BIOCHEMISTRY
including biophysical chemistry & molecular biology

Article

pubs.acs.org/biochemistry

Probing the Residual Structure of the Low Populated Denatured State of ADA2h under Folding Conditions by Relaxation Dispersion Nuclear Magnetic Resonance Spectroscopy

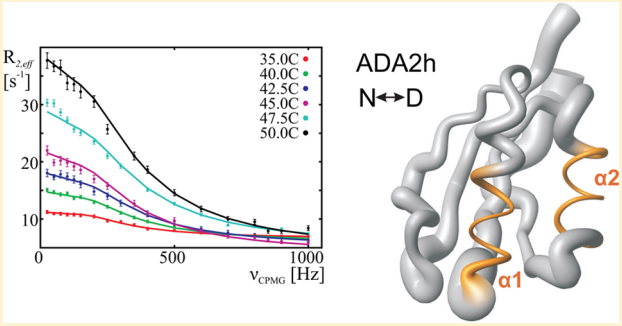
Yulia Pustovalova,[†] Predrag Kukic,[‡] Michele Vendruscolo,[‡] and Dmitry M. Korzhnev*[,†]

[†]Department of Molecular Biology and Biophysics, University of Connecticut Health Center, Farmington, Connecticut 06030, United States

[‡]Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, United Kingdom

Supporting Information

ABSTRACT: The structural characterization of low populated states of proteins with accuracy comparable to that achievable for native states is important for understanding the mechanisms of protein folding and function, as well as misfolding and aggregation. Because of the transient nature of these low populated states, they are seldom detected directly under conditions that favor folding. The activation domain of human procarboxypeptidase A2 (ADA2h) is an α/β-protein that forms amyloid fibrils at low pH, presumably initiated from a denatured state with a considerable amount of residual structure. Here we used Carr−Parcell−Meiboom−Gill relaxation dispersion (CPMG RD) nuclear magnetic resonance (NMR) spectroscopy to characterize the structure of the denatured state of the ADA2h I71V mutant under conditions that favor folding. Under these conditions, the lifetime of the denatured state of I71V ADA2h is on the order of milliseconds and its population is approximately several percent, which makes this mutant amenable to studies by CPMG RD methods. The nearly complete set of CPMG RD-derived backbone ^{15}N, ^{13}C, and ^{1}H NMR chemical shifts in the I71V ADA2h denatured state reveals that it retains a significant fraction (up to 50−60%) of nativelike α-helical structure, while the regions encompassing native β-strands are structured to a much lesser extent. The nativelike α-helical structure of the denatured state can bring together hydrophobic residues on the same sides of α-helices, making them available for intra- or intermolecular interactions. CPMG RD data analysis thus allowed a detailed structural characterization of the ADA2h denatured state under folding conditions not previously achieved for this protein.



The behavior of proteins can be described in terms of free energy landscapes, which include denatured, intermediate, and native states.[1,2] Much of our knowledge of protein structure and folding derives from studies of the highly populated native states using structural biology methods, including X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy, and a variety of biophysical approaches. By contrast, the properties of denatured[3] and intermediate[4−6] states, which are low populated under conditions that favor protein folding, still remain largely elusive. The transient nature and low populations of these "excited" states make them extremely difficult to study with the conventional experimental methods that are primarily used to probe folded states.

However, characterization of these excited states of proteins to the same extent as the native states is crucial for understanding the mechanisms of protein folding[7,8] and function,[9] as well as of misfolding and aggregation.[10,11] Transient unfolding events leading to the formation of sparsely populated non-native states with exposed hydrophobic regions may cause protein aggregation, which results in a loss of protein function and a gain of toxic properties.[12−15] Indeed, more than 40 human diseases have been associated with protein misfolding and aggregation into amyloid fibrils.[14,15]

NMR spectroscopy provides unique experimental tools for probing the structure and thermodynamics of excited states transiently populated under nativelike conditions caused by thermal fluctuations.[16−19] Although the signals of such states cannot be directly observed in the NMR spectra, the exchange between ground and excited states accompanied by changes in NMR chemical shifts may result in line broadening of the highly populated ground state. This line broadening is exploited in a range of NMR experiments that provide information about the exchange kinetics (rate constants), thermodynamics (populations of exchanging states), and structure of excited states (NMR chemical shifts).[16−19] Different types of NMR experiments can be used depending of the time scale of the exchange process. Thus, rotating frame $R_{1\rho}$ relaxation measurements can probe microsecond to millisecond processes with the

Received: April 1, 2015
Revised: June 25, 2015
Published: June 26, 2015

DOI: 10.1021/acs.biochem.5b00345
Biochemistry 2015, 54, 4611−4622

exchange rate constant k_{ex} ranging from 10^2 to 10^5 s^{-1} ($k_{ex} = k_{A\rightarrow B} + k_{B\rightarrow A}$ for A ↔ B interconversion).[20] Carr−Purcell−Meiboom−Gill[21,22] relaxation dispersion (CPMG RD) experiments are very effective for studying millisecond time-scale exchange with k_{ex} ranging from several hundreds to several thousands of inverse seconds.[16−19] Chemical-exchange saturation transfer (CEST) experiments are suitable for studying slower exchange processes with k_{ex} values of up to several hundred inverse seconds.[23−25]

In this context, CPMG RD methods have emerged as powerful tools for studying millisecond time-scale unfolding transitions in marginally stable proteins and destabilized protein mutants with unfolding free energies of 1−3 kcal/mol.[17,18] In CPMG RD experiments, the effective transverse relaxation rates of the protein ground sate ($R_{2,eff}$) are recorded as a function of the frequency of application of refocusing pulses in the CPMG pulse train (ν_{CPMG}). The dependences of $R_{2,eff}$ on ν_{CPMG}, called relaxation dispersion profiles, allow extraction of exchange rate constants, populations of the excited states (in particular denatured and intermediate states), and absolute values of chemical shift differences between states. A number of CPMG RD experiments have been proposed for nearly every type of NMR active nucleus of the protein backbone[26−34] and side chains,[35−39] which provide multiple independent probes of the unfolding−refolding equilibrium. These methods have been successfully used to probe the unfolding transitions in several small proteins and to structurally characterize low populated protein folding intermediates primarily based on the CPMG RD-derived chemical shifts.[7,8,10,40]

Here we use CPMG RD methods to study the folding process of the activation domain of human procarboxypeptidase A2 (ADA2h), which is an 81-residue domain released upon cleavage of an inactive precursor of the enzyme that consists of two α-helices packed against a four-stranded β-sheet.[41] The folding pathway of wild-type ADA2h has been extensively characterized,[42−47] showing that under nativelike conditions (pH 7.0, 25 °C) this protein domain folds by a reversible two-state mechanism via a transition state that consists of a partially structured core encompassing the two central β-strands (β1 and β3) and a partially folded α-helix 2.[44] On the other hand, thermal and chemical denaturation of ADA2h at acidic pH results in irreversible protein aggregation.[46,47]

ADA2h has been a subject of an increasing interest as a convenient model system, because of its ability to form amyloid fibrils upon denaturation under low-pH conditions that are not associated with any known disease.[46,47] It was proposed that amyloid formation by ADA2h proceeds through multiple aggregated states: the early soluble β-aggregates, more organized β-strand reach structures, and, finally, the well-structured amyloid fibrils.[47] A growing number of known proteins susceptible to β-aggregation has led to the development of methods for the prediction of aggregation tendencies of amino acid sequences, including TANGO[48,49] and Zyggregator.[50,51] In their work, Cerda-Costa et al.[47] used these algorithms to theoretically predict aggregation tendencies for a number of ADA2h mutants and then analyzed the ability of these mutants to form amyloids using pH-jump experiments. The computational predictions using the TANGO algorithm[48,49] identified the C-terminal part of ADA2h encompassing native β-strand 4 as the major contributor to amyloid formation. In agreement with these predictions, mutations of C-terminal residues S74, I75, E78, and D79 in the suggested core of the β-structured fibril were shown to hinder amyloid formation.[47] On the other hand, fibril formation was also inhibited by V64G, F65A, and E67K mutations on the outer surface of α-helix 2, suggesting that additional regions of ADA2h may be important for amyloid formation. These results are not surprising, because amyloid formation by the ADA2h is thought to be a multistep process that proceeds through one or several potentially nonfibrillar aggregates that may involve regions of ADA2h other than the C-terminus.

The denatured state of ADA2h under conditions that favor protein folding has an extensive residual structure. It has thus been proposed that early kinetics of amyloid formation is determined by the properties of the denatured state rather than the highly populated native conformer.[46] In this respect, the rate of amyloid formation by the ADA2h mutants shows little correlation with the overall protein stability; instead, aggregation kinetics seem to be determined by the stability of certain regions of residual structure in the denatured state. For example, stabilization of one or both α-helices, which makes intramolecular contacts in the denatured state more preferable than intermolecular ones, results in slower aggregation rates.[46] Hence, structural characterization of the ADA2h denatured state ensemble can provide important clues for understanding early events in amyloid formation by this protein.

In this work, we report a CPMG RD study of a transiently populated denatured state of a destabilizing I71V ADA2h mutant, which is predicted to have an aggregation tendency similar to that of wild-type ADA2h (according to the TANGO algorithm;[48,49] see Table 1 of ref 47), yet displays slower aggregation kinetics.[47] The unfolding free energy of 2.9 kcal/mol, corresponding to an ∼1% population of the denatured state, and the refolding rate of 450 s^{-1} for this mutant reported at 25 °C[44] make this state amenable to investigation by CPMG RD methods, which are best suited for studying protein millisecond time-scale transitions to high-energy states with populations in the range of 0.5−10%.[16−19] Similar to wild-type ADA2h, the I71V mutant can form amyloids at low pH and might as well serve as a model for studying the properties of a denatured state transiently populated under conditions that favor protein folding. The results presented in this work confirm that ADA2h folds reversibly by a two-state mechanism. The CPMG RD-derived chemical shifts for the backbone ^{15}N, $^{1}H^{N}$, $^{13}C^{\alpha}$, $^{1}H^{\alpha}$, and $^{13}C'$ nuclei in the denatured state allowed us to directly characterize the residual structure of this state, which mainly consists of the partially folded nativelike α-helices (residues 21−27 and 61−67). On the other hand, $^{1}H/^{2}H$ exchange experiments[52] reveal no regions of wild-type and I71V ADA2h with significantly elevated protection factors, suggesting that the denatured state is likely a dynamic ensemble of partially structured metastable conformers. These results are particularly relevant considering that at acidic pH ADA2h can form amyloid fibrils, a process that is presumably initiated from the denatured state.[46] This work is one of the first studies to report experimental data that allow characterization of a conformationally heterogeneous denatured state under folding conditions.

MATERIALS AND METHODS

Sample Preparation. The genes encoding wild-type and I71V ADA2h (residues 1−81) inserted after the cleavage site for TEV protease (ENLYFQG) were codon-optimized for expression in *Escherichia coli*, custom synthesized (GenScript), and subcloned into the pET28b(+) vector (Novagen) using

DOI: 10.1021/acs.biochem.5b00345
Biochemistry 2015, 54, 4611−4622

*Nde*I−*Bam*HI restriction sites. The protein samples used for temperature-dependent ^{15}N CPMG relaxation dispersion measurements and for ^{1}H/^{2}H exchange experiments were expressed in *E. coli* BL21(DE3) cells in H_2O-based M9 minimal medium supplemented with $^{15}NH_4Cl$ as the sole nitrogen source. $^{13}C'$,[32,33] ^{15}N,[28] and $^{1}H^{N}$ [29] CPMG RD experiments used for measurements of the backbone chemical shifts in the "excited" state were performed for the uniformly ^{15}N-, ^{13}C-, and ^{2}H-labeled I71V ADA2h sample expressed in 100% D_2O-based M9 medium supplemented with $^{15}NH_4Cl$ and [^{13}C]glucose. Uniformly ^{15}N-labeled, selectively $^{13}C^{\alpha}$-labeled I71V ADA2h for $^{13}C^{\alpha}$ CPMG measurements[29] was expressed in H_2O-based M9 medium supplemented with [2-^{13}C]glucose that allows labeling of α-carbons of all amino acids except Ile, Val, and Leu.[53] $^{1}H^{\alpha}$ CPMG dispersion experiments[34] were conducted on the I71V ADA2h sample produced in 50% D_2O/50% H_2O M9 medium supplemented with $^{15}NH_4Cl$ and [U-^{2}H/^{13}C]-glucose. This labeling scheme achieves a high level of deuteration at H^{β} positions and allows us to record relaxation dispersion profiles for all amino acids except Gly, Ser, and Thr. All proteins were expressed and purified using the same protocol. Cells were grown at 37 °C to an OD_{600} of 0.8−0.9. Protein expression was induced by 1 mM IPTG overnight at 20 °C. The protein was purified by Ni^{2+} affinity chromatography followed by cleavage with TEV protease overnight at 4 °C. The final step of purification included size-exclusion chromatography on a HighLoad Superdex 75 column followed by protein concentration. The final NMR samples contained 0.7−1.2 mM protein, 50 mM sodium phosphate, and 0.05% NaN_3. ^{15}N-labeled and ^{15}N/^{13}C/^{2}H-labeled I71V ADA2h samples were dissolved in 90% H_2O/10% D_2O, pH 7.7 buffer, whereas the samples used for $^{13}C^{\alpha}$ and $^{1}H^{\alpha}$ dispersion measurements were lyophilized from 100% H_2O, pH 7.6 buffer and then dissolved in the same volume of 100% D_2O. The sample used for the $^{13}C^{\alpha}$ CPMG RD experiment was also used for ^{1}H/^{2}H exchange measurements (see below).

NMR Spectroscopy. The backbone resonance assignment for the I71V ADA2h mutant was based on the existing assignment for wild-type ADA2h (BioMagResBank entry 5561)[41] and was performed using triple-resonance HNCA and HNCO spectra recorded for the ^{15}N/^{13}C/^{2}H-labeled sample and HNHA spectrum recorded for the $^{13}C^{\alpha}$-labeled sample at 11.7 T (500 MHz ^{1}H frequency) with an Agilent VNMRS spectrometer.[54]

Carr−Purcell−Meiboom−Gill relaxation dispersion (CPMG RD) experiments for the backbone ^{15}N,[28] $^{1}H^{N}$,[29] $^{13}C^{\alpha}$,[32] $^{1}H^{\alpha}$,[34] and $^{13}C'$[32,33] nuclei were recorded at 11.7 and 18.8 T (500 and 800 MHz ^{1}H frequencies) with Agilent VNMRS spectrometers. ^{15}N CPMG RD data for the ^{15}N-labeled protonated I71V ADA2h sample were collected at six temperatures ranging from 35 to 50 °C (35, 40, 42.5, 45, 47.5, and 50 °C). Additionally, ^{15}N, $^{1}H^{N}$, and $^{13}C'$ CPMG RD experiments were performed at 40 °C for the ^{15}N/^{13}C/^{2}H-labeled I71V ADA2h sample dissolved in 90% H_2O/10% D_2O buffer, and $^{13}C^{\alpha}$ and $^{1}H^{\alpha}$ CPMG RD experiments were performed at 40 °C using appropriately labeled I71V ADA2h samples dissolved in 100% D_2O buffer (see above).

All spectra were processed in NMRPipe and visualized in NMRDraw.[55] Relaxation dispersion profiles, $R_{2,\mathrm{eff}}(\nu_{\mathrm{CPMG}})$, were generated from peak intensities $I_1(\nu_{\mathrm{CPMG}})$ in a series of two-dimensional spectra recorded at different frequencies of application of the CPMG refocusing pulses, where $\nu_{\mathrm{CPMG}} = 1/(2\delta)$ (δ is the delay between 180° refocusing pulses in the CPMG sequence). Peak intensities were extracted from a series of two-dimensional spectra using the MUNIN approach[56,57] and were converted into the effective relaxation rates using the equation $R_{2,\mathrm{eff}}(\nu_{\mathrm{CPMG}}) = -1/T_{\mathrm{relax}} \ln[I_1(\nu_{\mathrm{CPMG}})/I_0]$, where T_{relax} is a constant relaxation delay and I_0 values are peak intensities obtained with relaxation delay T_{relax} omitted.[35] A T_{relax} of 40 ms was used for ^{15}N and $^{1}H^{N}$ and 30 ms for $^{1}H^{\alpha}$ and $^{13}C^{\alpha}$ CPMG RD experiments. Experimental errors were calculated as described previously;[58] each CPMG RD experiment included at least two duplicated spectra recorded at the same ν_{CPMG} that were used to estimate errors in the measured $R_{2,\mathrm{eff}}(\nu_{\mathrm{CPMG}})$ rates. Minimal errors of 2% or 0.2 s^{-1}, whichever is greater, were assumed for ^{15}N, $^{1}H^{N}$, and $^{13}C'$ CPMG RD data and 5% or 0.5 s^{-1}, whichever is greater, for $^{13}C^{\alpha}$ and $^{1}H^{\alpha}$ CPMG RD data. Larger errors assumed for $^{13}C^{\alpha}$ and $^{1}H^{\alpha}$ $R_{2,\mathrm{eff}}$ rates reflect the quality of our experimental data.

Relaxation Dispersion Data Analysis. ^{15}N CPMG RD profiles for uniformly ^{15}N-labeled protonated I71V ADA2h were recorded at $n_t = 6$ temperatures (35 to 50 °C) and at $n_f = 2$ magnetic fields (11.7 and 18.8 T) and were analyzed to extract kinetic and thermodynamic parameters that describe the I71V ADA2h unfolding−refolding transitions. The data for $n_r = 41$ residues with nonoverlapped cross-peaks over the considered range of temperatures were fit together to a model of two-site exchange between the "invisible" denatured (D) and ground native (N) states (D ↔ N) by minimization of the following χ^2 target function:

$$\chi^2(\zeta) = \sum \frac{(R_{2,\mathrm{eff}}^{\mathrm{clc}}(\zeta) - R_{2,\mathrm{eff}})^2}{(\Delta R_{2,\mathrm{eff}})^2} \quad (1)$$

where $R_{2,\mathrm{eff}}$ and $\Delta R_{2,\mathrm{eff}}$ are experimental effective relaxation rates and their uncertainties, respectively, $R_{2,\mathrm{eff}}^{\mathrm{clc}}(\zeta)$ values are predicted relaxation rates, ζ denotes the set of n_{par} adjustable model parameters, and the summation is over the number of experimental data points, n_{dat}. $R_{2,\mathrm{eff}}$ rates were predicted by modeling magnetization evolution during the CPMG sequence using numerical solution of Bloch−McConnell equations, as described previously.[7] A global fit of the temperature-dependent ^{15}N CPMG RD data was performed as described elsewhere[7] assuming that (i) chemical shift differences between D and N states, $\Delta\varpi_{\mathrm{DN}}$, are independent of temperature and (ii) temperature dependences of unfolding and refolding rates $k_{\mathrm{N\to D}}$ and $k_{\mathrm{D\to N}}$, respectively, follow the Eyring equation $k_{\mathrm{i\to j}} = (k_B \kappa T/h)/(-\Delta G_{\mathrm{ij}}^{\dagger}/RT)$, where i,j represents {D,N}, $\Delta G_{\mathrm{ij}}^{\dagger} = \Delta H_{\mathrm{ij}}^{\dagger} - T\Delta S_{\mathrm{ij}}^{\dagger}$ is the activation free energy, $\Delta H_{\mathrm{ij}}^{\dagger}$ and $\Delta S_{\mathrm{ij}}^{\dagger}$ are the activation enthalpy and entropy, respectively, κ is a transmission coefficient of 1.6×10^{-7} shown to be appropriate for protein folding,[59] k_B is Boltzmann's constant, h is Planck's constant, and R is the universal gas constant. The global optimization model included $n_{\mathrm{par}} = 4 + n_r + n_r n_f n_t$ adjustable parameters: equilibrium and activation enthalpies and entropies for the D ↔ N process, ΔH_{DN}, $\Delta H_{\mathrm{D\to N}}^{\dagger}$, ΔS_{DN}, $\Delta S_{\mathrm{D\to N}}^{\dagger}$, and n_r absolute values of chemical shift differences between D and N states, $|\Delta\varpi_{\mathrm{DN}}|$, for all considered residues, and $n_r n_f n_t$ intrinsic relaxation rates $R_{2,0}$ for all considered residues, temperatures, and magnetic fields. The errors in extracted exchange parameters were estimated using the covariance matrix method.[60] To estimate the order of magnitude of the temperature dependence of ^{15}N chemical shifts of the minor state, we have performed a separate analysis of temperature-dependent ^{15}N CPMG RD data assuming a linear temperature

DOI: 10.1021/acs.biochem.5b00345
Biochemistry 2015, 54, 4611−4622

Table 1. Results of the Global Fits of CPMG RD Data for ^{15}N, $^{1}H^{N}$, $^{13}C'$, $^{13}C^{\alpha}$, and $^{1}H^{\alpha}$ Nuclei of I71V ADA2h

nucleus[a]	temp	χ^2	$\sum\chi^2$	no. of residues	no. of degrees of freedom (#DF)	k_{ex} (s^{-1})[b]	p_D (%)[b]
^{15}N	35.0	949	809	66	1978	697.2 ± 9.2 (887)	0.77 ± 0.01 (0.67)
	40.0	1244	1098	62	1858	932.2 ± 9.5 (973)	1.23 ± 0.01 (1.21)
	42.5	1693	1556	60	1748	1085.1 ± 11.4 (1021)	1.63 ± 0.01 (1.62)
	45.0	2592	1666	60	1752	1050.4 ± 12.1 (1076)	2.34 ± 0.02 (2.20)
	47.5	2550	2050	55	1578	1211.0 ± 14.8 (1140)	2.85 ± 0.02 (3.04)
	50.0	3883	2849	44	1302	1203.5 ± 16.2 (1211)	4.15 ± 0.04 (4.17)
	35−50	13530	10276	41	7135		
^{15}N	40.0	958	715	66 (66)[c]	1912	731.9 ± 4.5	1.64 ± 0.01
$^{1}H^{N}$	40.0	1880	1397	64 (66)[c]	1662	766.0 ± 16.2	1.19 ± 0.02
$^{13}C'$	40.0	2005	1686	62 (63)[c]	1636	641.0 ± 10.0	1.68 ± 0.02
$^{13}C^{\alpha}$	40.0	1464	1263	40 (47)[c]	1700	2022.8 ± 92.9	1.37 ± 0.10
$^{1}H^{\alpha}$	40.0	291		6 (62)[c]	232	1144.1 ± 120.7	0.95 ± 0.07

[a]The ^{15}N-labeled protonated I71V ADA2h sample was used for all temperature-dependent ^{15}N CPMG RD experiments (top part). A $^{15}N/^{13}C/^{2}H$-labeled sample was used for ^{15}N, $^{1}H^{N}$, and $^{13}C'$ CPMG RD experiments (bottom part). Protein samples used for $^{13}C^{\alpha}$ and $^{1}H^{\alpha}$ CPMG RD measurements are described in Materials and Methods. [b]Exchange parameters (k_{ex} and p_D) predicted from equilibrium and activation entropies and enthalpies (see Materials and Methods) obtained in a global fit of ^{15}N CPMG RD data for all temperatures are shown in parentheses. [c]Number of residues used in the global data fit. Shown in parentheses is the total number of residues for which CPMG RD data were recorded and analyzed. For the residues that were not used in the global data fits, $|\Delta\varpi_{DN}|$ values were calculated from per-residue fits with k_{ex} and p_D fixed to the values obtained in the global data fits.

dependence of ^{15}N $\Delta\varpi_{DN}$, resulting in a mean $d|\Delta\varpi_{DN}|/dT$ of 0.03 ± 0.02 ppm/K that reflects modulation of transient structure in the denatured state ensemble.

^{15}N CPMG RD data for multiple residues of uniformly ^{15}N-labeled protonated I71V ADA2h recorded at magnetic fields of 11.7 and 18.8 T were also analyzed individually for each temperature to extract the population of the denatured state, p_D, and exchange rate constant ($k_{ex} = k_{N\to D} + k_{D\to N}$), n_r values of $|\Delta\varpi_{DN}|$, $n_r n_f$ intrinsic relaxation rates $R_{2,0}$ (a total of $n_{par} = 2 + n_r + n_r n_f$ parameters were extracted for each temperature; n_r values were different for different temperatures). Good agreement between model parameters extracted from these fits and the values obtained from the global fit of temperature-dependent ^{15}N CPMG RD data (Table 1) validates the assumptions made in the global fit of the temperature-dependent data. The quality of the data fit was assessed by comparison of the value of the χ^2 target function in its minimum with the number of degrees of freedom of the model, $\#DF = n_{dat} - n_{par}$ (Table 1, columns 3 and 6).[61] In addition, we have performed CPMG RD data fits individually for each residue; comparison of the sum of χ^2 target functions ($\sum\chi^2$) obtained in these fits (Table 1, column 4) with χ^2 values obtained in global analyses of the data for multiple residues (Table 1, column 3) was used as an additional criterion to assess the quality of data fit by the global two-state exchange model (see the text for details).

To characterize residual structure in the denatured state of I71V ADA2h, we have collected ^{15}N, $^{1}H^{N}$, $^{13}C'$, $^{13}C^{\alpha}$, and $^{1}H^{\alpha}$ CPMG RD data at 40 °C and analyzed these data separately for each type of nucleus to extract absolute values of chemical shift differences between the denatured (D) and native (N) states, $|\Delta\varpi_{DN}|$. These data were collected using three different I71V ADA2h samples with appropriate isotope labeling schemes [$^{15}N/^{13}C/^{2}H$ sample in 90% H_2O/10% D_2O buffer for ^{15}N, $^{1}H^{N}$, and $^{13}C'$ CPMG measurements;[28,29,32,33] ^{15}N/selectively $^{13}C^{\alpha}/^{1}H$ sample in 100% D_2O buffer for the $^{13}C^{\alpha}$ CPMG experiment;[32] $^{15}N/^{13}C$/partially ^{2}H sample in 100% D_2O buffer for the $^{1}H^{\alpha}$ CPMG experiment[34] (see above)]. First, the data for a subset of residues displaying pronounced relaxation dispersion profiles (different for each type of nucleus) were globally fit to the model of two-state exchange, D↔N, to extract the population of the denatured state, p_D, and exchange rate constant, k_{ex} (Table 1, bottom part). Then, the values of $|\Delta\varpi_{DN}|$ for all residues with nonoverlapped cross-peaks were extracted from data fits with p_D and k_{ex} fixed to previously determined values.

NMR spectra of the "invisible" denatured state were reconstructed by combining the observed chemical shifts of the ground native state, ϖ_N, with the chemical shift differences between states, $\Delta\varpi_{DN}$, obtained from the analysis of CPMG RD data: $\varpi_D = \varpi_N + \Delta\varpi_{DN}$. Signs of chemical shift differences, $\Delta\varpi_{DN}$, for ^{15}N, $^{1}H^{N}$, $^{13}C'$, and $^{13}C^{\alpha}$ nuclei were determined from differences in peak positions in ^{15}N or ^{13}C HSQC spectra recorded at magnetic field strengths of 11.7 and 18.8 T and/or from differences in peak positions in HSQC and HMQC spectra recorded at 18.8 T.[62,63] In cases where signs of the extracted ^{15}N, $^{1}H^{N}$, $^{13}C^{\alpha}$, and $^{13}C'$ chemical shift differences cannot be determined experimentally and for all $^{1}H^{\alpha}$ chemical shift differences, the $\Delta\varpi_{DN}$ signs were selected to be the same as signs of $\varpi_N - \varpi_{RC}$, where ϖ_{RC} is the random-coil chemical shift predicted using the CSI module of NMRView program.[64,65] The procedure used for measurement of $\Delta\varpi_{DN}$ signs is described in detail in the Supporting Information; the resulting ϖ_N, ϖ_D, and $\Delta\varpi_{DN}$ values are listed in Tables S1 and S2 of the Supporting Information.

The secondary structures of the native and denatured states of I71V ADA2h were predicted from the backbone chemical shifts using TALOS+[66] and δ2D;[67] the backbone flexibility was predicted from chemical shifts using the RCI approach.[68,69] Secondary structure populations were also predicted from the amino acid sequence using the s2D method.[70]

Hydrogen/Deuterium Exchange Measurements. $^{1}H/^{2}H$ exchange for the backbone amide groups of wild-type and I71V ADA2h was monitored by recording $^{1}H-^{15}N$ HSQC spectra of a protein sample lyophilized from 100% H_2O, 50 mM sodium phosphate, 0.05% NaN_3, pH 7.6 buffer and then dissolved in the same volume of 100% D_2O at 20 °C. Peak intensities in a series of HSQC spectra recorded as a function of

DOI: 10.1021/acs.biochem.5b00345
Biochemistry 2015, 54, 4611−4622

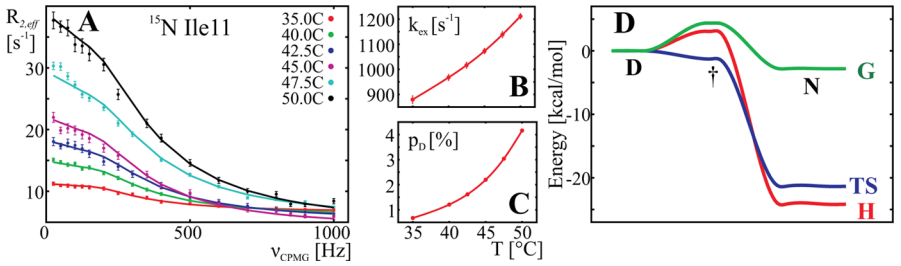


Figure 1. (A) ^{15}N CPMG relaxation dispersion profiles of residue Ile11 of the ^{15}N-labeled protonated I71V mutational variant of ADA2h recorded as a function of temperature at a magnetic field strength of 18.8 T (circles) and their best fit to the two-state exchange model (lines). Although the data are shown for one residue only, a global two-state data fit was performed using the data for multiple residues recorded at all temperatures and at two magnetic fields. (B and C) Temperature dependences of the exchange rate constant, $k_{ex} = k_{N\rightarrow D} + k_{D\rightarrow N}$, and population of the denatured state, p_D, respectively, obtained from the global fit of temperature-dependent ^{15}N CPMG RD data. (D) Free energy profiles [free energy G (green), enthalpic contribution H (red), and entropic contribution TS (blue)] along the pathway from the denatured to the native state calculated from equilibrium and activation thermodynamic parameters derived from the global fit of temperature-dependent ^{15}N CPMG RD data.

time were fit by an exponent to extract per-residue ^{1}H/^{2}H exchange rates, R_{exp}. The obtained R_{exp} rates were used to calculate logarithms of protection factors for the backbone amide groups, $\log_{10}(R_{int}/R_{exp})$, where R_{int} values are intrinsic random-coil ^{1}H/^{2}H exchange rates calculated assuming pD 7.6 (corrected) using the FBMME_HD.xls Microsoft Excel spreadsheet available at http://hx2.med.upenn.edu.[71,72]

RESULTS

The activation domain of human procarboxypeptidase A2 (ADA2h) is a small 81-residue protein that forms two α-helices packed against a four-stranded β-sheet.[41] Equilibrium and kinetic unfolding−refolding studies have shown that wild-type ADA2h folds in a two-state manner with a rate of 758 s^{-1} and has an unfolding free energy of 4.4 kcal/mol at neutral pH and 25 °C.[42,44] Although equilibrium protein unfolding−refolding transitions that occur on the millisecond time scale are generally amenable to studies by CPMG RD methods, the thermodynamic stability of wild-type ADA2h places it beyond the limit of sensitivity for this type of experiment.[16−19] Therefore, we studied a destabilized I71V ADA2h mutant with a reported unfolding free energy of 2.9 kcal/mol and a refolding rate of 450 s^{-1} at neutral pH and 25 °C,[44] which populates a denatured state (i.e., the lowest energy non-native state under physiological conditions[73,74]) at the level of several percent and displays excellent quality CPMG RD profiles at temperatures above 35 °C (see below). For this mutational variant we performed a detailed CPMG RD study of the denatured state transiently populated at native-like condition that was proposed to mediate early events in amyloid formation at acidic pH.[46,47]

I71V ADA2h folds by a two-state mechanism. To characterize equilibrium unfolding/refolding transitions of I71V ADA2h we performed ^{15}N CPMG RD measurements[28] at magnetic fields of 14.1 and 18.8 T and at six different temperatures ranging from 35 to 50 °C. Figure 1A shows examples of the backbone ^{15}N CPMG RD profiles for residue Ile11 recorded over a range of temperatures at a magnetic field of 18.8 T. More than half of the I71V ADA2h residues exhibit pronounced ^{15}N CPMG RD profiles characteristic of a millisecond time-scale exchange process accompanied by changes in NMR chemical shifts. Villegas et al.[42] reported that at neutral pH wild-type ADA2h folds by a reversible two-state mechanism. To verify these results for the I71V ADA2h mutant, we have performed the global fit of ^{15}N CPMG RD data recorded for multiple residues at different temperatures and magnetic fields to a model of two-state exchange between the native state (N) and the denatured state (D), N ↔ D. As protein folding is a global process that involves all residues in a protein, in the case of two-state unfolding−refolding transitions, ^{15}N CPMG RD data for all amide groups in a protein should report on the same exchange process. Inconsistencies between exchange parameters extracted for individual residues using the two-state model and/or poor quality of global fit where the data for all residues are analyzed together may provide an indication that the folding process is more complex.[7,58] For example, global analyses of CPMG RD data measured for multiple protein residues at different magnetic fields under different experimental conditions allowed the detection of previously undiscovered folding intermediates for the Fyn and Abp1p SH3 domain mutants.[7,10,40]

Table 1 summarizes the results of global fits of the backbone ^{15}N CPMG RD data for I71V ADA2h measured at six temperatures (35−50 °C) and two magnetic fields (11.7 and 18.8 T). The top of the table shows cumulative χ^2 target functions (eq 1) obtained in the global two-state fits of ^{15}N CPMG RD data for all residues and magnetic fields performed individually for each temperature, and globally for all temperatures together (see Materials and Methods). Provided that the model adequately fits the data, the χ^2 target function obtained during the least-squares data fit should correspond to the number of degrees of freedom of the model, #DF = $n_{dat} - n_{par}$, where n_{dat} and n_{par} are the numbers of experimental data points and adjustable model parameters, respectively.[61] This holds well for the global two-state ^{15}N CPMG RD data fits performed for individual temperatures (χ^2/#DF ratio range from 0.5 at 35 °C to 1.4 at 50 °C), and for the global two-state fit where the data for all temperatures were analyzed together (χ^2/#DF ∼ 1.3). This result suggests that under our experimental conditions the I71V ADA2h unfolding−refolding transitions are well-described by the two-state exchange model.

Note that χ^2 values obtained in least-squares data fits (eq 1) are sensitive to the selection of uncertainties of experimental data points: overestimated (underestimated) uncertainties result in lower (higher) χ^2 values. To eliminate the possibility that artificially low χ^2 values were obtained due to overestimated experimental errors, we compared the cumulative χ^2 target functions obtained in the global two-state fits of ^{15}N CPMG RD data for all residues together (Table 1, column 3) with the sum of χ^2 ($\sum\chi^2$) obtained in individual data fits performed separately for each residue (Table 1, column 4). Typically, ^{15}N CPMG RD profiles for individual residues

DOI: 10.1021/acs.biochem.5b00345
Biochemistry 2015, 54, 4611−4622

recorded at one or several magnetic field strengths can be well fitted to a model of two-state exchange even in cases where the actual exchange process involves more than two states.[58] Therefore, for multistate processes (e.g., N ↔ I ↔ D, where I is the intermediate), one could expect a significant increase in the global χ^2 obtained when the data for multiple residues are analyzed together as compared to the sum of per-residue χ^2 values, $\sum\chi^2$.[7,58] The global χ^2 values obtained in two-state ^{15}N CPMG RD data fits for I71 V ADA2h, on the other hand, are in good agreement with $\sum\chi^2$, with the ratios between the two values not exceeding 1.3. These results provide additional evidence that I71V ADA2h folds by a two-state mechanism.

Thermodynamics and Kinetics of I71V ADA2h Folding. The global analysis of ^{15}N CPMG RD data allows extraction of accurate conformational exchange parameters that describe interconversions between the N and D states of I71V ADA2h, including population of the D state, p_D, and the exchange rate constant, $k_{ex} = k_{N\to D} + k_{D\to N}$, where $k_{N\to D}$ and $k_{D\to N}$ are the unfolding and refolding rates, respectively. These parameters are directly related to the free energy difference between the D and N states, ΔG_{DN}, and the activation free energy for the folding process, $\Delta G_{D\to N}^{\dagger}$, allowing reconstruction of the free energy profile along the folding pathway. In turn, the unfolding and activation free energies (ΔG_{DN} and $\Delta G_{D\to N}^{\dagger}$, respectively) consist of enthalpic (ΔH_{DN} and $\Delta H_{D\to N}^{\dagger}$) and entropic ($-T\Delta S_{DN}$ and $-T\Delta S_{D\to N}^{\dagger}$) contributions, which can be obtained from the analysis of ^{15}N CPMG RD data measured as a function of temperature.

Table 1 (columns 7 and 8) and panels B and C of Figure 1 show temperature dependences of p_D and k_{ex}, for I71V ADA2h obtained from ^{15}N CPMG RD data. In the temperature range of 35−50 °C, p_D increases gradually from 0.8 to 4.2% while k_{ex} changes from 700 to 1200 s^{-1}. The exchange rates and populations obtained from global fits of ^{15}N CPMG RD data performed separately at each temperature (Table 1, columns 7 and 8) are in good agreement with the values predicted from equilibrium (ΔH_{DN} and ΔS_{DN}) and activation ($\Delta H_{D\to N}^{\dagger}$ and $\Delta S_{D\to N}^{\dagger}$) thermodynamic parameters obtained in the global analysis of the data measured at multiple temperatures (Table 1, columns 7 and 8, in parentheses). This observation once again confirms that the I71V ADA2h unfolding−refolding interconversions are described well by the two-state model.

Figure 1D shows the free energy (G) profile as well as enthalpic (H) and entropic (TS) contributions to the free energy along the folding pathway of I71V ADA2h calculated at 40 °C. These profiles were generated on the basis of thermodynamic parameters (ΔH_{DN}, ΔS_{DN}, $\Delta H_{D\to N}^{\dagger}$, and $\Delta S_{D\to N}^{\dagger}$) extracted from the global fit of temperature-dependent ^{15}N CPMG RD data performed under the assumption that temperature dependences of the refolding and unfolding rates follow the Eyring equation (see Materials and Methods). The relatively small ΔG_{DN} 2.8 ± 0.1 kcal/mol for the I71V mutant results from a large mutually compensating enthalpic ΔH_{DN} 24.2 ± 0.1 kcal/mol and an entropic $T\Delta S_{DN}$ 21.4 ± 0.1 kcal/mol contributions (Figure 1D). The observed entropy−enthalpy compensation is characteristic of the protein unfolding−refolding transition,[73] resulting from a loss of configurational entropy of the polypeptide chain upon protein folding, opposed by an increase in solvent entropy due to protein dehydration and a decrease in chain enthalpy due to the formation of favorable contacts between protein groups. On the other hand, the folding barrier between the D and N states, $\Delta G_{D\to N}^{\dagger}$, is primarily enthalpic (Figure 1D), which may indicate that the rate-limiting step for ADA2h folding is the formation of a critical number of intramolecular contacts in the protein core. This is consistent with the results of Φ-value analysis,[44] suggesting that ADA2h folds via a compact transition state where α-helix 2 docks against the two central β-strands.

Backbone NMR Chemical Shifts of the Denatured State. Perhaps the most important type of parameter that can be extracted from CPMG RD data is the absolute value of the NMR chemical shift difference between the denatured (D) and native (N) states, $|\Delta\varpi_{DN}|$. Chemical shifts report on local protein structure and provide very sensitive probes of changes in the chemical environment of NMR active nuclei.[75,76] The backbone NMR chemical shifts are routinely used for protein secondary structure prediction at early steps of protein structure calculation.[66] Furthermore, in some cases, the backbone chemical shifts can be used to predict three-dimensional structure of small proteins and protein domains with accuracy comparable to that of the conventional approach.[77−79] Secondary chemical shifts (i.e., differences between the observed chemical shifts of the native state and predicted random-coil shifts) can be also used to assess the local backbone flexibility.[68] Although the resonances of the low populated denatured state are not observed directly in NMR spectra of I71V ADA2h, the resonance assignment of this state can be obtained by combining the experimental chemical shifts of the native state, ϖ_N, with chemical shift differences between the N and D states, $\Delta\varpi_{DN} = \varpi_D - \varpi_N$, obtained from CPMG RD data fits. Note that CPMG RD data analysis allows extraction of only the absolute values of chemical shift differences between states, $|\Delta\varpi_{DN}|$, with signs of $\Delta\varpi_{DN}$ need to be measured using additional NMR experiments.[62,63,80]

To obtain the backbone chemical shifts of the low populated D state, we performed a series CPMG RD experiments for the backbone ^{15}N,[28] ^{1}H^{N},[29] ^{13}C$^{\alpha}$,[32] ^{1}H$^{\alpha}$,[34] and ^{13}C′[32,33] nuclei using appropriately isotope-labeled I71V ADA2h samples. Figure 2 shows examples of CPMG RD profiles for the five types of nuclei recorded at a 18.8 T field and 40 °C, with additional examples of CPMG RD data shown in the Supporting Information. Note that CPMG RD experiments for different backbone nuclei require different protein isotope labeling schemes. Thus, CPMG RD measurements for the backbone ^{15}N, ^{1}H^{N}, and ^{13}C′ nuclei used to quantify chemical shifts of the D state were performed on the uniformly ^{15}N/^{13}C/^{2}H-labeled (amide-protonated) I71V ADA2h sample dissolved in 90% H_2O/10% D_2O buffer. On the other hand, CPMG experiments for ^{13}C$^{\alpha}$ and ^{1}H$^{\alpha}$ nuclei were performed using the two I71V ADA2h samples with different isotope labeling schemes dissolved in 100% D_2O buffer: ^{13}C$^{\alpha}$ CPMG experiments performed with selectively ^{13}C$^{\alpha}$-labeled fully protonated sample[32] and ^{1}H$^{\alpha}$ CPMG data collected for uniformly ^{15}N/^{13}C/partially-^{2}H-labeled protein.[34]

CPMG RD data were fit to a model of two-state exchange, D ↔ N, independently for each type of nucleus. The resulting exchange parameters are listed at the bottom of Table 1. As expected, the fully deuterated ^{15}N/^{13}C/^{2}H sample used for ^{15}N, ^{1}H^{N}, and ^{13}C′ CPMG RD measurements is less thermodynamically stable and exhibits a refolding rate slower than that of the ^{15}N-labeled protonated protein used to record temperature-dependent ^{15}N CPMG data (Table 1, bottom part vs top part). This reflects the fact that deuteration of aliphatic side-chain positions preferentially destabilizes structured protein states because of weakening of van der Waals interactions that involve

DOI: 10.1021/acs.biochem.5b00345
Biochemistry 2015, 54, 4611−4622

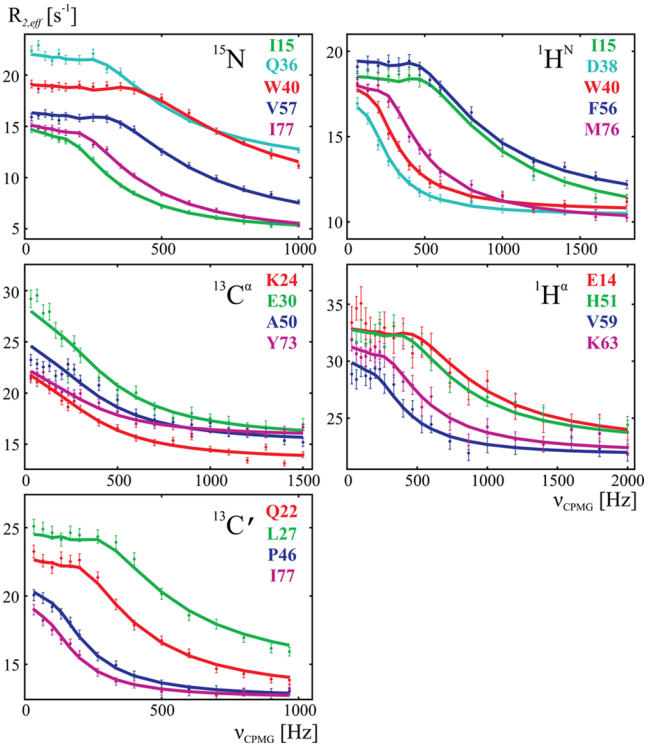


Figure 2. Examples of experimental CPMG relaxation dispersion profiles for selected ^{15}N, $^{1}H^{N}$, $^{13}C^{\alpha}$, $^{1}H^{\alpha}$, and $^{13}C'$ nuclei of I71V ADA2h measured at a 18.8 T magnetic field strength and 40 °C (circles) and their best fits to the two-state exchange model (lines).

deuterons as compared to protons.[81,82] On the other hand, the I71V ADA2h samples used for $^{13}C^{\alpha}$ and $^{1}H^{\alpha}$ measurements dissolved in 100% D_2O buffer exhibit refolding rates faster than those of samples studied in 90% H_2O/10% D_2O buffer. Previous studies addressed the effects of solvent on protein stability and in most cases reported moderate protein stabilization in D_2O as compared to H_2O, likely due to changes in the hydrophobic effect and hydrogen bonding.[83−85] Concomitant with this is an increase in the refolding rate as, for example, has been reported for the FF domain from HYPA/FBP11.[8]

The CPMG RD experiments described above, which were performed for the five types of backbone nuclei, allowed us to measure 77% of all possible backbone chemical shift differences between the D and N states of I71V ADA2h, $|\Delta\varpi_{DN}|$, with signs of $\Delta\varpi_{DN}$ obtained from the comparison of peak positions in HSQC and HMQC spectra recorded at two magnetic field strengths[62,63] (see Materials and Methods). In 93% of cases, the experimentally determined $\Delta\varpi_{DN}$ signs coincided with signs of $\varpi_{rc} - \varpi_N$, where ϖ_{rc} values are predicted random-coil chemical shifts. In cases where $\Delta\varpi_{DN}$ signs cannot be determined experimentally, we assumed that $\Delta\varpi_{DN}$ values have the same sign as $\varpi_{rc} - \varpi_N$. The full list of chemical shift differences $\Delta\varpi_{DN}$ and the backbone chemical shifts of the D state of I71V ADA2h calculated as $\varpi_D = \varpi_N + \Delta\varpi_{DN}$ can be found in the Supporting Information.

Residual Structure and Flexibility of the Denatured State from Chemical Shifts. A nearly complete set of backbone ^{15}N, $^{1}H^{N}$, $^{13}C^{\alpha}$, $^{1}H^{\alpha}$, and $^{13}C'$ chemical shifts of the transiently populated D state of I71V ADA2h (see the Supporting Information) allowed us to access its residual structure and the backbone flexibility. Figure 3A shows the

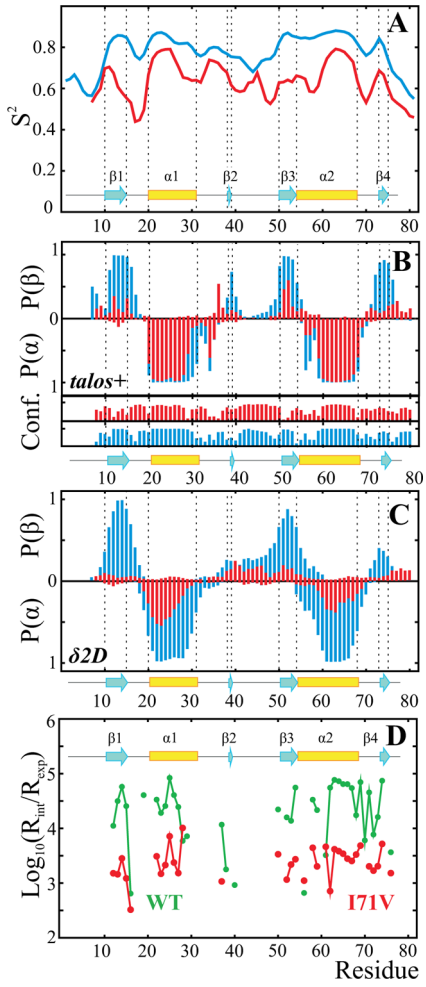


Figure 3. (A−C) Secondary structure populations and backbone dynamics of the native (blue lines) and denatured (red lines) states of I71V ADA2h. (A) Order parameters S^2 for the backbone amide groups of I71V ADA2h predicted from the backbone ^{15}N, $^{1}H^{N}$, $^{13}C^{\alpha}$, $^{1}H^{\alpha}$, and $^{13}C'$ chemical shifts using the RCI approach.[68,69] (B and C) Probabilities of secondary structure in the native (blue bars) and denatured (red bars) states of I71V ADA2h predicted from the backbone chemical shifts using (B) TALOS+[66] (confidence levels for TALOS+ predictions are shown at the bottom) and (C) δ2D.[67] (D) Logarithms of protection factors for the backbone amide protons of wild-type (green circles) and I71V (red circles) ADA2h, $\log_{10}(R_{int}/R_{exp})$, where R_{exp} values are experimental $^{1}H \rightarrow {}^{2}H$ exchange rates and R_{int} values are intrinsic random-coil $^{1}H \rightarrow {}^{2}H$ exchange rates calculated using the FBMME_HD.xls script downloaded from http://hx2.med.upenn.edu.[71,72] The secondary structure shown in each panel corresponds to that in the solution NMR structure of wild-type ADA2h (Protein Data Bank entry 1O6X).[41]

comparison of order parameters S^2 for the backbone amide groups of the D (red line) and N (blue line) states of I71V ADA2h predicted from the backbone chemical shifts using the random-coil index (RCI) approach.[68,69] $S^2 = 1(0)$ corresponds to a fully constrained (unrestricted) angular motion of the NH vector. As expected, predicted S^2 values for the N state reveal restricted dynamics for all secondary structure elements that are present in the NMR structure of wild-type ADA2h[41] (shown at the bottom of Figure 3A). The D state of I71V ADA2h is considerably more flexible. Notably, there are several regions of the D state with a lower predicted amplitude of backbone motions ($S^2 > 0.7$), including the N-terminal part of native α-

DOI: 10.1021/acs.biochem.5b00345
Biochemistry 2015, 54, 4611−4622

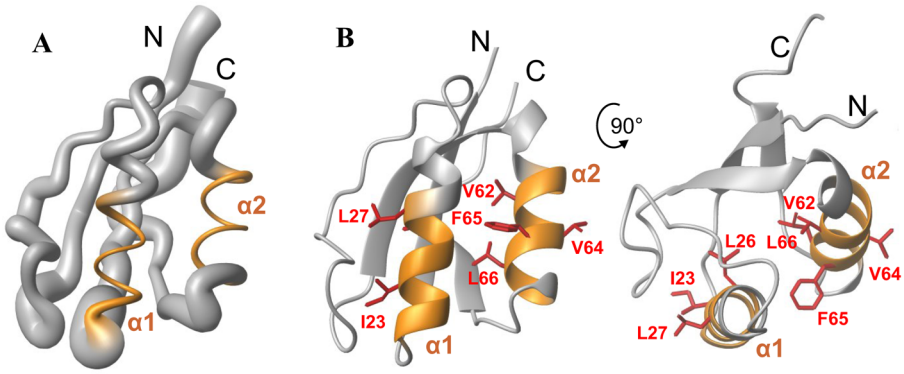


Figure 4. Residual structure and backbone dynamics of the denatured state of I71V ADA2h mapped onto the NMR structure of wild-type native ADA2h (Protein Data Bank entry 1O6X).[41] (A) Ribbon diagram of the ADA2h structure with the ribbon width proportional to the backbone amide order parameters $1 - S^2$ calculated using the RCI approach.[68,69] (B) Regions of the denatured state with a δ2D-predicted secondary structure population of >20% (residues 20−27 and 60−67) color mapped onto the structure of wild-type ADA2h.[41] These regions match those with TALOS +-predicted probabilities of finding residues in an α-helical conformation of >70% (20−28 and 60−67). Side chains of hydrophobic residues in this region (I23, L26, L27, A61, V62, V64, F65, and L66) that potentially form a cluster in the denatured state are highlighted.

helices 1 (residues 21−27), the C-terminal part of native α-helix 2 (residues 61−67), and an additional short stretch (residues 35−37) from the native loop between α-helix 1 and β-strand 2. The rest of the I71V ADA sequence, including regions of the native β-strands, is somewhat more dynamic in the D state, but there are few local increases in order parameters in more dynamic parts of the D state. To illustrate the backbone flexibility of the D state, Figure 4A shows a ribbon diagram of the wild-type ADA2h structure with the ribbon width proportional to $1 - S^2$ (a thick ribbon corresponds to more dynamic regions).

To further characterize the D state of I71V ADA2h, we calculated its secondary structure populations from the CPMG RD-derived backbone chemical shifts. Figure 3B shows probabilities of the I71V ADA2h residues adopting extended β-stand and α-helical conformations in the N (blue bars) and D (red bars) states calculated from the chemical shifts using TALOS+.[66] Predicted regions with high secondary structure probability in the D state (>70%) are in good agreement with the regions of restricted backbone mobility (Figure 3A), including parts of native α-helices 1 and 2 (residues 20−28 and 60−67, respectively). The exception is the stretch of residues 35−37 in which no secondary structure can be predicted with sufficient confidence (Figure 3A). On the other hand, no regions with significant β-structure population can be predicted with sufficient confidence in the D state of I71V ADA2h, including regions corresponding to native β-strands 1−4 (Figure 3A,B).

Figure 3C shows the secondary structure probability in the N (blue bars) and D (red bars) states of I71V ADA2h predicted from the chemical shifts using the δ2D program.[67] δ2D is particularly suitable for determining secondary structure populations of disordered proteins as it was trained and tested not only on native protein states but also on denatured, intermediate, and intrinsically disordered states.[67] Similar to TALOS+ prediction, δ2D identifies two regions within the D state of I71V ADA2h (residues 20−28 and 60−67) with a considerable (>20%) population of α-helical structure (Figures 3C and 4B). Within the two α-helices, residues 21−25, 62, and 63 showed an α-helical population of more than 40%. However, unlike secondary structure probabilities predicted by TALOS+,[66] the populations of the two α-helices in the denatured state determined by δ2D were only fractional and considerably lower than those in the native state.

The results obtained using δ2D[67] are consistent with relatively low protection factors measured for the two helical regions of I71V ADA2h (see below) and the decrease in their order parameters S^2 relative to those of the N state (Figure 3A). On the other hand, the population of β-structure in the D state was predicted to be considerably lower than in the N state. In particular, δ2D identified three regions of the D state corresponding to native β-strands 2−4 with the β-strand population above 10% (residues 38−46, 50−53, and 73−80). The polyproline II (PPII) population was found to be negligible in the D state. These results are also consistent with the intrinsic tendencies of the different regions of ADA2h to form secondary structures. In the D state, the Agadir sequence-based prediction indicates a 27% α-helix content for α-helix 1.[86] Using the s2D sequence-based predictions,[70] we obtain an approximately 50−70% population for α-helix 1, an approximately 30−50% population for α-helix 2, and an approximately 10−40% population for the region of β-strands 2 and 3. Overall, the measured chemical shifts imply that the D state retains considerable population of α-helices 1 and 2 present in the N state, whereas the β-strands showed an only minor population.

^{1}H/^{2}H Exchange in the Denatured State of Wild-Type and I71V ADA2h. ^{1}H/^{2}H exchange monitored by NMR provides additional information about the solvent exposure of the backbone amide groups in the low populated D state of ADA2h.[52] ^{1}H/^{2}H exchange rates, R_{exp}, can be calculated from exponential fits of peak intensities in a series of ^{1}H−^{15}N HSQC spectra recorded after lyophilized protein is dissolved in 100% D_2O. The measured R_{exp} rates can be used to estimate protection factors for the backbone amide groups in the D state, R_{int}/R_{exp}, where R_{int} values are predicted intrinsic random-coil ^{1}H/^{2}H exchange rates[71,72] (see Materials and Methods). Figure 3D shows logarithms of protection factors, R_{int}/R_{exp}, for the backbone amide groups of wild-type and I71V ADA2h. As expected, we were able to measure protection factors for only the regions that participate in regular secondary structure elements of the N state. The protection factors for wild-type and I71V ADA2h forms exhibit similar behavior, reflecting the fact that the protein transiently populates a low populated D state where the backbone amides are primarily solvent exposed.

DOI: 10.1021/acs.biochem.5b00345
Biochemistry 2015, 54, 4611−4622

As expected, uniformly higher protection factors, R_{int}/R_{exp}, were measured for the wild type (green circles) as compared to I71V ADA2h (red circles), which reflects the higher thermodynamic stability of the wild-type protein. On the other hand, no distinct protein regions with elevated protection factors were observed, pointing to a lack of significant protection from solvent exposure in the D state and suggesting that residual structure of this state is only marginally stable under our experimental conditions. This is consistent with δ2D[67] prediction of secondary structure probability in the D state from CPMG RD-derived chemical shifts (Figure 3C), which suggests that secondary structure elements in the D state of I71V ADA2h are populated only fractionally.

From the $^{1}H/^{2}H$ exchange data for the 10 most protected amide groups, we can estimate free energy differences between the D and N states, ΔG_{DN}, of 6.5 ± 0.1 and 4.9 ± 0.2 kcal/mol for wild-type (WT) and I71V ADA2h, respectively (20 °C, pD 7.6). To compare these values with those obtained by denaturation methods, the measured ΔG_{DN} values need to be corrected by −0.5 kcal/mol to take into account the effect of four proline residues in ADA2h,[41,87] resulting in ΔG_{DN} values of 6.0 ± 0.1 and 4.4 ± 0.2 kcal/mol for WT and I71V ADA2h, respectively. The ΔG_{DN} value measured for I71V ADA2h from $^{1}H/^{2}H$ exchange data is in reasonable agreement with the ΔG_{DN} of 4.2 kcal/mol calculated from the ΔH_{DN} of 24.2 kcal/mol and the ΔS_{DN} of 68.4 cal mol^{-1} K^{-1} obtained for this mutant from the analysis of temperature-dependent ^{15}N CPMG RD data (see above). The destabilizing effect of the I71V mutation of 1.6 kcal/mol obtained from our $^{1}H/^{2}H$ exchange data is also consistent with 1.5 kcal/mol destabilization reported for this mutant by Villegas et al.[44] The somewhat lower stability for WT and I71V ADA2h at 25 °C and pH 7.0 reported in previous studies[41,44] could be attributed to the different experimental conditions used in our work.

DISCUSSION

We have investigated the equilibrium unfolding−refolding transition of the I71V mutant of ADA2h using CPMG relaxation dispersion (RD) experiments.[16−19] We have found that the temperature-dependent ^{15}N CPMG RD profiles (Figure 1A) can be well fitted to a model of two-state exchange between the ground native state (N) and the low populated denatured state (D), D ↔ N, occurring on the millisecond time scale (Table 1 and Figure 1B,C). The two exchanging states are separated by a free energy barrier that is primarily enthalpic (Figure 2D). The conformational exchange parameters determined over a range of temperatures are listed in Table 1. Thus, at 40 °C, unfolding−refolding interconversions of ^{15}N-labeled fully protonated I71V ADA2h monitored by ^{15}N CPMG RD measurements occur with the rate constant $k_{ex} = k_{N\to D} + k_{D\to N} = 932.2 \pm 9.5\ s^{-1}$ and involve a transition to the denatured state with equilibrium population $p_D = 1.23 \pm 0.01\%$.

To characterize the residual structure and flexibility of the low populated D state, we have preformed a series of CPMG RD experiments for the backbone ^{15}N,[28] $^{1}H^{N}$,[29] $^{13}C^{\alpha}$,[32] $^{1}H^{\alpha}$,[34] and $^{13}C'$[32,33] nuclei using the I71V ADA2h samples with appropriate isotope labeling (Figure 2). The data were analyzed to extract chemical shift differences between the D and N states, $\Delta\varpi_{DN}$, that were subsequently used to obtain the backbone resonance assignment of the low populated D state (see the Supporting Information). The obtained nearly complete set of the backbone chemical shifts has allowed us to identify regions of the D state with restricted mobility (predicted using the RCI program[68,69]) (Figures 3A and 4A) and increased likelihood of secondary structure formation (predicted using TALOS+[66] and δ2D[67]) (Figure 3B,C). We have shown that the D state of I71V ADA2h has extensive regions of residual structure and restricted mobility encompassing the first two turns of native α-helix 1 (residues 21−27) and the last two turns of native α-helix 2 (residues 61−67). On the other hand, the regions corresponding to native β-strands 1−4 as well as the rest of the I71V ADA2h sequence are more disordered in the D state.

Villegas et al.[44] have previously reported the Φ-value analysis of the ADA2h folding that revealed a compact transition state, which has a nativelike secondary structure in the region encompassing the two central β-strands 1 and 3 and α-helix 2 and a hydrophobic core formed around residue V52 (from β-strand 3). On the basis of sparse Φ-value data obtained for native α-helix 1, the authors concluded that this helix does not contribute significantly to the stability of the ADA2h folding nucleus. This conclusion is supported by the fact that an isolated peptide comprising isolated α-helix 1 is soluble and partially folded without the need for stabilization by the rest of the protein[86] and that stabilization of α-helix 1 by N25K/Q28E/Q32K/E33K mutations does not lead to acceleration of ADA2h refolding.[43] These observations, however, do not exclude the possibility that some residues involved in α-helix 1 (such as I23 and L26) could participate in tertiary contacts in the transition state.[44] In contrast to α-helix 1, stabilization of α-helix 2 by Q60E/V64A/S68A/Q69H mutations results in a 4-fold acceleration of the refolding rate,[43] suggesting a key role for α-helix 2 in stabilization of the ADA2h transition state.

Our results suggest that the two N-terminal turns of α-helix 1 (of a total three) spanning residues 21−27 and the two turns of α-helix 2 (of a total of four) spanning residues 61−67 exhibit restricted mobility (Figures 3A and 4A) and are at least partially formed in the D state of I71V ADA2h (Figures 3B,C and 4). Hydrophobic side chains in a partially structured region of the D state, including nativelike α-helices 1 and 2 (I23, L26, L27, A61, V62, V64, F65, and L66), are grouped on the same sides of the helices and thus become available for tertiary interactions (Figure 4B). Presumably, these residues may form a hydrophobic cluster in the D state with additional participation of L35 and L37 from the native α-helix 1−β-strand 2 loop that display restricted mobility in the D state (Figures 3A and 4A) and/or other hydrophobic residues from the native β-strand regions. Conceivably, incorporation of additional hydrophobic side chains into a hydrophobic cluster formed by residues on the α-helix surfaces may induce transient formation of β-structure, which is known to be primarily dependent on the tertiary context.[88] Modulation of the stability of such a cluster, e.g., with appropriate mutations in the α-helices,[86] may facilitate efficient protein refolding[43] as well as favor intra- versus intermolecular interactions and thus reduce the tendency for ADA2h variants to undergo aggregation.[46]

ADA2h was shown to form well-ordered amyloid fibrils following thermal or chemical denaturation at pH 3.0.[46,47] Interestingly, amyloid formation by the ADA2h mutants was found to be uncorrelated with the overall stability of the native state.[46,47] On the other hand, mutations that stabilize one or both α-helices[86] significantly reduce the ADA2h propensity for amyloid formation.[46,47] These findings suggest that stabilization of local intramolecular interactions in the ADA2h D state may help prevent β-aggregation. Cerda-Costa et al.[47] reported an

DOI: 10.1021/acs.biochem.5b00345
Biochemistry 2015, 54, 4611−4622

extensive kinetic study of amyloid formation by the ADA2h mutants (including I71V) and proposed a model in which the ADA2h aggregation proceeds in three steps: (i) formation of soluble aggregates from the denatured state (occurring with rates of 0.40 s^{-1} for WT ADA2h and 0.07 s^{-1} for the I71V mutant), (ii) slower rearrangement and growth of the soluble aggregates (occurring with rates of 0.04 s^{-1} for the WT and 0.02 s^{-1} for the I71V mutant), and (iii) a final extremely slow step of conversion of the soluble aggregate into amyloid fibrils. Sequence specific predictions and comparison of the aggregation behavior of 27 ADA2h mutants led to the conclusion that β-strand 4 is the major contributor to amyloid formation.[47] In line with this conclusion, mutations of residues from native β-strand 4 resulted in significantly impaired amyloid formation. Remarkably, mutations on the outer surface of α-helix 2 (V64G, F65A, and E67K) also led to impaired amyloid formation similar to mutations in native β-strand 4. These findings imply that aggregation tendencies of the ADA2h mutants are at least in part determined by the residual structure of the D state. Our results suggest that the D state of I71V ADA2h retains a significant residual structure in the regions encompassing native α-helices 1 and 2. On the other hand, C-terminal β-strand 4 that was reported to be the major contributor to amyloid formation[47] is largely disordered in the D state. This finding implies that the primary event in amyloid formation by ADA2h may involve aggregation of a partially structured D state lacking significant β-structure followed by rearrangements of initial aggregates.

CONCLUSIONS

We have examined the residual structure of the low populated denatured state of the ADA2h I71V mutant using CPMG RD experiments, which have allowed us to obtain a nearly complete set of backbone chemical shifts in this state. The analysis of the denatured state chemical shifts revealed the presence of regions of restricted mobility and of significant residual structure corresponding to native α-helices 1 and 2, and to a much lesser degree to the central β-strands. These structural features create the possibility of forming a hydrophobic cluster in the denatured state that involves aliphatic residues grouped on the same sides of the α-helices with possible integration of additional hydrophobic side chains from other regions of the ADA2h sequence. Conceivably, modulation of the stability of this cluster with amino acid mutations may favor intra- versus intermolecular interactions (or vice versa) and thus decelerate (or accelerate) protein aggregation.

Denatured protein states transiently populated under nativelike conditions so far have been recalcitrant to experimental characterization. The nearly complete set of CPMG RD-derived backbone chemical shifts obtained in this work represents one of the first experimental data sets that allow the characterization of a conformationally heterogeneous denatured state under folding conditions that provides an important benchmark for computational modeling of conformational ensembles of these elusive denatured states.

ASSOCIATED CONTENT

Supporting Information

^{15}N, $^{1}H^{N}$, and $^{13}C'$ chemical shifts of the native, N, and denatured, D, states of I71V ADA2h measured using the uniformly $^{15}N/^{13}C/^{2}H$-labeled sample (Table S1), $^{13}C^{\alpha}$ and $^{1}H^{\alpha}$ chemical shifts of the native, N, and denatured, D, states of I71V ADA2h measured using uniformly ^{15}N/selectively-$^{13}C^{\alpha}$-labeled and uniformly $^{15}N/^{15}C$/partially-^{2}H-labeled samples, respectively (Table S2), additional examples of experimental CPMG relaxation dispersion profiles (Figure S1), and five additional references. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.biochem.5b00345.

AUTHOR INFORMATION

Corresponding Author

*E-mail: korzhniev@uchc.edu. Phone: (860) 679-2849. Fax: (860) 679-3408.

Funding

This work was supported by UCHC startup funds and a CNPq Science without Boarders grant to D.M.K.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Dill, K. A., and Chan, H. S. (1997) From Levinthal to pathways to funnels. *Nat. Struct. Biol. 4*, 10−19.
(2) Jahn, T. R., and Radford, S. E. (2005) The Yin and Yang of protein folding. *FEBS J. 272*, 5962−5970.
(3) Mittag, T., and Forman-Kay, J. D. (2007) Atomic-level characterization of disordered protein ensembles. *Curr. Opin. Struct. Biol. 17*, 3−14.
(4) Brockwell, D. J., and Radford, S. E. (2007) Intermediates: ubiquitous species on folding energy landscapes? *Curr. Opin. Struct. Biol. 17*, 30−37.
(5) Englander, S. W., Mayne, L., and Krishna, M. M. G. (2007) Protein folding and misfolding: mechanism and principles. *Q. Rev. Biophys. 40*, 287−326.
(6) Vendruscolo, M., and Dobson, C. M. (2013) Structural biology: Protein self-assembly intermediates. *Nat. Chem. Biol. 9*, 216−217.
(7) Korzhnev, D. M., Salvatella, X., Vendruscolo, M., Di Nardo, A. A., Davidson, A. R., Dobson, C. M., and Kay, L. E. (2004) Low-populated folding intermediates of Fyn SH3 characterized by relaxation dispersion NMR. *Nature 430*, 586−590.
(8) Korzhnev, D. M., Religa, T. L., Banachewicz, W., Fersht, A. R., and Kay, L. E. (2010) A Transient and Low-Populated Protein-Folding Intermediate at Atomic Resolution. *Science 329*, 1312−1316.
(9) De Simone, A., Aprile, F. A., Dhulesia, A., Dobson, C. M., and Vendruscolo, M. (2015) Structure of a low-population intermediate state in the release of an enzyme product. *eLife 4*, DOI: 10.7554/eLife.02777.
(10) Neudecker, P., Robustelli, P., Cavalli, A., Walsh, P., Lundstroem, P., Zarrine-Afsar, A., Sharpe, S., Vendruscolo, M., and Kay, L. E. (2012) Structure of an Intermediate State in Protein Folding and Aggregation. *Science 336*, 362−366.
(11) Korzhnev, D. M., Religa, T. L., and Kay, L. E. (2012) Transiently populated intermediate functions as a branching point of the FF domain folding pathway. *Proc. Natl. Acad. Sci. U. S. A. 109*, 17777−17782.
(12) Dobson, C. M. (2003) Protein folding and misfolding. *Nature 426*, 884−890.
(13) Powers, E. T., Morimoto, R. I., Dillin, A., Kelly, J. W., and Balch, W. E. (2009) Biological and chemical approaches to diseases of proteostasis deficiency. *Annu. Rev. Biochem. 78*, 959−991.
(14) Chiti, F., and Dobson, C. M. (2006) Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem. 75*, 333−366.
(15) Knowles, T. P. J., Vendruscolo, M., and Dobson, C. M. (2014) The amyloid state and its association with protein misfolding diseases. *Nat. Rev. Mol. Cell Biol. 15*, 496−496.
(16) Palmer, A. G., 3rd, Kroenke, C. D., and Loria, J. P. (2001) Nuclear magnetic resonance methods for quantifying microsecond-to-

DOI: 10.1021/acs.biochem.5b00345
Biochemistry 2015, 54, 4611−4622

millisecond motions in biological macromolecules. *Methods. Enzymol. 339*, 204−238.
(17) Korzhnev, D. M., and Kay, L. E. (2008) Probing invisible, low-populated states of protein molecules by relaxation dispersion NMR spectroscopy: An application to protein folding. *Acc. Chem. Res. 41*, 442−451.
(18) Neudecker, P., Lundstrom, P., and Kay, L. E. (2009) Relaxation Dispersion NMR Spectroscopy as a Tool for Detailed Studies of Protein Folding. *Biophys. J. 96*, 2045−2054.
(19) Palmer, A. G., 3rd (2014) Chemical exchange in biomacromolecules: past, present, and future. *J. Magn. Reson. 241*, 3−17.
(20) Palmer, A. G., and Massi, F. (2006) Characterization of the dynamics of biomacromolecules using rotating-frame spin relaxation NMR spectroscopy. *Chem. Rev. 106*, 1700−1719.
(21) Carr, H. Y., and Purcell, E. M. (1954) Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Phys. Rev. 94*, 630−638.
(22) Meiboom, S., and Gill, D. (1958) Modified spin-echo method for measuring nuclear relaxation times. *Rev. Sci. Instrum. 29*, 688−691.
(23) Bouvignies, G., and Kay, L. E. (2012) Measurement of proton chemical shifts in invisible states of slowly exchanging protein systems by chemical exchange saturation transfer. *J. Phys. Chem. B 116*, 14311−14317.
(24) Hansen, A. L., Bouvignies, G., and Kay, L. E. (2013) Probing slowly exchanging protein systems via $^{13}C^{\alpha}$ CEST: monitoring folding of the Im7 protein. *J. Biomol. NMR 55*, 279−289.
(25) 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*, 8148−8161.
(26) Loria, J. P., Rance, M., and Palmer, A. G. (1999) A relaxation-compensated Carr-Purcell-Meiboom-Gill sequence for characterizing chemical exchange by NMR spectroscopy. *J. Am. Chem. Soc. 121*, 2331−2332.
(27) Tollinger, M., Skrynnikov, N. R., Mulder, F. A. A., Forman-Kay, J. D., and Kay, L. E. (2001) Slow dynamics in folded and unfolded states of an SH3 domain. *J. Am. Chem. Soc. 123*, 11341−11352.
(28) Hansen, D. F., Vallurupalli, P., and Kay, L. E. (2008) An improved ^{15}N relaxation dispersion experiment for the measurement of millisecond time-scale dynamics in proteins. *J. Phys. Chem. B 112*, 5898−5904.
(29) Ishima, R., and Torchia, D. A. (2003) Extending the range of amide proton relaxation dispersion experiments in proteins using a constant-time relaxation-compensated CPMG approach. *J. Biomol. NMR 25*, 243−248.
(30) Orekhov, V. Y., Korzhnev, D. M., and Kay, L. E. (2004) Double- and zero-quantum NMR relaxation dispersion experiments sampling millisecond time scale dynamics in proteins. *J. Am. Chem. Soc. 126*, 1886−1891.
(31) Ishima, R., Baber, J., Louis, J. M., and Torchia, D. A. (2004) Carbonyl carbon transverse relaxation dispersion measurements and ms-μs timescale motion in a protein hydrogen bond network. *J. Biomol. NMR 29*, 187−198.
(32) Hansen, D. F., Vallurupalli, P., Lundström, P., Neudecker, P., and Kay, L. E. (2008) Probing chemical shifts of invisible states of proteins with relaxation dispersion NMR spectroscopy: How well can we do? *J. Am. Chem. Soc. 130*, 2667−2675.
(33) Lundström, P., Hansen, D. F., and Kay, L. E. (2008) Measurement of carbonyl chemical shifts of excited protein states by relaxation dispersion NMR spectroscopy: comparison between uniformly and selectively ^{13}C labeled samples. *J. Biomol. NMR 42*, 35−47.
(34) Lundstrom, P., Hansen, D. F., Vallurupalli, P., and Kay, L. E. (2009) Accurate measurement of alpha proton chemical shifts of excited protein states by relaxation dispersion NMR spectroscopy. *J. Am. Chem. Soc. 131*, 1915−1926.
(35) Mulder, F. A., Skrynnikov, N. R., Hon, B., Dahlquist, F. W., and Kay, L. E. (2001) Measurement of slow (μs-ms) time scale dynamics in protein side chains by ^{15}N relaxation dispersion NMR spectroscopy: application to Asn and Gln residues in a cavity mutant of T4 lysozyme. *J. Am. Chem. Soc. 123*, 967−975.
(36) Skrynnikov, N. R., Mulder, F. A. A., Hon, B., Dahlquist, F. W., and Kay, L. E. (2001) Probing slow time scale dynamics at methyl-containing side chains in proteins by relaxation dispersion NMR measurements: Application to methionine residues in a cavity mutant of T4 lysozyme. *J. Am. Chem. Soc. 123*, 4556−4566.
(37) Korzhnev, D. M., Kloiber, K., Kanelis, V., Tugarinov, V., and Kay, L. E. (2004) Probing slow dynamics in high molecular weight proteins by methyl-TROSY NMR spectroscopy: Application to a 723-residue enzyme. *J. Am. Chem. Soc. 126*, 3964−3973.
(38) Hansen, A. L., and Kay, L. E. (2011) Quantifying millisecond time-scale exchange in proteins by CPMG relaxation dispersion NMR spectroscopy of side-chain carbonyl groups. *J. Biomol. NMR 50*, 347−355.
(39) Hansen, A. L., Lundstrom, P., Velyvis, A., and Kay, L. E. (2012) Quantifying Millisecond Exchange Dynamics in Proteins by CPMG Relaxation Dispersion NMR Using Side-Chain ^{1}H Probes. *J. Am. Chem. Soc. 134*, 3178−3189.
(40) Korzhnev, D. M., Neudecker, P., Zarrine-Afsar, A., Davidson, A. R., and Kay, L. E. (2006) Abp1p and Fyn SH3 Domains Fold Through Similar Low-Populated Intermediate States. *Biochemistry 45*, 10175−10183.
(41) Jimenez, M. A., Villegas, V., Santoro, J., Serrano, L., Vendrell, J., Aviles, F. X., and Rico, M. (2003) NMR solution structure of the activation domain of human procarboxypeptidase A2. *Protein Sci. 12*, 296−305.
(42) Villegas, V., Azuaga, A., Catasus, L., Reverter, D., Mateo, P. L., Aviles, F. X., and Serrano, L. (1995) Evidence for a two-state transition in the folding process of the activation domain of human procarboxypeptidase A2. *Biochemistry 34*, 15105−15110.
(43) Viguera, A. R., Villegas, V., Aviles, F. X., and Serrano, L. (1997) Favourable native-like helical local interactions can accelerate protein folding. *Folding Des. 2*, 23−33.
(44) Villegas, V., Martinez, J. C., Aviles, F. X., and Serrano, L. (1998) Structure of the transition state in the folding process of human procarboxypeptidase A2 activation domain. *J. Mol. Biol. 283*, 1027−1036.
(45) Fernandez, A. M., Villegas, V., Martinez, J. C., Van Nuland, N. A., Conejero-Lara, F., Aviles, F. X., Serrano, L., Filimonov, V. V., and Mateo, P. L. (2000) Thermodynamic analysis of helix-engineered forms of the activation domain of human procarboxypeptidase A2. *Eur. J. Biochem. 267*, 5891−5899.
(46) Villegas, V., Zurdo, J., Filimonov, V. V., Aviles, F. X., Dobson, C. M., and Serrano, L. (2000) Protein engineering as a strategy to avoid formation of amyloid fibrils. *Protein Sci. 9*, 1700−1708.
(47) Cerda-Costa, N., Esteras-Chopo, A., Aviles, F. X., Serrano, L., and Villegas, V. (2007) Early kinetics of amyloid fibril formation reveals conformational reorganisation of initial aggregates. *J. Mol. Biol. 366*, 1351−1363.
(48) Linding, R., Schymkowitz, J., Rousseau, F., Diella, F., and Serrano, L. (2004) A comparative study of the relationship between protein structure and beta-aggregation in globular and intrinsically disordered proteins. *J. Mol. Biol. 342*, 345−353.
(49) Fernandez-Escamilla, A. M., Rousseau, F., Schymkowitz, J., and Serrano, L. (2004) Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins. *Nat. Biotechnol. 22*, 1302−1306.
(50) Pawar, A. P., Dubay, K. F., Zurdo, J., Chiti, F., Vendruscolo, M., and Dobson, C. M. (2005) Prediction of "aggregation-prone" and "aggregation-susceptible" regions in proteins associated with neurodegenerative diseases. *J. Mol. Biol. 350*, 379−392.
(51) Tartaglia, G. G., and Vendruscolo, M. (2008) The Zyggregator method for predicting protein aggregation propensities. *Chem. Soc. Rev. 37*, 1395−1401.
(52) Englander, S. W., Mayne, L., Bai, Y., and Sosnick, T. R. (1997) Hydrogen exchange: The modern legacy of Linderstrom-Lang. *Protein Sci. 6*, 1101−1109.

DOI: 10.1021/acs.biochem.5b00345
Biochemistry 2015, 54, 4611−4622

(53) Lundström, P., Teilum, K., Carstensen, T., Bezsonova, I., Wiesner, S., Hansen, F., Religa, T. L., Akke, M., and Kay, L. E. (2007) Fractional 13C enrichment of isolated carbons using [1-^{13}C]- or [2-^{13}C]-glucose facilitates the accurate measurement of dynamics at backbone Cα and side-chain methyl positions in proteins. *J. Biomol. NMR 38*, 199−212.
(54) Sattler, M., Schleucher, J., and Griesinger, C. (1999) Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Magn. Reson. Spectrosc. 34*, 93−158.
(55) Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPIPE - a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR 6*, 277−293.
(56) Korzhnev, D. M., Ibraghimov, I. V., Billeter, M., and Orekhov, V. Y. (2001) MUNIN: Application of three-way decomposition to the analysis of heteronuclear NMR relaxation data. *J. Biomol. NMR 21*, 263−268.
(57) Orekhov, V. Y., Ibraghimov, I. V., and Billeter, M. (2001) MUNIN: A new approach to multi-dimensional NMR spectra interpretation. *J. Biomol. NMR 20*, 49−60.
(58) Korzhnev, D. M., Neudecker, P., Mittermaier, A., Orekhov, V. Y., and Kay, L. E. (2005) Multiple-site exchange in proteins studied with a suite of six NMR relaxation dispersion experiments: An application to the folding of a Fyn SH3 domain mutant. *J. Am. Chem. Soc. 127*, 15602−15611.
(59) Hagen, S. J., Hofrichter, J., Szabo, A., and Eaton, W. A. (1996) Diffusion-limited contact formation in unfolded cytochrome c: Estimating the maximum rate of protein folding. *Proc. Natl. Acad. Sci. U. S. A. 93*, 11615−11617.
(60) Press, W. H. (1992) *Numerical recipes in C: The art of scientific computing*, 2nd ed., Cambridge University Press, Cambridge, U.K.
(61) Zar, J. H. (1984) *Biostatistical analysis*, 2nd ed., Prentice-Hall, Englewood Cliffs, NJ.
(62) Skrynnikov, N. R., Dahlquist, F. W., and Kay, L. E. (2002) Reconstructing NMR spectra of "invisible" excited protein states using HSQC and HMQC experiments. *J. Am. Chem. Soc. 124*, 12352−12360.
(63) Auer, R., Hansen, D. F., Neudecker, P., Korzhnev, D. M., Muhandiram, D. R., Konrat, R., and Kay, L. E. (2010) Measurement of signs of chemical shift differences between ground and excited protein states: a comparison between H(S/M)QC and $R_{1\rho}$ methods. *J. Biomol. NMR 46*, 205−216.
(64) Johnson, B. A., and Blevins, R. A. (1994) NMRView: A computer program for the visualization and analysis of NMR data. *J. Biomol. NMR 4*, 603−614.
(65) Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995) ^{1}H, ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids: I. Investigations of nearest-neighbor effects. *J. Biomol. NMR 5*, 67−81.
(66) Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009) TALOS plus: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR 44*, 213−223.
(67) Camilloni, C., De Simone, A., Vranken, W. F., and Vendruscolo, M. (2012) Determination of Secondary Structure Populations in Disordered States of Proteins Using Nuclear Magnetic Resonance Chemical Shifts. *Biochemistry 51*, 2224−2231.
(68) Berjanskii, M. V., and Wishart, D. S. (2005) A simple method to predict protein flexibility using secondary chemical shifts. *J. Am. Chem. Soc. 127*, 14970−14971.
(69) Berjanskii, M. V., and Wishart, D. S. (2008) Application of the random coil index to studying protein flexibility. *J. Biomol. NMR 40*, 31−48.
(70) Sormanni, P., Camilloni, C., Fariselli, P., and Vendruscolo, M. (2015) The s2D Method: Simultaneous Sequence-Based Prediction of the Statistical Populations of Ordered and Disordered Regions in Proteins. *J. Mol. Biol. 427*, 982−996.
(71) Bai, Y., Milne, J. S., Mayne, L., and Englander, S. W. (1993) Primary structure effects on peptide group hydrogen exchange. *Proteins: Struct., Funct., Genet. 17*, 75−86.
(72) Connelly, G. P., Bai, Y., Jeng, M.-F., and Englander, S. W. (1993) Isotope effects in peptide group hydrogen exchange. *Proteins: Struct., Funct., Genet. 17*, 87−92.
(73) Fersht, A. (1999) *Structure and mechanism in protein science: A guide to enzyme catalysis and protein folding*, W. H. Freeman, New York.
(74) Dill, K. A., and Shortle, D. (1991) Denatured states of proteins. *Annu. Rev. Biochem. 60*, 795−825.
(75) Wishart, D. S., and Case, D. A. (2001) Use of chemical shifts in macromolecular structure determination. *Methods Enzymol. 338*, 3−34.
(76) Wishart, D. S. (2011) Interpreting protein chemical shift data. *Prog. Nucl. Magn. Reson. Spectrosc. 58*, 62−87.
(77) Cavalli, A., Salvatella, X., Dobson, C. M., and Vendruscolo, M. (2007) Protein structure determination from NMR chemical shifts. *Proc. Natl. Acad. Sci. U. S. A. 104*, 9615−9620.
(78) Shen, Y., Lange, O., Delaglio, F., Rossi, P., Aramini, J. M., Liu, G. H., Eletsky, A., Wu, Y. B., Singarapu, K. K., Lemak, A., Ignatchenko, A., Arrowsmith, C. H., Szyperski, T., Montelione, G. T., Baker, D., and Bax, A. (2008) Consistent blind protein structure generation from NMR chemical shift data. *Proc. Natl. Acad. Sci. U. S. A. 105*, 4685−4690.
(79) Wishart, D. S., Arndt, D., Berjanskii, M., Tang, P., Zhou, J., and Lin, G. (2008) CS23D: a web server for rapid protein structure generation using NMR chemical shifts and sequence data. *Nucleic Acids Res. 36*, W496−W502.
(80) Auer, R., Neudecker, P., Muhandiram, D. R., Lundstrom, P., Hansen, D. F., Konrat, R., and Kay, L. E. (2009) Measuring the Signs of $^{1}H^{\alpha}$ Chemical Shift Differences Between Ground and Excited Protein States by Off-Resonance Spin-Lock $R_{1\rho}$ NMR Spectroscopy. *J. Am. Chem. Soc. 131*, 10832−10833.
(81) Brockwell, D., Yu, L., Cooper, S., McCleland, S., Cooper, A., Attwood, D., Gaskell, S. J., and Barber, J. (2001) Physicochemical consequences of the perdeuteriation of glutathione S-transferase from S. japonicum. *Protein Sci. 10*, 572−580.
(82) Mittermaier, A., Korzhnev, D. M., and Kay, L. E. (2005) Side-chain interactions in the folding pathway of a Fyn SH3 domain mutant studied by relaxation dispersion NMR spectroscopy. *Biochemistry 44*, 15430−15436.
(83) Makhatadze, G. I., Clore, G. M., and Gronenborn, A. M. (1995) Solvent isotope effect and protein stability. *Nat. Struct. Biol. 2*, 852−855.
(84) Kuhlman, B., and Raleigh, D. P. (1998) Global analysis of the thermal and chemical denaturation of the N-terminal domain of the ribosomal protein L9 in H_2O and D_2O. Determination of the thermodynamic parameters, ΔH degree, ΔS degree, and ΔC_p degree, and evaluation of solvent isotope effects. *Protein Sci. 7*, 2405−2412.
(85) Scholtz, J. M., Pace, C. N., and Huyghues-Despointes, B. M. P. (1999) Protein conformational stabilities can be determined from hydrogen exchange rates. *Nat. Struct. Biol. 6*, 910−912.
(86) Villegas, V. V., Viguera, A. R., Aviles, F. X., and Serrano, L. (1996) Stabilization of proteins by rational design of alpha-helix stability using helix/coil transition theory. *Folding Des. 1*, 29−34.
(87) Bai, Y., Milne, J. S., Mayne, L., and Englander, S. W. (1994) Protein stability parameters measured by hydrogen exchange. *Proteins: Struct., Funct., Genet. 20*, 4−14.
(88) Minor, D. L., Jr., and Kim, P. S. (1994) Context is a major determinant of beta-sheet propensity. *Nature 371*, 264−267.

DOI: 10.1021/acs.biochem.5b00345
Biochemistry 2015, 54, 4611−4622