Pure & Appl. Chem., Vol. 66, No. 1, pp. 65–69, 1994.
Printed in Great Britain.

Case study of protein structure, stability, and function: NMR investigations of the proline residues in staphylococcal nuclease

John L. Markley, Andrew P. Hinck,[a] Stewart N. Loh, [b] Kenneth Prehoda, Dagmar Truckses, William F. Walkenhorst,[c] and Jinfeng Wang

Department of Biochemistry, College of Agricultural and Life Sciences, University of Wisconsin- Madison, Madison WI 53706, USA
Present addresses: [a]*Department of Chemistry, University of Wisconsin-Madison, Madison WI 53706, USA;* [b]*Department of Biochemistry, Stanford University School of Medicine, Beckman B400, Palo Alto, CA 94305, USA;*
[c]*Institute for Cancer Research, Fox Chase Cancer Center, 7701 Burholme Ave., Philadelphia PA 19111, USA.*

Abstract

We have used NMR spectroscopy to determine peptide bond configurations and to measure the rates and equilibria of interconversion at individual Xaa-Pro peptide bond linkages in staphylococcal nuclease and several variants produced by site-directed mutagenesis. In general, tertiary interactions, rather than short-range interactions, have been found to be critical for stabilizing the *cis* linkage at Lys^{116}- Pro^{117} which predominates in the wild-type enzyme. A correlation has been found between the position of the *cis/trans* equilibrium at the Lys^{116}-Pro^{117} peptide bond and thermal stability of the variant. Enthalpic interactions that stabilize the folded protein appear to be present when the peptide bond is *cis* but not when it is *trans*. Hydrogen exchange protection factors correlate with the mole fraction of the protein that is in the *cis* configuration. Nuclease variants in which the peptide bond is predominantly *cis* are more stable against denaturation (by heat, pressure, or guanidinium chloride) than those that are predominantly *trans*. Disulfide bridges have been engineered and introduced by mutagenesis that stabilize certain conformational states; one of these shows coupling between the oxidation state of the engineered cysteine pair and the *cis/trans* configuration about the Lys^{116}-Pro^{117} peptide bond. A rough correlation is seen between the catalytic activities of mutants and the *cis/trans* ratio; the effect is primarily on k_{cat} rather than on K_m.

INTRODUCTION

Staphylococcal nuclease (nuclease) has been used widely as a model for studies of structure-function relationships and protein folding (ref. 1). Nuclease was among the first enzymes whose structure was determined by x-ray crystallography (ref. 2) and whose physical properties were investigated by NMR (ref. 3). More highly refined x-ray structures of nuclease have been published (refs. 4-5) as have nearly complete NMR assignments from two natural variants of nuclease (refs. 6-12). Recent NMR studies have shown that native nuclease (Foggi form) exists in solution as a slowly interconverting equilibrium mixture of four conformers (refs. 13-16): *tc* (72%), *cc* (18%), *tt* (8%), and *ct* (2%), where the first letter (c, *cis;* t, *trans*) refers to the configurational state of the His^{46}-Pro^{47} peptide bond (ref. 13), and the second letter to that of the Lys^{116}-Pro^{117} peptide bond (ref. 16). Only one of these four conformational states (*tc*) has been observed in x-ray structures of wild-type proteins.

PROTEIN FOLDING, STABILITY, AND CONFORMATIONAL PURITY

Proteins are synthesized in cells as linear polymers that must fold properly in order to achieve their activity. Folded forms of proteins are only slightly more stable than the unfolded protein; typically, only a few kcal/mole separate the two ($\sim$ 5.5 kcal/mole in the case of wild-type staphylococcal nuclease). Staphylococcal nuclease refolds spontaneously *in vitro*, thus it does not belong to the

classes of proteins that have been found to require assistance in achieving their active three-dimensional structure (e.g., from a chaperonin, disulfide isomerase, or prolyl peptide isomerase).

Most of the steps in protein folding are very fast--so rapid, in fact, that they have proved difficult to follow. Thus many studies of protein folding have concentrated on the two kinds of steps that are relatively slow: disulfide bond formation or interchange and peptide bond isomerism at Xaa-Pro peptide linkages. Although x-ray structures of crystalline proteins generally have found that only one or the other prolyl peptide bond configuration is populated at each site, NMR studies of proteins in solution have shown that appreciable populations of alternative configurations can coexist in proteins as has been found in many proline-containing peptides (ref. 17). This is not totally unexpected, since free energy would be expended in order to perturb the *cis* = *trans* equilibrium in one direction or another, and the only source of this energy is the free energy of folding. If one assumes that $K_{tc} \sim 0.25$ for a prolyl peptide bond in a random-coil, then approximately 0.8 kcal/mol or 1.5 kcal/mol would be needed to make the sample appear to be, respectively, "all *trans*" or "all *cis*" by NMR analysis. Since this energy would be needed at each proline, this effect can represent an appreciable fraction of the total free energy of folding.

Protein production and mutagenesis. Staphylococcal nuclease was cloned first by Shortle (ref. 18) who developed an overexpression system for the protein (ref. 19). We have an efficient overexpression procedure, which makes use of the T7 expression vector pET3a (ref. 20), that enables us to prepare large quantities of nuclease and to label it with stable isotopes (ref. 15). The mutants discussed here were prepared in our laboratory by site-directed mutagenesis (refs. 21-22). Oligonucleotides carrying the appropriate base substitutions, typically 21 base pairs in length, were used to generate appropriate mutations.

NMR SPECTROSCOPIC STRATEGIES

Multidimensional, multinuclear NMR methods were used to assign NMR spectra of the V8 nuclease (H124L) (refs. 8-12). The major peptide bond configuration at the 116-117 linkage was determined by isotope-edited NOE spectroscopy with selectively labeled nuclease molecules prepared biosynthetically. In order to detect the presence of a *trans* peptide bond, the protein was labeled with [98% $^{13}C^{\alpha}$]lysine and [98% $^{13}C^{\beta,\delta}$]proline. In order to detect the presence of a *cis* peptide bond, the protein was labeled with [98% $^{13}C^{\alpha}$]lysine and [98% $^{13}C^{\alpha}$]proline. Since the protein

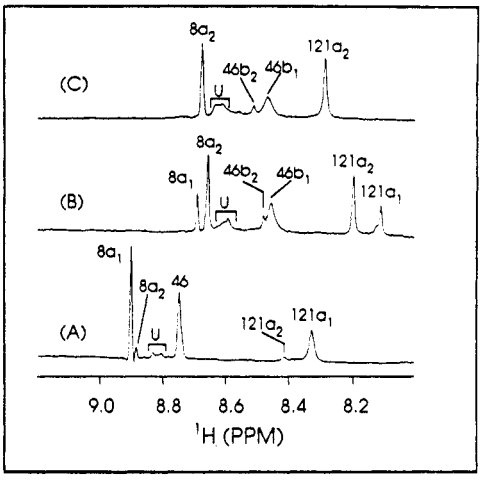

Fig. 1. Histidine $^1H^{\varepsilon 1}$ region of the 500 MHz ^1H NMR spectrum of three staphylococcal nuclease variants: (A) H124L; (B) G79S + H124L; and (C) D77A + H124L. Spectra were acquired at 38 C in ^2H_2O at pH* 5.5 (pH meter reading uncorrected for the deuterium isotope effect at the glass electrode). Histidine peaks from unfolded protein are indicated by the symbol U. Numbers refer to residue numbers. Peaks designated by a_1 and a_2 are assigned, respectively, to protein molecules with *cis* and *trans* 116-117 peptide bonds. Peaks designated by b_1 and b_2 are assigned, respectively, to protein molecules with *trans* and *cis* 46-47 peptide bonds. (From ref. 15.)

contains only a single Lys-Pro sequence at 116-117, this labeling scheme provides automatic NMR assignments as revealed by sequential isotope-edited NOE's and determines the predominant peptide bond configuration present in solution (ref. 23). *Cis/trans* ratios at the Lys^{116}-Pro^{117} peptide bond were determined by one- dimensional ^1H NMR spectroscopy at 500 MHz or 600 MHz (15) from chemical shifts of the $^1H^{\varepsilon 1}$ protons of His^8 and His^{121} (and His^{124} when present). Typical spectra are shown in Fig. 1.

Regions of the nuclease molecule that are affected by changes in configuration of the Lys116-Pro117 peptide bond. Peptide bond 116-117 in nuclease H124L is predominantly *cis* (90%). In nuclease G79S, however peptide bond 116-117 is predominantly *trans* (ref. 14). By comparing assigned chemical shifts of residues in these two proteins, it was possible to identify residues of potential importance in stabilizing the *cis* configuration over the *trans* (Fig. 2).

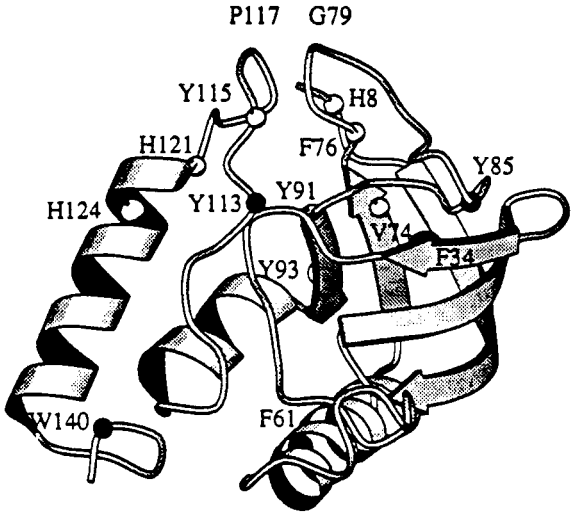

Fig. 2. Schematic representation of the main-chain fold of nuclease (ref. 5) drawn with the MOLSCRIPT program (ref. 24). Labels indicate the Type VI_a reverse turn loop containing Pro^{117} (P117) and the adjacent loop containing Gly^{79} (G79). Hollow spheres indicate the positions of α-carbons of aromatic residues whose chemical shifts in H124L and G79S differ by 0.1 ppm; filled spheres indicate the positions of α-carbons of aromatic residues whose chemical shifts in H124L and G79S differ by less than this value. Residues His^8, His^{121}, His^{124}, and Val^{74} are also indicated by hollow spheres since their chemical shifts depend on the isomerization state of peptide bond 116-117. (From ref. 15.)

Factors that affect the equilibrium between cis and trans peptide bond isomers at 116-117. Table 1 summarizes results with three variants of nuclease: wild-type Foggi nuclease (V8), a mutant of the V8 protein in which glycine-79 has been replaced with serine (G79S), and a mutant of the V8 protein in which aspartate-77 has been replaced by alanine. K_{tc} values were determined as shown in Fig. 1. The predominant configuration of the peptide bond was established by the isotope-edited NOE experiment described above.

Table 1. Peptide Bond Isomerism at Proline-117 in Staphylococcal Nuclease.

nuclease H124L variant	K_{tc} = [c]/[t]	residue 116	residue 117	inhibitor pdTp-Ca^{2+}	IE-NOE results
WT	10.5	K^{α}	P^{α}	-	+ (*cis*)
	10.5	K^{α}	$P^{\beta,\delta}$	-	- (*cis*)
G79S	0.4	K^{α}	P^{α}	-	- (*trans*)
	>12	K^{α}	P^{α}	+	+ (*cis*)
D77A	<0.02	K^{α}	$P^{\beta,\delta}$	-	+ (*trans*)
	6.1	K^{α}	$P^{\beta,\delta}$	+	- (*cis*)

Greek letters indicate ^{13}C labeled carbons. IE-NOE refers to isotope edited nuclear Overhauser effect measurements. (From refs. 15,23.)

Table 2. Correlation between Thermal Stability (T_m) and the Position of the Peptide Bond Isomerism at Proline-117 (K_{tc})

nuclease variant H124L +	T_m (°C)		K_{tc} = [c] / [t] (major species present)	
----	56.1		10.5 (c)	
G79S	46.5		0.4 (t > c)	
D77A	43.2		<0.025 (t)	
	oxid	*red*	*oxid*	*red*
Q80C + K116C	59.6	52.3	<0.02 (t)	5 (*c*)
G79S + Q80C + K116C	53.3	43.3	<0.02 (t)	0.1 (*t*)
D77C + N118C	53.2	29.3	<0.02 (t)	<0.02 (*t*)

Abbreviations: (c) *cis*; (t) *trans*. (From refs. 15,23.)

Correlation between thermal stability and the state of 116-117 peptide bond isomerization. Variants with higher thermal stability have been found to have higher values of K_{tc} (Table 2). This change in stability is also seen in hydrogen exchange rates monitored by NMR. Higher protection factors against hydrogen exchange are found in variants with higher values of K_{tc} or ternary complexes of nuclease with thymidine-3',5'-bisphosphate (pdTp) and calcium ion (ref. 25-26).

Correlation between kinetic parameters and the state of the 116-117 peptide bond. Table 3 shows that mutations that alter the *cis/trans* ratio have an effect on enzyme activity and that this effect is much greater on k_{cat} than on K_m.

Engineered disulfide bridges and their effect on peptide bond 116-117. Table 2 also contains results from three disulfide mutants: Q80C + K116C, G79S + Q80C + K116C, and D77C + N118C. Shown are unfolding temperatures at pH* 5.5 for protein mutants in which the disulfide bridges are oxidized or reduced. Closure of the disulfide bridge confers stability, but that gain in stability is much lower in Q80C + K116C than in the other disulfide mutants presumably because some of the energy goes into changing the configuration of the peptide bond from predominantly *cis*, when the disulfide is reduced, to predominantly *trans*, when the disulfide is oxidized (ref. 27).

Table 3. Comparison between the Peptide Bond Configuration at Position 116-117 and Kinetic Parameters for Staphylococcal Nuclease Variants (ref. 27).

nuclease (V8) variant H124L	K_{tc} = [c]/[t]	relative values		
		K_m	k_{cat}	k_{cat}/K_m
wild type	10.5	1.00	1.00	1.00
G79S	0.4	1.81	0.67	0.38
D77A	<0.02	1.78	0.23	0.13
Q80C + K116C (ox)	<0.02	1.22	0.022	0.018
Q80C + K116C (red)	5.0	3.70	1.04	0.28
G79S + Q80C + K116C (ox)	<0.02	2.31	0.0066	0.0029
G79S + Q80C + K116C (red)	0.1	1.13	0.22	0.19
D77C + N118C (ox)	<0.02	1.09	0.054	0.050
D77C + N118C (red)	<0.02	1.01	0.020	0.020

Substrate: 5-PNP-dT-3P (ref. 28).

Conditions: pH 9.5, 100 mM $CaCl_2$, 200 mM KCl, 100 mM CHES, 28°C.

High-pressure effects on nuclease variants. Fluorescence parameters (intensity, phase, modulation, and heterogeneity) on Trp^{140} in various nuclease variants have been determined as a function of pH and hydrostatic pressure up to 2.5 kbar (ref. 29). Parallel ^1H NMR studies have been carried out at 300 MHz at pressures up to 5 kbar (ref. 29). Pressure stability parallels thermal stability in that variants in which the 116-117 peptide bond is predominantly *cis* require higher pressures for unfolding.

SUMMARY AND FUTURE PROSPECTS

These studies have revealed interesting correlations among: thermal stability, GuHCl stability, pressure stability, K_{tc} for the Pro^{117} peptide bond, and enzymatic activity. Disulfide bridges have been designed that favor either a *trans* or *cis* Pro^{117} peptide bond. K_{tc} in these mutants is altered upon reducing the disulfide linkage. Thus these studies indicate how a disulfide bond can bias the population of protein configurational forms during protein folding. A critical goal of future studies is the determination of the full three-dimensional structures of nuclease mutants that have a *trans* Lys^{116}-Pro^{117} peptide bond and of disulfide mutants in the oxidized and reduced forms. Staphylococcal nuclease provides an excellent model system for pressure denaturation studies (ref. 27), and it is hoped that mutants can be designed to test hypotheses concerning the important factors that lead to protein unfolding at high pressures. Finally, the preliminary results presented here suggest a role for protein flexibility in the catalytic activity of nuclease that bears further investigation.

Acknowledgements

This work was supported by NIH grants GM35976 and RR02301. A.P.H. was supported in part by a Molecular Biophysics Training Grant, and S.N.L. was supported in part by a Cell and Molecular Biology Training Grant.

REFERENCES

1. C. B. Anfinsen, *Science,* **181**: 223-230 (1973).
2. F. A. Cotton, E. E. Hazen, Jr., and M. J. Legg, *Proc. Natl. Acad. Sci. U.S.A.*, **76** :2551-2555 (1979).
3. D. H. Meadows, J. L. Markley, J. S. Cohen, and O. Jardetzky, *Proc. Natl. Acad. Sci. U.S.A.*, **58**: 1307-1313 (1967).
4. P. J. Loll and E. E. Lattman, *Proteins: Struct. Funct. Genet.*, **5**: 183-201 (1989).
5. T. R. Hynes and R. O. Fox, *Proteins Struct. Funct. Genet.*, **10**: 92-105 (1991).
6. D. A. Torchia, S. W. Sparks, and A. Bax, *Biochemistry,* **28**: 5509-5524 (1989).
7. L. E. Kay, D. A. Torchia, and A. Bax, *Biochemistry,* **28**: 8972-8979 (1989).
8. J. Wang, D. M. LeMaster, and J. L. Markley, *Biochemistry*, **29**: 88-101 (1990).
9. J. Wang, A. P. Hinck, S. N. Loh, and J. L. Markley, *Biochemistry,* **29**: 102-113 (1990).
10. J. Wang, A. P. Hinck, S. N. Loh, and J. L. Markley, *Biochemistry*, **29**: 4242-4253 (1990).
11. J. Wang, E. S. Mooberry, W. F. Walkenhorst, and J. L. Markley, *Biochemistry* **31**: 911-920 (1992).
12. J. Wang, A. P. Hinck, S. N. Loh, D. M. LeMaster, and J. L. Markley, *Biochemistry,* **31**: 921-936 (1992).
13. S. N. Loh, C. W. McNemar, and J. L. Markley, In *Techniques in Protein Chemistry II,* (J.J. Villafranca, Ed.), Academic Press, New York, pp 275-282, 1991.
14. A. T. Alexandrescu, A. P. Hinck, and J. L. Markley, *Biochemistry,* **29**: 4516-4525 (1990).
15. A. P. Hinck, Ph.D. Thesis, University of Wisconsin-Madison, Madison, WI, USA (1993).
16. P. A. Evans, R. A. Kautz, R. O. Fox, and C. M. Dobson, *Biochemistry*, **28**: 362-370 (1989).
17. C. Grathwohl and K. Wüthrich, *Biopolymers,* **20**: 2623-2633 (1981).
18. D. Shortle, *Gene,* **22:** 181-189 (1983).
19. D. Shortle and B. Lin, *Genetics,* **110**:539-555 (1985).
20. F. W. Studier, A. H. Rosenberg, J. J. Dunn, and J. W. Dubendorff, *Methods Enzymol.*, **204**: 19- 43 (1987).
21. T. A. Kunkel, J. D. Roberts, and R. A. Zakour, *Methods Enzymol.*, **154**: 367-382 (1987).
22. F. W. Ausubel, R. Brent, R. E. Kensington, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl, *Current Protocols in Molecular Biology,* Willey, New York, 1987.
23. A. P. Hinck, E. S. Eberhardt, and J. L. Markley, *Biochemistry* (in press).
24. P. Kraulis, J. Appl. *Crystall.*, **24**: 946-950 (1991).
25. S. N. Loh, K.E. Prehoda, J. Wang and J. L. Markley, *Biochemistry*, **32**: 11022-11028 (1993)..
26. W. F. Walkenhorst, H. Roder, and J. L. Markley, unpublished results.
27. A. P. Hinck and J. L. Markley, unpublished results.
28. C.B. Grissom and J.L. Markley, *Biochemistry*, **28**: 2116-2124 (1989).
29. C. A. Royer, A. P. Hinck, S. N. Loh, K. E. Prehoda, X. Peng, J. Jonas, and J. L. Markley, *Biochemistry*, **32**: 5222-5232 (1993).