Conformation in Solution of the Fully Toxic Domain of Heat-Stable Enterotoxin (ST_p) Produced by Enterotoxigenic *Escherichia coli*[1)]

Hiroshi Ozaki,* Hirokazu Kubota, Takashi Sato, Yuji Hidaka, Haruhiko Tamaoki, Yuji Kobayashi, Yoshimasa Kyogoku, Takashi Sugimura,† Akira Tai,† and Yasutsugu Shimonishi
Institute for Protein Research, Osaka University, Yamada-oka 3-2, Suita, Osaka 565
†Faculty of Science, Himeji Institute of Technology, 2167 Shosha, Himeji, Hyogo 671-22
(Received November 27, 1990)

Heat-stable enterotoxin (ST_p) produced by a porcine strain of enterotoxigenic *Escherichia coli* is a peptide of 18 amino acid residues. The full toxicity is generated by a core peptide of 13 amino acids from residue 5 to 17 ($ST_p(5—17)$). Detailed conformational analysis of $ST_p(5—17)$ was performed by NMR spectroscopy and distance geometry calculations. The tertiary structure of $ST_p(5—17)$ was found to have a right-handed spiral fold throughout the whole molecule which was stabilized by the network of disulfide linkages and hydrogen bonds. This folding pattern was demonstrated in both organic and aqueous solutions.

Enterotoxigenic *E. coli* elaborates two kinds of heat-stable enterotoxins (named ST_p and ST_h) with 18 and 19 amino acid residues respectively, which are responsible for acute diarrhea in infants and domestic animals.[2-5)] The ST acts on ST-sensitive cells in the intestine to stimulate guanylate cyclase and to increase the level of cyclic GMP in the cells, resulting in fluid secretion from the intestinal epithelial cells.[6-9)] Recently, a selective and high-affinity binding protein(s) for ST on rat intestinal cell membranes has been identified and membrane bound guanylate cyclase has been suggested to be bound to, or present near this protein(s).[10-12)] The initial step in the biological reaction of ST is the formation of an ST-protein(s) complex. However, details of this step are not yet fully understood. For understanding the molecular basis of the interaction of ST with its receptor protein(s) and the role of amino acid residues of ST in generation of its toxicity, an important problem is to elucidate the molecular conformation of ST and to identify the part(s) of the ST molecule recognized by its receptor protein(s).

Our recent studies demonstrated that the full enterotoxigenic activities of ST_p and ST_h are generated by toxic domains (abbreviated as $ST_p(5—17)$ and $ST_h(6—18)$, respectively) consisting of 13 amino acid residues (boxed by a broken line in Fig. 1).[1)] These findings imply that this domain includes the site(s) for recognition of ST by its receptor protein(s) on the intestinal cell membrane.[13,14)] The conformation of the toxic domain of ST_h in solution was initially studied by us[15)] and by Gariepy et al.[16)] by ^1H NMR spectroscopy and distance geometry calculations. Both studies provided useful information on the conformation of the peptide backbone of the toxic domain of ST_h. However, the structure of the whole molecule could not be refined, because an efficient nuclear Overhauser effect (NOE) could not be observed and the exact disulfide pairings in ST were not known at that time. Thus, the conformation of ST in solution is still uncertain.

In this work, we investigated the conformation in solution of $ST_p(5—17)$, the toxic domain of ST_p, which differs from $ST_h(6—18)$ in only one amino acid residue, by ^1H NMR spectroscopy and distance geometry algorithm analysis.[15,17,18)] $ST_p(5—17)$ rather than $ST_h(6—18)$ was chosen for confomational study of ST, because ST_p is found more often than ST_h. In an organic solvent, $ST_p(5—17)$ was shown to have a spatial structure with a right-handed spiral stabilized by three intramolecular disulfide linkages and hydrogen bondings, and its structure in an aqueous solution

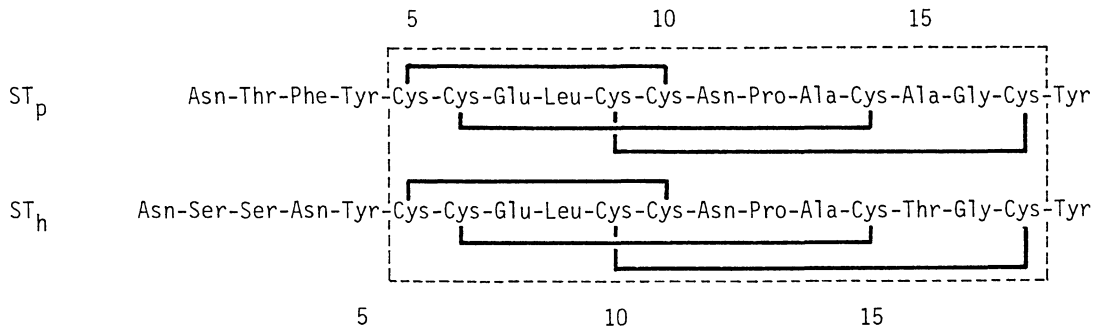

Fig. 1. Amino acid sequences and disulfide linkages of heat-stable enterotoxins (ST_p[5)] and ST_h[4,14)]) of enterotoxigenic *E. coli*.

was suggested to be almost the same. The conformation of $ST_p(5—17)$ elucidated here should be useful for understanding the molecular interaction of ST and its binding protein(s), which proceeds in an aqueous environment, and for examining the conformation-function relationship of the peptide.

Results and Discussion

Resonance Assignment: The ^1H NMR spectrum of

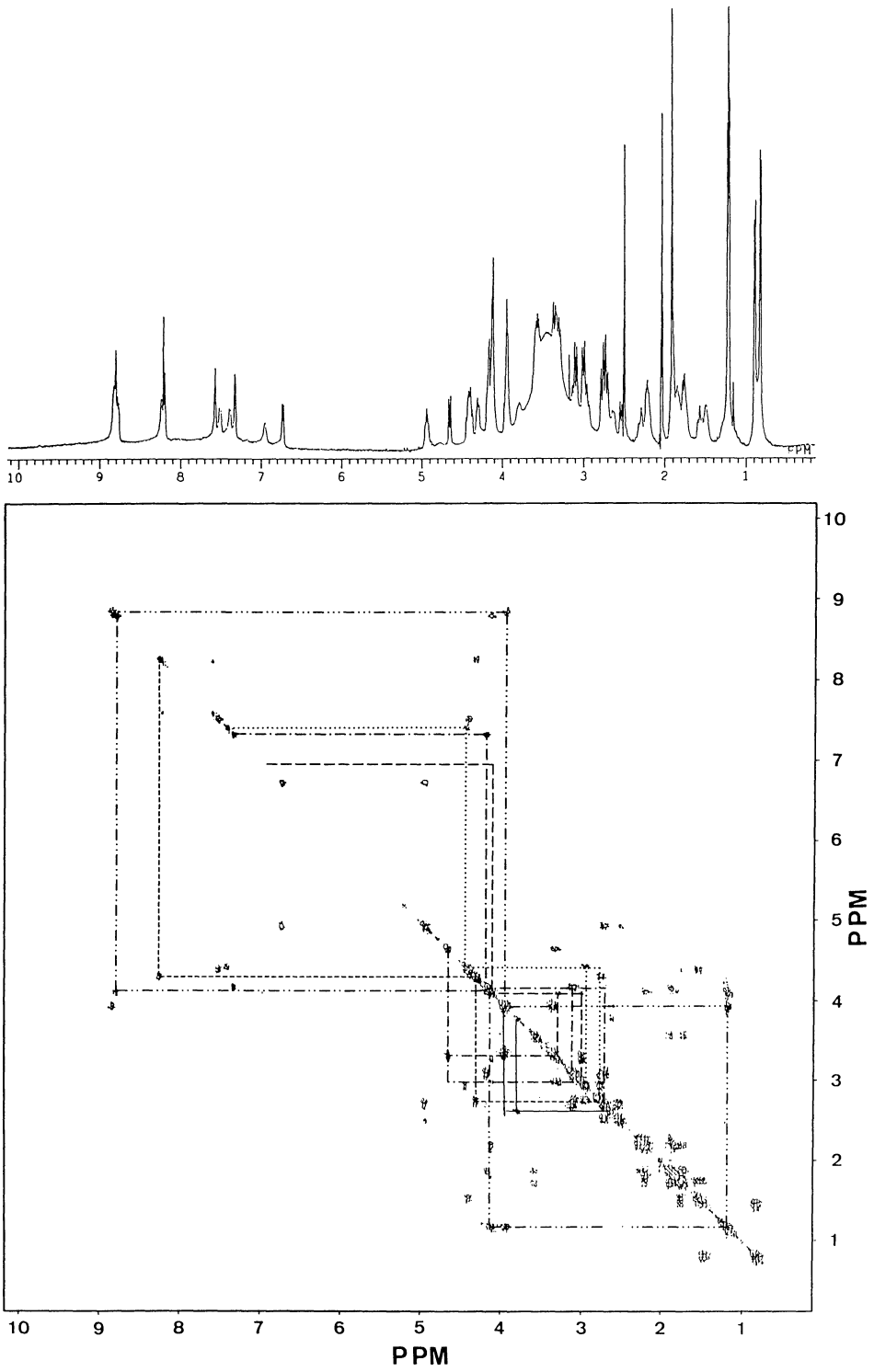

Fig. 2. DQF-COSY spectrum of ST_p (5—17) at 10 °C in DMSO-d_6/CH_3CN-d_3 (v/v, 85/15). Spin systems for six Cys and two Ala residues are indicated: Cys^5 (——), Cys^6 (– · – –), Cys^9 (------), Cys^{10} (— —), Cys^{14} (— · —), Cys^{17} (·····), Ala^{13} (— · · —), and Ala^{15} (— ··· —).

ST_p(5—17) was measured in a mixture of DMSO-d_6 and CH_3CN-d_3, as used in ^1H NMR spectroscopy of ST_h(6—19).[15] Peak and sequential assignment were achieved by analyzing spin systems of constituent amino acid residues. These were examined by through-bond connectivities in the DQF-COSY spectrum of ST_p(5—17) and sequential NOE connectivities between the NH proton of an amino acid residue (i+1) and the protons of the $C_\alpha H$, NH and $C_\beta H$ of the neighboring residue (i) according to the procedures described in Refs. 19 and 20, respectively.

Figures 2 and 3 show the DQF-COSY spectrum of the peptide and its NOESY spectrum in the $C_\alpha H$-NH cross peak region, respectively. The chemical shifts of all the constituent amino acid residues of ST_p(5—17) were assigned except for those of Cys^5 and Cys^6. These protons were assigned by comparison of the 1D spectrum of ST_p(5—17) with that of [β,β-d_2-Cys^5]ST_p(5—17), in which two protons on the C_β of Cys^5 were replaced by deuterium, as shown in Fig. 4. However, no sequential connectivity was detected between Cys^5 and Cys^6 or between Cys^6 and Glu^7 in the N-terminal segment Cys^5-Cys^6-Glu^7, because the NH protons were exchanged quickly with protons from a trace of water contaminating the solvent and so the NH protons of Cys^5 and Cys^6 could not be observed in the spectrum and the peak of the NH proton of Glu^7 became broad. The connectivities in the sequence from Asn^{11} to Ala^{13} were confirmed by through space NOEs between the $C_\alpha H$ of Asn^{11} and the $C_\delta H$ of Pro^{12} and between the $C_\delta H$ and $C_\gamma H$ of Pro^{12} and the NH of Ala^{13}, although Asn^{11} could not be connected to Pro^{12} by sequential NOE connectivities. The presence of the former NOE indicated that the peptide bond between Asn^{11} and Pro^{12} had the trans configuration, and this was confirmed by construction of a molecular model. Table 1 summarizes the resonance assignment of ST_p(5—17).

The resonance assignments of the four Cys residues at positions of 5, 6, 9, and 10 in ST_p(5—17) (corresponding to 6, 7, 10, and 11 in ST_h) were essential to provide the distance information in the N-terminal region, as discussed later, because all long-range NOEs observed in the region from Cys^5 to Cys^{10} were derived from the protons of these four Cys residues, as summarized in Table 2. In a previous study of ST_h(6—19),[16] the spin systems of these four Cys were not assigned, because no sequential NOEs in the sequence containing these amino acid residues were detected as the peptide was examined in aqueous

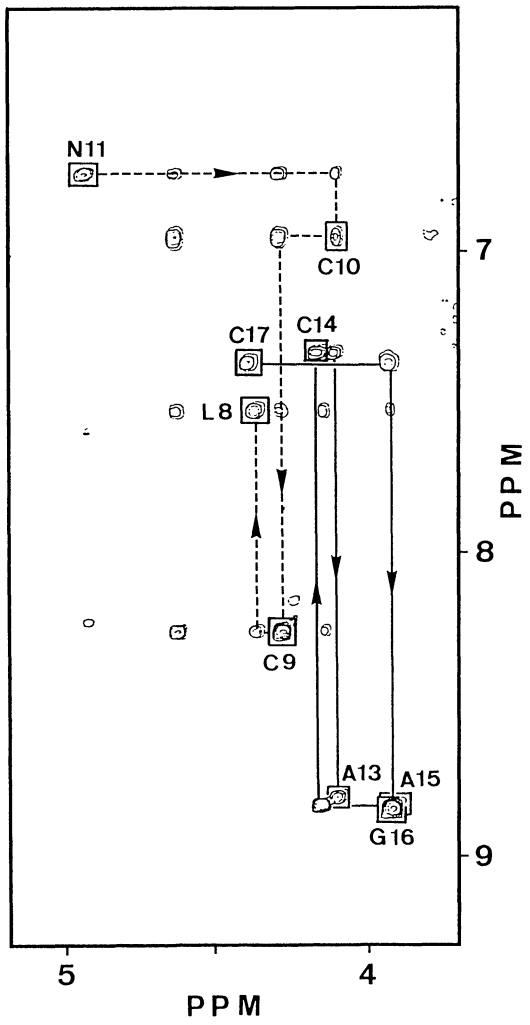

Fig. 3. NOESY spectrum of ST_p (5—17) in the $C_\alpha H$-NH NOE cross peak region. Sequential assignments of segments 8—11 (----) and 13—17 (——) are indicated by arrows according to the direction of the assignment. COSY cross peak positions for each residue are indicated by boxes.

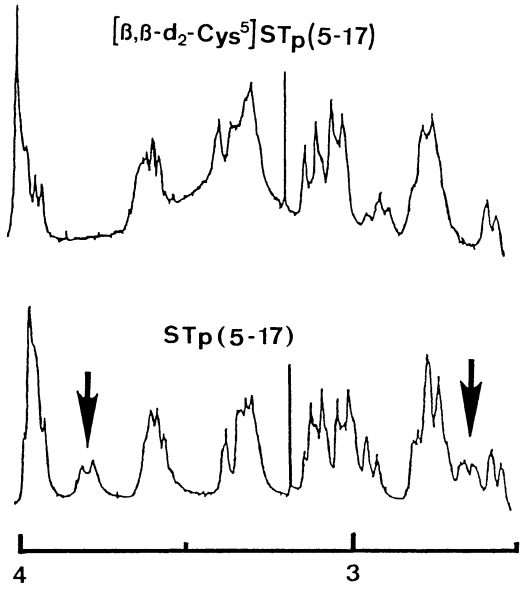

Fig. 4. 1D ^1H NMR spectra of ST_p (5—17) and [β,β-d_2-Cys^5]ST_p (5—17) in DMSO-d_6/CH_3CN-d_3 (v/v, 85/15). The peaks of $C_\beta H$ protons of Cys^5 are indicated by arrows.

Table 1. Resonance Assignment of ST_p (5-17) in DMSO-d_6/CH_3CN-d_3 (85/15)[a)]

Residue	Chemical shift/ppm					
	NH	$C_\alpha H$	$C_\beta H$	$C_\gamma H$	$C_\delta H$	β-$CONH_2$
Cys^5	—	3.95	3.80, 2.65			
Cys^6	—	4.65	3.01, 3.34			
		(4.88)	(3.06, 3.58)			
Glu^7	10.33	4.17	1.91—1.94	2.30		
	(9.66)	(4.35)	(2.16)	(2.46, 2.67)		
Leu^8	7.51	4.39	1.58, 1.78	1.48	0.82, 0.90	
	(7.38)	(4.69)	(1.59, 1.86)	(1.51)	(0.89, 0.94)	
Cys^9	8.22	4.31	2.77, 3.09			
	(8.35)	(4.41)	(2.91, 3.27)			
Cys^{10}	6.95	4.11	3.00, 3.30			
		(4.50)	(3.32, 3.51)			
Asn^{11}	6.73	4.94	2.55, 2.73			7.58, 8.20
	(7.12)	(5.16)	(2.79, 2.87)			(7.12, 7.90)
Pro^{12}		4.13	1.85, 2.22	1.74, 1.91	3.59	
		(4.34)	(1.99, 2.35)	(2.00)	(3.75)	
Ala^{13}	8.78	4.13	1.21			
	(8.49)	(4.32)	(1.37)			
Cys^{14}	7.32	4.17	2.70, 3.10			
	(7.70)	(4.37)	(3.14—3.16)			
Ala^{15}	8.81	3.94	1.20			
	(8.71)	(4.21)	(1.41)			
Gly^{16}	8.84	3.38, 3.96				
	(8.87)	(3.80, 4.11)				
Cys^{17}	7.34	4.42	2.77, 2.94			
	(7.63)	(4.81)	(2.98, 3.14)			

a) Values in parentheses indicate chemical shifts (ppm) in H_2O/D_2O (80/20) at 17°C and pH 5.6. Chemical shifts are shown with reference to internal tetramethylsilane in DMSO-d_6/CH_3CN-d_3 (85/15) and sodium 3-(trimethylsilyl)-1-propanesulfonate in H_2O/D_2O (8/2).

Table 2. Input Distance Constraints (in Å) for the Distance Geometry Calculations on ST_p (5—17)

Proton pair[1)]				Upper limit of distance[2)]/Å		Proton pair[1)]				Upper limit of distance[2)]/Å	
				in DMSO[3)]	in H_2O[4)]					in DMSO[3)]	in H_2O[4)]
Cys^5	$C_\alpha H$	Cys^{10}	$C_\beta P$	5.0[a]			NH	Asn^{11}	NH	3.5	
Cys^6	$C_\alpha H$	Leu^8	NH	3.5			NH	Cys^{10}	$C_\beta P$	4.1[b]	
	$C_\alpha H$	Cys^9	$C_\alpha H$	3.5	(4.0)	Asn^{11}	$C_\alpha H$	Ala^{13}	NH	4.0	
	$C_\alpha H$	Cys^9	NH	3.5			$C_\alpha H$	Pro^{12}	$C_\delta P$	3.5[a]	4.0[a]
	$C_\alpha H$	Cys^{10}	NH	3.5			NH	Cys^{14}	NH	4.0	
	$C_\alpha H$	Asn^{11}	NH	3.5	(4.0)		$C_\beta P$	Ala^{13}	NH		(5.0[a])
	$C_\alpha H$	Cys^{14}	$C_\beta P$	4.5[a]	(5.0[a])		NH	Asn^{11}	$C_\beta P$	4.1[b]	4.1[b]
Glu^7	$C_\alpha H$	Leu^8	NH	3.5		Pro^{12}	$C_\gamma P$	Ala^{13}	NH	4.5[a]	4.5[a]
	NH	Leu^8	NH	4.0	4.0		$C_\delta P$	Ala^{13}	NH	4.5[a]	4.5[a]
	$C_\beta P$	Leu^8	NH	5.0[a]			$C_\alpha H$	Cys^{17}	$C_\beta P$	5.0[a]	
	NH	Glu^7	$C_\beta P$		4.1[b]	Ala^{13}	$C_\alpha H$	Cys^{14}	NH	3.0	(4.0)
Leu^8	$C_\alpha H$	Leu^8	$C_\delta Q$	4.7[c]	4.7[c]		$C_\beta M$	Cys^{14}	NH	4.5[a]	(5.0[a])
	$C_\alpha H$	Leu^8	$C_\gamma H$	3.5			NH	Cys^{14}	NH	3.0	3.0
	NH	Leu^8	$C_\beta P$	4.1[b]	4.6[b]		NH	Ala^{13}	$C_\beta M$	3.6[b]	3.6[b]
	$C_\alpha H$	Cys^9	NH	3.5	4.0	Cys^{14}	$C_\alpha H$	Ala^{15}	NH	3.0	2.5
	NH	Cys^9	$C_\alpha H$	4.0			$C_\beta P$	Ala^{15}	NH	4.5[a]	4.5[a]
	NH	Cys^9	NH	3.0	3.5		NH	Cys^{14}	$C_\beta P$	3.6[b]	3.1[b]
	NH	Cys^{10}	NH	3.5		Ala^{15}	$C_\alpha H$	Gly^{16}	NH	3.0	2.5
Cys^9	$C_\alpha H$	Asn^{11}	NH	3.5	3.5		$C_\beta M$	Gly^{16}	NH		4.5[a]
	$C_\alpha H$	Cys^{10}	NH	3.5			NH	Ala^{15}	$C_\beta M$	3.6[b]	3.6[b]
	NH	Cys^{10}	NH	3.5			$C_\alpha H$	Cys^{17}	NH		4.0
	NH	Cys^9	$C_\beta P$	4.6[b]		Gly^{16}	$C_\alpha P$	Cys^{17}	NH	4.5[a]	4.5[a]
Cys^{10}	$C_\alpha H$	Asn^{11}	NH	3.5	4.0		NH	Cys^{17}	NH	3.5	3.5
	$C_\beta P$	Asn^{11}	NH	5.0[a]		Cys^{17}	NH	Cys^{17}	$C_\beta P$	4.1[b]	3.6[b]

1) P and M indicate pseudoatoms of methylene (including β-amido) and methyl protons, respectively. The pseudo position for methyl protons of leucine is Q in Ref. 25. 2) Upper limit of interatomic distances containing pseudoatoms include their corrections of a 1.0, b 0.6, and c 1.7 Å. 3) NOE cross peaks were collected from NOESY spectra in DMS0-d_6/CH_3CN-d_3 (85/15) at 10 °C. 4) NOE cross peaks were collected from NOESY spectra in H_2O/D_2O (80/20) at pH 5.6 and 17 °C. All these NOEs were also observed at 5 °C. Values in parentheses indicate upper constraints for NOEs detected only at 5 °C.

conditions. This was probably the reason for lack of the distance information in the N-terminal region, resulting in uncertainty about the conformation of $ST_h(6—19)$.

The ^1H NMR spectrum of $ST_p(5—17)$ was also measured in aqueous solution to obtain information about the conformation of the peptide in an aqueous environment. Chemical shifts were assigned by two procedures: sequential assignment as described above, and assignment of the NH and $C_\alpha H$ protons derived by titration of the peptide in DMSO-d_6/CH_3CN-d_3 with H_2O. Results on resonance assignment are listed in Table 1.

Folding Pattern of $ST_p(5—17)$: The data on interatomic distance constraints obtained from NOEs are summarized in Table 2. The folding pattern of the peptide backbone of $ST_p(5—17)$ was calculated from these data and disulfide pairings by distance geometry algorithm analysis.[15] Thirty calculations were performed using different random initial structures. The value of the target function (T-value)[15,17] should be zero if a calculated structure satisfies all constraints. Therefore, T-values were used for evaluating the calculated structures. In addition, the compatibility of the interatomic distances in the calculated structures with the input data was examined by checking individual NOE distance constraint violations. For non-bonded contact, 0.1 Å was an acceptable allowance within the range of possible contacts between two atoms. Among the 30 calculated structures, 14 exhibited no individual violations or non-bonded contacts of more than 0.1 Å. All these structures had T-values of less than 0.1. The convergency of these structures was estimated by calculating the RMSDs between these structures, because the average RMSD of calculated structures with the same folding pattern fell within 2.8 Å, as seen in Table 3. The averaged RMSD was 1.80 Å for atoms in the peptide backbone. All these structures showed the same folding pattern, as seen in Fig. 5(R).

For conformational analysis by distance geometry calculation, sufficient distance information is required to define the correct folding pattern of the peptide backbone.[21] In previous conformational studies of ST, its unique conformation could not be determined because of lack of sufficiently useful NOE data, as described above.[15,16] The question arose of how much interatomic distance information was necessary for defining the conformation of the ST molecule. Therefore, we examined the relationship between the amount and kind of distance information and the folding patterns of the calculated structures, as summarized in Table 3. Thirty calculations were carried out for all cases and T-values were used to evaluate the folding patterns of calculated structures. When the distance information was only that for disulfide pairings, the structure ($T<0.1$) were roughly classified into five folding patterns by viewing the folding patterns obtained by computer graphics; (1) a right-handed spiral fold (R), (2) a left-handed spiral fold (L), (3) R in the N-terminal region end L in the C-terminal region (RML), (4) L in the N-terminal and R in the C-terminal (LMR). and (5) a fold in which the N-terminal portion passes through the loop formed in the central region (C). These folding patterns are drawn as ribbon models in Fig. 5. Examination of the folding pattern of the peptide using intra-residual and sequentially connected NOEs and disulfide pairings excluded two folding patterns (LMR and C) and revealed three folding patterns (R, L and RML). By

Table 3. Amount of Distance Constraints Used and Fold Patterns of Calculated Structures

Constraints used	Number of conformers (x/n)	Total RMSD Å	Type of fold pattern[a]	RMSD in fold type Å
S-S	24/30	3.63	R (8)	2.10
			L (5)	2.40
			RML (3)	1.81
			LMR (3)	2.80
			C (5)	1.87
S-S + Short-range NOEs	9/30	3.50	R (3)	2.20
			L (4)	2.02
			RML (2)	1.90
S-S + All NOEs	14/30	1.80	R (14)	1.80
All NOEs	12/30	1.62	R (12)	1.62

a) R, right-handed spiral fold pattern; L, left-handed spiral fold pattern; RML, R in the N-terminal region and L in the C-terminal region; LMR, L in the N-terminal region and R in the C-terminal region; C, the peptide backbone passes through the loop formed by disulfide linkages. Values in parentheses indicate numbers of conformers to the respective folding patterns.

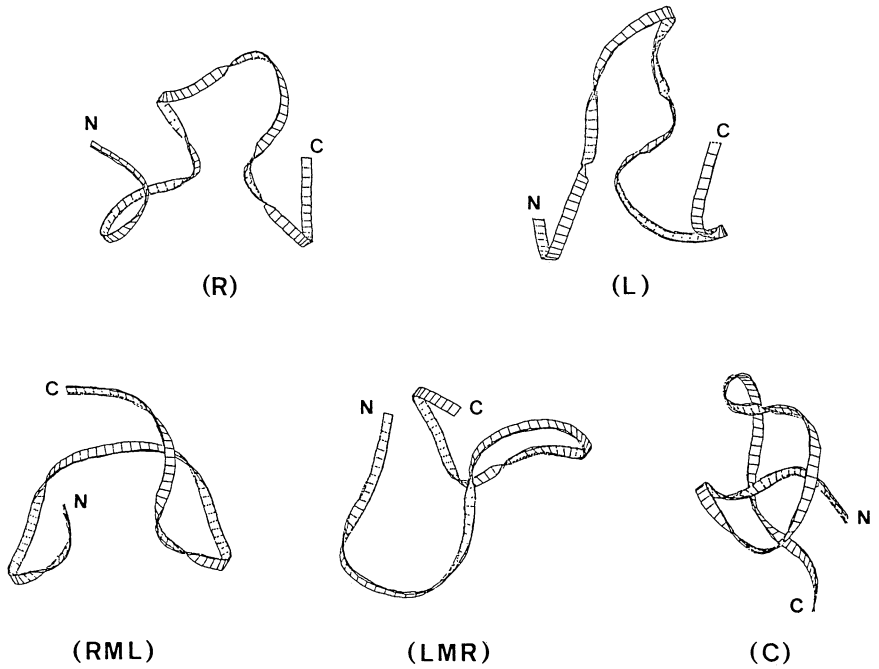

Fig. 5. Typical folding patterns of R, L, RML, LMR and C in Table 3. When the amounts of distance information were not sufficient, the single fold pattern could not be defined. The fold patterns of the peptide backbone are shown as ribbon models.

use of the distance information of all NOEs, all structures ($T<0.1$) gave the R-folding pattern. These results indicated that the distance information of NOEs shown in Table 2, especially the inter-residual NOE set, was indispensable for defining the folding pattern of $ST_p(5—17)$.

During the process of these calculations the possibility was suggested that disulfide pairings can be determined using only the distance information of NOEs, even if the positions of disulfide linkages are not known. Since $ST_p(5—17)$ contains six Cys residues, there are 15 possible combinations of disulfide linkages. Of these combinations, the disulfide pairings that satisfied the constraints of NOEs were searched; that is, the correct disulfide pairings should not affect the violation of NOE constraints and the structure of the peptide with $T<0.1$ should be obtained in the calculation using the distance information of NOE and a given combination of disulfide linkages as well as only the distance information of NOEs. Ten calculations were carried out using random initial structures for each combination of disulfide linkages. In each case, the structure with $T<0.1$ was obtained with only one combination of disulfide linkages, which was the same as that determined chemically.[14] This result demonstrates that with a sufficient amount of distance information, not only the unique folding pattern of $ST_p(5—17)$ but also its disulfide pairings can be determined. These processes may be useful not only for calculation of the folding pattern of ST, but also for determination of disulfide linkages in natural ST and synthetic ST analogs in which the disulfide pairings are unknown, if a NOE set such as that in Table 2, and especially inter-residual NOEs, can be obtained.

The distance constraints from NOEs of $ST_p(5—17)$ in aqueous solution are also summarized in Table 2. Almost all NOEs in an aqueous solvent were the same as those in an organic solvent, although only about 40% of the NOEs of the N-terminal region of the peptide observed in an organic solvent were detected in an aqueous solvent. From twenty calculations using constraints of disulfide linkages and NOEs in an aqueous solvent (Table 2), eight structures ($T<0.1$) were obtained and they all exhibited the R-folding pattern (not shown). Their average RMSD was 1.98 Å for backbone atoms. This result demonstrated that the folding pattern of $ST_p(5—17)$ was the same in an organic solvent as in an aqueous solvent.

Refinement of Tertiary Structure of $ST_p(5—17)$: From an H-D exchange experiment (not shown) five NH protons of Cys^9, Asn^{11}, Ala^{13}, Cys^{14}, and Cys^{17} were found to show slow H-D exchange, although four amide protons of Asn^{12}, Ala^{14}, Cys^{15}, and Cys^{18} of ST_h (corresponding to Asn^{11}, Ala^{13}, Cys^{14}, and Cys^{17}, respectively, in ST_p) were reported to be involved in hydrogen bondings in Ref. 16. The signals of these NH protons were still observed in the spectrum 6.5 h after addition of 500 μl of D_2O to the peptide solution in 500 μl of DMSO-d_6/CH_3CN-d_3 (v/v, 85/15). The NH proton of Cys^{17} was most labile among these protons. Taking account of the size of $ST_p(5—17)$,

these NH protons with slow H-D exchangeability were assumed to be involved in hydrogen bondings. The most plausible CO groups forming hydrogen bonds with these NH groups were searched according to the following criteria: (1) Structures with $T<0.1$ should be obtained for the correct pairing of a given NH and CO, because additional information on a hydrogen bonding does not increase interatomic distance violations. (2) Hydrogen bonding pairs with an angle ∠N-O-C of 100° to 140° are expected to be present in the tertiary structure of the peptide, because in the molecular structure of an analog of ST_p ([Mpr^5]ST_p(5—17)) the angles ∠N-O-C for all intramolecular hydrogen bonds are in this range.[22] (3) The NH protons may be hydrogen-bonded with the CO oxygens in the same loops, as seen in the folding pattern of ST_p(5—17) in Fig. 5(R). Furthermore, upper and lower constraints of 2.5 Å and 2.0 Å, respectively, were input into the calculation as the distance of the hydrogen bonding between an NH proton and a CO oxygen. Ten calculations were carried out from different random initial structures in all cases using the distance constraints of NOEs that were obtained from the measurement of ST_p(5—17) in an organic solvent (Table 2).

First we examined the hydrogen bonding of the NH proton of Cys^9 with either the CO group of Cys^5 or Cys^6 or Glu^7 in the segment from Cys^5 to Cys^{10} in the first loop from the N-terminus, as depicted in Fig. 5(R). Structures with $T<0.1$ were obtained for the CO groups of both Cys^5 and Cys^6 and the values of ∠N-O-C were 73°±6° and 100°±0.5° for Cys^5 and Cys^6, respectively. According to the second criterion described above, the acceptor of the NH proton of Cys^9 was concluded to be the CO group of Cys^6. Similarly, the acceptors of the NH protons of Cys^{14} and Cys^{17} in the segments from Cys^{10} to Cys^{14} and from Cys^{14} to Cys^{17} in the second and third loop, respectively, in Fig. 5(R) were examined and determined to be the CO of Asn^{11} and the CO of Cys^{14}, respectively. The CPK-model constructed from the calculated structures at this stage suggested that the acceptor in the hydrogen bonding with the NH proton of Asn^{11} was the CO group of Cys^5 or Cys^9. The CO of Cys^5 gave structures with $T<0.1$, while the CO of Cys^9 did not. However, the hydrogen bonding of the NH of Asn^{11} to the CO of Cys^5 was excluded by the second criterion. The NH proton of Asn^{11} might be present close to the CO of Cys^5 in the interior of the molecule. Finally, the acceptor for the NH of Ala^{13} was concluded to be the β-CO of Asn^{11}, because only this CO can be present in the space where the NH of Ala^{13} can interact. This type of hydrogen bonding between the side-chain and the main-chain has been reported.[23] These results (in Table 4) indicate that the hydrogen bondings were between the NH of Cys^9 and the CO of the Cys^6, the NH of Cys^{14} and the CO of Asn^{11}, the NH of Cys^{17} and the CO of Cys^{14}, and the NH of Ala^{13} and the β-CO of Asn^{11}.

Table 4. Pairing of Hydrogen Bonds and Bond Angles (∠N-O-C) in Calculated Structures

	Pairing of hydrogen bonds		Number of structures ($T<0.1$)	r^d (∠N-O-C)
	NH	CO		
1	Cys^9	Cys^5	3	73±6
		Cys^6	2	100±0.5
		Glu^7	0	
2[a]	Cys^{14}	Cys^{10}	0	
		Asn^{11}	2	113±0.5
		Pro^{12}	0	
		Gly^{16}	0	
3[b]	Cys^{17}	Ala^{13}	0	
		Cys^{14}	2	102+0.5
		Ala^{15}	2	83±2
4[c]	Asn$_{11}$	Cys^5	3	164±8
		Cys^9	0	
5[c]	Ala^{13}	Asn^{11}(β-CO)	2	117±1

Information on disulfide bonds and all NOEs were used for all calculations. a, b, and c: additional information on hydrogen bonds was used: a, the NH of Cys^9 and the CO of Cys^6; b, a and the NH of Cys^{14} and the CO of Asn^{11}; c, b and the NH of Cys^{17} and the CO of Cys^{14}. d: Averaged angles of ∠N-O-C under consideration are shown in degrees.

Thirty calculations were carried out using constraints from three disulfide bonds, all NOEs and the four hydrogen bonds determined above. Seven structures ($T<0.1$) were obtained and the average RMSD between them was 1.13 Å for all backbone atoms. Thus, the RMSD was further improved by using additional information of hydrogen bondings. Four structures with similar low T-values were superimposed on the structure with the lowest T-value, as shown in Fig. 6.

The tertiary structure of ST_p(5—17) had a right-handed spiral fold throughout the whole molecule with three loops along this spiral. The first loop (Cys^5-Cys^{10}) was fixed by the disulfide bond between Cys^5 and Cys^{10} and the hydrogen bond between the NH of Cys^9 and the CO of Cys^6. The second loop (Cys^{10}-Cys^{14}) was stabilized by the two disulfide bonds between Cys^5 and Cys^{10} and Cys^6 and Cys^{14} and the two hydrogen bonds between the NH of Cys^{14} and the CO of Asn^{11} and the NH of Ala^{13} and the β-CO of Asn^{11}. The third loop (Cys^{14}-Cys^{17}) was stabilized by the two disulfide bonds between Cys^6 and Cys^{14} and Cys^9 and Cys^{17} and the hydrogen bond between the NH of Cys^{17} and the CO of Cys^{14}. The formation and stabilization of the β-turn in each loop were attributed to the hydrogen bonds between the backbone atoms. Thus, the spatial structure of ST_p(5—17) was strongly stabilized by the network of disulfide and hydrogen bonds.

Comparison of the Conformations of ST_p(5—17) in Solution and the Solid State: Recently we determined the molecular structure of an analog of ST_p(5—

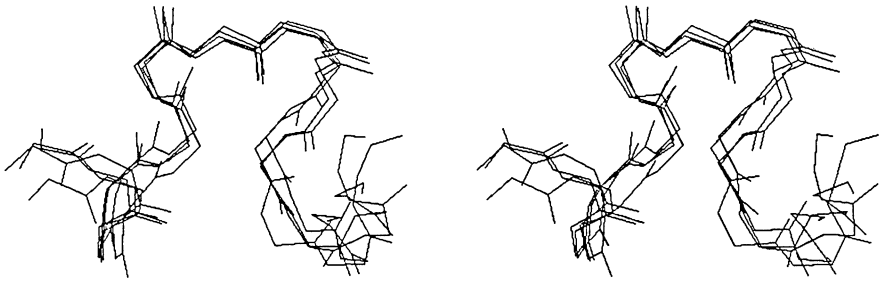

Fig. 6. Stereo view of backbone structures of ST_p (5—17) superimposed to give the minimum averaged RMSD. The all structures exhibited violations of less than 0.1 Å.

17) in which Cys was replaced by Mpr at position 5 ([Mpr^5]ST_p(5—17)).[22] We compared the structure of ST_p(5—17) in solution with that of [Mpr^5]ST_p(5—17) in the solid state, because [Mpr^5]ST_p(5—17) lacks only the N-terminal NH_2-group of ST_p(5—17) and therefore the structure of this Mpr^5-derivative should be very similar to that of ST_p(5—17). The backbone structure of ST_p(5—17) was superimposed on that of [Mpr^5]ST_p(5—17) in the solid state, as shown in Fig. 7. The two structures had the same folding pattern. Four intramolecular hydrogen bonds and three β-turns in the structure of ST_p(5—17) in solution have also been observed in the solid structure of [Mpr^5]ST_p(5—17). The common folding pattern and secondary structure in solution and the solid states might be caused by the tight stabilization by the net work of disulfide and hydrogen bonds. The RMSD values for backbone atoms in the first, second, and third loops were 0.92, 0.45, and 0.76 Å, respectively. The second loop had the smallest RMSD value and fitted extremely well to that in the solid state, implying that this region is scarcely perturbed in solution. The slightly larger deviations of the first and third loops of the structure in solution from those in the solid state might be due to fluctuations in these loops.

In conclusion, the toxic domain of ST_p in solution could be defined by ^{1}H NMR spectroscopy and distance geometry calculations. The conformation of ST_p(5—17) in solution was almost the same as that of [Mpr^5]ST_p(5—17) in the solid state, indicating that the toxic domain of ST_p has a stable conformation with low perturbation in solution. The present finding should be useful for understanding the part(s) of ST_p recognized by its receptor protein(s) and for conformational analyses of analogs of ST_p and its conformation-activity relationship.

Experimental

Boc-amino acids were purchased from the Peptide Institute Inc. (Minoh, Osaka). Deuterated solvents, DMSO-d_6 (99.96%), CH_3CN-d_3 (99.6%), and D_2O (99.95%), were obtained from CEA (Commissariat a L'Energie Atomique, Cedex, France). All other reagents were commercial products and were purified further before use in analytical

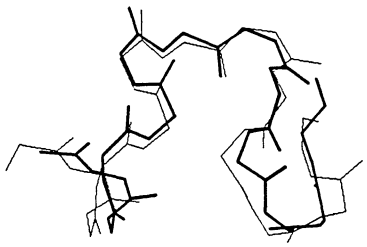

Fig. 7. Comparison of the backbone structure of ST_p (5—17) in solution (——) with that of [Mpr^5]ST_p (5—17) in the solid state (▬).

experiments.

Syntheses of ST_p(5—17) and [β,β-d_2-Cys^5]ST_p(5—17): ST_p(5—17) was synthesized by the solid-phase method.[14] [β,β-d_2-Cys^5]ST_p(5—17) was prepared by the same procedure as ST_p(5—17), except for the use of Boc-β,β-d_2-Cys(MeBzl) at position 5 in the sequence of ST_p instead of Boc-Cys-(MeBzl). The synthesis of Boc-β,β-d_2-Cys(MeBzl) was described previously.[24]

NMR Spectroscopy: 2D ^{1}H NMR spectra were measured in DMSO-d_6/CH_3CN-d_3 (v/v, 85/15) (8 mmol dm^{-3}) at 10 °C and in H_2O/D_2O (2.4 mmol dm^{-3}) at 3—20 °C in a JEOL GX-500 spectrometer. DMSO-d_6/CH_3CN-d_3 was chosen as solvent, because ST_p(5—17) was soluble in DMSO but was gradually destroyed under the conditions for ^{1}H NMR measurement, while ST_p(5—17) was stable in CH_3CN but not soluble enough in this solvent. 2D DQF-COSY and NOESY were performed in the phase-sensitive detection mode by the four-quadrant method.[25] Chemical shifts were measured relative to the methyl resonance of tetramethylsilane added as an internal standard. For NOESY in an organic solvent, a mixing time of 150 ms was adopted, because it was short enough to exclude the effects of spin diffusion and coherent magnetization transfer.[15] NOESY spectra in an aqueous solvent were recorded with mixing times of 100, 150, and 300 ms. From evaluation of secondary effects, NOEs for distance geometry calculation were collected from the NOESY spectra recorded with a mixing time of 150 ms. The sizes of the time domain data were 2K or 4K points for the t_2 direction and 256 points for the t_1 direction. The data size was zero-filled once or twice in the t_1 direction.

The H-D exchange experiment was performed by addition of D_2O to a sample solution in DMSO-d_6/CH_3CN-d_3. The time course of change in 1D spectra of this sample was

recorded in a JEOL GX-400 at 22 °C.

Distance Constraints: The intensities of the NOE peaks were translated into the coresponding interatomic distances based on the relation, NOE intensity $\propto 1/r^{6}$.[18] The aptitude of the interpretation method was discussed in a previous paper.[15,18] The intensity of NOE was evaluated from the number of lines in contour plots of the NOESY spectra of the peptide, and the intensity of the strongest NOE between two C_β protons of Cys was used as a standard intensity for the distance 1.75 Å. The relative intensities with 7—5, 4—3, 2 and 1 contour lines were estimated to be 2.5, 3.0, 3.5, and 4.0 Å, respectively. These values were used as the upper distance constraints. The lower distance constraints were assumed to be the sums of the core radii of the two protons, i.e., 2.0 Å. Information of three disulfide linkages provided an additional distance constraint of 2.04 Å on the S^i-S^j bond (where S^i and S^j are the sulfur atoms of Cys^i and Cys^j residues, which are linked by a disulfide bond) and 3.05 Å on the S^i-$C_\beta{}^j$ and S^j-$C_\beta{}^i$ distances across the disulfide bridge.[26] Since the dihedral angles χ_3 (C_β-S-S-C_β) of disulfide bonds in all proteins observed by high resolution X-ray crystallography are approximately $\pm 90° \pm 10°$,[27] this value of χ_3 was defined by a distance constraint of 3.86 Å on the $C_\beta{}^i$-$C_\beta{}^j$ distance.

Since protons on a methylene or methyl group could not be assigned stereospecifically, these protons were replaced by pseudoatoms as reference points by the method suggested by Wüthrich and his associates.[28]

Distance Geometry Calculations: The distance geometry program DADAS[15,17] was used to construct the tertiary structure of $ST_p(5—17)$. DADAS defines the target function (T) as an error function consisting of the square sum of the differences between atomic distances in calculated structures and the input distance constraints. The initial structure is constructed by generating the random values for dihedral angles. Then T is minimized by the algorism working in the dihedral angle space.[15,17] Calculations were performed with ACOS S2020 and SX-2N computers at the Computer Center of Osaka University. The computed structures were visualized and analyzed using a computer graphics program VENUS at the Research Center for Protein Engineering, Institute for Protein Research, Osaka University.

This work was supported in part by a Grant-in-Aid from the Ministry of Education, Science and Culture. We are indebted to Mrs. Kazuko Kawaguchi for preparation of the manuscript.

References

1) Presented partly in "Peptide Chemistry 1988," (Proceedings of the 26th Symposium on Peptide Chemistry), ed by M. Ueki, pp. 179—182. The abbreviations used in this paper are those recommended by the IUPAC-IUB [*J. Biol. Chem.*, **261**, 1 (1986)]. Additional abbreviations are: *E. coli*, *Escherichia coli*; DMSO, dimethyl sulfoxide; 1D and 2D, one and two dimensional, respectively; DQF, double quantum filtered; COSY, correlated spectroscopy; NOE, nuclear Overhauser enhancement; NOESY, NOE spectroscopy; RMSD, root mean square distance; Mpr, 3-mercaptopropionic acid; MeBzl, 4-methylbenzyl.

2) H. W. Smith and C. L. Gyles, *J. Med. Microbiol.*, **3**, 387 (1970).

3) R. N. Greenberg and R. L. Guerrant, *Pharmacol. Ther.*, **13**, 507 (1981).

4) S. Aimoto, T. Takao, Y. Shimonishi, S. Hara, T. Takeda, Y. Takeda, and T. Miwatani, *Eur. J. Biochem.*, **129**, 257 (1982).

5) T. Takao, T. Hitouji, S. Aimoto, Y. Shimonishi, S. Hara, T. Takeda, Y. Takeda, and T. Miwatani, *FEBS Lett.*, **152**, 1 (1983); M. R. Thompson and R. A. Giannella, *Infect. Immun.*, **47**, 834 (1985).

6) M. Field, L. H. Graf, Jr., W. J. Laird, and P. L. Smith, *Proc. Natl. Acad. Sci. U.S.A.*, **75**, 2800 (1978).

7) J. M. Hughes, F. Murad, B. Chang, and R. L. Guerrant, *Nature*, **271**, 755 (1978).

8) P. M. Newsome, M. N. Burgess, and N. A. Mullan, *Infect. Immun.*, **22**, 290 (1978).

9) M. C. Rao, S. A. Orellana, M. Field, D. C. Robertson, and R. L. Guerrant, *Infect. Immun.*, **33**, 165 (1981).

10) S. A. Waldman, T. Kuno, Y. Kamisaki, L. Y. Chang, J. Gariepy, P. O'Hanley, G. K. Schoolnik, and F. Murad, *Infect. Immun.*, **51**, 320 (1986).

11) J. Gariepy, A. K. Judd, and G. K. Schoolnik, *Proc. Natl. Acad. Sci. U.S.A.*, **84**, 8907 (1987).

12) T. Hirayama, Y. Oku, Y. Takeda, N. Iwata, S. Aimoto, and Y. Shimonishi, "Advances in Cholera and Related Diarrheas," ed by R. B. Sack and Y. Zinnaka, **1990**, Vol. 7. p. 105.

13) S. Yoshimura, H. Ikemura, H. Watanabe, S. Aimoto, Y. Shimonishi, S. Hara, T. Takeda, T. Miwatani, and Y. Takeda, *FEBS Lett.*, **181**, 138 (1985).

14) Y. Shimonishi, Y. Hidaka, M. Koizumi, M. Hane, S. Aimoto, T. Takeda, T. Miwatani, and Y. Takeda, *FEBS Lett.*, **215**, 165 (1987).

15) T. Ohkubo, Y. Kobayashi, Y. Shimonishi, Y. Kyogoku, W. Braun, and N. Gō, *Biopolymers*, **25**, S123 (1986).

16) J. Gariepy, A. Lane, F. Frayman, D. Wilbur, W. Robien, G. K. Schoolnik, and O. Jardetzky, *Biochemistry*, **25**, 7854 (1986).

17) W. Braun and N. Gō, *J. Mol. Biol.*, **186**, 611 (1985).

18) W. Braun, C. Bösch, L. R. Brown, N. Gō, and K. Wüthrich, *Biochim. Biophys. Acta*, **667**, 377 (1981).

19) G. M. Clore and A. M. Gronenborn, *Protein Eng.*, **1**, 275 (1987).

20) M. Billeter, W. Braun, and K. Wüthrich, *J. Mol. Biol.*, **155**, 321 (1982).

21) W. J. Metzler, D. R. Hare, and A. Pardi, *Biochemistry*, **28**, 7045 (1989).

22) H. Ozaki, T. Sato, H. Kubota, Y. Hata, Y. Katsube, and Y. Shimonishi, *J. Biol. Chem.*, in press (1991).

23) M. Marraud, S. Premilat, and A. Aubry, "Peptides," (Proceedings of the 11th American Peptide Symposium), ed by J. E. Rivier and G. R. Marshall, (1990) p. 535.

24) Y. Hidaka and Y. Shimonishi, *Bull. Chem. Soc. Jpn.*, **62**, 1986 (1989).

25) D. J. States, R. A. Haberkorn, and D. J. Reuben, *J. Magn. Reson.*, **48**, 286 (1982).

26) G. Nemethy, M. S. Pottle, and H. A. Scheraga, *J. Phys. Chem.*, **87**, 1883 (1983).

27) J. S. Richardson, *Adv. Protein Chem.*, **34**, 167 (1981).

28) K. Wüthrich, M. Billeter, and W. Braun, *J. Mol. Biol.*, **169**, 949 (1983).