Experimental Evidence for Millisecond–Timescale Structural Evolution Following the Microsecond–Timescale Folding of a Small Protein

C. Blake Wilson, Wai-Ming Yau, and Robert Tycko

Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520, USA

(Received 26 April 2023; accepted 27 October 2023; published 24 January 2024)

Prior work has shown that small proteins can fold (i.e., convert from unstructured to structured states) within 10 μs. Here we use time-resolved solid state nuclear magnetic resonance (ssNMR) experiments to show that full folding of the 35-residue villin headpiece subdomain (HP35) requires a slow annealing process that has not been previously detected. $^{13}$C ssNMR spectra of frozen HP35 solutions, acquired with a variable time $\tau_e$ at 30 °C after rapid cooling from 95 °C and before rapid freezing, show changes on the 3–10 ms timescale, attributable to slow rearrangements of protein sidechains during $\tau_e$.

DOI: 10.1103/PhysRevLett.132.048402

Research into mechanisms by which protein chains fold to their native three-dimensional structures is motivated by the close association between biological function and structure, by the involvement of protein misfolding in human diseases [1], and by the early realization that protein folding without a nonrandom mechanism would require exorbitant time periods [2]. An important finding of recent research is that protein folding can occur on the microsecond timescale, especially for certain small, single-domain proteins [3,4]. Observations of fast folding depend primarily on optical spectroscopies that probe structural changes around a fluorophore or local changes in protein backbone conformation following a rapid experimental perturbation. Since optical measurements may not be sensitive to all characteristics of a protein structure, especially the extent to which sidechain groups are conformationally ordered and optimally packed, it is conceivable that protein structures may not be fully formed on the timescales indicated by time-resolved optical signals.

Here we report the results of time-resolved solid state nuclear magnetic resonance (ssNMR) experiments [5–7] in which folding of the 35-residue villin headpiece subdomain (HP35) is triggered by a rapid drop in temperature ($T$), i.e., a rapid inverse $T$ jump. HP35 is one of the fastest folding proteins known [8–11], forming a three-helix structure at physiological temperatures with $\alpha$-helical segments H1 (residues 44–50), H2 (residues 55–58), and H3 (residues 63–72) (Fig. S1 [12]). As shown by circular dichroism (CD) spectroscopy, HP35 loses its helical structure at elevated temperatures, with an unfolding midpoint temperature $T_{\text{mid}} \approx 68 ^\circ C$ [Figs. 1(a) and 1(b)]. Kinetics of HP35 folding have been characterized by measurements of time-dependent fluorescence intensities following laser-induced positive $T$ jumps [11,21]. These measurements indicate a folding time of approximately 4 μs (folding rate $k_f \approx 2-3 \times 10^5 \text{ s}^{-1}$) for temperatures in the 35–83 °C range. Rapid folding is also supported by analyses of $^1$H NMR line-shapes in solutions of a 36-residue version (HP36) that yield $k_f \approx 0.5-2 \times 10^5 \text{ s}^{-1}$ over the 56–78 °C range [22], by time-resolved infrared absorbance measurements on HP35 and HP36 that yield $k_f \approx 1-3 \times 10^5 \text{ s}^{-1}$ over the 45–85 °C range [23–26], and by all-atom molecular dynamics (MD) simulations in which HP35 fluctuates between unfolded and folded structural states, with average unfolded state lifetimes of about 15–20 μs at simulation temperatures of 72–87 °C [27]. Other MD simulations also indicate timescales less than 20 μs [28,29].

Our experiments use an inverse $T$-jump apparatus described previously [5] and in the Supplemental Material [12] (see also Refs. [13–20] therein). Briefly, a solution containing $^{13}$C-labeled HP35 at 2.5 mM concentration in aqueous buffer with 20% v/v glycerol (initially at room temperature) is heated to 95 °C within 25 ms by flowing through a 2.5–6.0 cm section of 100 μm inner diameter (i.d.) copper capillary soldered to a heated copper plate. The solution then flows through a 9 mm section of 30 μm i.d. copper capillary soldered to a water-cooled copper plate, dropping its temperature to 30 °C. With a 0.63 ml/min flow rate, the average time $\tau_{\text{hot}}$ in the heated section is 25–40 ms, depending on the length of this section, which is much greater than the unfolding time of HP35 at 95 °C [11]. The average time $\tau_{\text{cold}}$ in the cooled section is 0.6 ms. The HP35 solution leaves the cold capillary as a jet, traveling a variable flight distance through air at 15 m/s, then striking the surface of a stirred isopentane bath at −145 °C. In the isopentane, the solution breaks into particles with 30–40 μm diameters that cool to temperatures below −35 °C, where the solution viscosity exceeds 30 cP [30], within approximately 150 μs [5,31] (see methods and Fig. S4 in Supplemental Material [12]). The slurry of frozen particles is packed into magic-angle
spinning (MAS) rotors for low-temperature ssNMR measurements (Fig. S3 [12]), with signal enhancements from dynamic nuclear polarization (DNP) [32].

Figure 1(c) shows double-quantum-filtered [33] one-dimensional (1D) $^{13}$C ssNMR spectra of frozen HP35 solutions with indicated values of the variable structural evolution time $\tau_e$, defined to be the sum of $\tau_{\text{cold}}$ and the variable flight time from the end of the cold capillary to the isopentane surface. Spectra with $\tau_e = 0$ and $\tau_e = \infty$ correspond to samples that were frozen directly from 95°C and 30°C, respectively, without an inverse $T$ jump. HP35 was $^{13}$C-labeled at all carbon sites of V50 (backbone CO and $C_a$, sidechain $C_\beta$ and $C_\gamma$) and at the CO and $C_a$ sites of G52 and L69 (Fig. S1a [12]). Spectra in Fig. 1(c) were recorded at 115 K, where methyl group rotation results in sharper V50 $C_\gamma$ signals than at lower temperatures (Fig. S5a [12]).

Center frequencies of $^{13}$C$_a$ and $^{13}$CO ssNMR lines in Fig. 1(c) (i.e., $^{13}$C chemical shifts, see Table I) are nearly independent of $\tau_e$ and are consistent with $\alpha$-helical structure at V50 and L69 (in H1 and H3, respectively) and the presence of G52 in a structured loop (between H1 and H2) in the folded state of HP35. However, signals from the two methyl ($C_\gamma$) sites of V50 depend on $\tau_e$, displaying a doublet structure with a 3.3 ppm splitting at $\tau_e = \infty$ that is absent or less pronounced at earlier values of $\tau_e$. To quantify the evolution of the $^{13}$C ssNMR lineshapes, we performed a principal component analysis of these spectra. As shown in Fig. 1(d), PC2 contains peaks above the noise that correspond to the downfield component of the V50 $^{13}$C$_\gamma$ signals and to a downfield shift in the V50 $C_\beta$ signal. The coefficient of PC2 is plotted as a function of $\tau_e$ in Fig. 1(e). An exponential fit yields a time constant $\tau_{PC2} = 5.7 \pm 1.1$ ms. These results suggest that conformational ordering of the V50 sidechain occurs relatively slowly at 30°C, on a timescale that is much longer than the timescale of the initial folding process. Additional measurements with solution temperatures of 37°C during $\tau_e$, rather than 30°C, show that the timescale for conformational ordering of the V50 sidechain is not strongly temperature-dependent (Fig. S5b [12]).

Two-dimensional (2D) $^{13}$C-$^{13}$C ssNMR spectra of the same samples were also recorded (Fig. 2 and Fig. S6 [12]), at measurement temperatures of 90 K for higher signal-tonoise. Although the 2D spectra do not show a strong dependence on $\tau_e$, subtle changes in the $^{13}$C$_a$/$^{13}$CO crosspeak lineshape for L69 are apparent [Fig. 2(a)]. To quantify these changes, we calculated difference spectra $S_\Delta(\tau_e) = S(\tau_e) - \lambda(\tau_e)S_F$ by subtracting the crosspeak signal $S_F$ of maximally folded HP35 [Fig. 2(b)] from the crosspeak signals $S(\tau_e)$. $S_F$ was obtained from an HP35 solution that contained 40% v/v glycerol and was frozen relatively slowly from room temperature by immersion in liquid nitrogen, producing a 5 K/s cooling rate. The scaling factor $\lambda(\tau_e)$ was adjusted to null $S_\Delta(\tau_e)$ at 55.2 ppm/176.6 ppm, which is the L69 $^{13}$C$_a$/$^{13}$CO crosspeak position at L69. An exponential fit to the ratio of crosspeak volumes in $S_\Delta(\tau_e)$ and $S(\tau_e)$ yields a time constant $\tau_\Delta = 2.6 \pm 0.4$ ms [Fig. 2(d)].

A second series of measurements was performed on samples with uniform $^{13}$C labeling at M53 (in the loop between H1 and H2), S56 and F58 (in H2), and G74 (in the disordered segment after H3) (Fig. S1b [12]). Figure 3(a)
TABLE I. $^{13}$C chemical shifts relative to tetramethylsilane (TMS) in ssNMR spectra of frozen HP35 solutions. Values in parentheses are random coil chemical shifts [34], adjusted to the TMS reference. $^{13}$CO and $^{13}$C$_{\alpha}$ chemical shifts greater than and $^{13}$C$_{\beta}$ chemical shifts less than random coil values indicate $\alpha$-helical structure [35–37].

<table>
<thead>
<tr>
<th>Residue</th>
<th>$^{13}$C$_{\alpha}$ (ppm)</th>
<th>$^{13}$CO (ppm)</th>
<th>$^{13}$C$_{\beta}$ (ppm)</th>
<th>$^{13}$C$_{\gamma}$ (ppm)</th>
<th>$^{13}$C$_{\delta}$ (ppm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>V50</td>
<td>62.6 ± 0.2 (60.5)</td>
<td>175.3 ± 0.2 (174.6)</td>
<td>29.8 ± 0.2 (31.2)</td>
<td>20.2 ± 0.2 (19.4)</td>
<td>...</td>
</tr>
<tr>
<td>G52</td>
<td>43.4 ± 0.2 (43.4)</td>
<td>170.1 ± 0.2 (173.2)</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>M53</td>
<td>50.7 ± 0.2 (53.7)</td>
<td>169.8 ± 0.3 (174.6)</td>
<td>31.7 ± 0.2 (31.2)</td>
<td>26.8 ± 0.2 (30.3)</td>
<td>14.0 ± 0.2 (15.2)</td>
</tr>
<tr>
<td>S56$^a$</td>
<td>59 ± 3 (56.6)</td>
<td>175 ± 3 (172.9)</td>
<td>61 ± 3 (62.1)</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>F58</td>
<td>58.6 ± 0.2 (56.0)</td>
<td>174.9 ± 0.4 (174.1)</td>
<td>37.4 ± 0.2 (37.9)</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>P62</td>
<td>59.7 ± 0.2 (61.7)</td>
<td>174.5 ± 0.3 (175.7)</td>
<td>30.7 ± 0.2 (30.4)</td>
<td>25.8 ± 0.2 (25.5)</td>
<td>47.0 ± 0.2 (48.1)</td>
</tr>
<tr>
<td>L69</td>
<td>55.2 ± 0.2 (53.4)</td>
<td>176.7 ± 0.2 (175.9)</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>G74</td>
<td>43.5 ± 1 (43.4)</td>
<td>172.2 ± 1 (173.2)</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

$^a$Poorly resolved in 2D spectra.

shows the 2D $^{13}$C-$^{13}$C spectrum of a sample with $\tau_e = 1.6$ ms. 2D spectra with other $\tau_e$ values are shown in Fig. S7 [12]. Reductions in crosspeak linewidths with increasing $\tau_e$ are observed at many sites, indicating an overall increase in conformational order on the millisecond timescale. Examples are shown in Figs. 3(b)–3(d), where 1D slices from the 2D spectra are plotted as a function of $\tau_e$. The slice at 128.1 ppm through signals from aromatic carbons of the F58 sidechain [Fig. 3(b)] shows a sharpening of signals between 135 and 140 ppm, consistent with increasing conformational order for the F58 sidechain. Signals in this slice at 174.9 and 58.6 ppm, arising from CO and C$_{\alpha}$ sites of F58, also become progressively narrower. The slice at 26.8 ppm through M53 $^{13}$C$_{\gamma}$ signals [Fig. 3(c)] shows a sharpening of $^{13}$C$_{\gamma}$/$^{13}$C$_{\beta}$ crosspeak signals between 30 and 35 ppm, consistent with increasing order in the M53 sidechain. The signal in this slice at 50.7 ppm, arising from the C$_{\alpha}$ site of M53, also becomes progressively narrower. The slice at 58.6 ppm [Fig. 3(d)] additionally shows that the combined $^{13}$C$_{\alpha}$/$^{13}$CO cross-peaks of S56 and F58 ($^{13}$CO signals around 174.7 ppm) become sharper with increasing $\tau_e$. Analyses of the dependences of linewidths on $\tau_e$ (Figs. S8 and S9 [12]) show that linewidths decrease by 10%–30% with time constants in the 3–10 ms range.

Overall, the time-resolved ssNMR data show that HP35 converts from a disordered state at 95°C to a state that is close to the fully folded state when the HP35 solution is frozen by injection into cold isopentane, vitrifying within approximately 150 ms. However, this state (defined as $\tau_e = 0$) contains residual disorder. The conformations and structural environments of amino acid sidechains become increasingly homogeneous as $\tau_e$ increases, leading to a progressive narrowing of ssNMR lines. We emphasize that the dependences on $\tau_e$ in Figs. 1–3 must reflect processes that occur at 30°C, in liquid solutions, since conditions for heating to 95°C, for rapid cooling from 95°C to 30°C, and for freezing after $\tau_e$ were kept constant. Similar results would be expected if the folding process was initiated by rapid dilution of denaturant, for example, rather than by a rapid inverse $T$ jump.
Our interpretation is depicted in Fig. 4. All-atom MD simulations by Piana et al. [27] show spontaneous unfolding and folding of HP35 multiple times over a 400 μs trajectory near \( T_{\text{mid}} \). Within time periods where HP35 is in a folded configuration (defined for Fig. 4 by backbone atom coordinates in the three helices being within a root-mean-squared distance of 1.5 Å from HP35 crystal structure coordinates [10]), amino acid sidechain conformations are dynamically disordered. In our inverse \( T \)-jump experiments, similar disorder is partially trapped as the solution is rapidly cooled from 95 °C to 30 °C (approximately 10^5 K/s), leading to a structural state that is not fully equilibrated. Optimization of sidechain conformations and packing requires a subsequent annealing process that is relatively slow at 30 °C, especially for sidechains (such as those of V50, M53, F58, and L69) that participate in the hydrophobic core of the folded structure.

In principle, cis-trans isomerization of the L61-P62 peptide bond could contribute to the annealing process indicated by our time-resolved ssNMR data [38]. To investigate this possibility, we prepared an HP35 sample in which only P62 was 13C-labeled and compared ssNMR spectra of solutions that were rapidly frozen after heating to 95 °C with \( \tau_{\text{hot}} = 25 \text{ ms} \) or \( \tau_{\text{hot}} = 20 \text{ min} \) (with \( \tau_e = 0 \)), rapidly frozen from 30 °C (with \( \tau_{\text{hot}} = 0 \)), and slowly frozen from 24 °C (Fig. S10 [12]). Population of the cis isomer would be expected to produce observable changes in 13C chemical shifts of CO, \( C_{\gamma} \), and/or \( C_{\alpha} \) sites of P62 [39]. No differences in the experimental spectra were observed, aside from linewidths reduced by ~0.6 ppm in spectra of the slowly frozen sample. We conclude that cis-trans isomerization does not contribute significantly to the dependences on \( \tau_e \) discussed above.

Lattice-model simulations of protein folding by Kussell et al. [40,41] showed a separation between the timescale for the protein backbone to adopt its native (i.e., folded) conformation and the timescale for sidechains to adopt their final, free energy-minimizing configuration. Kussell et al. attribute this behavior to trapping of sidechains (outside the folding nucleus) in partially ordered states after the main folding event, with energy barriers that...
inhibit sidechain transitions. In their simulations, optimization of sidechain configurations requires fluctuations of the backbone conformation that reduce these energy barriers, leading to slow equilibration [19,20]. Our experimental results for HP35 verify that the timescales for backbone folding and optimization of sidechain configurations can differ by factors greater than 100. Our observation of time-scales for changes in ssNMR signals in the 3–10 ms range may also be consistent with the residue-dependent variations in relaxation times reported by Kussell et al. [19,20].

Kubelka et al. estimated unfolding rates \( k_u \) for HP35 below \( T_{\text{mid}} \) from a two-state analysis, according to which \( k_u = f_u (k_a + k_f) \), where \( f_u \) is the fractional population of unfolded molecules, estimated from equilibrium CD and fluorescence measurements, and \( k_u + k_f \) is the equilibration rate determined from \( T \)-jump fluorescence data [11]. At 30°C, their estimated values extrapolate to \( k_u \approx 2 \times 10^3 \text{ s}^{-1} \). This raises an important point: How is it possible to observe slow structural annealing of HP35 molecules in their folded states if the folded states persist for only 0.5 ms on average, after which individual molecules unfold and then rapidly refold? Why do repeated unfolding/refolding events not disrupt the slow structural annealing process, resetting the structural ensemble to something that resembles the intermediate state depicted in Fig. 4(a)? A possible explanation is that \( f_u \) is substantially smaller at 30°C than estimated from the CD and fluorescence measurements, and \( k_u \) is actually less than 200 \( \text{ s}^{-1} \). Alternatively, unfolding rates may depend on the degree of structural order within folded molecules, such that HP35 molecules with sidechain conformations closer to their optimal states unfold less frequently. It is also possible that unfolded states at 30°C are substantially less conformationally disordered than unfolded states near \( T_{\text{mid}} \), so that unfolding or refolding events do not disrupt the annealing process.

In conclusion, we have used time-resolved ssNMR to identify a structural annealing process that occurs on the timescale of 3–10 ms after the microsecond–timescale initial folding process of the model protein HP35. This annealing process involves increased ordering of the conformations and environments of amino acid sidechains, as it is the ssNMR signals of sidechains that exhibit the largest changes with increasing evolution time following a rapid inverse \( T \)-jump from 95°C to 30°C. Although to our knowledge this is the first experimental evidence for such a slow annealing process in a rapidly folding protein, we expect that similar processes will be found in other biopolymer systems if similar experimental methods are applied.

This work was supported by the Intramural Research Program of the National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, under project No. 1-ZIA-DK029029-27. We thank Bernard Howder for his assistance with fabrication of equipment used in this work.

[38] W. J. Wedemeyer, E. Welker, and H. A. Scheraga, Biochemistry 41, 14637 (2002).