds from the

unfolded state with the same rate as the folding of the WT\*.

#### Populating $I_{eq}$ by a native-state hydrogen exchange-directed protein engineering

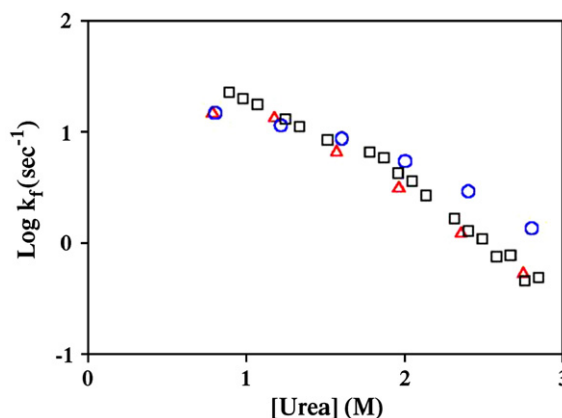
To measure the apparent folding rate from the unfolded state to the  $I_{eq}$ , we used the native-state hydrogen exchange-directed protein engineering to populate the  $I_{eq}$ .<sup>19</sup> Figure 7(a) illustrates the free



**Figure 7.** Illustration of a native-state hydrogen exchange-directed protein engineering for populating the partially unfolded intermediate. (a) Free energies and structures of the WT\*, intermediate, and unfolded state. The intermediate was identified by the native-state hydrogen exchange method. (b) Mutations in the N-terminal domain, which destabilize the native state, have no effect on the stability of the intermediate, make the intermediate the most stable species. Hydrophobic core residues in the N-terminal domain, including I17, I27, L33, L46, and L58, were mutated to Gly (see Figure 1).

energy and the structures of the fully unfolded state, the equilibrium intermediate, and the fully folded state, characterized by the native-state hydrogen exchange. Under native conditions, the native state is the most stable state and dominantly populates in solution. Although the intermediate and the unfolded state also exist in solution, they are not directly observable due to their miniscule population. To populate the intermediate without disrupting the folded region of the protein, mutations can be made in the unfolded region of the intermediate to destabilize the native state (see Figure 7(b)). In this case, we substituted five large hydrophobic residues (I17, I27, L33, L46, I58) in the N-terminal domain with glycine residues (see Figure 1).

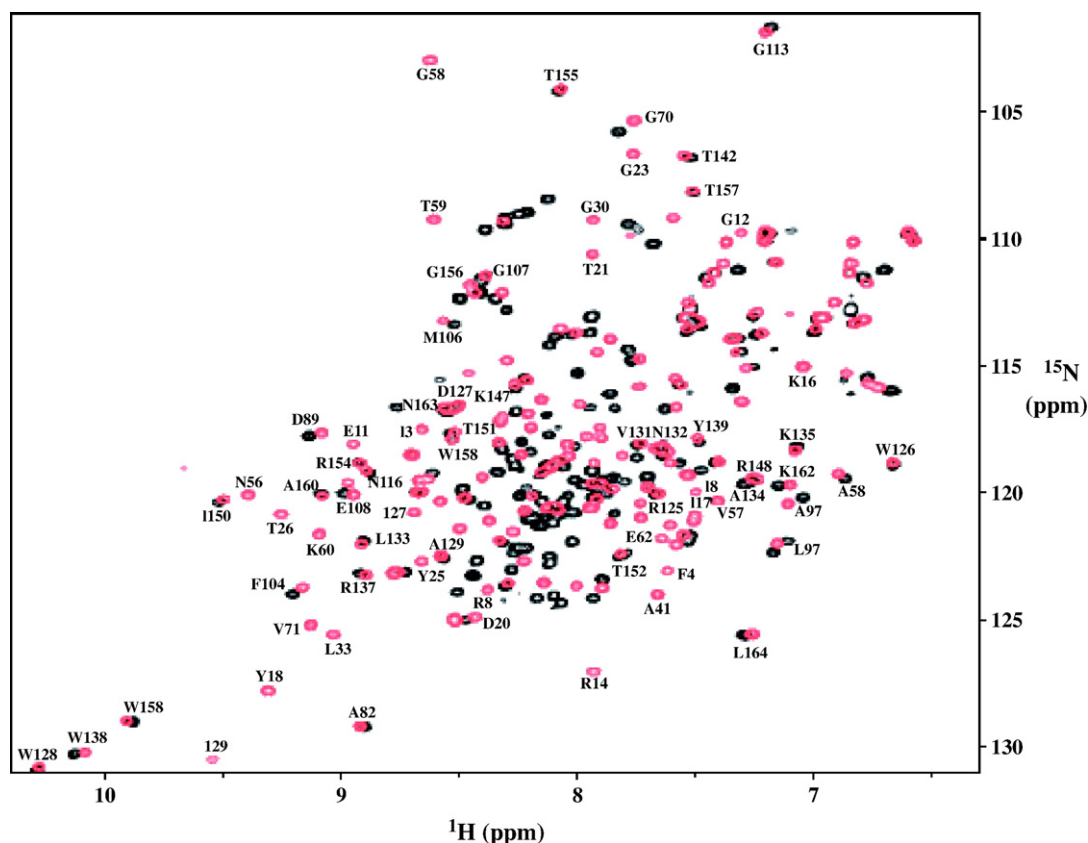
To test whether the intermediate has been populated, we assigned the resonances of the mutant by multi-dimensional NMR. Figure 8 shows the overlay of the  $^1\text{H}$ - $^{15}\text{N}$  heteronuclear single quantum coherence (HSQC) spectra of the WT\* and the mutant. The chemical shifts of the amides in the C-terminal domain of the mutant showed little change when compared with those of WT\*, whereas those of the N-terminal domain all move to the narrow region of  $8.0(\pm 0.5)$  ppm, a typical range for unfolded peptides. Thus, the NMR spectrum of this 5G-mutant is consistent with the structure of the intermediate detected by the native-state hydrogen exchange experiment.



**Figure 9.** The folding rate of WT\* (blue open circles), T4\_IM (red open triangle), 5G-mutant (black open square).

#### Kinetic folding rates of the $I_{eq}$ and the native-state

Figure 9 compares the logarithms of the folding and unfolding rates of the 5G-mutant and the WT\*. Both proteins have similar folding rate at low concentrations of denaturant, indicating that folding from the unfolded state to the  $I_{eq}$  is not slower than the folding from the unfolded state to the native state. Therefore, the  $I_{eq}$  is not kinetically isolated from the unfolded state under such conditions. This



**Figure 8.** Comparison of the HSQC spectra between WT\* (red) and the 5-Gly mutant (black).



conclusion was further confirmed with a new intermediate mimic T4\_IM that has the region from residues 17 to 58 substituted by a short glycine linker (see Figure 9 and accompanying article). For this second mutant, a high-energy native state does not exist.<sup>32</sup> Therefore, folding of the intermediate cannot go through the native state, providing more unequivocal evidence that the  $I_{eq}$  is an on-pathway intermediate and exists after the rate-limiting transition state. These results are consistent with an earlier experiment that showed that single mutations in the N-terminal domain in WT\* did not affect the folding rate.<sup>24</sup> In addition, we also performed the hydrogen exchange-competition experiment on T4\_IM. The protection factors of the amide protons that can be studied were very close to 1, suggesting that these protons do not form any stable structures on the submillisecond time scale (Figure 4(a)). It should be noted that, T4\_IM is less stable than WT\*. Only four amide protons in the E-helix (A97, A98, N101 and M102) can be studied in this experiment.

### Hidden folding intermediates

Most protein folding studies have focused on identifying and characterizing folding intermediates that exist before the rate-limiting transition states by studying the kinetic process of protein folding.<sup>33,34</sup> Hidden intermediates that exist after the rate-limiting transition state have been studied only recently. They are identified in two different ways. Some are observed directly in kinetic unfolding experiments under denaturing conditions, which include RNase A,<sup>35</sup> DHFR,<sup>36</sup> lysozyme,<sup>37</sup> and IL-1 $\beta$ .<sup>38</sup> Others are identified as equilibrium intermediates by native-state hydrogen exchange methods, which include Cyt *c*,<sup>10,39,40</sup> T4 Lysozyme,<sup>12</sup> Rd-apocyt *b*<sub>562</sub>,<sup>41</sup> a PDZ-3 construct,<sup>42</sup> and Barnase.<sup>43,44</sup> Structural information on hidden intermediates is mainly derived from native-state hydrogen exchange experiments. For T4 lysozyme, earlier protein engineering studies have also shown that the C-terminal region can fold independently.<sup>45</sup>

The observation of hidden intermediates in both single and multi-domain proteins with different folds suggests that they occur generally in protein folding, providing further evidence for the hypothesis that proteins use partially unfolded intermediates to solve the large-scale conformational search problem.<sup>30,46,47</sup>

### Conclusion

The kinetic folding of T4 lysozyme was re-examined using a number of methods, including stopped-flow CD and fluorescence, hydrogen exchange-folding competition, and protein engineering. No stable submillisecond folding intermediates were detectable at neutral pH. The intermediate identified in the earlier native-state hydrogen exchange experiment was populated by a

native-state hydrogen exchange-directed protein engineering method. It folded with the similar rate to WT\* at low concentrations of denaturant. These results suggest that the intermediate is on the folding pathway and exists after the rate-limiting step, consistent with the folding behavior of several other small proteins. The folding behavior of T4 lysozyme at high pH, however, is more complex and remains to be investigated.

## Materials and Methods

### Protein expression and purification

The expression system of a psuedo wild-type (C54T/C97A), WT\*, of T4 lysozyme was kindly provided by Dr Brian Matthews (University of Oregon). Glycine mutations were made using a quick-change mutation kit (Stratgen, CA). Proteins were expressed in *Escherichia coli* (BL21DE3) and purified using ion-exchange and reversed-phase HPLC. <sup>15</sup>N and <sup>13</sup>C labeled proteins were grown on minimum media with <sup>15</sup>NH<sub>4</sub>Cl and [<sup>13</sup>C]glucose as the sole sources for nitrogen and carbon.

### NMR characterization of intermediates

<sup>1</sup>H-<sup>15</sup>N NMR spectra for both WT\* and the 5G-mutant were collected at 25 °C and pH 5.0 on a Bruker (Bellerica, MA) DRX 500 MHz spectrometer equipped with a 5 mm *x*, *y*, *z*-shielded pulse field gradient triple resonance probe. The concentrations of the protein samples are  $\approx$ 1 mM. 3D CBCACONH and HNCACAB spectra were collected for backbone resonance assignment. The NMR spectra were processed using NMRPipe<sup>48</sup> and analyzed using NMRView<sup>49</sup> and Sparky $\ddagger$ .

### Kinetic folding and unfolding experiments

Kinetic folding and unfolding experiments were done by using a stopped-flow apparatus (Biologic, France) with a dead time of  $\approx$ 1 ms. In the kinetic folding experiments, proteins in 8 M urea or 6 M GdmCl were diluted quickly to initiate folding at different final concentrations of urea or GdmCl. The folding traces were fitted with two exponentials. Unfolding was initiated by mixing the folded proteins in the buffer solution with high concentrations of denaturant. A single exponential kinetics was observed for unfolding. The following equations were used to fit the measured folding rate constants of the fast phase and unfolding rate constants as a function of denaturant concentrations:

$$U \rightleftharpoons I \rightleftharpoons N \quad (8)$$

$$k_{\text{obs}} = (1/(1 + K_{IU}))k_{\text{IN}} + k_{\text{NI}} \quad (9)$$

$$k_{\text{obs}}((\text{den})) = (1/(1 + K_{IU}^{\text{H}_2\text{O}} \exp(m_{\text{IU}}^{\ddagger}(\text{Den})))) \times k_{\text{IN}}^{\text{H}_2\text{O}} \exp(m_{\text{IN}}^{\ddagger}(\text{Den})) + k_{\text{NI}}^{\text{H}_2\text{O}} \exp(m_{\text{NI}}^{\ddagger}[\text{den}]) \quad (10)$$

$\ddagger$  <http://www.cgl.ucsf.edu/home/sparky>

where  $K_{IU}^{H_2O}$  is the unfolding equilibrium constant of the putative intermediate in water,  $k_{IN}^{H_2O}$  and  $k_{NI}^{H_2O}$  are the values of the corresponding microscopic rate constants in the absence of denaturant, and  $m_{IN}^\ddagger$  and  $m_{NI}^\ddagger$  are the coefficients for the denaturant dependence of the microscopic rate constants. Den is either urea or GdmCl.

### Hydrogen exchange experiments

Hydrogen exchange experiments were initiated by loading 0.5 ml protein sample in  $^2H_2O$  at  $pD_{read}$  9.6 (100 mM glycine and 50 mM  $Na_3PO_4$ ) to the spin column (Sephadex 25, bed volume 3 ml) that was pre-equilibrated with  $H_2O$  at pH 9.6 (100 mM glycine and 50 mM  $Na_3PO_4$ ). The column was spun for 1 min at 4000 rpm using a bench top centrifuge. The eluted protein sample was immediately transferred into an NMR tube for collecting a series of  $^1H$ - $^{15}N$  HSQC spectra. The concentration of the protein sample was  $\approx 1$  mM. The NMR spectra were processed using NMRPipe<sup>47</sup> and analyzed using NMRView<sup>48</sup> to obtain the exchange rate constants.

### Hydrogen exchange-folding competition experiments

The competition experiments were performed with the same stopped-flow apparatus. Folding of denatured lysozyme ( $\approx 5$  mg/ml) in 8.5 M deuterated urea ( $pD$  6.8, 10 mM Mes/ $^2H_2O$ ) solution was initiated by tenfold dilution with a buffer solution (pH 6.8, 10 mM Mes/ $H_2O$ ). Reactions were allowed to proceed for several minutes before pH was adjusted to 3.1. Sample solutions were concentrated, and buffer was exchanged with a 10 %  $^2H_2O$  and 25 mM Gly-HCl (pH 3.0) solution with Amicon cells. The HSQC spectra were measured at 20 °C. Additional experiments were performed for T4-IM. In this case, the denatured protein in 6.0 M deuterated GdmCl ( $pD$  7.1, 10 mM Hepes/ $^2H_2O$ ) was diluted into a refolding buffer (pH 7.1, 10 mM Hepes/ $H_2O$ ) by tenfold. After the reaction, the buffer was exchanged with 10 %  $^2H_2O$  and 50 mM NaAc (pH 5.0). Peak intensities were used to measure the proton occupancies with side-chain amides of three Trp residues (W126, W138 and W158) averaged as internal reference. The following relationship holds when the competition proceeded until native proteins were completely formed:<sup>19</sup>

$$P_{occ}(H) = 0.8k_{ex}/(k_{ex} + k_f) + 0.2 \quad (11)$$

where  $P_{occ}(H)$  is proton occupancy and  $k_{ex}$  and  $k_f$  are exchange and folding rates, respectively. The factors 0.8 and 0.2 represent the fraction of the *trans* and *cis* proline isomers responsible for fast and slow folding phases, respectively. Once one measures  $P_{occ}(H)$  and  $k_f$ ,  $k_{ex}$  can be calculated by equation (9), and protection factors were calculated by  $k_{int}/k_{ex}$ .

### Dead-time pulse-labeling experiments

The pulse labeling experiments were performed with Bio-Logic QFM-4 instrument. Experiments were performed at 25 °C.  $^{15}N$ -labeled WT\* was initially dissolved in 8.5 M deuterated urea solution and left overnight at room temperature to complete the amide hydrogen exchange process. The initial protein concentration was  $\approx 5.0$  mg/ml. Refolding was initiated by a tenfold dilution with high pH buffer (0.5 M Caps, pH 10.2). After 13 ms, amide hydrogen exchange was quenched by diluting the sample 1:2.5 into the quench buffer (25% acetic acid, pH 1.5) so that the final pH was 3.1. The sample was concentrated using an Amicon

cell with 10,000 molecular weight cut-off and adjusted to pH 3.0 with buffer solution (10%  $^2H_2O$ , 25 mM Gly-HCl, pH 3.0). The  $^1H$ - $^{15}N$  HSQC spectra were collected at 20 °C. The reference sample for the quantitative measurement of proton occupancy was prepared by heating 10 mg/ml protein solution in 10 %  $^2H_2O$  for 1 h.

### Acknowledgements

We thank Dr Frederic Dahlquist for comments on the article. H.K. was supported by the Japan Society for Promotion of Science (JSPS) Research Fellowships for Young Scientists. This research is supported by the intramural research program of the NIH, National Cancer Institute, Center for Cancer Research.

### References

1. Baldwin, R. L. & Roder, H. (1991). Characterizing protein folding intermediates. *Curr. Biol.* **1**, 218–220.
2. Englander, S. W. & Mayne, L. (1992). Protein folding studied using hydrogen-exchange labeling and two-dimensional NMR. *Annu. Rev. Biophys. Biomol. Struct.* **21**, 243–265.
3. Englander, S. W. (2000). Protein folding intermediates and pathways studied by hydrogen exchange. *Annu. Rev. Biophys. Biomol. Struct.* **29**, 213–238.
4. Krishna, M. M., Hoang, L., Lin, Y. & Englander, S. W. (2004). Hydrogen exchange methods to study protein folding. *Methods*, **34**, 51–64.
5. Roder, H., Elove, G. A. & Englander, S. W. (1988). Structural characterization of folding intermediates in cytochrome c by H-exchange labelling and proton NMR. *Nature*, **335**, 700–704.
6. Udgaonkar, J. B. & Baldwin, R. L. (1988). NMR evidence for an early framework intermediate on the folding pathway of ribonuclease A. *Nature*, **335**, 694–699.
7. Miranker, A., Robinson, C. V., Radford, S. E., Aplin, R. T. & Dobson, C. M. (1993). Detection of transient protein folding populations by mass spectrometry. *Science*, **262**, 896–900.
8. Jennings, P. A. & Wright, P. E. (1993). Formation of a molten globule intermediate early in the kinetic folding pathway of apomyoglobin. *Science*, **262**, 892–896.
9. Raschke, T. M. & Marqusee, S. (1997). The kinetic folding intermediate of ribonuclease H resembles the acid molten globule and partially unfolded molecules detected under native conditions. *Nature Struct. Biol.* **4**, 298–304.
10. Bai, Y., Sosnick, T. R., Mayne, L. & Englander, S. W. (1995). Protein folding intermediates: native-state hydrogen exchange. *Science*, **269**, 192–197.
11. Chamberlain, A. K., Handel, T. M. & Marqusee, S. (1996). Detection of rare partially folded molecules in equilibrium with the native conformation of RNaseH. *Nature Struct. Biol.* **3**, 782–787.
12. Llinas, M., Gillespie, B., Dahlquist, F. W. & Marqusee, S. (1999). The energetics of T4 lysozyme reveal a hierarchy of conformations. *Nature Struct. Biol.* **6**, 1072–1078.
13. Yan, S., Kennedy, S. D. & Koide, S. (2002). Thermodynamic and kinetic exploration of the energy landscape of *Borrelia burgdorferi* OspA by native-state hydrogen exchange. *J. Mol. Biol.* **323**, 363–375.

14. Silverman, J. A. & Harbury, P. B. (2002). The equilibrium unfolding pathway of a beta/alpha8 barrel. *J. Mol. Biol.* **324**, 1031–1040.
15. Lu, J. & Dahlquist, F. W. (1992). Detection and characterization of an early folding intermediate of T4 lysozyme using pulsed hydrogen exchange and two-dimensional NMR. *Biochemistry*, **31**, 4749–4756.
16. Bai, Y. (1999). Kinetic evidence for an on-pathway intermediate in the folding of cytochrome c. *Proc. Natl Acad. Sci. USA*, **96**, 477–480.
17. Heidary, D. K., O'Neill, J. C., Jr, Roy, M. & Jennings, P. A. (2000). An essential intermediate in the folding of dihydrofolate reductase. *Proc. Natl Acad. Sci. USA*, **97**, 5866–5870.
18. Capaldi, A. P., Shastry, M. C., Kleanthous, C., Roder, H. & Radford, S. E. (2001). Ultrarapid mixing experiments reveal that Im7 folds via an on-pathway intermediate. *Nature Struct. Biol.* **8**, 68–72.
19. Takei, J., Pei, W., Vu, D. & Bai, Y. (2002). Populating partially unfolded forms by hydrogen exchange-directed protein engineering. *Biochemistry*, **41**, 12308–12312.
20. Krantz, B. A., Mayne, L., Rumbley, J., Englander, S. W. & Sosnick, T. R. (2002). Fast and slow intermediate accumulation and the initial barrier mechanism in protein folding. *J. Mol. Biol.* **324**, 359–371.
21. Chu, R. A., Takei, J., Barchi, J. J., Jr & Bai, Y. (1999). Relationship between the native-state hydrogen exchange and the folding pathways of barnase. *Biochemistry*, **38**, 14119–14124.
22. Silow, M. & Oliveberg, M. (1997). Transient aggregates in protein folding are easily mistaken for folding intermediates. *Proc. Natl Acad. Sci. USA*, **94**, 6084–6086.
23. Nawrocki, J. P., Chu, R. A., Pannell, L. K. & Bai, Y. (1999). Intermolecular aggregations are responsible for the slow kinetics observed in the folding of cytochrome c at neutral pH. *J. Mol. Biol.* **293**, 991–995.
24. Gassner, N. C., Baase, W. A., Lindstrom, J. D., Lu, J., Dahlquist, F. W. & Matthews, B. W. (1999). Methionine and alanine substitutions show that the formation of wild-type-like structure in the carboxy-terminal domain of T4 lysozyme is a rate-limiting step in folding. *Biochemistry*, **38**, 14451–14460.
25. Roder, H. & Wuthrich, K. (1986). Protein folding kinetics by combined use of rapid mixing techniques and NMR observation of individual amide protons. *Proteins: Struct. Funct. Genet.* **1**, 34–42.
26. Takei, J., Chu, R. A. & Bai, Y. (2000). Absence of stable intermediates on the folding pathway of barnase. *Proc. Natl Acad. Sci. USA*, **97**, 10796–10801.
27. Bai, Y., Milne, J. S., Mayne, L. & Englander, S. W. (1993). Primary structure effects on peptide group hydrogen exchange. *Proteins: Struct. Funct. Genet.* **17**, 75–86.
28. Silow, M. & Oliveberg, M. (1997). High-energy channeling in protein folding. *Biochemistry*, **36**, 7633–7637.
29. Chan, H. S., Shimizu, S. & Kaya, H. (2004). Cooperativity principles in protein folding. *Methods Enzymol.* **380**, 350–379.
30. Bai, Y. (2006). Energy barriers, cooperativity, and hidden intermediates in the folding of small proteins. *Biochem. Biophys. Res. Commun.* **340**, 976–983.
31. Hvidt, A. & Nielsen, S. O. (1966). Hydrogen exchange in proteins. *Advan. Protein Chem.* **21**, 287–386.
32. Zhou, Z., Huang, Y. & Bai, Y. (2005). An on-pathway hidden intermediate and the early rate-limiting transition state of Rd-apocytochrome b562 characterized by protein engineering. *J. Mol. Biol.* **352**, 757–764.
33. Kim, P. S. & Baldwin, R. L. (1982). Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. *Annu. Rev. Biochem.* **51**, 459–489.
34. Matthews, C. R. (1993). Pathways of protein folding. *Annu. Rev. Biochem.* **62**, 653–683.
35. Kiefhaber, T. & Baldwin, R. L. (1995). Kinetics of hydrogen bond breakage in the process of unfolding of ribonuclease A measured by pulsed hydrogen exchange. *Proc. Natl Acad. Sci. USA*, **92**, 2657–2661.
36. Hoeltzli, S. D. & Frieden, C. (1995). Stopped-flow NMR spectroscopy: real-time unfolding studies of 6-19F-tryptophan-labeled *Escherichia coli* dihydrofolate reductase. *Proc. Natl Acad. Sci. USA*, **92**, 9318–9322.
37. Kiefhaber, T., Bachmann, A., Wildegger, G. & Wagner, C. (1997). Direct measurement of nucleation and growth rates in lysozyme folding. *Biochemistry*, **36**, 5108–5112.
38. Roy, M. & Jennings, P. A. (2003). Real-time NMR kinetic studies provide global and residue-specific information on the non-cooperative unfolding of the beta-trefoil protein, interleukin-1beta. *J. Mol. Biol.* **328**, 693–703.
39. Sosnick, T. R., Mayne, L. & Englander, S. W. (1996). Molecular collapse: the rate-limiting step in two-state cytochrome c folding. *Proteins: Struct. Funct. Genet.* **24**, 413–426.
40. Bai, Y. & Englander, S. W. (1996). Future directions in folding: the multi-state nature of protein structure. *Proteins: Struct. Funct. Genet.* **24**, 145–151.
41. Chu, R., Pei, W., Takei, J. & Bai, Y. (2002). Relationship between the native-state hydrogen exchange and folding pathways of a four-helix bundle protein. *Biochemistry*, **41**, 7998–8003.
42. Feng, H., Vu, N. D. & Bai, Y. (2005). Detection of a hidden folding intermediate of the third domain of PDZ. *J. Mol. Biol.* **346**, 345–353.
43. Fersht, A. R. (2000). A kinetically significant intermediate in the folding of barnase. *Proc. Natl Acad. Sci. USA*, **97**, 14121–14126.
44. Vu, N. D., Feng, H. & Bai, Y. (2004). The folding pathway of barnase: the rate-limiting transition state and a hidden intermediate under native conditions. *Biochemistry*, **43**, 3346–3356.
45. Llinas, M. & Marqusee, S. (1998). Subdomain interactions as a determinant in the folding and stability of T4 lysozyme. *Protein Sci.* **7**, 96–104.
46. Levinthal, C. (1968). Are there pathways for protein folding? *J. Chim. Phys.* **65**, 44–45.
47. Wetlaufer, D. B. (1973). Nucleation, rapid folding, and globular intrachain regions in proteins. *Proc. Natl Acad. Sci. USA*, **70**, 697–701.
48. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR*, **6**, 277–293.
49. Johnson, B. A. & Blevins, R. A. (1994). NMRView: a computer program for the visualization and analysis of NMR data. *J. Biomol. NMR*, **4**, 603–614.

Edited by K. Kuwajima

(Received 1 June 2006; received in revised form 10 October 2006; accepted 16 October 2006)  
Available online 21 October 2006