Biophysical Journal
Article

Biophysical Society

Replacement of the native *cis* prolines by alanine leads to simplification of the complex folding mechanism of a small globular protein

Anushka Kaushik[1] and Jayant B. Udgaonkar[1,2,*]
[1]Indian Institute of Science Education and Research, Pune, India and [2]National Centre for Biological Sciences, Tata Institute of Fundamental Research, Bengaluru, India

ABSTRACT The folding mechanism of MNEI, a single-chain variant of naturally occurring double-chain monellin, is complex, with multiple parallel refolding channels. To determine whether its folding energy landscape could be simplified, the two native *cis*-prolines, Pro41 and Pro93, were mutated, singly and together, to Ala. The stability of P93A was the same as that of the wild-type protein, pWT; however, P41A and P41AP93A were destabilized by ~0.9 kcal mol^{-1}. The effects of the mutations on the very fast, fast, slow, and very slow phases of folding were studied. They showed that heterogeneity in the unfolded state arises due to *cis* to *trans* isomerization of the Gly92-Pro93 peptide bond. The Pro41 to Ala mutation abolished the very slow phase of folding, whereas surprisingly, the Pro93 to Ala mutation abolished the very fast phase of folding. Double-jump, interrupted folding experiments indicated that two sequential *trans* to *cis* proline isomerization steps, of the Gly92-Pro93 peptide bond followed by the Arg40-Pro41 peptide bond, lead to the formation of the native state. They also revealed the accumulation of a late native-like intermediate, N*, which differs from the native state in the isomeric status of the Arg40-Pro41 bond, as well as in a few tertiary contacts as monitored by near-UV CD measurements. The Pro to Ala mutations not only eliminated the *cis* to *trans* Pro isomerization reaction in the unfolded state, but also the two *trans* to *cis* Pro isomerization reactions during folding. By doing so, and by differentially affecting the relative stabilities of folding intermediates, the mutations resulted in a simplification of the folding mechanism. The two Pro to Ala mutations together accelerate folding to such an extent that the native state forms more than 1000-fold faster than in the case of pWT.

SIGNIFICANCE Proteins often fold via complex mechanisms that feature multiple folding pathways, and it is important to identify the origin of such heterogeneity in protein folding reactions. In proteins containing *cis* Pro residues, folding complexity is invariably attributed to proline isomerization, both during unfolding and during subsequent folding. In this study, it is shown that, even upon mutating both of the two *cis* Pro residues of monellin to Ala, two out of three folding pathways still remain operational, suggesting an alternative source of complexity. The mutations do, however, simplify the overall folding mechanism: the very slow phase of folding is eliminated and so is the slow phase of folding on one of the competing pathways that still operates, as a result of destabilization of a folding intermediate. As a consequence, monellin not having any *cis* Pro residue folds more than 1000-fold faster than the WT protein. The mutations also reveal how structure formation can be coupled to proline isomerization during folding. Importantly and surprisingly, the mutations also abolish the very fast phase of folding by their effect on the unfolded state. This study therefore demonstrates a new way in which substitution of Pro residues can reduce folding complexity by eliminating not only the slow phases but also the fast phases of folding.

INTRODUCTION

The mechanism of protein folding from an unstructured, random-coil unfolded (U) state to a uniquely structured native (N) state is still not fully understood. In particular, whereas various experimental, theoretical, and computational studies have highlighted the existence of parallel folding channels (1–3), little is understood about the sequence determinants responsible for multiple folding pathways. Multiple folding pathways can arise when the subensembles within unfolded or intermediate ensembles interconvert on timescales much slower than the subsequent

Submitted June 5, 2023, and accepted for publication August 15, 2023.
*Correspondence: jayant@iiserpune.ac.in
Editor: Sudipta Maiti.
https://doi.org/10.1016/j.bpj.2023.08.012

folding steps (2,4). Intermediate ensembles populated at different stages during the folding reaction may be composed of structurally distinct subensembles, resulting in multiple routes to the N state (5–7). Determination of the cause of heterogeneity at the early and intermediate stages of a folding reaction would lead to a better understanding of the interactions dictating the sequence of structure-forming events during folding. A major cause of the heterogeneity in the U-state ensemble for many proteins, which results in parallel channels of folding, is the isomerization of peptidyl-prolyl bonds (8–14). The *trans* isomer is much more stable than the *cis* isomer in the case of a non-prolyl bond, but it is only slightly more stable for a prolyl bond. Hence, both the isomeric states for the peptidyl-prolyl bonds may be populated significantly in the U-state ensemble (15–17). Of course, there can be causes of heterogeneity in the U-state ensemble other than proline isomerization (18–20).

Of the 20 naturally occurring amino acid residues found in proteins, proline is unique in that its side chain is covalently linked to the backbone nitrogen, thereby forming a cyclic ring. Thus, up to 25% of X-Pro bonds are found in the *cis* conformation in the N states of different proteins (21). They typically occur in tight turns and kinks essential for maintaining the structural integrity of the protein (21). In the U state, however, an X-Pro bond that is *cis* in the N state, can undergo *cis* to *trans* isomerization, resulting in significant population of the *trans* conformation. The important step of *trans* to *cis* prolyl isomerization then becomes inevitable along the folding pathway, as shown for many proteins (8,13,15,17,22,23). Such prolyl isomerization steps (either *cis* to *trans* or *trans* to *cis*) are usually much slower than the structure-forming steps, to which they are coupled (24–26). Thus, the occurrence of *cis* prolyl bonds in the N state, though structurally significant, can result in the retardation of the folding reaction. However, the interpretation of studies of the folding of proteins in which the rate-limiting Pro residues were replaced by other residues has not been straightforward (27–33). Nonetheless, it is worthwhile to explore the potential simplification of the folding mechanism upon replacing Pro residues and to thereby provide additional detail on the influence of the amino acid sequence on the mechanism of protein folding.

Single-chain monellin (MNEI) is a variant of naturally occurring heterodimeric monellin, in which the C-terminus of chain B is linked to the N-terminus of chain A via a Gly-Phe dipeptide linker (34). It is a small globular protein (97 amino acid residues) that consists of an α-helix packed against a five-stranded β-sheet (Fig. 1 *A*). Detailed characterization of the folding and unfolding mechanisms of MNEI using optical probes (34–37) and mass spectrometry (38) has been achieved in previous studies. Relevant to the current study, a previous fluorescence-based study (34) identified five distinct kinetic phases of refolding of MNEI: an ultrafast (>1000 s^{-1}) phase monitored using ANS fluorescence, very fast (~10 s^{-1}), fast (~1 s^{-1}), and slow (~0.1 s^{-1}) phases monitored using Trp4 fluorescence, and a very slow (<0.01 s^{-1}) phase monitored by double-jump, interrupted refolding experiments. The U-state ensemble was shown to consist of two subensembles, with U_1 folding in the very fast phase and U_2 folding in the competing fast and slow pathways, and a folding scheme (34) was proposed. A very recent trFRET study, which monitored the compaction of different structural segments of MNEI (7), also independently arrived at a similar folding scheme as shown in Fig. 1 *B*.

The intermediate, I_{F1}, populating on the fast phase arising from an early intermediate, I_{E1}, was shown to be heterogeneous, consisting of two subensembles that form in parallel (7). It was shown that the N-like intermediate, N*, formed directly from the U_1 subensemble, but via intermediates along two different parallel pathways from the U_2 subensemble. It was suggested that heterogeneity in the U-state ensemble arises due to different isomeric states of Pro residue(s). The molecules folding from the U_1 subensemble differ in the sequence of structure-forming steps from the molecules folding from the U_2 subensemble (7). A possible reason for this could be that different isomeric states of Pro residues in these subensembles provide different structural biases for subsequent folding events.

It is important to identify the Pro residue(s) responsible for the kinetic heterogeneity in the U-state ensemble, which gives rise to parallel pathways of folding.

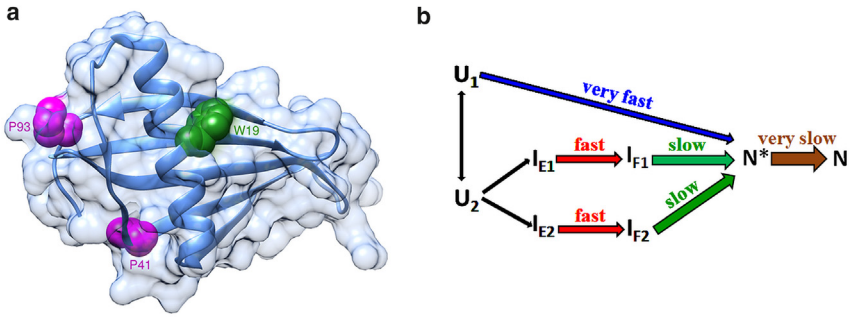

FIGURE 1 Structure and complex folding mechanism of the single-chain monellin (MNEI). (*a*) Structure of MNEI showing the locations of the two *cis* Pro residues mutated to Ala. The side chains of Pro41 and Pro93 are coloured in magenta, and are buried to extents of 97% and 40%, respectively. The side chain of Trp19 is coloured in green. The structure is drawn using Chimera (PDB ID: 1IV7). (*b*) The complex folding mechanism of MNEI with multiple parallel refolding channels (7,34). To see this figure in color, go online.

Peptidyl-prolyl bonds that are *cis* in the N state are more likely to cause heterogeneity in the U state than those present in the *trans* conformation(s) (15,17). Of the six Pro residues present in MNEI, Pro41 and Pro93 are in the *cis* conformation in the N state. Pro41 is present in the middle of the chain and participates in the formation of the β-sheet, whereas Pro93 is present in the C-terminus region (Fig. 1 *A*). It was shown in a previous study that the very slow rate of formation of the N state (Fig. 1 *B*) from the late intermediates is independent of the GdnHCl concentration in which the protein was refolded, as is known to be the case for proline isomerization reactions (34). Moreover, the rate constant of the formation of the N state was observed to lie in the range of 0.01 to 0.001 s^{-1}, which is that expected for prolyl bond isomerization (39). It is therefore crucial to demonstrate the role of the native *cis* prolines at the later stages of folding of MNEI and to segregate the slow proline isomerization events from the structure-forming folding steps.

In the current study, the potential roles of Pro41 and Pro93 in retarding the overall folding kinetics of pWT and also in giving rise to multiple pathways have been explored. Pro41 and Pro93 were replaced with Ala using mutagenesis, and the folding kinetics of the mutant variants P41A, P93A, and P41AP93A were studied with the aim of determining whether and how elimination of the Pro isomerization steps could result in a simplification of the folding mechanism of MNEI.

MATERIALS AND METHODS

Buffers and reagents

The native buffer used in all the experiments contained 20 mM Tris, 250 μM EDTA at pH 8.0. The unfolding buffer contained GdnHCl (ultrapure grade, 99.5+ % from Alfa Aesar) in native buffer. 1 mM DTT was added to both the buffers. The concentrations of GdnHCl solutions were determined from the measurement of refractive indices on an Abbe 3L refractometer from Milton Roy. All the buffers and solutions were filtered using 0.22-μm filters and degassed before use. All the experiments were carried out at room temperature (25°C).

Site-directed mutagenesis

The pseudo-wild-type, pWT, (W4YF19W) mutant variant of MNEI was used in this study (6). Mutant variants of pWT were generated using the Quickchange site-directed mutagenesis method developed by Stratagene. Three Pro mutant variants were prepared: P41A, P93A, and P41AP93A. The mutations were verified by DNA sequencing.

Protein expression and purification

pWT and the Pro mutant variants were expressed and purified as reported previously (34). The purity of each protein was confirmed by electrospray ionization mass spectrometry (Waters corporation). Each protein had its expected mass. Protein concentrations were determined by the measurement of absorbance at 280 nm and using an extinction coefficient of 18,300 M^{-1} cm^{-1} (6).

Fluorescence spectra

Fluorescence spectra were collected on a Fluoromax 4 (Horiba) spectrofluorimeter. Fluorescence was excited at 295 nm, and the emission spectra were collected from 305 to 400 nm, with a response time of 1 s, and excitation and emission bandwidths of 1 nm and 5 nm, respectively. The protein concentration was 10 μM. Each spectrum was recorded as the average of three fluorescence emission wavelength scans.

Far-UV CD and near-UV CD spectra

Measurements of CD spectra were carried out on a Jasco J-815 spectropolarimeter. Far-UV CD spectra were collected using a 0.2-mm pathlength cuvette, with a bandwidth of 1 nm, a response time of 1 s, and a scan speed of 20 nm min^{-1}. 30 wavelength scans were averaged for each spectrum.

Near-UV CD spectra were collected using a 10-cm pathlength cuvette, with a bandwidth of 1 nm, a response time of 1 s, and a scan speed of 20 nm min^{-1}. 50 wavelength scans were averaged for each spectrum. The protein concentration was ∼10 μM.

Near-UV CD monitored kinetic folding experiments

Folding kinetics was monitored at 270 nm, using a 10-cm pathlength cuvette, with a response time of 1 s and an integration time of 16 s. The protein was unfolded in 4 M GdnHCl for 6 h. Folding was initiated by dilution with the native buffer such that the final GdnHCl concentration and the protein concentration were 0.4 M and 10 μM, respectively. The data obtained were averaged for every 300 s. The data were fit to a single exponential.

Fluorescence-monitored equilibrium and kinetic folding experiments

For GdnHCl-induced equilibrium unfolding experiments, 10 μM of protein was incubated for >6 h in different concentrations of GdnHCl (0–4 M). The equilibrated protein samples were excited at 295 nm, and the fluorescence signals were monitored on an MOS-450 optical system (Biologic), using a 360 ± 10 nm band-pass filter (Asahi spectra). A two-state, N↔U model was used to fit the sigmoidal equilibrium unfolding transitions (40) to obtain the values for the free energy of unfolding in water, $\Delta G_{U,}$ and the midpoint of transition (Cm).

Millisecond kinetic folding experiments were carried out on an SFM-4 stopped-flow machine (Biologic). A mixing dead time of ∼10 ms was achieved using a cuvette of pathlength 0.20 cm, with a flow rate of 5 mL s^{-1}. The final protein concentration was ∼10 μM. The excitation wavelength was 295 nm, and the fluorescence emission was collected using a 360 ± 10 nm band-pass filter (Asahi spectra). The protein was unfolded in 4 M GdnHCl for 6 h. Folding was initiated by rapidly diluting the GdnHCl solution with the native buffer to obtain different final concentrations of GdnHCl. The kinetic traces obtained were fit to the minimum number of statistically significant exponentials.

Double-jump, interrupted folding experiment

Folding was initiated by manually mixing 30 μL of equilibrium-unfolded protein in 4 M GdnHCl with 270 μL of native buffer such that the final GdnHCl concentration was 0.4 M. The folding mixture was aged for different folding times, "t." After aging for time t, 300 μL of the aged-folding mixture was manually mixed with 300 μL of 8 M GdnHCl, and unfolding in ∼4.2 M GdnHCl was monitored by measurement of intrinsic Trp fluorescence. This procedure was followed for pWT and P93A. A similar protocol was used to carry out experiments for P41A, except that the unfolding was carried out at a final GdnHCl concentration of 3.5 M. The final

protein concentration was ~10 μM. For P41AP93A, a Biologic SFM-4 mixing module was used to mix 30 μL of equilibrium-unfolded protein in 4 M GdnHCl with 270 μL of native buffer, such that the final concentration of GdnHCl was 0.4 M. The folding mixture was aged for different durations in a delay loop of 190 μL volume. After different folding times, 50 μL of the solution in the delay loop was mixed with 305 μL of ~4.8 M GdnHCl unfolding buffer, and unfolding was monitored at a final GdnHCl concentration of ~4.2 M in a 0.15-cm pathlength cuvette. The final protein concentration was ~6 μM. A dead time of 35 ms was achieved using flow rates of 1.4 mL s^{-1}.

The kinetic traces for unfolding in 4.2 M GdnHCl (for pWT, P93A, and P41AP93A) and 3.5 M (for P41A) GdnHCl after first folding in 0.4 M GdnHCl for different lengths of time were fit to a single exponential for an exponential phase corresponding to the unfolding of their N state (the unfolding rate constants of the N state for all the four protein variants were determined separately from manual-mixing kinetic unfolding experiments). The signal change of the unfolding kinetic trace at different times of refolding was normalized with that obtained from unfolding of the refolded state, and the fractional change in signal was plotted as a function of the time at which the refolding was interrupted.

RESULTS

The Pro to Ala mutations have a minor effect on structure and do not significantly alter protein stability

As can be seen from the fluorescence spectra of the N state of all four protein variants (Fig. 2 *A*), there was no major change in the fluorescence properties upon mutation of Pro41 and Pro93 to Ala. Hence, the mutations did not significantly affect the environment of Trp19 in the different variants.

The far-UV CD spectrum of each of the four protein variants had a minimum and a maximum at 214 nm and 192 nm, respectively, characteristic of a β-sheet-rich protein (Fig. 2 *B*). This suggested that no significant changes in the polypeptide backbone have occurred upon mutation. The far-UV CD spectrum of P93A was very similar to that of pWT; however, an increase in the amplitude of the minimum and the maximum peak in the case of P41A and P41AP93A was seen. Also, the Pro41 to Ala mutation resulted in a minor change in the shape of the spectrum, which will be discussed later. The extra plateau-like small hump observed at ~205 nm in the far-UV CD spectra of pWT and P93A arises because of aromatic amino acid residue contribution (41,42). Since this peak was absent in the case of P41A and P41AP93A, it is speculated that some stabilizing interaction between Pro41 and an aromatic amino residue has been perturbed by the mutation.

The near-UV CD spectrum of P41AP93A was similar to that of pWT (Fig. 2 *C*). However, there were differences in the spectra of P41A and P93A, especially in the 265–285 nm wavelength region, indicating changes in the packing of one or more Tyr residues that occur upon Pro41 and Pro93 mutations.

Fig. 3 shows the fluorescence-monitored equilibrium unfolding transitions of all four protein variants. The stability

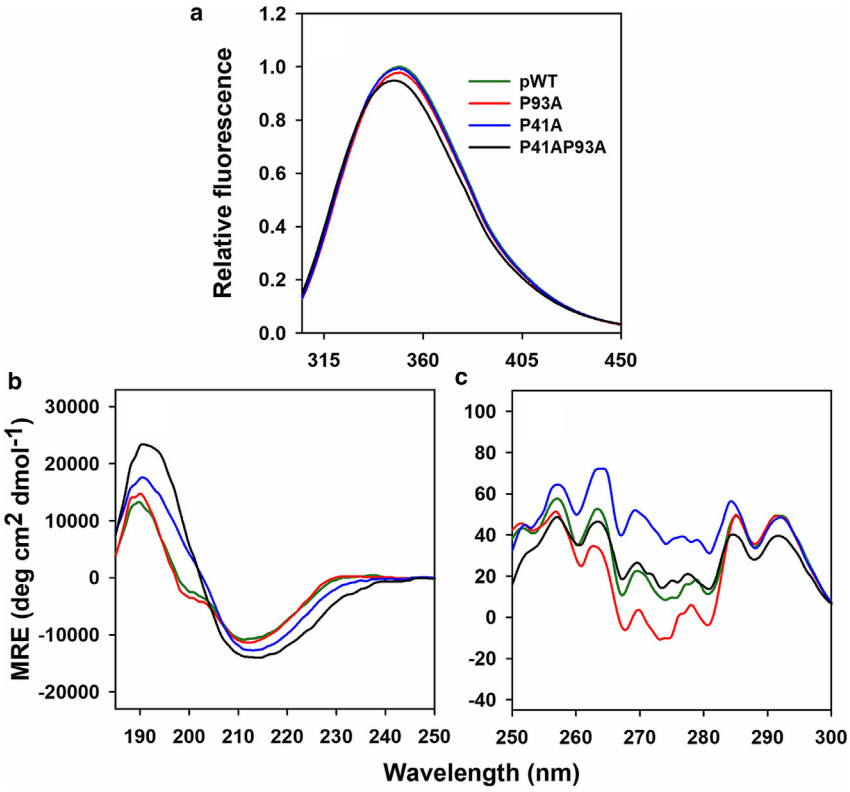

FIGURE 2 Spectroscopic characterization of pWT and the Pro mutant variants. Fluorescence (*a*), far-UV CD (*b*), and near-UV CD (*c*) spectra. The fluorescence spectra of the Pro mutant variants are normalized to that of pWT. The colours of the lines represent the different mutant variants as indicated. To see this figure in color, go online.

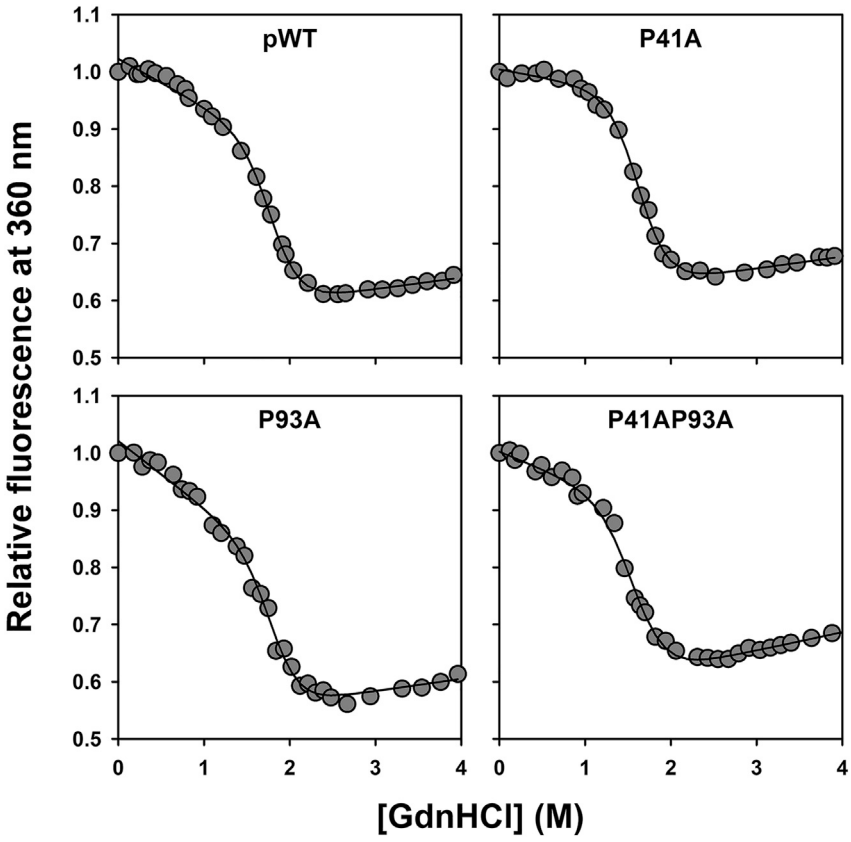

FIGURE 3 Equilibrium unfolding transitions of pWT and the Pro mutant variants. The circles represent the data points, and the solid lines through them are fits to an equation describing a two-state, $N \leftrightarrow U$ unfolding model (40). The fits gave values for ΔG_U of 6.3 ± 0.2, 5.4 ± 0.1, 6.2 ± 0.1, and 5.5 ± 0.1 kcal mol^{-1} for pWT, P41A, P93A, and P41AP93A, respectively. The mid-points (C_m) of the unfolding transition for pWT, P41A, P93A, and P41AP93A are at 1.88 ± 0.04, 1.56 ± 0.02, 1.81 ± 0.02, and 1.61 ± 0.02 M, respectively. The errors in the values of the thermodynamic parameters represent the spread obtained from at least two independent experiments.

of P93A was the same as that of pWT; however, P41A and P41AP93A were slightly destabilized. The stabilities of P41A and P41AP93A were ∼0.9 and ∼0.8 kcal mol^{-1} respectively, lower than that of pWT. Because Pro41 is present in the middle of the chain and participates in the β-sheet of the protein, mutation of this residue to Ala might have disrupted some stabilizing interactions, resulting in a slight decrease in conformational stability.

The very fast phase of folding is absent for P93A and P41AP93A

The time courses of the increase in Trp19 fluorescence for the four protein variants during folding in two different concentrations of GdnHCl are shown in Fig. 4. The kinetic folding traces for pWT and P41A were found to fit best to the sum of three exponentials, suggesting that the folding of these proteins in the millisecond to second time domain occurs in three distinct kinetic phases: very fast (∼10 s^{-1}), fast (∼1 s^{-1}), and slow (∼0.1 s^{-1}). However, in the case of P93A and P41AP93A, the folding traces fit best to the sum of two exponentials, indicating that folding occurs only in two distinct kinetic phases: fast (∼1 s^{-1}) and slow (∼0.1 s^{-1}). Inspection of the residuals for the "sum of three exponentials" fit versus the "sum of two exponentials" fit for the folding kinetic traces of the four protein variants also highlighted the absence of the very fast phase in the case of P93A and P41AP93A (Fig. 4). Fig. 5 shows the dependence of the observed rate constants on GdnHCl and the variation of the relative amplitudes of the kinetic phases in different GdnHCl concentrations for all four protein variants. The extrapolated rate constants of the three phases of folding in water and the dependence of the relative amplitudes on GdnHCl concentration of pWT were similar to that of WT, as reported earlier (6,34). Similar results were obtained for P41A, P93A, and P41AP93A, except that for P93A and P41AP93A, the very fast phase was not observed. The amplitude of the signal change observed in the kinetic folding experiment corresponded well to that observed in the equilibrium unfolding experiment (not shown).

The Pro to Ala mutations lead to overall faster refolding kinetics

Fig. 6 shows the kinetics of formation of the N state monitored by double-jump, interrupted folding experiments, for all four protein variants. The N state of pWT as well as of P93A formed with the apparent rate constants of 0.0009 ± 0.00001 s^{-1} and 0.0006 ± 0.00007 s^{-1}, respectively. The N states of P41A and P41AP93A formed with rate constants of 0.032 ± 0.002 s^{-1} (∼35-fold faster than pWT) and 1.1 ± 0.1 s^{-1} (∼1200-fold faster than pWT),

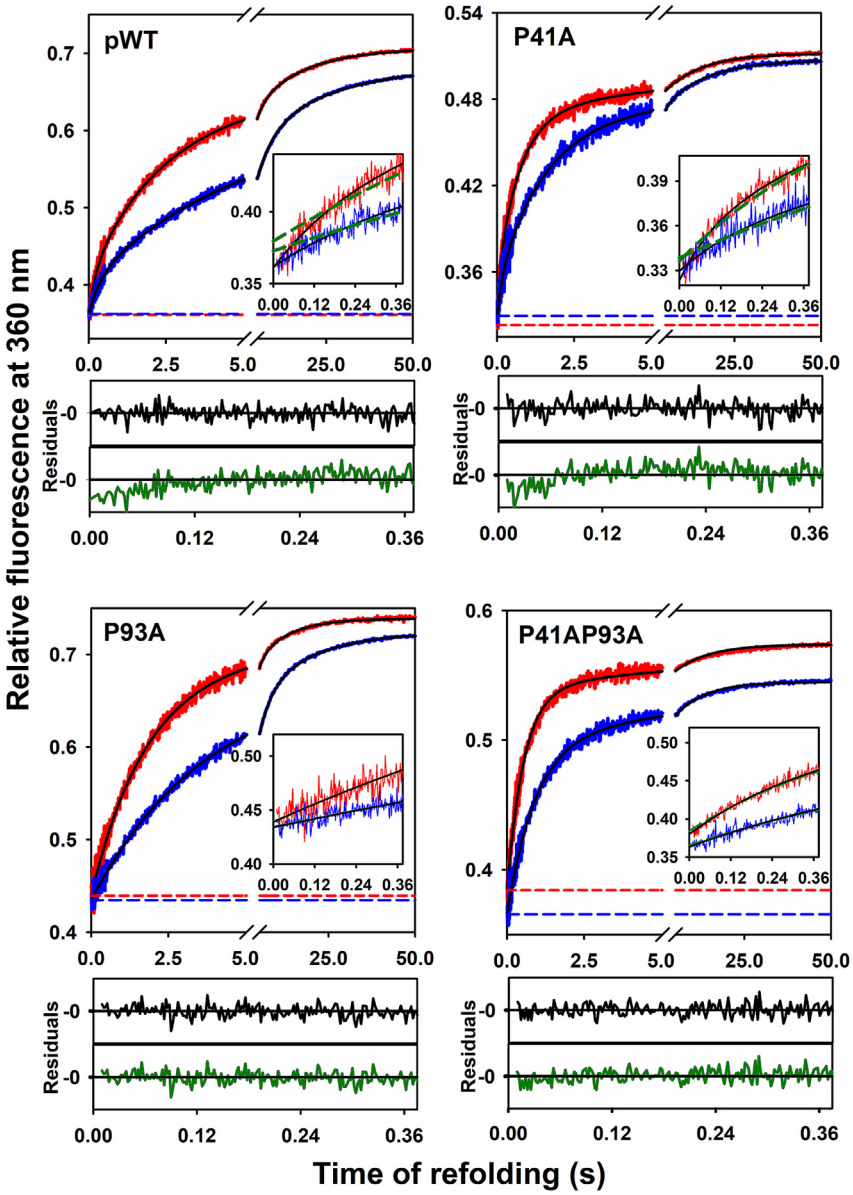

FIGURE 4 Representative folding kinetic traces of pWT and the Pro mutant variants. Relative fluorescence at 360 nm is plotted against the time of refolding. In each panel, the kinetic traces of folding in 0.4 M (red) and 0.6 M (blue) GdnHCl are shown. The extrapolated unfolded protein signals at 0.4 and 0.6 M GdnHCl are shown as red and blue dashed lines, respectively. The fast and slow phases of folding are shown, and the inset in each panel shows the fluorescence change at early times of folding. In the inset and the residuals (for refolding in 0.4 M GdnHCl) below, the dark green and the black colors represent the fit to the "sum of two exponentials" and the "sum of three exponentials," respectively. Inspection of the fits and the residuals shows that the data for pWT and P41A fit well to the sum of three exponentials, with the insets showing the very fast phase of folding. However, the folding traces of P93A and P41AP93A fit well to the sum of two exponentials. To see this figure in color, go online.

respectively. These results indicate that the slowest and, hence, the rate-determining step on the folding pathway of pWT is the last step leading directly to the formation of the N state with a rate constant of 0.0009 $\pm$ 0.00001 s^{-1}. A similar rate constant for the formation of the N state was observed in the case of P93A. The rate constant of formation of the N state in the case of P41A and P41AP93A has increased many fold and appears to lack the rate-determining step seen for pWT and P93A: the mutation of Pro41 to Ala has undoubtedly led to overall faster refolding kinetics. Mutant variants lacking native Pro isomers have been shown to accelerate the (un)folding kinetics for many proteins (43–46). The results shown in Fig. 6 indicate the involvement of Pro41 at the later stages of the folding of pWT. In the case of other proteins too (8,13,15,17,22,23), proline isomerization has been shown to play a role late in folding.

Tertiary contacts are formed in the rate-determining step of refolding

To identify any structural changes accompanying the slowest step on the folding pathway, manual-mixing kinetic studies of folding monitored by far-UV and near-UV CD were carried out. No change was observed in the far-UV CD monitored kinetics (data not shown), suggesting that the secondary structure of the protein had formed early, and that the slowest step does not involve any structural change at the secondary level. For pWT and P93A, the near-UV CD signal was seen to decrease with rate constants

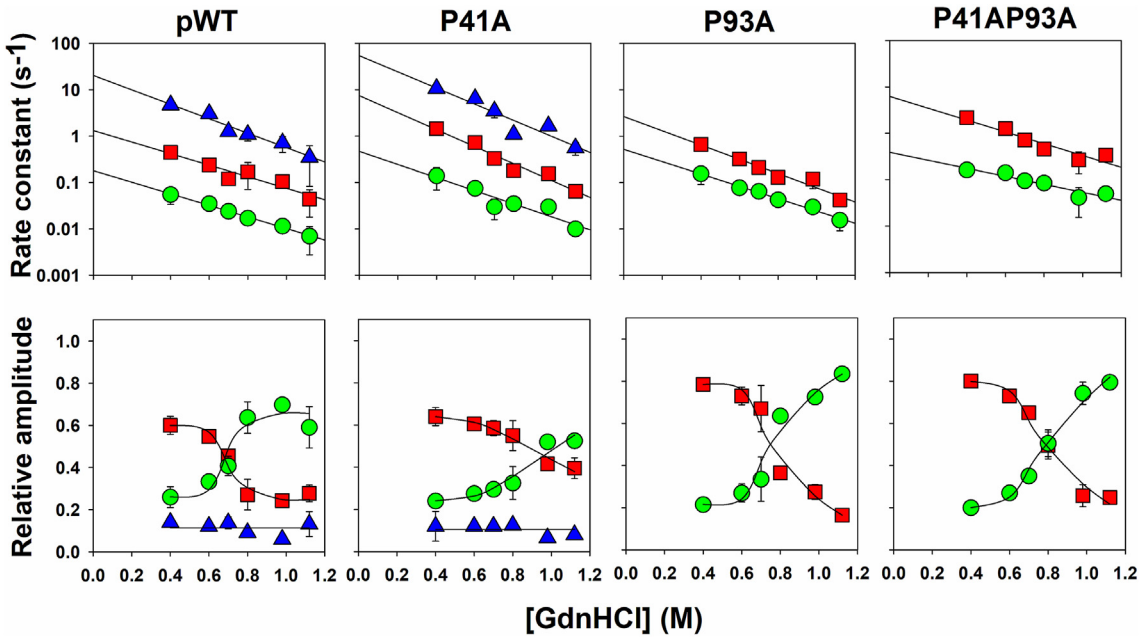

FIGURE 5 Folding kinetics of the MNEI variants. The dependences on GdnHCl concentration of the observed rate constants (*top row*) and the relative amplitudes (*bottom row*) for pWT and the Pro mutant variants are shown. The very fast, fast, and slow kinetic phases of folding monitored by measurement of the change in Trp fluorescence are shown by the blue triangles, red squares, and green circles, respectively. The solid lines through the data points were drawn to guide the eye. The error bars represent the spread in the data obtained from at least two independent experiments. To see this figure in color, go online.

of 0.0008 ± 0.0003 s^{-1} and 0.0005 ± 0.00001 s^{-1}, respectively, suggesting that tight packing of some aromatic amino residues occurs concurrently with the isomerization of Pro41 during the slowest step of refolding. No change in the near-UV CD signal was observed for P41A and P41AP93A (Fig. 7), on the timescale it was observed in the case of pWT and P93A.

DISCUSSION

Effect of the Pro to Ala mutations on protein structure and stability

Replacement of a *cis* Pro residue by an Ala residue that adopts the *trans* conformation might be expected to affect the structure of MNEI. Although the local environment of Trp19 is affected to only a minor extent, as reflected in small differences in the fluorescence spectra between pWT and the different Pro to Ala mutant variants (Fig. 2 *A*), changes in secondary structure as reflected in the far-UV CD spectra (Fig. 2 *B*) appear to be more significant. It is possible that some of the differences in the far-UV CD spectra do not reflect changes in secondary structure but changes in the environment of aromatic residues that can also contribute to the shape of far-UV CD spectra (41,42). Such an explanation is probable, as the near-UV CD spectra, for which the major contribution is from aromatic residues, are seen to be affected significantly (Fig. 2 *C*).

The stabilities of pWT, P41A, P93A, and P41AP93A are, however, not very different (Fig. 3). It is difficult to predict the effect of substituting Pro with another residue. It could be expected that because a Pro residue confers rigidity to the polypeptide chain, substituting it with an Ala residue should entropically stabilize the U state and, hence, reduce protein stability. Moreover, substituting a Pro residue with the smaller Ala should perturb packing interactions and enthalpically destabilize the N state, again reducing protein stability. In the case of MNEI, it also appears that the side chain of Pro41, which is nearly completely buried in the core of the protein (Fig. 1 *A*), forms a CH/π interaction with the aromatic side chain of Tyr64 (Fig. 8 *A*): the far-UV CD spectra of pWT and P93A, but not of P41A and P41AP93A in which Pro41 is substituted with Ala, show a positive plateau-like hump at 205 nm (Fig. 2 *B*) that is known to arise from contributions of aromatic residues (41,42). The stabilizing CH/π interaction would therefore be absent in P41A. Pro93 is the second residue of the hydrophobic GPVPPP sequence segment at the C-terminal end of MNEI, and it appears to have stabilizing interactions with hydrophobic residues, including Phe72, in the β-sheet, which are enabled by it being in the *cis* conformation (Fig. 8 *B* and *C*). The stabilizing hydrophobic interactions appear to be absent in P93A (Fig. 8 *D*). The observation that the Pro93 to Ala mutation barely affects the stability, and that the Pro41 to Ala mutation reduces stability by only 0.9 kcal mol^{-1}, suggests that the destabilizing effect of replacing the *cis* Pro residue must be offset by the stabilizing effect of substituting in the *trans* Ala residue. It seems that the N states of the mutant variants are stabilized by at least one additional hydrogen bond involving the main

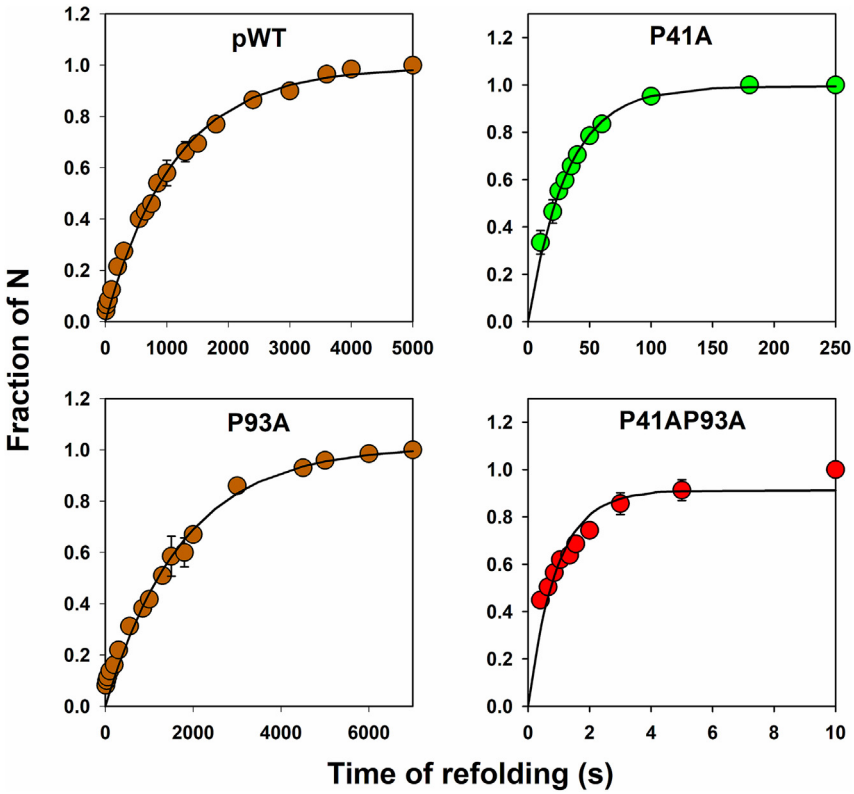

FIGURE 6 Kinetics of formation of the N state. Folding was initiated by effecting a drop in GdnHCl concentration from 4 to 0.4 M. Folding was interrupted at different times, with a jump in GdnHCl concentration from 0.4 to 4.2 M for pWT, P93A, and P41AP93A and from 0.4 to 3.5 M for P41A. The solid lines in all the panels describe the formation of N according to a single-exponential process, with apparent rate constants of 0.0009 ± 0.00001 s^{-1}, 0.032 ± 0.002 s^{-1}, 0.0006 ± 0.00007 s^{-1}, and 1.1 ± 0.1 s^{-1} for pWT, P41A, P93A, and P41AP93A, respectively. The error bars represent the spread in the data obtained from at least two independent experiments. To see this figure in color, go online.

chain, due to the replacement of Pro with Ala, than is the N state of pWT. Such stabilization has been observed when Pro has been substituted with other residues in the case of other proteins (47–49).

Proline residues and heterogeneity in the unfolded state

The presence of two proline residues, Pro41 and Pro93, in their *cis* conformations in the N state suggests that *cis* to *trans* proline isomerization should lead to the population of at least four distinct conformations in slow equilibrium with each other in the U state. U_{C41C93}, U_{C41T93}, U_{T41C93}, and U_{T41T93} are expected to be populated to extents of about 3%, 14%, 14%, and 69%, respectively (11,50). The equilibria between these four U-state conformations would be slow because of the slow rate constant of proline isomerization, which typically falls in the range of 0.01 to 0.001 s^{-1} (39). Since the folding reaction starting from each of these unfolded conformations would be much faster, each conformation would fold via an independent pathway, and the flux on each pathway would be determined by the population of the starting unfolded conformation.

The folding of MNEI in the millisecond to second time domain occurs in three kinetic phases—very fast, fast, and slow—when monitored by changes in Trp fluorescence, when a single Trp residue is present at either residue position 4 in WT MNEI (34,51) or residue position 19 in pWT MNEI (Fig. 4). Double-jump, interrupted folding experiments indicated that the very fast phase of folding arises from a separate population of unfolded molecules (U_1 in Fig. 1 *B*). These experiments also showed that the fast and slow kinetic phases of folding arise from another population of unfolded molecules (U_2 in Fig. 1 *B*). The observation that the fraction of molecules folding via the very fast phase is about 14% (Figs. 4 and 5) suggests that the very fast phase arises from either U_{C41T93} or U_{T41C93}. Since the very fast phase is eliminated upon mutation of Pro93 to Ala, U_1 must correspond to U_{T41C93}. In the case of other proteins too, it has been shown that mutation of Pro to other residues results in a decrease in U-state heterogeneity (44,45,52,53).

U_2 would therefore be expected to be a mixture of U_{C41C93}, U_{C41T93}, and U_{T41T93}. U_{C41C93} is expected to populate to too low an extent (∼3%) for its folding process to be separately detected. U_{C41T93} is, however, expected to be populated sufficiently for its folding to be separately observable. It is possible that U_{C41T93} is less populated in the U-state ensemble, because of the presence of destabilizing interactions that are present when the Arg40-Pro41 bond is in the *cis* conformation and which are absent in the *trans* conformation. It is known that the ratio of *cis* to *trans* isomers can depend not only on the residue preceding Pro (54,55) and on the initial denaturing conditions (56,57) but also on the local environment (58). In any case, U_2 appears to consist predominantly of U_{T41T93}.

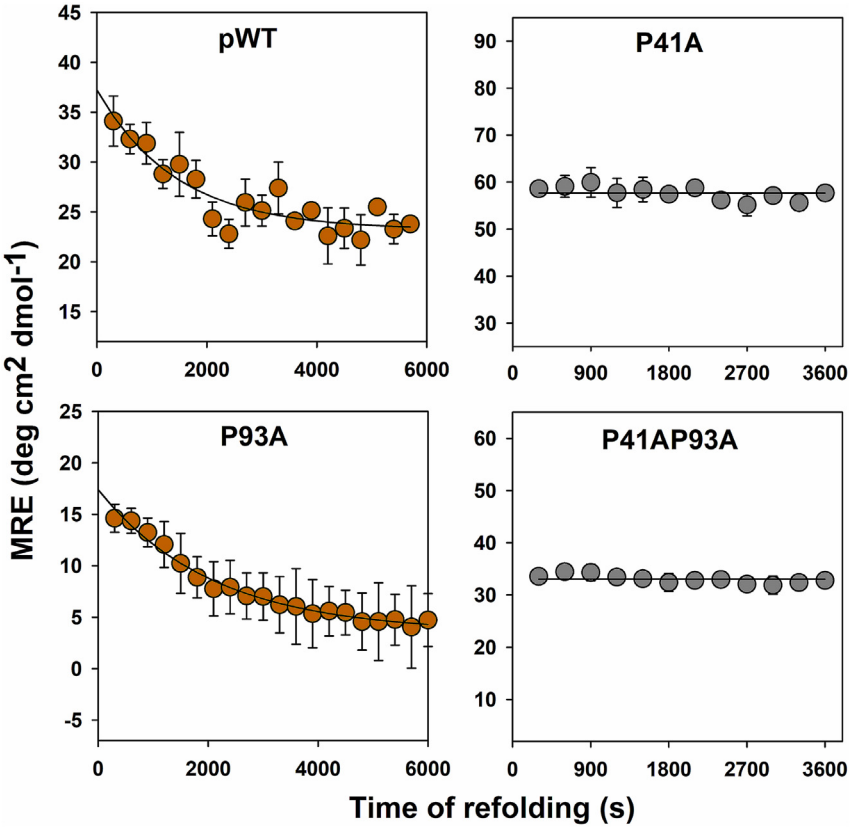

FIGURE 7 Kinetics of formation of N monitored by measurement of near-UV CD at 270 nm for pWT and the Pro mutant variants. The solid lines through the data points for pWT and P93A represent single-exponential processes with apparent rate constants of 0.0008 ± 0.0003 s^{-1} and 0.0005 ± 0.00001 s^{-1}, respectively, whereas the solid lines through the data points for P41A and P41AP93A are drawn to guide the eye. The error bars represent the standard errors obtained from at least two different experiments. To see this figure in color, go online.

cis Pro93 enables faster folding than *trans* Pro93

Previously, the very fast phase of folding had been attributed to the folding of U_1, and the fast and slow phases of folding were attributed to the folding of U_2 (Fig. 1 *B*) (34). The current study validates this aspect of the folding scheme shown in Fig. 1 *B*: the mutation of Pro93 to Ala eliminates only the very fast phase of folding and has no effect on the fast and slow phases (Fig. 5).

The current study also suggests that when Pro93 is in the *cis* conformation, the folding of MNEI is faster because some stabilizing interactions form early. It is possible that these stabilizing interactions are the same hydrophobic interactions that appear to be present in the N state of MNEI (see above), which appear to be enabled by the conformational rigidity provided by the *cis* Gly92-Pro93 bond. In the case of ribonuclease A, an early intermediate stabilized by hydrophobic interactions is populated on the very fast pathway originating from an unfolded form having all Pro residues in their native conformations (59). In the case of yeast phosphoglycerate kinase too, a native *cis* Pro residue speeds up folding (60). A recent study has shown that Pro residues can facilitate the compaction of disordered protein chains (61), in agreement with a much earlier study that showed that native isomers of certain Pro residues can facilitate the association of residues on folding pathways via hydrophobic interactions (62). Specific hydrophobic clustering mediated by Pro during the initial folding steps is also known for CRABP1 (63). In the case of MNEI, a collapsed globule having solvent-exposed hydrophobic patches is known to form very early during folding (37), and if stabilizing hydrophobic interactions are absent early during the folding of U_2, just as they appear to be absent in the N state of the P93A variant in which Ala93 is in the *trans* conformation (Fig. 8 *D*), that could be the reason why the folding of U_2 occurs more slowly than that of U_1 in the case of pWT. For several proteins whose folding is accompanied by *trans* to *cis* Pro isomerization, it has been shown that the unfolded form having the nonnative *trans* Pro conformation folds slower than the unfolded form having the native *cis* Pro conformation (52,53,64), but this is not always the case (13,22,65).

Proline isomerization during folding

Structure formation reactions during protein folding, including chain collapse (66,67) and secondary structure formation (2), are fast compared with the timescale of proline isomerization. It is therefore not surprising that for many proteins that contain *cis* Pro residues in their native conformations, *trans* to *cis* isomerization occurs as the last step in folding (17,22–24,68,69). In many (22,63,68,69) but not all (23,24,70) cases, the *trans* to *cis* isomerization

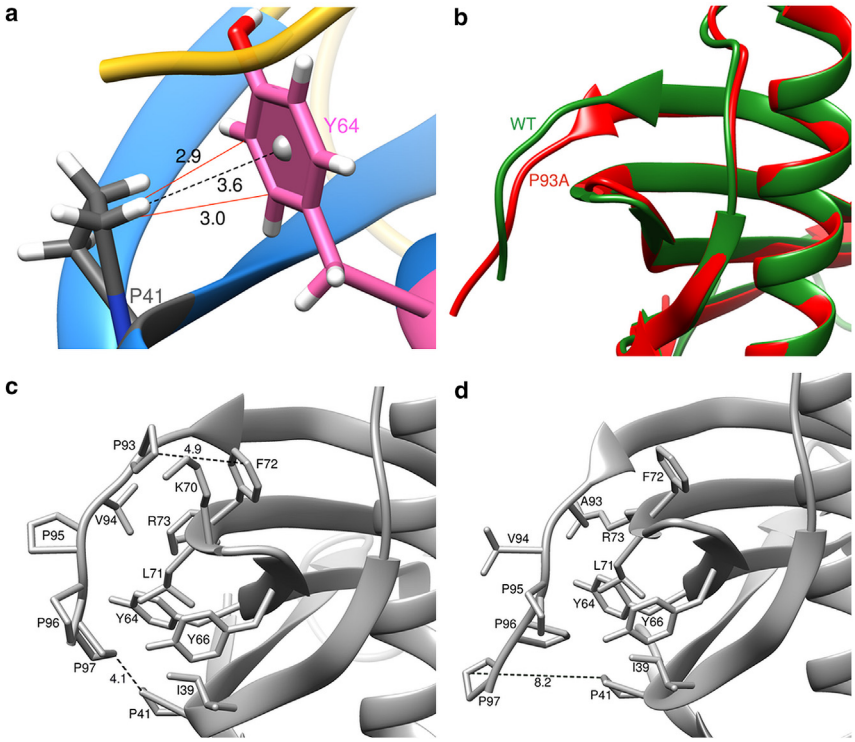

FIGURE 8 Structural interactions of Pro41 and Pro93. (*a*) Structure showing the apparent CH/π interaction between Pro41 (in the β2-strand) and Tyr64 (in the β3-strand). The solid red lines represent the pseudo-bonds given by the software (Chimera). The dashed line represents the distance separating the centroid of the Tyr ring to the δ hydrogen of Pro41. (*b*) Structures of MNEI (PDB ID: 1IV7) and P93A (PDB ID: 7EUA) overlapped to highlight the difference at the C-terminus region. The ribbon colors represent the mutant variants as indicated. (*c*) The WT structure highlights the hydrophobic interactions between the residues of the C-terminus "GPVPPP" sequence with those in the β-sheet. (*d*) Structure of the P93A showing the apparent loss of the hydrophobic interactions upon mutation of Pro93 to Ala. The distance separating Pro41 from Pro97 is seen to increase from 4.1 to 8.2 Å when Pro93 is mutated to Ala. All the distances are shown in angstroms. To see this figure in color, go online.

step appears to be silent to the commonly used (Fig. 4) spectroscopic probes, indicating that major structure acquisition precedes the proline isomerization step. For example, in the case of MNEI, the very fast, fast, and slow phases of folding can be monitored by measuring changes in fluorescence or CD (34,51), but not the very slow phase whose rate constant corresponds to that of proline isomerization. The very slow phase can only be monitored by measurements of ANS fluorescence (34,37) and near-UV CD (Fig. 7 *A*), as well as by double-jump, interrupted folding experiments (34). No lag phase is observed in the formation of the N. A possible reason for this could be the presence of multiple folding pathways, with native molecules being formed more rapidly along the folding pathway arising from the U_{C41C93} subensemble, which is populated to 3%–5% (discussed above). Proline isomerization would not be rate limiting during the folding of the U_{C41C93} subensemble.

Isomerization of Pro41

In the case of pWT, the folding of U_2 to N must be accompanied by the *trans* to *cis* isomerization of both Pro41 and Pro93. The very slow phase of folding of pWT, which leads to the formation of N, is about threefold slower than that of WT MNEI. In the latter case, it had been shown previously that the very slow phase of folding is independent of denaturant concentration (34). In the case of many other proteins, it has been shown that folding steps involving proline isomerization reactions can be independent of denaturant concentration (13,71,72). Hence, the very slow phase of the folding of WT MNEI had been attributed to proline isomerization (34). The observation, from double-jump, interrupted refolding experiments, that the N state forms with a rate constant of about $0.0009 \pm 0.00001\ s^{-1}$ in the case of pWT, and that the rate constant of the very slow phase is about the same for the P93A variant (Fig. 6), suggests that this very slow phase of folding must correspond to the *trans* to *cis* isomerization of Pro41. This conclusion is validated by the observation that when Pro41 is mutated to Ala in P41A and P41AP93A, the very slow phase of folding is abolished (Fig. 6), and the N state forms much more rapidly.

In the N state of pWT and WT MNEI, Pro41 introduces a sharp bend of about 90° in the second β-strand, which allows tight packing between the β-sheet and the helix of MNEI (Fig. 1 *A*). The side chain of Pro41 is itself nearly completely buried. It is likely that during the folding of pWT, Pro41 remains solvent accessible in its *trans* conformation until the last step of folding, when the Arg40-Pro41 bond undergoes *trans* to *cis* isomerization, and the side chain of the *cis* Pro41 isomer gets inserted into the hydrophobic core of the protein. This would require that at least a few of the hydrophobic core residues remain solvent accessible until the last step. It is therefore not surprising that the hydrophobic dye ANS can remain bound to folding MNEI molecules until the commencement of the very slow phase of folding (34,37). It has been shown, both in the case of WT MNEI (34,37) and pWT (data not shown), that ANS becomes completely unbound only during the very slow

phase of folding that leads directly to the formation of N. It is also not surprising that the very slow phase of the folding of both pWT and P93A involves a final consolidation of tertiary structure measurable by a change in near-UV CD, given that it involves accommodation of the side chain of Pro41 in the hydrophobic core of the protein. Not surprisingly, no change in near-UV circular dichroism is detectable in the last step leading to the formation of N in the case of the P41A and P41AP93A variants in which Pro41 is replaced by Ala (Fig. 7). Furthermore, it is also not surprising to find that the *trans* to *cis* isomerization of Pro41 occurs in the compact and highly structured, native-like intermediate N* (7) as it transits to N (Fig. 1 *B*). Indeed, proline isomerization has been found to occur in native states of proteins (73–76).

Isomerization of Pro93

It was possible that *trans* to *cis* isomerization of Pro93 occurs with the same very slow rate constant as that of *trans* to *cis* isomerization of Pro41 during the folding of pWT. Such a possibility could be discounted in the current study, which shows that the very slow phase of folding is absent in the case of the P41A variant that possesses *cis* Pro93 in its N state. The observed rate constant of the slow phase of folding (Fig. 4) is within the range of rate constants with which proline isomerization can occur during folding. The observation from double-jump, interrupted refolding experiments that the N state of P41A forms during the slow phase of folding indicates that *trans* to *cis* isomerization of Pro93 must occur during the slow phase of folding of pWT, as well as of P93A.

Hence, the *trans* to *cis* isomerization reactions of the Gly92-Pro93 and Arg40-Pro41 peptide bonds occur sequentially during the folding of pWT, with the former occurring during the slow phase and the latter during the very slow phase of folding. It is perhaps not surprising that the *trans* to *cis* isomerization of the Arg40-Pro41 bond is slower than that of the Gly92-Pro43 bond, as it is known that when a bulky residue (such as Arg) immediately precedes a Pro residue, isomerization is retarded in comparison to when a smaller residue (such as Gly) is the preceding residue (58).

More likely, the *trans* to *cis* isomerization reaction of Pro93 is speeded up significantly because it is accompanied by substantial structure formation: the slow phase is known to lead to substantial formation of secondary and tertiary structure (34,37). The formation of stabilizing structure would stabilize the transition state that slows down *trans* to *cis* isomerization of Pro93. In contrast, the *trans* to *cis* isomerization reaction of Pro41 is not accompanied by any secondary structure formation, and it is accompanied by the formation of only minor tertiary structure. In the case of both staphylococcal nuclease (25) and CRABP1 (77), a kinetic phase known to involve proline isomerization was not abolished upon replacing the Pro residue, indicating that the kinetic phase also involved structure formation. For other proteins too, proline isomerization has been shown to be faster when accompanied by structure formation (78–80), although it is also possible for structure in an intermediate to slow down proline isomerization (64). It appears that in the case of MNEI, the slow phase of structure formation and the *trans* to *cis* isomerization reaction of Pro93 occur with about the same rate constant, as the replacement of the *cis* Gly92-Pro93 bond with the *trans* Gly92-Ala93 bond in P93A does not significantly affect the rate constant of the slow phase of folding.

Simplification of the folding mechanism upon mutation of the Pro residues

The current study indicates that in the case of pWT, *trans* to *cis* isomerization of Pro93 occurs during the slow phase leading to the formation of the native-like intermediate, N* (Fig. 1 *B*), and that *trans* to *cis* isomerization of Pro41 occurs during the very slow phase corresponding to the transition of N* to N. The current study also shows that the U-state heterogeneity seen in the case of pWT originates in the *cis* to *trans* isomerization of Pro93. Fig. 1 *B* also shows that the fast and slow phases of folding arise because of the transient population of intermediates, and it is known that these intermediates are structurally distinct in having different structural segments collapsed (7). Thus, it could be expected that mutations might differentially affect the stabilities of the folding intermediates and hence determine which intermediates are populated sufficiently to be observable. Hence, it would not be surprising if the mutations of Pro41 and Pro93 to Ala, singly as well as together, lead to apparently simpler folding mechanisms.

Folding mechanism of the P41A variant

Since U_1 is populated, the very fast phase is observed when folding is monitored using the fluorescence of Trp19 as the probe. The fast and slow phases of folding are also observed. The Pro41 to Ala mutation appears to destabilize the native-like intermediate, N*, seen in the case of pWT so that it is populated to a negligible extent. The very slow phase of folding is absent, and the N state appears to form as a product of the slow phase of folding. The very fast, fast, and slow phases of fluorescence change are observed. *Trans* to *cis* isomerization of the Gly92-Pro93 peptide bond occurs during the slow phase of folding. Thus, the folding mechanism of pWT (Fig. 1 *B*) simplifies upon mutation of Pro41 to Ala (Fig. 9 *A*).

Folding mechanism of the P93A variant

Since U_1 is absent, the very fast phase of folding is not observed. Only the fast and slow phases are observed when folding is monitored using the fluorescence of Trp19 as the probe. *Trans* to *cis* isomerization of the Arg40-Pro41 peptide bond occurs during the very slow phase of folding as N*

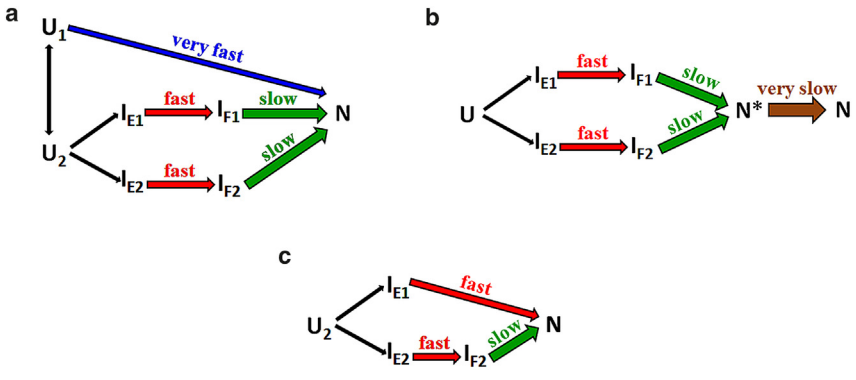

FIGURE 9 Simplification of the complex folding mechanism of MNEI achieved upon replacement of the native *cis* Pro by Ala. The folding schemes of (*a*) P41A, (*b*) P93A, and (*c*) P41AP93A. To see this figure in color, go online.

transitions to N. Thus, the folding mechanism of pWT (Fig. 1 *B*) simplifies upon mutation of Pro93 to Ala (Fig. 9 *B*).

Folding mechanism of the P41AP93A variant

U_1 is absent; hence, the very fast phase is not observed when folding is monitored using Trp fluorescence as the probe. Only the fast and slow phases of folding are observed. Folding is no longer limited by *trans* to *cis* proline isomerization as both the *cis* Pro41 and *cis* Pro93 have been replaced by Ala residues that adopt *trans* conformations. N forms via two pathways on which the slow folding intermediates differ in having different structural segments compacted (7), which would result in them differing in their stabilities. It appears that on one of these pathways, the slow folding intermediate is destabilized sufficiently by the mutations so that it is no longer populated significantly; hence, the N state forms during the fast phase of folding on that pathway. On the other pathway, N forms during the slow phase of folding. The observed rate constant for the formation of N, about 1.1 s^{-1}, is the sum of the amplitude-weighted rate constants for the formation of N along the two pathways. Thus, the folding mechanism of pWT (Fig. 1 *B*) simplifies significantly upon replacement of both the native *cis* Pro to Ala (Fig. 9 *C*).

Comparison of the folding mechanisms of the single-chain and double-chain variants of monellin

The folding mechanism of naturally occurring double-chain monellin (dcMN) appears to be simpler than that of its artificially created single-chain counterpart, MNEI (34,81). Although the structures of the two variants are virtually identical, it was possible that the increased complexity observed for the folding of MNEI was due to changes in enthalpy as well as lower chain entropy caused by the covalent linkage of the two chains in MNEI. The current study suggests that this is not the case. The folding mechanism of MNEI has been shown to be complex because of proline isomerization, and it reduces to a much simpler mechanism similar to that of dcMN upon mutation of the two *cis* Pro residues to Ala. What remains surprising is why the same two Pro residues, which are also *cis* in dcMN, do not introduce similar complexity into the folding mechanism of dcMN. Future studies will investigate the role of proline isomerization in the folding of dcMN.

Proline isomerization and protein folding: Lessons from MNEI

The current study has highlighted several aspects of proline isomerization reactions coupled to protein folding, which need to be taken into consideration in any protein folding study. 1) It is invariably assumed, when parallel pathways of folding are observed for a protein containing Pro residues, that their origin lies in the heterogeneity arising because of the presence of the Pro residues. The observation in the case of MNEI that competing pathways remain even when both Pro residues are replaced suggests that the complexity of folding of Pro-containing proteins can arise for reasons other than proline isomerization. 2) Proline isomerization is known to occur in the unfolded state (15–17), and substitution of Pro with other residues invariably results in the abolishment of slow phases of folding. It seems now that fast phases of folding may also be abolished: the mutation of Pro93 to Ala leads to complete abolishment of the very fast phase of folding of MNEI (Fig. 9 *B*). 3) Proline isomerization reactions are known to occur late during the course of protein folding (8,13,15,17,22,23), and it is therefore not surprising that in the case of MNEI, the very slow phase of folding is abolished upon mutation of Pro41 to Ala (Fig. 9 *A*). What is surprising is that Pro to Ala mutations may differentially affect the stabilities of folding intermediates in a manner that leads to the abolishment of a slow folding phase on one pathway but not a competing pathway (Fig. 9 *C*). 4) When multiple *cis* Pro residues are present in a protein, their *trans* to *cis* isomerization reactions during folding may occur not concurrently but sequentially with very different rate constants. 5) Pro residues may play a role in preventing aggregation of folding intermediates. In the case of MNEI, substitution of Pro41 with Ala leads to aggregation during folding at higher (>60 μM) protein concentrations at which the WT protein is fully soluble.

CONCLUSION

The Pro to Ala mutations simplify the folding mechanism of MNEI in multiple ways. By reducing unfolded state heterogeneity, the Pro93 to Ala mutation results in the elimination of a folding pathway. The Pro41 to Ala mutation results in the elimination of the slowest proline isomerization step, in the destabilization of the late native-like intermediate, N*, and in the faster formation of N. Together, the Pro41 to Ala and Pro93 to Ala mutations result in a simpler folding mechanism in which elimination of the proline isomerization steps and destabilization of folding intermediates lead to more than 1000-fold acceleration of folding.

AUTHOR CONTRIBUTIONS

A.K. carried out the experiments and data analysis. J.B.U. designed the research and oversaw the experiments and analysis. A.K. and J.B.U. wrote the paper.

ACKNOWLEDGMENTS

We thank members of our laboratory for discussions and Dr. Sandhya Bhatia for making the single Pro to Ala mutations and for preliminary experiments. A.K. is a recipient of Prime Minister Research's fellowship (PMRF) awarded by the Ministry of Human Resource Development, Government of India. J.B.U. is a recipient of the J.C. Bose National Fellowship from the Government of India. This work was funded by the Indian Institute of Science Education and Research Pune and a grant from the Science and Engineering Research Board of the Government of India.

DECLARATION OF INTERESTS

The authors declare no competing financial interest.

REFERENCES

1. Wallace, L. A., and C. R. Matthews. 2002. Sequential vs. parallel protein-folding mechanisms: experimental tests for complex folding reactions. *Biophys. Chem.* 101–102:113–131.
2. Udgaonkar, J. B. 2008. Multiple routes and structural heterogeneity in protein folding. *Annu. Rev. Biophys.* 37:489–510.
3. Aksel, T., and D. Barrick. 2014. Direct observation of parallel folding pathways revealed using a symmetric repeat protein system. *Biophys. J.* 107:220–232.
4. Eaton, W. A., and P. G. Wolynes. 2017. Theory, simulations, and experiments show that proteins fold by multiple pathways. *Proc. Natl. Acad. Sci. USA.* 114:E9759–E9760.
5. Chavez, L. L., S. Gosavi, …, J. N. Onuchic. 2006. Multiple routes lead to the native state in the energy landscape of the β-trefoil family. *Proc. Natl. Acad. Sci. USA.* 103:10254–10258.
6. Bhatia, S., G. Krishnamoorthy, …, J. B. Udgaonkar. 2019. Observation of continuous contraction and a metastable misfolded state during the collapse and folding of a small protein. *J. Mol. Biol.* 431:3814–3826.
7. Bhatia, S., G. Krishnamoorthy, and J. B. Udgaonkar. 2021. Mapping distinct sequences of structure formation differentiating multiple folding pathways of a small protein. *J. Am. Chem. Soc.* 143:1447–1457.
8. Brandts, J. F., H. R. Halvorson, and M. Brennan. 1975. Consideration of the possibility that the slow step in protein denaturation reactions is due to cis-trans isomerism of proline residues. *Biochemistry.* 14:4953–4963.
9. Garel, J.-R., B. T. Nall, and R. L. Baldwin. 1976. Guanidine-unfolded state of ribonuclease A contains both fast-and slow-refolding species. *Proc. Natl. Acad. Sci. USA.* 73:1853–1857.
10. Kiefhaber, T., R. Quaas, …, F. X. Schmid. 1990. Folding of ribonuclease T1. 1. Existence of multiple unfolded states created by proline isomerization. *Biochemistry.* 29:3053–3061.
11. Kiefhaber, T., H.-H. Kohler, and F. X. Schmid. 1992. Kinetic coupling between protein folding and prolyl isomerization: I. Theoretical models. *J. Mol. Biol.* 224:217–229.
12. Kiefhaber, T., and F. X. Schmid. 1992. Kinetic coupling between protein folding and prolyl isomerization: II. Folding of ribonuclease A and ribonuclease T1. *J. Mol. Biol.* 224:231–240.
13. Shastry, M. C., V. R. Agashe, and J. B. Udgaonkar. 1994. Quantitative analysis of the kinetics of denaturation and renaturation of barstar in the folding transition zone. *Protein Sci.* 3:1409–1417.
14. Wu, Y., and C. R. Matthews. 2002. Parallel channels and rate-limiting steps in complex protein folding reactions: prolyl isomerization and the alpha subunit of Trp synthase, a TIM barrel protein. *J. Mol. Biol.* 323:309–325.
15. Kim, P. S., and R. L. Baldwin. 1982. Specific intermediates in the folding reactions of small proteins and the mechanism of protein folding. *Annu. Rev. Biochem.* 51:459–489.
16. Eberhardt, E. S., S. N. Loh, and R. T. Raines. 1993. Thermodynamic origin of prolyl peptide bond isomers. *Tetrahedron Lett.* 34:3055–3056.
17. Wedemeyer, W. J., E. Welker, and H. A. Scheraga. 2002. Proline cis–trans isomerization and protein folding. *Biochemistry.* 41:14637–14644.
18. Bhuyan, A. K., and J. B. Udgaonkar. 1998. Multiple kinetic intermediates accumulate during the unfolding of horse cytochrome c in the oxidized state. *Biochemistry.* 37:9147–9155.
19. Pierce, M. M., and B. T. Nall. 2000. Coupled kinetic traps in cytochrome c folding: His-heme misligation and proline isomerization. *J. Mol. Biol.* 298:955–969.
20. Welker, E., W. J. Wedemeyer, …, H. A. Scheraga. 2001. Coupling of conformational folding and disulfide-bond reactions in oxidative folding of proteins. *Biochemistry.* 40:9059–9064.
21. Stewart, D. E., A. Sarkar, and J. E. Wampler. 1990. Occurrence and role ofcis peptide bonds in protein structures. *J. Mol. Biol.* 214:253–260.
22. Schreiber, G., and A. R. Fersht. 1993. The refolding of cis-and trans-peptidylprolyl isomers of barstar. *Biochemistry.* 32:11195–11203.
23. Wu, Y., and C. R. Matthews. 2002. A cis-prolyl peptide bond isomerization dominates the folding of the alpha subunit of Trp synthase, a TIM barrel protein. *J. Mol. Biol.* 322:7–13.
24. Veeraraghavan, S., B. T. Nall, and A. L. Fink. 1997. Effect of prolyl isomerase on the folding reactions of staphylococcal nuclease. *Biochemistry.* 36:15134–15139.
25. Maki, K., T. Ikura, …, K. Kuwajima. 1999. Effects of proline mutations on the folding of staphylococcal nuclease. *Biochemistry.* 38:2213–2223.
26. Bhuyan, A. K., and J. B. Udgaonkar. 1999. Observation of multistate kinetics during the slow folding and unfolding of barstar. *Biochemistry.* 38:9158–9168.
27. Herning, T., K. Yutani, …, M. Kikuchi. 1991. Effects of proline mutations on the unfolding and refolding of human lysozyme: the slow refolding kinetic phase does not result from proline cis-trans isomerization. *Biochemistry.* 30:9882–9891.
28. Tweedy, N. B., S. K. Nair, …, D. W. Christianson. 1993. Structure and energetics of a non-proline cis-peptidyl linkage in a proline-202. fwdarw. alanine carbonic anhydrase II variant. *Biochemistry.* 32:10944–10949.
29. Mayr, L. M., D. Willbold, …, F. X. Schmid. 1994. Generation of a non-prolyl cis peptide bond in ribonuclease T1. *J. Mol. Biol.* 240:288–293.
30. Odefey, C., L. M. Mayr, and F. X. Schmid. 1995. Non-prolyl cis-trans peptide bond isomerization as a rate-determining step in protein unfolding and refolding. *J. Mol. Biol.* 245:69–78.

31. Vanhove, M., X. Raquet, …, J. M. Frère. 1996. The rate-limiting step in the folding of the cis-Pro167Thr mutant of TEM-1 β-lactamase is the trans to cis isomerization of a non-proline peptide bond. Proteins: Structure, Function. *Proteins.* 25:104–111.
32. Walkenhorst, W. F., S. M. Green, and H. Roder. 1997. Kinetic evidence for folding and unfolding intermediates in staphylococcal nuclease. *Biochemistry.* 36:5795–5805.
33. Pappenberger, G., H. Aygün, …, T. Kiefhaber. 2001. Nonprolyl cis peptide bonds in unfolded proteins cause complex folding kinetics. *Nat. Struct. Biol.* 8:452–458.
34. Patra, A. K., and J. B. Udgaonkar. 2007. Characterization of the folding and unfolding reactions of single-chain monellin: evidence for multiple intermediates and competing pathways. *Biochemistry.* 46:11727–11743.
35. Kimura, T., T. Uzawa, …, T. Fujisawa. 2005. Specific collapse followed by slow hydrogen-bond formation of β-sheet in the folding of single-chain monellin. *Proc. Natl. Acad. Sci. USA.* 102:2748–2753.
36. Jha, S. K., D. Dhar, …, J. B. Udgaonkar. 2009. Continuous dissolution of structure during the unfolding of a small protein. *Proc. Natl. Acad. Sci. USA.* 106:11113–11118.
37. Goluguri, R. R., and J. B. Udgaonkar. 2016. Microsecond rearrangements of hydrophobic clusters in an initially collapsed globule prime structure formation during the folding of a small protein. *J. Mol. Biol.* 428:3102–3117.
38. Malhotra, P., and J. B. Udgaonkar. 2016. Secondary structural change can occur diffusely and not modularly during protein folding and unfolding reactions. *J. Am. Chem. Soc.* 138:5866–5878.
39. Schmid, F. X., and R. L. Baldwin. 1979. The rate of interconversion between the two unfolded forms of ribonuclease A does not depend on guanidinium chloride concentration. *J. Mol. Biol.* 133:285–287.
40. Agashe, V. R., and J. B. Udgaonkar. 1995. Thermodynamics of denaturation of barstar: evidence for cold denaturation and evaluation of the interaction with guanidine hydrochloride. *Biochemistry.* 34:3286–3299.
41. Woody, R. W. 1978. Aromatic side-chain contributions to the far ultraviolet circular dichroism of peptides and proteins. *Biopolymers.* 17:1451–1467.
42. Chakrabartty, A., T. Kortemme, …, R. L. Baldwin. 1993. Aromatic side-chain contribution to far-ultraviolet circular dichroism of helical peptides and its effect on measurement of helix propensities. *Biochemistry.* 32:5560–5565.
43. Brandts, J. F., M. Brennan, and L.-N. Lin. 1977. Unfolding and refolding occur much faster for a proline-free proteins than for most proline-containing proteins. *Proc. Natl. Acad. Sci. USA.* 74:4178–4181.
44. Kiefhaber, T., H. P. Grunert, …, F. X. Schmid. 1990. Replacement of a cis proline simplifies the mechanism of ribonuclease T1 folding. *Biochemistry.* 29:6475–6480.
45. Wu, Y., and C. R. Matthews. 2003. Proline replacements and the simplification of the complex, parallel channel folding mechanism for the alpha subunit of Trp synthase, a TIM barrel protein. *J. Mol. Biol.* 330:1131–1144.
46. Roderer, D. J., M. A. Schärer, …, R. Glockshuber. 2015. Acceleration of protein folding by four orders of magnitude through a single amino acid substitution. *Sci. Rep.* 5:1–16.
47. Pal, D., and P. Chakrabarti. 1999. Cis peptide bonds in proteins: residues involved, their conformations, interactions and locations. *J. Mol. Biol.* 294:271–288.
48. Jain, R., R. Jain, …, S. Ramakumar. 2002. Structural consequences of replacement of an α-helical Pro residue in Escherichia coli thioredoxin. *Protein Eng.* 15:627–633.
49. Bajaj, K., M. S. Madhusudhan, …, R. Varadarajan. 2007. Stereochemical criteria for prediction of the effects of proline mutations on protein stability. *PLoS Comput. Biol.* 3:e241.
50. Bradley, C. M., and D. Barrick. 2005. Effect of multiple prolyl isomerization reactions on the stability and folding kinetics of the notch ankyrin domain: experiment and theory. *J. Mol. Biol.* 352:253–265.
51. Goluguri, R. R., and J. B. Udgaonkar. 2015. Rise of the Helix from a Collapsed Globule during the Folding of Monellin. *Biochemistry.* 54:5356–5365.
52. Nakano, T., L. C. Antonino, …, A. L. Fink. 1993. Effect of proline mutations on the stability and kinetics of folding of staphylococcal nuclease. *Biochemistry.* 32:2534–2541.
53. Napolitano, S., A. Pokharna, …, A. D. Gossert. 2021. The trans-to-cis proline isomerization in E. coli Trx folding is accelerated by trans prolines. *Biophys. J.* 120:5207–5218.
54. MacArthur, M. W., and J. M. Thornton. 1991. Influence of proline residues on protein conformation. *J. Mol. Biol.* 218:397–412.
55. Reimer, U., G. Scherer, …, G. Fischer. 1998. Side-chain effects on peptidyl-prolyl cis/trans isomerisation. *J. Mol. Biol.* 279:449–460.
56. Houry, W. A., D. M. Rothwarf, and H. A. Scheraga. 1994. A very fast phase in the refolding of disulfide-intact ribonuclease A: implications for the refolding and unfolding pathways. *Biochemistry.* 33:2516–2530.
57. Pertinhez, T. A., D. Hamada, …, C. M. Dobson. 2000. Initial denaturing conditions influence the slow folding phase of acylphosphatase associated with proline isomerization. *Protein Sci.* 9:1466–1473.
58. Grathwohl, C., and K. Wüthrich. 1981. NMR studies of the rates of proline cis–trans isomerization in oligopeptides. *Biopolymers.* 20:2623–2633.
59. Houry, W. A., and H. A. Scheraga. 1996. Structure of a Hydrophobically Collapsed Intermediate on the Conformational Folding Pathway of Ribonuclease A Probed by Hydrogen− Deuterium Exchange. *Biochemistry.* 35:11734–11746.
60. Osváth, S., and M. Gruebele. 2003. Proline can have opposite effects on fast and slow protein folding phases. *Biophys. J.* 85:1215–1222.
61. Mateos, B., C. Conrad-Billroth, …, R. Pierattelli. 2020. The ambivalent role of proline residues in an intrinsically disordered protein: from disorder promoters to compaction facilitators. *J. Mol. Biol.* 432:3093–3111.
62. Lewis, P. N., F. A. Momany, and H. A. Scheraga. 1973. Chain reversals in proteins. *Biochim. Biophys. Acta.* 303:211–229.
63. Eyles, S. J., and L. M. Gierasch. 2000. Multiple roles of prolyl residues in structure and folding. *J. Mol. Biol.* 301:737–747.
64. Kiefhaber, T., R. Quaas, …, F. X. Schmid. 1990. Folding of ribonuclease T1. 2. Kinetic models for the folding and unfolding reactions. *Biochemistry.* 29:3061–3070.
65. Shastry, M. C., and J. B. Udgaonkar. 1995. The folding mechanism of barstar: evidence for multiple pathways and multiple intermediates. *J. Mol. Biol.* 247:1013–1027.
66. Haran, G. 2012. How, when and why proteins collapse: the relation to folding. *Curr. Opin. Struct. Biol.* 22:14–20.
67. Udgaonkar, J. B. 2013. Polypeptide chain collapse and protein folding. *Arch. Biochem. Biophys.* 531:24–33.
68. Kelley, R. F., and F. M. Richards. 1987. Replacement of proline-76 with alanine eliminates the slowest kinetic phase in thioredoxin folding. *Biochemistry.* 26:6765–6774.
69. Pappenberger, G., A. Bachmann, …, T. Kiefhaber. 2003. Kinetic mechanism and catalysis of a native-state prolyl isomerization reaction. *J. Mol. Biol.* 326:235–246.
70. Schmid, F. X., R. Grafl, …, J. J. Beintema. 1986. Role of proline peptide bond isomerization in unfolding and refolding of ribonuclease. *Proc. Natl. Acad. Sci. USA.* 83:872–876.
71. Cook, K. H., F. X. Schmid, and R. L. Baldwin. 1979. Role of proline isomerization in folding of ribonuclease A at low temperatures. *Proc. Natl. Acad. Sci. USA.* 76:6157–6161.
72. Ogasahara, K., and K. Yutani. 1997. Equilibrium and kinetic analyses of unfolding and refolding for the conserved proline mutants of tryptophan synthase α subunit. *Biochemistry.* 36:932–940.
73. Hodel, A., L. M. Rice, …, A. T. Brünger. 1995. Proline cis-trans isomerization in staphylococcal nuclease: Multi-substate free energy perturbation calculations. *Protein Sci.* 4:636–654.

74. Andreotti, A. H. 2003. Native state proline isomerization: an intrinsic molecular switch. *Biochemistry.* 42:9515–9524.

75. Mukaiyama, A., T. Nakamura, …, K. Kuwajima. 2013. Native-state heterogeneity of β2-microglobulin as revealed by kinetic folding and real-time NMR experiments. *J. Mol. Biol.* 425:257–272.

76. Schmidpeter, P. A. M., and F. X. Schmid. 2015. Prolyl isomerization and its catalysis in protein folding and protein function. *J. Mol. Biol.* 427:1609–1631.

77. Burns-Hamuro, L. L., P. M. Dalessio, and I. J. Ropson. 2004. Replacement of proline with valine does not remove an apparent proline isomerization-dependent folding event in CRABP I. *Protein Sci.* 13:1670–1676.

78. Jackson, S. E., and A. R. Fersht. 1991. Folding of chymotrypsin inhibitor 2. 1. Evidence for a two-state transition. *Biochemistry.* 30:10428–10435.

79. Texter, F. L., D. B. Spencer, …, C. R. Matthews. 1992. Intramolecular catalysis of a proline isomerization reaction in the folding of dihydrofolate reductase. *Biochemistry.* 31:5687–5691.

80. Plaxco, K. W., C. Spitzfaden, …, C. M. Dobson. 1996. Rapid refolding of a proline-rich all-beta-sheet fibronectin type III module. *Proc. Natl. Acad. Sci. USA.* 93:10703–10706.

81. Bhattacharjee, R., and J. B. Udgaonkar. 2022. Differentiating between the sequence of structural events on alternative pathways of folding of a heterodimeric protein. *Protein Sci.* 31, e4513.