Protein Science (1996), 5:204–211. Cambridge University Press. Printed in the USA.

Are turns required for the folding of ribonuclease T1?

JAMES B. GARRETT, LEISHA S. MULLINS, AND FRANK M. RAUSHEL
Department of Chemistry, Texas A&M University, College Station, Texas 77843

(RECEIVED September 19, 1995; ACCEPTED November 13, 1995)

Abstract

Ribonuclease T1 (RNase T1) is a small, globular protein of 104 amino acids for which extensive thermodynamic and structural information is known. To assess the specific influence of variations in amino acid sequence on the mechanism for protein folding, circularly permuted variants of RNase T1 were constructed and characterized in terms of catalytic activity and thermodynamic stability. The disulfide bond connecting Cys-2 and Cys-10 was removed by mutation of these residues to alanine (C2,10A) to avoid potential steric problems imposed by the circular permutations. The original amino-terminus and carboxyl-terminus of the mutant (C2,10A) were subsequently joined with a tripeptide linker to accommodate a reverse turn and new termini were introduced throughout the primary sequence in regions of solvent-exposed loops at Ser-35 (cp35S1), Asp-49 (cp49D1), Gly-70 (cp70G1), and Ser-96 (cp96S1). These circularly permuted RNase T1 mutants retained 35–100% of the original catalytic activity for the hydrolysis of guanylyl(3′ → 5′)cytidine, suggesting that the overall tertiary fold of these mutants is very similar to that of wild-type protein. Chemical denaturation curves indicated thermodynamic stabilities at pH 5.0 of 5.7, 2.9, 2.6, and 4.6 kcal/mol for cp35S1, cp49D1, cp70G1, and cp96S1, respectively, compared to a value of 10.1 kcal/mol for wild-type RNase T1 and 6.4 kcal/mol for (C2,10A) T1. A fifth set of circularly permuted variants was attempted with new termini positioned in a tight β-turn between Glu-82 and Gln-85. New termini were inserted at Asn-83 (cp83N1), Asn-84 (cp84N1), and Gln-85 (cp85Q1). No detectable amount of protein was ever produced for any of the mutations in this region, suggesting that this turn may be critical for the proper folding and/or thermodynamic stability of RNase T1.

Keywords: circular permuted proteins; protein folding; ribonuclease T1

The mechanism by which proteins fold from a linear chain of amino acids to a specific tertiary structure has been of interest to biological chemists for some time. Even though there are several overlapping theories to explain the protein-folding process, current experimental data from a number of laboratories indicate that many proteins fold along a small number of sequential pathways and form a finite number of transient intermediates (Jennings & Wright, 1993; Bai et al., 1995). It appears evident that protein folding begins with the initial formation of discreet secondary structural units that subsequently undergo reordering or further interaction to form the framework for the final tertiary structure. If the initial steps along the folding pathway involve the nucleation of secondary structural units, then it is important to address what role the linear organization of the primary amino acid sequence has on these steps. This feature of the protein-folding process can be examined by circularly permuting the primary amino acid sequence of a protein such that the original amino- and carboxyl-termini are covalently linked and new termini are created at alternate sites within the sequence.

The primary amino acid sequence of a protein can be divided into individual segments that, in the folded structure, form small secondary structural units such as α-helices and β-sheets. Connecting these segments are regions that ultimately form loops, turns, or areas of nonordered structure in the folded protein. The simplest way to probe the importance of the order of the amino acid sequence on protein folding is to keep the secondary structural segments of the primary sequence intact while systematically creating new termini in the nonordered or loop regions. In this way, secondary structural elements can be separated from their original environment, allowing insights into the role of loops or turns in the overall folding process.

Since the pioneering experiments of Goldenberg and Creighton (1983), circular variants of a number of proteins have been created and characterized (Luger et al., 1989; Buchwalder et al., 1992; Horlick et al., 1992; Yang & Schachman, 1993; Zhang et al., 1993; Kreitman et al., 1994; Mullins et al., 1994; Vignais et al., 1995). In addition, a systematic circular permutation of all three turns in the α-spectrin protein has recently been reported by Viguera et al. (1995). Although most examples have

Reprint requests to: Frank M. Raushel, Department of Chemistry, Texas A&M University, College Station, Texas 77843; e-mail: raushel@tamu.edu.

Abbreviations: GpC, guanylyl(3′→5′)cytidine; RNase T1, ribonuclease T1; BSA, bovine serum albumin; MES, *N*-morpholinoethanesulfonic acid; IPTG, isopropyl β-D-thiogalactopyranoside.

involved single-domain monomeric proteins, circular variants have also been made of multidomain proteins such as T4 lysozyme and aspartate transcarbamoylase. Regardless of the topographical fold of the protein, all circular variants reported to date have shown properties similar to the native protein. We have now extended these investigations by measuring the effects of creating new termini in every major loop that connects the secondary structural elements in ribonuclease T1.

RNase T1 is a small globular protein of 104 amino acids that cleaves RNA at guanosine sites. The protein has been extensively characterized, which makes it an excellent model for protein-folding studies. The native structure, as determined by X-ray crystallography, consists of a 4.5-turn α-helix packed against a central 5-strand antiparallel β-sheet (Martinez-Oyanedel et al., 1991). The native protein is stabilized by two disulfide bonds. Both oxidized and reduced forms of RNase T1 have been well characterized thermodynamically and kinetically (Pace et al., 1988). The fully reduced protein can fold and adopt a native-like structure under favorable solvent conditions. The removal of one of the disulfide bonds, that between Cys-2 and Cys-10, slightly reduces the conformational stability of the protein. The remaining disulfide bond, between Cys-6 and Cys-103, effectively constrains the native structure into a large loop. These characteristics make RNase T1 an excellent model for use in a systematic study into the role of turns on the mechanism of protein folding.

A ribbon diagram for the native enzyme is shown in Figure 1. The secondary structural elements of RNase T1 are connected by five major loops. It was possible to create circular variants of RNase T1 that have new termini in each of the loops except for the tight β-turn between residues 83 and 84. The genes for all five possible circular variants of RNase T1 have been created and the structural, thermodynamic, and kinetic properties of the expressed proteins have been characterized.

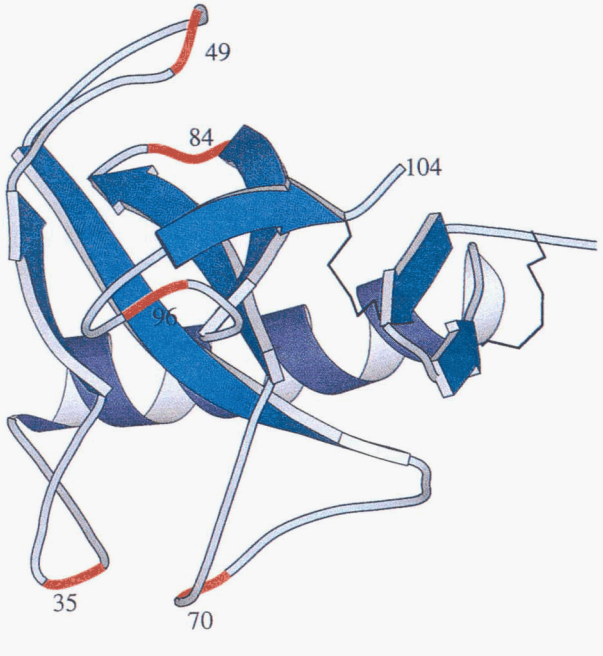

Fig. 1. Ribbon diagram for the three-dimensional structure of RNase T1 (Martinez-Oyanedel et al., 1991). Red regions indicate positions in the wild-type structure that were chosen as sites to create new termini in the circularly permuted variants.

Results

Six new proteins were designed, isolated, and characterized to examine the role of surface loops and turns on the mechanism of protein folding. The gene for RNase T1 was circularly permuted to produce mutant proteins in which the original termini were joined together with a suitable peptide linker, and new amino- and carboxyl-terminal ends were created within solvent-exposed regions of the native structure. This objective was accomplished using PCR to amplify specific regions of the native gene for RNase T1 and then these gene fragments were recombined in reverse sequence for the purpose of constructing a series of circularly permuted proteins for these investigations.

Proper linker design is an essential requirement for the fabrication of circularly permuted proteins. In the wild-type protein, the N- and C-terminal ends of RNase T1 are 11 Å apart and structurally constrained due to the presence of a nearby disulfide bond. Theoretically, this span could be bridged by the addition of a peptide linker of four or five amino acids in length. However, the disulfide bond that links the cysteine residues at positions 2 and 10 is not required for proper protein folding or enzymatic activity (Pace & Creighton, 1986) and thus it was removed by mutation of these two residues with alanine replacements. Elimination of the disulfide bond at this site was anticipated to increase the flexibility of the amino-terminus and thus reduce the effective distance between the original termini to a gap that could be spanned by a shorter three-amino acid linker. Two linking peptides (Gly-Pro-Gly and Gly-Gly-Gly) were designed from amino acids with high frequencies for occurring in β-turns (Levitt, 1978). The overall effect on the thermodynamic stability for each of these linker peptides was assessed by incorporation of these two linkers into the cp35S1 protein.[1] The overall stability of cp35S1-ggg was greater than cp35S1-gpg with a $\Delta\Delta G$ of 1.1 kcal/mol (Fig. 2). Therefore, the rest of the circularly permuted mutants of RNase T1 incorporated the Gly-Gly-Gly linker.

The secondary structural elements of the wild-type RNase T1 are separated by five solvent-exposed loops. These surface loops and turns were chosen as the targets for our investigation into the contribution of these structural features on the mechanism of protein folding. Solvent-exposed loops in the native enzyme occur between residues 29 and 40, residues 42 and 57, residues 61 and 76, residues 81 and 86, and between residues 91 and 101. In order to make circularly permuted variants of RNase T1, new termini were introduced roughly halfway between the pre-existing secondary structural elements of the wild-type protein at residues Ser-35, Asp-49, Asn-84, Gly-70, and Ser-96.

Four of the circularly permuted variants (cp35S1, cp49D1, cp70G1, and cp96S1) were expressed and purified to homogeneity as determined by PAGE. No new protein was detectable by PAGE, kinetic activity, or immunological blotting when attempts were made to express the cp84N1 mutant. Subsequent

[1] Proteins from these genetic constructions were designated by the notation cpXyZ, where cp stands for circularly permuted, X is the wild-type position of the new N-terminus, y is the amino acid that was moved to the new amino-terminus, and Z is the new location of that amino acid in the circularly permuted sequence.

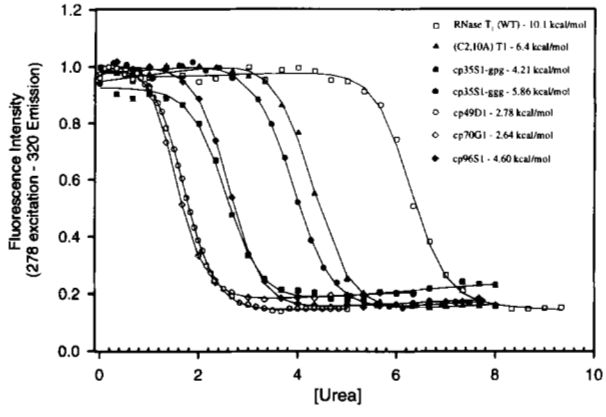

Fig. 2. Urea denaturation curves for the unfolding of RNase T1 and the circularly permuted variants.

attempts to express cp83N1 and cp85Q1 also failed repeatedly. The typical purification yields of the other circularly permuted proteins averaged 40 mg of homogeneous protein from approximately 180 g of cell paste. Native RNase T1, the parent mutant (C2,10A), and the four circularly permuted variants were subsequently characterized by PAGE (SDS and native gels at 4 °C and 25 °C), CD, N-terminal amino acid sequencing, urea denaturation, and catalytic activity.

The authenticity of the five circularly permuted variants of RNase T1 was confirmed by N-terminal amino acid sequencing. Each of these proteins was subjected to at least 10 cycles of Edman degradation. The derived N-terminal amino acid sequence of the isolated proteins (Table 1) matches exactly with the predicted sequence based on the genetic reconstruction.

Urea denaturation experiments were utilized to determine the relative thermodynamic stabilities of each of these RNase T1 mutants. The urea denaturation profiles are illustrated in Figure 2 and the corresponding ΔG_{H_2O} values at pH 5.0, 25 °C are presented in Table 1. All of the mutants are less stable than the wild-type enzyme. Those mutants with new termini created nearer the center of the native sequence (cp49D1 and cp70G1) are less stable than the corresponding mutations made closer to the original ends of the native enzyme (cp35S1 and cp96S1).

The parent protein (C2,10A) and all four of the circularly permuted proteins were catalytically active in the hydrolysis of RNA and the dinucleotide GpC. The k_{cat} values ranged from 36 to 100% of the wild-type value and the K_m values for the GpC substrate ranged from 210 to 2,500 μM. The kinetic constants are listed in Table 1.

Alterations in the tertiary structure of the circularly permuted proteins were probed by CD. The CD spectra of the wild-type protein and the five mutants are shown in Figure 3. In comparison to the wild-type and (C2,10A) proteins, all of the mutant enzymes displayed CD spectra that are consistent with somewhat less secondary structure with more random coil and/or alterations in the environment of the aromatic residues. This is consistent with the mutants still retaining significant catalytic activity. Variants cp49D1 and cp70G1 showed the largest differences in the CD spectra relative to the native enzyme.

Discussion

Several circularly permuted variants of RNase T1 have been created by rearrangement of the secondary structural elements within the native protein sequence. Perturbations in the overall thermodynamic stability and kinetic parameters of the mutant proteins were measured in an attempt to evaluate the effect of this type of mutation on the structure and function of RNase T1. By systematically creating new termini at all of the solvent-exposed loops and turns of this protein, we sought to determine the essential character of these structural features for the folding of RNase T1.

Circularly permuted protein variants are formed when the original amino- and carboxyl-termini are joined by a suitable peptide linker and a single amide bond is broken at another location. Several considerations were kept in mind during the design of these mutants. The termini of the native protein must be in close proximity to one another and the new peptide linker must not significantly disrupt the overall protein structure by exerting conformational strain at the linkage site. In the crystal structure of the native RNase T1, the two termini lie 11 Å apart. In order to minimize the gap for bridging the two ends, the disulfide bond between Cys-2 and Cys-10 was removed by replacement of the cysteine residues with alanines. This was done in an attempt to eliminate the need for additional amino acids in the peptide linker by bringing the two ends into closer proximity. The disulfide bond connecting residues 2 and 10 has been shown previously to be nonessential for the protein to assume a native-like conformation and catalytic activities (Pace et al., 1988).

Table 1. *Kinetic and thermodynamic characteristics of wild-type and mutant RNase T1 proteins*

Enzyme	N-terminal sequence	Stability in urea (kcal/mol)	m Value (cal mol^{-1} M^{-1})	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (μM^{-1} s^{-1})
RNase T1	–	10.1	−1,600	1,400 ± 90	320 ± 40	4.4 ± 0.3
(C2,10A)	AADYTCGSNA	6.4	−1,500	1,200 ± 30	210 ± 10	5.6 ± 0.2
cp35S1-gpg	ASNSYPHKYN	4.6	−1,650	–	–	–
cp35S1-ggg	ASNSYPHKYNN	5.7	−1,450	760 ± 40	510 ± 50	1.5 ± 0.1
cp49D1-ggg	ADFSVSSPYY	2.9	−1,790	1,400 ± 100	2,500 ± 200	0.6 ± 0.01
cp70G1-ggg	AGGSPGADRV	2.6	−1,840	970 ± 90	320 ± 50	3.0 ± 0.3
cp96S1-ggg	ASGNNFVECT	4.6	−1,760	510 ± 50	390 ± 60	1.3 ± 0.1

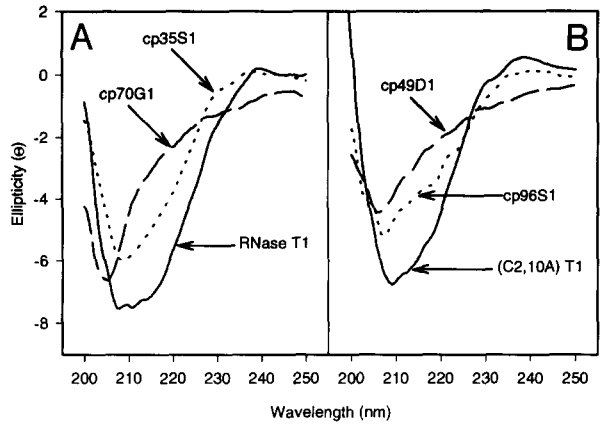

Fig. 3. CD spectra for RNase T1 and the circularly permuted variants.

In designing the linking peptide, it was important to utilize an amino acid sequence that would assume a reverse turn because the amino- and carboxyl-termini are antiparallel to one another. Two sequences were designed from amino acids with high frequencies for occurring in β-turns (Levitt, 1978). The first linker (Gly-Pro-Gly) tested was found to impart a lower stability (1.1 kcal/mol) than an alternate linker composed of Gly-Gly-Gly. Presumably, the increase in stability for the all-glycine linker arises because of the greater conformational flexibility of the middle glycine relative to the more restrictive proline residue. Therefore, all subsequent mutants were constructed with the Gly-Gly-Gly linker. Experiments to further optimize the sequence of the linking peptide are in progress.

The specific locations for the new termini created within the solvent-exposed loops were selected to be as far as possible from the individual units of protein secondary structure to avoid disrupting the local structure of the β-strands or α-helix. Therefore, each of the new termini was inserted approximately halfway between the ends of the existing units of protein secondary structure. These positions correspond to the sites selected for the circularly permuted proteins cp35S1, cp49D1, cp70G1, cp84N1, and cp96S1.

The circularly permuted variant, cp35S1, proved to be the most stable of all the circularly permuted variants created for this investigation. The solvent-accessible loop, located between wild-type protein residues 29 and 40, was disconnected between residues 34 and 35. Serine-35 was chosen as the new amino-terminus because it was located on the most solvent-exposed loop and was not near the active site. This cleavage site occurs between the lone α-helix and the first strand of the β-sheet of the wild-type protein. This mutant was only 0.7 kcal/mol less stable than the parent (C2,10A) mutant. The Michaelis constant for the GpC substrate was elevated by a factor of approximately 2 relative to either the wild-type protein or the protein (C2, 10A). This increase in K_m and corresponding decrease in V_{max} is most likely due to the increase in conformational flexibility near the new termini. In addition, the side-chain phenol of Tyr-38 has been shown (Martinez-Oyanedel et al., 1991) to hydrogen bond to the phosphate group of the substrate and it is quite likely that this interaction has been perturbed somewhat.

The second site for circular permutation occurs on the solvent-exposed loop between residues 42 and 57. In cp49D1, the K_m for the substrate GpC increased by almost an order of magnitude, although V_{max} remained essentially unchanged. In this mutant, the new termini are located quite close to the active site of the protein and to the residues involved in substrate specificity and binding. The significant increase in the K_m for GpC is due in part to the proximity of the circular permutation to one of the residues (Glu-46) involved directly in substrate specificity and binding (Steyaert et al., 1991a). The side chain of Glu-46 has been proposed to hydrogen bond to the NH_2 group at C-2 and the N(1)-H of the guanine moiety of the substrate. The K_m for mutants of the wild-type enzyme increase by an order of magnitude when Glu-46 is replaced with either an alanine or glutamine residue (Steyaert et al., 1991a). The circular permutation at position 49 most likely imparts a great deal more flexibility to the binding site and lowers the overall protein stability significantly from that of (C2, 10A) T1. The CD spectrum of cp49D1 is also measurably different from that of the wild-type protein in that it exhibits more random coil-like structure.

The third site for circular permutization of RNase T1 occurs between residues 61 and 76. In the circularly permuted variant designated as cp70G1, the thermodynamic stability was slightly less than that of cp49D1, but the mutant retained kinetic parameters that more closely resembled those of the wild-type enzyme. In this variant, the decreased stability may be attributed to a change in the hydrogen bonding pattern for two of the structurally conserved water molecules (Pletinckx et al., 1994). Within the RNase T1 structure, there exists a chain of 10 hydrogen bonded, structurally conserved water molecules (Malin et al., 1991). One of the ends of this chain occurs near the site of the cp70G1 mutation and hydrogen bonds Tyr-68, Trp-59, and Pro-60 with one water molecule of the chain. Another important water molecule increases local stability within the loop by hydrogen bonding to Asp-76, Cys-6, Ser-8, Asn-9, and Thr-93 (Pace et al., 1991). By increasing the mobility of the residues close to these conserved water molecules, the structure might be perturbed enough to lose one or more of these stabilizing interactions and decrease stability.

The fourth solvent-exposed loop to be examined occurs between residues 81 and 86. Asn-84 was chosen as the initial site for circular permutization at this turn, but no protein could be detected by any of the assay methods. Two other sites, Asn-83 and Glu-85, also occurring within this loop, were subsequently investigated as possible termini, but again the expression system failed to produce significant amounts of detectable protein by activity assays, SDS-PAGE, or immunoblots. The most likely explanation for the apparent lack of production of any of these mutants is the extreme loss of thermodynamic stability under the expression conditions and subsequent proteolysis. However, we cannot rule out unidentified problems with the biosynthetic apparatus during the expression of mutants within this region. Nevertheless, the loop between residues 81 and 86 is a very tight β-turn. This region of the protein is the only site tested where no mutant could be expressed and thus this region appears to be critical for proper folding and/or thermodynamic stability. This turn occurs between two β-strands that form the center of the hydrophobic core of the protein. Mullins et al. (1993) found that the hydrogen bonds formed between these two strands are among the earliest to be formed during the folding of wild-type RNase T1. Moreover, in the native enzyme, the amide hydrogens involved in the hydrogen bonding between these two strands are among the most resistant to exchange in the whole

protein (L.S. Mullins, unpubl. obs.). These results are consistent with a working model where the turn composed of amino acid residues 82–85 acts as a nucleation site for formation of the antiparallel β-sheet and the folding of the rest of the protein.

The remaining solvent-exposed loop in the protein occurs between residues 90 and 103. The N-terminus of the circularly permuted protein at this site was positioned at Ser-96. The cp96S1 mutant showed the lowest catalytic rate of all the circularly permuted variants expressed, although the k_{cat} was reduced by only a factor of about 2.5. This loss in activity is most likely due to the location of the new termini because it is quite close to His-92 and Asn-98. His-92 has been proposed to act as a general acid/base during the cleavage of RNA (Heinemann & Saenger, 1982; Steyaert et al., 1990) and Asn-98 has been proposed to facilitate the binding of the guanine moiety by backbone hydrogen bonding with the N(2)-H (Steyaert et al., 1991b).

In summary, four of the five solvent-exposed loops of RNase T1 can be exploited to make circularly permutized mutants of the native structure. The isolated mutant proteins are all catalytically active and reasonably stable and thus these sites cannot be essential for proper folding of RNase T1. These results demonstrate the overall flexibility of the fully folded protein for significant alterations for the specific locations of the chain termini. Similar results have been obtained by Viguera et al. (1995). They were able to construct three circularly permuted variants of α-spectrin, including a site located within a very tight β-turn. However, the complete lack of any mutant constructed within the tight β-turn of RNase T1 suggests that this particular site in RNase T1 may be significantly different. It is suggested that this site may participate in the nucleation of protein folding of RNase T1 by the initiation of a localized antiparallel β-sheet. Experiments designed to measure the specific effects of these circular permutations on the kinetics of protein folding are in progress.

Materials and methods

Materials

Urea, GpC, all buffers, and Type II-C ribonucleic acid core were purchased from Sigma. All other reagents and restriction enzymes were purchased from either Promega, Perkin-Elmer, Stratagene, or United States Biochemical Corp. The oligonucleotides that were used for mutagenesis and sequencing were synthesized by the Gene Technology Laboratory in the Biology Department at Texas A&M University. The plasmid pMc5TPRTQ (Steyaert et al., 1990) and *Escherichia coli* strain WK6 (Zell & Fritz, 1987) were generous gifts from Professor C.N. Pace of Texas A&M University.

Construction of the circularly permuted genes

The gene for RNase T1 is encoded in the plasmid, pMc5TPRTQ, immediately downstream from the signal peptide portion of the alkaline phosphatase gene, *phoA*, and under the transcriptional control of the *tac* promoter. The whole assembly is bounded by two unique restriction sites. Upstream of the *tac* promoter is an *Eco*R I restriction site, whereas a *Hin*d III restriction site is found at the 3′ end of the RNase T1 gene. During expression, the protein is directed to the periplasmic space where the leader peptide is cleaved from the fusion protein and RNase T1 accumulates in the intermembrane space.

To facilitate the construction of the genes encoding the circularly permuted proteins, two modifications were made to the plasmid, pMc5TPRTQ, encoding the wild-type RNase T1. First, pMc5TPRTQ was modified using the PCR overlap extension method (Ho et al., 1989) to include a silent *Sty* I restriction site in the *phoA* leader sequence prior to the beginning of the RNase T1 gene. The resulting plasmid, pKW01, was then used as a template to create a gene for a modified RNase T1 in which the disulfide bond between cysteines 2 and 10 was removed by replacement of both cysteines with alanine residues. The plasmid containing the gene for the modified protein was designated pKW02 and was used as the first template in the PCR amplification steps for the construction of the genes of the circular variants.

The procedure to construct the genes encoding the circularly permuted variants of RNase T1 used four oligonucleotide primers in three PCR overlap extension steps as described previously by Mullins et al. (1994). This procedure is detailed in Figure 4. The primers, designated cpA, cpD, linker B, and linker C, are shown in Table 2. The cpA and cpD primers are the circularly permuting primers encoding the new amino- and carboxyl-termini, respectively, whereas the linker B and linker C primers encode the residues that link the original amino- and carboxyl-termini. The cpA primer consists of a portion of the *phoA* leader sequence and a *Sty* I restriction site, followed by a codon for an extra alanine, and then the codons for the first five amino acids of the desired circularly permuted protein. The extra alanine was added to ensure the proper processing of the mature circularly permuted protein from the *phoA* fusion product. The linker B primer contains the codons for Gly-97–Thr-104 of the wild-type protein, the codons for the three-residue linker, and the codons for the first two residues of the (C2,10A) protein. The analogous primer, linker C, contained the codons for Cys-103 and Thr-104, the codons for the three-residue linker, and the codons for the first eight residues of the (C2,10A) protein. The cpD primer contained the codons for the last five residues of the desired circularly permuted protein, a stop codon, and the *Hin*d III restriction site.

As seen in Figure 4, the first PCR step involved the amplification of pKW02 with the primers cpA and linker B to create

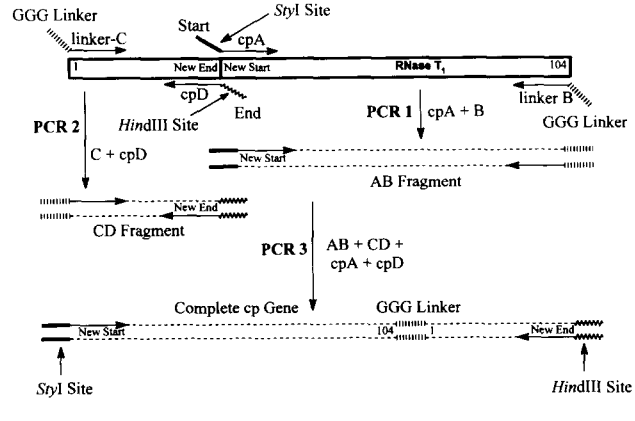

Fig. 4. Schematic diagram of the PCR strategy used for the construction of circularly permuted variants of RNase T1.

Table 2. *PCR primers used in the creation of RNase T1 mutants*[a]

Primer	Sequence 5′ → 3′												
Linker B-gpg	AGC	AGC	ACC	GGC	ACC	TGT	ACA	TTC	AAC	GAA	GTT	GTT	ACC
Linker C-gpg	TGT	ACA	GGT	CCG	GGT	GCT	GCT	GAC	TAC	ACT	TGC	GGT	TCT
Linker B-ggg	AGC	AGC	ACG	GCC	ACC	TGT	ACA	TTC	AAC	GAA	GTT	GTT	ACC
Linker C-ggg	TGT	ACA	GGT	GGC	GGT	GCT	GCT	GAC	TAC	ACT	TGC	GGT	TC
cp35A	AGA	CGG	TGC	CAA	GGC	GGC	TTC	CAA	TTC	TTA	CCC	ACA	C
cp35D	TGG	GTA	AGA	AGC	TTA	TCC	AAC	AGT	TTC	ACC	GTC		
cp49A	AAC	AAC	TCC	AAG	GCG	GCT	GAT	TTC	TCT	GTG	AGC	TCT	CCC
cp49D	GCT	CAC	AGA	AGC	TTA	AAA	ACC	TTC	GTA	GTT	GTT		
cp70A	TCG	AGC	GGT	GCC	AAG	GCG	GCT	GGT	GGG	TCC	CCG	GGT	GCT
cp70D	ACC	CGG	GGA	AGC	TTA	AGA	GTA	AAC	ATC	ACC	GCT		
cp83A	CGT	GTC	GCC	AAG	GCG	GCT	AAC	AAC	CAA	CTA	GCT	GGT	
cp83D	AGC	TAG	TTA	AGC	TTA	TTC	GTT	GAA	GAC	GAC	ACG		
cp84A	GTC	GTC	TCC	AAG	GCG	GCT	AAC	CAA	CTA	GCT	GGT	GTT	
cp84D	ACC	AGC	TAA	AGC	TTA	GTT	TTC	GTT	GAA	GAC	GAC		
cp85A	GTC	TTC	ACC	AAG	GCG	GCT	CAA	CTA	GCT	GGT	GTT	ATC	
cp85D	AAC	ACC	AAA	AGC	TTA	GTT	GTT	TTC	GTT	GAA	GAC		
cp96A	GTT	ATC	ACT	CCC	AAG	GCG	GCT	TCT	GGT	AAC	AAC	TTC	GTT
cp96D	GAA	GTT	GTA	AGC	TTA	AGC	ACC	AGT	GTG	AGT	GAT		

[a] Underlined nucleotides indicate mutagenic sites.

a DNA fragment encoding a region from the new amino-terminus through the linker. In the second PCR step, pKW02 was amplified with the cpD and linker C primers yielding a DNA fragment encoding the linker region to the new carboxyl-terminus. The amplified fragments from the first two PCR steps were then isolated and combined in a third PCR step with the cpA and cpD primers to create an amplified fragment containing the entire circularly permuted sequence. The circularly permuted fragment was then restricted with *Sty* I and *Hind* III, purified, and cloned into a pKW02 fragment that had also been restricted with these same enzymes and purified. The newly constructed, circularly permuted variant of RNase T1 thus encodes a protein with 108 amino acids. The RNase T1 genes within pKW01, pKW02, and all the circularly permuted plasmids (pKW03, pJG35, pJG49, pJG70, pJG83, pJG84, pJG85, and pJG96) were completely sequenced in order to ensure that no unwanted mutations were incorporated during the PCR amplification steps.

Immunological blots

Wild-type RNase T1, (C2,10A) T1, and all the circular variants were screened by immuno-blotting. The serum from a rabbit with a titer to wild-type RNase T1 was used for western blot analysis of RNase T1 and the mutants. Both SDS and native polyacrylamide gels were used to separate the proteins present in the supernatant fluid from an osmotic shock of *E. coli* WK6 cells. Proteins were resolved using a 15% polyacrylamide gel, after which the proteins were transferred to nitrocellulose paper by overnight electrophoresis at 12 V in the running buffer (Towbin et al., 1979; Harlow & Lane, 1988). The nitrocellulose paper was then blocked with BSA and probed with rabbit antibodies. The attached rabbit antibodies were then probed with goat anti-rabbit Ig antibodies conjugated to horseradish peroxidase. Color development was done with hydrogen peroxide and 4-chloro-1-naphthol according to the procedure described in the literature from Promega.

Amino-terminal sequencing

The amino-terminal protein sequences of (C2,10A) T1, cp35S1-gpg, cp35S1-ggg, cp49D1-ggg, cp70G1-ggg, and cp96S1-ggg were determined with an Applied Biosystems 470A sequencer in the Biotechnology Support Laboratory of Texas A&M University.

Cell growth and enzyme purification

Wild-type RNase T1, (C2,10A) T1, and the circular variants were expressed and purified as described in Mayr and Schmid (1993) and Shirley and Laurents (1990), with the following modifications. All the plasmids were transformed into *E. coli* WK6 for cell growth and subsequent protein purification. The transformed WK6 cells were grown in a Lab Line/S.M.S. Hi-density Fermentor in TB media (37 °C, 10 μg/mL chloramphenicol, air at 15 $L \cdot min^{-1}$) after inoculation with a 6-mL cell culture (OD_{600} greater than 1.0), which was grown in LB media (20 μg/mL chloramphenicol). Cell growth was monitored until the OD_{600} was greater than 2.5, at which time protein production was induced by the addition of 200 mg IPTG concomitant with reduction of the temperature to 25 °C. The cells were allowed to continue to grow for an additional 12 h and then harvested by centrifugation for 15 min at 8,000 × *g*.

All purification steps were performed at 4 °C unless noted otherwise. Pellets were resuspended in 300 mL of 50 mM Tris-HCl, pH 7.5, 20% sucrose, 10 mM EDTA, and stirred for 45 min. Cells were centrifuged (30 min, 8,000 × *g*) and the supernatant fluid was saved. Supernatant was diluted to 4 L with 400 mL of 0.5 M sodium phosphate buffer and deionized water, and the pH adjusted to 7.1. The solution was then applied to a Zeta-prep® QAE cartridge preequilibrated with 50 mM sodium phos-

phate buffer, pH 7.1. Upon completion of sample loading, the cartridge was washed with 5 volumes of 50 mM sodium phosphate buffer, pH. 7.1, and then the protein was eluted with 100 mM sodium phosphate buffer, pH 2.7. Fractions containing RNase T1 activity were pooled and then concentrated using diafiltration. Samples were desalted by continual diafiltration against deionized H_2O until a dilution factor of 10^9 was reached. Protein was then assayed, concentrated, and stored at 4 °C.

Enzyme assay and kinetics

Enzyme activity was determined by the catalytic hydrolysis of RNA using a continuous assay (Oshima et al., 1976) with Type II-C ribonucleic acid core. All kinetic constants were determined at 25 °C with GpC as the substrate in 50 mM MES, pH 6.5, and 100 mM KCl. GpC concentrations were determined at 278 nm using a molar extinction coefficient of 1.26×10^4 M^{-1} cm^{-1} (Zabinski & Walz, 1976). Initial velocities were determined by measuring the increase in absorbance at 280 nm with a Gilford 260 spectrophotometer using a pathlength of 0.1 cm. RNase T1 concentrations were determined using a molar extinction coefficient of 1.7 at 278 nm for a 1 mg/mL solution in a 1.0-cm-pathlength cell (Hu et al., 1992). GpC concentrations were varied between 37 μM and 637 μM. The protein concentrations used in the assays were 9 nM for RNase T1 and (C2,10A) and 18 nM for the circularly permuted proteins. Kinetic parameters were determined by fitting the data to Equation 1 using the HYPERO program (Cleland, 1967) from Savanna Shell Software. In this equation, v is the initial velocity, V_{max} is the maximal velocity, $[S]$ is the substrate concentration, and K_m is the Michaelis constant.

$$v = \frac{V_{max}[S]}{K_m + [S]} \quad (1)$$

Thermodynamic measurements

Urea denaturation curves were determined for RNase T1, (C2,10A) T1, cp35S1-gpg, cp35S1-ggg, cp49D1-ggg, cp70G1-ggg, and cp 96S1-ggg by measuring the intrinsic fluorescence intensity of tryptophan with a thermostated SLM Aminco Series II spectrofluorometer. The fluorescence intensity (278 nm excitation and 320 nm emission) of solutions containing approximately 3 μM protein were measured after incubation in denaturant for 24 h. The free energy of unfolding was calculated by the linear extrapolation method and fit to Equation 2 using iterative calculations (Santoro & Bolen, 1988). In this equation, Y_O is the observed fluorescence intensity, Y_N is the fluorescence intensity of the native conformation, Y_U is the fluorescence intensity of the unfolded molecule, m_N is the slope of the increase in fluorescence intensity of the native molecule, m_U is the slope of the increase in fluorescence intensity of the unfolded molecule, $[D]$ is the denaturant concentration, ΔG_{N-U} is the change in free energy in going from the folded to unfolded state, m_G is the slope of the ΔG_{N-U} extrapolation, R is the universal gas constant, and T is the absolute temperature:

$$Y_0 = \frac{\left[\begin{array}{l}(Y_N + m_N[D]) + (Y_U + m_U[D]) \\ \quad \times \exp(-(\Delta G_{N-U} + m_G[D])/RT)\end{array}\right]}{1 + \exp(-(\Delta G_{N-U} + m_G[D])/RT)} \quad (2)$$

CD

The proteins were analyzed on a Jasco J600 spectropolarimeter in deionized H_2O using a 0.2-mm cell. The collected data was averaged over 30 acquisitions with a 0.2-nm wavelength increment and a 2-nm bandwidth.

Acknowledgments

We thank Dr. C. Nick Pace for the generous use of his spectrofluorometer and many helpful discussions. We also thank Dr. Thomas O. Baldwin for the use of his spectropolarimeter. This work was supported in part by the NIH (GM49706).

References

Bai Y, Sosnick T, Mayne L, Englander SW. 1995. Protein folding intermediates: Native-state hydrogen exchange. *Science 269*:192–197.

Buchwalder A, Szadkowski H, Kirschner K. 1992. A fully active variant of dihydrofolate reductase with a circularly permuted sequence. *Biochemistry 31*:1621–1630.

Cleland WW. 1967. The statistical analysis of enzyme kinetic data. *Adv Enzymol Related Areas Mol Biol 29*:1–32.

Goldenberg DP, Creighton TE. 1983. Circular and circularly permuted forms of bovine pancreatic trypsin inhibitor *J Mol Biol 165*:407–413.

Harlow E, Lane D. 1988. Immunoblotting. In: *Antibodies: A laboratory manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory. pp 471–510.

Heinemann U, Saenger W. 1982. Specific protein-nucleic acid recognition in ribonuclease T1-2′-guanylic acid complex: An X-ray study. *Nature 299*:27–31.

Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR. 1989. Site directed mutagenesis by overlap extension using the polymerase chain reaction. *Gene 77*:51–59.

Horlick RA, George HJ, Cooke GM, Tritch RJ, Newton RC, Dwivedi A, Lische M, Salemme FR, Weber PC, Horuk R. 1992. Permuteins of interleukin 1β-A simplified approach for the construction of permuted protein having new termini. *Protein Eng 5*:427–431.

Hu CQ, Sturtevant JM, Erickson RE, Pace CN. 1992. Thermodynamics of ribonuclease T1 denaturation. *Biochemistry 31*:4876–4882.

Jennings PA, Wright PE. 1993. Formation of a molten globule intermediate early in the kinetic folding pathway of apomyoglobin. *Science 262*: 892–896.

Kreitman R, Puri RK, Pastan I. 1994. A circularly permuted recombinant interleukin-4 toxin with increased activity. *Proc Natl Acad Sci USA 91*:6889–6893.

Levitt M. 1978. Conformational preferences of amino acids in globular proteins. *Biochemistry 17*:4277–4285.

Luger K, Hommel U, Herold M, Hofsteenge J, Kirschner K. 1989. Correct folding of circularly permuted variants of a $\beta\alpha$ barrel enzyme in vivo. *Science 243*:206–210.

Malin R, Zielenkiewicz P, Saenger W. 1991. Structurally conserved water molecules in ribonuclease T1. *J Biol Chem 266*:4848–4852.

Martinez-Oyanedel J, Choe HW, Heinemann U, Saenger W. 1991. Ribonuclease T1 with free recognition and catalytic site: Crystal structure analysis at 1.5 Å resolution. *J Mol Biol 222*:335–352.

Mayr LM, Schmid FX. 1993. A purification method for labile variants of ribonuclease T1. *Protein Purif Expr 4*:52–58.

Mullins LS, Pace CN, Raushel FM. 1993. Investigation of ribonuclease T1 folding intermediates by hydrogen-deuterium amide exchange two-dimensional NMR spectroscopy. *Biochemistry 32*:6152–6156.

Mullins LS, Wesseling K, Kuo JM, Garrett JB, Raushel FM. 1994. Transposition of protein sequences: Circular permutation of ribonuclease T1. *J Am Chem Soc 116*:5529–5533.

Oshima T, Uenishi N, Imahori K. 1976. Simple assay methods for ribonuclease T1, T2 and nuclease P1. *Anal Biochem 71*:632–634.

Pace CN, Creighton TE. 1986. The disulphide folding pathway of ribonuclease T1. *J Mol Biol 188*:11820–11825.

Pace CN, Grimsley GR, Thompson JA, Barnet BJ. 1988. Conformational stability and activity of ribonuclease T1 with zero, one, and two intact disulfide bonds. *J Biol Chem 263*:11820–11825.

Pace CN, Heinemann U, Hahn U, Saenger W. 1991. Ribonuclease T1: Structure, function, and stability. *Angewandte Chemie 30*:343–454.

Pletinckx J, Steyaert J, Zegers I, Choe HW, Heinemann U, Wyns L. 1994.

Crystallographic study of Glu 58 Ala-RNase T1 2′-guanosine monophosphate at 1.9-Å resolution. *Biochemistry 33*:1654–1662.

Santoro MM, Bolen DW. 1988. Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl α-chymotrypsin using different denaturants. *Biochemistry 27*: 8063–8068.

Shirley BA, Laurents DV. 1990. Purification of recombinant ribonuclease T1 expressed in *Escherichia coli*. *J Biochem Biophys Methods 20*:181–188.

Steyaert J, Haikal AF, Wyns L, Stanssens P. 1991a. Subsite interactions of ribonuclease T1: Asn 36 and Asn 98 accelerate GpN transesterification through interactions with the leaving nucleoside N. *Biochemistry 30*: 8666–8670.

Steyaert J, Hallenga K, Wyns L, Stanssens P. 1990. Histidine-40 of ribonuclease T1 acts as base catalyst when the true catalytic base, glutamic acid-58, is replaced by alanine. *Biochemistry 29*:9064–9072.

Steyaert J, Opsomer C, Wyns L, Stanssens P. 1991b. Quantitative analysis of the contribution of Glu 46 and Asn 98 to the guanosine specificity of ribonuclease T1. *Biochemistry 30*:494–499.

Towbin H, Staehelin T, Gordon J. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: Procedure and some applications. *Proc Natl Acad Sci USA 76*:4350–4354.

Vignais ML, Corbier C, Mulliert G, Branlant C, Branlant G. 1995. Circular permutation within the coenzyme binding domain of the tetrameric glyceraldehyde-3-phosphate dehydrogenase from *Bacillus stearothermophilus*. *Protein Sci 4*:994–1000.

Viguera AR, Blanco FJ, Serrano L. 1995. The order of secondary structure elements does not determine the structure of a protein but does affect its folding kinetics. *J Mol Biol 247*:670–681.

Yang YR, Schachman HK. 1993. Aspartate transcarbamoylase containing circularly permuted catalytic polypeptide chains. *Proc Natl Acad Sci USA 90*:11980–11984.

Zabinski M, Walz FG. 1976. Subsites and catalytic mechanism of ribonuclease T1: Kinetic studies using GpC and GpU as substrates. *Arch Biochem Biophys 175*:558–561.

Zell R, Fritz HJ. 1987. DNA mismatch-repair in *Escherichia coli* counteracting the hydrolytic deamination of 5-methyl-cytosine residues. *EMBO J 6*:1809–1815.

Zhang T, Bertelsen E, Benvegnu D, Alber T. 1993. Circular permutization of T4 lysozyme. *Biochemistry 32*:12311–12318.