Biochemistry

Article

Cite This: *Biochemistry* 2018, 57, 4776−4787

pubs.acs.org/biochemistry

Identification and Rationalization of Kinetic Folding Intermediates for a Low-Density Lipoprotein Receptor Ligand-Binding Module

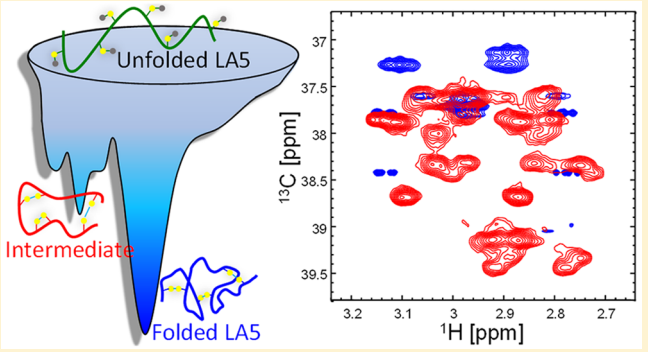
Or Szekely,[†] Gad Armony,[‡] Gregory Lars Olsen,[†] Lavi S. Bigman,[‡] Yaakov Levy,[‡] Deborah Fass,*[,‡] and Lucio Frydman*[,†]

[†]Department of Chemical and Biological Physics, Weizmann Institute of Science, Rehovot 7610001, Israel

[‡]Department of Structural Biology, Weizmann Institute of Science, Rehovot 7610001, Israel

Supporting Information

ABSTRACT: Many mutations that cause familial hypercholesterolemia localize to ligand-binding domain 5 (LA5) of the low-density lipoprotein receptor, motivating investigation of the folding and misfolding of this small, disulfide-rich, calcium-binding domain. LA5 folding is known to involve non-native disulfide isomers, yet these folding intermediates have not been structurally characterized. To provide insight into these intermediates, we used nuclear magnetic resonance (NMR) to follow LA5 folding in real time. We demonstrate that misfolded or partially folded disulfide intermediates are indistinguishable from the unfolded state when focusing on the backbone NMR signals, which provide information on the formation of only the final, native state. However, ^{13}C labeling of cysteine side chains differentiated transient intermediates from the unfolded and native states and reported on disulfide bond formation in real time. The cysteine pairings in a dominant intermediate were identified using ^{13}C-edited three-dimensional NMR, and coarse-grained molecular dynamics simulations were used to investigate the preference of this disulfide set over other non-native arrangements. The transient population of LA5 species with particular non-native cysteine connectivities during folding supports the conclusion that cysteine pairing is not random and that there is a bias toward certain structural ensembles during the folding process, even prior to the binding of calcium.



One of the strengths of nuclear magnetic resonance (NMR) is its ability to illuminate transiently or sparsely populated states of proteins. This ability is expanding because of technological developments in sample preparation,[1,2] instrumentation, and data collection.[3−11] Real-time NMR holds particular promise for monitoring relatively slow biological processes, such as RNA folding[12−14] and peptide bond *cis/trans* isomerization in proteins.[15,16] Another process that may be accessible to real-time NMR is the folding of proteins coupled to the formation of intramolecular covalent cross-links. Specifically, for certain proteins, cysteine amino acids must be paired and oxidized to form disulfide bonds with the proper connectivity, before the folded conformation of the protein becomes the thermodynamically stable state of the system. The folding of these proteins is thus limited by the rates at which these covalent bonds form, typically occurring on multisecond time scales. Transient occupation of non-native intermediate states with alternative cysteine pairing may significantly delay folding. The inherently slower and more readily tunable folding of proteins that rely on cysteine oxidation suggests that these proteins may be suitable for experiments using NMR, to follow kinetic folding reactions in real time. Residue-resolution observation of slow-folding, disulfide-bonded proteins by NMR, may further illuminate sources of ruggedness in protein folding landscapes in general.

This study exploits these methodological opportunities to monitor the folding of a previously characterized disulfide-bonded, calcium-binding protein domain from the low-density lipoprotein receptor (LDLR).[17] The amino-terminal region of LDLR controls the uptake of cholesterol-carrying lipoproteins into cells. It consists of seven cysteine-rich modules of ∼40 amino acids each. Mutations in LDLR, particularly in the fifth ligand-binding domain, LA5, cause familial hypercholesterolemia.[18,19] This condition is characterized by high levels of cholesterol in the blood, which in turn cause premature cardiovascular disease.[20] Disease-causing mutations in LDLR can exert their effects either by directly affecting lipoprotein binding[18] or by undermining domain folding and thus function.[19] A subset of LA5 residues has been shown to be key to folding of the domain,[21] and as long as these are preserved, LA5 folding is achieved at equilibrium.

The kinetics and routes to LDLR folding, in addition to the factors that affect equilibrium stability, are of interest because

Received: April 22, 2018
Revised: July 5, 2018
Published: July 6, 2018

DOI: 10.1021/acs.biochem.8b00466
Biochemistry 2018, 57, 4776−4787

of their relevance to the *in vivo* function of the receptor.[22] The folding kinetics of LA5, including formation of its three disulfide bonds, have been studied previously using reversed-phase high-performance liquid chromatography (rpHPLC). In this batched, discrete method, species with different numbers or connectivities of disulfide bonds were quantified as a function of time after initiating folding, or at equilibrium.[19,23] Such experiments revealed intermediate species containing one or two disulfide bonds, as well as three-disulfide species with non-native cysteine connectivities.[23] The identities of these species and their putative substructures, however, are not known. Insight into the nature of the transient, non-native disulfide-bonded LA5 species and the site-specific changes accompanying LA5 folding, would be an important advance in characterizing LDLR. In this study, we took advantage of fast NMR methods and of the NMR resonance assignments that were available for the LA5 folded state,[24−26] to bridge the gap between the detection of partially folded or misfolded LA5 species by rpHPLC and the structural information available from an in-depth analysis of the fully oxidized protein native state. Kinetics of folding were monitored by changes in NMR chemical shifts and peak intensities as a function of time after initiating cysteine oxidation, and different NMR observables provided different perspectives on structural and dynamics changes occurring during the folding process. Specifically, complementing backbone NMR with NMR experiments monitoring the kinetics of covalent bonding between cysteine side chains, provided unique insight into LA5 folding. A course-grained computational model facilitated the analysis of the NMR results and allowed us to explore the factors that promoted the observed non-native intermediates over other possible structures and disulfide connectivities. Complementing backbone NMR by monitoring the kinetics of covalent bonding between cysteine side chains thus provided unique insight into LA5's folding.

METHODS

Protein Expression and Purification. LA5, spanning amino acid residues 193−232 of human LDLR (UniProt entry P01130), was produced from a DNA sequence codon-optimized for *Escherichia coli* inserted into a pMMHb vector.[27] A His_6 tag and a TrpΔLE tag for driving downstream proteins into inclusion bodies were included and were separated from the LA5 coding sequence by a methionine for cleavage using cyanogen bromide. The recombinant vector was transformed into *E. coli* strain BL21 (DE3). Cells were grown at 37 °C in LB to an optical density of 0.5−0.6 at 600 nm. Isopropyl 1-thio-β-D-galactopyranoside (IPTG) was added to a final concentration of 0.5 mM, and the cultures were grown overnight at 37 °C. Uniformly ^{15}N-labeled LA5 was grown in M9 medium with $^{15}NH_4Cl$ as the sole nitrogen source. For labeling LA5 with ^{13}C at the C_β cysteine atoms, the vector was transformed into cysteine auxotroph *E. coli*[28,29] and grown in rich defined medium supplemented with [$^{13}C_\beta$]cysteine.[30] Cells were harvested and resuspended in 20 mM Tris (pH 8) and 1 mM ethylenediaminetetraacetic acid (EDTA), sonicated on ice, and centrifuged for 1 h at 40000*g*. Inclusion bodies from the pellet were washed three times with 20 mM Tris (pH 8), 1 mM EDTA, and 0.1% Triton X-100 and then once with 20 mM Tris (pH 8) and 1 mM EDTA. The washed inclusion bodies were dissolved overnight in 50 mM Tris (pH 8), 6 M guanidine, and 30 mM 2,2′-dithiodiethanol. Dissolved inclusion bodies were applied to a Ni-NTA column pre-equilibrated with 20 mM Tris (pH 8) and 6 M guanidine, and the LA5 construct was eluted with 0.2% acetic acid and 6 M guanidine. LA5 was cleaved from the TrpΔLE tag at the methionine by a 100-fold molar excess of cyanogen bromide in the Ni-NTA elution fraction supplemented with 0.2 M HCl overnight in the dark at room temperature. The cleavage reaction mixture was dialyzed against 5% acetic acid to remove salts and then lyophilized to dryness. LA5 was extracted by repeatedly resuspending the dry proteins in NH_4/AcOH (pH 9), centrifuging, and collecting the supernatants; the supernatants were then lyophilized to dryness. LA5 was then reduced with excess dithiothreitol (DTT) and subjected to rpHPLC, using a C-18 column and eluting with a gradient of acetonitrile (from 25 to 34% in 26 min) in water with 0.1% trifluoroacetic acid (TFA). The fractions containing reduced unfolded LA5 were lyophilized.

Monitoring LA5 Folding Kinetics by rpHPLC. The LA5 folding reaction was monitored by rpHPLC to optimize the conditions for the dynamic NMR experiments. Unfolded, reduced LA5 was dissolved in 50 mM Tris (pH 7.4 or 8.4) with excess $CaCl_2$, and then the folding reaction was started by addition of oxidized glutathione (GSSG). The reaction was quenched by adding TFA to a final concentration of 3%, the mixture was applied to a C-18 column and eluted with a gradient of acetonitrile (from 15 to 35% over 50 min) in water with 0.1% TFA. LA5 folding intermediates were isolated by folding LA5 with GSSG as described above but with EDTA instead of Ca^{2+}. The folding reaction was allowed to continue for 10 min before being quenched with TFA. The folding intermediates were purified from the quenched reaction mixture by rpHPLC using a shallow gradient of acetonitrile (from 20.5 to 23% over 60 min or from 20.5 to 24.5% over 100 min) in water with 0.1% TFA.

Mass Spectrometry. LA5 isolated intermediate samples were lyophilized to dryness and reconstituted in 50% organic solvent (Acetonitrile + TFA or Methanol) before mass spectrometry (MS) analysis. The LA5 folding reaction was quenched with excess N-ethylmaleimide (NEM); the mixture was acidified with HCl, and salts were removed by dialysis against water. Samples were diluted in methanol to the appropriate concentration before MS analysis. The samples were infused at a rate of 10 μL/min to a Waters SQD2 instrument equipped with an ESI source. Mass spectra were collected in positive mode with a capillary voltage of 3.50 kV and a cone voltage of 40 V.

NMR Spectroscopy. NMR experiments were conducted using 5 mm "inverse" NMR probes. The 5 mm probes included a liquid nitrogen-cooled Prodigy probe in a 14.1 T Bruker magnet interfaced to a Bruker Avance III console, a cryogenically cooled probe in a 18.8 T Bruker magnet interfaced to a Bruker Avance III console, and a room-temperature HCN probe in a 14.1 T Oxford magnet interfaced to a Varian VNMRS console. For static NMR measurements (Figure 2), folded oxidized ^{15}N-labeled LA5 was dissolved in 10 mM Tris buffer (pH 7.4) in the presence of 1 mM $CaCl_2$[25,31] and 10% D_2O. For dynamic NMR experiments, unfolded, reduced, labeled LA5 was dissolved in 50 mM Tris (either protonated or deuterated, pH 8.4) with 6 mM $CaCl_2$ and either 10 or 100% D_2O for ^{15}N- or ^{13}C-based experiments, respectively. The folding reaction was initiated by mixing the LA5 solution with GSSG in an NMR tube, immediately inserting it into the spectrometer, and collecting sequential NMR spectra: one-dimensional (1D) ^1H (^{15}N-filtered) or two-

DOI: 10.1021/acs.biochem.8b00466
Biochemistry 2018, 57, 4776−4787

dimensional (2D) $^{1}H-^{15}N$ for ^{15}N-labeled LA5 and 1D ^{13}C or 2D $^{1}H-^{13}C$ for $^{13}C_{\beta}$-Cys-labeled LA5. The mixing was done manually or using an automated injection system. This system employs a two-state valve[32,33] to control the filling of the NMR tube and a three-port accessory involving both forward and backward gas pressures as described,[34] and is controlled by Arduino-based software.[35] To perform as expected, both these gas inlets must exceed normal atmospheric pressure and operate under a programmable unit controlling events in real time, and obtaining feedback from an optical sensor. Unless otherwise indicated, all measurements were taken at 20 °C. Additional experimental parameters for NMR acquisitions can be found in the figure captions and in the Supporting Information.

Computational Methods. The formation of disulfide bonds among the six Cys residues of LA5 was studied using a coarse-grained molecular dynamics (MD) model, in which each residue was represented by two beads. The backbone atoms of each residue were represented by a single bead placed at the position of each C_{α} atom (Protein Data Bank entry 1AJJ), and side chains were represented by another bead at the position of each C_{β} atom, which was bonded to the backbone bead. Disulfide bonds were modeled as Lennard-Jones (LJ) interactions with an optimal distance of 3.8 Å between two Cys C_{β} atoms. In addition to the disulfide bonds, electrostatic and hydrophobic interactions were incorporated in some cases. Electrostatic interactions between charged residues were modeled using the Debye−Hückel potential,[36] and the beads representing side chains of positively (Arg and Lys) and negatively (Asp and Glu) charged amino acids were assigned positive and negative unit charges, respectively. Hydrophobic interactions between aromatic residues (Phe and Trp) were modeled using the LJ potential (with an optimal distance of 6 Å). When included, binding of Ca^{2+} was modeled by introducing, using harmonic potentials, 15 distance constraints between the residues that participate in Ca^{2+} binding. The simulations were performed in a modular fashion whereby disulfide bonds, electrostatic interactions, hydrophobic interactions, and Ca^{2+} binding, were added independently. The polymeric properties of the protein were controlled by three bond-related input parameters: bond lengths, bond angles, and dihedral angles. The parameters of these energies were adopted from similar models used to study protein folding.[37] We note that in contrast to topology-based models, the current model does not have any bias toward the native structure of LA5. Accordingly, no native contacts apart from Ca^{2+} binding were included in the model. Also worth noting is the fact that although all-atom MD simulations would in principle be more realistic, they would be computationally too expensive to allow us to decipher the role of the individual factors in the formation of various structural intermediates.

The six cysteines in LA5 can form 15 [i.e., $6!/(2^3 \times 3!)$] different combinations of three disulfide bonds. In this study, we examined independently seven combinations that spanned the full entropic range of the system; additional combinations appeared unlikely to shed further light and hence are not discussed here. The molecular motions of the systems with different disulfide combinations were simulated by Langevin dynamics with a friction coefficient γ of 0.01, as described previously.[36] The temperature in all the simulations was set such that $k_B T = 0.25$. Each simulation included at least 5×10^7 elementary steps. The simulation trajectories yielded conformations with different extents to which the predefined disulfide bonds were present, defined by a distance of <4.5 Å between the two corresponding C_{β} atoms. The probability for the formation of a specific combination of three disulfide bonds was calculated according to $P_{s-s} = N_{ss}/N_{tot}$, where N_{ss} is the number of conformations in which all three disulfide bonds were formed and N_{tot} is the total number of sampled conformations in the simulation.

RESULTS

Driving Oxidative Folding under Suitable NMR Conditions. A favorable feature for studying LA5 folding kinetics by NMR is the high solubility of the unfolded state, an outcome of the paucity of hydrophobic residues and the prevalence of aspartates for Ca^{2+} binding. This feature allowed us to study the folding process at 2−5 mM protein concentrations. However, these concentrations are incompatible with an oxidative protein folding process driven by dissolved oxygen, which at saturation is present in aqueous solutions at approximately 280 μM. By comparison, forming three disulfide bonds for a 5 mM protein would require electrons at a 30 mM concentration to be transferred from the cysteine thiols to an electron acceptor. Enzyme catalysis of disulfide bond formation does not solve this problem, as sulfhydryl oxidases also use molecular oxygen as an electron acceptor. Introduction of additional oxygen into the NMR tube during measurement was not practical because of bubble formation. Therefore, alternative electron acceptors were required to drive disulfide bond formation and complete LA5 folding. Oxidized glutathione (GSSG) was chosen because of its solubility and because of its presence in the endoplasmic reticulum during the natural folding of LDLR. Complete oxidation of millimolar concentrations of LA5 could be driven by GSSG, as evidenced by rpHPLC (Figure 1). Good folding yields and rates were achieved with 20 mM GSSG, excess $CaCl_2$, and a pH 8.4 buffer. A temperature of 20 °C was chosen as it provided the best NMR signal. In agreement with previous results collected at lower protein concentrations under similar solution conditions,[23] most of the

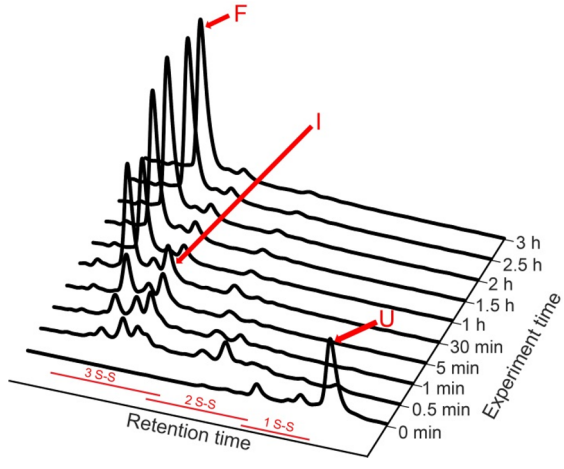


Figure 1. Successive rpHPLC traces recorded during oxidative folding of LA5. The reaction was performed with 1 mM LA5 in Tris-HCl buffer (pH 8.4) and 20 mM $CaCl_2$. "U" and "F" correspond to unfolded and folded LA5, respectively. "I" indicates peaks corresponding to species with non-native disulfide connectivities. Regions of the rpHPLC traces containing species with one (1 S−S), two (2 S−S), and three (3 S−S) disulfides are indicated.

DOI: 10.1021/acs.biochem.8b00466
Biochemistry 2018, 57, 4776−4787

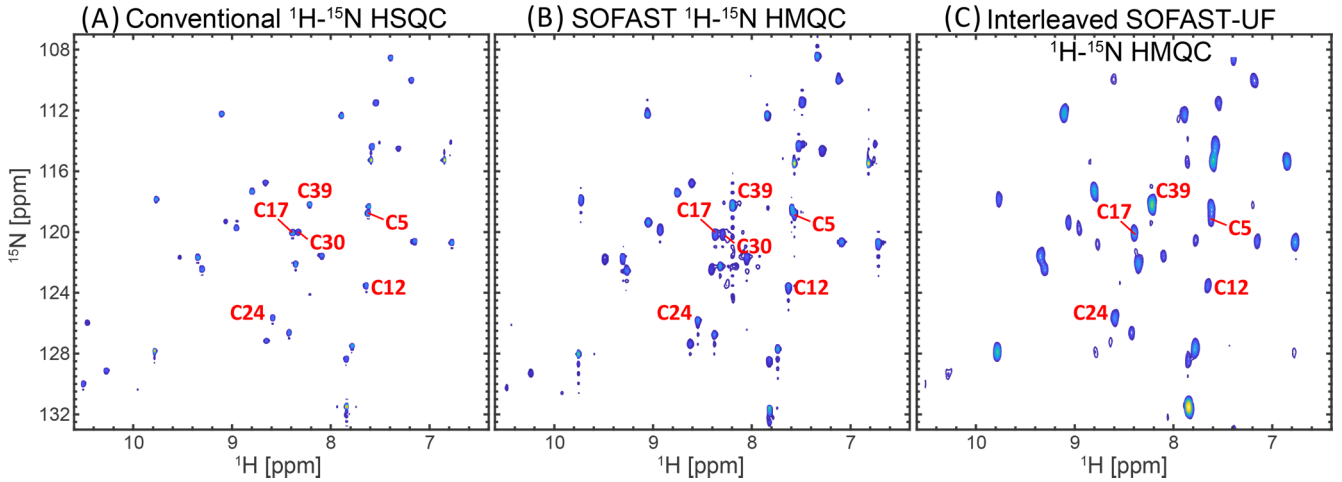


Figure 2. ^{1}H−^{15}N amide correlation maps measured on ^{15}N-labeled LA5 on a 600 MHz Varian NMR spectrometer. (A) Conventional 2D ^{1}H−^{15}N HSQC spectrum. The spectrum was collected using four scans (N_S), 128 hypercomplex[40] t_1 increments, an acquisition time of 262 ms, and a relaxation delay (d_1) of 2 s. The total experimental time was 40 min. (B) 2D SOFAST HMQC spectrum. The flip angle was 120° for optimal sensitivity, and the spectrum was collected using two scans for each of the 64 hypercomplex t_1 increments, an acquisition time of 85.2 ms, and a relaxation delay (d_1) of 0.5 s. The total experimental time was 155 s. (C) Phase-cycled, four-step interleaved ultraSOFAST ^{1}H−^{15}N HMQC[39] spectrum (pulse sequence shown in Figure S1): eight scans total, a d_1 of 0.7 s, and an experimental time of 6 s. Measurements were taken at 25 °C (HSQC, UltraSOFAST HMQC) or 20 °C (SOFAST HMQC). Assignment of cysteine residues (labeled in red) was done on the basis of the work of Koduri et al.[25] and was assisted by additional ^{15}N-edited TOCSY and NOESY experiments.

LA5 reached the native state within 1.5 hours, and partial or non-native disulfide-bonded species were predominant only within the first few minutes of folding (Figure 1).

Optimizing Acquisition Schemes and Pulse Sequences. Methods recently developed for real-time NMR[3−5,38,39] were adapted for the study of LA5 folding. These methods were tested on folded LA5 to determine which was most appropriate, considering the folding time frame and available signal-to-noise ratio (SNR). Figure 2 presents the full 2D ^{1}H−^{15}N correlation spectrum of ^{15}N-labeled LA5, measured using different acquisition schemes and pulse sequences. Figure 2A shows a conventional 2D heteronuclear single-quantum correlation (HSQC) spectrum recorded in ∼40 min. Figure 2B shows a SOFAST 2D heteronuclear multiple-quantum correlation (HMQC) spectrum, with a total measurement time of ∼2 min.[4,5] Because selective pulses are used in this experiment, a faster sampling of the indirect domain data and greater averaging of scans per unit time become feasible. Figure 2C shows a 2D HMQC spectrum of LA5 recorded using the ultrafast NMR technique in combination with the SOFAST effect,[3,39] demonstrating that a 2D spectrum of LA5 can also be obtained with full spectral widths in 6 s. Ultrafast NMR could have enabled us to monitor folding of the LA5 kinetically with a high time resolution. However, because the domain folds on a time scale of minutes to hours under the solution conditions selected, we prioritized SNR over speed and monitored the folding by means of 1D NMR and 2D SOFAST HMQC.

Folding Intermediates Are Invisible in Amide-Based Heteronuclear NMR. In general, the spectra of folded proteins contain sharp, well-dispersed peaks, whereas spectra of unfolded proteins show substantial signal overlap and peak broadening due to conformational flexibility. NMR spectral dispersion can thus be used as an indicator of protein folding. Utilizing a 1D version of ^{15}N-edited SOFAST HMQC that focused on the amide spectral region, folding was monitored with good SNR in real time after injecting GSSG into a solution of unfolded LA5 (Figure 3). Over the course of the experiment, the intensities of sharp peaks were observed to

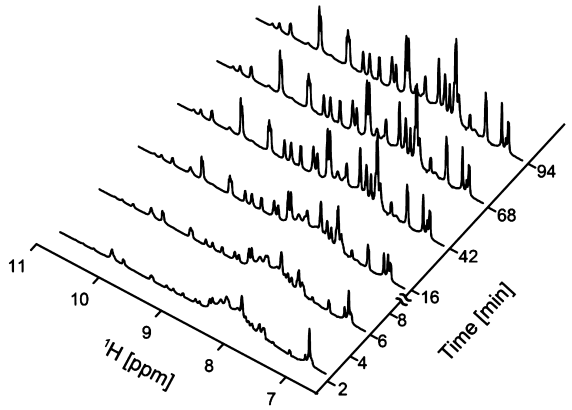


Figure 3. Time course for a 1D ^{15}N-edited HMQC, measured on ^{15}N-labeled LA5 (pH 8.4). A spectral width of 9014.4 Hz was covered, and the flip angle was 120° for optimal sensitivity. Sixteen scans were collected using an acquisition time of 454 ms, a relaxation delay of 0.7 s, and an experimental time of ∼16 s per spectrum. Oxidative folding of ∼3 mM LA5 was initiated by injecting 20 mM GSSG into a buffered solution of reduced, unfolded LA5 [50 mM Tris buffer (pH 8.4) and 10% D_2O] in the presence of Ca^{2+} ions. Measurements were taken at 20 °C.

increase while the intensities of broad peaks decreased. However, no peaks were seen to appear and then decay. The monotonic peak behavior observed in these experiments thus provided no supporting evidence for the presence of folding intermediates.

To determine whether the lack of detection of folding intermediates in the 1D SOFAST HMQC experiment was due to limited spectral resolution, kinetic experiments were repeated using 2D SOFAST ^{1}H−^{15}N HMQC NMR. Folding was initiated by injection of GSSG and followed by collecting SOFAST spectra approximately every 4 min (Figure 4A−C).

DOI: 10.1021/acs.biochem.8b00466
Biochemistry 2018, 57, 4776−4787

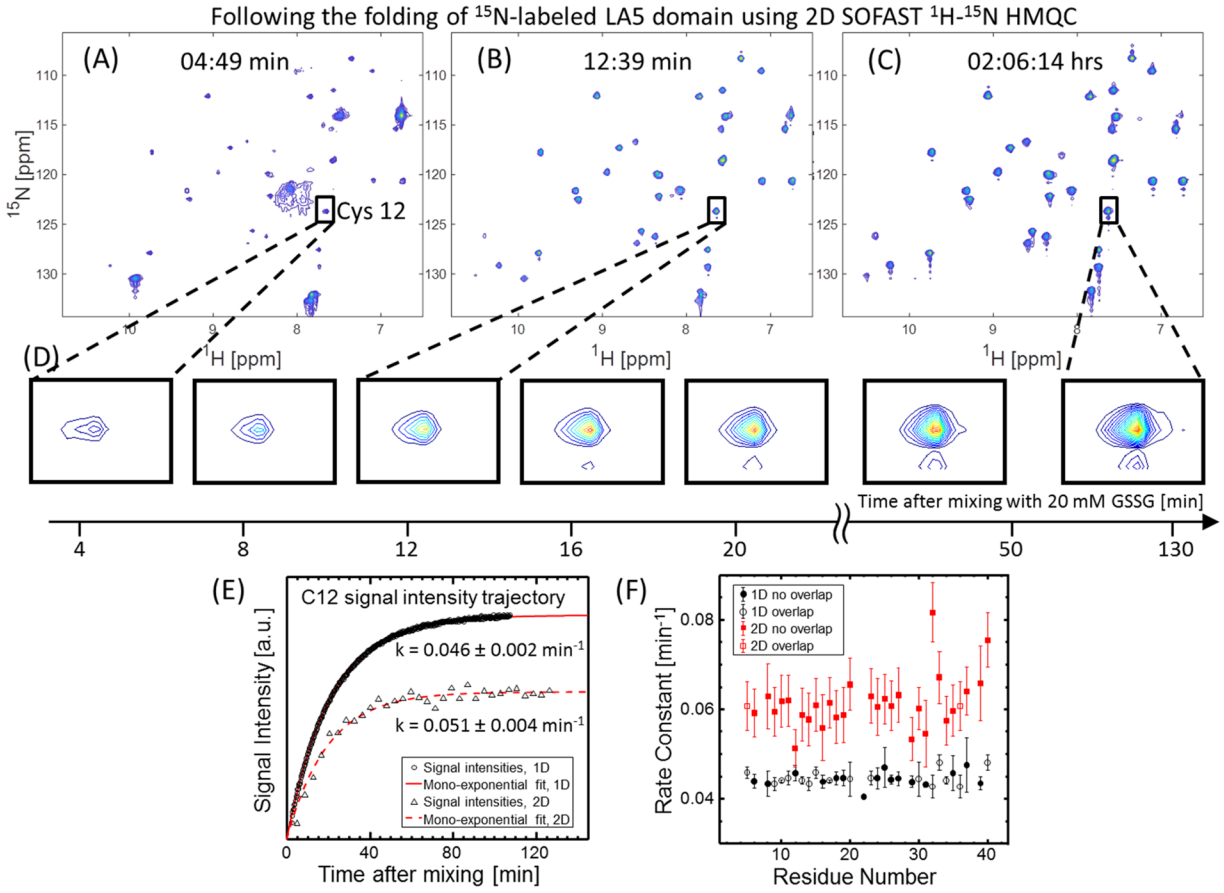


Figure 4. Dynamics of LA5 folding, followed by a series of 2D SOFAST $^{1}H-^{15}N$ HMQC spectra. Oxidative folding of 2.9 mM ^{15}N-labeled LA5 (pH 8.4) was initiated by injecting 20 mM GSSG into a buffered solution of unfolded LA5 [50 mM Tris buffer (pH 8.4) and 10% D_2O] in the presence of 10 mM Ca^{2+} ions. (A–C) Selected spectra recorded along the folding time course, with times after GSSG injection indicated. The flip angle was 120° for optimal sensitivity, and the spectra were collected using two scans for each of the 64 hypercomplex t_1 increments, an acquisition time of 85.2 ms, a relaxation delay of 0.82 s, and an experimental time of 237 s per spectrum. Measurements were taken at 20 °C. (D) Selected spectral region containing one of the peaks (corresponding to Cys12) and showing its intensity buildup as a function of time. (E) Cys12 peak intensities and monoexponential fits to these data, extracted from the 1D spectra of Figure 3 (empty circles, red curve) or from the 2D spectra (empty triangles, red dashed curve). These fits provide an apparent rate constant k_{C12} of 0.046 ± 0.002 min^{-1} (1D) or 0.051 ± 0.004 min^{-1} (2D). (F) Folding rate constants for the amides in the LA5 sequence. Red squares represent the rate constants extracted from a 2D SOFAST $^{1}H-^{15}N$ HMQC experiment, while black circles represent the rate constants extracted from the 1D dynamic experiment (filled symbols, clearly resolved peaks in the spectrum; empty symbols, data obtained from partially overlapping peaks).

In the first spectrum, recorded 5 min after injection, we observed a small population of protein that appeared to be fully folded according to good peak correspondence with the final spectrum (Figure 4C). Once again, only peaks belonging to the folded conformation of LA5 or resonating within the poorly resolved region corresponding to unfolded protein were detected throughout the folding time course; no other transient peaks with distinct chemical shifts, which could potentially reflect folding intermediates, were observed.

Panels E and F of Figure 4 summarize the kinetic folding studies of LA5 by 1D and 2D $^{1}H-^{15}N$ NMR. The protein appears to fold uniformly, showing a monoexponential buildup of the folded-state signature for cysteine (e.g., Cys12 in Figure 4E) and other residues as a function of time post-injection. The folding rate constants appear to be slightly, systematically lower in the 1D experiment than in the 2D counterpart, but most of the amide peaks build up with comparable monoexponential rate constants within each experiment (Figure 4F). Multiexponential increases in peak intensities corresponding to the folded state, which might have indirectly indicated the presence of folding intermediates, were not detected, and non-native peaks arising from such intermediates were not directly apparent in the spectra.

The apparent discrepancy between the observation of kinetic folding intermediates by rpHPLC (Figure 1; also ref 23) and the lack of evidence for intermediates in amide-based $^{15}N-^{1}H$ NMR experiments may reflect the nature of these intermediates. Intermediate conformations with non-native disulfide bonds might be flexible, and their $^{1}H-^{15}N$ correlation maps may not be sensitive enough to observe them. Furthermore, solvent-exposed amide protons in such species would exchange rapidly with H_2O, possibly broadening their resonances beyond detection. If that is the case, the folding process would appear monoexponential by either 1D or 2D $^{1}H-^{15}N$ NMR. In essence, these NMR experiments would report solely on the final step in the folding process, despite the actual existence of intermediate steps and species.

^{13}C-Based NMR Experiments Reveal Folding Intermediates. In contrast to the $^{1}H-^{15}N$ acquisition schemes, ^{13}C NMR is free from some of the aforementioned broadening effects. Structural disorder may still contribute to the broadening of side chain resonances, but the protons bonded

DOI: 10.1021/acs.biochem.8b00466
Biochemistry 2018, 57, 4776–4787

to the C_α or side chain C_β atoms will not exchange with the solvent even in an open flexible conformation of the polypeptide chain. We therefore turned to the ^{13}C NMR of cysteine side chains, to probe for folding intermediates. The chemical shifts of $^{13}C_\beta$ in cysteine residues are particularly sensitive to redox state and to disulfide bond connectivity.[41] The chemical shift difference between reduced and disulfide-bonded cysteines is substantial, ~15 ppm, so free and bonded cysteines can be readily distinguished by ^{13}C NMR.[41−46] The $^{13}C_\beta$ chemical shifts of reduced cysteines in unfolded LA5 appear in the region of 25−26.5 ppm, whereas the $^{13}C_\beta$ chemical shifts of cysteines in disulfides in folded LA5 appear in the region of 33−40 ppm (Figure 5A).

After reduced, unfolded, $^{13}C_\beta$ cysteine-labeled LA5 had been mixed with GSSG, 1D ^{13}C NMR showed the buildup of peaks present in the folded conformation of LA5 and a set of peaks that appeared and decayed (red asterisks in Figure 5B). This behavior clearly points to a folding intermediate. Given the location of these peaks in the region corresponding to disulfide bonds and the fact that they are distinct from peaks corresponding to native pairings, they represent non-native disulfides that subsequently rearrange. Because of the long time required to collect the first spectrum in this experiment (~7:30 min), only the decay and not the buildup of these peaks was observed. Nevertheless, these peaks were absent from the spectrum of the unfolded state, so they must have appeared during the experimental dead time. To obtain information about the formation of the non-native disulfide species, faster mixing and a more sensitive acquisition scheme were required.

The disulfide intermediates were observed with better spectral and temporal resolution by acquiring more sensitive 2D $^1H-^{13}C$ NMR data. A series of 2D $^1H-^{13}C$ HSQC spectra arising from the labeled cysteine residues were collected after GSSG was injected using an automatic, pressure-driven, two-state valve system. The decrease in the mixing dead time to a few seconds and the added ^1H sensitivity enabled detection of the buildup and decay of disulfide intermediates. The first post-injection spectrum already showed intermediate peaks but at lower intensities than in later spectra (Figure 6). Although the sensitivity and the time resolution of this experiment could provide only approximate rates for the formation and decay of these intermediates, the transition from the unfolded state to the intermediate was on the order of 0.7 min^{-1}, while the transition from the intermediate to the folded state occurred on the order of 0.05 min^{-1}. The latter rate is similar to the monoexponential buildup rates arising from Figures 3 and 4, on the basis of amide group NMR experiments.

Identifying the Disulfide Bond Connectivity of the LA5 Folding Intermediate. The intermediate species observed in these kinetics experiments was transient and not highly populated, so its structure could not be fully characterized in real time. To obtain sufficient amounts of the major folding intermediate to identify its disulfide bond connectivity and additional properties, ^{13}C-labeled LA5 was oxidized in the presence of EDTA instead of Ca^{2+}.[19] Oxidation was initiated using GSSG and allowed to continue for 10 min before quenching with 3% TFA, and the resulting species were then separated to the extent possible using rpHPLC. The purified intermediate was stable in the absence of an added reducing agent, enabling its characterization. The purified intermediate was a mixture of species; but one, containing three disulfide bonds, was predominant (Figure S3). This major species had the same rpHPLC retention time as the kinetic intermediate observed during folding in the presence of Ca^{2+} (not shown). Figure 7 compares the 2D $^1H-^{13}C$ HSQC spectrum of this intermediate with the spectrum measured 6 min post-mixing in the kinetic experiment (Figure 6). The

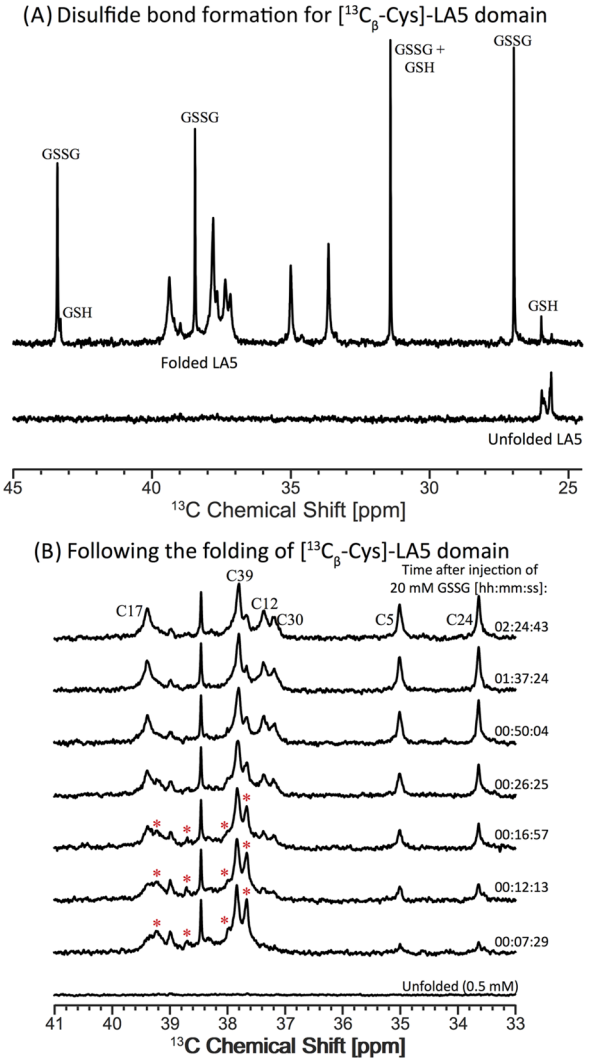


Figure 5. LA5 folding intermediates revealed by 1D ^{13}C NMR. (A) 1D spectra of unfolded and folded $^{13}C_\beta$-Cys-labeled LA5. Measurement of 1D ^{13}C spectra was aided by magnetization transfers from ^1H atoms using a refocused INEPT sequence[47] and by applying ^1H decoupling during data acquisition. The sample was 0.5 mM labeled protein in 50 mM Tris buffer (pH 8.4) and 10% D_2O in the presence of 6 mM Ca^{2+}. Data were collected immediately after dissolution of unfolded LA5 in this solution and after initiating folding with 20 mM GSSG. The labeled sharp peaks arise from the indicated (natural-abundance but high-concentration) glutathione species. For unfolded LA5 (bottom trace), 16 scans were collected using an acquisition time of 271.4 ms, a relaxation delay of 2 s, and an experimental time of 55 s. For folded LA5 (top trace), 2048 scans were collected using an acquisition time of 799.7 ms, a relaxation delay of 2 s, and an experimental time of 97 min. (B) Close-ups of selected 1D ^{13}C spectra recorded along the folding time course of 1.36 mM $^{13}C_\beta$-Cys-labeled LA5, initiated by injection of 20 mM GSSG. One hundred scans were collected using an acquisition time of 799.7 ms, a relaxation delay of 2 s, and an experimental time of 284 s per spectrum. The assignment[26] of the cysteine residues in the folded conformation is indicated on the topmost spectrum. Folding intermediate peaks are indicated by red asterisks in the bottom three traces. All measurements were taken at 20 °C.

DOI: 10.1021/acs.biochem.8b00466
Biochemistry 2018, 57, 4776−4787

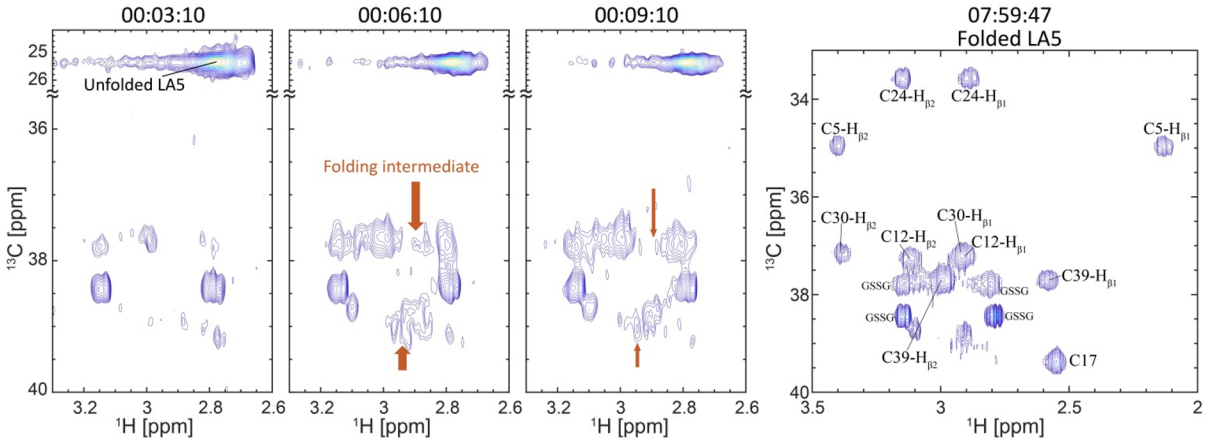


Figure 6. Folding intermediates for the $^{13}C_{\beta}$-Cys-labeled LA5 revealed using 2D $^{1}H-^{13}C$ NMR. 2D $^{1}H-^{13}C$ HSQC spectra tracked the folding of ~1.6 mM $^{13}C_{\beta}$-Cys-labeled LA5 in 50 mM Tris-d_{11} buffer (pH 8.4) (100% D_2O) and 6 mM Ca^{2+}, initiated by injection of GSSG to a concentration of 4.7 mM. The spectra were collected on a Bruker 600 MHz NMR spectrometer at 20 °C, using two phase-cycled scans for each of the 64 t_1 hypercomplex increments, an acquisition time of 177.6 ms, a relaxation delay of 0.5 s, and an experimental time of 3 min per spectrum. Selected spectra are shown, with times after GSSG injection indicated as hh:mm:ss. Additional spectra at other time points are displayed in Figure S2. Selected spectral regions of early time points are shown for the sake of clarity, highlighting peaks corresponding to the folding intermediate; a broader region is shown for the final time point to encompass native-state peaks (assigned). Positions of peaks corresponding to the folding intermediate are indicated by orange arrows, with arrow size reflecting peak intensity.

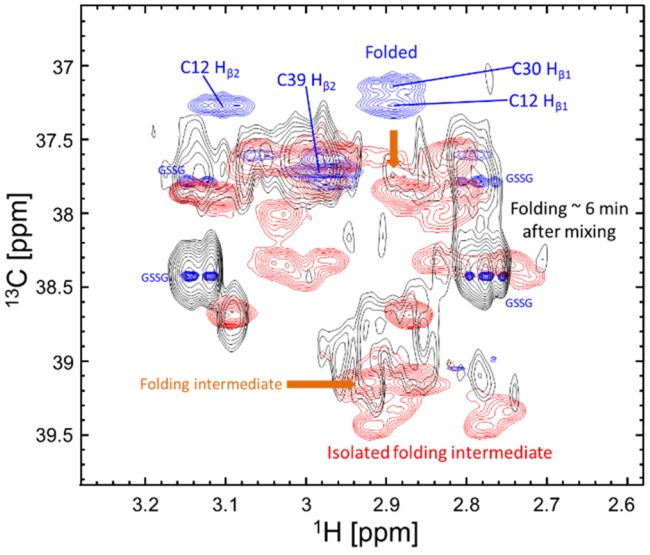


Figure 7. Comparison between 2D $^{1}H-^{13}C$ HSQC spectra of the isolated (red) and kinetic (black; ~6 min after initiating folding) LA5 intermediates. Also shown are peaks in this spectral region from folded LA5 (blue). The folded LA5 spectrum was collected on a 600 MHz NMR spectrometer and TCI probe at 20 °C. The kinetic intermediate was recorded as indicated in Figure 6. The isolated folding intermediate spectrum was collected[31] a 0.61 mM $^{13}C_{\beta}$-Cys-labeled LA5 solution in 50 mM Tris-d_6 buffer (pH 8.4) (100% D_2O). Its spectrum was collected on an 800 MHz NMR spectrometer and TCI cryoprobe at 25 °C using eight scans 64 hypercomplex t_1 increments, an acquisition time of 127.8 ms, a relaxation delay of 2 s, and an experimental time of 2246 s. GSSG peaks, present in both the folded LA5 solution and in the kinetic experiment, are labeled. Orange arrows show peak correspondence between the isolated (red) and kinetic (black) folding intermediates. Though the isolated intermediate showed some additional peaks that were not seen in the kinetic experiments, these were present at lower intensities and did not compromise the analysis.

major peaks in the purified intermediate correspond well to those from the folding intermediate generated in real time (Figure 7). In light of this correspondence, the disulfide bonding pattern of the purified intermediate at equilibrium was determined by NMR to shed light on the nature of the species observed transiently during folding.

To determine the disulfide connectivities, ^{13}C-edited nuclear Overhauser effect spectroscopy (NOESY) experiments were performed on the isolated intermediate. In the 2D version of such an experiment, magnetization is transferred through space between nearby protons; however, because their starting points are filtered through the $^{13}C_{\beta}$ nuclei in the six labeled residues of LA5, the ensuing spectra will be considerably simpler than the crowded NOESY spectrum one would measure either on a uniformly labeled protein or without ^{13}C editing. Figure S4 shows results of this 2D experiment for folded LA5. The dominant cross peaks are in all cases those between the inequivalent, prochiral H_{β} protons bonded to the ^{13}C-labeled sites; strong cross peaks between the cysteine H_{β} protons involved in the native disulfide bonds,[31] i.e., C5−C17, C24−C39, and C12−C30, were also observed. Cross peaks to other protons also appeared in the spectrum (Figure S4B), because these protons are in the proximity of the H_{β} of the labeled cysteines.

When the same strategy was used to determine the disulfide bond connectivity in the folding intermediate, the resolution of the 2D ^{13}C-edited NOESY was not sufficient to distinguish between the different diagonal and cross peaks (Figure S4C). Complicating the situation was the aforementioned fact that the isolated intermediate solution contained more than one species, leading to further signal overlap onto the relatively weak $^{1}H-^{1}H$ NOESY peaks. To improve the resolution, full three-dimensional (3D) $^{1}H-^{13}C$ NOESY-HSQC experiments, yielding 2D NOESY planes for every ^{13}C chemical shift were performed. The pulse sequence for this experiment is shown in Figure S5. Additional experiments were performed using non-uniform sampling (NUS) in the two indirect dimensions, whereby a higher resolution was obtained by sampling only 25% of the conventional data.[10,11] Figure 8A shows selected $^{1}H-^{1}H$ planes arising from these experiments, with the corresponding ^{13}C chemical shifts indicated. With this

DOI: 10.1021/acs.biochem.8b00466
Biochemistry 2018, 57, 4776−4787

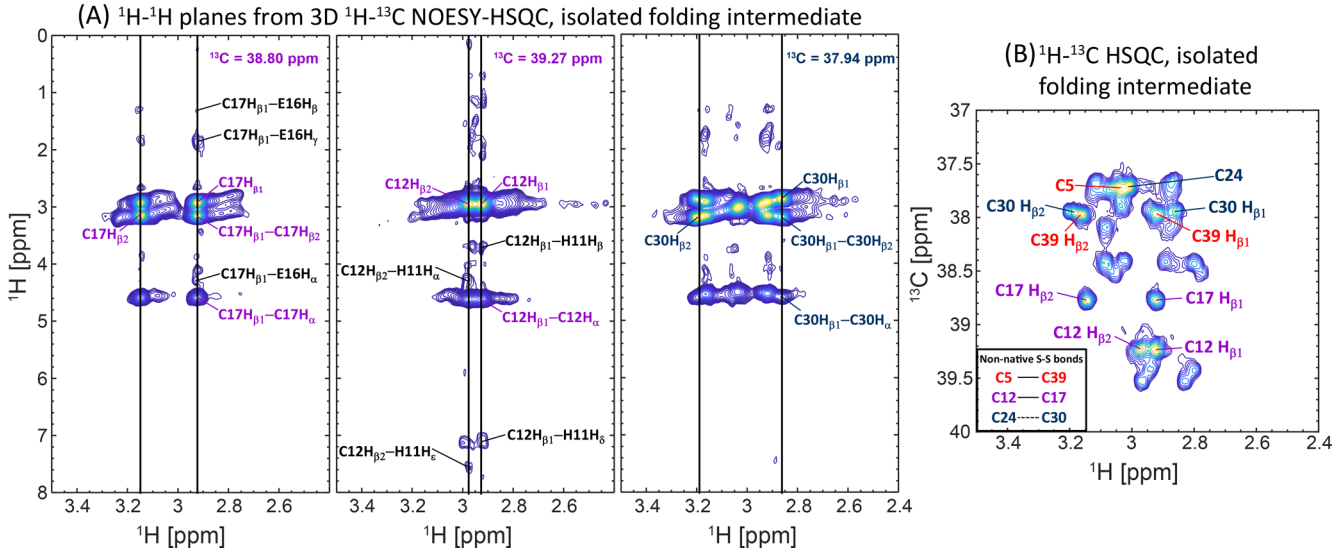


Figure 8. Characterization and assignment of the isolated folding intermediate by 3D NMR. (A) ^{1}H−^{1}H planes from 3D NOESY-HSQC (regularly sampled) of the isolated LA5 folding intermediate [0.61 mM $^{13}C_{\beta}$-Cys-labeled LA5 in 50 mM Tris-d_6 buffer (pH 8.4) and 100% D_2O]. The corresponding ^{13}C chemical shift is indicated in each panel. Black lines indicate the chemical shifts of Cys H_{β} atoms. This pattern of cross peaks with neighboring residues was used to assign the peaks to their corresponding cysteine residues. The 3D spectrum was collected on an 800 MHz NMR spectrometer and cryoprobe at 25 °C using 16 scans 65 hypercomplex t_1 increments for the ^{1}H dimension using STATES acquisition,[40] 48 hypercomplex t_2 increments for the ^{13}C dimension using an Echo-Antiecho acquisition,[48] an acquisition time of 127.8 ms, and a relaxation delay of 1.5 s. (B) 2D ^{1}H−^{13}C HSQC spectrum of the isolated LA5 folding intermediate. The spectrum was recorded as specified in Figure 7 (red). The indicated assignment of the major component was determined by 3D NOESY-HSQCs measured regularly and non-uniformly sampled, as described in the text. The inset shows the established disulfide connectivity of the intermediate, with the solid lines indicating directly observed disulfides and the dashed line the remaining disulfide connectivity.

information available, the peaks of cysteine residues could be individually assigned according to the pattern of cross peaks displayed by their neighboring amino acids in the sequence (Figure 8A). Following assignment, disulfide bond connectivity was determined by identifying cross peaks between the H_{β} protons of the bonded cysteines. These cross peaks were obtained from a 3D ^{1}H−^{13}C HSQC-NOESY experiment (Figure S6), which showed greater resolution of the relevant peaks but overall lower SNR compared to that of the NOESY-HSQC. Through this procedure, the connectivity indicated in the inset of Figure 8B was established. Cross peaks identifying the C5−C39 and C12−C17 linkages were directly observed, while cross peaks for the third linkage, C24−C30, could not be resolved. We assumed the existence of the C24−C30 disulfide bond given the *a priori* knowledge that there are three disulfide bonds in the major folding intermediate (Figure S3). Therefore, we conclude that the major three-disulfide kinetic intermediate in the real-time experiments contained the C5−C39/C12−C17/C24−C30 disulfide pattern.

Computational Insight into the Transient Disulfide Isomers of LA5. Previous studies of LA5 showed the existence of kinetic folding intermediates[23] but did not reveal their connectivities. The demonstration and characterization here of a specific, non-native kinetic intermediate formed during LA5 folding provide an opportunity to investigate the factors that lead to transient formation of this particular species over others. To assess these factors, an LA5 computational model was subjected to a simplified MD simulation comparing the stability of the C5−C39/C12−C17/C24−C30 connectivity versus that of selected disulfide pattern alternatives (Figure 9A). To evaluate the likelihood of these various forms, restraints corresponding to the different disulfide pairings were imposed, and the extents to which the protein backbone sampled conformations consistent with these restraints were converted to probabilities; high probabilities indicated compatibility of the polypeptide chain with the given disulfide pattern. When only bond energy terms including disulfide (SS) interactions were considered, the kinetic intermediate is slightly more likely to occur than the final, native conformation (Figure 9B, blue). Moreover, connectivities different from the one corresponding to the kinetic intermediate, i.e., variants 3 and 4 in Figure 9, showed even higher probabilities. This indicates that the experimentally observed transient intermediate is not preferred based only on LA5 backbone dynamics or on the entropy costs for closing disulfide loops. However, when electrostatic interactions (Es) and interactions between hydrophobic (HP) side chains were incorporated into the model (Figure 9B, red), the probability of encountering the C5−C39/C12−C17/C24−C30 connectivity increased, while the probability of retaining other connectivities decreased. Furthermore, when Ca^{2+} was added to the model by imposing distance restraints between the Ca^{2+}-coordinating residues, the probability of finding the peptide in what ends up being its native disulfide connectivity increased sharply, while the non-native species became strongly disfavored (Figure 9C). These observations suggest that the disulfide connectivity of the dominant three-disulfide intermediate is favored by weak side chain interactions, which are a dominating factor while the Ca^{2+} is not yet bound and the polypeptide chain is still in a flexible, dynamic state. However, once the Ca^{2+} ion encounters its binding site, it strongly promotes a shift of the equilibrium toward the native configuration of LA5, which finally dominates the thermodynamic landscape.

DOI: 10.1021/acs.biochem.8b00466
Biochemistry 2018, 57, 4776−4787

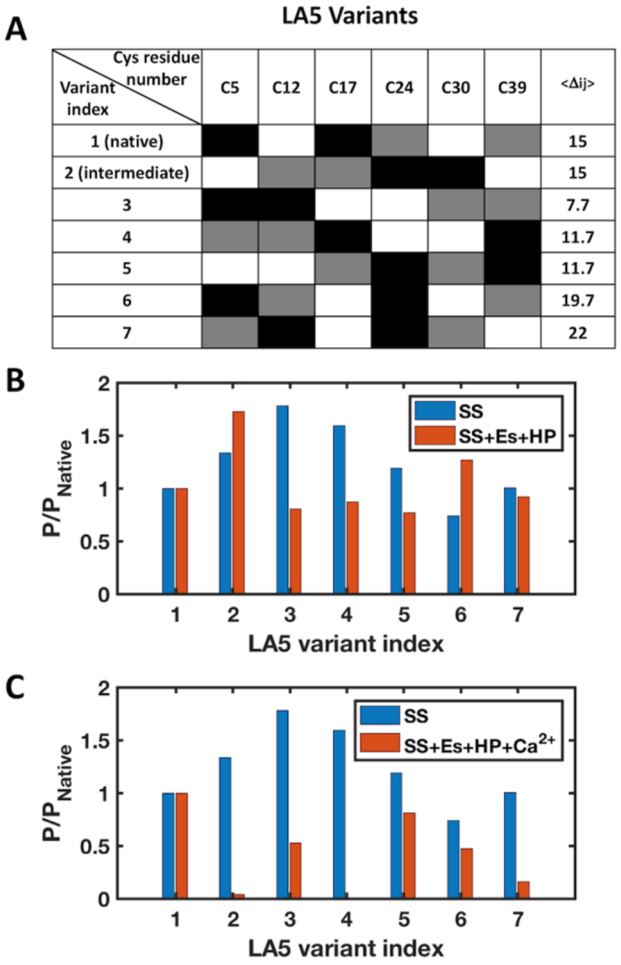


Figure 9. Evaluating disulfide bond connectivities in LA5 using coarse-grained MD simulations. (A) Seven combinations of three disulfide bonds were modeled. The numbered columns represent the positions of the six cysteine residues in the LA5 sequence. Each row represents a different connectivity: two cells with the same color represent a disulfide bond between the two corresponding cysteines. For example, in the native set of disulfide bonds (LA5 variant index 1), the connectivity is C5−C17 (black), C12−C30 (white), and C24−C39 (gray). The connectivity in each isoform is quantified by the average number of intervening amino acids $\langle \Delta ij \rangle$ (unitless number denoting the average number of inter-cysteine residues for each particular variant) between linked cysteines. (B) Probability for the formation of the indicated variants according to a model incorporating only disulfides (SS) and backbone flexibility (blue) vs a model that includes also electrostatic (Es) and hydrophobic (HP) interactions (red). The probabilities are normalized to the native connectivity (variant 1). (C) Normalized probabilities for the formation of the indicated variants according to a model incorporating only disulfides and backbone flexibility (blue) vs a model that includes also electrostatic, hydrophobic, and Ca^{2+}-mediated (Ca^{2+}) interactions (red).

DISCUSSION

Proteins with tandem repeats of small, disulfide-rich modules constitute a significant fraction of cell-surface receptors and extracellular matrix proteins.[49] How these proteins achieve the correct disulfide pairings during folding and whether non-native species are populated prior to acquisition of the native state are important questions relating to these proteins' biosynthesis and function.[22] LDLR is a key member of this class and, as such, has become a model system for developments in studying the structure and folding of disulfide-rich proteins. The presence of Ca^{2+} ions drives the folding of LDLR modules such as LA5 toward a compact native state, but even in the presence of Ca^{2+}, non-native LA5 species containing three disulfide bonds are known to arise during the folding reaction (e.g., Figure S3 and ref 23). This study set out to determine the kinetics of formation of such species, to identify their structures using an NMR-based experimental approach combining dynamic and static multi-dimensional measurements, and to further characterize the folding landscape of LA5 by combining this information with insight arising from MD calculations.

NMR has been used previously to study protein folding intermediates, generally by taking one of two approaches: monitoring directly the kinetics of folding, or trapping folding intermediate analogues for an in-depth equilibrium analysis. Most such studies have focused on proteins that fold slowly and exhibit multiphase folding kinetics due to *cis*/*trans* isomerization of prolines. In some cases, folding intermediates were shown to be highly flexible,[50] while in other cases, major regions of native-like structure were observed.[15,51] For proteins containing disulfide bonds, the disulfides were generally left intact during *in vitro* unfolding, and they were therefore present at the start of the refolding reactions monitored by NMR.[52,53] In the study presented here, NMR was used to monitor the entire process of oxidative folding, including disulfide bond formation. The first question addressed was whether backbone amide chemical shift information collected during LA5 folding revealed information about intermediate species. No evidence of non-native peaks from intermediates could be distinguished; this could be either because no such intermediates existed (contradicting other, non-NMR measurements) or because the non-native LA5 isomers did not stabilize the peptide backbone hydrogen bonding pattern sufficiently, to produce distinct amide NMR signatures. Instead, single-exponential kinetics were observed for the formation of native-state backbone NMR signals, for all amino acids in LA5.

By contrast, methods such as rpHPLC (Figure 1) and MS of intact LA5 (Figure S7) clearly demonstrated the presence of intermediates in disulfide formation, even if these methods provided no further structural characterization. For other proteins, MS of proteolytic fragments has been used to assign paired cysteine residues and thereby obtain maps of disulfide bonding patterns.[54] Such approaches can also be used to follow changes in disulfide connectivity as a function of time.[55] However, because of the compact and highly cross-linked nature of disulfide-bonded LA5, obtaining the necessary proteolytic cleavages to isolate the linked peptides for direct identification of cysteine pairing by MS is challenging (unpublished observations). Therefore, alternate approaches to identifying cysteine pairings and characterizing folding intermediates were needed.

NMR of selectively ^{13}C-labeled LA5 was found to be particularly suitable for mapping of the disulfide connectivity, since there was no need for polypeptide cleavage and isolation of bonded peptides; at the same time, ^{13}C NMR also avoided some of the limitations observed when focusing on the backbone resonances by ^{15}N NMR. In contrast to the evident spectral overlap characterizing backbone $^{1}H-^{15}N$ NMR experiments in incompletely oxidized LA5 and in reduced (unfolded) LA5, the ^{13}C spectral region corresponding to disulfide bonds, whether native or non-native, is distinct from

DOI: 10.1021/acs.biochem.8b00466
Biochemistry 2018, 57, 4776−4787

that of the free thiols present at the start of the oxidative folding reaction. Furthermore, peaks corresponding to disulfides in the natively folded LA5 were highly dispersed and unambiguously separated from the set of more poorly dispersed disulfide peaks arising from kinetic intermediates.

Despite the poor dispersion of peaks corresponding to disulfide intermediates, the dependence of cysteine ^{13}C chemical shifts on the residues' environment provided sufficient differences to assign peaks to specific cysteines. This assignment required a combination of real-time folding and equilibrium 3D NMR experiments, with the latter performed on suitably trapped intermediates. These experiments revealed different kinds of disulfide bonds present in the major non-native LA5 intermediate, which was found to contain two shorter-range disulfides (C12−C17 and C24−C30) and one longer-range disulfide (C5−C39). This species, "variant 2" in Figure 9, could be compared with other possible non-native species containing only short-range disulfides; this was done, for instance, by setting up *in silico* "variant 3" (C5−C12, C17−C24, C30−C39). This latter species is more likely to arise than other alternatives when only backbone dynamics are considered (Figure 9), yet it was not detected experimentally. Interestingly, when the nature of amino acid side chains was added to the computational model by including electrostatics and hydrophobic interactions between aromatic residues, the likelihood of the various species was altered. The probability of encountering the experimentally observed species increased, while the probabilities of finding the unobserved ones decreased. This result suggests that, despite having a highly dynamic backbone, the observed LA5 intermediate becomes the preferred variant because of nonspecific side chain interactions. Finally, MD calculations also found that the appearance of Ca^{2+} can override these preferences and tilt the balance in favor of the native disulfide bond connectivity over other variants.

It is interesting to consider further why and how the three-disulfide kinetic intermediate arises. As mentioned, this intermediate corresponds to the most favored of the species expected to arise when oxidizing LA5 in the absence of Ca^{2+}. Because Ca^{2+} is unlikely to be stably bound to LA5 until a suitable Ca^{2+}-binding pocket is at least partially formed, it is not surprising that some intermediate stages of LA5 folding resemble folding in the absence of Ca^{2+}. In very early stages, it is reasonable to assume that one- and two-disulfide species will be populated. Such species will rapidly isomerize, because free cysteine thiols with high effective concentrations are present on the same polypeptide and will facilitate thiol−disulfide exchange.[56] The transient and dynamic nature of one- and two-disulfide species is thus consistent with their lack of detection by NMR. Upon formation of three disulfide bonds, however, interconversion between species is expected to occur less readily, as isomerization will depend on the availability of an exogenous reductant, i.e., GSH. Indeed, the non-native intermediate observed during LA5 folding is a species containing three disulfide bonds, but with a connectivity differing from that of native LA5. This species lacks an appropriate Ca^{2+}-binding pocket, and the rate-limiting factor in its conversion to the native state is disulfide rearrangement. Though disulfide isomers that support formation of the Ca^{2+}-binding pocket (i.e., the native connectivity) are initially less populated, they are kinetically sampled to some extent. The presence of Ca^{2+} is thus expected to shift the equilibrium away from the initially dominant disulfide connectivity and toward the connectivity eventually observed in the polypeptide's native state.

It was previously shown that a limited number of side chains in LA5 were necessary to specify the native disulfide bond connectivity and the native fold.[21] The findings presented here support the idea that side chains contribute to the stabilization of non-native states prior to Ca^{2+} incorporation, leading to preferred folding intermediates being observed among the many alternatives. It would be interesting if some of the same side chains and properties that promote formation of the native state are also those that stabilize the intermediate, as this might explain why a smoother landscape for LA5 folding did not arise.

In summary, NMR spectroscopy was used, in combination with computational techniques, to reveal and analyze structural and dynamic aspects of transient LA5 oxidative folding intermediates. While ^{15}N backbone NMR showed that all intermediates are dynamic and failed to reveal their presence, ^{13}C NMR of the cysteine side chains revealed the preferred non-native species that accumulates. The disulfide bond connectivity of this species was not the one with the smallest average distance between disulfide-bonded cysteines, which could have been expected to be most populated based on polymer dynamics (Figure 9B,C). Instead, the transiently populated intermediate that was observed was apparently stabilized by side chain contributions compatible with a disulfide connectivity containing certain remote cysteine pairings (i.e., C5−C39); the incorporation of the Ca^{2+} ion then shifted the final equilibrium toward dominance by the native state. This work focused on the LA5 domain of LDLR, but similar studies can shed light on the energy landscapes of many proteins that fold in an analogous multistep manner, including other systems leading to diseases related to protein misfolding.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.8b00466.

Detailed experimental procedures, additional figures for MS experiments, NMR pulse sequences, additional time points for 2D $^{1}H-^{13}C$ kinetics, and 2D ^{13}C-edited NOESY spectra (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: deborah.fass@weizmann.ac.il. Telephone: +972-8-9343214.

*E-mail: lucio.frydman@weizmann.ac.il. Telephone: +972-8-9344903.

ORCID

Yaakov Levy: 0000-0002-9929-973X

Deborah Fass: 0000-0001-9418-6069

Lucio Frydman: 0000-0001-8208-3521

Funding

This work was supported by the European Research Council under the European Union's Seventh Framework Programme, Grant 310649 to D.F., and by the I-CORE Program of the Planning and Budgeting Committee and The Israel Science Foundation (Grant 1775/12) to D.F. and L.F. The Kimmel

Institute of Magnetic Resonance (Weizmann Institute) and the Perlman Family Foundation are gratefully acknowledged.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors are grateful to K. Zibzener (Weizmann Institute) for technical assistance. The authors thank The Chemical analysis and Mass spectrometry unit (Chemical Research Support, Weizmann Institute) for performing the mass spectrometry experiments on LA5.

■ REFERENCES

(1) Ollerenshaw, J. E., Tugarinov, V., and Kay, L. E. (2003) Methyl TROSY: explanation and experimental verification. *Magn. Reson. Chem. 41* (10), 843−852.
(2) Tugarinov, V., and Kay, L. E. (2005) Methyl Groups as Probes of Structure and Dynamics in NMR Studies of High-Molecular-Weight Proteins. *ChemBioChem 6* (9), 1567−1577.
(3) Frydman, L., Scherf, T., and Lupulescu, A. (2002) The acquisition of multidimensional NMR spectra within a single scan. *Proc. Natl. Acad. Sci. U. S. A. 99* (25), 15858−15862.
(4) Schanda, P., and Brutscher, B. (2005) Very Fast Two-Dimensional NMR Spectroscopy for Real-Time Investigation of Dynamic Events in Proteins on the Time Scale of Seconds. *J. Am. Chem. Soc. 127* (22), 8014−8015.
(5) Schanda, P., Kupče, Ē., and Brutscher, B. (2005) SOFAST-HMQC Experiments for Recording Two-dimensional Heteronuclear Correlation Spectra of Proteins within a Few Seconds. *J. Biomol. NMR 33* (4), 199−211.
(6) Schanda, P., Van Melckebeke, H., and Brutscher, B. (2006) Speeding Up Three-Dimensional Protein NMR Experiments to a Few Minutes. *J. Am. Chem. Soc. 128* (28), 9042−9043.
(7) Baldwin, A. J., and Kay, L. E. (2009) NMR spectroscopy brings invisible protein states into focus. *Nat. Chem. Biol. 5* (11), 808−814.
(8) Tal, A., and Frydman, L. (2010) Single-scan multidimensional magnetic resonance. *Prog. Nucl. Magn. Reson. Spectrosc. 57* (3), 241−292.
(9) Vallurupalli, P., Bouvignies, G., and Kay, L. E. (2012) Studying "Invisible" Excited Protein States in Slow Exchange with a Major State Conformation. *J. Am. Chem. Soc. 134* (19), 8148−8161.
(10) Mobli, M., and Hoch, J. C. (2014) Nonuniform sampling and non-Fourier signal processing methods in multidimensional NMR. *Prog. Nucl. Magn. Reson. Spectrosc. 83* (Suppl. C), 21−41.
(11) Billeter, M. (2017) Non-uniform sampling in biomolecular NMR. *J. Biomol. NMR 68* (2), 65−66.
(12) Fürtig, B., Buck, J., Manoharan, V., Bermel, W., Jäschke, A., Wenter, P., Pitsch, S., and Schwalbe, H. (2007) Time-resolved NMR studies of RNA folding. *Biopolymers 86* (5−6), 360−383.
(13) Buck, J., Fürtig, B., Noeske, J., Wöhnert, J., and Schwalbe, H. (2009) Time-Resolved NMR Spectroscopy: Ligand-Induced Refolding of Riboswitches. In *Riboswitches: Methods and Protocols* (Serganov, A., Ed.) pp 161−171, Humana Press, Totowa, NJ.
(14) Lee, M.-K., Gal, M., Frydman, L., and Varani, G. (2010) Real-time multidimensional NMR follows RNA folding with second resolution. *Proc. Natl. Acad. Sci. U. S. A. 107* (20), 9192−9197.
(15) Killick, T. R., Freund, S. M. V., and Fersht, A. R. (1999) Real-time nmr studies on a transient folding intermediate of barstar. *Protein Sci. 8* (6), 1286−1291.
(16) Corazza, A., Rennella, E., Schanda, P., Mimmi, M. C., Cutuil, T., Raimondi, S., Giorgetti, S., Fogolari, F., Viglino, P., Frydman, L., Gal, M., Bellotti, V., Brutscher, B., and Esposito, G. (2010) Native-unlike Long-lived Intermediates along the Folding Pathway of the Amyloidogenic Protein β2-Microglobulin Revealed by Real-time Two-dimensional NMR. *J. Biol. Chem. 285* (8), 5827−5835.
(17) Jeon, H., and Blacklow, S. C. (2005) Structure and physiologic function of the low-density lipoprotein receptor. *Annu. Rev. Biochem. 74* (1), 535−562.
(18) Hobbs, H. H., Russell, D. W., Brown, M. S., and Goldstein, J. L. (1990) The LDL Receptor Locus in Familial Hypercholesterolemia: Mutational Analysis of a Membrane Protein. *Annu. Rev. Genet. 24* (1), 133−170.
(19) Blacklow, S. C., and Kim, P. S. (1996) Protein folding and calcium binding defects arising from familial hypercholesterolemia mutations of the LDL receptor. *Nat. Struct. Mol. Biol. 3*, 758.
(20) Goldberg, A. C., Hopkins, P. N., Toth, P. P., Ballantyne, C. M., Rader, D. J., Robinson, J. G., Daniels, S. R., Gidding, S. S., de Ferranti, S. D., Ito, M. K., McGowan, M. P., Moriarty, P. M., Cromwell, W. C., Ross, J. L., and Ziajka, P. E. (2011) Familial Hypercholesterolemia: Screening, diagnosis and management of pediatric and adult patients. *J. Clin. Lipidol. 5* (3), S1−S8.
(21) Abdul-Aziz, D., Fisher, C., Beglova, N., and Blacklow, S. C. (2005) Folding and Binding Integrity of Variants of a Prototype Ligand-Binding Module from the LDL Receptor Possessing Multiple Alanine Substitutions. *Biochemistry 44* (13), 5075−5085.
(22) Jansens, A., van Duijn, E., and Braakman, I. (2002) Coordinated Nonvectorial Folding in a Newly Synthesized Multidomain Protein. *Science 298* (5602), 2401−2403.
(23) Arias-Moreno, X., Arolas, J. L., Aviles, F. X., Sancho, J., and Ventura, S. (2008) Scrambled Isomers as Key Intermediates in the Oxidative Folding of Ligand Binding Module 5 of the Low Density Lipoprotein Receptor. *J. Biol. Chem. 283* (20), 13627−13637.
(24) North, C. L., and Blacklow, S. C. (1999) Structural Independence of Ligand-Binding Modules Five and Six of the LDL Receptor. *Biochemistry 38* (13), 3926−3935.
(25) Koduri, V., and Blacklow, S. C. (2001) Folding Determinants of LDL Receptor Type A Modules. *Biochemistry 40* (43), 12801−12807.
(26) Guttman, M., and Komives, E. A. (2011) Structure, dynamics and binding of the LA45 module pair of the Low Density Lipoprotein Receptor suggest an important role for LA4 in ligand release. *Biochemistry 50* (51), 11001−11008.
(27) Staley, J. P., and Kim, P. S. (1994) Formation of a native-like subdomain in a partially folded intermediate of bovine pancreatic trypsin inhibitor. *Protein Sci. 3* (10), 1822−1832.
(28) Denk, D., and Bock, A. (1987) L-cysteine biosynthesis in Escherichia coli: nucleotide sequence and expression of the serine acetyltransferase (cysE) gene from the wild-type and a cysteine-excreting mutant. *Microbiology 133* (3), 515−25.
(29) Mueller, S., Senn, H., Gsell, B., Vetter, W., Baron, C., and Boeck, A. (1994) The Formation of Diselenide Bridges in Proteins by Incorporation of Selenocysteine Residues: Biosynthesis and Characterization of (Se)2-Thioredoxin. *Biochemistry 33* (11), 3404−3412.
(30) Crane, E. J., Vervoort, J., and Claiborne, A. (1997) ^{13}C NMR Analysis of the Cysteine-Sulfenic Acid Redox Center of Enterococcal NADH Peroxidase. *Biochemistry 36* (28), 8611−8618.
(31) Fass, D., Blacklow, S., Kim, P. S., and Berger, J. M. (1997) Molecular basis of familial hypercholesterolaemia from structure of LDL receptor module. *Nature 388* (6643), 691−693.
(32) Bowen, S., and Hilty, C. (2008) Time-Resolved Dynamic Nuclear Polarization Enhanced NMR Spectroscopy. *Angew. Chem., Int. Ed. 47* (28), 5235−5237.
(33) Bowen, S., and Hilty, C. (2010) Rapid sample injection for hyperpolarized NMR spectroscopy. *Phys. Chem. Chem. Phys. 12* (22), 5766−5770.
(34) Olsen, G., Markhasin, E., Szekely, O., Bretschneider, C., and Frydman, L. (2016) Optimizing water hyperpolarization and dissolution for sensitivity-enhanced 2D biomolecular NMR. *J. Magn. Reson. 264*, 49−58.
(35) Katsikis, S., Marin-Montesinos, I., Pons, M., Ludwig, C., and Günther, U. L. (2015) Improved Stability and Spectral Quality in Ex Situ Dissolution DNP Using an Improved Transfer Device. *Appl. Magn. Reson. 46* (7), 723−729.

DOI: 10.1021/acs.biochem.8b00466
Biochemistry 2018, 57, 4776−4787

(36) Azia, A., and Levy, Y. (2009) Nonnative Electrostatic Interactions Can Modulate Protein Folding: Molecular Dynamics with a Grain of Salt. *J. Mol. Biol. 393* (2), 527−542.
(37) Clementi, C., Nymeyer, H., and Onuchic, J. N. (2000) Topological and energetic factors: what determines the structural details of the transition state ensemble and "en-route" intermediates for protein folding? an investigation for small globular proteins. *J. Mol. Biol. 298* (5), 937−953.
(38) Gal, M., Mishkovsky, M., and Frydman, L. (2006) Real-Time Monitoring of Chemical Transformations by Ultrafast 2D NMR Spectroscopy. *J. Am. Chem. Soc. 128* (3), 951−956.
(39) Gal, M., Schanda, P., Brutscher, B., and Frydman, L. (2007) UltraSOFAST HMQC NMR and the Repetitive Acquisition of 2D Protein Spectra at Hz Rates. *J. Am. Chem. Soc. 129* (5), 1372−1377.
(40) States, D. J., Haberkorn, R. A., and Ruben, D. J. (1982) A two-dimensional nuclear overhauser experiment with pure absorption phase in four quadrants. *J. Magn. Reson. (1969-1992) 48* (2), 286−292.
(41) Sharma, D., and Rajarathnam, K. (2000) ^{13}C NMR chemical shifts can predict disulfide bond formation. *J. Biomol. NMR 18* (2), 165−171.
(42) Arnesano, F., Balatri, E., Banci, L., Bertini, I., and Winge, D. R. (2005) Folding Studies of Cox17 Reveal an Important Interplay of Cysteine Oxidation and Copper Binding. *Structure 13* (5), 713−722.
(43) Sahu, D., Debnath, P., Takayama, Y., and Iwahara, J. (2008) Redox properties of the A-domain of the HMGB1 protein. *FEBS Lett. 582* (29), 3973−3978.
(44) Mobli, M., and King, G. F. (2010) NMR methods for determining disulfide-bond connectivities. *Toxicon 56* (6), 849−854.
(45) Banci, L., Bertini, I., Cefaro, C., Ciofi-Baffoni, S., and Gallo, A. (2011) Functional Role of Two Interhelical Disulfide Bonds in Human Cox17 Protein from a Structural Perspective. *J. Biol. Chem. 286* (39), 34382−34390.
(46) Váradi, G., Tóth, G. K., Kele, Z., Galgóczy, L., Fizil, Á., and Batta, G. (2013) Synthesis of PAF, an Antifungal Protein from P. chrysogenum, by Native Chemical Ligation: Native Disulfide Pattern and Fold Obtained upon Oxidative Refolding. *Chem. - Eur. J. 19* (38), 12684−12692.
(47) Morris, G. A., and Freeman, R. (1979) Enhancement of nuclear magnetic resonance signals by polarization transfer. *J. Am. Chem. Soc. 101* (3), 760−762.
(48) Davis, A. L., Keeler, J., Laue, E. D., and Moskau, D. (1992) Experiments for recording pure-absorption heteronuclear correlation spectra using pulsed field gradients. *J. Magn. Reson. (1969-1992) 98* (1), 207−216.
(49) Uings, I. J., and Farrow, S. N. (2000) Cell receptors and cell signalling. *Mol. Pathol. 53* (6), 295−299.
(50) Reader, J. S., Van Nuland, N. A. J., Thompson, G. S., Ferguson, S. J., Dobson, C. M., and Radford, S. E. (2001) A partially folded intermediate species of the β-sheet protein apo-pseudoazurin is trapped during proline-limited folding. *Protein Sci. 10* (6), 1216−1224.
(51) Balbach, J., Steegborn, C., Schindler, T., and Schmid, F. X. (1999) A protein folding intermediate of ribonuclease T_1 characterized at high resolution by 1D and 2D real-time NMR spectroscopy. *J. Mol. Biol. 285* (2), 829−842.
(52) Udgaonkar, J. B., and Baldwin, R. L. (1988) NMR evidence for an early framework intermediate on the folding pathway of ribonuclease A. *Nature 335*, 694.
(53) Houry, W. A., and Scheraga, H. A. (1996) Structure of a Hydrophobically Collapsed Intermediate on the Conformational Folding Pathway of Ribonuclease A Probed by Hydrogen−Deuterium Exchange. *Biochemistry 35* (36), 11734−11746.
(54) Borges, C. R., and Sherma, N. D. (2014) Techniques for the Analysis of Cysteine Sulfhydryls and Oxidative Protein Folding. *Antioxid. Redox Signaling 21* (3), 511−531.
(55) Heldman, N., Vonshak, O., Sevier, C. S., Vitu, E., Mehlman, T., and Fass, D. (2010) Steps in reductive activation of the disulfide-generating enzyme Ero1p. *Protein Sci. 19* (10), 1863−1876.
(56) Weissman, J. S., and Kim, P. S. (1992) The pro region of BPTI facilitates folding. *Cell 71* (5), 841−851.

DOI: 10.1021/acs.biochem.8b00466
Biochemistry 2018, 57, 4776−4787