Journal of Biomolecular NMR, 5 (1995) 297–305
ESCOM

J-Bio NMR 234

Three-dimensional structure of gurmarin, a sweet taste-suppressing polypeptide

Katsuaki Arai[a], Rieko Ishima[a], Soichi Morikawa[b], Akiko Miyasaka[c], Toshiaki Imoto[c], Shoko Yoshimura[d], Saburo Aimoto[d] and Kazuyuki Akasaka[e,*]

[a]*Department of Chemistry, Faculty of Science, Kyoto University, Kyoto 606, Japan*
[b]*Protein Engineering Research Institute, Suita 565, Japan*
[c]*Faculty of Medicine, Tottori University, Yonago 683, Japan*
[d]*Protein Research Institute, Osaka University, Suita 565, Japan*
[e]*Department of Chemistry, Faculty of Science, Kobe University, Kobe 657, Japan*

Received 2 June 1994
Accepted 16 September 1994

Keywords: Three-dimensional structure; Two-dimensional NMR; Simulated annealing; Sweet taste-suppressing polypeptide; *Gymnema sylvestre*

Summary

The solution structure of gurmarin was studied by two-dimensional proton NMR spectroscopy at 600 MHz. Gurmarin, a 35-amino acid residue polypeptide recently discovered in an Indian-originated tree *Gymnema sylvestre*, selectively suppresses the neural responses of rat to sweet taste stimuli. Sequence-specific resonance assignments were obtained for all backbone protons and for most of the side-chain protons. The three-dimensional solution structure was determined by simulated-annealing calculations on the basis of 135 interproton distance constraints derived from NOEs, six distance constraints for three hydrogen bonds and 16 dihedral angle constraints derived from coupling constants. A total of 10 structures folded into a well-defined structure with a triple-stranded antiparallel β-sheet. The average rmsd values between any two structures were 1.65 ± 0.39 Å for the backbone atoms (N, C^{α}, C) and 2.95 ± 0.27 Å for all heavy atoms. The positions of the three disulfide bridges, which could not be determined chemically, were estimated to be Cys^{3}-Cys^{18}, Cys^{10}-Cys^{23} and Cys^{17}-Cys^{33} on the basis of the NMR distance constraints. This disulfide bridge pattern in gurmarin turned out to be analogous to that in ω-conotoxin and *Momordica charantia* trypsin inhibitor-II, and the topology of folding was the same as that in ω-conotoxin.

Introduction

It is well known that chewing a piece of the leaves of *Gymnema sylvestre*, a folkloric remedy for diabetes mellitus in India, causes complete loss of sweet taste sensation (Edgeworth, 1847–48). The active substance was first extracted as a mixture of acidic compounds, and was named gymnemic acid (Hooper, 1887). The extent of anti-sweet effect of gymnemic acid varies greatly among species, even in mammals (Glaser et al., 1984). For instance, no significant effect was observed in pig, rabbit and rat (Hellekant, 1976; Hellekant and Gopal, 1976).

Recently, a new substance which significantly suppresses the neural responses of rat to sweet taste stimuli was isolated from the hot water extract of the leaves of *Gymnema sylvestre*. This substance was proven to be a polypeptide and was named gurmarin (Imoto et al., 1991). Gurmarin inhibits the sweet taste sensation of rat, but shows no or only a very weak effect on humans. In contrast, gymnemic acid does not affect the sweet taste sensation of rat, but affects that of humans. Gurmarin is the first peptide which shows a strong and specific inhibitory effect on mammalian sweet taste responses.

Gurmarin consists of 35 amino acid residues with three

*To whom correspondence should be addressed.

Abbreviations: DQF-COSY, double-quantum-filtered correlated spectroscopy; HOHAHA, homonuclear Hartmann–Hahn spectroscopy; NOESY, nuclear Overhauser enhancement spectroscopy; ppm, parts per million; rmsd, root-mean-square deviation; TSP, 3-(trimethylsilyl)-2,2,3,3-tetra-deutero-propionate.

0925-2738/$ 6.00 + 1.00 © 1995 ESCOM Science Publishers B.V.

disulfide bridges, and its primary sequence has been determined (Kamei et al., 1992). However, despite considerable efforts, the positions of the disulfide bridges have not been determined chemically, because of the unfortunate situation that gurmarin has a consecutive Cys-Cys sequence (Cys^{17} and Cys^{18}) and is stable against selective cleavage by proteases, even under conditions such as high temperature, low pH and the presence of urea (Kamei, K., private communication). Even so, determination of the three-dimensional structure of this peptide in solution was considered worth trying, in view of its importance in understanding the molecular mechanism of action, hence the present NMR study was undertaken.

First, sequence-specific resonance assignments were obtained for most of the ^1H NMR signals in solution. Then the three-dimensional structure of gurmarin was determined by simulated-annealing calculations (Nilges et al., 1988) on the basis of distance constraints obtained from NMR experiments, without information on the positions of the disulfide bridges. The positions of the three disulfide bridges were estimated from the calculated structures.

Materials and Methods

Sample preparation and NMR measurements

Gurmarin was purified from leaves of *Gymnema sylvestre* (Imoto et al., 1991). Alternatively, gurmarin was chemically synthesized according to the known sequence. Both preparations gave essentially the same NMR patterns. For 2D NMR measurements, the samples were prepared from lyophilized preparations of gurmarin, purified from *Gymnema sylvestre*, and contained 3 mM of gurmarin in either 90% H_2O/10% D_2O or 100% D_2O. The solubility of gurmarin is quite low under slightly acidic conditions, because the isoelectric point of the peptide is 4.5–4.6. The pH (pD) values were adjusted to 2.9 (as direct pH meter reading) by adding little amounts of DCl. The ^1H NMR measurements were made at 20 °C and at 35 °C, with 3-(trimethylsilyl)-2,2,3,3-tetradeuteropropionic acid sodium salt (TSP-d_4) as internal reference for chemical shifts. All 2D spectra were recorded on a Bruker AM600 spectrometer, and processed using the program UXNMR. The water signal was suppressed by continuous selective irradiation. DQF-COSY (Rance et al., 1983), HOHAHA (Bax and Davis, 1985) with a mixing time of 55 ms and NOESY (Jeener et al., 1979; Macura et al., 1981) experiments with mixing times of 50, 100 and 200 ms were performed. All spectra were recorded with a spectral width of 8064.52 Hz in the phase-sensitive mode, using quadrature detection in ω_1 by the time-proportional phase of the first pulse, the carrier frequency being placed at the center of the spectrum. The DQF-COSY spectra were recorded with 512 (ω_1) × 4096 (ω_2) data points and a relaxation delay of 1.6 s, and the HOHAHA and NOESY spectra with 512 (ω_1) × 1024 (ω_2) data points and a relaxation delay of 1.9 s. The coupling constants were measured from the cross sections parallel to ω_2 of the cross peaks in the DQF-COSY spectrum, which were Fourier transformed into 1K (ω_1) × 32K (ω_2) data points after inverse Fourier transformation. Coupling constants $J_{H^\alpha H^\beta}$ and J_{NH^α} were obtained from D_2O and H_2O DQF-COSY spectra, respectively.

Structure calculations

The distance constraints employed in the structure calculations were collected from the NOESY spectra with mixing times of 50 and 100 ms at 35 °C. The cross-peak intensities were classified into three categories with upper bounds of 2.7, 3.5 and 5.0 Å, respectively. For the NOESY spectra in D_2O, geminal β-methylene protons were used as calibration, with a distance of 1.75 Å. For the NOESY spectra in H_2O, the largest $d_{\alpha N}$ cross peak was used as calibration, with a distance of 2.2 Å. Lower bounds of the distance restraints were set to the sum of the van der Waals radii. A distance of 0.5 Å was added to the upper bounds for the distance constraints involving methyl protons, in consideration of the stronger intensity of methyl resonances. Distances of 1.8, 4.8 and 4.8 Å were added to the upper bounds of the distance constraints involving methylene, methyl and aromatic protons, respectively, for which stereospecific assignments could not be made. In addition to the above constraints, distance constraints for hydrogen bonds were also included. In an antiparallel β-sheet, identification of a hydrogen bond between the amide proton of residue i on one strand and the carbonyl oxygen of residue j on the other strand was made from the observation of three NOEs: NH(i)-NH(j), NH(i)-C^αH(j+1) and C^αH(i−1)-C^αH(j+1). The distances of the identified hydrogen bonds were constrained to the ranges $2.8 \leq r_{N(i)-O(j)} \leq 3.3$ Å and $1.8 \leq r_{NH(i)-O(j)} \leq 2.3$ Å. Furthermore, distance constraints from ϕ and χ^1 dihedral angles were also included.

On the basis of interproton distances and torsion-angle restraints, structure calculations were performed with the program EMBOSS, v. 4.0 (Nakai et al., 1993). Calculations were started from random coil conformations as initial structures and the structures were optimized by simulated-annealing calculations (Nilges et al., 1988) and energy minimization. In structure calculations, all protons were treated explicitly, and the parameters of the covalent geometry were those of the AMBER all-atom force field (Weiner et al., 1986). All calculations were performed on an IRIS Indigo workstation (Silicon Graphics).

Results

Assignment of the ^1H NMR signals of gurmarin

The ^1H NMR signals of gurmarin were assigned by standard 2D ^1H NMR spectroscopy techniques. The

sequence-specific assignment of the backbone ^{1}H NMR signals of gurmarin was made by conventional sequential assignment techniques (Wüthrich, 1986). Through-bond connectivities were identified by 2D DQF-COSY and HOHAHA in H_2O and D_2O (Wüthrich, 1986). Complete spin systems were identified for one glycine, two valine, three leucine, two isoleucine, two proline, four AM(PT)X (glutamate and glutamine) and 14 AMX systems, while spin systems were identified incompletely for five lysine, one AM(PT)X and one AMX system.

Figure 1 shows the fingerprint region observed in the DQF-COSY spectrum in H_2O at 35 °C. From the amino acid sequence of gurmarin, which contains two prolines, one glycine and the N-terminal pyroglutamic acid, 34 NH-C$^\alpha$H cross peaks were expected in this region. The DQF-COSY spectrum in H_2O at 35 °C revealed 23 cross peaks. Some of the NH-C$^\alpha$H cross peaks could not be observed, either because of the saturation of C$^\alpha$H resonances or because of rapidly exchanging NH resonances upon irradiation of the H_2O signal. Among these, the NH-C$^\alpha$H cross peaks of Glu^8, Tyr^{14}, Glu^{19}, Glu^{22} and Trp^{29} were observed by DQF-COSY in H_2O at 20 °C. The NH-C$^\alpha$H cross peak of <Glu^1 (pyroglutamic acid) could not be observed, probably due to a small coupling constant. The NH-C$^\alpha$H cross peak of Tyr^{13} was not observed at 20 °C, nor at 35 °C, because of broadening of the peak. The backbone signals of these residues could not be observed in the H_2O DQF-COSY spectrum, but were assigned by HOHAHA and the sequential assignment was made with the NOESY spectra in H_2O.

Figure 2 shows the HOHAHA spectrum in H_2O at 35 °C. Gurmarin contains two valine residues. The spin

TABLE 1
PROTON RESONANCE ASSIGNMENTS OF GURMARIN

Amino acid	Chemical shift (ppm)			
	HN	H$^\alpha$	H$^\beta$	Others
<Glu^1	7.80[a]	4.37	2.51, 2.06	H$^\gamma$ 2.42, 2.34
Gln^2	8.51	4.43	2.12, 2.00	H$^\gamma$ 2.40, –; NH_2 7.51[a], 6.81[a]
Cys^3	8.07	4.92	3.14, 3.03	
Val^4	8.76	4.01	1.91	H$^\gamma$ 1.07, 0.88
Lys^5	7.51	4.13	1.99, 1.63	H$^\gamma$ 1.57[a], 1.41; H$^\delta$ 1.70, –; H$^\varepsilon$ 3.05, –; NH_3 7.48
Lys^6	7.71	3.64	1.67, –	H$^\gamma$ 1.40[a], 1.34; H$^\delta$ 1.73, –; H$^\varepsilon$ 2.99, –; NH_3 –
Asp^7	9.70	4.13	3.16[a], 3.11[a]	
Glu^8	7.65	4.62	2.18, 2.10	H$^\gamma$ 2.35, 2.30
Leu^9	8.02	4.77	1.74, 1.59	H$^\gamma$ 1.82; H$^\delta$ 1.04, 0.95
Cys^{10}	7.98	5.02	3.36, 3.23	
Ile^{11}	8.56	4.36	1.70	H$^\gamma$ 1.45, 1.12; H$^{\gamma(CH3)}$ 0.88; H$^\delta$ 0.86
Pro^{12}		3.92	1.92, 1.67	H$^\gamma$ 1.67, 1.34; H$^\delta$ 3.28, 2.91
Tyr^{13}	8.15	4.08	3.11, 3.03	2,6H 6.88; 3,5H 6.78
Tyr^{14}	7.93	4.57	2.91, 2.64	2,6H 7.15; 3,5H 6.86
Leu^{15}	8.52	4.51	1.64, 1.36	H$^\gamma$ 1.42; H$^\delta$ 0.86, 0.81
Asp^{16}	8.39	4.75	2.82, 2.72	
Cys^{17}	8.78	4.91	H$^{\beta2}$ 2.83; H$^{\beta3}$ 2.41	
Cys^{18}	9.75	4.58	H$^{\beta3}$ 3.28; H$^{\beta2}$ 2.52	
Glu^{19}	8.87	4.52	2.06[a], 2.01	H$^\gamma$ 2.62, 2.57
Pro^{20}		4.89	2.55, 2.02	H$^\gamma$ 2.01, 1.60; H$^\delta$ 3.56, 3.27
Leu^{21}	8.64	4.17	H$^{\beta2}$ 2.02; H$^{\beta3}$ 1.13	H$^\gamma$ 1.58; H$^\delta$ 0.84, 0.66
Glu^{22}	8.51	4.59	1.99, 1.84	H$^\gamma$ 2.49, 2.37
Cys^{23}	8.85	4.78	H$^{\beta2}$ 3.36; H$^{\beta3}$ 2.88	
Lys^{24}	8.68	4.77	1.84, 1.65	H$^\gamma$ 1.42[a], 1.32[a]; H$^\delta$ 1.64, –; H$^\varepsilon$ 2.95, –; NH_3 –
Lys^{25}	8.67	3.94	1.76, 1.30	H$^\gamma$ 0.98[a], 0.94[a]; H$^\delta$ 1.42, –; H$^\varepsilon$ 2.81, 2.71; NH_3 –
Val^{26}	8.08	4.36	2.14	H$^\gamma$ 0.90, 0.71
Asn^{27}	7.83	4.42	3.07, 2.95	HN_2 7.73, 7.02
Trp^{28}	7.87	4.06	2.73, 2.50	N1H 10.01; 2H 6.67; 4H 7.33; 5H 7.10; 6H 7.19; 7H 7.44
Trp^{29}	7.72	4.70	2.96, –	N1H 10.14; 2H 7.04; 4H 7.45; 5H 7.19; 6H 7.23; 7H 7.47
Asp^{30}	8.22	4.99	2.43, 2.36	
His^{31}	8.56	5.39	H$^{\beta2}$ 3.40; H$^{\beta3}$ 3.07	2H 8.78; 4H 7.01
Lys^{32}	8.93	5.24	1.44, –	H$^\gamma$ 1.16, 1.14[a]; H$^\delta$ 1.52, 1.47; H$^\varepsilon$ 2.80, 2.76; NH_3 –
Cys^{33}	7.83	5.15	H$^{\beta2}$ 3.08; H$^{\beta3}$ 2.61	
Ile^{34}	9.27	4.72	1.95	H$^\gamma$ 1.41, 1.31; H$^{\gamma(CH3)}$ 0.96; H$^\delta$ 0.81
Gly^{35}	8.22	4.01, 3.88		

Chemical shifts are in ppm relative to internal TSP at pH 2.9 and 35 °C. <Glu: pyroglutamic acid; –: not assigned; $^{\beta2}$ and $^{\beta3}$: stereospecific assignment of β-methylene protons (IUB-IUPAC representation).
[a] Ambiguous assignment.

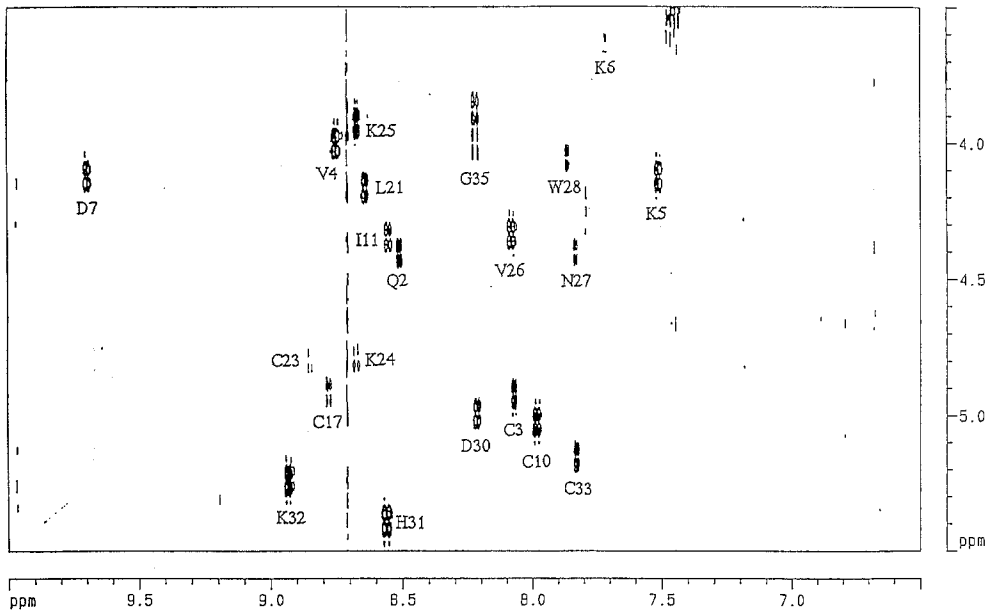

Fig. 1. The fingerprint region of a DQF-COSY spectrum of gurmarin in H_2O at pH 2.9 and 35 °C. The assignments of the NH-C$^{\alpha}$H cross peaks are indicated. The cross peak of Tyr^{13} does not show up due to its large line width.

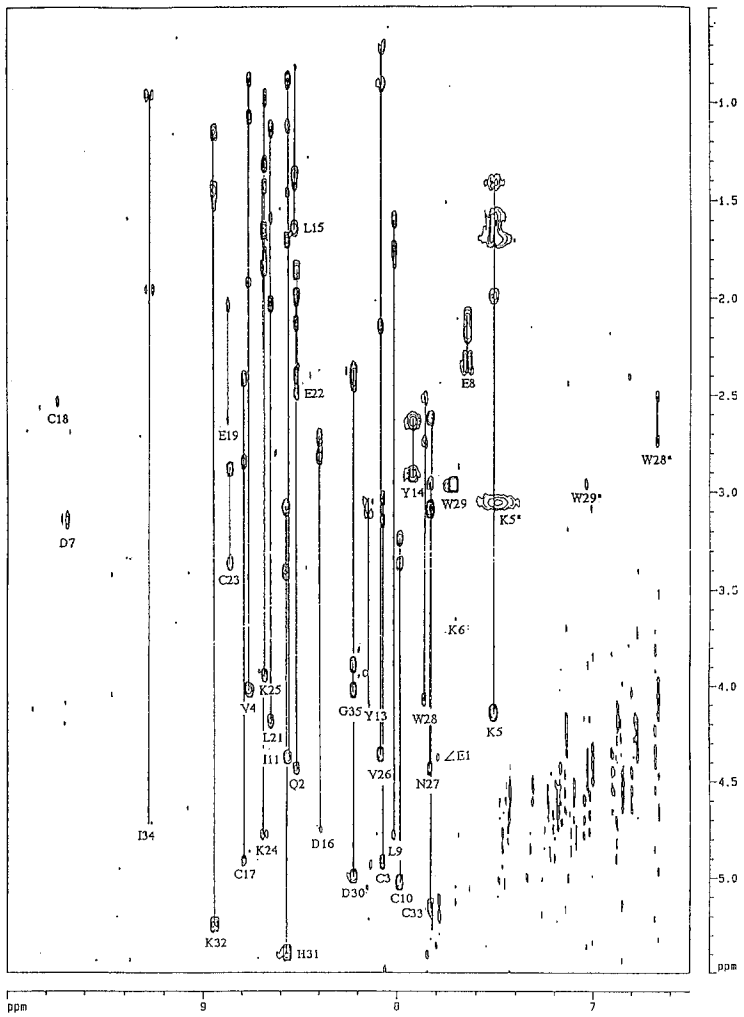

Fig. 2. A portion of a HOHAHA spectrum (τ_m = 55 ms) of gurmarin in H_2O at pH 2.9 and 35 °C. This region shows the cross peaks between NH protons and the other intraresidual protons. The assignments of the relayed cross peaks and the NH-C$^{\alpha}$H cross peaks are indicated.

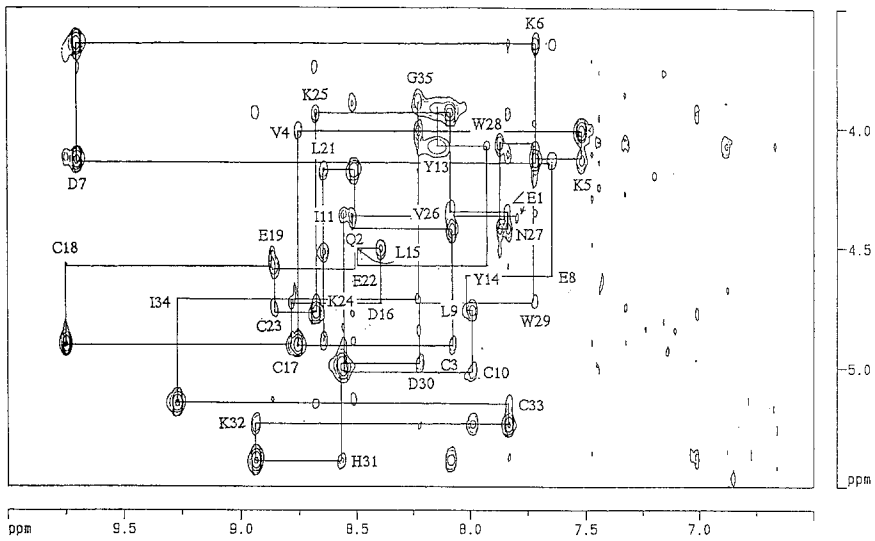

Fig. 3. The fingerprint region of a NOESY spectrum (τ_m = 200 ms) of gurmarin in H_2O at pH 2.9 and 35 °C. The positions of the NH-C$^\alpha$H cross peaks are indicated by residue name and number and the sequential connectivities are represented by solid lines.

systems of these residues were easily identified from the DQF-COSY and HOHAHA spectra. Figure 3 shows the NOESY spectrum in H_2O at 35 °C. The sequential assignments started with these valine residues and were conducted in the NH-C$^\alpha$H cross-peak region, as shown in Fig. 3. Here again, the NH-C$^\alpha$H cross peak of Tyr[13] turned out to have a very large line width. The reason for this broadening is not clear, but some conformational heterogeneity or dynamics may be related to it. All backbone and most of the side-chain proton signals of gurmarin were assigned, as shown in Table 1. The side-chain protons of the lysines, Gln[2] and Trp[29] could not be assigned. Gurmarin contains two tyrosines, two tryptophans and one histidine. Aromatic proton signals of these residues were assigned from the NOESY cross peaks between the corresponding aromatic protons and the C$^\beta$ protons. The side-chain amide protons of Asn[27] were assigned from the NOESY cross peaks between the amide protons and the C$^\beta$ protons. There were no single NOESY cross peaks connecting the side-chain amide protons with the C$^\gamma$ protons in Gln[2]. The unassigned HOHAHA cross peak was considered to correspond to the side-chain amide protons of Gln[2].

For some residues, determination of the dihedral angles χ^1 and stereospecific assignments of β-methylene protons were performed by the combined use of $^3J_{H\alpha H\beta}$ coupling constants and intraresidue sequential NOEs in a manner similar to that described before (Wagner et al., 1987). Residues were investigated for which two $^3J_{H\alpha H\beta}$ values were both larger than 11 Hz, or one $^3J_{H\alpha H\beta}$ value larger than 11 Hz and the other smaller than 5 Hz. The dihedral angles χ^1 were determined for 10 residues, among which stereospecific assignments of β-methylene protons were made for six residues, i.e., Cys[17], Cys[18], Leu[21], Cys[23], His[31] and Cys[33] (Table 1). The dihedral angles χ^1 of all the cysteine residues were determined, which implies the existence of three disulfide bridges. The dihedral angles ϕ were constrained on the basis of the $^3J_{NH\alpha}$ coupling constants. The $^3J_{NH\alpha}$ coupling constants larger than 9 Hz were converted into the range $-160° \le \phi \le -80°$, and those smaller than 5 Hz were converted into the range $-90° \le \phi \le -40°$.

Gurmarin contains two prolines. The Ile[11]-Pro[12] peptide bond was determined to be in the trans conformation from the characteristic C$^\alpha$H(i)-C$^\delta$H$_2$(i+1) sequential NOE and the Glu[19]-Pro[20] peptide bond was determined to be cis

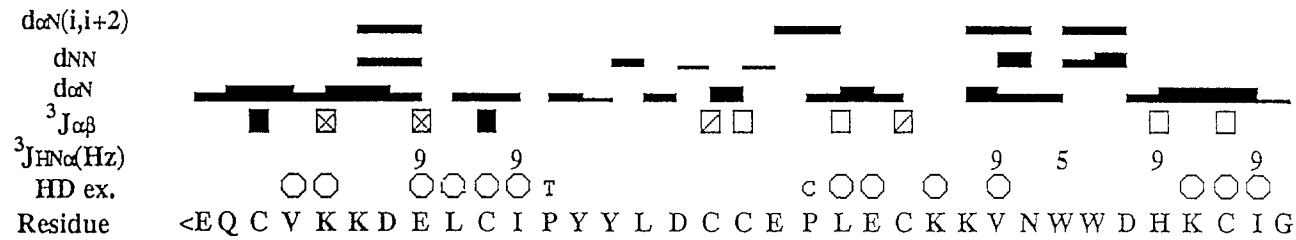

Fig. 4. Summary of the sequential and short-range connectivities, the coupling constants and the slow-exchanging amide protons of gurmarin. NOEs are classified as weak, medium and strong, and the thickness of the bar represents the intensity of NOE. □, ▨, ■ and ⊠ identify constraints on $\chi 1$ of $-60° \pm 60°$, $180° \pm 60°$, $60° \pm 60°$, and $-60° \pm 60°$ or $180° \pm 60°$, respectively. '9' indicates that $^3J_{HN\alpha}$ is larger than 9 Hz and the dihedral angle ϕ was constrained in the range $-160° \le \phi \le -80°$. '5' indicates that $^3J_{HN\alpha}$ is smaller than 5 Hz and ϕ was constrained in the range $-90° \le \phi \le -40°$. Open circles indicate slow exchange rates of amide protons. The *cis*-Pro and *trans*-Pro are marked C and T, respectively.

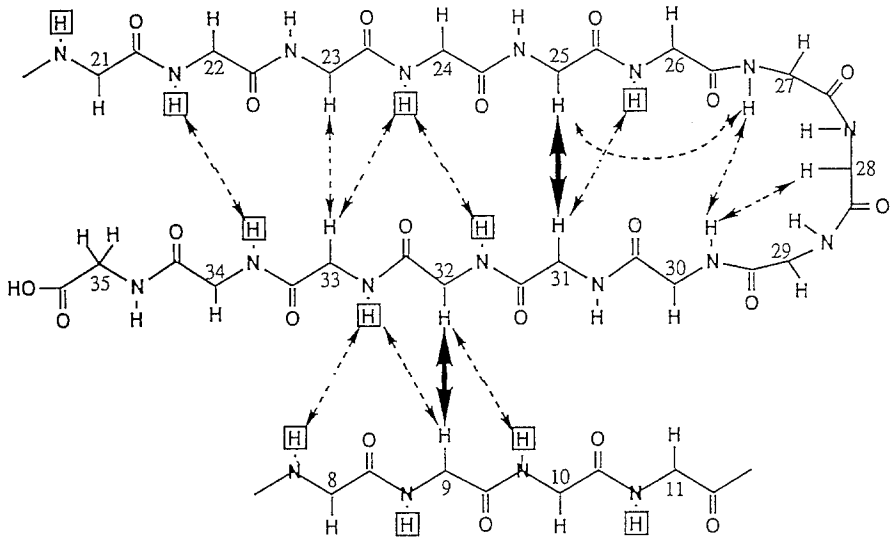

Fig. 5. The β-sheet structure of gurmarin. Solid and dashed arrows indicate strong and medium or weak NOEs, respectively. The slowly exchanging amide protons are squared.

from the $C^{\alpha}H(i)$-$C^{\alpha}H(i+1)$ sequential NOE. Figure 4 summarizes the sequential and short-range connectivities of the NOEs involving NH and $C^{\alpha}H$, the dihedral angles χ^1 and ϕ, and the amide hydrogen exchange rates.

Identification of secondary structure elements

The circular dichroism (CD) spectrum of gurmarin showed that the peptide contains hardly any α-helices (data not shown). This result coincided with the NMR data in the sense that the pattern of short-range NOE connectivities and the $^3J_{NH\alpha}$ values characteristic of an α-helix did not occur. Figure 4 shows the pattern of short-range $d_{\alpha N}$, d_{NN} and $d_{\alpha N}(i, i+2)$ NOE connectivities, and the $^3J_{NH\alpha}$ values. By analyzing the pattern of $d_{\alpha N}$ connectivities and the $^3J_{NH\alpha}$ values, three β-strands of gurmarin could be identified as β1 (residues 8–11), β2 (residues 22–26) and β3 (residues 30–34). From the NOE connectivities between the β-strands their arrangement was determined, as shown in Fig. 5. A total of 13 NH-$C^{\alpha}H$ cross peaks were observed in the HOHAHA spectrum with a mixing time of 40 ms immediately after dissolving gurmarin in D_2O. Several of these slowly exchanging NH protons are associated with the hydrogen-bond networks of the β-sheet, as shown in Fig. 4.

Input data for structure calculations

The distance constraints employed in the structure calculations were collected from the NOESY spectra with mixing times of 50 and 100 ms. A total of 135 distance constraints were derived mostly from the data with τ_m = 100 ms, while the data with τ_m = 50 ms, lacking spin diffusion, were used to supplement the 100 ms data. From the observation of three NOEs, NH(i)-NH(j), NH(i)-$C^{\alpha}H(j+1)$ and $C^{\alpha}H(i-1)$-$C^{\alpha}H(j+1)$, three hydrogen bonds were identified for the β-sheets of gurmarin as (i,j) = (24,32), (32,24) and (33,8). Furthermore, 16 distance constraints from ϕ and χ^1 dihedral angles were obtained, as shown in Fig. 4. The elements of the 157 distance constraints are summarized in Table 2.

Structure calculations

Random-coil conformations of the initial structures were optimized by simulated-annealing calculations and energy minimization. To enable a long-time simulation,

TABLE 2
CONSTRAINTS FOR STRUCTURE CALCULATIONS OF GURMARIN

Upper bound (Å)	Intraresidue	Sequential	Medium-range[a]	Long-range[b]	Total
2.7	2	16	–	2	20
3.5	2	36 (5)[c]	8 (3)	28 (12)	74 (20)
5.0	–	18 (8)	4 (2)	19 (8)	41 (18)
Total	4	70 (13)	12 (5)	49 (20)	135 (38)

[a] Medium-range NOE. The relation between two residues i and j is $2 \leq |i-j| \leq 4$.
[b] Long-range NOE. The relation between two residues i and j is $|i-j| \geq 5$.
[c] The number in parentheses indicates the number of distance constraints between side-chain protons, except β-protons.
In addition to these NOE constraints, six distance constraints corresponding to three hydrogen bonds, six ϕ and 10 χ^1 angle constraints were used in the structure calculations.

Fig. 6. Stereoview of the backbone atoms (N, C^{α}, C) of the 10 converged structures of gurmarin. The individual structures are superimposed on the backbone structure with the smallest total distance violations.

large mass weights of 1000 Da were used for all atoms. Since the positions of the three disulfide bridges of gurmarin have not been determined chemically, structure calculations were carried out on the condition that there were no disulfide bridges, but all the sulfur atoms in the six cysteine residues were considered to be in the same state as the sulfur atoms in disulfide bridges. Out of 100 calculated structures, 10 with the smallest total distance violations were selected (Fig. 6). The rmsd value for any two structures was 1.65 ± 0.39 Å for the backbone atoms (N, C^{α}, C) and 2.95 ± 0.27 Å for all heavy atoms. None of the 10 structures contained distance violations greater than 0.15 Å, all except one violation falling within 0.1 Å. Structural parameters for the 10 converged structures showed very small deviations from ideal values (Table 3).

Discussion and Conclusions

Characteristic features of the determined structure

The secondary structure elements of gurmarin are composed of a triple-stranded antiparallel β-sheet, β1 (residues 8–11), β2 (residues 22–26) and β3 (residues 30–34). Three loops, *l*1 (residues 4–7), *l*2 (residues 12–21) and *l*3 (residues 27–29), connect Cys^3 with β1, β1 with β2 and β2 with β3, respectively.

Figure 7a shows the number of distance constraints at each residue and Fig. 7b shows the rmsd value at each residue between the structure with the smallest total distance violations and the other nine structures for the backbone atoms and all heavy atoms. In residues 4–11, 17–25 and 30–34, the rmsd values are smaller than 1.3 Å for the backbone atoms, and in residues 3–10, 17–18, 20–23, 25–26, 31 and 33–34, they are smaller than 2.0 Å for all heavy atoms. These residues contain three β-strands and three disulfide bridges. The loop *l*1 (residues 4–7) is rather well converged, because of the occurrence of a strong NOE between $C^{\alpha}H$ of Lys^6 and NH of Asp^7. On the other hand, for part of the N-terminal region (residues 1–2), part of *l*2 (residues 12–16) and part of *l*3 (residues 27–29), the numbers of distance constraints are small (Fig. 7a) and the divergence of these residues is very large (Fig. 7b). The fact that the backbone NH protons of these residues exchanged rapidly with deuterons in D_2O suggests that the segments containing these residues are more mobile than the rest of the polypeptide chain.

TABLE 3
STRUCTURAL STATISTICS FOR THE CALCULATED STRUCTURES

Parameter	Value
Bond length (Å)	0.00401 ± 0.000028
Angle (°)	0.749 ± 0.0078
Chiral volume ($Å^3$)	0.0212 ± 0.00044

Averages ± standard deviations were calculated for 10 converged structures. Each value indicates the deviation from the ideal value.

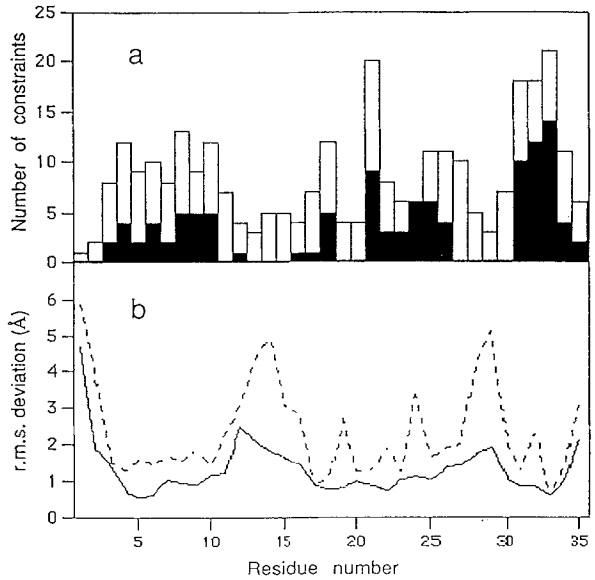

Fig. 7. The number of distance constraints and rmsd value at each residue of gurmarin. (a) The bars represent total constraints; filled parts represent long-range ($|i-j| \geq 5$) constraints. (b) The solid and dotted lines show the rmsd values of the backbone atoms (N, C^{α}, C) and all heavy atoms, between the structure with the smallest total distance violations and the other nine structures, respectively.

TABLE 4
ALL DISTANCES BETWEEN SULFUR ATOMS

Combination of cysteine residues	Distance[a] (Å)
(3, 10)	14.85 ± 0.63
(3, 17)	8.04 ± 0.61
(3, 18)[b]	3.98 ± 0.42
(3, 23)	12.45 ± 1.49
(3, 33)	9.61 ± 0.80
(10, 17)	9.75 ± 0.75
(10, 18)	14.23 ± 0.84
(10, 23)[b]	4.26 ± 0.72
(10, 33)	7.83 ± 0.48
(17, 18)	5.72 ± 0.51
(17, 23)	6.93 ± 0.87
(17, 33)[b]	3.91 ± 0.40
(18, 23)	11.17 ± 1.01
(18, 33)	8.60 ± 0.55
(23, 33)	6.69 ± 0.80

[a] Averages and standard deviations were calculated for 10 converged structures.
[b] The combinations which are considered to be the positions of the disulfide bridges.

Positions of the three disulfide bridges

As already mentioned, structure calculations were carried out without information on the positions of the disulfide bridges. All distances between the sulfur atoms in the six cysteine residues in 10 converged structures are listed in Table 4. It turned out that the sulfur-to-sulfur distances were exceptionally small between Cys^3 and Cys^{18}, between Cys^{10} and Cys^{23} and between Cys^{17} and Cys^{33}. Since the proximities of these sulfur atoms are judged to be more than accidental, we consider it reasonable to conclude that the disulfide bridges are formed in these pairs. We note that the Cys^3-Cys^{18} bridge connects the N-terminal segment and *l*2, the Cys^{10}-Cys^{23} bridge connects β1 and β2, and the Cys^{17}-Cys^{33} bridge connects *l*2 and β3. The arrangement of the above three disulfide bridges in gurmarin was found to be very similar to that in ω-conotoxin, a neurotoxin from marine snails (Nishiuchi et al., 1986), and in *Momordica charantia* trypsin inhibitor-II (MCTI-II), a serine protease inhibitor isolated from bitter gourd seeds (Hara et al., 1989). For the latter two compounds, the positions of the three disulfide bridges have been chemically determined (Fig. 8).

Implications of the determined structure for the function of gurmarin

Although at present we have no direct knowledge of what and where the target molecule of gurmarin is, the distinct feature of gurmarin action, i.e., high specificity to sweet taste stimuli with species-specific properties, suggests that gurmarin interacts with a so-called 'sweet taste receptor' on the taste cell membrane.

Figure 9 depicts the structure with the smallest total distance violations, showing representative side-chain positions. One of the interesting features of the structure is that the loops *l*2 (residues 12–16) and *l*3 (residues 27–29) contain two tyrosines and two tryptophans, whose side chains are all directed outwardly (the bottom part of Fig. 9). Together with the side chains of Leu^9, Ile^{11} and Pro^{12}, they appear to form a cluster that is rich in hydrophobic residues. These loops are also characterized by very poor convergence (Fig. 7b) and high hydrogen exchange rates. It appears probable that they act as the site for interaction with a receptor protein.

Comparing the 3D structure of gurmarin with that of the ω-conotoxin determined recently by NMR (Pallaghy et al., 1993), we notice that they are very similar, with the same topology of folding. While gurmarin was found to be a potent inhibitor of the sweet taste response of rat (Imoto et al., 1991), ω-conotoxin, produced by fish-eating marine snails of the *Conus genus*, is neurotoxic against

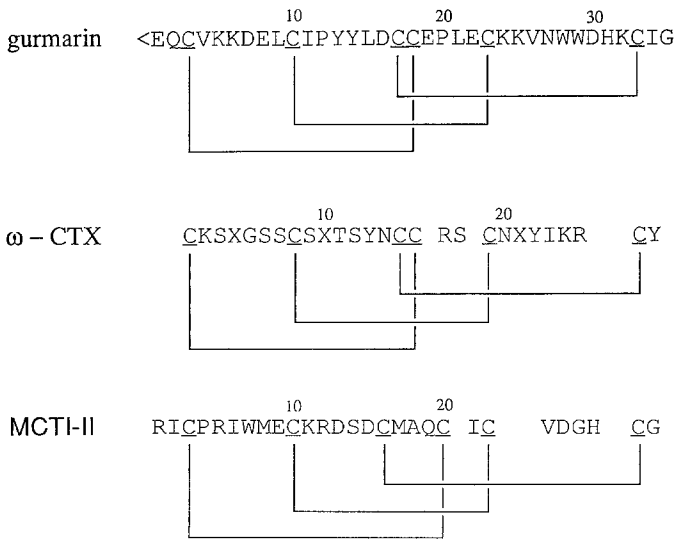

Fig. 8. Amino acid sequences and the positions of three disulfide bridges of gurmarin, ω-conotoxin and MCTI-II.

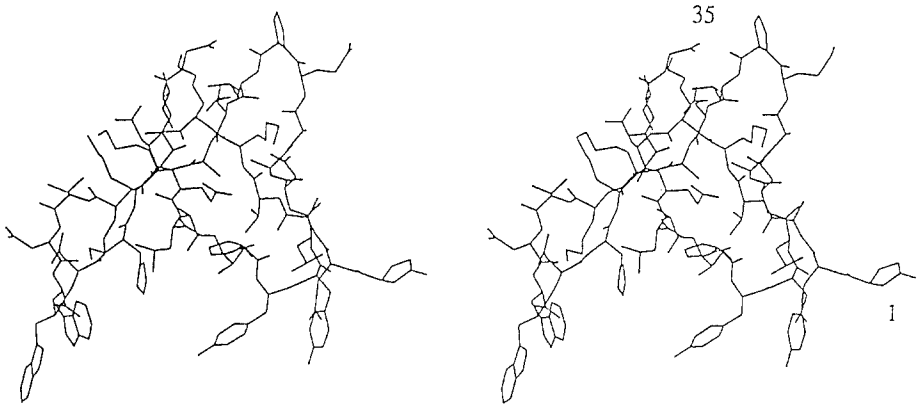

Fig. 9. Stereoview of all heavy atoms of the gurmarin structure with the smallest total distance violations.

mammals by blocking calcium channels (Nishiuchi et al., 1986; McCleskey et al., 1987). The similarity in the 3D structures of gurmarin and ω-conotoxin is quite interesting, in view of the fact that their origins as well as functions are totally different. The question whether the function of gurmarin is merely inhibition against the sweet taste response of rat or if it also has some other functions, remains to be clarified.

Acknowledgements

The 600 MHz ^{1}H NMR spectra were measured and processed at the university facility at the Faculty of Pharmaceutical Sciences, Kyoto University. We thank A. Nosaka and K. Kanaori of Ciba-Geigy Japan, Ltd. for helpful discussions on the NMR measurements, and H. Nakamura and T. Nakai at the Protein Engineering Research Institute for providing the program EMBOSS and for valuable suggestions on the method of calculation. Figures 6 and 9 were prepared by kind cooperation of N. Gō with the help of S. Sunada at Kyoto University. We thank T. Konno and T. Yamaguchi for comments and suggestions during the course of the work. This work was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Science and Culture of Japan (to K.A.).

References

Bax, A. and Davis, D.G. (1985) *J. Magn. Reson.*, **65**, 355–360.
Edgeworth, P. (1847–48) *Proc. Linnean Soc. London*, **7**, 351–352.
Glaser, D., Hellekant, G., Brouwer, J.N. and Van der Wel, H. (1984) *Chem. Senses*, **8**, 367–374.
Hara, S., Makino, J. and Ikenaka, T. (1989) *J. Biochem.*, **105**, 88–92.
Hellekant, G. (1976) *Chem. Senses Flavour*, **2**, 85–95.
Hellekant, G. and Gopal, V. (1976) *Acta Physiol. Scand.*, **98**, 136–142.
Hooper, D. (1887) *Nature*, **35**, 565–567.
Imoto, T., Miyasaka, A., Ishima, R. and Akasaka, K. (1991) *Comput. Biochem. Physiol. A*, **100**, 309–314.
Jeener, T., Meier, B.H., Bachmann, P. and Ernst, R.R. (1979) *J. Chem. Phys.*, **71**, 4546–4553.
Kamei, K., Takano, R., Miyasaka, A., Imoto, T. and Hara, S. (1992) *J. Biochem.*, **111**, 109–112.
Macura, C., Huang, Y., Suter, D. and Ernst, R.R. (1981) *J. Magn. Reson.*, **43**, 259–281.
McCleskey, E.W., Fox, A.P., Feldman, D.H., Cruz, L.J., Olivera, B.M., Tsien, R.W. and Yoshikami, D. (1987) *Proc. Natl. Acad. Sci. USA*, **84**, 4327–4331.
Nakai, T., Kidera, A. and Nakamura, H. (1993) *J. Biomol. NMR*, **3**, 19–40.
Nilges, M., Clore, G.M. and Gronenborn, A.M. (1988) *FEBS Lett.*, **229**, 317–324.
Nishiuchi, Y., Kumagaye, K., Noda, Y., Watanabe, T. and Sakakibara, S. (1986) *Biopolymers*, **25**, S61–68.
Pallaghy, P.K., Duggan, B.M., Pennington, M.W. and Norton, R.S. (1993) *J. Mol. Biol.*, **234**, 405–420.
Rance, M., Sørensen, O.W., Bodenhausen, G., Wagner, G., Ernst, R.R. and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.*, **117**, 479–485.
Wagner, G., Braun, W., Havel, T.F., Schaumann, T., Gō, N. and Wüthrich, K. (1987) *J. Mol. Biol.*, **196**, 611–639.
Weiner, S.J., Kollman, P.A., Nguyen, D.T. and Case, D.A. (1986) *J. Comput. Chem.*, **7**, 230–252.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York, NY.