BIOCHEMISTRY
including biophysical chemistry & molecular biology
Article

Biochemistry **2009,** *48,* 8449–8457 8449
DOI: 10.1021/bi900596j

The Folding Pathway of Onconase Is Directed by a Conserved Intermediate†

Cindy Schulenburg,‡ Christian Löw,§,∥ Ulrich Weininger,§ Carmen Mrestani-Klaus,‡ Hagen Hofmann,‡,⊥ Jochen Balbach,§ Renate Ulbrich-Hofmann,‡ and Ulrich Arnold*,‡

‡*Institute of Biochemistry and Biotechnology, Martin-Luther University Halle-Wittenberg, Kurt-Mothes-Strasse 3, 06120 Halle, Germany, and* §*Institute of Physics, Martin-Luther University Halle-Wittenberg, Betty-Heimann-Strasse 7, 06120 Halle, Germany.* ∥*Present address: Department of Medical Biochemistry and Biophysics, Karolinska Institutet, Scheeles väg 2, 17177 Stockholm, Sweden.* ⊥*Present address: Institute of Biochemistry, University of Zurich, Winterthurer Str. 190, 8051 Zurich, Switzerland*

Received April 7, 2009; Revised Manuscript Received August 4, 2009

ABSTRACT: A promising approach to unravel the relationship between sequence information, tertiary structure, and folding mechanism of proteins is the analysis of the folding behavior of proteins with low sequence identity but comparable tertiary structures. Ribonuclease A (RNase A) and its homologues, forming the RNase A superfamily, provide an excellent model system for respective studies. RNase A has been used extensively as a model protein for folding studies. However, little is known about the folding of homologous RNases. Here, we analyze the folding pathway of onconase, a homologous protein from the Northern leopard frog with great potential as a tumor therapeutic, by high-resolution techniques. Although onconase and RNase A significantly differ in the primary structure (28% sequence identity) and in thermodynamic stability ($\Delta\Delta G$ = 20 kJ mol^{-1}), both enzymes possess very similar tertiary structures. The present folding studies on onconase by rapid mixing techniques in combination with fluorescence and NMR spectroscopy allow the structural assignment of the three kinetic phases observed in stopped-flow fluorescence spectroscopy. After a slow peptidyl−prolyl *cis*-to-*trans* isomerization reaction in the unfolded state, ONC folds via an on-pathway intermediate to the native state. By quenched-flow hydrogen/deuterium exchange experiments coupled with 2D NMR spectroscopy, 31 amino acid residues were identified to be involved in the structure formation of the intermediate. Twelve of these residues are identical in the RNase A sequence, which is a significantly higher percentage (39%) than the overall 28% sequence identity. Moreover, the structure of this intermediate closely resembles two of the intermediates that occur early during the refolding of RNase A. Obviously, in spite of considerable differences in their amino acid sequence the initial folding events of both proteins are comparable, guided by a limited number of conserved residues.

Despite great progress in structural biology, the interrelation of primary structure, folding pathways, tertiary structure, and stability of proteins is not yet sufficiently understood. Proteins with high sequence identity have been found to differ in structure and function (*1*, *2*). Conversely, numerous proteins, which significantly differ in sequence, were found to show comparable tertiary structures, and even the basic folding mechanism can be conserved within a fold family (*3*−*7*). Promising model systems for scrutinizing the similarities and differences of folding pathways are homologous proteins such as the members of the ribonuclease A (RNase A)[1] superfamily. Besides the well-known RNase A (from bovine pancreas; EC 3.1.27.5), this family comprises, for instance, bovine seminal RNase, human pancreatic RNase 1, and human angiogenin with 82%, 68%, and 36% sequence identity, respectively. In recent years, onconase (ONC; EC 3.1.27.5; Alfacell Corp., Somerset, NJ), an RNase A homologue from the oocytes and early embryos of the Northern leopard frog (*Rana pipiens*) (*8*), has gained increasing attention. Because of its antitumor activity ONC is a promising anticancer therapeutic that has reached phase IIIb of clinical trials (*9*). Interestingly, although ONC shows only 28% sequence identity with RNase A (Supporting Information Figure S1) and is more stable by 20 kJ mol^{-1} (*10*), it shares the typical kidney-shaped tertiary structure with RNase A (*11*) (Figure 1). While three of the four disulfide bonds are conserved, the fourth disulfide bond in RNase A (C65−C72) is located in an extended loop but tethers the C-terminus to the protein body in ONC (C87−C104) (Figure 1).

RNase A is one of the best studied enzymes in protein folding (*12*−*19*), and various parallel folding pathways have been described for this enzyme. Four folding intermediates were structurally characterized (*14*, *16*, *20*), among others by H/D exchange experiments. The protection of hydrogen-bonded main-chain amide protons against exchange with solvent deuterium

†The Land Sachsen-Anhalt is acknowledged for financial support of C.S. Moreover, we are grateful for the support of this work by the Graduiertenkolleg 1026 of the German Research Foundation (DFG), Bonn.

*Corresponding author. Telephone: +49 345 5524865. Fax: +49 345 5527303. E-mail: ulrich.arnold@biochemtech.uni-halle.de.

[1]Abbreviations: GdnDCl, guanidine deuterochloride; GdnHCl, guanidine hydrochloride; GdnSCN, guanidine thiocyanate; H/D, hydrogen/deuterium; HSQC, heteronuclear single-quantum coherence; I_{ONC}, intermediate state during the folding of ONC; k^F_{fast}, k^F_{medium}, and k^F_{slow}, rate constants of the fast, medium, and slow refolding phases; k^U_{fast} and k^U_{slow}, rate constants of the fast and slow unfolding phases; N, native state; NaOD, deuterated NaOH; NOESY, nuclear Overhauser enhancement spectroscopy; ONC, onconase; RNase A, ribonuclease A; TOCSY, total correlation spectroscopy; TROSY, transverse relaxation optimized spectroscopy; U, unfolded state.

Published on Web 08/05/2009

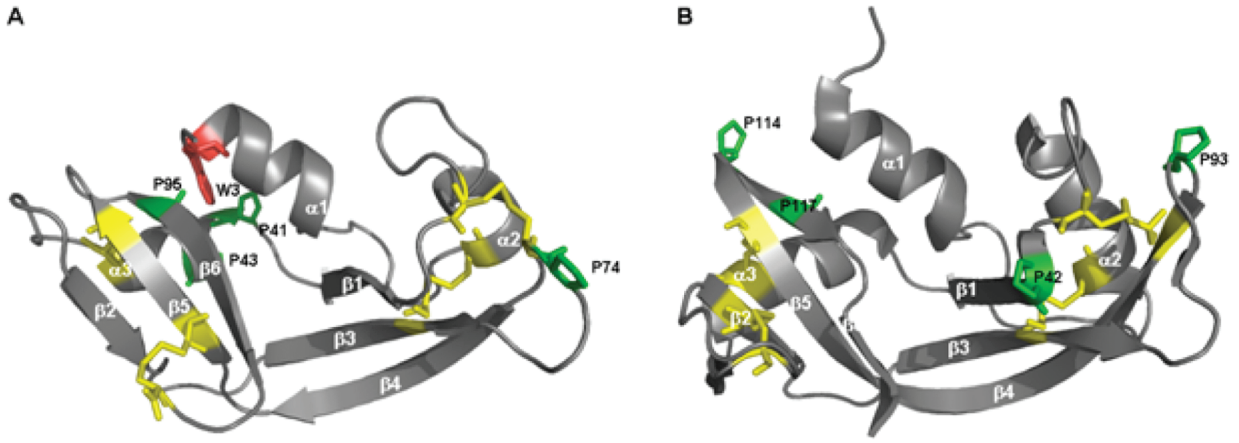

FIGURE 1: Ribbon diagram of ONC (A) and RNase A (B). The models (1ONC, 7RSA) were taken from the Brookhaven Protein Data Bank and drawn with PyMol (DeLano Scientific LLC, Palo Alto, CA). The disulfide bonds are indicated in yellow, and the tryptophan and proline residues are labeled.

atoms allows the identification of hydrogen bond pattern in less populated intermediates (*21*, *22*). In contrast, little is known about the folding of other members of the RNase A superfamily. From the comparison of folding kinetics of various pancreatic RNases, which show a sequence identity of > 85% but vary in the number of proline residues, a conserved refolding pathway was proposed (*23*) including a native-like intermediate (*24*). A comparative study on the folding of the homologues RNase A, ONC, and angiogenin addressed the influence of the number of *cis* prolines in the native state of these proteins on their folding kinetics (*25*). The refolding of the proteins was followed by single-jump absorbance spectroscopy, and the number of *cis* prolines was found to correlate with complexity of the folding behavior. However, in no case are structural data on intermediates available, which would allow the comparison with the intermediate structures of RNase A.

In contrast to RNase A, which contains two *cis* and two *trans* prolines and, thus, shows a rather complex folding behavior (*14*−*16*), all of the four proline residues of native ONC are in *trans* conformation. As in unfolded peptides the *trans* state dominates (*26*), a less complex folding mechanism is expected for ONC in comparison to RNase A.

Recently, we have shown that under native-like conditions three kinetic phases occur during the refolding of ONC (*27*). Here, we analyze the folding mechanism of ONC in more detail by rapid (single- and double-) mixing techniques in combination with fluorescence and NMR spectroscopy, which identified an folding intermediate and allowed an assignment of the kinetic phases to a peptidyl−prolyl *cis*-to-*trans* isomerization reaction in the unfolded state, followed by a sequential folding to N via an on-pathway intermediate (I_{ONC}). Moreover, quenched-flow hydrogen/deuterium (H/D) exchange experiments followed by multidimensional NMR spectroscopy were applied to yield structural information on I_{ONC} by identifying the amino acid residues, which are involved in the formation of I_{ONC}. From the comparison with the intermediates in the folding pathway of RNase A we conclude that regardless of sequence dissimilarities a limited number of conserved amino acid residues account for the first interactions in the protein folding reaction and guide the folding to the common tertiary structure.

MATERIALS AND METHODS

Materials. Deuterated GdnHCl (GdnDCl) and NaOH (NaOD) were obtained by repeated (three to five times) dissolving in D_2O followed by lyophilization. D_2O (99.9%) was purchased from Euriso-Top, Saint-Aubin Cedex, France, and $^{15}NH_4Cl$ was from Cambridge Isotope Laboratories, Saarbrücken, Germany. All chemicals used were of the purest grade commercially available.

Expression, Renaturation, and Purification of ONC. Expression, renaturation, and purification of ONC were performed as described previously (*10*). For 2D and 3D NMR spectra, ^{15}N-labeled ONC was purified from *Escherichia coli* cultures grown in M9 medium at 37 °C containing $^{15}NH_4Cl$. The cells were harvested 13 h after induction with isopropyl β-D-1-thiogalactopyranoside (1 mM). Renaturation and purification procedures were the same as described for unlabeled ONC (*10*).

Fluorescence Spectroscopy. All kinetic experiments were carried out in 100 mM sodium acetate buffer, pH 5.5, at 20 °C if not stated otherwise. Folding reactions initiated by manual mixing were carried out using a Fluoro-Max-2 spectrometer (Jobin Yvon, Munich, Germany) as described previously (*10*). Fast kinetics were followed by stopped-flow fluorescence spectroscopy using an Applied Photophysics BioSequential DX.20 MV stopped-flow spectrometer as described previously (*27*).

Unfolding of ONC was initiated by 11-fold dilution of native protein into buffer containing 5−7 M GdnHCl. Only one unfolding phase was detectable by stopped-flow fluorescence spectroscopy, which was too slow to come to completion in the time frame of the stopped-flow technique. The data, however, could unequivocally be described by a single-exponential equation with the rate constant of the slow phase from manual mixing experiments (*10*).

To follow the formation of N, unfolded ONC (in 6 M GdnHCl) was diluted 11-fold to initiate the refolding in double-mixing experiments (*14*). After defined delay times, refolding was interrupted by transferring the solution into 100 mM glycine buffer, pH 2.0, and 5 M guanidine thiocyanate (GdnSCN) to ensure fast and complete unfolding of ONC. The observed amplitudes of the unfolding reaction as a function of delay times were used as a measure for the formation of native protein.

To follow the formation and decay of I_{ONC}, unfolded ONC (in 6 M GdnHCl) was diluted 11-fold to initiate refolding. After defined delay times (0.025−5 s), refolding was interrupted by 5-fold dilution into a final denaturant concentration of 6.0 M GdnHCl. Unfolding kinetics were fitted to either a single-exponential equation or a double-exponential equation with

the preset rate constant of the slow unfolding phase. The observed amplitude of the fast unfolding phase as a function of the delay time was used as a measure of the formation and decay of I_{ONC}.

To follow the unfolding of I_{ONC}, double-mixing experiments were carried out at a constant delay time of 0.8 s (maximal population of I_{ONC} during refolding) with an 11-fold dilution of unfolded ONC (in 6 M GdnHCl) in the first mixing step and final denaturant concentrations between 2 and 4.4 M GdnHCl in the second mixing step. Data were fitted to either a single-exponential equation or a double-exponential equation with the preset rate constant of the slow unfolding phase. The final protein concentration for all interrupted refolding double-mixing experiments was 2 μM.

To determine the activation energy of the slow refolding reaction of ONC, refolding of ONC (initiated by an 11-fold dilution from 6 M GdnHCl) was followed between 10 and 45 °C. The resulting rate constants were fitted to the Arrhenius equation:

$$k = Ae^{-E_a/RT}$$

where A is the pre-exponential factor, E_a the activation energy, T the temperature, and R the gas constant.

Analysis of the Kinetic Data. The data of the chevron plot were fitted according to a sequential folding model using the analytical solutions (*28*, *29*) by use of the program Scientist (MicroMath).

NMR Measurements. For resonance assignment of native ONC (Supporting Information Table S2), 3D ^{15}N-NOESY-HSQC and 3D ^{15}N-TOCSY-HSQC NMR spectra were recorded in 50 mM sodium phosphate buffer, containing 10% D_2O at pH 7.0, on an Avance 400 NMR spectrometer (Bruker Biospin GmbH, Rheinstetten, Germany) at 20 °C. The assignment is based on the data for the ONC variant M-1/Q1/M23L-ONC (*30*). All other spectra were recorded in 50 mM sodium acetate buffer, pH 5.5 or pD 5.5, respectively, containing either 10% (v/v) or 100% (v/v) D_2O. All data were either processed and analyzed using FELIX-ND (PC version; Accelrys, San Diego, CA) or processed using NMRPipe (*31*) and analyzed using NMRView (*32*).

Real-time unfolding was followed on a Bruker Avance 400 NMR spectrometer. Unfolding was initiated by dissolving ONC (lyophilized from an aqueous solution) in 5.5 M GdnHCl, pH 5.5 (final concentration of ONC = 1.7 mM), at 20 °C and followed by a series of 40 2D ^{15}N-TROSY-HSQC NMR spectra. The dead time of the reaction was 12 min. The peak volume as a function of unfolding time was fitted by a single-exponential equation. The assignment was transferred to the 2D ^{15}N-TROSY-HSQC NMR spectrum recorded at 5.5 M GdnHCl by a series of 2D ^{15}N-TROSY-HSQC NMR spectra recorded at increasing GdnHCl concentrations.

1D real-time refolding experiments (in 0.6 M GdnHCl, pH 5.5, at 10 °C) were carried out on an Avance II 600 NMR spectrometer (Bruker Biospin GmbH). The reaction was started by a rapid 10-fold dilution of unfolded ONC (in 6 M GdnHCl). The dead time of the detection was 12 s. During the refolding reaction, 32 1D ^{1}H NMR spectra were recorded. The peak volume as a function of the time of refolding was fitted by a single-exponential equation.

To assess native state H/D exchange in ONC, a series of 2D ^{15}N-HSQC spectra was recorded for 3 days on an Avance II 600 NMR spectrometer (Bruker) after dissolving ONC (lyophilized from an aqueous solution) in D_2O buffer. Only residues that show no exchange over 24 h were analyzed in the quenched-flow experiments described below.

Quenched-Flow Analyses. Quenched-flow experiments were carried out at 20 °C using a rapid mixing quenched-flow machine (Bio Logic, Claix, France). In the first mixing step, refolding of protonated, unfolded ONC (in 6 M GdnHCl) was initiated by a 9-fold dilution into 50 mM sodium acetate buffer, pH 5.5. After defined refolding times (60−5000 ms), the samples were diluted 10-fold into D_2O buffer (200 mM glycine buffer, 0.667 M GdnHCl, pD 11.0) to initiate the H/D exchange. The exchange was terminated after 8.3 ms by lowering the pH by a 2-fold dilution with D_2O, containing 10 M acetic acid and 0.667 M GdnHCl. ONC proved to be stable at pH 2−12 as determined by recording GdnHCl-induced transition curves as described previously (*10*) at various pH values (data not shown). The protein was concentrated to 500 μL of a 0.3 mM solution of ONC using Vivaspin centrifugal concentrators (Sigma-Aldrich, Steinheim, Germany), and the buffer was exchanged to 50 mM sodium acetate buffer, pD 5.5.

As a reference, protonated protein was refolded, exchanged, and quenched with buffer containing H_2O instead of D_2O. The exchange was monitored by 2D ^{15}N-HSQC NMR spectra (Supporting Information Figure S2) on a Bruker Avance II 600 NMR spectrometer. The percentage of proton occupancy was calculated by integration of the peak volumes in relation to the reference. To assess residual structures in the unfolded protein, protonated ONC was incubated in 100% deuterated 50 mM sodium acetate buffer containing 6 M guanidine deuterochloride (GdnDCl), pD 5.5, for 12 h, refolded in 100% deuterated 50 mM sodium acetate buffer, pD 5.5, and analyzed by a 2D ^{15}N-HSQC NMR spectrum.

RESULTS

Assignment of the Kinetic Phases to Folding Events. In the refolding reaction of ONC, three phases with the observed rate constants k^F_{fast}, k^F_{medium}, and k^F_{slow}, differing by 1 order of magnitude each, had been detected by fluorescence and CD spectroscopy (*27*) (Figure 2). In the unfolding reaction, one phase with the observed rate constant k^U_{slow} had been detected by single-mixing experiments (*27*), but an additional fast unfolding phase with the observed rate constant k^U_{fast} was observed by double-mixing experiments (see below). In the following, these kinetic phases will be assigned to definite folding events.

(A) The Slow Refolding Phase Is Caused by a Peptidyl−Prolyl Cis-to-Trans Isomerization Reaction in Unfolded ONC. The rate constant of k^F_{slow} in the refolding of ONC is independent of the guanidine hydrochloride (GdnHCl) concentration (Figure 2), which is a characteristic of peptidyl−prolyl *cis*/*trans* isomerization reactions (*33*). The assignment of the slow phase to a peptidyl−prolyl *cis*-to-*trans* isomerization reaction is supported by the temperature dependence of k^F_{slow} (Supporting Information Figure S3). The resulting Arrhenius activation energy of (74.5 ± 0.2) kJ mol^{-1} is comparable to the typical value for a peptidyl−prolyl *cis*/*trans* isomerization reaction of 80 kJ mol^{-1}, whereas conformational folding reactions are characterized by significantly lower activation energies (*34*).

Moreover, the refolding reaction of ONC was followed by 1D real-time NMR spectroscopy in 0.6 M GdnHCl at 10 °C. Under these conditions, the reaction proceeding with k^F_{slow} ((2.6 ± 1.2) × 10^{-2} s^{-1}, determined by fluorescence spectroscopy) is sufficiently slow to be monitored by real-time NMR spectroscopy. Although

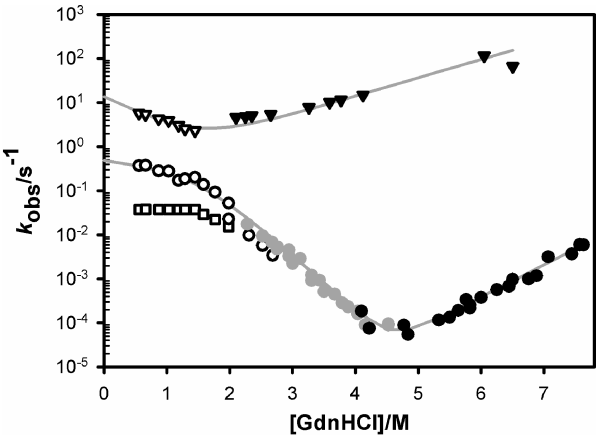

FIGURE 2: Chevron plot of the observed folding rate constants of ONC. Rate constants k_{obs} were determined as a function of GdnHCl concentration by stopped-flow single-mixing (▽, k^F_{fast}; ○, k^F_{medium}; □, k^F_{slow}), manual mixing (shaded circle, k^F_{medium}; ●, k^U_{slow}), or interrupted refolding stopped-flow (▼, k^U_{fast}) fluorescence spectroscopy experiments, respectively. The lines represent a fit of the data according to a sequential three-state model. All data were obtained in 100 mM sodium acetate buffer, pH 5.5, at 20 °C as described in Materials and Methods.

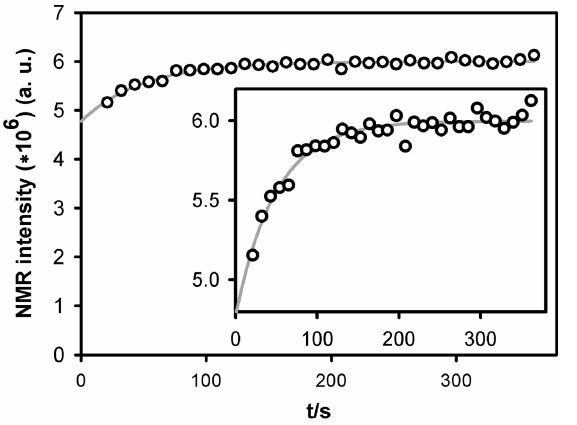

FIGURE 3: Refolding kinetics of ONC monitored by real-time NMR spectroscopy. The formation of N is represented by the integral of the NMR signal intensities of the side-chain resonances at 0.2 ppm (Figure 4). The data were fitted by a single-exponential equation, with a rate constant of $(2.1 \pm 1.0) \times 10^{-2}$ s^{-1}. During the dead time, (78.3 ± 2.8)% of N are formed. The insert shows the same data with expanded ordinate. NMR spectra were obtained in 100 mM sodium acetate buffer and 10% D_2O, pH 5.5, at 10 °C.

the first spectrum recorded after 12 s already resembles the spectrum of native ONC, a further increase in the intensities of native signals (amide protons with chemical shifts downfield to 8.5 ppm or methyl groups with shifts highfield to 0.5 ppm) could be observed. Both the amplitude ((21.7 ± 2.8)%) and the rate constant of the slow phase ($(2.1 \pm 1.0) \times 10^{-2}$ s^{-1}; Figure 3) coincide with the corresponding values for the slow phase in the stopped-flow fluorescence spectroscopy experiments (23% and 2.6×10^{-2} s^{-1}; see above).

Figure 4A shows the high-field part of the 1D ^{1}H NMR spectrum < 1.4 ppm of native ONC. Within the dead time of the refolding experiment (12 s), the fast and medium refolding phases, which yield ~75% native ONC as known from folding experiments followed by fluorescence spectroscopy (*27*), are completed. In Figure 4B, the spectrum of ONC after refolding for 12 s, reduced by 75% of the signal intensity of the spectrum of native ONC, is shown. This difference spectrum, which now shows resonances of the slow folding species only, still resembles the spectrum of unfolded ONC (no signal < 0.4 ppm), indicating that the reaction with k^F_{slow}, which has been characterized as a *cis*-to-*trans* isomerization reaction (see above), occurs in the unfolded state.

(*B*) *The Medium-Rate Refolding Phase Results in Native ONC*. To identify the refolding phase that results in the formation of native protein, interrupted refolding experiments (N-assay) (*14*) were carried out and followed by fluorescence spectroscopy as described in Materials and Methods. The refolding reaction of unfolded ONC (in 6.0 M GdnHCl) was initiated by 11-fold dilution in the first mixing step. The resulting concentration of 0.55 M GdnHCl represents native conditions for ONC. After defined delay times, the refolding reaction was interrupted, and the protein was subjected again to unfolding conditions. Since the unfolding reaction of native ONC is very slow under common unfolding conditions (e.g., $(1.34 \pm 0.2) \times 10^{-4}$ s^{-1} in 5.5 M GdnHCl, pH 5.5, at 20 °C), harsher unfolding conditions were chosen (glycine buffer, pH 2.0, containing 5 M GdnSCN) to accelerate the unfolding reaction of native ONC.

The amplitude of the monitored single-exponential unfolding reaction reflects the percentage of native protein formed after a given refolding time (Figure 5, circles). The rate constant of the formation of native molecules ((0.54 ± 0.08) s^{-1}) is in good accordance with k^F_{medium} at the corresponding GdnHCl concentration ((0.48 ± 0.01) s^{-1}), showing that native ONC is formed during the refolding phase characterized by k^F_{medium}.

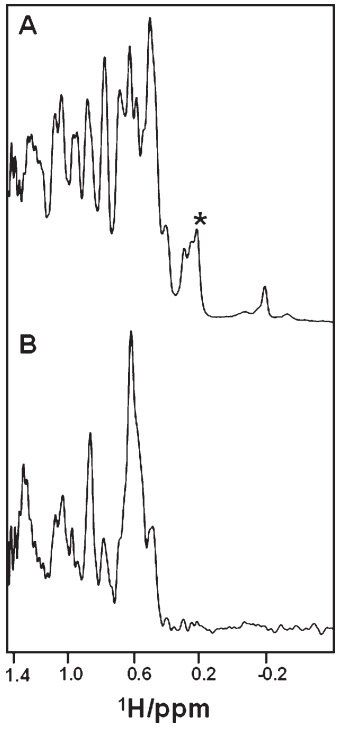

FIGURE 4: High-field region of the 1D ^{1}H NMR spectra of ONC. (A) Native ONC with all proline residues in *trans* conformation. The star marks the signal used to follow the refolding reaction in Figure 3. (B) Extracted spectrum of ONC with proline residues in *cis* conformation. In the first spectrum of refolding (t = 12 s, Figure 3), the fast and medium phases (75% of total amplitude) (*27*) are completed. By subtraction of 75% of the spectrum of native ONC, the spectrum of unfolded ONC with at least one proline residue in the non-native *cis* conformation was calculated.

(*C*) *The Fast Refolding Phase Leads to an On-Pathway Intermediate*. From the results of sections A and B above it can

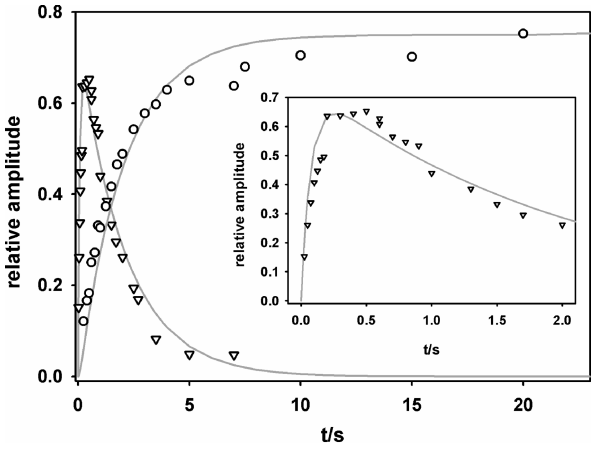

FIGURE 5: Relative amplitudes of the unfolding reactions of N and I_{ONC} as a function of the delay time. Amplitudes of the unfolding reaction of N (○) and I_{ONC} (▽) were determined by interrupted refolding experiments followed by fluorescence spectroscopy as described in Materials and Methods. The lines represent a simulation of the data using the intrinsic unfolding and refolding rates at 0.55 M GdnHCl derived from the global analysis of the chevron plot (Figure 2).

be concluded that k_{fast}^{F} describes the formation of an folding intermediate, termed I_{ONC}. To analyze the formation and the decay of I_{ONC}, interrupted refolding experiments were carried out and followed by fluorescence spectroscopy as described in Materials and Methods. As I_{ONC} was expected to be less stable than N and, thus, should also unfold more rapidly than N, the experiments were performed in the standard buffer containing GdnHCl instead of GdnSCN. Refolding of unfolded ONC (in 6.0 M GdnHCl) was initiated by 11-fold dilution in the first mixing step, and unfolding was followed in 6.0 M GdnHCl after varying delay times. Depending on the delay time, the unfolding kinetics followed either a single-exponential (delay times < 0.8 s, yielding the rate constant of the fast unfolding reaction k_{fast}^{U}) or a double-exponential function (delay times > 0.8 s, yielding the unfolding reaction rate constants k_{fast}^{U} and k_{slow}^{U}; data not shown), where k_{slow}^{U} reflects the unfolding of native molecules (*10*) formed during longer delay times. However, this unfolding reaction is too slow to be followed to completion by stopped-flow measurements, but the reaction can be well described by a double-exponential function with k_{slow}^{U} determined by manual mixing experiments (Figure 2) (*10*).

In the plot of the relative amplitudes of k_{fast}^{U} vs delay time (Figure 5, triangles), the increase of the amplitudes reflects the formation of I_{ONC}. The increase in amplitude without lag phase indicates a direct coupling of U and I_{ONC} without further detectable intermediate states. The rate constant of the formation of I_{ONC} ($(9.04 \pm 0.12)\ s^{-1}$) coincides with k_{fast}^{F} in the stopped-flow fluorescence spectroscopy experiments under comparable conditions ($(5.74 \pm 0.10)\ s^{-1}$, Figure 2). Furthermore, the rate constant of the decay of I_{ONC} ($(0.66 \pm 0.09)\ s^{-1}$) equals k_{medium}^{F} under corresponding conditions ($(0.48 \pm 0.01)\ s^{-1}$), indicating a direct connection between I_{ONC} and N. The direct coupling of I_{ONC} to both N and U characterizes I_{ONC} as an on-pathway intermediate.

Additionally, interrupted refolding experiments were carried out to follow the unfolding of I_{ONC} as a function of the GdnHCl concentration. The refolding reaction in 0.55 M GdnHCl was interrupted after 0.8 s, where I_{ONC} is maximally populated (Figure 5), and the unfolding reaction was initiated by subjecting the samples into solutions of varying GdnHCl concentration. At high concentrations of GdnHCl, only one phase is observed, which reflects the unfolding of I_{ONC} to U (Supporting Information Figure S4A). In the presence of low concentrations of GdnHCl (2.35−4.2 M GdnHCl), however, the fluorescence signal initially increases but decreases after 2 s (Supporting Information Figure S4B). The first phase again reflects the unfolding of I_{ONC} to U with k_{fast}^{U}, but the second phase is caused by the simultaneous folding of I_{ONC} to N (with k_{medium}^{F}), which is populated under these conditions. By varying the concentration of GdnHCl in the second mixing step, the unfolding limb of k_{fast}^{U} in the chevron plot (Figure 2, filled triangles) was determined.

(*D*) *ONC Unfolds without Detectable Intermediates before the Rate-Limiting Step*. Since for RNase A intermediates have been reported to occur before the rate-limiting step of unfolding (*15*, *35*, *36*), the unfolding process of ONC was explored at the level of single amino acid residues by real-time NMR spectroscopy in 5.5 M GdnHCl. Under these conditions, single-jump initiated unfolding of N occurs with one phase (*10*) and is slow enough to be followed by 2D NMR spectroscopy ($k_{slow}^{U} = (1.34 \pm 0.20) \times 10^{-4}\ s^{-1}$). In the first NMR spectrum ($t = 17$ min; [N] = 87%), the same 95 out of the 104 amino acid residues could be assigned as in native ONC, indicating that no local structural changes have occurred in the dead time of the experiment.

The traces for all peaks, independent of decreasing or increasing signals (i.e., decay of N or formation of U, respectively; Supporting Information Figure S5), could be fitted to a single-exponential equation with a mean rate constant of $(1.24 \pm 0.09) \times 10^{-4}\ s^{-1}$, which coincides with k_{slow}^{U} determined by fluorescence spectroscopy. This indicates an unfolding reaction for ONC without any detectable intermediates before the rate-limiting step (Supporting Information Figure S6B).

Global Analysis of the Kinetic Data of ONC. From the observed rate constants of the folding and unfolding reaction (Figure 2), the intrinsic rate constants were determined by a global fit for a simple sequential folding pathway (Scheme 1) without consideration of the peptidyl−prolyl *cis*/*trans* isomerization in the unfolded state. The intrinsic rate constants (Supporting Information Table S1) were used to calculate the population profiles for the native and intermediate state in the double-mixing experiments (Figure 5, gray curves). These profiles are consistent with the experimentally detected formation of N as well as the transient population of I_{ONC}. Consequently, the folding pathway of ONC is sufficiently described by the sequential model even though a more intricate folding behavior (such as according to a triangular mechanism) cannot be excluded.

The intrinsic rate constants also allowed the calculation of the thermodynamic stability of I_{ONC} in the absence of denaturant ($\Delta G_I^\circ = (9.2 \pm 2.1)$ kJ mol^{-1}), which contributes to the total thermodynamic stability of ONC ($\Delta G_N^\circ = (61.1 \pm 3.2)$ kJ mol^{-1}, determined by GdnHCl-induced transition curves) (*10*).

Structural Characterization of I_{ONC}. To identify individual amino acid residues involved in the formation of I_{ONC}, quenched-flow H/D exchange experiments coupled with 2D NMR spectroscopy were carried out as described in Materials and Methods. For the exchange experiments presented here, refolding of fully protonated ONC was performed in protonated

Scheme 1

$$U \rightleftharpoons I_{ONC} \rightleftharpoons N$$

buffer as deuteration can influence the protein stability (*37*). By these quenched-flow H/D exchange experiments, backbone amide protons, which are located in folded structures of the intermediate, can be identified. As they are already protected from exchange with bulk water (D_2O), they yield an NMR signal in contrast to unprotected protons. After defined delay times, the protein was diluted 10-fold into D_2O, pH 10, to initiate the H/D exchange, which was subsequently quenched by a rapid decrease of the pH. The extent of exchange was determined by 2D ^{15}N-HSQC NMR spectroscopy. The peak volumes of the NMR signals were integrated and related to the NMR signals of ONC that was identically refolded, exchanged, and quenched with buffer containing H_2O instead of D_2O. To probe stable residual structures in the unfolded state, which might distort the data analysis, ONC was denatured and refolded in solutions containing 100% D_2O. No peaks have been detected in the 2D ^{15}N-HSQC NMR spectrum, indicating that unfolded ONC does not contain any residual structures.

Furthermore, to exclude interference by exchange in the native state, native-state H/D exchange of ONC was followed for 3 days by 2D NMR spectroscopy (Supporting Information Figure S7). Residues that showed significant exchange during one day were excluded from the quenched-flow analysis (Figure 6, red residues).

Forty-two residues showed no exchange during this time and were followed by the quenched-flow H/D exchange experiment. All of them showed an increase in proton occupancy with increasing delay times. Interestingly, the residues cluster in two groups differing in the kinetics of the proton occupancy. Thirty-one residues are already highly protected after very short delay times (< 250 ms, Figure 7A,B). The gray curves in Figure 7A,B show the simulation according to a double-exponential equation with values for the intrinsic rate constants $k_{U\rightarrow I_{ONC}} = 4.72\ s^{-1}$ and $k_{I_{ONC}\rightarrow N} = 0.38\ s^{-1}$ (values at 0.67 M GdnHCl, calculated from the parameters in Supporting Information Table S1). All residues showing this type of kinetics of the proton occupancy are involved in the structure formation of I_{ONC}

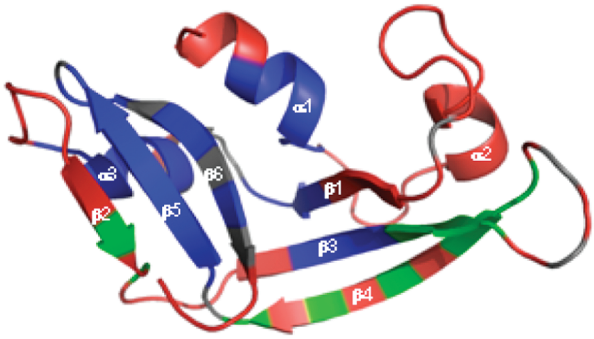

FIGURE 6: Residues of ONC with occupied protons in the individual refolding stages. Residues in blue show high proton occupancy within very short delay times of the quenched-flow experiments and are involved in the formation of I_{ONC}; residues in green show increasing proton occupancy during the medium refolding phase only, i.e., in the formation of N. All residues in red show H/D exchange in native ONC within 24 h and were not used for quenched-flow analysis. Residues in gray either are proline residues or cannot be evaluated due to signal overlap or missing assignment. The ribbon diagram is drawn with PyMol.

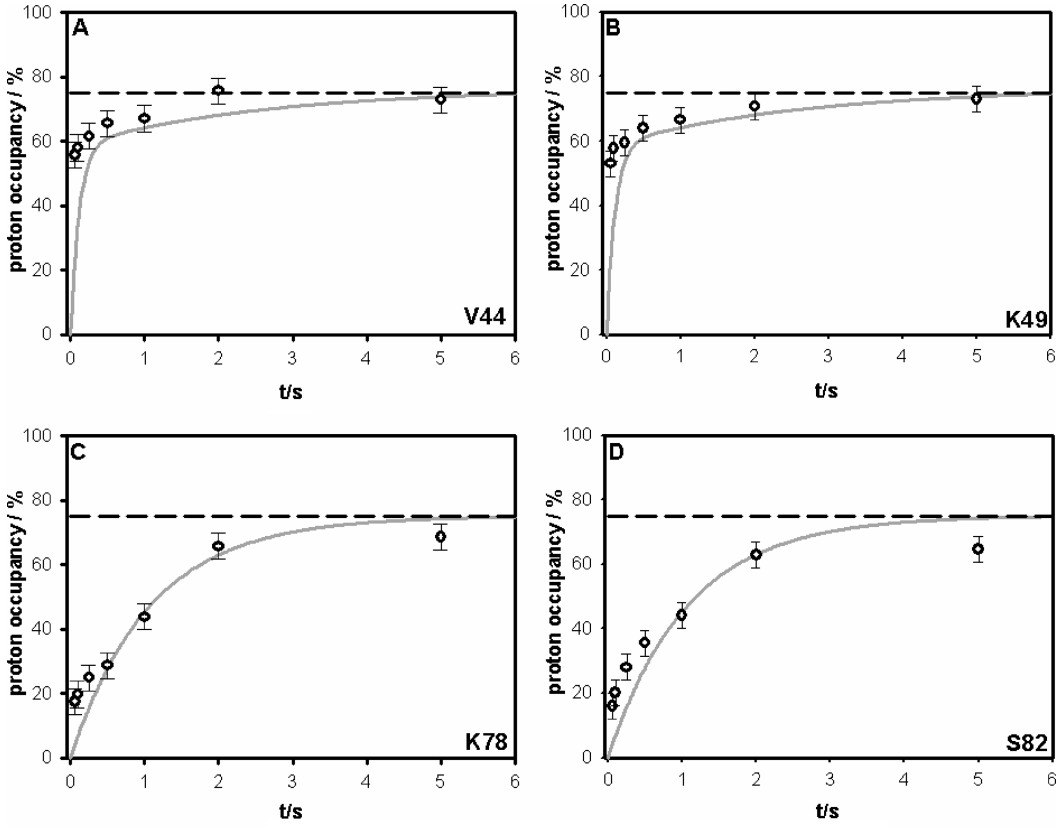

FIGURE 7: Time courses of amide proton occupancy. Amide proton occupancy was determined by quenched-flow H/D exchange as described in Materials and Methods. The gray curves represent the time course according to a double-exponential function with $k_{U\rightarrow I_{ONC}} = 4.72\ s^{-1}$ (58% amplitude) and $k_{I_{ONC}\rightarrow N} = 0.38\ s^{-1}$ (17% amplitude) for panels A (V44) and B (K49). In panels C (K78) and D (S82), the gray curves represent the time course according to a single-exponential function with $k_{I_{ONC}\rightarrow N} = 0.38\ s^{-1}$ (75% amplitude). The line at 75% marks the maximum of proton occupancy that can be reached within the fast and medium refolding phase.

(with k^F_{fast}) and the consecutive reaction to N (with k^F_{medium}). Obviously, these protons are less protected against exchange in I_{ONC} than in N. Such a phenomenon has been described previously for folding intermediates of other proteins (*20*, *38*). These residues, which are located in the C-terminal part of the first α-helix, in the third α-helix, and in the fifth and sixth β-strands, are labeled in blue in Figure 6. In these regions, there are numerous hydrophobic amino acid residues that form an extended hydrophobic cluster in the native protein (*39*). Moreover, several residues in the first (F36, I37, Y38, S39) and third β-strands (L65, S66, D67) become protected.

The second group (the remaining 11 residues) shows an increase in proton occupancy with the rate constant of the formation of native molecules (k^F_{medium}). The gray lines in Figure 7C,D demonstrate the simulation according to a single-exponential equation with the intrinsic rate constant $k_{I_{ONC}\rightarrow N}$ = 0.38 s^{-1} (values at 0.67 M GdnHCl, calculated from the parameters in Supporting Information Table S1). The residues which become protected within this process do still exchange in I_{ONC} and are protected in N only. These residues are labeled green in Figure 6. Interestingly, alternating residues of the second and fourth β-strand become protected concomitantly with the formation of N, reflecting the characteristic hydrogen bond pattern of exterior β-strands.

None of the observable residues reaches a proton occupancy of 100% within 5 s, which is due to the slow peptidyl−prolyl *cis*-to-*trans* isomerization reaction in unfolded ONC (k^F_{slow}; see section A above).

DISCUSSION

Despite the low sequence identity of 28% (*11*) (Supporting Information Figure S1), RNase A and ONC show a similar tertiary structure (Figure 1) and distribution of polar and hydrophobic regions (*39*). The presence of the three conserved disulfide bonds, which tether the ONC and RNase A molecule, respectively, might imply a comparable folding behavior of these two proteins. However, the large stability difference ($\Delta\Delta G$ = 20 kJ mol^{-1}) (*10*) as well as the *trans* conformation of all proline residues in the native state of ONC make differences in the unfolding and/or refolding between ONC and RNase A very probable. In fact, significant differences are revealed by the relatively simple sequential folding mechanism of ONC in contrast to the complex folding mechanism of RNase A (*14*, *16*, *20*).

Folding of ONC has been described to proceed via two phases at pH 8, 15 °C, 0.55 M GdnHCl (with rate constants of 3.85 × 10^{-1} s^{-1} and 1.45 × 10^{-2} s^{-1}) (*25*) or three phases at pH 5.5, 20 °C, 0.55 M GdnHCl (with rate constants of 5.74 s^{-1}, 4.83 × 10^{-1} s^{-1}, and 8.46 × 10^{-2} s^{-1}) (*27*), however, without any structural information on intermediate states hitherto.

The more detailed analysis of the folding pathway of ONC in this paper, using rapid mixing methods coupled with fluorescence as well as with NMR spectroscopy, provides deeper insight into the folding mechanism of ONC. The slow phase has been suggested to originate from peptidyl−prolyl *cis*/*trans* isomerization previously (*25*) even though this reaction could not be accelerated by the peptidyl−prolyl *cis*/*trans* isomerases cyclophilin 18 or *Thermus thermophilus* SlyD (*27*). By following the folding process of ONC by real-time NMR spectroscopy, we show that the slow phase (with an amplitude of about 25%, Figure 3) originates from a conformational conversion reaction in the unfolded state (Figure 4). An activation energy of (74.5 ± 0.2) kJ mol^{-1} (Supporting Information Figure S3) characterizes k^F_{slow} as a peptidyl−prolyl *cis*/*trans* isomerization reaction.

As concluded from the interrupted refolding experiments with the unfolding step under extreme conditions (Figure 5), native ONC is formed with k^F_{medium}. Interrupted refolding experiments (Figure 5) as well as the global analysis of all kinetic data (Figures 2 and 5) demonstrate that native ONC is formed via an on-pathway intermediate (I_{ONC}). As the reactivation of ONC ($(9.3 \pm 0.9) \times 10^{-2}$ s^{-1} at 10 °C) (*27*) coincides with the formation of native ONC ($k^F_{medium} = (9.7 \pm 0.1) \times 10^{-2}$ s^{-1} at 10 °C), I_{ONC} is suggested to be catalytically inactive, which is supported by the fact that the catalytically essential K31 is still unprotected in I_{ONC} (Figure 6).

The unfolding reaction of ONC is monophasic, and no intermediates were detectable by 2D NMR spectroscopy before the rate-limiting folding reaction (Supporting Information Figures S5 and S6B). In contrast, in the unfolding process of the homologous enzyme RNase A, local unfolding has been described to occur before the global unfolding reaction (*15*, *35*, *36*).

Due to the presence of two proline residues in *cis* conformation in the native state of RNase A (Figure 1B) and their isomerization in the unfolded state, the folding process of RNase A is rather complex and results in multiple parallel folding pathways including various intermediates (*12*−*14*, *16*, *20*). The best characterized, native-like intermediate I_N (*14*, *34*) is preceded by the early intermediate I_1, which is already structured and shows strong protection for 14 backbone amide protons (*20*). On a parallel route, a hydrophobically collapsed intermediate I_Φ becomes populated, and the analysis of 21 of the 124 amino acid residues revealed that I_Φ has well-established secondary structure elements (*16*). The protected regions, the β-sheets, particularly in the C-terminal part of the RNase A molecule (β5 and β6 in Figure 1B), are largely identical in both intermediates and support the suggestion of residues 106−118 as the folding initiation site of RNase A (*40*).

The folding of ONC follows a much simpler sequential process when monitored by fluorescence or NMR spectroscopy (alone or in combination with quenched-flow H/D exchange experiments). A rate-limiting peptidyl−prolyl *cis*-to-*trans* isomerization reaction occurs here in the unfolded state and, thus, does not cause the population of a folding intermediate. The NMR spectra of the quenched-flow experiments reveal both native-like (31 residues, blue in Figure 6) and still unstructured regions in I_{ONC} (Figure 6). Interestingly, the fraction of identical amino acid residues between ONC and RNase A, which already possess a native-like conformation in I_{ONC} (39%), is significantly larger than the total fraction of identical residues (28%, Supporting Information Figure S1) (*11*). The number of identical residues increases even further to 66%, if the section of the proposed folding initiation site (I106−V118 in RNase A (*40*) corresponding to C87−V96 in ONC) is considered only (Supporting Information Figure S1). The structural assembly of I_{ONC} (Figure 6) is in excellent accordance with the characteristics of I_1 (*20*) and I_Φ (*16*) in the refolding of RNase A. In all of these intermediates large parts of the β-sheets (particularly in the C-terminal region of the proteins) are structured.

Interestingly, also during the oxidative folding of reduced ONC those regions that are structured in I_{ONC} become structured initially. In the first isolated intermediate (I_1) (*41*) the disulfide bonds C48−C90 and C87−C104 (Figures 1A and 6), both of

which are located within the structured regions of I_{ONC}, are already formed.

In summary, our results reveal that, despite a general low sequence identity and clear differences in the folding behavior, the initial interactions of a limited number of conserved amino acid residues are analogous in the folding reaction of ONC and RNase A. Therefore, structurally comparable, early folding intermediates restrict the conformational search of the unfolded polypeptide chain to achieve the common native tertiary structure. They seem to be more crucial than a high overall sequence identity or a conserved folding mechanism. This interpretation would explain the evolutionary robustness of the tertiary structure of homologous proteins against mutations.

ACKNOWLEDGMENT

The authors are very grateful to PD Dr. Christian Lücke (Max-Planck Research Unit for Enzymology of Protein Folding, Halle, Germany) for help with the assignment of ONC proton NMR signals and thank PD Dr. Hauke Lilie (Martin-Luther University, Halle, Germany) for helpful discussions. Prof. Ronald T. Raines is acknowledged for the gift of the plasmid pET-22b(+) containing the gene for the expression of ONC.

SUPPORTING INFORMATION AVAILABLE

Figures S1−S7 and Tables S1 and S2. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. He, Y., Chen, Y., Alexander, P., Bryan, P. N., and Orban, J. (2008) NMR structures of two designed proteins with high sequence identity but different fold and function. *Proc. Natl. Acad. Sci. U.S.A. 105*, 14412–14417.
2. Dalal, S., and Regan, L. (2000) Understanding the sequence determinants of conformational switching using protein design. *Protein Sci. 9*, 1651–1659.
3. Chiti, F., Taddei, N., White, P. M., Bucciantini, M., Magherini, F., Stefani, M., and Dobson, C. M. (1999) Mutational analysis of acylphosphatase suggests the importance of topology and contact order in protein folding. *Nat. Struct. Biol. 6*, 1005–1009.
4. Martinez, J. C., and Serrano, L. (1999) The folding transition state between SH3 domains is conformationally restricted and evolutionarily conserved. *Nat. Struct. Biol. 6*, 1010–1016.
5. Friel, C. T., Capaldi, A. P., and Radford, S. E. (2003) Structural analysis of the rate-limiting transition states in the folding of Im7 and Im9: similarities and differences in the folding of homologous proteins. *J. Mol. Biol. 326*, 293–305.
6. Travaglini-Allocatelli, C., Gianni, S., and Brunori, M. (2004) A common folding mechanism in the cytochrome *c* family. *Trends Biochem. Sci. 29*, 535–541.
7. Chi, C. N., Gianni, S., Calosci, N., Travaglini-Allocatelli, C., Engstrom, K., and Jemth, P. (2007) A conserved folding mechanism for PDZ domains. *FEBS Lett. 581*, 1109–1113.
8. Mikulski, S. M., Ardelt, W., Shogen, K., Bernstein, E. H., and Menduke, H. (1990) Striking increase of survival of mice bearing M109 Madison carcinoma treated with a novel protein from amphibian embryos. *J. Natl. Cancer Inst. 82*, 151–153.
9. Schulenburg, C., Ardelt, B., Ardelt, W., Arnold, U., Shogen, K., Ulbrich-Hofmann, R., and Darzynkiewicz, Z. (2007) The interdependence between catalytic activity, conformational stability, and cytotoxicity of onconase. *Cancer Biol. Ther. 6*, 1233–1239.
10. Arnold, U., Schulenburg, C., Schmidt, D., and Ulbrich-Hofmann, R. (2006) Contribution of structural peculiarities of onconase to its high stability and folding kinetics. *Biochemistry 45*, 3580–3587.
11. Mosimann, S. C., Ardelt, W., and James, M. N. (1994) Refined 1.7 Å X-ray crystallographic structure of P-30 protein, an amphibian ribonuclease with anti-tumor activity. *J. Mol. Biol. 236*, 1141–1153.
12. Cook, K. H., Schmid, F. X., and Baldwin, R. L. (1979) Role of proline isomerization in folding of ribonuclease A at low temperatures. *Proc. Natl. Acad. Sci. U.S.A. 76*, 6157–6161.
13. Lin, L. N., and Brandts, J. F. (1983) Isomerization of proline-93 during the unfolding and refolding of ribonuclease A. *Biochemistry 22*, 559–563.
14. Schmid, F. X. (1983) Mechanism of folding of ribonuclease A. Slow refolding is a sequential reaction via structural intermediates. *Biochemistry 22*, 4690–4696.
15. Kiefhaber, T., and Baldwin, R. L. (1995) Kinetics of hydrogen bond breakage in the process of unfolding of ribonuclease A measured by pulsed hydrogen exchange. *Proc. Natl. Acad. Sci. U.S.A. 92*, 2657–2661.
16. 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*, 11734–11746.
17. Arnold, U., Hinderaker, M. P., Köditz, J., Golbik, R., Ulbrich-Hofmann, R., and Raines, R. T. (2003) Protein prosthesis: a non-natural residue accelerates folding and increases stability. *J. Am. Chem. Soc. 125*, 7500–7501.
18. Tam, A., Arnold, U., Soellner, M. B., and Raines, R. T. (2007) Protein prosthesis: 1,5-disubstituted[1,2,3]triazoles as *cis*-peptide bond surrogates. *J. Am. Chem. Soc. 129*, 12670–12671.
19. Marshall, G. R., Feng, J. A., and Kuster, D. J. (2008) Back to the future: ribonuclease A. *Biopolymers 90*, 259–277.
20. Udgaonkar, J. B., and Baldwin, R. L. (1990) Early folding intermediate of ribonuclease A. *Proc. Natl. Acad. Sci. U.S.A. 87*, 8197–8201.
21. Bycroft, M., Matouschek, A., Kellis, J. T., Jr., Serrano, L., and Fersht, A. R. (1990) Detection and characterization of a folding intermediate in barnase by NMR. *Nature 346*, 488–490.
22. Radford, S. E., Dobson, C. M., and Evans, P. A. (1992) The folding of hen lysozyme involves partially structured intermediates and multiple pathways. *Nature 358*, 302–307.
23. Krebs, H., Schmid, F. X., and Jaenicke, R. (1983) Folding of homologous proteins. The refolding of different ribonucleases is independent of sequence variations, proline content and glycosylation. *J. Mol. Biol. 169*, 619–635.
24. Krebs, H., Schmid, F. X., and Jaenicke, R. (1985) Native-like folding intermediates of homologous ribonucleases. *Biochemistry 24*, 3846–3852.
25. Pradeep, L., Shin, H. C., and Scheraga, H. A. (2006) Correlation of folding kinetics with the number and isomerization states of prolines in three homologous proteins of the RNase family. *FEBS Lett. 580*, 5029–5032.
26. Reimer, U., Scherer, G., Drewello, M., Kruber, S., Schutkowski, M., and Fischer, G. (1998) Side-chain effects on peptidyl-prolyl *cis*/*trans* isomerisation. *J. Mol. Biol. 279*, 449–460.
27. Schulenburg, C., Martinez-Senac, M. M., Löw, C., Golbik, R., Ulbrich-Hofmann, R., and Arnold, U. (2007) Identification of three phases in onconase refolding. *FEBS J. 274*, 5826–5833.
28. Ikai, A., and Tanford, C. (1973) Kinetics of unfolding and refolding of proteins. I. Mathematical analysis. *J. Mol. Biol. 73*, 145–163.
29. Bachmann, A., and Kiefhaber, T. (2005) Kinetic mechanisms in protein folding , Vol. 1, Wiley-VCH, Weinheim.
30. Gorbatyuk, V., Chen, Y. C., Wu, Y. J., Youle, R. J., and Huang, T. H. (1999) Sequence-specific ^{1}H, ^{13}C and ^{15}N resonance assignments of recombinant onconase/P-30 protein. *J. Biomol. NMR 15*, 343–344.
31. 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.
32. Johnson, B. A. (2004) Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol. Biol. 278*, 313–352.
33. Balbach, J., and Schmid, F. X. (2000) Proline isomerization and its catalysis in protein folding , Oxford University Press, Oxford.
34. Schmid, F. X. (1981) A native-like intermediate on the ribonuclease A folding pathway. 1. Detection by tyrosine fluorescence changes. *Eur. J. Biochem. 114*, 105–109.
35. Arnold, U., Rücknagel, K. P., Schierhorn, A., and Ulbrich-Hofmann, R. (1996) Thermal unfolding and proteolytic susceptibility of ribonuclease A. *Eur. J. Biochem. 237*, 862–869.
36. Juneja, J., and Udgaonkar, J. B. (2002) Characterization of the unfolding of ribonuclease A by a pulsed hydrogen exchange study: evidence for competing pathways for unfolding. *Biochemistry 41*, 2641–2654.
37. Makhatadze, G. I., Clore, G. M., and Gronenborn, A. M. (1995) Solvent isotope effect and protein stability. *Nat. Struct. Biol. 2*, 852–855.

38. Eyles, S. J., Radford, S. E., Robinson, C. V., and Dobson, C. M. (1994) Kinetic consequences of the removal of a disulfide bridge on the folding of hen lysozyme. *Biochemistry 33*, 13038–13048.
39. Kolbanovskaya, E. Y., Terwisscha van Scheltinga, A. C., Mukhortov, V. G., Ardelt, W., Beintema, J. J., and Karpeisky, M. Y. (2000) Localization and analysis of nonpolar regions in onconase. *Cell. Mol. Life Sci. 57*, 1306–1316.
40. Matheson, J. R. R., and Scheraga, H. A. (1978) A method for predicting nucleation sites for protein folding based on hydrophobic contacts. *Macromolecules 11*, 819–829.
41. Gahl, R. F., Narayan, M., Xu, G., and Scheraga, H. A. (2008) Dissimilarity in the oxidative folding of onconase and ribonuclease A, two structural homologues. *Protein Eng. Des. Sel. 21*, 223–231.