Linear Free Energy Analysis of Hydrogen Bonding in Proteins

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

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

J. Kevin Judice

Department of Protein Engineering, Genentech, Inc.
460 Point Bruno Boulevard
South San Francisco, California 94080

Received September 9, 1994

The analysis of substituent effects has proven an extremely valuable tool for the physical organic chemist in the study of the physical properties and chemical reactivities of small molecules. However, the use of substituent effects to study protein structure and function has been limited both by the constraints of the protein biosynthetic machinery and by the difficulties associated with the synthesis or semisynthesis of large proteins (>10 kDa). We now report the free energies of denaturation for a series of Staphylococcal nuclease (SNase) mutants in which the strength of a tyrosine−glutamate hydrogen bond has been modulated by varying the acidity of the tyrosine hydroxyl group using fluorine ring substitutions. These mutants were designed to determine the degree to which hydrogen bonding preferentially stabilizes the folded state of proteins while minimizing the perturbations resulting from the deletion of one member of a hydrogen-bonded pair.[1,2] A simple linear free energy correlation between $\Delta\Delta G_{H_2O}$ of the mutant proteins and the pK_a of the tyrosine hydroxyl group provides strong evidence that intramolecular hydrogen bonds in folded proteins can be more stable than the corresponding hydrogen bonds to water in the unfolded state.

Tyr^{27} is located in the five-stranded β-barrel of SNase and participates in an intramolecular hydrogen bond between the side chain hydroxyl group and the carboxylate side chain of Glu^{10}.[3] A series of fluorinated tyrosine analogs, 2-fluorotyrosine (**1**), 3-fluorotyrosine (**2**), and 2,3,5,6-tetrafluorotyrosine (**3**) (ΔpK_a = 4.5 pK_a units), were substituted to Tyr^{27}.[4] Because of fluorine's relatively small size (van der Waals radius of 1.35 Å relative to 1.2 Å for hydrogen),[5] the steric perturbation relative to tyrosine should be minimal. Incorporation of the fluorinated tyrosine analogs **1**−**3**, as well as 2,3,5,6-tetrafluoro-L-phenylalanine (**4**), L-phenylalanine (**5**), and L-tyrosine (**6**), at position 27 of SNase was accomplished by *in vitro* suppression of a Tyr^{27} → TAG stop mutation with a chemically aminoacylated suppressor tRNA derived from yeast $tRNA^{Phe}$.[6] At 10 mM added magnesium acetate, amino acids **1**−**6** were incorporated with suppression efficiencies ranging from 15% (**2** and **3**) to 60% (**6**). In contrast, when $tRNA_{CUA}$ was omitted from the *in vitro* reaction or did not carry an amino acid, less than 1% full-length SNase was produced compared to expression of wild-type SNase by pKJSN1.[7] Mutant SNases were purified to homogeneity from 5.0 mL in vitro protein synthesis reactions.[7]

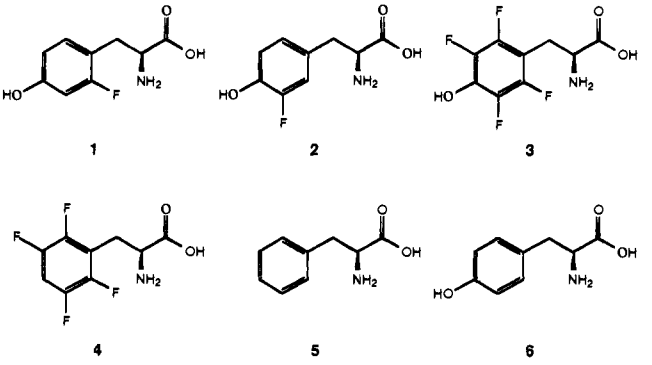
All mutants in the present study have specific activities, K_m values, and V_{max} values within experimental error of those of wild-type SNase, indicating an absence of significant configurational abnormalities in any of the mutants studied (Table 1).[7] The K_{app}'s, an apparent equilibrium constant of denaturation, were determined with purified protein by monitoring the intrinsic fluorescence of the single tryptophan residue at position 140 as a function of denaturant concentration.[8] The values of *m* (Table 1), the rate of change of the free energy of denaturation with

Table 1. Amino Acids Incorporated at Position 27 of SNase and the Experimental Thermodynamic Stabilities and Kinetic Constants of the Resulting Enzymes

amino acid	pK_a[a]	C_m[b]	m_{GuHCl}[c]	$\Delta\Delta G_{H_2O}$[d]	K_m[e]	V_{max}[f]
1	9.3	0.82	0.94	0.34 (0.5)	6.0	7.1
2	8.8	0.93	1.0	0.99 (1.15)	7.0	6.7
3	5.3	1.03	0.99	1.59 (2.23)	5.0	6.8
4		0.50	0.88	−2.1	6.0	5.8
5		0.58	0.92	−1.49	7.6	6.8
6	10.0	0.77	1.0	0.0	9.0	6.9

[a] Due to the lack of a pK_a determination for 2-fluorotyrosine, the large discrepancy (pK_a's ranging from 8.3 to 9.21) among pK_a determinations for 3-fluorotyrosine,[12] and the very close relationship between the hydroxyl pK_a of tetrafluorotyrosine (pK_a = 5.4)[13] and tetrafluorophenol (pK_a = 5.3)[14] and between that of tyrosine (pK_a = 10.07)[15] and phenol (pK_a = 9.95), the p*Ka* values adapted for this discussion are based upon the determined values of fluoro-substituted phenols.[16] [b] Midpoint concentration of GuHCl in molarity. [c] 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. [d] $\Delta\Delta G_{H_2O}$ is expressed in units of kcal mol^{-1}. The values in parentheses correspond to corrections applied from the comparison of 2,3,5,6-tetrafluorophenylalanine and phenylalanine substitutions. Positive values denote increased stability with respect to wild-type suppressed **6**. [e] 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}). [f] V_{max} is expressed in units of A_{260} μg^{-1} min^{-1}.

Scheme 1. Amino Acid Substitutions for Tyr^{27}



* To whom corrspondence should be addressed.

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

(2) (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.

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

(4) Amino acid substitutions for Tyr^{27}. Amino acids **1**−**3**, **5**, and **6** are commercially available. Amino acid **4** was synthesized via lithium aluminum hydride reduction of 2,3,5,6-tetrafluorobenzoic acid followed by conversion to the corresponding bromide with HBr. The bromide was subsequently coupled with diethyl acetamidomalonate followed by hydrolysis to give the corresponding amino acid HBr salt. Although the precursors to unnatural amino acids **1**−**4** were DL, previous studies indicate that only the L-amino acid can be utilized for *in vitro* suppression (D. Mendel, unpublished material).

(5) West, R. C., Ed. In *Handbook of Chemistry and Physics*; CRC Press: Cleveland, OH, 1976; p D178.

(6) (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.

(7) Judice, J. K.; Gamble, T. R.; Murphy, E. C.; de Vos, A. M.; Schultz, P. G. *Science* **1993**, *261*, 5128.

0002-7863/95/1517-1157$09.00/0

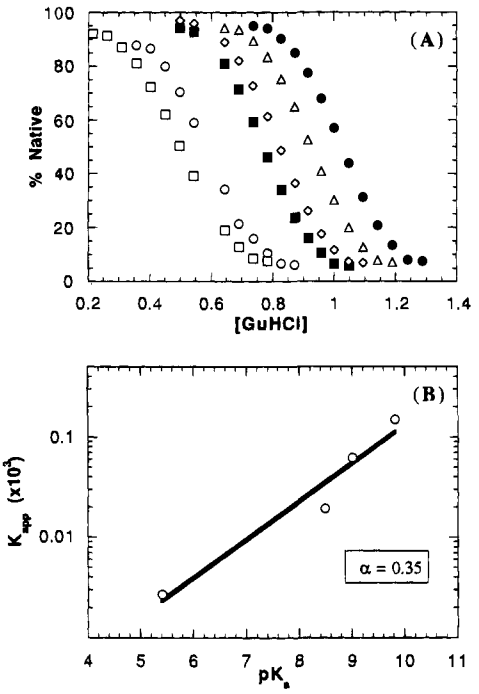


Figure 1. (A) Free energy plot for the stabilization of SNase by increasing the acidity of the proton donor in the Tyr^{27}—Glu^{10} hydrogen bond pair. The slope of the line (α) through the data is 0.35. (B) Denaturation curves for SNase variants containing the following amino acids at position 27: (□) 2,3,5,6-tetrafluoro-L-phenylalanine (**4**); (○) L-phenylalanine (**5**); (■) L-tyrosine (**6**); (◇) 2-fluoro-L-tyrosine (**1**); (△) 3-fluoro-L-tyrosine (**2**); and (●) 2,3,5,6-tetrafluoro-L-tyrosine (**3**).

respect to the denaturant concentration (which is approximately proportional to the difference in solvent accessible surface area between the denatured and native states),[9] are in close agreement for the fluorotyrosine mutants, indicating that **1**−**3** are isosteric with Tyr. The variance of m from wild-type protein found for the phenylalanine derivatives **4** and **5** suggests that volume or hydrophobic corrections may be required for correct interpretation of these substitutions relative to wild-type protein.

The change in stability of the mutant proteins, $\Delta\Delta G_{H_2O}$, increases with fluorine substitution, the Tyr^{27} → tetrafluorotyrosine (**3**) mutant being some 1.59 kcal mol^{-1} more stable than wild-type SNase. In order to determine whether this effect is associated solely with increased acidity of the tyrosine hydroxyl group, and concomitant increased hydrogen bond strength,[10] or whether other steric/electronic perturbations contribute, the relative stabilities of the Tyr^{27} → Phe and Tyr^{27} → tetrafluorophenylalanine (**4**) mutants were compared. The Tyr^{27} → **4** mutant is destabilized by 0.64 kcal mol^{-1} with respect to **5**, affording a correction factor of 0.16 kcal mol^{-1} per fluorine substituent, which translates to an overall stabilization energy of 2.23 kcal mol^{-1} for the tetrafluorotyrosine mutant. A linear free energy analysis of the log K_{app} versus pK_a provides a good linear relationship, indicating that K_{app} correlates well with the acidity of the proton donor. The value of α from a linear free energy correlation, log $K_{app} = \alpha(pK_a) + C$ where α reflects the degree of proton transfer between Tyr^{27} and Glu^{10}, is 0.35. This value indicates that the shared proton resides in a relatively symmetrical orientation with a slight bias for the proton donor of the Tyr^{27}—Glu^{10} donor—acceptor pair.

A variety of model studies have failed to give an unequivocal answer to the long-standing question of whether the formation of intramolecular hydrogen bonds in folded proteins makes a net favorable contribution to protein stability versus the alternative hydrogen bond to solvent.[1b] Some of these studies have even suggested that hydrogen bonds with water would be favored, strengthening the idea that hydrophobic interactions are the dominant force in protein folding.[2a] Because values of hydrogen bond stabilization determined from mutational studies of proteins only measure the overall change in protein stability, it is difficult to dissect out the effects of a given mutation in terms of the hydrogen bonding, van der Waals interactions, conformational entropy, and solvation effects associated with a given steric or electronic change.[2] Interpretation of mutagenesis data can be further complicated by the introduction of additional destabilizing interactions resulting from unpaired hydrogen bond donors or acceptors. Here we show that increasing the acidity of the tyrosine hydroxyl group by fluorine ring substitution contributes an average 0.56 kcal mol^{-1} per fluorine, translating to a total stabilization energy of 2.23 kcal mol^{-1} for the Tyr^{27} → **3** mutant. Although the existence of any localized structure in proximity to Tyr^{27} in the unfolded state is unclear, the fact that an increase in the acidity of the Tyr^{27} hydroxyl group simultaneously increases the $\Delta\Delta G_{H_2O}$ provides strong evidence that hydrogen-bonding networks may play an important role in stabilizing the folded state of proteins. Additional studies at other sites will be required to determine the generality of this result.

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

JA942990S

(8) The denaturation studies were performed in triplicate in 25 mM sodium phosphate, 100 mM sodium chloride, pH 7.0, according to the procedure of Shortle and Meeker with a Perkin-Elmer LS-5B luminescence spectrometer.[11] Protein concentrations used in the denaturation measurements were 6 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 fitted to log K_{app} versus [GuHCl] by using a linear least-squares method. From this plot was extrapolated the value of K_{app} at [GuHCl] = 0, and ΔG_{H_2O} was determined from the equation $\Delta G = -RT \ln K_{app}$.

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

(10) (a) Gordy, W.; Stanford, S. C. *J. Chem. Phys.* **1961**, *9*, 204. (b) Arnett, E. M. *Prog. Phys. Org. Chem.* **1963**, *1*, 223.

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

(12) (a) Carne, T. J.; Huber, R. E.; Davitt, P.; Edwards, L. A. *J. Chromatogr.* **1986**, *367*, 393. (b) Dettman, H. D.; Weimer, J. H.; Sykes, B. D. *Can. J. Biochem. Cell Biol.* **1985**, *63*, 1120.

(13) Filler, R.; Ayvanger, N. R.; Gatowski, W.; Karg, M. M. *J. Org. Chem.* **1969**, *34*, 534.

(14) Forbes, E. J.; Richardson, R. D.; Stacey, M.; Tatlow, J. C. *J. Chem. Soc.* **1959**, 2019.

(15) Barber, R. In *Organic Chemistry of Biological Compounds*; Prentice-Hall: Engelwood Cliffs, NJ, 1971; p 57.

(16) Fasman, G. D., Ed. In *Handbook of Biochem. and Mol. Biol.: Physical and Chemical Data*; CRC Press: Boca Raton, FL, 1976; Vol. I, p 314.