J|A|C|S
COMMUNICATIONS

Published on Web 08/19/2004

Backbone Thioester Exchange: A New Approach to Evaluating Higher Order Structural Stability in Polypeptides

Matthew G. Woll and Samuel H. Gellman*

Department of Chemistry, University of Wisconsin−Madison, 1101 University Avenue, Madison, Wisconsin 53706

Received May 26, 2004; E-mail: gellman@chem.wisc.edu

Understanding the forces that control higher order structure in proteins remains a topic of broad interest despite extensive research over the past half-century. Protein folding thermodynamics are generally evaluated via structural disruption. A common approach is to monitor a spectroscopic indicator of higher order structure (e.g., circular dichroism (CD)) as conditions are varied from structure-promoting ("native") to structure-destabilizing ("denaturing").[1] This variation can be achieved by raising the temperature or the concentration of a denaturing additive such as guanidinium chloride (GdmCl). Results obtained under partially denaturing conditions are typically extrapolated to native conditions in order to determine how changes in sequence are correlated with changes in stability.[1] One concern with this approach is that the limiting conformational states, native and denatured, may vary as a function of temperature or GdmCl concentration. In other words, higher order structure in the native and/or denatured states may differ between the conditions under which data are acquired and the conditions to which the results are extrapolated (e.g., room temperature and zero denaturant concentration). Indeed, there is evidence for residual secondary structure and local clustering of hydrophobic side chains in denatured proteins;[2] such residual structure could be affected by chemical denaturants or high temperature.

We describe a new strategy for evaluating higher order structural stability in proteins. This approach is expected to provide information complementary to that available from thermal or chemical denaturation studies because all measurements are made under native conditions. Our strategy offers the prospect of achieving separation between the contributions of autonomously formed secondary structure and those of higher order structure to net conformational stability.

Our approach involves replacing a backbone amide linkage in a polypeptide with a thioester linkage. The thioester-containing molecule can participate in thioester/thiol exchange reactions, which achieve equilibrium rapidly in aqueous solution at neutral pH.[3] The "backbone thioester exchange" (BTE) equilibrium constant, measured by HPLC, provides insight on the conformational stability of the thioester-containing protein analogue in its native state.[4] We illustrate the BTE method by application to bovine pancreatic polypeptide (bPP), a 36-residue protein that adopts a well-defined tertiary structure (Figure 1) and a dimeric quaternary structure.[5] bPP is a good test case for the BTE method because secondary structure, especially the long C-terminal α-helix, is likely to be populated under native conditions even in the absence of tertiary structure.

BTE will be most useful if the amide-to-thioester swap causes little or no change in the geometry or stability of the folded conformation relative to the pure polypeptide backbone. In an important precedent, Kent et al. have shown that HIV protease activity is retained after replacement of a backbone amide group with a thioester;[6] this result implies that the thioester-containing analogue adopts native tertiary and quaternary structure (native HIV protease is a homodimer). Secondary amide and thioester groups have comparable conformational properties,[7] but they differ in their hydrogen bonding abilities. Therefore, we chose a backbone amide group in bPP that does not appear to form intramolecular hydrogen bonds, the linkage between residues 9 and 10, as the site for thioester replacement. This site occurs in a short loop that connects the N-terminal polyproline II segment (residues 1−8) with the C-terminal α-helix segment (residues 15−31). Our thioester contains a variation corresponding to the $Asp_{10} \rightarrow Gly$ mutation in bPP because we used a thioglycolic acid residue at position 10 rather than a thiomalic acid residue to generate the backbone thioester group. In addition, we changed Leu_3 of the bPP sequence to Arg, to promote aqueous solubility of the N-terminal segment. This mutation was not expected to alter conformational stability because the Leu_3 side chain is solvent-exposed in the native tertiary fold[5a] and not involved in the putative dimer interface.[5b]

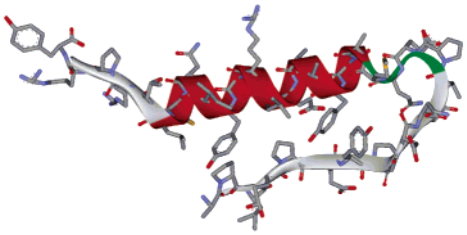

Figure 1. (top) Solution structure of bPP (structure 1BBA in the PDB) determined by NMR;[5a] (bottom) amino acid sequence of bPP, m-bPP, and t-bPP (**X** = thioglycoyl).

Before examining thioester exchange reactions, we used CD to compare the structures and stabilities of three molecules: bPP, m-bPP (the double-mutant polypeptide, $Leu_3 \rightarrow$ Arg and $Asp_1 \rightarrow$ Gly), and t-bPP (the 9/10 thioester analogue of m-bPP).[8] All three molecules display similar CD signatures at room temperature (50 μM), suggesting that all three adopt comparable folding patterns.[9] In contrast, the CD signatures of bPP variants containing single mutations that are expected to be destabilizing such as $Ala_{22} \rightarrow$ Gly or $Tyr_{27} \rightarrow$ Ala are noticeably different from that of bPP itself.[9] Monitoring CD at 225 nm indicates that bPP, m-bPP, and t-bPP display cooperative denaturation behavior upon addition of GdmCl. Linear extrapolation to zero GdmCl concentration provides comparable $\Delta G_{fold/Gdm}$ values for these three molecules: −3.0 kcal/mol for bPP, −2.7 kcal/mol for m-bPP, and −2.5 kcal/mol for t-bPP.[9] Thus, the pair of mutations $Leu_3 \rightarrow$ Arg and $Asp_{10} \rightarrow$ Gly is slightly destabilizing (ca. 0.3 kcal/mol) for the pure polypeptide backbone, and there is only a small further destabilization (ca. 0.2 kcal/mol) upon replacing the Gly_9-Gly_{10} amide linkage with a thioester linkage.

10.1021/ja046891i CCC: $27.50

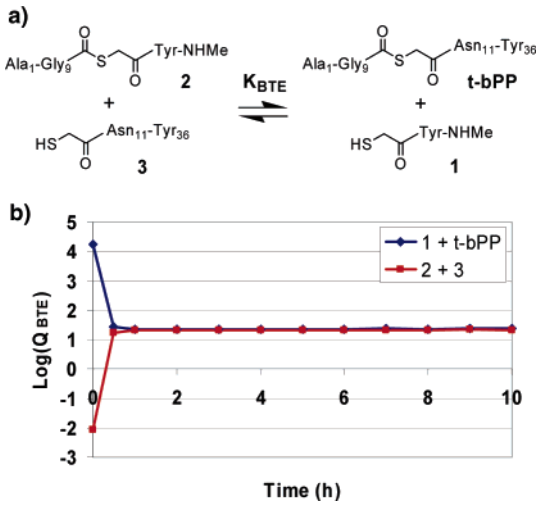

Figure 2. (a) Equation 1. (b) Graph illustrating BTE equilibration starting from the left side of eq 1 (red squares) or the right side (blue diamonds). Q_{BTE} is the value measured for [t-bPP][**1**]/[**2**][**3**] at a given time. At equilibrium, $Q_{BTE} = K_{BTE}$.

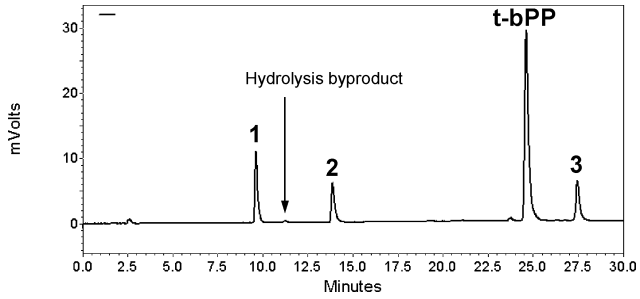

Figure 3. HPLC chromatogram of the species shown in Figure 2a at equilibrium (reaction time = 2.5 h). Conditions: UV detection at 275 nm, C_{18} column, eluent = CH_3CN/H_2O gradient (0.1% CF_3CO_2H).

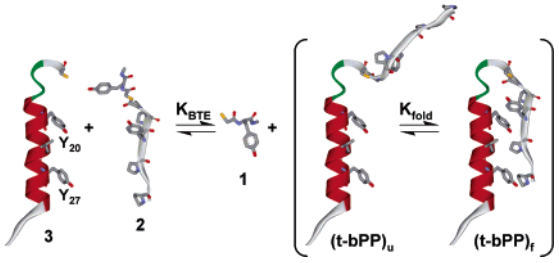

Figure 4. Model for thermodynamic analysis of BTE. Only selected t-bPP side chains are shown, including Y_{20} and Y_{27} on the α-helical segment.

To assess the stability of the folded state of t-bPP via BTE, we combined 0.1 mM t-bPP with 0.1 mM thioglycoyltyrosine *N*-methyl amide (**1**) in aqueous buffer (50 mM phosphate, pH 7). This solution was degassed and contained 2 mM tris-carboxyethylphosphine to suppress disulfide formation. HPLC showed that thiol/thioester exchange reached equilibrium within 1 h, as verified by monitoring the reaction started with **2** + **3** rather than **1** + t-bPP (Figure 2). Other than a trace of thioester hydrolysis, only thiol/thioester exchange occurred over several hours (Figure 3).

We used the model in Figure 4 to estimate folding thermodynamics for t-bPP on the basis of the HPLC-derived value for K_{BTE}. Formation of higher order structure in t-bPP is associated with equilibrium constant K_{fold} in this model. This folding process, by definition, involves tertiary contacts between the N-terminal polyproline II segment and the C-terminal α-helical segment; the native structure, $(t\text{-bPP})_f$, also includes quaternary contacts that lead to dimer formation, although this aspect of the structure is not indicated in the cartoon. The cartoon is not meant to imply that **2** and **3** and the corresponding peptide segments in $(t\text{-bPP})_u$ retain full secondary structure; CD data indicate that **3** is only partially α-helical,[9] and we have no evidence regarding secondary structure in **2**. Since we define $(t\text{-bPP})_u$ as lacking tertiary contacts between the N- and C-terminal segments, we can conclude that $[(t\text{-bPP})_u][\mathbf{1}] = [\mathbf{2}][\mathbf{3}]$ if we make two reasonable assumptions: (1) the C-terminal Tyr residue of **2** has no tertiary interactions with the N-terminal segment,[9] and (2) the thioester bonds in t-bPP and **2** are energetically equivalent. In this case, we can further conclude that $K_{BTE} = K_{fold} + 1$;[9] $\Delta G_{fold/BTE}$ is defined as $-RT \ln K_{fold}$.

Table 1. Comparison of ΔG_{Fold} Obtained by BTE and GdmCl Denaturation for bPP, t-bPP, and Various Mutants[a]

peptide	$\Delta G_{fold/BTE}$	$\Delta G_{fold/Gdm}$
bPP		−3.0
bPP (L3R, D10G)		−2.7
t-bPP	−1.4	−2.5
t-bPP (Y20A)	+0.8	
t-bPP (Y27A)	+0.1	
t-bPP (Y20,27A)	+1.1	
t-bPP (A22G)	−0.7	−0.8

[a] Values are reported in kcal/mol. We estimate the error in the BTE measurements to be ca. ±0.1 kcal/mol.

The measured K_{BTE} for t-bPP implies that $\Delta G_{fold/BTE} = -1.4$ kcal/mol, which is significantly smaller than $\Delta G_{fold/Gdm}$ for t-bPP (−2.5 kcal/mol; vide supra). We attribute this difference to the fact that GdmCl-induced unfolding leads to disruption of secondary, tertiary, and quaternary structure, while the BTE process (Figure 4) allows at least partial retention of secondary structure (a test of this hypothesis is described below). In addition, error may be introduced into $\Delta G_{fold/Gdm}$ by linear extrapolation of GdmCl-induced unfolding data to [GdmCl] = 0. The $\Delta G_{fold/BTE}$ value provides a useful perspective with regard to the forces that determine higher order structure in t-bPP because this number represents an upper limit (in terms of absolute value) on the free energy benefit of tertiary/quaternary structure formation under native conditions.

We examined the effects of several mutations to t-bPP on structural stability. The NMR-derived three-dimensional structure reported for bPP suggests that the side chains of Tyr_{20} and Tyr_{27} play a critical role in tertiary packing of the C-terminal α-helix against the N-terminal PPII helix,[5a] and removal of these side chains would therefore be predicted to destabilize the native state. CD data suggest that the $Tyr_{20} \rightarrow$ Ala, $Tyr_{27} \rightarrow$ Ala, and $Tyr_{20} \rightarrow$ Ala/$Tyr_{27} \rightarrow$ Ala mutants of t-bPP are not fully folded in the absence of GdmCl.[9] In addition, the CD signal at 225 nm for the double mutant is very similar to that for the corresponding α-helical segment alone when [GdmCl] = 0, and the CD_{225} values become even more similar as GdmCl is added.[9] Thus, it is not possible to evaluate accurately the formation of tertiary/quaternary structure in these mutants via chemical denaturation. In contrast, the BTE approach allows such studies (Table 1). $\Delta G_{fold/BTE} = +0.1$ kcal/mol for the $Tyr_{27} \rightarrow$ Ala mutant of t-bPP, implying that the unfolded and folded states have nearly identical free energy when the Tyr_{27} side chain is no longer available to participate in tertiary and quaternary contacts. For the $Tyr_{20} \rightarrow$ Ala mutant, $\Delta G_{fold/BTE} = +0.8$ kcal/mol, which suggests that the side chain of Tyr_{20} plays an even more important structure-stabilizing role than does the side chain of Tyr_{27}.

We attribute the deleterious effects of Tyr side chain truncation on t-bPP stability to loss of favorable tertiary and possible quaternary contacts in the folded state because these mutations should enhance secondary structural stability (Ala has a higher α-helical propensity than Tyr[10]). To probe the importance of

secondary structural stability on overall t-bPP folding, we mutated Ala_{22} to Gly. In the native state of bPP, Ala_{22} lies on the side of the α-helix that does not make contact with the N-terminal PPII helix,[5a] and Ala_{22} is not expected to be involved in quaternary contacts.[5b] Gly has a much lower α-helical propensity than Ala,[10] and the $Ala_{22} \rightarrow$ Gly mutation is therefore expected to affect overall conformational stability exclusively via secondary structure. CD data suggest that the $Ala_{22} \rightarrow$ Gly mutant of t-bPP is not quite fully folded at room temperature in the absence of GdmCl; we estimated $\Delta G_{fold/Gdm}$ for this mutant by using the fully folded CD_{225} value from t-bPP itself. This approach yielded $\Delta G_{fold/Gdm} =$ ca. -0.7 kcal/mol. BTE analysis of the $Ala_{22} \rightarrow$ Gly mutant gave $\Delta G_{fold/BTE} = -0.8$ kcal/mol. We attribute the similarity of these values to the fact that the C-terminal α-helix is not populated in either the Gdm-denatured state or the unfolded state of the full-length thioester that is accessed under native conditions (analogous to $(t\text{-}bPP)_u$ in Figure 4). This result supports our hypothesis that the difference between $\Delta G_{fold/BTE}$ and $\Delta G_{fold/Gdm}$ for t-bPP itself reflects the inclusion of C-terminal α-helix stability in the latter value but not (or at least not entirely) in the former.

We have shown that backbone thioester exchange provides a new way to evaluate the thermodynamics of higher order structure formation in polypeptides under native conditions. Judicious choice of the site for backbone amide-to-thioester modification in bPP leads to minimal perturbation of conformational stability but allows the use of chemical equilibration to provide insight on folding preferences. (It remains to be seen how generally such thioester replacements will be tolerated.) The resulting data are complementary to those obtained via classical chemical denaturation techniques,[1] and BTE can be used for molecules that are not fully folded in the absence of denaturant.[11] $\Delta G_{fold/BTE}$ can be determined with a single HPLC measurement using small amounts of material; therefore, this technique may be efficient for mapping out sequence-stability relationships in small polypeptides, for which the necessary peptide fragments can be synthesized and analyzed rapidly in parallel fashion.

Acknowledgment. This research was supported by the NIH (GM61238). M.W. was supported in part by a Fellowship from the Organic Division of the American Chemical Society, sponsored by Eli Lilly and Company. We thank Prof. Victor Muñoz, Prof. Tom Record, Jonathan Cannon, Michael Gelman, and Prof. Joel Schneider for helpful comments.

Supporting Information Available: CD data, chemical denaturation data, HPLC chromatograms, and algebraic derivations (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) For leading references, see: (a) Pace, C. N.; Shaw, K. L. *Proteins: Struct. Funct. Genet.* Suppl. 4, **2000**, 1. (b) Huyghues-Despointes, B. M. P.; Scholtz, J. M.; Pace, C. N. *Nat. Struct. Biol.* **1999**, *6*, 910.

(2) Dill, K. A.; Shortle, D. *Annu. Rev. Biochem.* **1991**, *60*, 795 and references therein.

(3) Thioester exchange reactions involving Cys side chains are a key element of the native chemical ligation (NCL) method of protein synthesis; for leading references, see: Dawson, P. E.; Kent, S. B. H. *Annu. Rev. Biochem.* **2000**, *69*, 923. Blaschke, U. K.; Silberstein, J.; Muir, T. W. *Methods Enzymol.* **2000**, *328*, 478−496. Hofmann, R. M.; Muir T. W. *Curr. Opin. Biotech.* **2002**, *13*, 297−303. For a conformationally assisted case of native chemical ligation, see: Beligere, G. S.; Dawson, P. E. *J. Am. Chem. Soc.* **1999**, *121*, 6332. In contrast to NCL, which is a preparative method, BTE is primarily a tool for gathering information on conformational stability.

(4) BTE can be viewed as a manifestation of "dynamic covalent chemistry": Rowan, S. J.; Cantrill, S. J.; Cousins, G. R. L.; Sanders, J. K. M.; Stoddart, J. F. *Angew. Chem., Int. Ed.* **2002**, *41*, 898. Thiol/disulfide exchange has been used in the context of protein structure; however, this method requires careful control of redox conditions, and disulfide linkages cannot mimic the protein backbone. For leading references, see: O'Shea, E. K.; Rutkowski, R.; Kim, P. S. *Science* **1989**, *243*, 538. Bilgicer, B.; Xing, X.; Kumar, K. *J. Am. Chem. Soc.* **2001**, *123*, 11815. Erlanson, D. A.; McDowell, R. S.; He, M. M.; Randal, M.; Simmons, R. L.; Kung, J.; Waight, A.; Hansen, S. K. *J. Am. Chem. Soc.* **2003**, *125*, 5602. Krishnan-Ghosh, Y.; Balasubramaniam, S. *Angew. Chem., Int. Ed.* **2003**, *42*, 2171.

(5) (a) Solution structure of bPP: Li, X.; Sutcliffe, M. J.; Schwartz, T. W.; Dobson, C. M. *Biochemistry* **1992**, *31*, 1245. Lerch, M.; Gafner, V.; Bader, R.; Christen, B.; Folkers, G.; Zerbe, O. *J. Mol. Biol.* **2002**, *322*, 1117. Roelfes, G.; Hilvert, D. *Angew. Chem., Int. Ed.* **2003**, *42*, 2275. (b) Crystal structure of a natural analogue, aPP: Blundell, T. L.; Pitts, J. E.; Tickle, I. J.; Wood, S. P.; Wu, C.-W *Proc. Natl. Acad. Sci. U.S.A.* **1981**, *78*, 4175. The dimerization mode in the crystal is expected to represent the dimerization mode of aPP and bPP in solution. (c) Clever use of this folding motif for design of functional miniproteins: Zondlo, N. J.; Schepartz, A. *J. Am. Chem. Soc.* **1999**, *121*, 6938. Chin, J. W.; Schepartz, A. *J. Am. Chem. Soc.* **2001**, *123*, 2929.

(6) Schnölzer, M.; Kent, S. B. H. *Science* **1992**, *256*, 221.

(7) Zacharias, D. E.; Murray-Rust, P.; Preston, R. M.; Glusker, J. P. *Arch. Biochem. Biophys.* **1983**, *222*, 22. Pawar, D. M.; Khalil, A. A.; Hooks, D. R.; Collins, K.; Elliott, T.; Stafford, J.; Smith, L.; Noe, E. A. *J. Am. Chem. Soc.* **1998**, *120*, 2108.

(8) We used BTE to prepare t-bPP. Fragment Ala_1-Gly_9 was synthesized on PAM-PS resin via standard Fmoc protocols and then cleaved from the resin using ethanethiol in the presence of $AlMe_3$, affording a C-terminal ethyl thioester (see: Sewing, A.; Hilvert, D. *Angew. Chem., Int. Ed.* **2001**, *40*, 3395). Before HPLC purification, the ethane thiol fragment was exchanged for methyl thioglycolate (large excess). The purified thioester was combined with **3** (prepared with standard Fmoc SPPS, the N-terminal residue being $Trt\text{-}SCH_2CO_2H$) in aqueous (pH 7) solution, which allowed thioester exchange. The desired peptide thioester (t-bPP) was collected by HPLC. Although the yield depends on K_{BTE} (ca. 12 for this reaction), the starting materials can be collected and recycled. Peptide thioester **2** was obtained in a similar fashion, with thiol **1** used in the final exchange step instead of **3**.

(9) Please see data in Supporting Information.

(10) Chou, P. Y.; Fasman, G. D. *Biochemistry* **1974**, *13*, 211. Muñoz, V.; Serrano, L. *Proteins: Struct. Funct. Genet.* **1994**, *20*, 301.

(11) Concept of partially folded proteins has recently become quite exciting; see: Garcia-Mira, M. M.; Sadqi, M.; Fischer, N.; Sanchez-Ruiz, J. M.; Muñoz, V. *Science* **2002**, *298*, 2191.

JA046891I