Forcing Thermodynamically Unfolded Proteins to Fold*
(Received for publication, December 18, 1997, and in revised form, January 14, 1998)

Ilia Baskakov and D. Wayne Bolen‡
From the Department of Human Biological Chemistry and Genetics, Sealy Center for Structural Biology, University of Texas Medical Branch, Galveston, Texas 77555-1052

A growing number of biologically important proteins have been identified as fully unfolded or partially disordered. Thus, an intriguing question is whether such proteins can be forced to fold by adding solutes found in the cells of some organisms. Nature has not ignored the powerful effect that the solution can have on protein stability and has developed the strategy of using specific solutes (called organic osmolytes) to maintain the structure and function cellular proteins in organisms exposed to denaturing environmental stresses (Yancey, P. H., Clark, M. E., Hand, S. C., Bowlus, R. D., and Somero, G. N. (1982) Science 217, 1214–1222). Here, we illustrate the extraordinary capability of one such osmolyte, trimethylamine N-oxide (TMAO), to force two thermodynamically unfolded proteins to fold to native-like species having significant functional activity. In one of these examples, TMAO is shown to increase the population of native state relative to the denatured ensemble by nearly five orders of magnitude. The ability of TMAO to force thermodynamically unstable proteins to fold presents an opportunity for structure determination and functional studies of an important emerging class of proteins that have little or no structure without the presence of TMAO.

A growing number of biologically important proteins have been identified as fully or partially disordered under physiological conditions (e.g., different classes of DNA-binding proteins (1), transactivation domains of transcription factors (2–6), non-Aβ component of Alzheimer’s disease amyloid plaque precursor implicated in Alzheimer’s disease (7), and others (8, 9). The issue of shifting a protein or domain from an unfolded to a folded ensemble is a topic of interest not only for these proteins, but also for a host of marginally stable proteins. A question of interest is whether such proteins can be induced to adopt unique and functionally important ordered structures by addition of solutes found in the cells of some organisms.

According to Anfinsen, “The native conformation of protein is determined by the totality of interatomic interactions and by the amino acid sequence, in a given environment” (10). Although the statement by Anfinsen acknowledges the importance of both amino acid sequence and the physiological milieu in defining the native (Gibbs energy minimum) conformation of proteins, the overwhelming emphasis in the protein folding field has been on the interatomic interaction aspect of the process (11). Nature, however, has not ignored the powerful effect that the solution can have on protein stability and has developed the strategy of using specific solutes (called organic osmolytes) to maintain the structure of proteins in cells exposed to denaturing environmental stresses (12). Thus, through the power of natural selection, solutes were evolved that have exceptional ability to promote the native states of proteins in the presence of denaturing stresses. The implication is that in the absence of denaturing stresses, osmolytes continue to exert a force to fold proteins that are highly unstable in an aqueous environment.

Two examples of proteins that exist in the unfolded state in buffer are considered. These two unfolded proteins are forced to fold cooperatively into native-like species by the presence of trimethylamine N-oxide (TMAO),1 a solute found in the cells of elasmobranchs that stabilizes the intracellular proteins against the presence of urea (12, 13). The two proteins are: 1) reduced and carboxyamidated ribonuclease T1 (RCAM-T1), a chemical modification that releases conformational constraints and greatly stabilizes the unfolded state relative to the native state, and 2) staphylococcal nuclease mutant protein (T62P), in which replacement of a threonine with proline in a helix greatly destabilizes the native state relative to the unfolded ensemble. We show that TMAO folds these two highly unfolded proteins to species that acquire functional activity and secondary and tertiary structures much like that of wild type or unmodified protein.

EXPERIMENTAL PROCEDURES

Reduction and carboxyamidation of RNase T1 was performed using iodoacetamide, in the manner described by Mücke and Schmid (14). Ribonuclease T1 with all four cysteines carboxyamidated migrated as a single band in native polyacrylamide gel electrophoresis and as a single peak when chromatographed on a Phenomenex Biosep SEC-S3000 gel-filtration column. The assay for free thiols with Ellman’s reagent was negative. RNase T1 and RCAM-T1 concentrations were determined spectrophotometrically at 278 nm (1 mg/ml = 1.9 absorbance units). Activity measurements of RNase T1 and RCAM-T1 are described in Ref. 15. Proteins equilibrated at different concentrations of TMAO (in 30 mM MOPS, 0.1 M NaCl, 2 mM EDTA, 0.5 mM mg/bene serum albumin, pH 7.0, 25 °C) for 4 h was diluted 20-fold into the same TMAO buffer containing 75 μg GpC at 25 °C. The resulting increase in absorbance at 257 nm caused by cleavage of GpC was recorded and the initial velocity presented in terms of ΔA257/min/mg of protein.

TMAO-induced folding of RCAM-T1 (10 μg/ml) was monitored using a Spex FluoroMax spectrofluorimeter and involved intrinsic fluorescence (278 or 295 nm excitation and 319 nm emission) measurements in solution (30 mM MOPS, 0.1 M NaCl, 2 mM EDTA, pH 7.0) as a function of TMAO concentration. All measurements were made using 1-cm square cuvettes thermostated at 25 °C, and all data were corrected for the contribution of the respective soluble concentrations. CD spectra were recorded at 40 nm/min in 0.1-cm cuvettes (0.3 mg/ml protein) for the peptide region (<250 nm), and in 1-cm cuvettes (1.5 mg/ml protein concentration) in the aromatic region (310–260 nm) in 10 mM Tris-HCl, pH 7.0, buffer at 25 °C. The bandwidth was 1.5 nm, and each spectrum shown is the result of eight spectra accumulated and averaged. All spectra were corrected for the contributions of the respective buffers.

1 The abbreviations used are: TMAO, trimethylamine N-oxide; RCAM-T1, reduced and carboxyamidated ribonuclease T1; MOPS, 4-morpholinepropanesulfonic acid; wt, wild type; SNase, staphylococcal nuclease T62P, SNase mutant protein.

‡ To whom correspondence should be addressed. Tel.: 409-772-0754; Fax: 409-747-4751; E-mail: wbolen@hbgc.utmb.edu.

* This work was supported by NIGMS Grant 49760 and The Sealy Foundation. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

© 1998 by The American Society for Biochemistry and Molecular Biology, Inc.
Printed in U.S.A.

This paper is available on line at http://www.jbc.org
Prior to their injection, RNase T1 and RCAM-T1 samples (10
absence or presence of the indicated concentration of TMAO at 25 °C.

50
mg/ml
samples were collected upon excitation at 278 nm (in 25 mM Tris-HCl,
electrospray mass spectrometry. The fluorescence spectra of SNase
acrylamide gel electrophoresis, and molecular weight was confirmed by

Size exclusion chromatography was carried out using a Phenomenex
Biosep SEC-S3000 HPLC gel filtration column 300 × 7.80 mm, equili-
brated in buffer (30 mM MOPS, 0.2 mM NaCl, 2 mM EDTA, pH 7.0) in
the absence or presence of the indicated concentration of TMAO at 25 °C.
Prior to their injection, RNase T1 and RCAM-T1 samples (10 µg/ml)
were incubated using the buffer and [TMAO] conditions for chromatog-
raphy equilibration. In all experiments the samples were incubated at
25 °C at least 4 h before measurement.

Purification and specific activity measurements of wild type SNase
and T62P mutant protein were performed as described in Ref. 16.
Protein solutions equilibrated at different concentrations of TMAO (in
25 mM Tris-HCl, 0.1 mM NaCl, 10 mM CaCl₂, pH 8.8) for 30 min were then
diluted 60-fold into assay solutions (thermostatted at 25 °C) containing
the same buffer and solute concentrations as the sample. The resulting
increase in absorbance at 260 nm caused by the cleavage of DNA was
recorded and the initial slope presented in terms of

increase in absorbance at 260/min/mg was
used as a measure of specific activity. Wild type SNase and T62P
mutant concentrations were determined by UV absorption at 280 nm (1
mg/ml = 0.93 absorbance units), purity was confirmed by SDS-poly-
acrylamide gel electrophoresis, and molecular weight was confirmed by
electrospray mass spectrometry. The fluorescence spectra of SNase
samples were collected upon excitation at 278 nm (in 25 mM Tris-HCl,
0.1 mM NaCl, 10 mM CaCl₂ containing 50 µg/ml protein (1-cm path
length) thermostatted at 25 °C in a Spex FluoroMax spectrofluorim-
eter). All spectra were corrected for contributions from the solution.

RESULTS AND DISCUSSION

Here, we provide two examples of proteins whose unfolded
ensemble dominates in buffer, but in the presence of TMAO are
forced to fold into native-like species having significant func-
tional activity. The first of these two protein examples is
RCAM-T1, whose unfolded state is favored by a large gain in
conformational entropy derived from disruption of disulfide
bonds. With RNase T1, it has been estimated that reduction
and carboxyamidation of its two disulfide bonds (RCAM-T1)
destabilizes the protein by nearly 9 kcal/mol, resulting in a
chemically modified protein that is extensively unfolded (17).

Size exclusion chromatography in Fig. 1A show, from the elution volumes of RNase T1 and RCAM-T1, that
RCAM-T1 has a significantly expanded structure relative to
RNase T1. At TMAO concentration above 2.5 M, both RNase T1
and RCAM-T1 are found to have identical elution volumes (see Fig.
1B), indicating that the two proteins have identical de-
grees of structural compactness under these conditions. The
close correspondence of the fluorescence emission spectra
of RCAM-T1 in buffer and RNase T1 in 6 M guanidine chloride
strongly indicates extensive unfolding in both cases (Fig. 2A).
But in the presence of 2.7 M TMAO, the fluorescence emission
spectrum of RCAM-T1 changes dramatically to that of native
RNase T1. This result shows that both wt RNase T1 and
RCAM-T1 are folded in the presence of 2.7 M TMAO and that
both are essentially equivalent in terms of their fluorescence
properties. The fluorescence emission data of RCAM-T1 given
as a function of TMAO concentration in Fig. 2B illustrates that
TMAO induces RCAM-T1 to fold in a cooperative manner,
reaching a maximum in folding at concentrations above 2.5 M
TMAO.

Using GpC as a substrate, the specific activity of RNase T1 is
not observed to change substantially as a function of TMAO
concentration (see Fig. 2C). By contrast, the specific activity
of RCAM-T1 increases some 50-fold from its level of around 0.4%
of the specific activity of wt RNase T1 in the absence of TMAO,
to a specific activity in 2.7 M TMAO that is approximately 20%
of the wt RNase T1 enzyme under these conditions. So, despite
the fact that the folding of RCAM-T1 accommodates four car-
boxyamido groups into its compact structure, a substantial
amount of catalytic activity is observed for this protein. It is
important to note that regardless of the parameter used in
monitoring the stability of the protein (fluorescence or specific
activity), the sign and magnitude of the ΔG° values derived
from the data in Fig. 2, B and C, show RCAM-T1 is thermody-
namically unstable in buffer solution (ΔG° = −2.23 kcal/mol
(m = −1.77) from fluorescence data, ΔG° = −2.48 kcal/mol
(m = −1.70) from activity measurement).

In terms of secondary and tertiary structure, Figs. 3, A and
B, illustrate that spectrally, RCAM-T1 is thermodynamically
unstable in buffer solution (ΔG° = −2.23 kcal/mol
(m = −1.77)). In comparison with the wt SNase
obtained in the presence of 6 M guani-
dinium chloride (see Fig. 4A) illustrates that T62P is intrins-
ically unfolded in buffer solution. Upon gel filtration chroma-
tography, the elution volume of T62P in buffer is very much like
that of wt SNase obtained in the presence of 6 M guani-
dinium chloride (see Fig. 4A) illustrates that T62P is intrins-
ically unfolded in buffer solution. Upon gel filtration chroma-
tography, the elution volume of T62P in buffer is equivalent to
a protein with an apparent molecular mass of 34 kDa, while the
elution volume of wt SNase (molecular mass of 17 kDa) is
equivalent to a protein of 16 kDa (data not shown). However,
upon dialyzing T62P against 2.5 M TMAO in buffer, the emis-
sion spectrum of mutant protein changes dramatically, giving a
fluorescence wavelength maximum equal to that of wt SNase
native protein (Fig. 4A). In comparison with the wt SNase
and SA max is the maximum specific activity represented by the total concentration of TMAO and expressed as RCAM-T1 monitored at 319 nm emission with excitation at 278 nm (background contributions. B, CD spectra in the near-UV region: RCAM-T1 in 2.7 M TMAO (●); RCAM-T1 in the absence of TMAO (○); RNase T1 in 2.7 M TMAO (△); and RNase T1 in the absence of TMAO (○).)

The lower quantum yield of T62P arises from protein-protein association and may also be due to differences in the folded protein, SAmin is the specific activity in the absence of TMAO, and the fraction of denatured protein (U/N) can be determined from (U = (SA – SAmin)/(SAmax – SAmin)). Apparent equilibrium constants for the N = U equilibrium can be obtained from the U/N ratio at each concentration of TMAO and expressed as ∆G versus TMAO as given by Pace (29).

FIG. 3. A, CD spectra in the far-UV region: RNase T1 in 2.7 M TMAO (●), RCAM-T1 in 2.7 M TMAO (○), and RCAM-T1 in the absence of TMAO (△). The spectrum of RNase T1 in the absence of TMAO was the same as in the presence of TMAO (data not shown). B, CD spectra in the near-UV region: RCAM-T1 in 2.7 M TMAO (●); RCAM-T1 in the absence of TMAO (○); RNase T1 in 2.7 M TMAO (△); and RNase T1 in the absence of TMAO (○).

The specific activity of T62P in 2.5 M TMAO (—) and in the absence of TMAO (— — — —); wild type SNase in 2.5 M TMAO (— — —) and in 6 M GuHCl (— — — —). Fluorescence spectra are corrected for solution contributions. B, effect of TMAO on the specific activity of wild type SNase (△) and T62P mutant (○) of SNase. Also shown are the data for TMAO-induced folding of T62P as monitored by increase in specific activity, which has been normalized (□) with respect to the effect of TMAO on the activity of wild type enzyme. The solid line is a nonlinear least squares best fit of the data as described in Fig. 2. All experiments were carried out at 30 °C in 25 mM Tris-HCl, 100 mM NaCl, 10 mM CaCl2, pH 8.8. The assay in evaluating specific activity uses single strand DNA obtained by thermal denaturation of double stranded calf thymus DNA prior to the activity measurement. TMAO stabilizes double stranded DNA (the melting point of double strand DNA increases 13 °C per 1 mM TMAO (31), and the strong tendency of TMAO to favor double strand character will decrease the availability of substrate (single strand DNA), resulting in an apparent decrease in wt SNase activity as a function of TMAO concentration. A problem in availability of single strand DNA substrate would affect activities of wt SNase and T62P the same.

FIG. 4. A, fluorescence emission spectra (278 nm excitation) of T62P in 2.5 M TMAO (—) and in the absence of TMAO (— — — —); wild type SNase in 2.5 M TMAO (— — —) and in 6 M GuHCl (— — — —). Fluorescence spectra in the near-UV region: RNase T1 in 2.7 M TMAO (●), RNase T1 in 2.7 M TMAO (△); and RNase T1 in the absence of TMAO (○).

structures of the mutant and wt SNase. Regardless of the reasons for fluorescence differences between wt SNase and T62P, these data strongly suggest that TMAO forces T62P to fold.

The specific activity of T62P in the absence of TMAO is 0.08% of wt SNase, but in the presence of increasing amounts of TMAO, the specific activity of T62P increases some 2 orders of magnitude. The specific activity of T62P increases to about 90% of the specific activity of wt SNase in 2.7 M TMAO, while the specific activity of wt SNase itself decreases by about an order of magnitude in going from 0 to 2.7 M TMAO (see Fig. 4B). If the specific activity of T62P is normalized to the specific activity of wt SNase, the normalized data in Fig. 4B can be fitted in the same manner as described for RCAM-T1. Fitting gives $\Delta G^0 = -4.02 \text{ kcal/mol}$ ($m = -2.55$) and the large negative $\Delta G^0$ shows that the fraction of folded (active) T62P is extremely small in the absence of TMAO. Moreover, $\Delta G$ changes from $-4.0 \text{ kcal/mol}$ in buffer solution to $-2.9 \text{ kcal/mol}$ in 2.7 M TMAO, which represents a shift in equilibrium from the un-
folded ensemble to “native” state of nearly 5 orders of magnitude.

TMAO induces both RCAM-T1 and T62P to fold, despite the fact that their folding motifs are very different, and their reasons for being unfolded (loss of disulfide bonds in RCAM-T1 and destabilization of the native state in T62P) are also very different. The ability to force proteins to fold, regardless of the source of the instability, suggests TMAO is a versatile folding agent, a property compatible with its role in elasmobranchs (13).

By means of transfer free energy measurements, we have recently shown that the ability of TMAO to increase the driving force for protein folding is due to its solvophobic effect on the peptide backbone exposed in the unfolded state (13). Folding of thermodynamically unstable protein can be forced by a number of naturally occurring osmolytes, and the efficacy of these osmolytes in folding is related to the relative strengths of their solvophobic effect on the peptide backbone, with TMAO being the most effective osmolyte. The recent demonstration that synthetic polymers assume helical structures through the action of solvophobic effects highlights the importance of this force (18). In addition, it has long been known that organic solutes (e.g. chloroethanol and trifluoroethanol) can drive formation of helices in peptides and proteins, but this leads to non-native species (19). The driving force for helix formation has been attributed to the solvophobic (highly unfavorable) interaction of alcohol with the peptide backbone, in combination with favorable alcohol-side chain interactions (20). Our measurements show that the exposed (hydrophobic) side chains are little affected by TMAO, and in fact, the propensities of hydrophobic groups to interact with solvent are essentially the same in water as they are in TMAO solution (13). The commonly held explanation of spontaneous folding of proteins in dilute buffer stresses hydrophobic interactions as important in folding to the native state (11), and because TMAO has little effect on hydrophobic interactions, the rules for protein folding that occur in dilute buffer are unchanged by the presence of TMAO. The solvophobic effect of TMAO on the peptide backbone makes the unfolded state of protein in osmolyte solution very unfavorable relative to the folded state (21), and it is this strongly destabilizing effect of TMAO on the unfolded state that forces the protein to fold.

Based on our results and the results of others, there are two distinctly different mechanisms to force intrinsically unstable proteins to fold: one is to lower the Gibbs energy of the native state by using the binding energy of ligands to drive folding (U_{N} and N_{L} in Scheme 1) (1, 22) and the other is to make use of the solvophobic effect of osmolytes on the peptide backbone to raise the free energy of unfolded state higher than that of native protein (U_{TMAO} and N_{TMAO} in Scheme 1) (13). Examples of folding being driven by ligand binding include proteins with disordered domains that are induced to fold on binding DNA (1, 23, 24), the case of the (unstructured) prodomain of subtilisin being induced to fold when complexed with subtilisin (25, 26), and several cases of preferential anion and/or cation binding to the native state of some proteins (23, 27), shifting the equilibrium in favor of the folded state. According to the induced-fit model, ligand binding provides the driving forces for folding and the ligand serves as a template for the disordered domain to adopt its final conformation (1, 28). Since TMAO induces folding not by binding but by solvophobic effects on the backbone, it may be possible to use TMAO to determine the extent to which the induced-fit model of folding is applicable.

For the purpose of forcing proteins to fold, there is considerable advantage in leaving the side chain forces alone while using the unfavorable interaction between osmolyte and backbone as an additional force for folding. The principal advantages are: 1) the backbone is the most numerous functional group in proteins so focusing on the backbone ensures that the effect of the osmolyte is generic in scope, and 2) when the side chain forces are the same in TMAO as in buffer, the tendency will be to fold the protein to the same native species that prevails in buffer. As can be seen from the examples, significant advantage can be taken of these properties for the study of proteins that are intrinsically unstable or partially folded.

Acknowledgments—We are indebted to Dr. N. Pace for providing ribonuclease T1 and to Dr. W. Stites for providing the strains of E. coli for overexpression of SNase mutant T62P.

REFERENCES

1. Spolar, R. S., and Record, M. T. (1994) Science 263, 777–784
2. Hoy, M. V., Leuther, K. K., Kodadek, T., and Johnston, S. A. (1993) Cell 72, 587–594
3. Dahlan-Wright, K., Baumann, H., McEwan, I. J., Almlof, T., Wright, A. P. H., Gustafsson, J. A., and Hard, T. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 1699–1703
4. Cho, H. S., Liu, C. W., Damberger, F. F., Pelton, J. G., Nelson, H. C. M., and Wenner, D. E. (1996) Protein Sci. 5, 262–269
5. Schmitz, M. L., Silva, M. A. S., Altmann, H., Caicedo, M., Holak, T. A., and Baeuerle, P. A. (1994) J. Biol. Chem. 269, 25613–25620
6. Donaldson, L., and Capone, P. J. (1992) J. Biol. Chem. 267, 1411–1414
7. Tostvoie, P. H., Zhou, W., Tsau, A. W., Conway, K. A., and Lansbury, P. T. (1996) Biochemistry 35, 13709–13715
8. Zardos, J., Sanz, J. M., Gonzalez, C., Rico, M., and Balesta, J. P. G. (1997) Biochemistry 36, 9625–9635
9. Gasiot, K., Damaaschun, H., Eckert, K., Schulze-Forter, K., Maurer, H., and Muller-Frohne, M. (1995) Biochemistry 34, 13211–13218
10. Anfinsen, C. B. (1973) Science 181, 223–230
11. Dill, K. (1990) Biochemistry 29, 7153–7153
12. Yancey, P. H., Clark, M. E., Hand, S. C., Bowls, R. D., and Somero, G. N. (1982) Science 217, 1214–1222
13. Wang, A., and Bolen, D. W. (1997) Biochemistry 36, 9101–9108
14. Mucke, M., and Schmid, F. X. (1992) Biochemistry 31, 7848–7854
15. Kiehahin, T., Quans, R., Hahn, U., and Schmid, F. X. (1990) Biochemistry 29, 3053–3061
16. Shortle, D., and Meeke, A. K. (1989) Biochemistry 28, 936–944
17. Pace, C. N., Grimesley, G. R., Thomson, J. A., and Barnett, B. J. (1988) J. Biol. Chem. 263, 11820–11825
18. Nelson, J. C., Saven, J. G., Moore, J. S., and Wolynes, P. G. (1996) Science 277, 1793–1796
19. Timasheff, S. N. (1970) Acc. Chem. Res. 3, 62–68
20. Nuski, Y., and Tanford, C. (1971) J. Biol. Chem. 246, 2211–2217
21. Lin, T.-Y., and Timasheff, S. N. (1994) Biochemistry 33, 12095–12071
22. Frankel, A. S., and Kim, P. S. (1991) Cell 65, 717–719
23. Pan, B., Deng, Z., Liu, D., Ghosh, S., and Mullen, G. P. (1997) Protein Sci. 6, 1277–1247
24. Weiss, M. A., Ellenberger, T., Wobbe, C. R., Lee, J. P., Harrison, S. C., and Struhl, K. (1990) Nature 347, 575–578
25. Bryan, P., Wang, L., Hoskins, J., Kovinov, S., Strausberg, S., Alexander, P., Almog, O., Gilliland, G., and Gallagher, T. (1995) Biochemistry 34, 10310–10318
26. Shinde, U. P., Liu, J. L., and Inouye, M. (1997) Nature 389, 520–522
27. Oshatake, M., Takahashi, S., and Ots, T. (1970) J. Biochem. (Tokyo) 68, 65–70
28. Koshal, D. E. (1958) Proc. Natl. Acad. Sci. U. S. A. 54, 98–104
29. Pace, C. (1975) CRC Crit. Rev. Biochem. 3, 1–43
30. Santoro, M. M., and Bolen, D. W. (1988) Biochemistry 27, 8063–8068
31. Buche, A., Colson, P., and Housier, C. (1993) J. Biol. Mol. Struct. Dyn. 11, 85–119
