Circular Permutation of RNase T1 Through PCR Based Site-Directed Mutagenesis

Jane M. Kuo, Leisha S. Mullins, James B. Garrett, and Frank M. Raushel
Department of Chemistry, Texas A&M University
College Station, TX 77843

I. Introduction

The prediction of the three-dimensional structure of a folded protein from its primary sequence has long been of interest to biological chemists studying the relationship between protein structure and function. One approach towards a better understanding of this problem involves studying the effects that rearrangements in the primary amino acid sequence will have on the kinetics and thermodynamics of a protein-folding pathway. Goldenberg and Creighton were the first to circularly permute a protein through chemical coupling methods on purified bovine pancreatic trypsin inhibitor (BPTI) (1). The circular permutation of a protein at the genetic level was pioneered by Kirshner's laboratory, which characterized variants of phosphoribosyl anthranilate isomerase (PRAI) and dihydrofolate reductase (DHFR) (2,3). More recently, a circularly permuted variant of T4 lysozyme has also been characterized (4). Current methods for creating circularly permuted variants through alterations in the genetic sequence of a protein are laborious and involve a large number of recombinant DNA manipulations. Our laboratory has developed a much simpler procedure through the use of the polymerase chain reaction (PCR). The method involves amplifying two separate segments of a gene by PCR and combining them in reverse order in a third PCR step (Figure 1). This general method of creating circularly permuted proteins through the use of PCR technology can be applied to any cloned gene, and furthermore, can be extended to make even more complex rearrangements of protein sequences. We have applied this general technique, which requires only four oligonucleotide primers and three PCR steps, to construct a circularly permuted variant of ribonuclease T1 (RNase T1).

Copyright © 1995 by Academic Press, Inc.
All rights of reproduction in any form reserved.

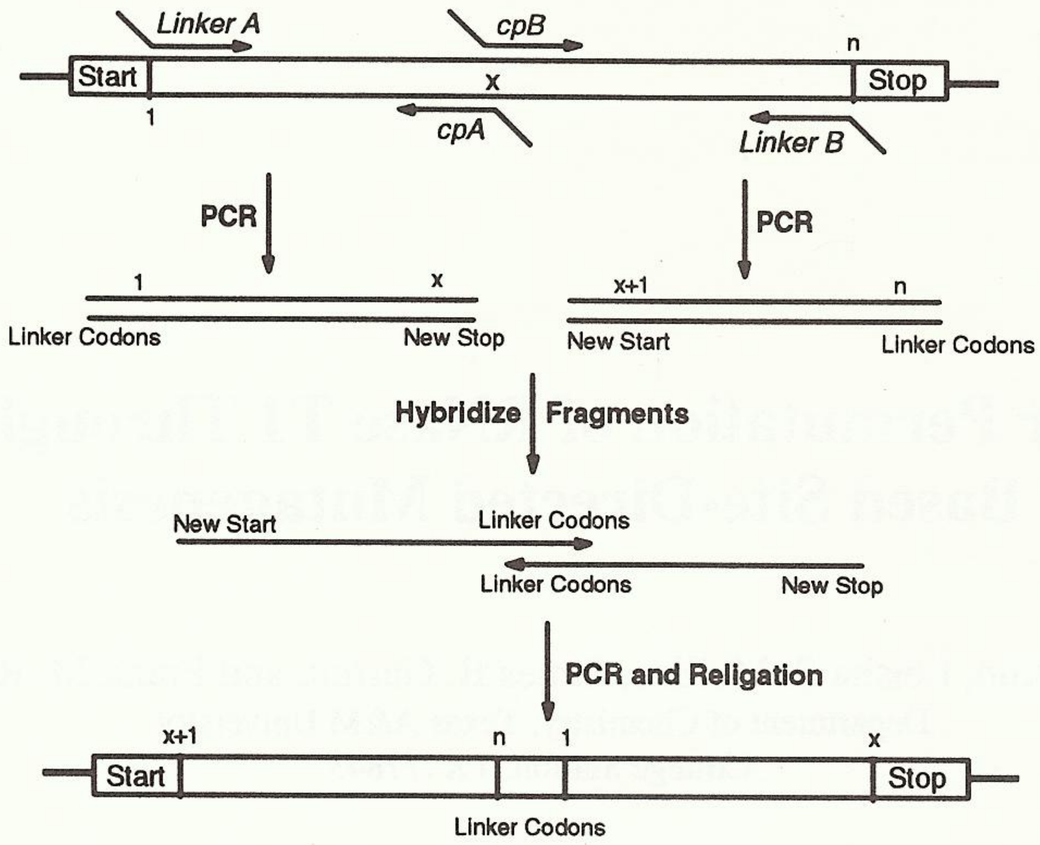

Figure 1. General PCR protocol for the creation of circularly permuted proteins.[5]

RNase T1 provides an ideal model for investigating the effects of a circularly permuted sequence on protein folding and stability, since the folding pathway of the native oxidized form of this well characterized 104 residue α,β globular protein has already been determined using the hydrogen-deuterium amide exchange 2D NMR technique (6,7). The three-dimensional structure of RNase T1 as determined by x-ray crystallography is also known to 1.5 Å resolution, and the complete NMR spectrum has previously been assigned (8,9). The kinetics and thermodynamics of the native RNase T1, which possesses two disulfide bonds, have been well characterized (10,11). Initial characterizations of both a mutated form of RNase T1, in which the 2-10 disulfide bond has been eliminated, and a circularly permuted variant of RNase T1 suggest that the two mutant proteins fold to conformational states similar to that of the native enzyme, since both forms maintain partial activity towards the catalytic hydrolysis of RNA. Thermodynamic studies indicate that the mutant forms are less stable than the native protein.

II. Materials and Methods

A. Materials

The plasmid pMac5-8 and the *Escherichia coli* strain WK6 were generous gifts from Dr. C. N. Pace (Texas A&M University). All oligonucleotides used for mutagenesis and sequencing were synthesized by the Gene Technology Laboratory of the Biology Department at Texas A&M University. Restriction

enzymes and Magic Miniprep and Magic PCR Prep DNA purification kits were purchased from Promega Corp. GeneClean DNA purification kits were purchased from Bio 101, Inc. Taq polymerase and GeneAmp PCR reagent kits were purchased from Perkin-Elmer-Cetus. Sequenase Version 2.0 sequencing kits were purchased from United States Biochemical Corp. Urea, hen egg lysozyme, and Type II-C ribonucleic acid core were purchased from Sigma.

B. Construction of Silent StyI Restriction Site

The gene for RNase T1 is encoded in the plasmid, pMac5-8, immediately downstream from the signal peptide portion of the alkaline phosphatase gene, *phoA* (12), and the protein is expressed as a *phoA*-RNase T1 precursor which undergoes posttranslational cleavage of the leader peptide. The pMac5-8 plasmid contains a unique *Hin*dIII restriction site immediately following the termination codon for the RNase T1 gene and a unique *Eco*RI restriction site upstream from the promoter region. In order to facilitate the transposition of segments of the RNase T1 gene without involving other portions of the genome, a unique and silent *Sty*I restriction site was inserted into the *phoA* leader sequence using the overlap extension PCR method for site-directed mutagenesis (13). The four primers used for this procedure, *primer A*, *StyI-B*, *StyI-C*, and *primer D2*, were designed such that *primer A* and *primer D2* flank the RNase T1 gene and anneal upstream of the *Eco*RI site and downstream of the *Hin*dIII site, respectively. *StyI-B* and *StyI-C* are complementary oligonucleotides and contain the base changes necessary for creating the silent *Sty*I site. Using the pMac5-8 plasmid as the template, two PCR amplifications were performed, one with *primer A* and *StyI-B* and one with *primer D2* and *StyI-C*. The fragments generated in these first two PCR amplifications, fragments AB and CD, were combined in a third PCR step to generate a full length AD fragment containing the newly created *Sty*I site within the *phoA*-RNase T1 gene flanked by the *Eco*RI and *Hin*dIII sites. This fragment was then cloned into the pMac5-8 plasmid between the *Eco*RI and *Hin*dIII sites, replacing the original *phoA*-RNase T1 gene and creating the new plasmid, pKW01.

C. Construction of (C2A, C10A) Mutant

The pKW01 plasmid was used as the template to create a Cys2→Ala, Cys10 →Ala double mutant for the purpose of eliminating one of the disulfide bonds in RNase T1. It was possible to accomplish this in one PCR step, because the two mutation sites were close enough to be encompassed by one mutagenic primer, *C2,10A*, which contained the base changes necessary for replacing the 2 and 10 cysteines with alanines. Primer D2 was used as the corresponding PCR primer. The resulting fragment was cloned into the pKW01 plasmid between the *Sty*I and *Hin*dIII sites, replacing the gene for the native RNase T1 with the gene for the (C2A, C10A) mutant. This plasmid was named pKW02.

D. Construction of Circularly Permuted cp35S1 Mutant

The segment between the *Sty*I and *Hin*dIII sites of the pKW02 plasmid was restricted and purified for use as the template in the first two PCR steps towards the creation of a circularly permuted variant of RNase T1. The four primers required for this procedure have been designated *cpA*, *cpB*, *linker A*, and *linker B*. The sequence (5'→3') of the *cpB* primer consisted of a portion of the *phoA* leader sequence containing the newly engineered *Sty*I site, followed by the codon for an extra alanine and the codons for Ser-35 through His-40 of RNase T1. An extra alanine was included as the first residue for the new protein to insure the proper processing of the mature protein from the *phoA* fusion product. The *linker B* primer sequence (3'→5') consisted of the complements to the codons for Gly-97 through Thr-104, a Gly-Pro-Gly linker, and Ala-1 through Ala-2. Primers *cpB* and *linker B* were combined in the first PCR step to create a fragment consisting of the first half of the gene for the circularly permuted variant. The second PCR step, which was performed in tandem with the first, utilized primers *cpA* and *linker A* to create a fragment consisting of the second half of the gene for the circularly permuted variant of RNase T1. The sequence (5'→3') of the *linker A* primer consisted of the codons for Cys-103 through Thr-104, the Gly-Pro-Gly linker, and Ala-1 through Ser-8 of the (C2A, C10A) double mutant. The *cpA* primer sequence (3'→5') consisted of the complements to the codons for Asp-29 through Gly-34 and a termination codon, followed by a *Hin*dIII site. The two fragments created in these first two PCR steps were purified and combined in a third PCR amplification to form a full length gene for the circularly permuted variant of RNase T1. Since the two fragments shared a portion of the same sequence for the Gly-Pro-Gly linker region, they acted as both primer and template for each other. Extra *cpA* and *cpB* primers were added to further amplify the final product. This circularly permuted gene was cloned into pKW02 between the *Sty*I and *Hin*dIII sites, yielding the new plasmid, pKW03. The circularly permuted protein (cp35S1) encoded by the new gene consists of 108 amino acid residues, beginning with an alanine, followed by Ser-35 through Thr-104 of the original protein, the Gly-Pro-Gly linker, and Ala-1 through Gly-34 of the (C2A, C10A) double mutant. All three constructions, contained in pKW01, pKW02, and pKW03, were completely sequenced to ensure that no unwanted mutations were incorporated during the PCR amplification steps. The construction of these plasmids from pMac5-8 is illustrated in Figure 2.

E. Enzyme Purification and Assay

The RNase T1, (C2A, C10A), and cp35S1 proteins were expressed in *E. coli* WK6 cells transformed with the plasmids pKW01, pKW02, and pKW03, respectively. The enzymes were purified according to the method described by Shirley and Laurents (14). Specific activities were determined by the continuous

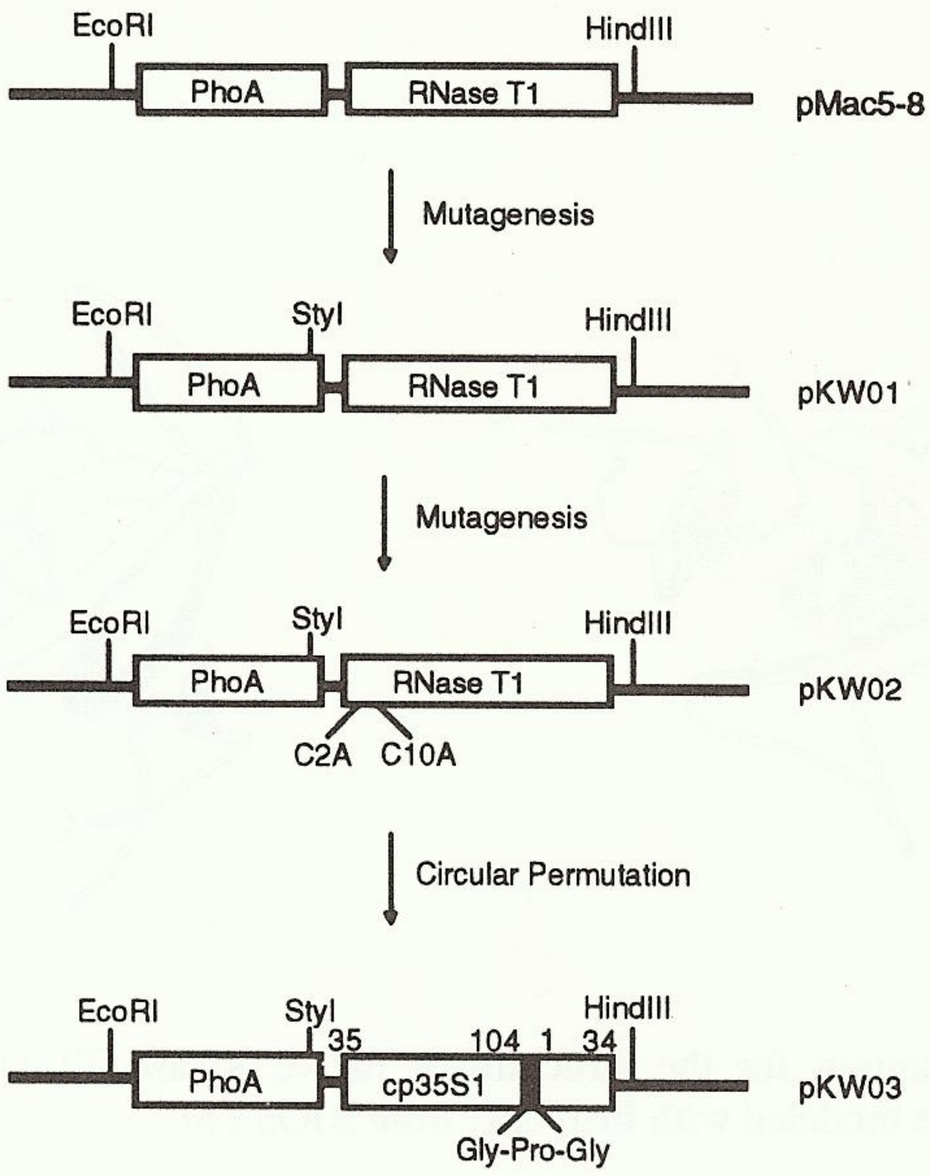

Figure 2. Construction of plasmids pKW01, pKW02, and pKW03 from pMac5-8.[5]

assay method described by Oshima *et al.* (15). Native and SDS polyacrylamide gels were run at room temperature according to the procedures described by Pace and Creighton (16).

F. Thermodynamic Measurements

Urea denaturation curves were determined by measuring the intrinsic fluorescence intensity (278 nm excitation and 320 nm emission) of solutions containing approximately 0.9 μM protein in sodium acetate/acetic acid, pH 5.0 buffer and increasing concentrations of urea in a temperature regulated Perkin-Elmer MPF 44 B spectrophotometer. Solutions were incubated at 25 °C for 24 h before measurements were taken. The free energy of unfolding was calculated by the linear extrapolation method (17). The error in ΔG was ± 0.6 kcal/mol.

G. Mass Spectrometry

The molecular masses of the RNase T1 proteins were determined by matrix-assisted laser desorption ionization (MALDI) mass spectrometry on a home-built spectrometer constructed in the laboratory of Dr. D. H. Russell (Dept. of Chemistry, Texas A&M University). The matrix solution was α-cyano-4-hydroxycinnamic acid (15 mg/mL) in methanol. Samples were ionized in a 20

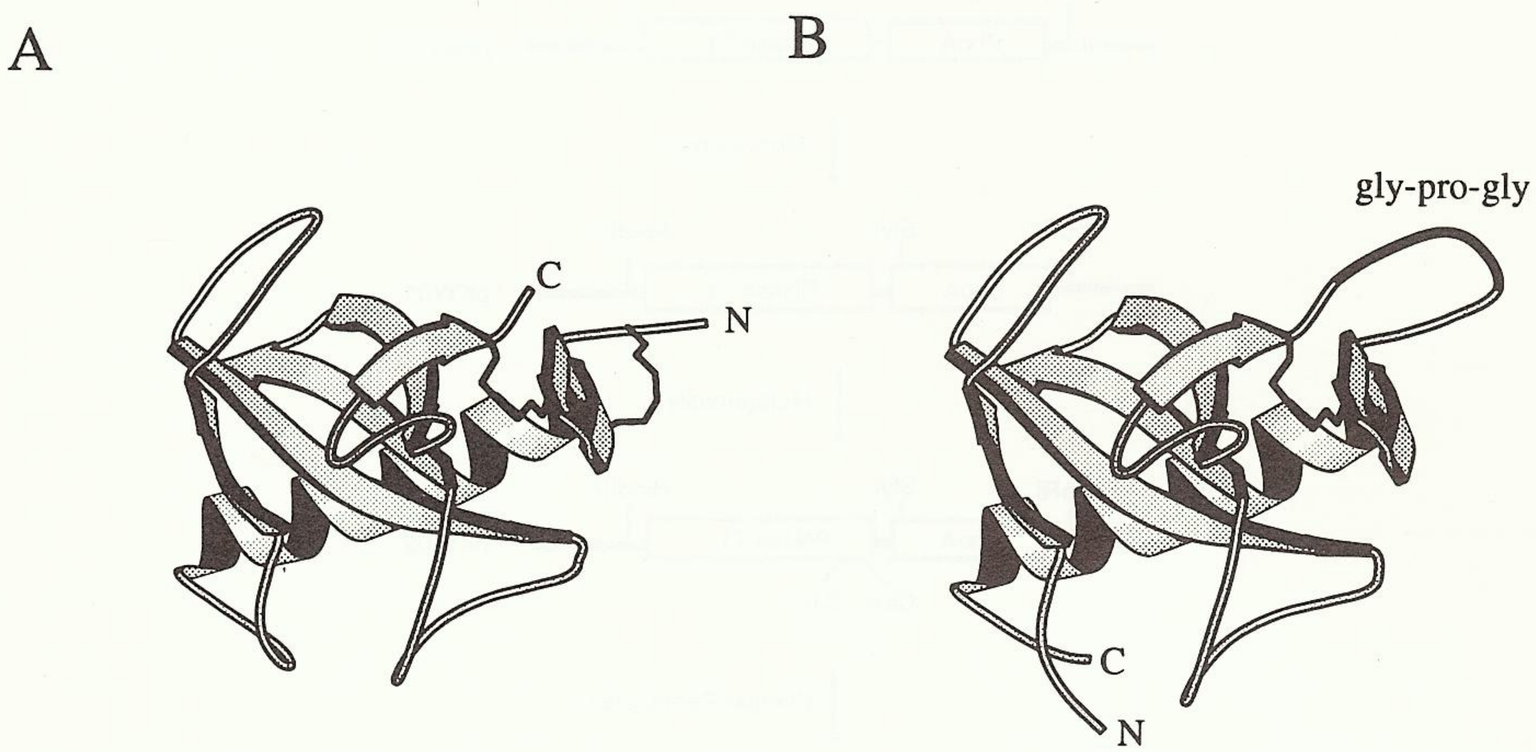

Figure 3. Ribbon diagram for the structure of native RNase T1 (A) and the circularly permuted cp35S1 (B) as modeled with Insight II from BIOSYM.[5]

kV electric field. The instrument was calibrated with hen egg lysozyme (1 mg/mL) in 50% methanol. Samples were prepared with a final protein concentration of 1-2 mg/mL and internally standardized with hen egg lysozyme (2 mg/mL, 50% methanol).

H. Amino-Terminal Sequencing

The amino-terminal sequences of the (C2A, C10A) and cp35S1 mutants were determined with an Applied Biosystems 470A sequencer in the Biology Support Laboratory at Texas A&M University.

III. Results and Conclusion

The three-dimensional structure of the native RNase T1 is shown in Figure 3A. A major and a minor loop are formed due to the two disulfide bridges (Cys-6 and Cys-103, Cys-2 and Cys-10). Since it has previously been shown that breaking the disulfide bond forming the smaller loop (Cys-2 and Cys-10) results in a protein that remains folded in a conformation very similar to that of the native enzyme (16), it was decided to remove this disulfide bridge in order to introduce greater flexibility at the amino terminus so that a peptide linker of only three amino acids would be sufficient to bridge the gap between the N- and C-terminal ends of RNase T1 (18). The Gly-Pro-Gly linker was chosen because these residues have a relatively high propensity to accommodate β-turns (19). Ser-35 and Gly-34 were chosen as the new N- and C-terminal ends for the

Table I. Comparison of Wild-type and Mutant Variants of RNase T1

Protein	Specific Activity[a] (U/mg)	Molecular Mass (calc.)[b] (daltons)	Molecular Mass (obs.)[c] (daltons)	Stability (urea, pH 5.0) (kcal/mol)	N-terminal Sequence
RNase T1	19,000	11,085	11,088	10.1	ACDYTXGSNCYS
(C2A, C10A)	17,000	11,023	11,004	6.4	AADYTXGSNAYS
cp35S1	5,400	11,305	11,303	4.2	ASNSYPHKYNNY

[a] From the continuous assay described by *Oshima et al.* Units are defined as 0.01 change in absorbance at 298.5 nm/min at 37 °C.

[b] Calculated molecular masses are based on average molecular weight for the individual amino acids.

[c] Represents the molecular mass in the +1 protonation state.

circularly permuted variant, because they are located in an exposed loop that has not been implicated in either substrate binding or catalytic activity (6). The serine at the new amino terminus is preceded by an alanine for correct polypeptide expression and processing. A computer-generated model of the circularly permuted variant of RNase T1, cp35S1, is shown in Figure 3B.

The native RNase T1, (C2A, C10A), and cp35S1 proteins were characterized by specific activity, polyacrylamide electrophoresis (SDS and native), thermodynamic stability, matrix-assisted laser desorption ionization (MALDI) mass spectrometry, and amino-terminal sequencing. Results are shown in Table I. The (C2A, C10A) and cp35S1 mutants were 89% and 28% as active as RNase T1, respectively, in hydrolyzing RNA. Electrophoresis on 15% native gels indicated that the larger cp35S1 migrated slower that RNase T1 and (C2A, C10A). Urea denaturation studies indicated that the thermodynamic stabilities of (C2A, C10A) and cp35S1 were reduced by 3.7 and 5.9 kcal/mol, respectively, relative to the native RNase T1. The *m* values for wild-type RNase T1, (C2A, C10A), and cp35S1 were 1600, 1500, and 1650 cal mol^{-1} M^{-1}, respectively. The observed molecular weights as determined by mass spectrometry for all three proteins were consistent with the predicted values based on amino acid composition. N-terminal protein sequencing verified the initial sequences of both the (C2A, C10A) and cp35S1 mutants.

The specific activity and thermodynamic stability of the (C2A, C10A) mutant confirm that the Cys-2 to Cys-10 disulfide bond imparts thermodynamic stability but has little effect on catalytic activity. Hence this mutant was selected as the starting point for constructing a circularly permuted form of RNase-T1 so that as short a linker as possible could be used to bridge the original N- and C-termini. The activity and stability of the circularly permuted variant indicate that it adopts an overall tertiary fold very similar to that of the native protein. Therefore, transposing the first 34 residues to the C-terminus has little effect on the overall folding to the final tertiary structure. The real effect, however, may be more evident in the kinetics of the specific folding pathway.

We have demonstrated the utility of a general PCR approach for the creation of circularly permuted proteins through the initial characterization of cp35S1. This circularly permuted variant of RNase T1 can be expressed and purified in a

catalytically active form. Multiple circularly permuted analogs of the same protein can easily be made by varying the *cpA* and *cpB* primers. Conversely, multiple peptide linkers can be designed by varying the *linker A* and *linker B* primers. These approaches will be used to make other circularly permuted variants of RNase T1 in order to investigate the impact of specific alterations in the amino acid sequence on the protein-folding pathways.

References

1. Goldenberg, D. P., and Creighton, T. E. (1984) *J. Mol. Biol.* **179**, 527-545.
2. Luger, K., Hommel, U., Herold, M., Hofsteenge, J., and Kirschner, K. (1989) *Science* **243**, 206-243.
3. Buchwalder, A., Szadkowski, H., and Kirschner, K. (1992) *Biochemistry* **31**, 1621-1630.
4. Zhang, T., Bertelsen, E., Benvegnu, D., and Alber, T. (1993) *Biochemistry* **32**, 12311-12318.
5. Mullins, L. S., Wesseling, K., Kuo, J. M., Garrett, J. B., and Raushel, F. M. (1994) *JACS* **116**, 5529-5533.
6. Pace, C. N., Heinemann, U., Hahn, U., and Saenger, W. (1991) *Angew. Chem.* **30**, 343-454.
7. Mullins, L. S., Pace, C. N., and Raushel, F. M. (1993) *Biochemistry* **32**, 6152-6156.
8. Martinez-Oyanedel, J., Heinemann, U., and Saenger, W. (1991) *J. Mol. Biol.* **222**, 335-352.
9. Hoffman, E., and Ruterjans, H. (1988) *Eur. J. Biochem.* **177**, 539-560.
10. (a) Shirley, B. A., Stanssens, P., Hahn, U., and Pace, C. N. (1992) *Biochemistry* **31**, 725-732. (b) Pace, C. N. (1990) *TIBS* **15**, 14-17.
11. (a) Kiefhaber, T., Quaas, R., Hahn, U., and Schmid, F. X. (1990) *Biochemistry* **29**, 3053-3061. (b) Kiefhaber, T., Quaas, R., Hahn, U., and Schmid, F. X. (1990) *Biochemistry* **29**, 3061-3070. (c) Kiefhaber, T., Grunert, H.-P., Hahn, U., and Schmid, F. X. (1990) *Biochemistry* **29**, 7475-6480. (d) Kiefhaber, T., Grunert, H.-P., Hahn, U., and Schmid, F. X. (1992) *Proteins, Struct. Funct. Genet.* **12**, 171-178. (e) Kiefhaber, T., Schmid, F. X., Willaert, K., Engelbroughs, Y., and Chaffottee, A. (1992) *Protein Sci.* **1**, 1162-1166.
12. Quaas, R., McKeown, Y., Stanssens, P., Frank, R., Blocker, H., and Hahn, U. (1988) *Eur. J. Biochem.* **273**, 617-622.
13. Ho, S. N., Hunt, H. D., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) *Gene* **77**, 51-59.
14. Shirley, B. A., and Laurents, D. V. (1990) *J. Biochem. Biophys. Methods* **20**, 181-188.
15. Oshima, T., Uenishi, N., and Imahori, K. (1976) *Anal. Biochem.* **71**, 632-634.
16. (a) Pace, C. N., and Creighton, T. E. (1986) *J. Mol. Biol.* **188**, 477-486. (b) Pace, C. N., Grimsley, G. R., Thomson, J. A., and Barnett, B. J. (1988) *J. Biol. Chem.* **263**, 11820-11825.
17. Santoro, M. M., and Bolen, D. W. (1988) *Biochemistry* **27**, 8063-8068.
18. Distances were determined by molecular modeling with Insight II from BIOSYM using the coordinates from the crystal structure determined by Martinez-Oyanedel *et al.* (7).
19. Creighton, T. E. (1993). *In* "Proteins-Structure and Molecular Principles" (W. H. Freeman and Company, New York) 201-259.