Conotoxin gm9a, a putative 27-residue polypeptide encoded by Conus gloriamaris, was recently identified as a homologue of the “spasmodic peptide”, tx9a, isolated from the venom of the mollusk-hunting cone shell Conus textile (Lirazan, M. B., Hooper, D., Corpuz, G. P., Ramilo, C. A., Bandyopadhyay, P., Cruz, L. J., and Olivera, B. M. (2000) Biochemistry 39, 1583–1588). The C. gloriamaris spasmodic peptide has been synthesized, and the refolded polypeptide was shown to be biologically active using a mouse bioassay. The chemically synthesized gm9a elicited the same symptomatology described previously for natively folded tx9a, and gm9a and tx9a were of similar potency, implying that neither the two γ-carboxyglutamate (Gla) residues found in tx9a (Ser⁸ and Ala¹³ in gm9a) nor Gly¹ (Ser¹ in gm9a) are crucial for biological activity. We have determined the three-dimensional structure of gm9a in aqueous solution and demonstrated that the molecule adopts the well known inhibitory cystine knot motif constrained by three disulfide bonds involving Cys²-Cys¹⁶, Cys⁶-Cys¹⁸ and Cys¹²-Cys²⁰. Based on the gm9a structure, the sites of Gla substitution in tx9a are in loops located on one surface of the molecule, which is unlikely to be involved directly in receptor binding. Because this is the first structure reported for a member of the newly defined P-superfamily conotoxins, a comparison has been made with structurally related conotoxins. This shows that the structural scaffold that characterizes the P-conotoxins has the greatest potential for exhibiting structural diversity among the robust inhibitory cystine knot containing conotoxins, a finding that has implications for functional epitope mimicry and protein engineering.

A novel Conus polypeptide was recently purified and characterized from the venom of Conus textile, a Conus species known to hunt other snails (1). The new polypeptide defined a new class of venom components, the P-conotoxin superfamily, which has a distinctive cysteine framework, -CXS⁶CX⁶CX⁶CX⁶⁰C-, where X can be any amino acid. Conopeptides with three disulfide linkages and no half-cystines adjacent to each other have not been characterized previously (1, 2). The polypeptide purified from C. textile venom designated the “spasmodic peptide” is unusual in two respects. First, when injected into mice it phenocopied a well known mutant, the spasmodic mouse (3, 4). Second, in addition to the standard amino acids, this polypeptide has two residues of γ-carboxyglutamate (Gla),¹ a distinctive post-translationally modified amino acid.

A second member of the P-superfamily was identified by molecular techniques from a different Conus species, Conus gloriamaris. The sequence of the latter was inferred from a cDNA clone (1). The predicted mature C. gloriamaris sequence is largely identical to the C. textile spasmodic peptide sequence, but it does not have the two glutamate codons presumably encoding Glu residues post-translationally modified to Gla. This polypeptide has not yet been isolated from C. gloriamaris venom. The sequences of the two mature P-superfamily polypeptides are shown in Fig. 1.

In this work, we have synthesized and refolded the C. gloriamaris polypeptide, gm9a, and assessed its activity in the same bioassay used to characterize tx9a. This has been done as a first step in understanding the molecular basis of the symptomatology elicited by tx9a, which is unique among known conotoxins. Of particular interest is the significance of the post-translationally modified amino acids found in tx9a but not in gm9a. Further to this end, we have determined the three-dimensional structure of gm9a in aqueous solution, which is expected to be homologous to that of tx9a considering the extent of sequence homology between the two polypeptides.

Because of their small size, robustness, and scope for structural diversity, conotoxins can be regarded as scaffolds that can be engineered for the mimicry of functional epitopes (5). In this context, the structures of the P-conotoxins are of interest because of their novel cystine framework. In this work, we explore the structural consequences of the additional degrees of freedom in the molecular structure that result when no Cys
residues are directly adjacent in the polypeptide sequence. The gm9a structure is considered in comparison with structurally related conotoxins, particularly the well-characterized O-superfamily conotoxins (6, 7).

**EXPERIMENTAL PROCEDURES**

**Peptide Synthesis and Folding**—Peptides were synthesized using standard solid-phase Fmoc protocol and standard side chain protections. The polypeptides were cleaved from solid support using reagent K (trifluoroacetic acid-water-ethanedithiol-phenoxythioisole, 90:5:2.5:7.5:5 by volume, respectively) for 5 h at room temperature. The crude polypeptide was precipitated with cold methyl tert-butyl ether and washed three times with methyl tert-butyl ether. The linear form was purified using semi-preparative reversed-phase C18 HPLC. The identity of the linear polypeptide was confirmed by electrospray ionization mass spectrometry. Folding reactions were carried out at 0 °C and contained 0.1 M Tris-HCl, 0.1 mM EDTA, 1 mM oxidized glutathione, 2 mM reduced glutathione, pH 8.7, and 20 μM linear polypeptide. After 24 h, the reaction was quenched by adding formic acid to 5% final concentration, and the mixture was separated by reversed-phase C18 HPLC in a linear gradient of acetonitrile in 0.1% trifluoroacetic acid from 9 to 31.5% in 20 min. For analytical and semi-preparative C18 columns, flow rates of 1 and 5 ml/min were used, respectively. All HPLC chromatography was monitored by UV absorption at 220 nm.

**NMR Spectroscopy**—A sample of gm9a was prepared for NMR analysis by dissolving the polypeptide in 0.30 ml of 90% H2O, 10% D2O. The pH was adjusted to 5.5 with small additions of 0.5 M NaOD, and the solution was transferred to a H2O-matched Shigemi tube (5-mm outer diameter, 8-mm-bottom length). Reported pH values were recorded at room temperature with no correction made for isotope effects. 1H chemical shifts were referenced to 2,2-dimethyl-2-silapentane-5-sulfonate at 0 ppm via the chemical shift of 1,4-dioxane at 3.75 ppm (8).

NMR spectra were recorded on Bruker AMX-500 or DRX-600 spectrometers. All two-dimensional spectra were recorded in phase-sensitive mode using the time-proportional phase incrementation method (9). Water suppression was achieved either by pulsed field gradients in a 90°/90°, were applied in both dimensions prior to Fourier transformation. Sine-squared window functions, phase-shifted by 8992.8 Hz at 500 and 600 MHz, respectively. Spectra were processed and 5 ml/min were used, respectively. All HPLC chromatography was monitored by UV absorption at 220 nm.

**Structure of P-conotoxin gm9a**

**Verification of Disulfide Connectivities**—To determine the pairing of Cys residues forming native disulfide bonds, we used the strategy described by Gray (21). An equal volume of 20 mM TCEP (Fierce, Rockford, IL) in 0.2 M citrate buffer, pH 3.0, and folded gm9a in HPLC eluent were mixed and allowed to react for 20 min at room temperature. Typically, the reaction contained 5–10 nmol of peptide in 0.1–0.2 ml. Resulting products were separated by reversed-phase HPLC using analytical C18 column (Vydac) in a linear gradient of acetonitrile (0.1% trifluoroacetic acid) from 13.5% to 22.5% in 20 min. The HPLC peaks were immediately subjected to rapid alkylation using saturated iodocetamide solution (0.2 M iodoacetic acid and 0.2 ml of 0.5 M Tris acetate buffer, 2 mM EDTA, pH 8.0, and the solution was heated to 65 °C until the solid was dissolved. 10 μl of 100 mM ascorbic acid was added to the iodocetamide solution to scavenge possible traces of iodine, and then 0.3 ml of the partially reduced intermediate in HPLC eluent was rapidly injected into the cooled iodocetamide solution using a fine-tipped syringe and mixed vigorously by gentle pipetting. The mixture was then quickalkylated with 0.5 M citric acid (0.4 ml), and the mixture was separated on a C18 analytical HPLC column. The following elution conditions were developed to separate alkylation products: 20 min of 9% acetonitrile in 0.1% trifluoroacetic acid (isocratic) followed by a linear gradient of acetonitrile (9–45% for 20 min) and flow rate of 1 ml/min. The resulting disulfide intermediates, which contained some pairs of carboxymethylated Cys residues, were further reduced with dithiothreitol and alkylated with 4-vinylpyridine. The HPLC eluent containing the intermediates was adjusted to pH 8.0 by adding 0.5 M Tris base solution. The intermediates were reduced in 10 mM dithiothreitol at 65 °C for 15 min. After the solution was cooled, 3 μl of 4-vinylpyridine was added for every 0.5 ml of the sample volume. The alkylation reaction was carried out in the dark at room temperature for 25 min and then diluted with 0.1% trifluoroacetic acid and injected on a C18 analytical HPLC column. The HPLC conditions were similar to those described above with the exception of the linear acetonitrile gradient applied (9–27% in 20 min). The final alkylation products were sequenced using a standard Edman protocol. Positions of carboxymethylated and pyridylethylated cysteines were determined from the sequencing data.
To obtain a cyclic polypeptide with disulfide bridges, we used a direct oxidation method in which the polypeptide containing six unprotected cysteines was allowed to form disulfides under oxidizing conditions. The linear polypeptide was treated with the mixture of oxidized and reduced glutathione. The folding reaction contained 0.1 M Tris-HCl, 0.1 mM EDTA, 1 mM oxidized glutathione, 2 mM reduced glutathione, pH 8.7, and linear polypeptide (20 μM). Folding was carried out at 0 °C for 24 h. The acid-quenched reaction mixture was separated by reversed-phase C18 HPLC. A major folding product was purified and characterized using the mouse bioassay previously employed for the C. textile spasmodic peptide (1). Qualitatively, the biological activity of purified synthetic gm9a was similar to that described previously for tx9a, the spasmodic peptide isolated from the venom of C. textile (1). Both peptides elicited the spasmody symptomatology described previously (1). A direct comparison between the two peptides demonstrated that both caused hypersensitivity to touch in mice when injected at a dose of 2 nmol and hyperactivity and convulsions at 5 nmol.2

NMR Spectroscopy—NMR spectra for structural analysis were collected on the gm9a polypeptide in 90% H2O, 10% D2O at pH 5.5. Broad NH chemical shift dispersion in one-dimensional spectra indicated that the polypeptide was well structured, and standard two-dimensional spectra recorded at 5, 15, and 25 °C showed that only one conformation was present in aqueous solution at these temperatures. Most amide peaks were well resolved at all of these temperatures, and many values of the 3J,HN-C,H coupling constants were apparent from the one-dimensional spectra at 500 MHz. Sequence-specific chemical shift assignments were made at 600 MHz for backbone and side-chain protons in gm9a with the aid of standard two-dimensional homonuclear NMR experiments. These assignments are provided as supplementary data (Table S-I) and have been deposited with BioMagResBank (22) under accession number BMRB-6086. Distance constraints were taken from the volumes of NOE cross-peaks in NOESY spectra acquired at 5 °C at 600 MHz.

In general, spectral overlap was minimal, but overlap in the regions of the Cys CßH and NH resonances made it difficult to assign unambiguously a small number of NOE cross-peaks that might have assisted in the identification of Cys-Cys bonds. Most of these ambiguities were resolved by conducting experiments at different temperatures and different pH values to disperse chemical shifts of key protons. Despite this, direct identification of Cys-Cys connectivities remained problematic as a number of Cys CßH and CßH had through-space NOE cross-peaks with more than one other Cys, indicative of a cluster of Cys side-chains in the polypeptide.

A suite of NMR experiments yielded distance and torsion angle constraints for use in structural calculations. Initial calculations were performed in DYANA with no disulfide or hydrogen bonds present. The family of 20 best structures obtained with no disulfide bonds imposed was used as a reference to determine the disulfide-bonding pattern in gm9a that was most consistent with the NMR data.

Structural constraints were well satisfied for every combination involving either the Cys2-Cys16 or Cys8-Cys18 pairings (Table S-II from supplementary data), and the lowest target function was obtained when both of these bonds were in place, i.e., the combination of Cys2-Cys16, Cys8-Cys18, and Cys12-Cys23. The structures derived with this combination of disulfides used to constrain the molecule showed that gm9a adopts the inhibitory cystine knot (ICK) motif (7, 23), a common structural motif incorporating a cystine knot and a triple-stranded β-sheet in toxic and inhibitory polypeptides.

Chemical Verification of Disulfide Connectivity in gm9a—The disulfide connectivities consistent with the NMR data were confirmed chemically before the final structure calculations were undertaken. The general strategy involved a partial reduction of the folded gm9a at acidic pH followed by a primary alkylation of the resulting Cys thiols with iodoacetamide. The alkylation products were isolated by reversed-phase HPLC, further reduced with dithiothreitol, and subjected to a secondary alkylation with 4-vinylpyridine. The Cys connectivities were deduced using standard Edman sequencing by identification of positions of carboxymethyl- and pyridylethyl-modified cysteines.

Fig. 2 shows the HPLC separation of products after TCEP reduction of the folded gm9a for 20 min at pH 3.0. Three major intermediates were purified and alkylated with iodoacetamide as described under “Experimental Procedures.” Two intermediates (Fig. 2, peaks labeled A and C) yielded clear alkylation products. However, intermediate B apparently underwent a rearrangement during alkylation as judged by multiple alkylation products observed after HPLC analysis (data not shown). Therefore, only alkylated intermediates A and C were further purified using reversed-phase HPLC.

2 E. C. Jimenez, unpublished results.
The polypeptides were reduced for 15 min at 65 °C using 10 mM dithiothreitol followed by addition of 4-vinylpyridine. After secondary alkylation, the polypeptides were purified by HPLC and sequenced. Intermediate A contained carboxymethylated cysteines at positions 6 and 18 and pyridylethylated cysteines at positions 2, 12, 16, and 23. Intermediate C contained carboxymethylated cysteines at positions 2, 6, 16, and 18 and pyridylethylated cysteines at positions 12 and 23. Based on these data, we deduce that the disulfide connectivity of biologically active gm9a is Cys2–Cys16, Cys6–Cys18, and Cys12–Cys23, which is the combination most consistent with the NMR data.

**NMR Assessment of Hydrogen Bonds**—The temperature dependence of the chemical shift of each gm9a amide proton was determined to probe for hydrogen bonding. However, all amide protons exhibiting a temperature coefficient of magnitude ≤3 ppm°C (Table S-III from supplementary data) were found to be from either cysteines (16, 18, 23) or residues directly adjacent to cysteines (5, 7, 11, 24). Because structure calculations indicated that the three disulfides form the core of folded gm9a, the low susceptibility of these amide chemical shifts to a change in temperature presumably reflects the location of these protons in the polypeptide core rather than their involvement in hydrogen bonding (24). Amide exchange experiments conducted at 5 °C and pH 5.5 indicated that all amide protons with the exception of those of Gln7 and Ile11 exchanged with solvent deuterium within minutes of dissolution of the polypeptide in 2H2O. Amide protons from these two residues were apparent in NMR spectra for several hours following dissolution; however, again these residues are directly adjacent to core Cys residues. Because the experimental indicators of hydrogen bonding could have been influenced by reduced solvent accessibility, no hydrogen bonds were used as distance constraints in structural calculations.

**Solution Structure of Conotoxin gm9a**—Parameters characterizing the final 20 structures of gm9a and structural statistics are summarized in Table I, and stereoviews of the structures superimposed over the backbone are shown in Fig. 3. These structures were calculated with the chemically determined disulfide bonds intact. Table I indicates that the final 20 structures fit well with experimentally derived distance and angle constraints and are well defined over the entire length of the polypeptide. The structures of gm9a have been deposited with the Protein Data Bank (PDB accession number 1IXT) (25).

Different features of the closest-to-mean gm9a structure are highlighted in Fig. 4. In terms of overall topology, the molecule has a sulfur-rich core of three disulfides that form an inhibitory cystine knot as illustrated in Fig. 4E and a backbone made up of three extended sections encompassing residues 1–6, 15–19, and 22–27 and three loops centered on residues 7–11, 12–14, and 20–21. Procheck (20) was used to identify elements of secondary structure in the final family of gm9a structures. The most prominent feature of the structures is a β-hairpin in the COOH-terminal half of the molecule (Fig. 4, A and B) with the first extended β-strand encompassing His15–Thr19, and the second extended β-strand encompassing Gly22–Ala26. The loop at the center of the hairpin (Phe20 and Arg21) was identified as a hydrogen-bond-stabilized Type I turn. Procheck also identified an isolated β-bridge at Asn4 that is hydrogen-bonded through its carbonyl to Cys25 in the second β-strand (26). However, all of the residues in the range Cys2–Cys6 occupied the β-sheet region of the Ramachandran plot, suggesting that gm9a contains a triple-stranded anti-parallel β-sheet in common with many ICK-containing conopeptides (7). The fact that this extended region is constrained at either end by a disulfide-bonded Cys (Cys2 and Cys8) may distort the backbone from ideal β-strand geometry. The final structures also contain hydrogen bonds between strands and across all three loops (Table S-III from supplementary data). The disulfides formed by Cys2–Cys16 and Cys12–Cys23 adopt a left-handed conformation, whereas the remaining disulfide, Cys6–Cys18, is right-handed.

Table I: Structural statistics for gm9a

| Constraint type                                | No. | S.D. | Average |
|------------------------------------------------|-----|------|---------|
| NOE-derived distance constraints               | 439 | 0.0048 ± 0.0002 | 0.015 |
| Intraresidual and sequential                   | 242 | 0.481 ± 0.046 | 0.028 |
| Medium range                                   | 30  | 0.373 ± 0.053 | 0.036 |
| Long range backbone                            | 12  | 0.70 ± 0.17  | 0.06 |
| Long range                                     | 137 | 1.30 ± 0.18  | 0.16 |
| Ramachandran plot                              | 56.5| 13.0 |
| % in additional allowed region                 | 30.4| 11.0 |
| % in generously allowed region                 |     |       |

* Average values ± S.D. calculated for the 20 final structures.


distances violations > 0.2 Å were 0, and no dihedral violations > 5° were 0.

No. of dihedral angle constraints

| Angle violation type                  | No. | S.D. |
|---------------------------------------|-----|------|
|Total                                  | 26  | 0.055 ± 0.174 |
|Medium range                           | 27  | 0.025 ± 0.001 |
|Long range                             |     |       |
|Ramachandran plot                      | 56.5| 13.0 |
|% in additional allowed region         | 30.4| 11.0 |
|% in generously allowed region         |     |       |

* Average values ± S.D. calculated for the 20 final structures.

**Average pairwise r.m.s. deviations of atomic coordinates Cα, Cα, and N for residues 1–27 across the 20 final structures.**

**Average pairwise r.m.s. deviations of all atoms excluding hydrogen for residues 1–27 across the 20 final structures.**

| R.m.s. deviations from idealized geometry | No. | S.D. |
|------------------------------------------|-----|------|
|Bonds (Å)                                 | 0.0048 ± 0.0002 | 0.015 |
|Angles (°)                                | 0.481 ± 0.046 | 0.028 |
|Improper (°)                              | 0.373 ± 0.053 | 0.036 |
|Backbone atoms r.m.s. deviations (Å)      | 0.70 ± 0.17  | 0.06 |
|All heavy atoms r.m.s. deviations (Å)     | 1.30 ± 0.18  | 0.16 |

**A limited pH titration was undertaken to determine the pKa values for the imidazolium groups of His9 and His15 and the carboxyl of Asp11. As summarized in Table S-IV from supplementary data, the imidazolium pKa values were 4.9 and 6.0, respectively, and the carboxyl pKa value was 4.8. In small histidine-containing peptides the imidazolium pKa was 7.0 at 35 °C (28), whereas in the uncharged model compound Ac-HisNHMe, it was 6.38 at 37 °C (29), 6.43 in H2O, and 6.54 in 2H2O, each containing 0.1 M NaCl at 30 °C (30). The intrinsic pKa values cited by Shire et al. (31) is 6.3. The pKa of His15 is only slightly
Fig. 3. gm9a structures. Stereoviews of the family of 20 final structures determined for gm9a superimposed over backbone heavy atoms from all 27 residues. For clarity, amino acid side-chains have been eliminated from the family of structures in the top view. Side-chain orientations are shown for the closest-to-mean structure in the bottom view. Disulfide bonds are shown in gray. The letters N and C refer to the amino and carboxyl termini, respectively.

Fig. 4. Aspects of the gm9a structure. A and B, ribbon diagrams representing the closest-to-mean gm9a structure where the two perspectives are related by a 180° rotation along the y axis, and yellow arrows indicate β-strands. C and D, simple charge distribution on the surface of gm9a and are shown in the same orientation as A and B, respectively. Electrostatic potentials are color-coded blue for the positively charged residues Ser1 and Arg21 and red for the negatively charged residue Asp11. E, a ribbon representation of the gm9a backbone, which demonstrates the inhibitory cystine knot motif where the Cys12-Cys23 disulfide bond passes through a loop constrained by the Cys2-Cys16 and Cys6-Cys18 disulfides. In E, disulfide bonds are shown in yellow, and the backbone of residues forming the loop are shown in violet. F, the closest-to-mean structure obtained for gm9a with the location of Ser8 and Ala13 indicated, which are sites of Gla substitution in the tx9a polypeptide. In F, Arg21 and Ser1 are in blue, hydrophobic residues (Ile17, Phe20, Ala25, and Val26) are shown in coral (excluding Ala13), Asp11 is in red, and the sites of Gla substitution are in cyan. The letters N and C refer to the amino and carboxyl termini, respectively.
lower than the model values, consistent with its largely solvant-exposed environment. The significantly lower pKₐ for His³ is consistent with its location on the surface but in close contact with several non-polar side chains to which it shows numerous NOEs (Ile¹² and Cys¹⁸). These interactions are expected to favor the neutral form and thus reduce the pKₐ. It is probable that protonation of the imidazole side chain causes some local conformational changes associated with movement of the positively charged ring to a more solvant-exposed environment and that changes in the chemical shifts of Ile¹² methyl resonances at this titration occurs probably reflect this. His³ is not close to the Arg²¹ side chain or the N-terminal ammonium group, so direct effects of either of these positively charged groups are unlikely to contribute to its low pKₐ.

The C⁵⁰H resonances of Asp¹¹ in gm9a titrate with a pKₐ of 4.8 (Table S-IV from supplementary data) significantly higher than the carboxyl group of Asp in small peptides with a pKₐ of 3.9 (28, 32). Even though the direction and magnitude of the chemical shift changes associated with this titration are as observed in small peptides (28), it is nevertheless possible that this pKₐ reflects the titration of His³ rather than the Asp¹¹ carboxyl. The two side chains are not in close proximity; however, both residues are located on a reverse turn (distorted Type I involving His³ and Ser¹⁰), so it is possible that local conformational changes associated with the titration of His³ could affect the environment of Asp¹¹. If the pKₐ of 4.8 does reflect the carboxyl titration, then it is unusually high. As in the case of His³, Asp¹¹ is located on the surface but is in close contact with several other side chains (residues 6, 7, 8, 10, and 12) to which it shows numerous NOEs. These interactions would be expected to favor the neutral form and thereby raise the carboxyl pKₐ.

**DISCUSSION**

Lirazan et al. (1) recently identified conotoxin tx9a from the venom of the cone shell C. textile and argued that it defined a new class of conotoxins, the P-superfamily, on the grounds that it had a novel cysteine framework (−C₅X₅C₅X₅−C₅X₅C₅−) and precursor structure and was pharmacologically distinct from other conotoxins. A second polypeptide that had high sequence identity with tx9a including the distinctive Cys framework was recently identified from the seabed cone Conus textile (33). This conopeptide is functionally related to tx9a. We and others have synonymized this polypeptide as gm9a. A second polypeptide that had high sequence identity with tx9a including the distinctive Cys framework was recently identified from the seabed cone Conus textile (33). This conopeptide is functionally related to tx9a. We and others have synonymized this polypeptide as gm9a.

**Functional Diversity among Conotoxins—**One of the tenets of structural genomics is that structure provides clues regarding function, but it has been pointed out that this nexus generally is weak in the case of polypeptide toxins (33). This is certainly true of the conotoxins where functional diversity is achieved by a combination of sequence variation and structural variation. As a result, structural analysis reveals few clues as to the potential molecular target of a given polypeptide. Fig. 5 shows ribbon representations of several conotoxin structures that contain six Cys residues and exhibit the same disulfide-bonding pattern exhibited by gm9a, namely C₁-C₂/C₅-C₆/C₇-C₈. Further details of these conopeptides are given in Table II.

Table II. The structures were downloaded from the Protein Data Bank (25) and correspond as follows: A, GIIIA (35); B, GIIIB (34); C, P-IIIE (36); D, TVIIIA, I; E, conotoxin GS (38); F and G, trans- and cis-isomers of EVLA, respectively (PDB codes 1G1P/1G1Z); H, SO3 (PDB code 1FGO); I, MVIIIA (41); J, GVIA (42); K, PVIIA (43); L, gm9a. The location of each cysteine is indicated in black.

![Fig. 5. Experimental structures of conotoxins with the same disulfide-bonding pattern as gm9a.](image-url)
Structure of P-conotoxin gm9a

Description of various conopeptides that have the same disulfide-bonding pattern as gm9a

| Peptide        | Sequence     | Cys - Cys, a | Ref.       |
|----------------|--------------|--------------|------------|
| A-GHIB         | APCCTXXYRCDEKKCKXGGGCA | 0, 5, 4, 4, 0 | (34)       |
| B-GHIB         | APCCTXXYRCDEKKCEGXGGGCA | 0, 5, 4, 4, 0 | (35)       |
| C-PHIE         | HXKCCYLCRCRSSYKCGSSACCGQR            | 0, 4, 5, 4, 0 | (36)       |
| D-TVIIA        | SCGSGDRSCEKXCGMGLCRSCRGCVSTYFEG     | 6, 3, 0, 4, 4 | (37)       |
| E-GS           | ASGGSCXCGGGLCGLIRCGRGPOCVICGGLG      | 6, 3, 0, 4, 7 | (38)       |
| F-EVIA         | DCXETYGYFCSPL1PLINGCSCGACGVCNAD     | 6, 9, 0, 3, 3 | (PDB 1GIZ) |
| G-EVIA         | DCXETYGYFCSPL1PLINGCSCGACGVCNAD     | 6, 9, 0, 3, 3 | (PDB 1GIZ) |
| H-SO3          | CKAKGKPSRIAYNGTCSRSGKC              | 6, 6, 0, 3, 4 | (41)       |
| I-MVIIA        | CKGGKACSRLMYDCSTCSSRCGSKGC          | 6, 6, 0, 2, 6 | (42)       |
| J-GVIA         | CKGXSScrSTWEYSRCNRCRXYRKCY          | 6, 6, 0, 3, 4 | (43)       |
| K-PVIIA        | CSR1XNCDFQMLDECCNGCRRNFCNGC        | 3, 5, 3, 1, 4 | (1)        |
| L-gm9a         | SCNNCGQCAHSHCCTFCPEGGAVY            | 3, 5, 3, 1, 4 | (1)        |
| M-tv9a         | SCNNCGQCAHSHCCTFCPEGGAVY            | 3, 5, 3, 1, 4 | (1)        |

a Cysteines are shown emboldened, residues with charged side chains at physiological pH are underlined, γ represents Gla, and X represents hydroxyproline.

b Each number series refers to the number of residues between sequential cysteines in the polypeptide sequence.

4, and 7 and a structure that differs from the ω- and κ-conotoxins. In TVIIA (D), these spacings are 6, 3, 0, 4, and 4, and as a consequence, it has essentially the same structure as conotoxin GS over the disulfide-linked region of the molecule with the exception that the loop between C5 and C6 is approximately half as large in TVIIA (Fig. 5). A similarly significant departure from the classic ω-conotoxin fold is observed in EVIA (G), which has the spacings of 6, 9, 0, 3, and 3. When this type of analysis is applied to the P-conotoxin gm9a in which 3, 5, 3, 1, and 4 residues separate sequential Cys residues, it is readily apparent why the structure of this polypeptide is unique among those shown in Fig. 5.

In light of this analysis, two points can be made with regard to the newly defined P-conotoxins. First, because the number of residues between sequential half-cystines is a significant indicator of structural variation and as new members of this family of conopeptides are identified, this source of structural diversity can be used to assess the extent of their homology with gm9a. Second, because the P-conotoxins have no cysteines directly adjacent to one another, the scaffold underlying this superfAMILY of conopeptides has the greatest scope for exhibiting structural and functional diversity among all of the six Cys-containing conotoxins. This latter point has significant implications for protein engineering and functional epitope mimicry (40, 45). An inspection of the structure of gm9a suggests that the loops between the second and third half-cysteines and the fifth and sixth are potentially the most suitable for the introduction of new sequences.

Implications for tv9a from C. textile—Chemically synthesized gm9a, the C. gloriamaris spasmodyic peptide, has biological activity comparable to that elicited by the C. textile spasmodyic peptide in an in vivo assay in mice (1). The two spasmodyic peptide sequences from C. textile (tv9a) and C. gloriamaris (gm9a) differ in only three positions, with Ser1, Ser6, and Ala10 in gm9a substituted for Gly, Glu, and Gla, respectively, in tv9a. We expect tv9a to adopt a conformation similar to that of gm9a, because the location of the six Cys residues is identical in the two polypeptides, the substitution at the flexible N terminus is not likely to be structurally significant, and the two Cys sequences Gln and Gla, neither glycine nor proline residues, mutations that might be expected to yield structural variation. In addition, it was demonstrated by Hill et al. (38) that substituting Gla for Glu in conotoxin GS resulted in no change in the backbone conformation of that polypeptide.

The nature of the molecular target of gm9a and tv9a remains undefined, but the symptomatology elicited by both polypeptides is reminiscent of spasmodyic and spastic mouse mutants (3, 4), which were characterized as being defective in glycine receptors. Although the receptor for these polypeptides is still to be determined, it is apparent from the comparable results of the mouse assays that the two γ-carboxyglutamates in tv9a are not involved directly in receptor binding.

An analysis of the gm9a structure shows that residues 8 and 13 (Fig. 4F) are found on the same face of the molecule, flanking the only acidic residue Asp11. One consequence of this finding is that on one face of the molecule, tv9a is substantially more electronegative than gm9a. This is significant because in the main, conotoxins are rich in basic side-chains (see Table II) that interact with anionic sites in their ion channel receptors. This also implies that this region of the molecular surface does not constitute the receptor-binding interface of these polypeptides. Rather, it can be inferred that the polar face of these molecules is solvent-exposed when the polypeptides bind their target and that the binding epitope is probably on the hydrophobic/cationic face of the molecule.

γ-Carboxyglutamates have been implicated in Ca2+ binding by Prorok et al. (39), who demonstrated that the Glu-rich Conus-derived conantokins T and G were susceptible to conformational change as a result of calcium binding to Glu residues. In these polypeptides, however, Glu residues are in close proximity to one another with some being adjacent in the sequences. Hill et al. (38) considered the possibility that the single Glu residue in conotoxin GS might be involved in Ca2+ binding. They demonstrated by NMR that there was no conformational change in the polypeptide observable in the presence of calcium ions and found no evidence of line broadening or perturbation of Glu32 side-chain resonances that might indicate calcium binding. It was concluded that this single residue might not play a crucial role in either the structure or function of conotoxin GS.

tv9a has two Glu residues, but inspection of the gm9a structures indicates that across the family of 20 structures, residues 8 and 13 are always >10 Å apart. Clearly, this is too large a distance for these residues to cooperatively chelate Ca2+ ions in the folded structure. However, it is conceivable that one or other of the Glu residues and the intervening Asp11 could participate in Ca2+ binding in tv9a. It is also possible that these Glu residues could be involved in the folding of the tv9a polypeptide. Prorok et al. (39) found that Ca2+ binding induced an increase in the α-helical content of the conantokins considered in their work. In gm9a and presumably in tv9a, residues 8 and 13 are found in the loops of the molecule. In tv9a, the formation of these loops could conceivably be induced by Ca2+ binding.
Acknowledgments—We thank Richard DelaCruz for preparation of synthetic polypeptide for NMR work and Elsie Jimenez for bioassays.

REFERENCES

1. Lirazan, M. B., Hooper, D., Corpuz, G. P., Ramilo, C. A., Bandyopadhyay, P., Cruz, L. J., and Olivera, B. M. (2000) Biochemistry 39, 1583–1588
2. Balaji, R. A., Ohtake, A., Sato, K., Gopalakrishnakone, P., Kini, R. M., Seow, K. T., and Bay, B. H. (2000) J. Biol. Chem. 275, 39516–39522
3. Ryan, S. G., Buckwalter, M. S., Lynch, J. W., Handford, C. A., Segura, L., Shiang, R., Wasmuth, J. J., Camper, S. A., Schofield, P., and O’Connell, P. (1994) Nat. Genet. 7, 131–135
4. Saul, B., Schmieden, V., Kling, C., Mulhardt, C., Gass, P., Kuhse, J., and Becker, C. M. (1994) FEBS Lett. 350, 71–76
5. Pallaghy, P. K., and Norton, R. S. (2000) Biopolymers 54, 173–179
6. Norton, R. S., Pallaghy, P. K., Baell, J. B., Wright, C. E., Lew, M. J., and Pallaghy, P. K., and Norton, R. S. (1998) Biochem. Biophys. Res. Commun. 153, 2577–2583
7. Norton, R. S., and Pallaghy, P. K. (1998) Toxicon 36, 1573–1583
8. Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995) J. Biomol. NMR 6, 139–140
9. Marion, D., and Wuthrich, K. (1983) Biochem. Biophys. Res. Commun. 113, 967–974
10. Sklenar, V., Piotto, M., Leppik, R., and Saudek, V. (1983) J. Magn. Reson. 102, 241–245
11. Kumar, A., Ernst, R. R., and Wu, J. (1981) J. Magn. Reson. 47, 409–422
12. Brunschwiler, L., and Ernst, R. (1983) J. Magn. Reson. 53, 521–528
13. Levy, M. H., Freeman, R., and Frenkel, T. (1983) Adv. Magn. Reson. 11, 47
14. Rance, M., Sörensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., and Wuthrich, K. (1983) Biochem. Biophys. Res. Commun. 117, 479–485
15. Griesinger, C., Sörensen, O. W., and Ernst, R. R. (1987) J. Magn. Reson. 75, 474–492
16. Bartels, C., Xin, T. H., Billeter, M., Güntert, P., and Wüthrich, K. (1995) C. R. Acad. Sci. 319, 139–146
17. Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) J. Mol. Biol. 273, 283–298
18. Bune, C., T., Adams, P. D., Clare, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszezewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Acta Crystallogr. Sect. D 54, 905–921
19. Linge, J. P., and Nilges, M. (1999) J. Biomol. NMR 13, 51–59
20. Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) J. Biomol. NMR 8, 477–486
21. Gray, W. R. (1993) Protein Sci. 2, 1732–1748
22. Seavey, B. R., Farr, E. A., Westler, W. M., and Markley, J. L. (1991) J. Biomol. NMR 1, 217–226
23. Pallaghy, P. K., Nielsen, K. J., Craik, D., and Norton, R. S. (1994) Protein Sci. 3, 1833–1839
24. Baxter, N. J., and Williamson, M. P. (1997) J. Biomol. NMR 9, 359–369
25. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) Nucleic Acids Res. 28, 235–242
26. Kabsch, W., and Sander, C. (1983) Biopolymers 22, 2577–2637
27. BioMagResBank (BMRB) database. www.bmrb.wisc.edu
28. Bundi, A., and Wuthrich, K. (1979) Biopolymers 18, 285–297
29. Tanokura, M., Tosu, M., and Miyazawa, T. (1976) Biopolymers 15, 393–401
30. Perutz, M. F., Gronenborn, A. M., Clore, G. M., Fogg, J. H., and Shih, D. T. (1985) J. Mol. Biol. 183, 491–498
31. Shire, S. J., Hanania, G. I., and Gurd, F. R. (1974) Biochemistry 13, 2967–2974
32. Keim, P., Vigna, R. A., Morrow, J. S., Marshall, R. C., and Gurd, F. R. (1975) J. Biol. Chem. 248, 7811–7818
33. Norton, R. S. (2002) in Perspectives in Molecular Toxicology (Ménez, A., ed) pp. 159–173, John Wiley & Sons, Inc., New York
34. Wakamatsu, K., Kohda, D., Hatanaka, H., Lancelin, J. M., Ichida, Y., Oya, M., Nakamura, H., Inagaki, F., and Sato, K. (1992) Biochemistry 31, 12577–12584
35. Hill, J. M., Alewod, P. F., and Craik, D. J. (1996) Biochemistry 35, 8824–8835
36. Mitchell, S. S., Shon, K. J., Foster, M. P., Davis, D. R., Gobus, B. M., and Ireland, C. M. (1998) Biochemistry 37, 1215–1220
37. Hill, J. M., Alewod, P. F., and Craik, D. J. (2000) Eur. J. Biochem. 267, 4649–4657
38. Hill, J. M., Alewod, P. F., and Craik, D. J. (1997) Structure 5, 571–583
39. Prorok, M., Warder, S. E., Blandl, T., and Castellino, F. J. (1996) Biochemistry 35, 16528–16534
40. Mourier, G., Servent, D., Zinn-Justin, S., and Ménez, A. (2000) Protein Eng. 13, 217–225
41. Nielsen, K. J., Thomas, L., Lewis, R. J., Alewod, P. F., and Craik, D. J. (1996) J. Mol. Biol. 267, 297–310
42. Pallaghy, P. K., and Norton, R. S. (1999) J. Pept. Res. 53, 343–351
43. Savarin, P., Gueneugueux, M., Gilsain, B., Lamthanh, H., Gasparini, S., Zinn-Justin, S., and Ménez, A. (1998) Biochemistry 37, 5407–5416
44. Sato, K., Ishida, Y., Wakamatsu, K., Kato, R., Honda, H., Ohizumi, Y., Nakamura, H., Oya, M., Lancelin, J. M., and Inagaki, F. (1991) J. Biol. Chem. 266, 16989–16991
45. Vita, C., Drakopoulou, E., Vizzavona, J., Rochette, S., Martin, L., Ménez, A., Roumestand, C., Yang, Y. S., Tsissiastigi, L., Benjoud, A., and Gluckman, J. C. (1999) Proc. Natl. Acad. Sci. U. S. A. 96, 13091–13096
Structure of a Novel P-superfamily Spasmodic Conotoxin Reveals an Inhibitory Cystine Knot Motif

Luke A. Miles, Catherine Y. Dy, Jake Nielsen, Kevin J. Barnham, Mark G. Hinds, Baldomero M. Olivera, Grzegorz Bulaj and Raymond S. Norton

J. Biol. Chem. 2002, 277:43033-43040.
doi: 10.1074/jbc.M206690200 originally published online August 21, 2002

Access the most updated version of this article at doi: 10.1074/jbc.M206690200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2002/11/04/277.45.43033.DC1

This article cites 43 references, 4 of which can be accessed free at http://www.jbc.org/content/277/45/43033.full.html#ref-list-1