Amide–Amide and Amide–Water Hydrogen Bonds: Implications for Protein Folding and Stability

Eric S. Eberhardt and Ronald T. Raines*

Department of Biochemistry
University of Wisconsin—Madison
Madison, Wisconsin 53706

Received November 1, 1993

As a protein folds, many of its main-chain amide groups exchange hydrogen bonds with water for hydrogen bonds with other main-chain amides. The energetic contribution of this exchange to the folding and stability of proteins is unclear.[1] Theoretical,[2] calorimetric,[3] and spectroscopic[4,5] studies indicate that amide–amide hydrogen bonds form readily in nonpolar media. In contrast, amide–amide hydrogen bonds form only at extremely high amide concentrations in water.[6] Extensive efforts[7] to evaluate the contribution of amide–amide hydrogen bonds to the aqueous stability of a particular receptor–ligand complex ultimately failed to exclude contributions to binding from other forces.[8] To assess the importance of amide–amide hydrogen bonds in protein folding and stability, we have determined the relative strength of amide–amide and amide–water hydrogen bonds.

Our analyses were performed on the simple peptide Ac-Gly-[β,δ-^{13}C]Pro-OMe (**1**) and the related amide [^{13}C=O]Ac-Pro-OMe (**2**).[9] In a previous study, the kinetic barrier to prolyl peptide bond isomerization of **1** was shown to depend on the ability of the solvent to donate a hydrogen bond to the amidic carbonyl group.[10] Here, the effects of amide solvents and water on this same kinetic barrier were determined using inversion transfer ^{13}C NMR spectroscopy.[11] Solvent effects on the amide I vibrational mode of **2** were determined using IR spectroscopy.[12] The amide solvents studied mimic amide groups found in proteins.[13]

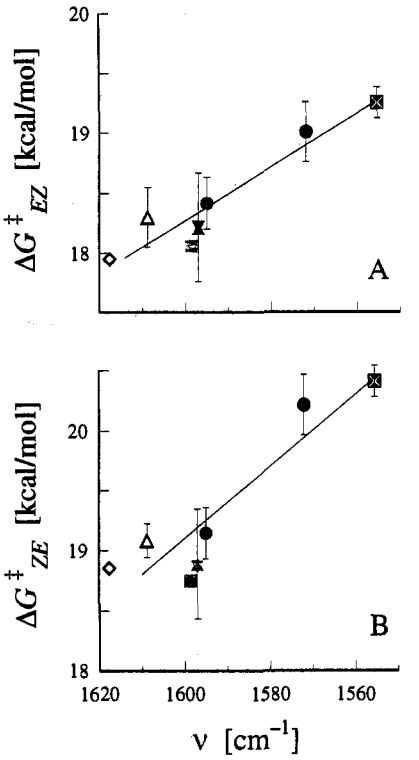
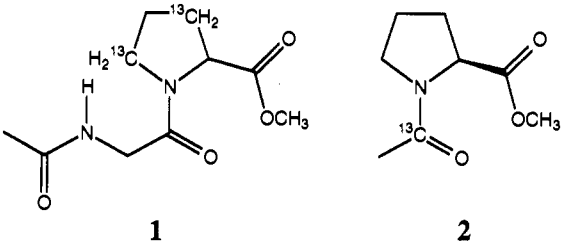


Figure 1. Plots of $\Delta G^{\ddagger}$ for isomerization of **1** vs ν of amide I vibrational mode of **2** in different solvents. The solvents (neat concentration, M; pK_a in Me_2SO, if known[24]) were as follows: ◇, dioxane (11.7); △, *N*,*N*-dimethylformamide (13.0); gray ■, *N*-methylpropionamide (10.7); gray ▲, *N*-ethylacetamide (10.8; 26.1); gray ▼, *N*-methylacetamide (13.1; 25.9); gray ●, *N*-methylformamide (17.1); ●, formamide (25.2; 23.45); ⊠, water (55.5; 32). A: *cis* to *trans*. B: *trans* to *cis*. Unweighted linear regression analysis gives slopes –0.022 ± 0.003 (A) and –0.030 ± 0.006 (B).



1 **2**

The relationship between the free energy of activation for the isomerization of **1** and the frequency of the amide I absorption band of **2** is shown in Figure 1. The amide I vibrational mode absorbs at lower frequencies with increasing strength of a hydrogen bond to the amide oxygen.[14] Also, the rate of prolyl peptide bond isomerization is related inversely to the strength of hydrogen bonds formed to the amide oxygen.[10] The axes in Figure 1 report

* Author to whom correspondence should be addressed.

(1) (a) Dill, K. A. *Biochemistry* **1990**, *29*, 7133–7155. (b) Ben-Naim, A. *J. Phys. Chem.* **1991**, *95*, 1437–1444. (c) Jeffery, G. A.; Saenger, W. *Hydrogen Bonding in Biological Structures*; Springer-Verlag: New York, 1991. (d) Shirley, B. A.; Stanssens, P.; Hahn, U.; Pace, C. N. *Biochemistry* **1992**, *31*, 725–732. (e) Schultz, J. M.; Baldwin, R. L. *Annu. Rev. Biophys. Biomol. Struct.* **1992**, *21*, 95–118. (f) Rose, G. D.; Wolfenden, R. *Annu. Rev. Biophys. Biomol. Struct.* **1993**, *22*, 381–415.

(2) (a) Dreyfus, M.; Maigret, B.; Pullman, A. *Theor. Chim. Acta* **1970**, *17*, 109–119. (b) Dreyfus, M.; Pullman, A. *Theor. Chim. Acta* **1970**, *19*, 20–37. (c) Jorgensen, W. L. *J. Am. Chem. Soc.* **1989**, *111*, 3770–3771. (d) Sneddon, S. F.; Tobias, D. J.; Brooks, C. L. *J. Mol. Biol.* **1989**, *209*, 817–820. (e) Dado, G. P.; Gellman, S. H. *J. Am. Chem. Soc.* **1992**, *114*, 3139–3141.

(3) (a) Spencer, J. N.; Garrett, R. C.; Mayer, F. J.; Merkle, J. E.; Powell, C. R.; Tran, M. T.; Berger, S. K. *Can. J. Chem.* **1980**, *58*, 1372–1375. (b) Bloemendal, M.; Somsen, G. *J. Solution Chem.* **1988**, *17*, 1067–1079.

(4) For leading references, see: (a) Tsuboi, M. *Bull. Chem. Soc. Jpn.* **1954**, *24*, 75–77. (b) Klotz, I. M.; Franzen, J. S. *J. Am. Chem. Soc.* **1962**, *84*, 3461–3466. (c) Eaton, G.; Symons, M. C. R.; Rastogi, P. P. *J. Chem. Soc., Faraday Trans. 1* **1989**, *85*, 3257–3271. (d) Nikolić, A. D.; Rozsa-Tarjani, M.; Komaromi, A.; Csanadi, J.; Petrović, S. D. *J. Mol. Struct.* **1992**, *267*, 49–54.

(5) For an especially thorough analysis, see: Dado, G.; Gellman, S. H. *J. Am. Chem. Soc.* **1993**, *115*, 4228–4245.

(6) (a) Schellman, J. A. *C. R. Trav. Lab. Carlsberg, Ser. Chim.* **1955**, *29*, 223–229. (b) Susi, H.; Timasheff, S. N.; Ard, J. S. *J. Biol. Chem.* **1964**, *239*, 3051–3054.

(7) (a) Williams, D. H.; Cox, J. P. L.; Doig, A. J.; Gardner, M.; Gerhard, U.; Kaye, P. T.; Lal, A. R.; Nicholls, I. A.; Salter, C. J.; Mitchell, R. C. *J. Am. Chem. Soc.* **1991**, *113*, 7020–7030. (b) Williams, D. H. *Aldrichimica Acta* **1991**, *24*, 71–80. (c) Doig, A. J.; Williams, D. H. *J. Am. Chem. Soc.* **1992**, *114*, 338–343.

(8) Williams, D. H. *Aldrichimica Acta* **1992**, *25*, 9.

(9) Racemic **1** was synthesized as described in ref 10 and in Hinck et al. (Hinck, A. P.; Eberhardt, E. S.; Markley, J. L. *Biochemistry* **1993**, *32*, 11810–11818). Amide **2** was synthesized in solution using standard methods (Bodanszky, M. *Peptide Chemistry*; Springer-Verlag: New York, 1988; pp 66–68).

(10) Eberhardt, E. S.; Loh, S. N.; Hinck, A. P.; Raines, R. T. *J. Am. Chem. Soc.* **1992**, *114*, 5437–5439.

(11) (a) Forsén, S.; Hoffman, R. A. *J. Chem. Phys.* **1963**, *39*, 2892–2901. (b) Led, J. J.; Gesmar, H. *J. Magn. Reson.* **1982**, *49*, 444–463. NMR experiments were performed using a Bruker AM500 instrument (125.68 MHz) and an external deuterium lock. All samples contained 0.1 M **1** in neat solvent. The presence of ^{13}C in **1** increased substantially the precision of the NMR measurements.

(12) IR experiments were done on a Nicolet 5PC spectrometer at 25 °C (except *N*-methylacetamide: 40 °C) using a ZnSe crystal. Samples contained 0.01 M **2** in neat solvent. The frequency of the amide I vibrational mode was determined to within 2 cm^{-1} and was not altered significantly by raising the temperature to 60 °C. The presence of ^{13}C in **2** decreased the frequency of its amide I vibrational mode (by 41 cm^{-1} in dioxane; by 50 cm^{-1} in water) and thereby avoided overlap with amide solvents.

(13) Amide solvents (Aldrich Chemical; St. Louis, MO) were distilled from either 4-Å molecular sieves (formamides) or CaH_2 (secondary amides). Acetamide, the simplest primary amide, was not studied due to its high melting point (82.3 °C). δ-Valerolactam, which has a relatively unhindered *cis* amide bond, was not studied because its amide I vibrational mode obscured that of **2**.

(14) Krimm, S.; Bandekar, J. *Adv. Protein Chem.* **1986**, *38*, 181–364.

0002-7863/94/1516-2149$04.50/0

independent measures of the ability of a solvent to donate a hydrogen bond to an amide oxygen.[15]

The data in Figure 1 show that water donates a strong hydrogen bond to an amidic carbonyl group. The analogous ability of secondary amide solvents, which resemble the main chain of proteins, to donate a hydrogen bond is dramatically less. The concentration of each solvent studied was >10 M, which is likely to exceed the effective concentration of peptide bonds to one another, at least during the early stages of protein folding. The relative strength of hydrogen bonds observed here therefore suggests that amide–amide hydrogen bond formation alone is unlikely to drive protein folding.[16]

The data in Figure 1 also show that formamide, which mimics the primary amide in the side chains of asparagine and glutamine residues, is a significantly better hydrogen bond donor than are any of the secondary amides studied, and it is almost as good as water.[17] This result suggests that side-chain–main-chain hydrogen bonds can contribute more to protein stability than can main-chain–main-chain hydrogen bonds. This idea is consistent with asparagine, glutamine, and glycine being preferred residues at the C-terminus of α-helices.[18] There, an amide side chain can donate a hydrogen bond to a main-chain carbonyl group, and a glycine residue can maximize the exposure of a main-chain carbonyl group to solvent water.[19]

What is the origin of the dramatic difference observed between the hydrogen bond donating abilities of secondary amides and formamide? An important contribution may arise from the effective concentration of donors, since an additional potential donor is always proximal to every hydrogen bond donated by formamide. Alternatively, the observed difference may result largely from steric constraints that restrict the number or geometry of hydrogen bonds donated by secondary amides, as has been proposed for large alcohols.[4c,20] Regardless of its origin, the observed difference in hydrogen bond donating abilities is likely to be manifested during protein folding and in folded proteins.

Approximately $^3/_4$ of the main-chain amides in globular proteins form hydrogen bonds with other main-chain amides.[21] Although the formation of such intramolecular amide–amide hydrogen bonds in water can be exothermic,[22] the results presented here and elsewhere[6] indicate that amides form stronger intermolecular hydrogen bonds with water than with other amides. We conclude that main-chain–main-chain hydrogen bonds can form only in a cooperative process, which is likely to be facilitated by the hydrophobic collapse of the unfolded protein and the consequent shedding of water molecules from main-chain amides.[1a] We also suggest that the desolvation of individual main-chain amides diminishes the stability of folded proteins.[1b,23]

Acknowledgment. We thank Professor R. L. Baldwin, Dr. G. Dado, and Dr. A. S. Edison for helpful discussions. NMR experiments were performed at NMRFAM [Grant RR02301 (NIH)]. E.S.E. is a Wharton Predoctoral Fellow. R.T.R. is a Presidential Young Investigator (NSF), Searle Scholar (Chicago Community Trust), and Shaw Scientist (Milwaukee Foundation).

(15) Further support for this interpretation of Figure 1 comes from the solvent dependence of the frequency of the ester carbonyl stretching vibration of **2**, which is related to that of the amide I vibrational mode by $\nu_{ester C=O} = (0.36 \pm 0.07)\nu_{amide I} + (1.2 \pm 0.1) \times 10^3$ for the eight solvents studied here.

(16) This study does not explicitly address any entropic contribution to protein folding and stability that may arise from the release of water molecules upon formation of amide–amide hydrogen bonds. Since the released water molecules form hydrogen bonds with bulk water, this entropic contribution is likely to be small.[1b,8]

(17) The solvent effect on the *cis–trans* equilibrium constant of the prolyl peptide bond of **1** was small ($K_{eq} = k_{ZZ}/k_{ZE} = 3.8 \pm 1.4$) and similar to that observed previously.[10] Arrhenius and Van't Hoff analyses indicated that the kinetic and thermodynamic barriers between the *cis* and *trans* isomers of **1** arise largely from enthalpic differences in all solvents studied, as was observed previously. See: Eberhardt, E. S.; Loh, S. N.; Raines, R. T. *Tetrahedron Lett.* **1993**, *34*, 3055–3056.

(18) (a) Richardson, J. S.; Richardson, D. C. *Science* **1988**, *240*, 1648–1652. (b) Dasgupta, S.; Bell, J. A. *Int. J. Pept. Protein Res.* **1993**, *41*, 499–511. (c) Chakrabartty, A.; Doig, A. J.; Baldwin, R. L. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 11 332–11 336.

(19) (a) Serrano, L.; Heira, J.-L.; Sancho, J.; Fersht, A. R. *Nature* **1992**, *356*, 453–455. (b) Fersht, A. R.; Serrano, L. *Curr. Opin. Struct. Biol.* **1993**, *3*, 75–83.

(20) (a) Kobayashi, M.; Nishioka, K. *J. Phys. Chem.* **1987**, *91*, 1247–1251. (b) Eaton, G.; Symons, M. C. R. *J. Chem. Soc., Faraday Trans. 1* **1989**, *85*, 3257–3271. (c) Eaton, G.; Symons, M. C. R.; Rastogi, P. P.; O'Duinn, C.; Waghorne, W. E. *J. Chem. Soc., Faraday Trans. 1* **1992**, *88*, 1137–1142.

(21) Stickle, D. F.; Preasta, L. G.; Dill, K. A.; Rose, G. D. *J. Mol. Biol.* **1992**, *226*, 1143–1159.

(22) Scholtz, J. M.; Marqusee, S.; Baldwin, R. L.; York, E. J.; Stewart, J. M.; Santoro, M.; Bolen, D. W. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 2854–2858.

(23) For methods to engineer protein mainchains, see: (a) Schnolzer, M.; Kent, S. B. H. *Science* **1992**, *256*, 221–225. (b) Chung, H. H.; Benson, D. R.; Schultz, P. G. *Science* **1993**, *259*, 806–809.

(24) Bordwell, F. G. *Acc. Chem. Res.* **1988**, *21*, 456–463. F. G. Bordwell, personal communication.