Characterization of folding the four-helix bundle protein Rop by real-time NMR

Nico A.J. van Nuland¹, Christopher M. Dobson² and Lynne Regan^{3,4,5}

¹Department of Physical Chemistry and Institute of Biotechnology, University of Granada, 18071 Granada, Spain, ²Oxford Center for Molecular Sciences, University of Oxford, South Parks Road, Oxford OX1 3QT, UK and Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK, ³Molecular Biophysics and Biochemistry and ⁴Chemistry, Yale University, 266 Whitney Avenue, New Haven, CT 06520, USA

 ^5To whom correspondence should be addressed. E-mail: lynne.regan@yale. edu

Rop is a four-helix bundle protein composed of two identical helix-loop-helix monomers. Protein folding monitored by stopped-flow fluorescence or CD exhibits biphasic kinetics when folding to low final denaturant concentrations. As the final concentration of denaturant is increased, the amplitude of the fast phase decreases, until at the highest concentrations the kinetics appear monophasic. We propose that the fast phase represents the formation of an intermediate. Here, we use real-time NMR to detect the formation of this intermediate and to characterize its structural features.

Keywords: protein folding/real-time NMR

Introduction

Rop is an RNA-binding protein that is involved in the regulation of copy number in ColE1 plasmids (Cesareni et al., 1991; Altschul et al., 1990; Cesareni et al., 1982; Eberle et al., 1991; Eguchi and Tomizawa, 1990, 1991; Gregorian and Crothters, 1995; Marino et al., 1995; Polisky, 1995; Predki et al., 1995; Fitzwater et al., 1998; Guijarro et al., 1998). It is a four-helix bundle protein formed by the association of two identical helix-loop-helix monomers, each of which is 63 amino acids long (Banner et al., 1987) (Fig. 1). Rop contains no proline residues, disulfide crosslinks or co-factors. The folding and unfolding of Rop are unusual in that they are substantially slower than that expected for such a small, apparently simple protein (Gittelman and Matthews, 1990; Milla and Sauer, 1994; Wendt et al., 1995; Zitzewitz et al., 1995; Sauer et al., 1996; Munson et al., 1996, 1997; Nagi and Regan, 1997; Nagi et al., 1999; Clarke et al., 1997; Plaxco et al., 1997, 1998; Itzhaki and Evans, 1996; Ferguson et al., 1999; Wolynes, 1997; Riddle et al., 1999; Willis et al., 2000).

We have performed extensive studies of the folding and unfolding of wild-type Rop and its variants, using fluorescence and CD (Munson *et al.*, 1996, 1997; Nagi and Regan, 1997; Nagi *et al.*, 1999; Dalal *et al.*, 2008). When folding of wild-type Rop is initiated to low final concentrations of denaturant, biphasic folding kinetics are observed (Fig. 2). Both the amplitude and rate of the fast phase increase with increasing protein concentration. The amplitude of the fast phase decreases as the concentration of denaturant in the final folding solution is increased. Eventually, at a final concentration of GuHCl of 1.75 M or higher, the amplitude of the fast phase has decreased to such an extent that the kinetics appear monophasic. These observations, together with the results of double-jump experiments and studies of the effect on protein folding of the length of the connecting loop between the helices, led us to suggest that the biphasic kinetics reflect the formation of a dimeric intermediate in which there is helicity and hairpin formation, but the final, fully folded form of the protein has not been achieved (Nagi and Regan, 1997; Nagi et al., 1999; Dalal et al., 2008). The intermediate is unstable at higher concentrations of GuHCl, and hence the folding kinetics appear monophasic under such conditions.

Although the results of the fluorescence and CD experiments allow us to propose a model for Rop's folding, they are limited to monitoring the overall behavior of the protein, rather than providing information on a residue-by-residue basis. In contrast, appropriate NMR methods allow folding to be monitored in real time on residue-by-residue basis (Schanda et al., 2007; Zeeb and Balbach, 2004; Canet et al., 2003; Mizuguchi et al., 2003; Roy and Jennings, 2003; Buevich and Baum, 2002; Killick et al., 1999; Balbach et al., 1999; Chiti et al., 1999; Van Nuland et al., 1998a). Here, we describe the results of a series of real-time NMR studies in which we investigate Rop's folding and the nature of the proposed intermediate in greater detail. In these experiments, protein folding is initiated by dilution of a concentrated solution of denatured protein into refolding buffer, directly in the NMR tube. Experimental conditions are such that 1D spectra, with good signal to noise, can be acquired approximately every 2 s, with the first spectrum completed 2.32 s after initiation of folding. It is feasible to follow the folding of Rop on such a time scale because its folding and unfolding kinetics are substantially slower than those of other small proteins of comparable size. The results support a model in which an intermediate is populated when refolding is initiated into low final denaturant concentrations and provide more detailed information as to its character.

Materials and methods

Material

Rop protein was expressed and purified as described previously (Munson *et al.*, 1994).

Real-time NMR spectroscopy

NMR spectra were recorded using a home-built NMR spectrometer operating at 600.2 MHz, and processed using

Abbreviations: CD, circular dichroism; GuHCl, guanidine hydrochloride; GuDCl, deuterated GuHCl; NMR, nuclear magnetic resonance.

[©] The Author 2008. Published by Oxford University Press. All rights reserved. For Permissions, please e-mail: journals.permissions@oxfordjournals.org



Fig. I. Ribbon representation of the structure of Rop. N and C termini of the two homodimers are indicated.



Fig. 2. Illustrative example of the biphasic kinetics observed when folding to low final concentration of GuHCl. Identical results are observed if folding is monitored by stopped-flow fluorescence or by stopped-flow CD.

FELIX (Hare Research). The spectral width was 8000 Hz. The residual water was saturated by weak on-resonance irradiation during the 0.8 s relaxation delay. Chemical shifts are expressed relative to sodium-2,2-dimethyl-2-silapentane-

5-sulphonate (DSS). For all refolding experiments, 50 µl of 160 mg/ml Rop in 7.88 M GuDCl (Pace, 1986) was diluted into a 475 µl volume of refolding buffer containing the appropriate GuDCl concentrations. Final conditions were 1.25 or 1.75 M GuDCl in 100 mM NaPi, 200 mM NaCl, 1 mM DTT, pH 7.0. The final protein concentration in all NMR samples was 1.3 mM. One hundred and twenty-eight 1D NMR spectra with two scans each were recorded for refolding at 20°C with an acquisition time of 0.256 s. To initiate the refolding of Rop, denatured protein was injected at a fixed time point during the recording of the 128 1D FIDs (Balbach et al., 1995). The first spectrum was acquired 2.32 s after injection and subsequent spectra were then recorded at 2.32 s intervals. Refolding kinetics were fitted to a single exponential function using the Origin software (Microcal Software, Inc.).

Results

Figure 3 shows a series of 1D NMR spectra recorded during the refolding of Rop. Close-ups of two informative regions of the spectra are shown. Denatured protein (that had been completely exchanged with D_2O) was rapidly injected into refolding buffer to a final GuDCl concentration of either 1.25 M (Fig. 3A) or 1.75 M (Fig. 3B). A final concentration of 1.25 M GuDCl provides refolding conditions in which biphasic kinetics are observed by fluorescence and CD, with a significant fast phase amplitude. A final concentration of 1.75 M GuDCl provides refolding conditions in which the kinetics observed by fluorescence and CD are monophasic. The spectra were recorded at 2.32 s intervals, the first scan being completed 1.16 s after injection.

The first and last spectra from the experiments described above are shown in Fig. 4. Figure 4A is the spectrum at 2.32 s (average of two scans) for refolding to 1.75 M GuDCl, whereas Fig. 4C is the spectrum at 2.32 s (average of two scans) for refolding to 1.25 M GuDCl. Figure 4B and D is the final spectra, obtained at long refolding times for refolding to 1.75 and 1.25 M GuDCl, respectively. The final spectra (Fig. 4B and D) for both refolding conditions are identical to each other and to that of the native state, indicating that under both conditions the protein has folded completely by the final time point. In contrast, the first spectra (Fig. 4A and C) are different depending on whether folding is initiated to a final concentration of 1.75 GuDCl or 1.25 M GuDCl. Regions of the spectra where these differences are most clear are indicated on the figure.

We used the increase in intensity of the resonances in the high-field region of the spectrum, representing methyl groups close to aromatic residues in the folded structure, as indicators of the appearance of the native state of the protein. These resonances are readily identified because ring current shifts cause substantial perturbations in their chemical shift values relative to random coil values. Figure 5 shows a plot of the increase in intensity of such resonances as a function of time. For refolding to both 1.25 M GuDCl and 1.75 M GuDCl, folding kinetics monitored in this fashion shows single-exponential behavior.

Fitting the data to a single exponential yield values of $0.162 \pm 0.009 \text{ s}^{-1}$ and $0.078 \pm 0.003 \text{ s}^{-1}$ for the rate constants of folding into 1.25 M GuDCl and 1.75 M GuDCl, respectively. The kinetic data derived from these



Fig. 3. Stacked plot of 1D ¹H NMR spectra of Rop in D_2O at 20°C. Spectra were recorded at time points between 2.32 and 278 s after initiation of refolding from 8.0 M GuDCl in 100 mM NaPi, 200 mM NaCl, 1 mM DTT, pH 7.0 to final GuDCl concentrations of 1.75 M (**A**) and 1.25 M (**B**). The region shown at the top contains resonances from aromatic groups and that at the bottom from methyl and methylene groups. Lines at the earlier time points are slightly broader, reflecting residual inhomogeneity of the mixture at early times or slow conformational averaging on the NMR time scale (Van Nuland *et al.*, 1998b).



Fig. 4. Comparison of the first and last spectra for Rop refolding to final GuDCl concentrations of 1.25 or 1.75 M. (A) Spectrum recorded 2.32 s (average over two scans) after initiating refolding to a final GuDCl concentration of 1.75 M. (B) Spectrum recorded following complete refolding to a final GuDCl concentration of 1.75 M. (C) Spectrum recorded 2.32 s (average over two scans) after initiating refolding to a final GuDCl concentration of 1.25 M. (D) Spectrum recorded following complete refolding to a final GuDCl concentration of 1.25 M.

experiments, in which formation of fully folded protein is monitored by following the up-field shifted resonances, were then used to reconstruct all NMR spectra at intermediate time points between the first and the final spectrum by a linear combination of the first and last recorded spectra.



Fig. 5. Refolding kinetics of Rop determined from the change in intensity of the native signals between 0.7 and 0.4 ppm as a function of time. The solid line represents the best fit to a single-exponential of data points up to 150 s. The fitting procedure yields values of $0.078 \pm 0.003 \text{ s}^{-1}$ and $0.162 \pm 0.009 \text{ s}^{-1}$ for the rate constants of folding into 1.75 M GuDCl and 1.25 M GuDCl, respectively.

Figure 6A and B shows such a comparison for a single time point during refolding to 1.75 M GuDCl and 1.25 M GuDCl, respectively.

For both sets of data, the reconstructed spectra are closely similar to the experimental spectra, regardless of the time at which the analysis was performed. Using the first and last of the recorded spectra, and the parameters from the kinetic analysis, the spectra at time zero can be reconstructed (Fig. 7). The spectra at the shorter times after initiating refolding in 1.75 M GuDCl and the time zero extrapolated



Fig. 6. Reconstruction of NMR spectra during the refolding of Rop, using a linear combination of the first and last recorded spectra, and the rate constants determined in Fig. 5. The reconstructed spectra at 13.9 and 6.95 s after initiating refolding in 1.75 M GuDC1 (A) and 1.25 M GuDC1 (B) are shown as examples. Both low- and high-field spectral regions are shown and are plotted in absolute intensity. In both A and B, from top to bottom: simulated spectrum, experimental spectrum, difference between simulated and experimental spectrum.



Fig. 7. Reconstructed spectrum of an intermediate. (A) Spectrum of Rop in 5.0 M GuDCl, the unfolded state. (B) Spectrum of Rop at t = 0, reconstructed using the first spectrum, recorded 2.32 s after initiation of folding in 1.75 M GuDCl and subtracting the last spectrum weighted by the population of native molecules predicted from the kinetics shown in Fig. 5. (C) Spectrum of Rop at t = 0, reconstructed using the first spectrum, recorded 2.32 s after initiation of folding in 1.25 M GuDCl and subtracting the last spectrum, recorded 2.32 s after initiation of folding in 1.25 M GuDCl and subtracting the last spectrum weighted by the population of native molecules expected from the kinetics shown in Fig. 5. The spectra shown in A and B of this figure are very similar, whereas the spectrum shown in C is clearly different, as indicated.

spectrum (Fig. 7B) are similar to that of the unfolded spectrum shown in Fig. 5A. Chemical shift dispersion is similar, but the lines are slightly broader, reflecting residual inhomogeneity of the mixture at these early times. Data obtained for refolding to 1.75 M GuDCl indicate that all spectral changes observed during refolding can be attributed to a single, highly co-operative two-state transition from the initial unfolded state into the final native state, in agreement with results obtained using optical probes to monitor the refolding process.

The time zero extrapolated spectrum when refolding is performed in 1.25 M GuDCl, however (Fig. 7C), is clearly different from both the extrapolated spectrum when refolding is performed in 1.75 M and the spectrum of fully unfolded protein. The time zero extrapolated spectrum, when refolding is performed in 1.25 M GuDCl, shows some of the unfolded spectral features, but in addition, its spectrum contains resonances that are typical of the native state. Particularly striking examples are the low-field shifted His peaks and the unassigned peak at 6.95 ppm. Analysis of the intensity of these signals shows that they have reached their final native intensity at the first time point after initiating refolding. In contrast, other aromatic resonances and those in the high-field region of the spectrum, which correspond to methyl groups close to aromatic residues in the native structure, have intensities substantially lower and are as expected from the kinetic analysis shown in Fig. 5.

It is clear, therefore, that the first spectrum after initiating folding to 1.25 M GuDCl spectrum cannot simply be reconstructed using a linear combination of the unfolded spectrum and the native state spectrum. This contrasts dramatically with the data for refolding to 1.75 M GuDCl, where the first spectrum is a simple linear combination of the spectra of the folded and unfolded states. This result indicates that after initiating refolding to 1.25 M GuDCl, in the dead-time of the experiment, a state is formed which is not completely unstructured, but which has acquired some, but not all, of its native conformation. As shown in Fig. 6, all spectra at intermediate time points can be reconstructed using the first and final spectra, indicating that refolding occurs in a co-operative manner from this initial state to the final fully native state. This observation is in concurrence with the observation of a fast phase when refolding is monitored by stopped-flow optical techniques (Fig. 2).

An alternative way to present the data is to compare the kinetics of appearance of the folded state His resonances and aromatic-shifted methyl resonances for folding to 1.25 M GuDCl. Figure 8 shows a comparison of the kinetics of appearance of the folded state His resonances. It is clear that when refolding to 1.25 M GuHCl, at the first time point these resonances are already at their final positions, whereas the aromatic-shifted methyl resonances appear more slowly.

Finally, real-time NMR can also be used to monitor the unfolding of Rop. Figure 9 shows a series of 1D spectra taken at the indicated times after initiation of unfolding to a final concentration of 5 M GuDCl. Unfolding is slow, indeed at 12.2 s after injection into 5 M GuDCl, no indication of unfolding is evident. At longer times the protein unfolds, and resonances for both the folded and the unfolded state His residues are evident at intermediate times. There is no evidence of the formation of any unfolding intermediates—all the spectral changes are concerted, and at each time point the observed spectrum corresponds to a weighted sum of the folded and unfolded spectra.



Fig. 8. Kinetics of folding to 1.25 M GuDCl, following different resonances. Comparison of the change in intensity of the native state His resonances (closed squares) and the native state aromatic-shifted methyl resonances (open circles) for refolding to 1.25 M GuDCl.



Fig. 9. Rop unfolding to a final concentration of 5 M GuDCl. Unfolding was initiated into 5 M GuDCl and 1D spectra acquired at the times indicated. The positions of the His resonances in the folded and unfolded states are indicated.

Discussion

In previous studies, we observed that refolding Rop to low denaturant concentrations gave rise to biphasic kinetics, whereas refolding to higher denaturant concentrations gave rise to monophasic kinetics. In both refolding conditions, the kinetics are identical, whether monitored by stopped-flow fluorescence or CD. These observations, combined with extensive additional measurements, including double-jump experiments and the study of natural and engineered Rop variants, led us to propose that the behavior we observed reflected the formation of an intermediate species. Here, we report the results of real-time NMR experiments in which we are able to detect and characterize the properties of this intermediate.

In the 1D ¹H spectra of Rop. a number of resonances are sufficiently resolved that their appearance or disappearance can be followed individually during folding. When refolding is initiated to a low final concentration of GuDCl, conditions in which we propose an intermediate is populated, not all resonances associated with the folded state of the protein appear with the same kinetics. We detect a species in which certain resonances have achieved their final, native position and amplitude, whereas others have not. In contrast, when refolding is initiated to a final concentration of denaturant at which no fast phase is detected by fluorescence or CD, all the resonances associated with the native state of the protein appear with the same kinetics. These observations provide direct support for the proposal that an intermediate is formed during folding to low final concentrations of denaturant. The real-time NMR results are important, because they reveal unique spectral characteristics that are associated with the intermediate, but with neither the fully folded or fully unfolded states.

Acknowledgements

We are grateful to members of the Regan lab for gifts of purified Rop. We thank Denis Carnet for invaluable assistance with combined stopped-flow fluorescence and CD experiments. We would also like to thank David Baker, Athena Nagi and Jenny Yang for useful discussions.

Funding

N.A.J.van N. was supported by the TMR program of the European Community. This is a contribution from the Oxford Centre for Molecular Sciences, which is supported by EPSRC, BBSRC and MRC. The research of C.M.D. is supported in part by an International Research Scholars award from the Howard Hughes Medical Institute. We acknowledge NATO for a travel grant and NIH GM49146-01A1. L.R. was a Dreyfus Teacher-Scholar.

References

- Altschul,S.F., Gish,W., Miller,W., Myers,E.W. and Lipman,D.J. (1990) *J. Mol. Biol.*, **215**, 403–410.
- Balbach, J., Forge, V., van Nuland, N.A.J., Winder, S.L., Hore, P.J. and Dobson, C.M. (1995) Nat. Struct. Biol., 2, 865–870.
- Balbach, J., Steegborn, C., Schindler, T. and Schmid, F.X. (1999) J. Mol. Biol., 285, 829–842.
- Banner, D.W., Kokkinidis, M. and Tsernoglou, D. (1987) J. Mol. Biol., 196, 657-675.
- Buevich, A.V. and Baum, J. (2002) J. Am. Chem. Soc., 124, 7156-7162.
- Canet, D., Lyon, C.E., Scheek, R.M., Robillard, G.T., Dobson, C.M., Hore, P.J. and van Nuland, N.A. (2003) J. Mol. Biol., 330, 397–407.
- Cesareni,G., Muesing,M.A. and Polisky,B. (1982) Proc. Natl Acad. Sci. USA, **79**, 6313–6317.
- Cesareni,G., Helmer-Citterich,M. and Castagnoli,L. (1991) *Trends Genet.*, 7, 230–235.
- Chiti, F., Webster, P., Taddei, N., Clark, A., Stefani, M., Ramponi, G. and Dobson, C.M. (1999) *Proc. Natl Acad. Sci. USA*, **96**, 3590–3594.
- Clarke, J., Hamill, S.J. and Johnson, C.M. (1997) J. Mol. Biol., 270, 771-718.
- Dalal,S., Canet,D., Kaiser,S.E., Dobson,C.M. and Regan,L. (2008) PEDS, 1-9.
- Eberle, W., Pastore, A., Sander, C. and Rosch, P. (1991) J. Biomol. NMR, 1, 71–82.
- Eguchi, Y. and Tomizawa, J. (1990) Cell, 60, 199-209.
- Eguchi, Y. and Tomizawa, J. (1991) J. Mol. Biol., 220, 831-842.
- Ferguson, N., Capaldi, A.P., James, R., Kleanthous, C. and Radford, S.E. (1999) J. Mol. Biol., 286, 1597–1608.

- Fitzwater, T., Zhang, X.Y., Elble, R. and Polisky, B. (1988) *EMBO J.*, 7, 3289–3297.
- Gittelman, M.S. and Matthews, C.R. (1990) Biochemistry, 29, 7011-7020.
- Gregorian, R.S. Jr and Crothers, D.M. (1995) J. Mol. Biol., 248, 968-984.
- Guijarro, J.I., Morton, C.J., Plaxco, K.W., Campbell, I.D. and Dobson, C.M. (1998) *J. Mol. Biol.*, **276**, 657–667.
- Itzhaki, L.S. and Evans, P.A. (1996) Protein Sci., 5, 140-146.
- Killick,T.R., Freund,S.M. and Fersht,A.R. (1999) Protein Sci., 8, 1286–1291.
- Marino, J.P., Gregorian, R.S. Jr, Csankovszki, G. and Crothers, D.M. (1995) Science, 268, 1448–1454.
- Milla, M.E. and Sauer, R.T. (1994) Biochemistry, 33, 1125-1133.
- Mizuguchi, M., Kroon, G.J., Wright, P.E. and Dyson, H.J. (2003) J. Mol. Biol., 328, 1161–1171.
- Munson, M., Predki, P.F. and Regan, L. (1994) Gene, 144, 59-62.
- Munson, M., Balasubramanian, S., Fleming, K.G., Nagi, A.D., O'Brien, R., Sturtevant, J.M. and Regan, L. (1996) Protein Sci., 5, 1584–1593.
- Munson, M., Anderson, K. and Regan, L. (1997) Fold Des, 2, 77-87.
- Nagi,A.D. and Regan,L. (1997) Fold Des, 2, 67-75.
- Nagi, A.D., Anderson, K.S. and Regan, L. (1999) J. Mol. Biol., 286, 257-265.
- Pace, C.N. (1986) Methods Enzymol., 131, 266–280.
- Parker, M.J. and Clarke, A.R. (1997) Biochemistry, 36, 5786-5794.
- Pelton, J.G. and Comolli, L.R. (1998) J Biomol NMR, 11, 463-464.
- Plaxco,K.W., Spitzfaden,C., Campbell,I.D. and Dobson,C.M. (1997) J. Mol. Biol., 270, 763–770.

Plaxco,K.W., Simons,K.T. and Baker,D. (1998) J. Mol. Biol., 277, 985–994. Polisky,B. (1988) Cell, 55, 929–932.

- Predki, P.F., Nyak, L.M., Gottlieb, M.B.C. and Regan, L. (1995) *Cell*, **80**, 41–50.
- Riddle,D.S., Grantcharova,V.P., Santiago,J.V., Alm,E., Ruczinski,I. and Baker,D. (1999) *Nat. Struct. Biol.*, **6**, 1016–1024.
- Roy, M. and Jennings, P.A. (2003) J. Mol. Biol., 328, 693-703.
- Sauer, R.T., Milla, M.E., Waldburger, C.D., Brown, B.M. and Schildbach, J.F. (1996) FASEB J., 10, 42–48.
- Schanda, P., Forge, V. and Brutscher, B. (2007) Proc. Natl. Acad. Sci. USA, 104, 11257–11262.
- Van Nuland, N.A.J., Chiti, F., Taddei, N., Raugei, G., Ramponi, G. and Dobson, C.M. (1998a) J. Mol. Biol., 283, 883–891.

Van Nuland,N.A.J., Forge,V., Balbach,J. and Dobson,C.M. (1998b) Acc. Chem. Res., 32, 773–780.

- Wendt,H., Berger,C., Baici,A., Thomas,R.M. and Bosshard,H.R. (1995) Biochemistry, 34, 4097–4107.
- Willis, M.A., Bishop, B., Regan, L. and Brunger, A.T. (2000) *Structure*, 8, 1319–1328.
- Wolynes, P.G. (1997) Proc. Natl Acad. Sci. USA, 94, 6170-6175.
- Zeeb, M. and Balbach, J. (2004) Methods, 34, 65-74.
- Zitzewitz, J.A., Bilsel, O., Luo, J., Jones, B.E. and Matthews, C.R. (1995) Biochemistry, 34, 12812–12819.

Received November 27, 2007; accepted November 28, 2007

Edited by Regina Murphy