A hairpin-loop conformation in tandem repeat sequence of the ice nucleation protein revealed by NMR spectroscopy

Sakae Tsuda[a,*], Ai Ito[a], Norio Matsushima[b]

[a]*Bioscience and Chemistry Division, Hokkaido National Industrial Research Institute (HNIRI), Toyohira, Sapporo 62, Japan*
[b]*Department of Biophysics, School of Health Sciences, Sapporo Medical University, Sapporo 60,Japan*

Received 11 February 1997; revised version received 8 April 1997

Abstract The ^{1}H-NMR spectrum of a synthetic 24-residue peptide (A^1-G-V-D-S-S-L-I-A-G-Y-G-S-T-Q-T-S-G-S-D-S-A-L-T^{24}; INP24), comprising three repeats of the 8-residue consensus sequence of *Pseudomonas syringae* ice nucleation protein, was fully assigned using 2-dimensional (2D) NMR spectroscopy at 4°C and 30°C. Close proximity of the aliphatic protons between Leu^7, Ile^8, Ala^9, and the ring-protons of Tyr^{11} was indicated from the observation of the inter-molecular nuclear Overhauser enhancement (NOE) effect. Hydrogen-bonding was strongly suggested for the NH group of Leu^7 from its extremely low-temperature coefficient estimated from the temperature dependence of the chemical shift. These results indicate the formation of a hairpin-loop conformation constructed by a hexapeptide segment of INP24, -Leu^7-Ile^8-Ala^9-Gly^{10}-Tyr^{11}-Gly^{12}.

Key words: Ice nucleation protein; *Pseudomonas syringae*; Two-dimensional NMR; Hairpin-loop conformation

1. Introduction

Freezing of water is initiated by formation of an ice crystal by assembly of water molecules onto an embryo nucleation particle [1]. Ice-nucleation protein (INP) located at the outer membrane of Gram-negative epiphytic bacteria is one of the most effective nucleating agents for freezing, which frost tea plants and other crops [2]. INP from *Pseudomonas syringae* is a single polypeptide (residues = 1200; MW = ~120 kDa) comprising three distinct domains: the N-terminal domain (~19 kDa), a central domain (~94 kDa), and the C-terminal domain (~7 kDa) [3]. Among these, the central domain is uniquely composed of about 20 times repeats of a 48-residue high-fidelity consensus sequence. This 48-residue sequence is subdivided into three 16-residue medium-fidelity repeats and is further divided into six 8-residue low-fidelity repeats [1]. It has been thought that this tandemly repetitive amino acid sequence of INP mimics an ice-like surface: a nucleation site for freezing [4–6]. The N- and C-domains contain no such repetitive sequence and are thought to be involved in assembly and/or stabilization of the central domain to the membrane [1].

The structural information elucidated by NMR and X-ray diffraction is indispensable for understanding the structure and function of INP and engineering the molecule with improved activity. However, no direct spectroscopic data of INP has been presented so far, while the 3D-structure models have been proposed [4–6]. These models assume that the tandem repeat domain constructs a non-globular super-secondary structure having many β-turns in high periodicity [1,7]. The proposed models, however, differ in the assignment of the lower-order repetitive sequence conformation, as well as in the spacial packing into the higher-order sequence [4–7]. Thus, it seems that the first priority should be a study of the lower-order tandem repeat sequence of INP.

In the present study we attempted to clarify the conformational features of INP24, a synthetic 24-residue peptide comprising three tandem repeats of the low-fidelity 8-residue consensus sequence (A^1-G-V-D-S-S-L-I-A-G-Y-G-S-T-Q-T-S-G-S-D-S-A-L-T^{24}), which corresponds to the amino acid residues 360–383 (or 456–479) of *Pseudomonas syringae* INP [3]. All ^{1}H-resonances of INP24 at 4°C and 30°C were assigned by 2D-NMR experiments. The NMR conformational parameters of temperature coefficient [8], inter-nuclear Overhauser effect (NOE) [9], and $^3J_{HN-H\alpha}$ coupling constants [10], were used to reveal a formation of a hairpin-loop structure at a local portion in the INP24 polypeptide.

2. Material and methods

The 24-residue repetitive peptide INP24 was synthesized using a Milligen 9050 peptide synthesizer and was further purified by HPLC column chromatography. The INP24 sample was dissolved at a final concentration of 2.5 mM in 500 μl of either 99% D_2O or 90% H_2O (10% D_2O for the lock) containing 25 mM KCl. The pH value of these sample solutions was adjusted to 6.7 (not deuterium corrected). All NMR experiments were carried out on a JEOL JNM-Alpha500 (500 MHz) spectrometer in the temperature range 4–30°C. All ^{1}H-resonances at two different temperatures (4°C and 30°C) were assigned by acquiring the following four sets of 2D-NMR data: (1) DQF-COSY [11]; (2) TOCSY (mixing time = 35–75 ms) [12]; (3) NOESY for 4°C (mixing time = 100–500 ms) [13]; and (4) ROESY for 30°C (mixing time = 100–500 ms) [14]. All of the 2D experiments were acquired in the phase-sensitive mode [15], in 256-512 t_1 increments, with a pre-saturation delay of 1.5–2.0 s, using a DANTE-pulse sequence [16]. The temperature coefficient ($-\Delta d/\Delta T$, ppb K^{-1}) was estimated from the change in NH-resonance chemical shift with temperature (4–30°C). The NMR data processing was done on a SGI Indigo2 work-station using the NMRPipe software [17]. The 2D-NMR data were zero filled to $2K \times 2K$ complex points, and a shifted sine-square window function was applied for resolution enhancement in both ω_1 and ω_2 dimensions. The $^3J_{HN-H\alpha}$ coupling constants were estimated from DQF-COSY (30°C) with resolution enhancement by $16K \times 2K$ zero-filling using a shifted sine-bell window function. Chemical shifts were measured from the internal standard, 2,2-dimethyl-2-silapentane-5-sulfonic acid (DSS).

*Corresponding author. Fax: (81) 11-857-8983

Abbreviations: INP24, a 24-residue peptide of ice-nucleation protein; NOE, nuclear Overhauser enhancement; COSY, chemical shift correlated spectroscopy; NOESY, 2D NOE spectroscopy; ROESY, 2D rotating-frame Overhauser spectroscopy

PII S0014-5793(97)00515-2

3. Results

3.1. Spectral assignment of INP24

The first assignment step was to classify the pattern of TOCSY cross-peaks according to the spin-system of an amino acid with reference to the empirically known chemical shift positions [18]. Fig. 1 shows an expansion of the TOCSY spectrum for INP24 at 4°C. Peak assignments are indicated by lines connecting the cross-peaks with the amino acid residual number of INP24. The cross-peaks of Ile(8), Leu(7, 23), Thr(24, 16, 14), Val(3), and Ala(22, 1, 9) were identified on the basis of their characteristic methyl-containing spin systems. Among these, the strings of cross-peaks of Ala^1 and Ala^9 are overlapped at 8.50 ppm (ω_1) at this temperature, but are separated by 0.04 ppm at 30.0°C (see Table 1). The cross-peaks of the AMX-type amino acids of INP24, Tyr(11) and Asp(4, 20), were identified separately. Only the former is linked with the aromatic δ- and ε-ring proton resonances (6.82 and 7.18 ppm) as determined by NOESY experiments. The Gln(15) residue was identified by the NH-$C_\beta H_2$ and NH-$C_\gamma H_2$ cross-peaks observed at their typical resonance positions (2.02, 2.14, and 2.37 ppm)[18]. Cross-peaks originating from two Gly(10, 18) and two Ser(13, 17) were identified separately, while those from two other Gly(2, 12) were overlapped with the NH-$C_\beta H_2$ cross-peaks of two other Ser(6, 17). This overlapping problem was overcome by comparison with the DQF-COSY spectrum obtained at 4°C (Fig. 2a), which contains no Ser NH-$C_\beta H_2$ cross-peaks. The spectrum allowed us to identify the NH-$C_\alpha H$ cross-peaks of all Gly and Ser without overlapping.

The second assignment step was to link the DQF-COSY (NH, $C_\alpha H$) cross-peak of residue i with that of residue i+1 sequentially along the main-chain of INP24. In this step, $d_\alpha N$-, dNN-, and $d_\beta N$-NOE connectivities [19] between neighboring residues were measured from NOESY (4°C) and ROESY (30°C). A total of 24 cross-peaks originated from the backbone protons of INP24 are observed dispersively in the finger-print region of the DQF-COSY as shown in Fig. 2a (cross-peak assignment indicated). The cross-peak patterns of Gly^{10} and Gly^{12} are different from those of the other two glycines, Gly^2 and Gly^{18}. This is a consequence of the small $^3J_{HNH\alpha}$ coupling constant of these glycines between their NH- and one of the $C_\alpha H$ protons, which is due to a motional restriction in a local conformation. The sequential $d_\alpha N$-NOE connectivities along the main chain of INP24 was indicated in the NOESY spectrum shown in Fig. 2b (the same region as Fig. 2a), in which the assignments of COSY-type (NH, $C_\alpha H$) cross-peaks are indicated by the residual numbers. The vertical lines represent the NOESY $d_\alpha N$ connectivities from residue i to residue i+1. The horizontal lines then locate the COSY-type cross-peak of the residue i+1. Consequently, the assignment spiral for all 24 residues of INP24 is obtained. This data, together with the observation of the sequential dNN- and $d_\beta N$-NOE connectivities, enable assignment of all the ^{1}H-resonance of INP24 unambiguously.

3.2. NMR conformational analysis of INP24

The secondary structure prediction of INP24 by Chou and Fasman's method [20,21] showed that INP24 mostly favors the β-turn and/or β-strand conformation. Conformational information can be examined by the following ^{1}H-NMR parameters:

1. the chemical shift index (CSI) [18,22]
2. the temperature coefficient ($-\Delta d/\Delta T$, ppb K^{-1}) of the NH-resonance [8]
3. the $^3J_{HN-C\alpha H}$ coupling constant [10]
4. the intra- and inter-residue NOE [23,24]

CSI is defined by the difference between the measured chemical shift for $C_\alpha H$ and the chemical shift reported for its random coil position [22]; CSI is strongly dependent on secondary structure (α-helix, β-strand, or coil). The relatively low NH-resonance temperature coefficient (e.g. < 4.0 ppb K^{-1})

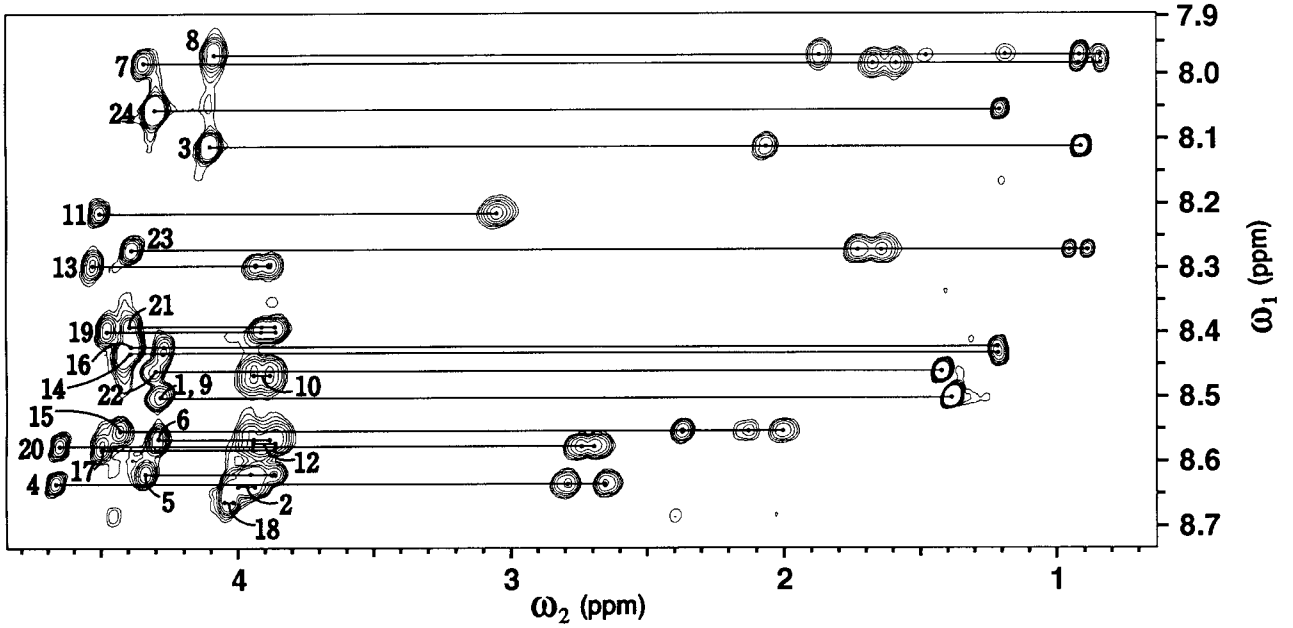

Fig. 1. An expansion of TOCSY (mix = 75 ms) spectrum of INP24 at 4°C (500 MHz). The assignment are indicated by the residual number labeled beside the strings of cross-peaks, which represent the spin system of each amino acid.

Table 1
Conformational parameters ($-\Delta\delta/\Delta T$, $J_{HN-H\alpha}$, CSI) and the ^{1}H-NMR chemical shifts of the 24-residue consensus peptide of *P. syringae* INP

Residue	No.	$-\Delta\delta/\Delta T$	$J_{HN-H\alpha}$	CSI	HN	Hα	Hβ	Hγ	Others
Ala	1	7.6	8.1	0	8.30	4.32	1.39		
					8.50	4.28	1.39		
Gly	2	7.3	-	0	8.44	3.94/3.94			
					8.63	3.96/3.96			
Val	3	6.5	8.4	1	7.94	4.15	2.08		0.92/0.92
					8.11	4.10	2.07		0.92/0.92
Asp	4	7.3	8.4	0	8.44	4.66	2.75/2.68		
					8.63	4.67	2.80/2.65		
Ser	5	8.8	8.1	-1	8.39	4.36	3.95/3.90		
					8.62	4.33	3.92/3.86		
Ser	6	4.2	8.4	-1	8.45	4.36	3.97/3.90		
					8.56	4.28	3.89/3.86		
Leu	7	2.3	6.5	1	7.92	4.35	1.67/1.62		0.94/0.85
					7.98	4.34	1.67/1.59		0.93/0.83
Ile	8	5.7	4.1	1	7.82	4.11	1.87	1.47/1.19	0.90/0.84
					7.97	4.09	1.88	1.48/1.20	0.93/0.83
Ala	9	9.2	7.7	0	8.26	4.28	1.38		
					8.50	4.28	1.39		
Gly	10	8.4	-	0	8.24	3.90/3.90			
					8.46	3.94/3.86			
Tyr	11	5.7	4.1	0	8.06	4.55	3.09/3.02		
					8.21	4.51	3.06/3.06		
Gly	12	6.5	-	0	8.39	3.92/3.92			
					8.56	3.97/3.83			
Ser	13	4.6	8.4	0	8.17	4.55	3.92/3.89		
					8.29	4.53	3.91/3.88		
Thr	14	6.9	8.4	0	8.25	4.40	4.28	1.22	
					8.43	4.40	4.27	1.22	
Gln	15	5.3	8.7	0	8.41	4.45	2.15/2.01	2.38/2.38	
					8.55	4.43	2.14/2.02	2.37/2.37	
Thr	16	6.5	7.7	0	8.25	4.40	4.28	1.22	
					8.42	4.40	4.27	1.22	
Ser	17	6.9	8.7	0	8.40	4.52	3.95/3.90		
					8.58	4.50	3.91/3.88		
Gly	18	6.9	-	0	8.47	4.05/4.05			
					8.65	4.04/4.04			
Ser	19	5.0	8.8	0	8.26	4.50	3.96/3.90		
					8.39	4.48	3.91/3.85		
Asp	20	5.3	8.8	0	8.44	4.66	2.75/2.68		
					8.58	4.66	2.76/2.69		
Ser	21	6.1	7.3	0	8.23	4.42	3.97/3.90		
					8.39	4.39	3.91/3.87		
Ala	22	5.7	8.1	0	8.31	4.33	1.42		
					8.46	4.31	1.42		
Leu	23	6.1	8.8	1	8.11	4.39	1.72/1.66	0.96/0.89	
					8.27	4.39	1.74/1.65	0.96/0.89	
Thr	24	5.3	8.8	0	7.91	4.31	1.21		
					8.05	4.31	1.21		

Two sets of the chemical shifts are given for each amino acid: top line, value at 30°C; bottom line, value at 4°C.

indicates hydrogen bonding. The $^3J_{HN-C\alpha H}$ coupling constant provides the dihedral φ angle information. For example, ^3J is approximately 3.9 Hz for helices (φ = -57°), and 8.9 Hz for antiparallel β-sheets (φ = -139°) [10]. NOE provides proximity information on the protons located within about 5 Å of each other. These criteria were utilized in the present study to examine the conformation of INP24.

Table 1 lists the temperature coefficient ($-\Delta\delta d/\Delta T$), the $^3J_{HN-H\alpha}$ coupling constant, the CSI, and the assignment of the ^{1}H-resonances (30°C and 4°C) of each amino acid residue of INP24. The $^3J_{HN-C\alpha H}$ coupling constants for the amino acid residues from Ser^{19} to Thr^{24} are relatively large (~8.8 Hz), suggesting β-sheet formation. However, CSI values obtained for most of these residues are equal to zero, indicating coil, rather than β-sheet formation. CSI values of +1 or -1 indicate α-helix or β-sheet conformation in the peptide, but no such values were found for any segment of INP24. An extremely low-temperature coefficient was obtained only for Leu^7 (2.3 ppb K^{-1}), which suggests strongly that the NH group of Leu^7 is hydrogen-bonded to the CO group of another residue. This hydrogen bond is presumably involved in INP24's local structure.

Fig. 3 shows an expansion of the NOESY spectrum of INP24 (4°C). The horizontal lines indicate the positions of the δ- and ε-ring proton resonances of Tyr^{11}. The vertical lines indicate the positions of the aliphatic proton resonances of Leu^7, Ile^8, Ala^9, Gly^{10}, and Tyr^{11} as labeled. In Fig. 3, cross-peaks observed at the line intersections represent the intra- and inter-residue NOEs originating from Leu^7, Ile^8, Ala^9, Gly^{10}, and Tyr^{11}. This result implies that the ring-protons of Tyr^{11} ($C_\delta H$, $C_\varepsilon H$) are proximal to the main- and side-chain protons of Leu^7, Ile^8, Ala^9, and Gly^{10}. As for the other residues of INP24, no significant inter-residue NOE was detected.

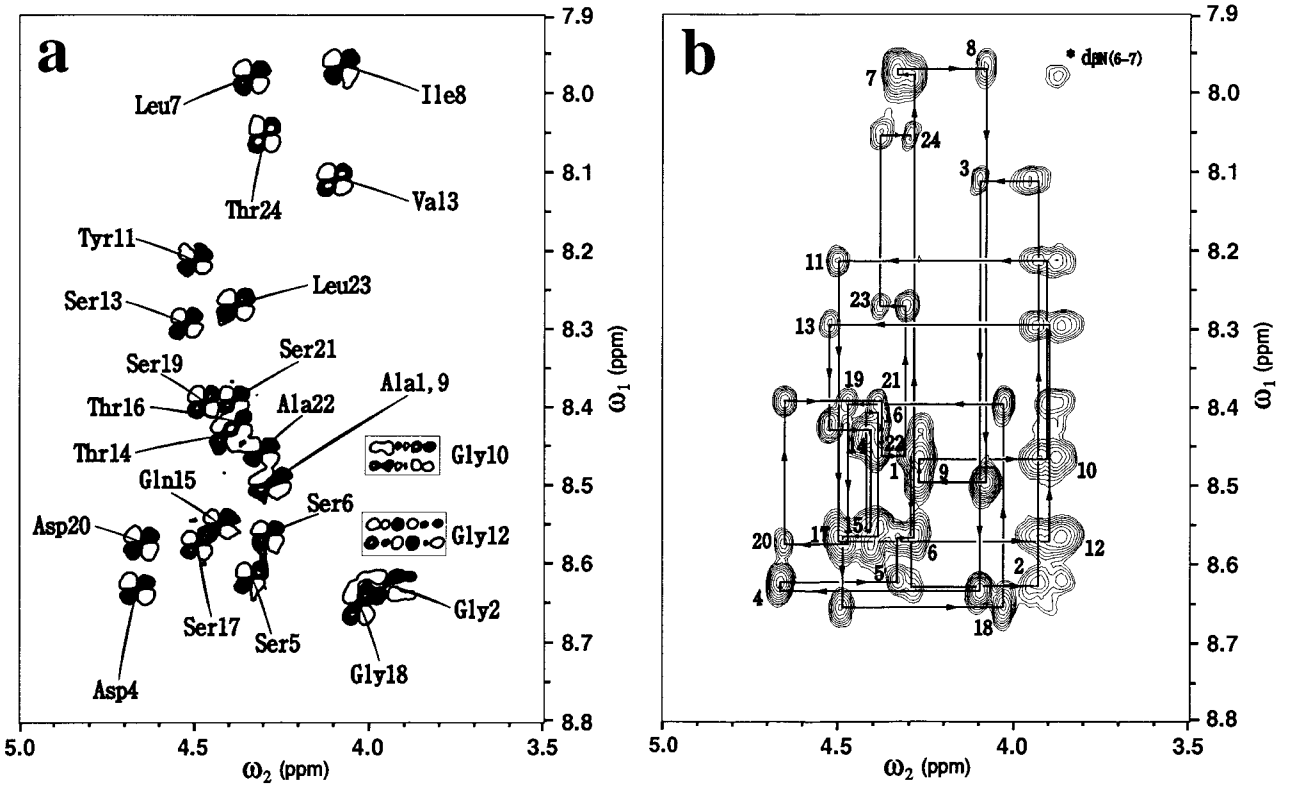

Fig. 2. The finger-print region of DQF-COSY spectrum (a) and the corresponding region of NOESY spectrum (b) of INP24 at 4°C (500 MHz). The full assignment of the (NH, $C_\alpha H$) cross-peaks are indicated in DQF-COSY. In the NOESY vertical lines represent the $d_\alpha N$-NOE connectivities from residue i to residue $i+1$. The horizontal lines then locate the COSY-type (NH, $C_\alpha H$) NOE cross-peak of residue $i+1$. The $d_\beta N$-NOE cross-peak between Ser^6 and Leu^7 is indicated by an asterisk.

4. Discussion

Previous modeling studies assumed the formation of β-turns (I, II, and III) in the tandem repeat sequence of INP [4–6]. These types of β-turns commonly consist of four residues with a hydrogen bond between the NH group of the first residue and the CO group of the fourth residue (denoted 1→4 hydrogen bond) [25,26]. However, the present NMR study shows

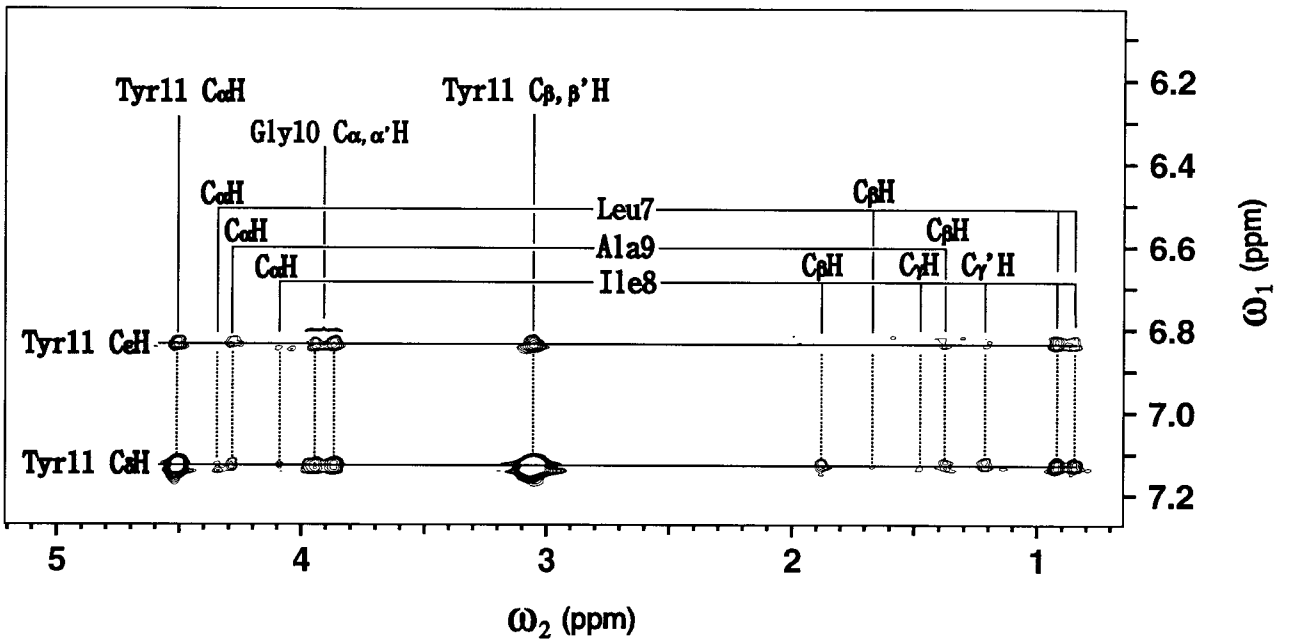

Fig. 3. An expansion of NOESY (mix = 100 ms) spectrum of INP24 at 4°C (500 MHz). The horizontal lines indicate the positions of the δ- and ε-ring proton resonances of Tyr^{11}. The vertical lines indicate the positions of the aliphatic proton resonances of Leu^7, Ile^8, Ala^9, Gly^{10}, and Tyr^{11}. The cross-peaks observed at the line intersections represent the intra- and inter-residue NOE originated from these amino acid residues.

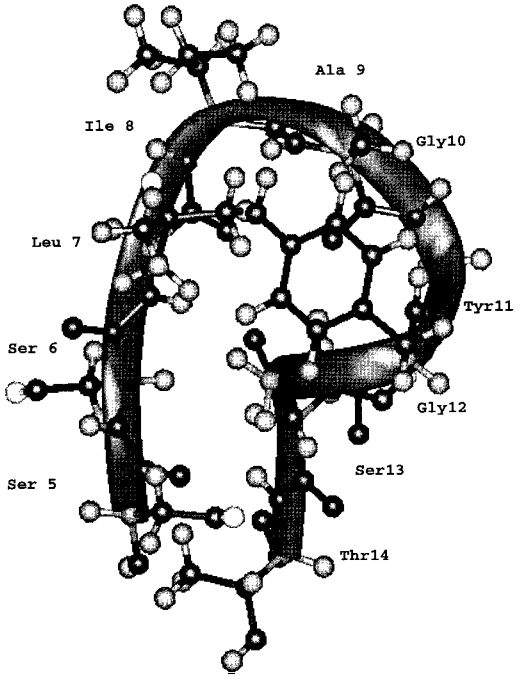

Fig. 4. A possible local conformation of a segment of INP24 containing the hairpin-loop conformation contributed by -Leu^{7}-Ile^{8}-Ala^{9}-Gly^{10}-Tyr^{11}-Gly^{12}-. The superposed solid tube represents the backbone conformation of the peptide.

that INP24 adopts a turn conformation, not the β-turn proposed in the model. Fig. 4 shows a new model for the local structure of INP24 where a turn conformation is formed from a 6-residue segment, -Leu^{7}-Ile^{8}-Ala^{9}-Gly^{10}-Tyr^{11}-Gly^{12}-. In this turn, the hydrophobic side chains of Leu^{7}, Ile^{8}, and Tyr^{11} are close to each other (Fig. 3), and the conformational restrictions of Gly^{10} and Gly^{12} (Fig. 2a) are accounted. One of the important characteristics of this type of turn is the formation of a $1 \rightarrow 6$ hydrogen bond. This is strongly suggested for the NH group of Leu^{7} (Table 1), whose hydrogen-bonding partner is presumably assigned to the CO group of Gly^{12}. This is in contrast with the previous β-turn model presented by Kajava and Lindow [6]. They assumed that Ile^{8} NH is hydrogen-bonded to Tyr^{11} CO. In Fig. 4, the dihedral φ angles of Leu^{7}, Ile^{8}, Ala^{9}, Gly^{10}, Tyr^{11}, and Gly^{12} are -167, -54, -83, -63, -60, and -95°, respectively, which are almost in accordance with the values estimated from the ^{3}J coupling constant (Table 1). This type of turn was named hairpin-loop conformation [26] which has been found in the X-ray structure of α-cobratoxin [27].

In order to predict the overall 3-dimensional (3D) structure of the tandem repeat domain of INP, it is necessary to know the feature of the unit conformation and its periodicity as well. In the proposed 3D-structures by Mizuno [5] and Kajava and Lindow [6], one β-turn-containing conformation was assumed for every octapeptide. As shown in Fig. 4, the hairpin loop appears to require the hydrophobic interactions between the residues, Leu and Tyr, which are located with 16-residues periodicity in the tandem repeat domain of INP. It could be assumed that this hydrophobic region constructed on one side of the loop participates in the self-stacking of the repetitive domain, which leads to the construction of the overall 3D structure. In addition, the conformational bending around the -T-S-G-S- portion which was assumed in the 3D models [5,6], is not observed by the present study and by our preliminary NMR results for another 24-residue peptide of INP (not shown). Therefore, the hairpin loop is thought to be constructed with a 16-residue periodicity if the loop is an essential conformation unit of the repetitive domain of INP.

Acknowledgements: The authors are grateful to Dr. Andy Schmitz for careful reading of the present paper.

References

[1] C.L. Hew, D.S.C. Yang, Eur. J. Biochem. 203 (1992) 33–42.
[2] S.E. Lindow, Annu. Rev. Physiol. 15 (1983) 59–78.
[3] L.V. Corotto, P.K. Wolber, G.J. Warren, EMBO J. 5 (1986) 231–236.
[4] G. Warren, L. Corotto, P. Wolber, Nucl. Acids Res. 14 (1986) 8047–8059.
[5] H. Mizuno, Proteins 5 (1989) 47–65.
[6] A.V. Kajava, S.E. Lindow, J. Mol. Biol. 232 (1993) 709–717.
[7] Kajava, A.V. (1996) in: Biological ice nucleation and its applications (Lee Jr., R.E., Warren, G.J., and Gusta, L.V. eds.), Ch. 6, pp. 101-114, APS Press, Minnesota.
[8] H.J. Dyson, M. Rance, R.A. Houghten, R.A. Lerner, P.E. Write, J. Mol. Biol. 201 (1988) 161–200.
[9] S. Macura, Y. Huang, D. Suter, R.R. Ernst, J. Magn. Reson. 43 (1981) 259–281.
[10] A. Pardi, M. Billeter, K. Wüthrich, J. Mol. Biol. 180 (1984) 741–751.
[11] M. Rance, O.W. Sorensen, G. Bodenhausen, G. Wagner, R.R. Ernst, K. Wüthrich, Biochem. Biophys. Res. Commun. 117 (1983) 479–485.
[12] L. Braunschweiler, R.R. Ernst, J. Magn. Reson. 55 (1983) 301–315.
[13] S. Macura, R.R. Ernst, Mol. Phys. 41 (1980) 95–117.
[14] G. Otting, K. Wüthrich, J. Am. Chem. Soc. 111 (1989) 1871–1875.
[15] D.J. States, R.A. Haberkorn, D.J. Ruben, J. Magn. Reson. 48 (1982) 286–292.
[16] E.R.P. Zuiderweg, K. Hallenga, E.T. Olejniczak, J. Magn. Reson. 70 (1986) 336–343.
[17] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, A. Bax, J. Biomol. NMR 6 (1995) 277–293.
[18] D.S. Wishart, B.D. Sykes, F.M. Richards, J. Mol. Biol. 222 (1992) 311–333.
[19] K. Wüthrich, M. Billeter, W. Braun, J. Mol. Biol. 180 (1984) 715–740.
[20] P.Y. Chou, G.D. Fasman, Adv. Enzymol. 47 (1978) 45–148.
[21] P.Y. Chou, G.D. Fasman, Ann. Rev. Biochem. 47 (1978) 251–276.
[22] D.S. Wishart, B.D. Sykes, F.M. Richards, Biochemistry 31 (1992) 1647–1651.
[23] S.L. Gordon, K. Wüthrich, J. Am. Chem. Soc. 100 (1978) 7094–7096.
[24] G. Wagner, K. Wüthrich, J. Magn. Reson. 33 (1979) 675–680.
[25] C.M. Wilmot, J.M. Thornton, J. Mol. Biol. 203 (1988) 221–232.
[26] A.V. Efimov, Prog. Biophys. Mol. Biol. 60 (1993) 201–239.
[27] M.D. Walkinshaw, W. Saenger, A. Maelicke, Proc. Natl. Acad. Sci. USA 77 (1980) 2400–2404.