Analysis of Hydrogen Bonding Strengths in Proteins Using Unnatural Amino Acids

Jon S. Thorson, Eli Chapman, and Peter G. Schultz*

Howard Hughes Medical Institute
Department of Chemistry, University of California
Berkeley, California 94720

Received March 20, 1995

Although hydrogen bonds clearly play an important role in determining the secondary and tertiary structure of proteins,[1] the magnitude of their contribution to protein stability has been difficult to assess.[2] Values of hydrogen bond stabilization have been determined from mutational studies in which one member of a hydrogen bonded pair is deleted.[3] One concern in all such mutational studies is the possibility of introducing additional destabilizing interactions such as poor solvation of the remaining hydrogen bonded partner or decreased side chain packing interactions. In an effort to minimize the steric/electronic perturbations associated with these substitutions, we have independently mutated each member of both the Tyr^{27}-Glu^{10} and Tyr^{93}-Glu^{75} hydrogen bonded pairs in Staphylococcal nuclease (SNase) to an isosteric group with weak hydrogen bonding character, i.e., Glu → γ-nitro-L-glutamate and Tyr → *p*-amino-L-phenylalanine. The free energies of denaturation of these mutants are compared with those of mutants in which the tyrosine hydroxyl group is removed (Tyr → Phe) or replaced with a repulsive lone pair−lone pair interaction (Tyr → *p*-fluoro-L-phenylalanine).

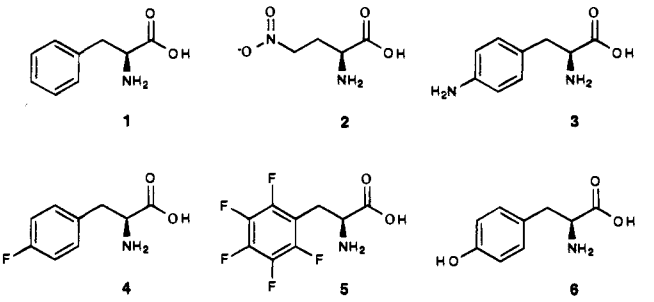
Tyr^{27} and Tyr^{93} are located in the five-stranded β-barrel of SNase and participate in an intramolecular hydrogen bond between the side chain hydroxyl group and the carboxylate side chain of Glu^{10} and Glu^{75}, respectively.[4] The side chains of Tyr^{27} and Glu^{10} are partially solvent exposed, while the side chains of Tyr^{93} and Glu^{75} are buried within the protein core. Since the nitro group has been shown to have poor hydrogen bonding ability compared to the isosteric carboxylate group,[5] γ-nitro-L-glutamate (**2**) was substituted for Glu^{10} and Glu^{75}. Similarly, the weakly acidic aromatic amino group of *p*-amino-L-phenylalanine (**3**) ($pK_a \sim 18$) should be a poor hydrogen bond donor when compared with the hydroxyl group of Tyr^{27} or Tyr^{93}.[6] Incorporation of amino acids **1**−**6** (Chart 1)[7] was accomplished by *in vitro* suppression of the corresponding amber mutants with a chemically aminoacylated suppressor tRNA derived from yeast $tRNA^{Phe}$.[8] At 10 mM added magnesium acetate, the aromatic amino acids **1**, **3**−**6** were incorporated with suppression efficiencies ranging from 25% (**5**) to 60% (**6**). The glutamate analog **2** was incorporated with 10% suppression efficiency, consistent with earlier reports that small, highly charged amino acids have low suppression efficiencies.[8] When $tRNA_{CUA}$ was omitted from the *in vitro* reaction or did not carry an amino acid, <1% full-length SNase was produced compared to expression of wild-type SNase by pKJSN1.[9a] Wild-type and mutant proteins were purified to homogeneity from 5.0 mL *in vitro* protein synthesis reactions.[9]

Table 1. SNase Tyr^{27}-Glu^{10} and Tyr^{93}-Glu^{75} Hydrogen Bond Variants and the Experimental Thermodynamic Stabilities and Kinetic Constants of the Resulting Enzymes

amino acid	C_m[a]	m_{GuHCl}[b]	$\Delta\Delta G_{H_2O}$[c]	K_m[d]	V_{max}[e]
Tyr^{27} → **1**	0.49	0.92	−1.26	7.6	6.8
Glu^{10} → **2**	0.50	0.98	−1.79	14.6	12.8
Tyr^{27} → **3**	0.59	0.80	−1.87	6.4	5.3
Tyr^{27} → **4**	0.39	0.90	−2.66	10.1	13.3
Tyr^{27} → **5**	0.31	0.87	−3.35	6.0	5.8
Tyr^{27} → **6**[f]	0.77	1.00		9.0	6.9
Tyr^{93} → **1**	0.63	0.89	−1.39	8.5	6.3
Glu^{75} → **2**	0.55	0.96	−1.96	10.6	12.4
Tyr^{93} → **3**	0.63	0.80	−1.78	9.4	5.3
Tyr^{93} → **4**	0.43	0.89	−2.56	14.1	16.3
Tyr^{93} → **5**	0.34	0.86	−3.13	8.0	7.3
Tyr^{93} → **6**[f]	0.77	1.00		9.0	6.9

[a] Midpoint concentration of GuHCl in molarity. [b] Units are relative to the wild-type *in vitro* suppressed value of 4.97 kcal mol^{-1} M^{-1} which has been normalized to 1.00. [c] $\Delta\Delta G_{H_2O}$ is expressed in units of kcal mol^{-1}. Negative values denote decreased stability with respect to WT suppressed **5**. [d] K_m is expressed in units of μg mL^{-1}. The assay mixture for kinetic determinations contained 40 mM sodium glycinate at pH 9.9, 10 mM Ca^{2+}, with varying amounts of single-stranded calf thymus DNA (7−50 μg mL^{-1}). [e] V_{max} is expressed in units of A_{260} μg^{-1} min^{-1}. [f] Reference 9b.

Chart 1



All mutants in the present study have specific activities, K_m, and V_{max} values very close to those of wild-type SNase, indicating an absence of significant structural abnormalities in any of the mutants (Table 1). The K_{app} values, an apparent equilibrium constant of denaturation, were determined with purified protein by monitoring the intrinsic fluorescence of the single tryptophan residue at position 140 as a function of guanidine hydrochloride concentration in 25 mM sodium phosphate, 100 mM NaCl, pH 7.0 aqueous buffer at 20 °C (Figure 1).[10] The values of *m* (Table 1), the rate of change of the free energy of denaturation with respect to the denaturant

* To whom correspondence should be addressed.

(1) (a) Pauling, L.; Corey, R. B. *Proc. Natl. Acad. Sci. U.S.A.* **1951**, *37*, 729. (b) Pauling, L.; Corey, R. B.; Branson, H. R. *Proc. Natl. Acad. Sci. U.S.A.* **1951**, *37*, 205.

(2) (a) Jeffrey, G. A.; Saenger, W. *Hydrogen Bonding in Biological Molecules*; Springer-Verlag: New York, 1991. (b) Rose, G. D.; Wolfenden, R. *Annu. Rev. Biophys. Biomol. Struct.* **1993**, *22*, 381.

(3) (a) Dill, K. A. *Biochemistry* **1990**, *29*, 7133. (b) Chen, Y. W.; Fersht, A. R.; Henrick, K. *J. Mol. Biol.* **1993**, *234*, 1158. (c) Shirley, B. R.; Stanssens, P.; Hahn, U.; Pace, C. N. *Biochemistry* **1992**, *31*, 725. (d) Fersht, A. R. *Trends Biochem. Sci.* **1987**, *12*, 301. (e) Alber, T.; *et al. Nature (London)* **1987**, *330*, 41. (f) Fersht, A. R.; *et al. Nature (London)* **1985**, *314*, 235. (g) Meiring, E. M.; Serrano, L.; Fersht, A. R. *J. Mol. Biol.* **1992**, *225*, 585. (h) Dang, L. X.; Merz, K. M.; Kollman, P. A. *J. Am. Chem. Soc.* **1989**, *111*, 8505.

(4) Hynes, T. R.; Fox, R. O. *Proteins: Struct., Funct., Genet.* **1991**, *10*, 92.

(5) Kelly, T. R.; Kim, M. H. *J. Am. Chem. Soc.* **1994**, *116*, 7072−7080.

(6) West, R. C., Ed. *Handbook of Chemistry and Physics*; CRC Press: Cleveland, OH, 1976; p D176.

(7) Amino acids **1**−**4**, and **6** are commercially available. Amino acid **5** was synthesized by converting 2,3,4,5,6-pentafluorobenzyl alcohol to the corresponding bromide with HBr. The bromide was subsequently coupled with diethyl acetamidomalonate, followed by hydrolysis, to give the corresponding racemic amino acid HBr salt. Although many of the precursors to the unnatural amino acids used for this study were racemic, previous studies indicate that only the L-amino acid is incorporated by *in vitro* suppression (D. Mendel, unpublished material).

(8) (a) Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C.; Schultz, P. G. *Science* **1989**, *244*, 182. (b) Ellman, J. A.; Mendel, D.; Anthony-Cahill, S. J.; Noren, C. J.; Schultz, P. G. *Methods Enzymol.* **1991**, *202*, 301. (c) Noren, C. J.; *et al. Nucleic Acids Res.* **1990**, *18*, 83. (d) Robertson, S. A.; Ellman, J. A.; Schultz, P. G. *J. Am. Chem. Soc.* **1991**, *113*, 2722. (e) Cornish, V. W.; Mendel, D.; Schultz, P. G. *Angew. Chem., Int. Ed. Engl.*, in press.

(9) (a) Judice, J. K.; Gamble, T. R.; Murphy, E. C.; de Vos, A. M.; Schultz, P. G. *Science* **1993**, *261*, 5128. (b) Thorson, J. S.; Chapman, E.; Murphy, E. C.; Schultz, P. G.; Judice, K. J. *J. Am. Chem. Soc.* **1995**, *117*, 1157.

0002-7863/95/1517-9361$09.00/0

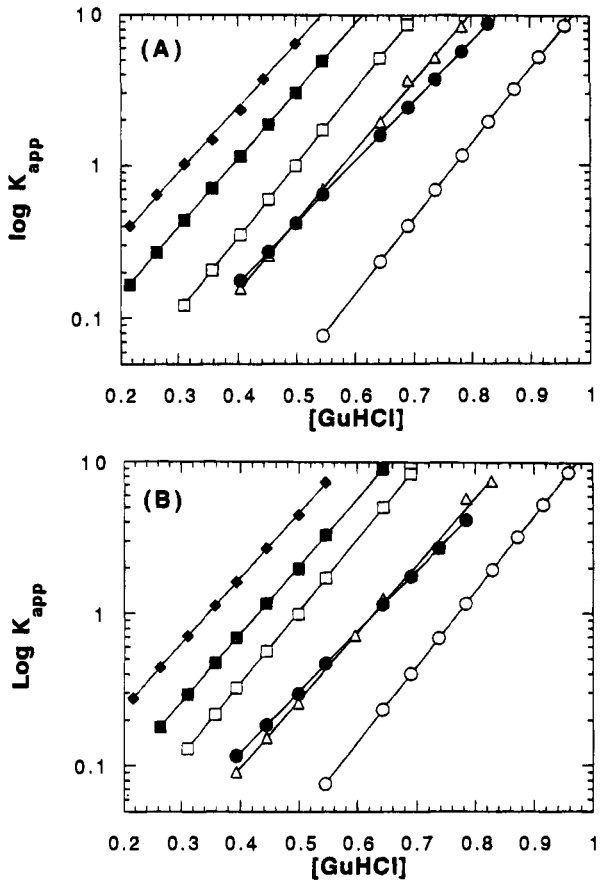


Figure 1. (A) Denaturation plots of log K_{app} versus guanidine hydrochloride concentration for the following SNase variants: (○) Tyr^{27} → L-tyrosine (**6**); (△) Tyr^{27} → L-phenylalanine (**1**); (●) Tyr^{27} → *p*-amino-L-phenylalanine (**3**); (■) Tyr^{27} → *p*-fluoro-L-phenylalanine (**4**); (◆) Tyr^{27} → 2,3,4,5,6-pentafluoro-L-phenylalanine (**5**); and (□) Glu^{10} → γ-nitro-L-glutamate (**2**). (B) Denaturation plots of log K_{app} versus guanidine hydrochloride concentration for following SNase variants: (○) Tyr^{93} → L-tyrosine (**6**); (△) Tyr^{93} → L-phenylalanine (**1**); (●) Tyr^{93} → *p*-amino-L-phenylalanine (**3**); (■) Tyr^{93} → *p*-fluoro-L-phenylalanine (**4**); (◆) Tyr^{93} → 2,3,4,5,6-pentafluoro-L-phenylalanine (**5**); and (□) Glu^{75} → γ-nitro-L-glutamate (**2**).

concentration,[11] are in relatively close agreement for wild-type and mutant proteins.

Substitution of Glu^{10} or Glu^{75} with the nitro analog **2** decreases the stability of SNase ($\Delta\Delta G_{H_2O}$) by 1.79 or 1.96 kcal mol^{-1}, respectively. The nitro group, while an isosteric substitution for the carboxylate group, is a very poor hydrogen bond acceptor in such solvents as DMSO and $CHCl_3$.[5] Consequently, the value of $\Delta\Delta G_{H_2O}$ is likely a very accurate measure of the strength of the Glu-Tyr hydrogen bond, with minimum perturbation from packing, entropy, or medium effects. Substitution of either tyrosine hydroxyl with the weak hydrogen bond donating aromatic amino group (Tyr^{27} → **3**, $\Delta\Delta G_{H_2O}$ = −1.87 kcal mol^{-1}; Tyr^{93} → **3**, $\Delta\Delta G_{H_2O}$ = −1.78 kcal mol^{-1}) leads to very similar decreases in protein stability. Since the lone pair of each tyrosine hydroxyl group does not appear to be involved in any interaction in the protein, this decreased stability again is largely associated with the substitution of a good hydrogen bonding group with a weak one. For comparison, removal of the hydroxyl group from either Tyr^{27} or Tyr^{93} leads to a decrease in protein stability of 1.26 or 1.39 kcal mol^{-1}, respectively. Perturbing the hydrogen donor in the opposite direction, i.e., substituting the hydrogen bond with a repulsive lone pair−lone pair interaction, Tyr^{27} → **5** or Tyr^{93} → **5**, destabilizes the protein by 2.66 or 2.56 kcal mol^{-1}, respectively.[12]

The stabilities of the Glu^{10} → **2** and Tyr^{27} → **3** mutants were also determined at pH 5.0 and 9.5, respectively and compared to those of wild-type protein.[13] The $\Delta\Delta G_{H_2O}$ value for **2** ($\Delta\Delta G_{H_2O}$ = −1.75 kcal mol^{-1}) at pH 5.0 is similar to that determined at pH 7.0, suggesting that the γ-nitro group is present in neutral form at pH 7.0, as expected. A slight decrease in $\Delta\Delta G_{H_2O}$ is observed for **3** ($\Delta\Delta G_{H_2O}$ = −1.75 kcal mol^{-1}) at pH 9.5, indicating that a small amount of the protonated species may be present at pH 7.0. Comparison of the 2,3,4,5,6-pentafluoro-L-phenylalanine (Tyr^{27} → **5**, $\Delta\Delta G_{H_2O}$ = −3.35 kcal mol^{-1}; Tyr^{93} → **5**, $\Delta\Delta G_{H_2O}$ = −3.13 kcal mol^{-1}) mutants and the corresponding *p*-fluorophenylalanine mutants illustrates that aromatic substitution other than at the donor/acceptor interface contributes only an average of 0.16−0.18 kcal mol^{-1} per fluorine, indicating only minimal perturbation due to π effects.

The $\Delta\Delta G_{H_2O}$ values for the Tyr → *p*-amino-L-phenylalanine and Glu → γ-nitro-L-glutamate mutants are remarkably similar, despite the differences in local environment and the fact that in one case the hydrogen bond donor and in the other case the acceptor is being mutated. These results support the notion that intramolecular hydrogen bonds in folded proteins make a net favorable contribution of 1−2 kcal mol^{-1} to protein stability relative to the corresponding hydrogen bonds to solvent in the denatured state. Moreover, the similarity of the stabilities of the γ-nitro-L-glutamate and *p*-amino-L-phenylalanine mutants to those of the Tyr → Phe mutants, in which the tyrosine hydroxyl group is deleted, suggest that the *net* perturbations associated with deletion of the tyrosine hydroxyl groups, other than the loss of the hydrogen bond, do not significantly affect protein stability. Thus, in this case, conventional mutagenesis provides a reasonable measure of intramolecular hydrogen bond strengths.

Acknowledgment. This work was supported in part by a grant from the National Institutes of Health (GM49220). J.S.T. is a Merck Postdoctoral Fellow of the Helen Hay Whitney Foundation.

JA9509083

(10) Denaturation studies were performed in triplicate in 25 mM sodium phosphate, 100 mM sodium chloride, pH 7.0 aqueous buffer according to the procedure of Shortle and Meeker with a Hitachi F-4500 fluorescence spectrophotometer.[14] Protein concentrations used in the denaturation measurements were 3 μg mL^{-1}. The apparent equilibrium constant, K_{app}, for reversible denaturation, assuming a two-state model, was determined for a series of GuHCl concentrations incremented in steps of 0.047 M by using the equation $K_{app} = (I_n - I)/(I - I_d)$, where I is the measured intrinsic fluorescence, I_n is the extrapolated value of fluorescence for the native state, and I_d is the extrapolated value for the denatured state. To obtain ΔG_{H_2O} and m_{GuHCl}, a straight line was fit to log K_{app} versus [GuHCl] by using a linear least-squares method. From this plot, the value of K_{app} at [GuHCl] = 0, was extrapolated, and ΔG_{H_2O} was determined from the equation $\Delta G = -RT \ln K_{app}$.

(11) Dill, K. A.; Shortle, D. *Annu. Rev. Biochem.* **1991**, 795.

(12) These π effects are consistent with those previously observed.[9b]

(13) Denaturation studies for **2** were performed in triplicate in 25 mM sodium citrate, 25 mM glycine, 100 mM sodium chloride, pH 5.0 aqueous buffer at 20 °C, while GuHCl titrations for **3** were performed in triplicate in 25 mM sodium citrate, 25 mM glycine, 100 mM sodium chloride, pH 9.5 aqueous buffer at 20 °C, as previously described.[10]

(14) Shortle, D.; Meeker, A. K. *Proteins: Struct., Funct., Genet.* **1986**, *1*, 81.