Three-Dimensional Structure of a γ-Carboxyglutamic Acid-Containing Conotoxin, Conantokin G, from the Marine Snail *Conus geographus*: The Metal-Free Conformer[†,‡]

Alan C. Rigby,[§,||] James D. Baleja,[*,||] Barbara C. Furie,[§,||,⊥] and Bruce Furie[*,§,||,⊥]

Marine Biological Laboratory, Woods Hole, Massachusetts 02543, and Center for Hemostasis and Thrombosis Research, Division of Hematology-Oncology, New England Medical Center, and Department of Medicine and Department of Biochemistry, Tufts University School of Medicine, Boston, Massachusetts 02111

Received February 12, 1997; Revised Manuscript Received April 3, 1997[⊗]

ABSTRACT: Conantokin G is a γ-carboxyglutamic acid-containing conotoxin from the venom of the marine cone snail *Conus geographus*. The 17-residue peptide, which contains five γ-carboxyglutamic acid (Gla) residues and an amidated C-terminal asparagine amide, was synthesized chemically in a form identical to the natural conantokin G. To gain insight into the role of γ-carboxyglutamic acid in the structure of this peptide, we determined the three-dimensional structure of conantokin G by ^{1}H NMR and compared its structure to other conotoxins and to the γ-carboxyglutamic acid-containing regions of the vitamin K-dependent blood-clotting proteins. Complete resonance assignments were made by two-dimensional ^{1}H NMR spectroscopy in the absence of metal ions. NOE cross-peaks $d_{\alpha N}$, d_{NN}, and $d_{\beta N}$ provided interproton distance information, and vicinal spin−spin coupling constants $^3J_{HN\alpha}$ were used to calculate ϕ torsion angles. Distance geometry and simulated annealing methods were used to derive 20 convergent structures from a set of 227 interproton distance restraints and 13 torsion angle measurements. The backbone rmsd to the geometric average for 20 final structures is 0.8 ± 0.1 Å. Conantokin G consists of a structured region commencing at Gla 3 and extending through arginine 13. This structure includes a partial loop centered around Gla 3 and Gla 4, a distorted type I turn between glutamine 6 and glutamine 9, and two type I turns involving Gla 10, leucine 11, and isoleucine 12 and arginine 13. Together, these two turns define approximately 1.6 turns of a distorted 3_{10} helix. The observed structure possesses structural elements similar to those seen in the disulfide-linked conotoxins.

γ-Carboxyglutamic acid (Gla) is a unique amino acid that is synthesized via a posttranslational mechanism that involves the γ-carboxylation of specific glutamic acid residues. This modification of a substrate polypeptide is catalyzed by the vitamin K-dependent γ-glutamyl carboxylase in a reaction that requires molecular oxygen, carbon dioxide, and reduced vitamin K (Furie & Furie, 1990; Suttie, 1993). Although this amino acid may be broadly distributed, its functional and structural role has been most completely characterized in three distinct protein classes: (1) the mammalian vitamin K-dependent blood-clotting and regulatory proteins, including prothrombin, factor IX, factor VII, factor X, protein C, and protein S (Furie & Furie, 1988); (2) mammalian proteins of mineralized tissue, including osteocalcin (bone Gla protein) and matrix Gla protein (Hauschka et al., 1989); and (3) certain peptides in the venom of the marine snail of the genus *Conus* (Olivera et al., 1990). Since its discovery in prothrombin in 1974 (Stenflo et al., 1974; Nelsestuen et al., 1974), γ-carboxyglutamic acid has been shown to be a metal binding amino acid (Sperling et al., 1978) which is required for the phospholipid membrane binding properties of the vitamin K-dependent blood coagulation proteins (Esmon et al., 1975). The determination of the three-dimensional structure of the γ-carboxyglutamic acid-rich domain (Gla domain)−Ca^{2+} complex of bovine prothrombin (Soriano-Garcia et al., 1992) demonstrated the importance of calcium ions bound to γ-carboxyglutamic acid in stabilizing a unique protein conformer. Similarly, the Gla domain of factor VIIa in the presence of calcium ions exhibits a homologous structure while bound to tissue factor (Banner et al., 1996). Calcium ions perform a parallel role in defining the structure of factor IX (Freedman et al., 1995a,b), with stabilization of a conformer that exposes the phospholipid binding site (Freedman et al., 1996). The function of γ-carboxyglutamic acid in proteins from mineralized tissues is less certain, but it may be involved in the regulation of bone formation (Ducy et al., 1996). The function of γ-carboxyglutamic acid in the conotoxins from the marine snail is unknown, although the presence of γ-carboxyglutamic acid has been shown to be a requirement for biological activity of these peptides (McIntosh et al., 1984; Chandler et al., 1993; Zhou et al., 1996).

Cone snails (*Conus*) are widely dispersed venomous marine gastropods which use biologically active peptides to paralyze fish, marine worms, and mollusks. The peptide toxins, known as conotoxins, are injected into their prey (Olivera et al., 1990, 1991; Cruz et al., 1992). Venom is synthesized in the venom duct, a tissue that contains abundant

[†] This work was supported in part by funds from the Marine Biological Laboratory and by grants from the National Institutes of Health (HL38216 and HL42443).

[‡] Atomic coordinates have been deposited with the Brookhaven Protein Data Bank under file name 1ad7.

[§] New England Medical Center and Department of Medicine, Tufts University School of Medicine.

[||] Department of Biochemistry, Tufts University School of Medicine.

[⊥] Marine Biological Laboratory.

[⊗] Abstract published in *Advance ACS Abstracts,* May 15, 1997.

S0006-2960(97)00321-8 CCC: $14.00

quantities of γ-carboxyglutamic acid (Hauschka et al., 1988). Many of these toxins have been isolated, and their neuropharmacology has been defined. As a family, most are small peptides characterized by a high density of disulfide bonds. Cysteine can represent up to 50% of the amino acid residues in some conotoxins. In addition, these sequences are characterized by C-terminal amides. These peptides target nicotinic acetylcholine receptors, sodium channels, and calcium channels and thus behave as potent receptor antagonists. Three major conopeptide classes are defined by the disulfide framework: the four-loop ω-conopeptides, the three-loop μ-conopeptides, and the two-loop α-conopeptides. The three-dimensional structures of a number of conotoxins have been determined by two-dimensional ^{1}H NMR spectroscopy (Kobayashi et al., 1989; Pardi et al., 1989; Pallaghy et al., 1993; Farr-Jones et al., 1995; Kohno et al., 1995; Hill et al., 1996) and X-ray crystallography (Guddat et al., 1996). A feature of the structures of these short polypeptides is the presence of distorted 3_{10} helices and type I β turns which are characterized by atypical φ and ψ angles.

In contrast, the conantokin G (sleeper peptide) from the cone snail *Conus geographus* lacks disulfide bonds but contains γ-carboxyglutamic acid (McIntosh et al., 1984). Conantokin G and conantokin T (Haack et al., 1990), (the latter isolated from *Conus tulipa*) are peptide neurotoxins of 17 and 21 residues, respectively. A conopeptide from *Conus textile* contains both γ-carboxyglutamic acid and cysteine (Fainzilber et al., 1991; Nakamura et al., 1996). Chemically synthesized peptides based upon the conantokin sequences are pharmacologically active neurotoxins (Rivier et al., 1987; Haack et al., 1990). Substitution of γ-carboxyglutamic acid by glutamic acid in chemically synthesized conantokin analogs yields pharmacologically inert peptides (Chandler et al., 1993). Conantokin G and conantokin T in their deamidated forms bind to calcium ions, although the role of calcium ions in biological function is unknown (Prorok et al., 1996). The recent report of a conantokin G−Ca^{2+} complex structure further indicates a role for metal ions in the stabilization of secondary structure (Skjaerbaek et al., 1997). Thus, γ-carboxyglutamic acid is critical to the biologic activity of these peptides, but the function of γ-carboxyglutamic acid remains uncertain.

In order to explore the structural role and functional contribution of γ-carboxyglutamic acid in proteins other than those involved in blood coagulation, we have studied the biosynthesis and the role of γ-carboxyglutamic acid in the marine cone snail. Venom duct from *Conus* contains a vitamin K-dependent carboxylase that shows functional similarity with the bovine carboxylase (Czerwiec et al., 1996). To elucidate the role of γ-carboxyglutamic acid in conotoxins, we have determined the three-dimensional structure of a synthetic form of conantokin G that has a chemical structure identical to that of conantokin G derived from *C. geographus*. We demonstrate that in the absence of metal ions most of this peptide is highly structured. The γ-carboxyglutamic acid residue side chains are exposed to solvent, suggesting a role for their malonate-like side chains in the formation of ion pairs with the conantokin receptor or in the interaction with metal ions.

MATERIALS AND METHODS

Peptide Synthesis and Identification. Conantokin G was synthesized using solid phase Fmoc [*N*-(9-fluorenyl)methoxycarbonyl/*N*-methylpyrrolidone] chemistry on an Applied Biosystems model 430A peptide synthesizer as described previously (Jacobs et al., 1994). However, to obtain C-terminal amidation, rink amide MBHA resin was used (NovaBiochem) and asparagine was double-coupled to the MBHA resin under standard conditions. The peptide was cleaved from the resin in triethylsilane/1,2-ethanedithiol/TFA (5:5:90) for 2.5 h at 25 °C with constant stirring. The peptide sample was lyophilized and then purified by reverse phase HPLC using a preparative reverse phase C_{18} Bio-Rad Hi-Pore 318 column (RPC_{18}, 21.5 mm × 250 mm) and a Beckman System Gold HPLC system. The column was developed with a linear gradient from 10 to 40% acetonitrile in the presence of 0.1% trifluoroacetic acid at a flow rate of 8.0 mL/min and monitored at 214 nm. The major peak containing 75% of the applied peptide was pooled and rechromatographed on a reverse phase HPLC system using a Vydac 218TP 5 μm column (4.6 mm × 250 mm) using the same linear gradient at a flow rate of 1.0 mL/min. The purified peptide was subjected to MALDI-TOF mass spectroscopy on a Voyager Linear MALDI-TOF spectrometer (PerSeptive Biosystems). The mass spectroscopy analysis was performed with a nitrogen laser at 337 nm, employing either linear mode positive or negative ionizations. The accelerating voltage was 30 kV, and spectra were generated from the sum of 37 averaged scans. The peptide was sequenced by automated Edman degradation using a Milligen 6600 Prosequencer.

NMR Spectroscopy. The sample contained 7.5 mg of conantokin G dissolved in 450 μL of H_2O (7.02 mM, pH 5.60) and 25 mM d_4-acetate with 10% D_2O as the deuterium lock signal. Samples prepared in D_2O were initially lyophilized from 99.8% D_2O and then redissolved in 99.96% D_2O. One-dimensional spectra were acquired at 25 °C with 16K real data points, 256 summed scans, and a spectral width of 8064.52 Hz and then processed by applying a squared sine bell window function shifted by 30°. All samples were pretreated with Chelex 100 to remove trace metal ions. Spectra were collected on a Bruker AMX-500 spectrometer with a proton frequency of 500.14 MHz. The carrier frequency was set on the water resonance, which was suppressed using presaturation or by using jump-and-return methodology (Plateau & Guéron, 1982). NOESY (Jeener et al., 1979) spectra were recorded with mixing times of 150, 250, and 500 ms at 25 °C and 250 ms at 10 °C. A total of 2048 (or 4096) real data points were acquired in t_2, with 512 TPPI (or States−TPPI) increments in t_1, with a spectral width of 8064.52 Hz in the F_2 dimension. A total of 128 summed scans was collected with a relaxation delay of 1.3 s. Spectra were multiplied with a sine bell window function shifted by 30° in t_2 (applied over 1024 points) and a sine bell window function shifted by 30° in t_1 (applied over all 512 acquired points) and zero filled to a 2K by 1K (real) matrix using the Bruker NMR processing program. NOESY cross-peak intensities were converted into three distance classes (1.7−3.0 Å, strong; 1.7−4.0 Å, medium; and 3.0−5.0 Å, weak) and calibrated using published methods (Hyberts et al., 1992). Non-stereospecifically assigned atoms were treated as pseudoatoms and given correction distances

(Wüthrich, 1986). Distance restraint information was extracted from NOESY spectra with different mixing times. Comparison of the short and long mixing time spectra was used to control for spin diffusion effects. TOCSY (Braunschweiler & Ernst, 1983) spectra were recorded and processed using parameters identical to those of the NOESY, the exception being that a 35 ms mixing time was used in collecting 128 summed scans employing the MLEV-17 spin lock sequence (Bax & Davis, 1985). A DQF-COSY (Rance et al., 1983) spectrum was recorded with 4096 real t_2 points, 40 summed scans, and 768 TPPI increments. The spectrum was multiplied by sine bell window functions shifted by 30° in t_2 and 30° in t_1 and zero filled to a 2K by 1K (real) matrix. In addition, a similar set of experiments (DQF-COSY, TOCSY, and NOESY) was performed on a 10 mM peptide sample in 99.96% D_2O at a noncorrected pH of 5.60.

Sequence-specific resonance assignments were made in two stages. The intraresidue spin systems were identified using the $^1H-^1H$ through-bond connectivities found in TOCSY and DQF-COSY spectra. The sequential assignment of residues was completed on the basis of sequential $d_{\alpha N}$ and d_{NN} NOE connectivities (Wüthrich, 1986). The NOESY spectrum collected on a sample in D_2O was used to reveal protons that were near the resonance of H_2O protons. The NOE contacts were classified into "intraresidue" for NOEs within a residue, "sequential and short range" for contacts between the backbone and side chain protons of residue i with residue $i + 1$, "medium range" for i to $i + 2$ connnectivities, and "long range" for NOEs between protons on residues separated by three or more amino acid positions in the sequence.

The vicinal spin−spin coupling constants $^3J_{HN\alpha}$ were used to calculate ϕ torsion angles (Pardi et al., 1984). The coupling constants were measured from the splitting of amide cross-peaks in a NOESY spectrum that was resolution-enhanced by multiplying with a squared sine bell window function shifted by 20° and applied over 2048 (real) points in t_2. Structure determination used a set of 227 distance restraints (204 intraresidue and sequential, 23 medium and long range) and 13 ϕ angles. A combination of distance geometry and simulated annealing methods (Havel, 1991) was used to generate 25 structures (convergence of 20) using the DGII program of InsightII (Biosym Technologies, San Diego, CA). The simulated annealing protocol has been described elsewhere (Freedman et al., 1995a). The final total error function value was 0.08 ± 0.002 kcal/mol. Nonconverged structures had energies of >0.20 kcal/mol. The average structure for the ensemble was calculated using the Analysis program of InsightII. Average root-mean-square deviation (rmsd) values following superimposition of the backbone atoms of each structure with the geometric average reflected the quality of the structures determined. In addition, the coherence of torsion angles among different structures was evaluated. The average torsion angle was measured by a vector addition method (Hyberts et al., 1992). An order parameter, S, equals 1.0 when the torsion angle is the same in all structures, whereas an order parameter near zero indicates disorder at that position.

RESULTS

The conantokin G peptide was prepared by solid phase Fmoc peptide synthesis. The rink amide MBHA resin was used in a procedure to ensure amidation of the carboxy-terminal asparagine. The peptide was cleaved under conditions that preserve the C-terminal asparagine amide since amidation is required for biological activity of the peptide (Chandler et al., 1993). This synthetic 17-residue peptide corresponds to the entire polypeptide sequence of conantokin G as isolated from the venom of *C. geographus*: NH_2-Gly-Glu-Gla-Gla-Leu-Gln-Gla-Asn-Gln-Gla-Leu-Ile-Arg-Gla-Lys-Ser-Asn-NH_2. Following purification of the peptide, the synthetic conantokin G was analyzed by analytical reverse phase HPLC. The HPLC chromatogram (Figure 1A) indicates the presence of a single dominant peptide species, eluting at 27% acetonitrile and accounting for greater than 95% of the area under the five peaks observed. This peptide was subjected to automated Edman degradation, where the expected amino acid sequence was confirmed. MALDI-TOF analysis of this peptide in the positive ion mode produced a series of peaks, with the largest molecular mass determined to be 2268.4 Da (calculated, M − H = 2264.2 Da) for the fully carboxylated peptide. The other peaks seen in the positive ion mode were displaced by 44 Da each. Since decarboxylation of a single Gla lowers the molecular mass by 44 Da, the other species represent partially decarboxylated peptides generated during mass spectroscopy. The mass of the completely decarboxylated peptide was determined using linear mode negative ionization. A single peak with a

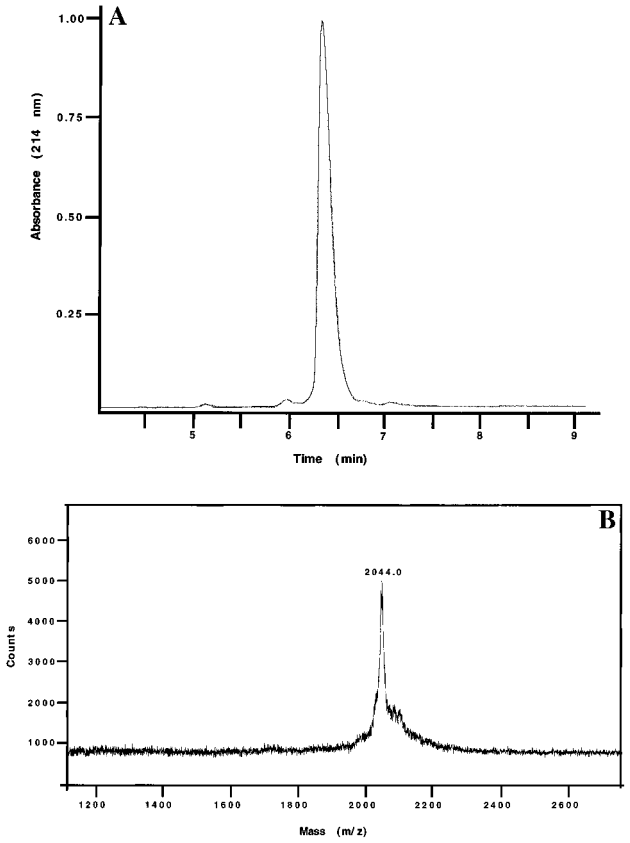

FIGURE 1: Characterization of conantokin G. (A) HPLC chromatogram of conantokin G. The peptide was isolated from the cleavage reaction mixture by preparative reverse phase HPLC and then rechromatographed using a Vydac C_{18} reverse phase column. The peptide, which eluted at 27% acetonitrile (0.1% TFA), was subjected to amino acid sequencing and mass spectroscopy. (B) Linear mode, negative ion MALDI spectra of conantokin G. The molecular mass is 2044 Da, comparable to the theoretical mass of 2044.2 Da for the decarboxylated peptide. Under the conditions of analysis, Gla-containing peptides are decarboxylated.

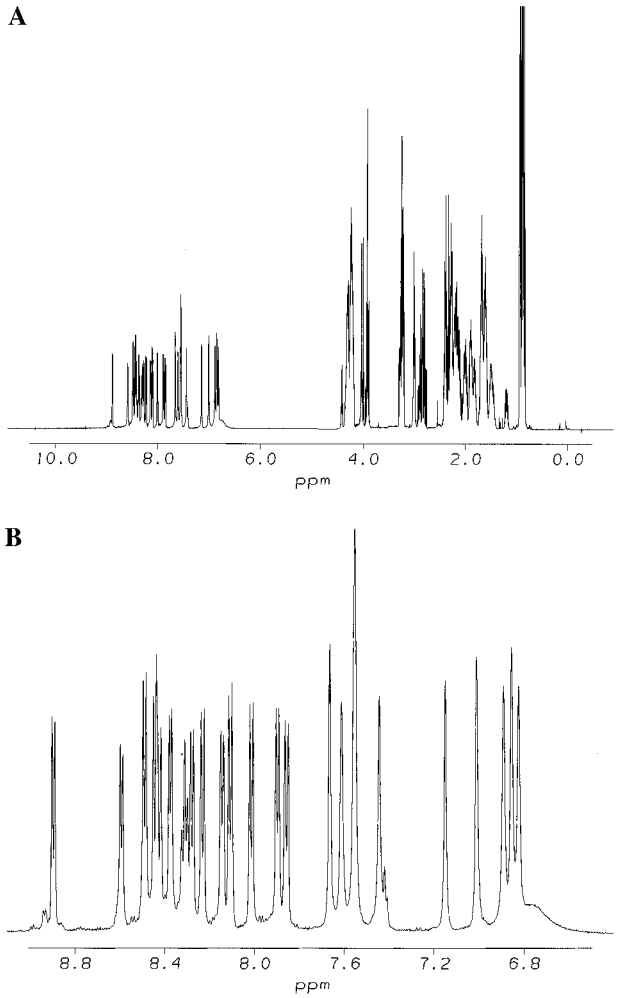

FIGURE 2: One-dimensional ^{1}H NMR spectra of conantokin G. The sample contained 3.6 mM metal-free conantokin G and 25 mM d_4-acetate in H_2O/D_2O (90:10) at pH 5.6. The spectrum was acquired with 256 scans. (A) One-dimensional spectrum of conantokin G. (B) Detail of the downfield NH region of the spectrum in panel A. The spectrum was processed with a sine bell window function shifted by 30°, over 16K real data points, and the solvent peak was removed with a sine bell convolution function.

molecular mass of 2044.0 Da (calculated, M − H = 2044.2 Da) was observed (Figure 1B). These results identify the peptide as a homogeneous species of the correct amino acid sequence and molecular mass.

The peptide was rendered metal-free (apo) by treatment with Chelex 100 prior to NMR spectral analysis. Repeated treatment with Chelex 100 did not alter the NMR spectra. The one-dimensional NMR spectra of the metal-free conantokin G at pH 5.6 were collected at 25 °C (Figure 2A). The proton resonances are spectrally disperse, including resonances within the amide region (Figure 2B). This indicates the presence of an ordered structure. Many of the amide and α proton resonances are shifted from their random coil values (Wüthrich, 1986). For example, the ^{1}Hα resonances for amino acids 9 through 14 possess chemical shifts that are shifted upfield by greater than 0.1 ppm from random coil values seen in unstructured peptides (Wüthrich, 1986). Similarly, the amide protons are not present as a single broad resonance of spectrally degenerate peaks near 8.4 ppm but appear to be dispersed between 7.8 and 9.0 ppm, shifted from their random coil values (Wüthrich, 1986). The amide protons are in chemically distinct environments as a result of structural features inherent in the metal-free

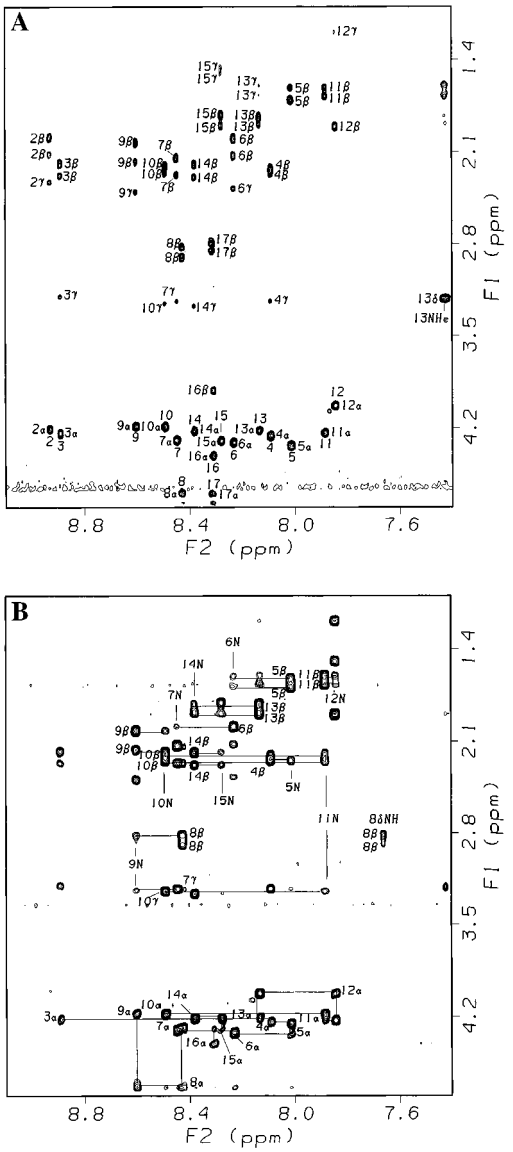

FIGURE 3: Two-dimensional ^{1}H NMR spectra of conantokin G. (A) A region of the two-dimensional TOCSY spectrum collected with a 35 ms mixing time. Intraresidue side chain connectivities are illustrated for the various amino acids, identified by residue number and proton type. Sample conditions are given in Figure 2. The data were processed as described in Materials and Methods. (B) The same region of a two-dimensional NOESY spectrum (500 ms mixing time). Line connectivities are illustrated between intraresidue and sequential α−NH (i,i+1) cross-peaks throughout the peptide, and the intraresidue NH−α connectivities are labeled with the amino acid number. In addition, interresidue and short range connectivities are identified.

conantokin G peptide. The spectral dispersion within the amide proton region of the spectra (Figure 2B) facilitated measurement of $^3J_{NH\alpha}$ coupling constants, which were used to calculate ϕ torsion angle constraints.

This peptide sample was employed to determine the structure of conantokin G in the absence of metal ions by two-dimensional homonuclear ^{1}H NMR spectroscopy. Standard NOESY, TOCSY, and DQF-COSY data were collected at pH 5.6 and 25 °C, facilitating the assignment of spin system resonances (Bax & Davis, 1985; Wüthrich, 1986). To remove overlapping resonance ambiguity, NOESY spectra were also collected at 10 °C. In the TOCSY experiment, a mixing time of 35 ms was used, permitting successive single, double, and multiple through-bond connectivities to

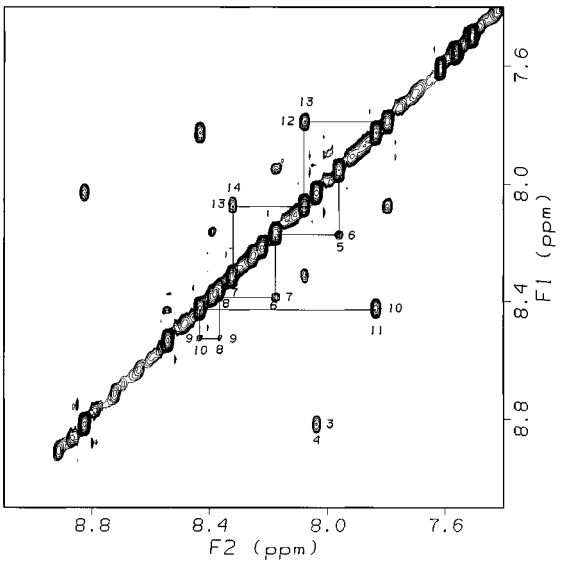

FIGURE 4: Amide−amide region of the NOESY spectrum of conantokin G. Sample conditions are described in Figure 2. All cross-peaks involving sequential amide protons are labeled with the residue number. The two regions of amide−amide connectivities are delineated above and below the diagonal with the C-terminal region commencing at Gla 10 and extending through Gla 14 identified above the diagonal (the leucine 11−isoleucine 12 cross-peak is obscured by the diagonal).

be observed. The TOCSY spectra allowed the identification of entire spin systems for all residues including the five γ-carboxyglutamic residues located at positions 3, 4, 7, 10, and 14 (Figure 3A). Using this information, all remaining non-intraresidue, sequential, and nonsequential NOE cross-peaks could be identified (Figure 3B). In the amide−amide region of the NOESY spectrum (Figure 4), sequential NH−NH contacts were identified from residue 3 and extending through to residue 14, with interruptions at Gla 4−leucine 5 and leucine 11−isoleucine 12. The strongest sequential NH−NH connectivities were identified between Gla 3 and Gla 4 and between Gla 10 and Leu 11, Ile 12 and Arg 13, and Arg 13 and Gla 14. The assignment of all proton resonances facilitated the identification of longer range interactions from the NOESY spectrum. The proton resonance assignments for the conantokin G peptide are presented in Table S-1 of the Supporting Information. Coupling constants between vicinal α and amide protons were measured to identify ϕ torsion angle constraints. These angles were restrained to the negative ϕ angle region due to the size of the coupling constant and the intensity of the intraresidue and sequential NOE cross-peaks (Ludvigsen & Poulsen, 1992).

The structure of conantokin G was determined by converting NOESY cross-peaks into a set of interproton distance restraints. This was achieved using the number of contours within a cross-peak to establish upper and lower limit distances and internally calibrated using the NH−α distances as standards (Hyberts et al., 1992). Structure calculations by distance geometry and simulated annealing methods employing the NMR data were used to generate 20 final structures. This data set included 227 distance restraints (intraresidue, 126; sequential, 78; and medium and long range, 23) and 13 torsion angle restraints. Of 25 generated structures, the 20 structures which converged are overlayed in Figure 5. The pairwise rmsd to the average structure is 0.8 ± 0.1 Å for the well-defined backbone atoms between residues 4 and 13 of the 20 final structures. To assess the superimposition of the structures, average torsion angles were obtained from the values of each structure, following vector addition methodology (Hyberts et al., 1992). The correlation and mean ϕ and ψ torsion angles for the family of 20 structures were measured at each residue (Hyberts et al., 1992). Many of the ϕ and ψ angles are clustered in the negative ϕ and positive ψ angle region, suggesting type I and II turns (Richardson, 1981). These backbone torsion angles possessed correlation values (*S*) near 1, indicative of an ordered, well-defined region from residues 4 to 14. The

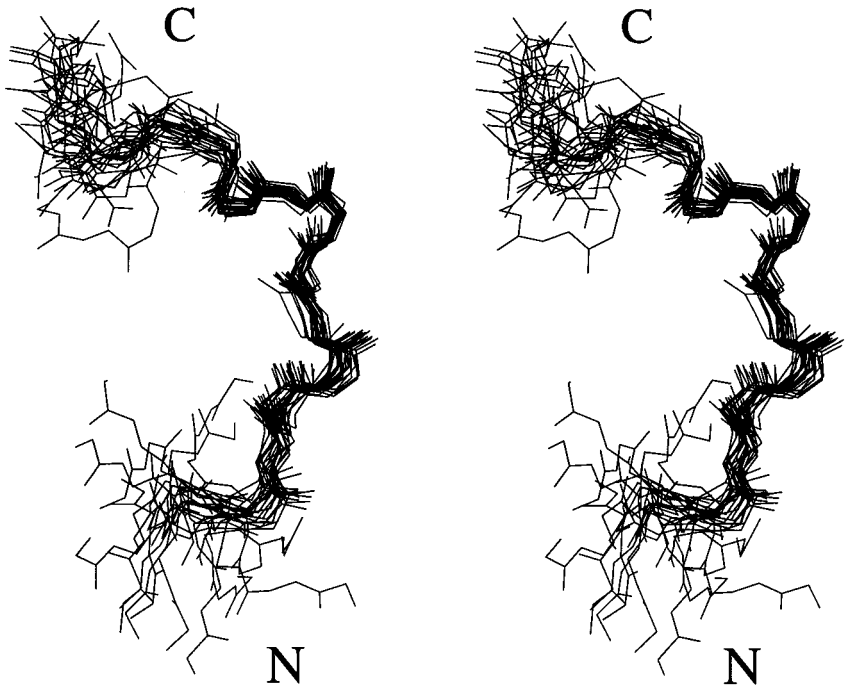

FIGURE 5: Stereoview overlay of the 20 calculated structures for conantokin G. All structures are shown superimposed using the backbone atoms of the well-defined residues with the geometric average. The polypeptide backbone root-mean-squared deviation compared to the average is 0.8 ± 0.1 Å. The N and C termini are labeled.

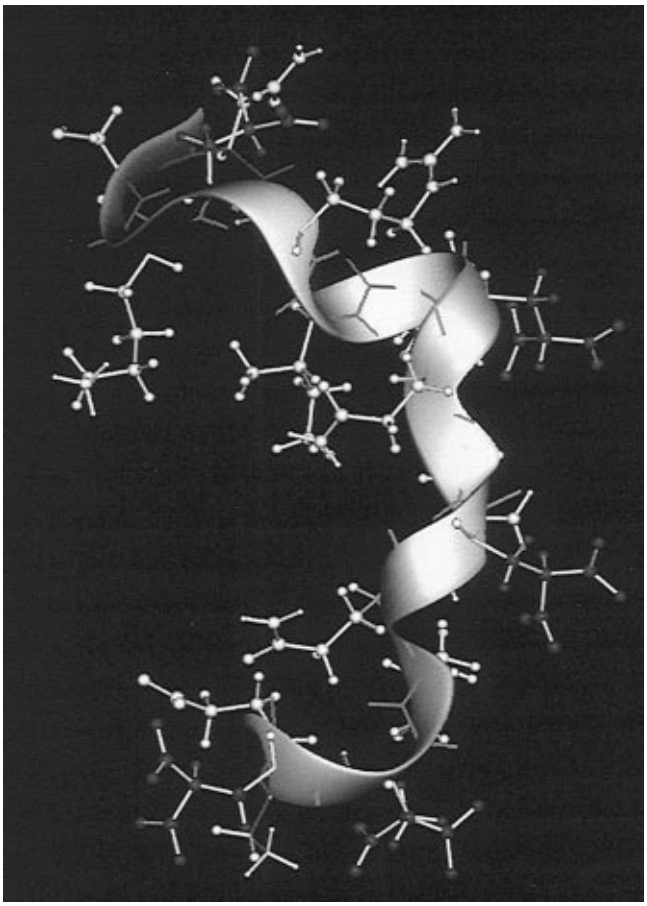

FIGURE 6: Ribbon diagram of conantokin G illustrating secondary structural features. The backbone of the lowest-energy structure of metal-free conantokin G is represented as a ribbon. The structure is oriented with the N terminus pointing down. The Gla residues are red, while all other residues are white. The loosely defined loop centered around Gla 3 and Gla 4 can be seen, as can the distorted type I β turn between glutamine 6 and glutamine 9 and the 3_{10} helix identified at the C terminus (Gln 9−Gla 14).

N and C termini have low correlation (*S*) values, consistent with the lack of structure in these regions.

Qualitative analyses of short and medium range NOEs, $^3J_{NH}$ coupling constants, and ϕ and ψ angles were used to characterize secondary structural elements of conantokin G (Figure 6). A strong amide−amide connectivity between Gla 3 and Gla 4 combined with $^3J_{HN\alpha}$ coupling constants of less than 6 Hz at residues 2 and 3 defines a loop region with partial tight β turn or helical characteristics. Poorly defined ϕ and ψ angles within this region, and a lack of NOEs, are due to flexibility of the N terminus.

A secondary structure best classified as a distorted type I β turn is located between glutamine 6 and glutamine 9 (Figure 6) (Richardson, 1981; Wüthrich, 1986). Weak sequential amide−amide connectivities are observed in this region, and the presence of a type I turn is supported by the $^3J_{HN\alpha}$ coupling constants measured for asparagine 8 ($^3J_{HN\alpha} > 8$ Hz) and for glutamine 9 ($^3J_{HN\alpha} < 6$ Hz). Although these residues are well-defined in terms of average backbone rmsd and angular order (*S*) correlations, the average ϕ and ψ angles within this segment are atypical for type I β turns ($\phi = -57 \pm 40°$, $\psi = 0$ to -180) (Richardson, 1981). In the average structure, a hydrogen bond is observed between the carbonyl oxygen of glutamine 6 and the amide nitrogen of asparagine 8, possibly explaining the strained ϕ and ψ torsion angles in this distorted turn.

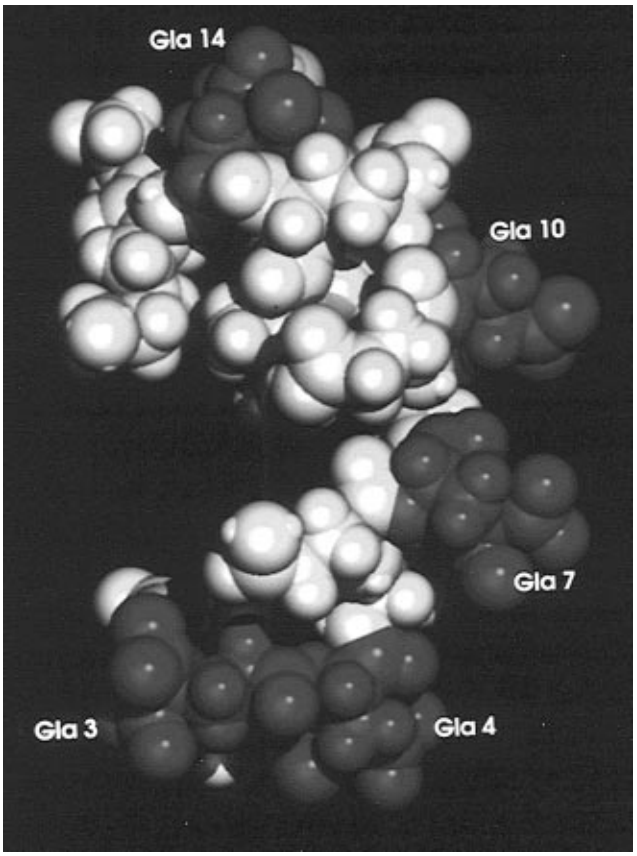

FIGURE 7: Space-filling model of conantokin G emphasizing the orientation of the Gla residues. The lowest-energy structure is illustrated as a space-filling cpk model. The five Gla residues are red and are labeled with the residue number. The Gla residues show a preferential solvent-exposed location on the same face of the conantokin G structure.

Two additional type I β turns (Richardson, 1981) centered around Gla 10 and leucine 11 and isoleucine 12 and arginine 13 were identified (Figure 6). The average ϕ and ψ torsion angles within this region are within the energetically favorable regions defined by a Ramachandran plot for helical elements and type I β turns (Richardson, 1981). The presence of strong amide−amide NOEs between Gla 10 and Gla 14 (the leucine 11−isoleucine 12 amide−amide is obscured by the diagonal) and side chain 9α−12β (*i*,*i*+3) NOEs within this region suggests this region is helical. The distortion in the helix centers on isoleucine 12 despite the fact that is has well-defined ϕ and ψ angles. The presence of a helical element is substantiated by $^3J_{HN\alpha}$ coupling constants of <6 Hz measured for residues 9−11, 13, and 14. Additional support for the presence of a helical element is provided by analysis of Hα chemical shifts for sequential residues 9−14 (Wishart et al., 1991). The Hα resonances of these residues show characteristic upfield helical shifts from their random coil values.

The orientation of the five γ-carboxyglutamic acid residues in conantokin G can be observed in a space-filling model (Figure 7). The malonate side chains of the five γ-carboxyglutamic acid residues (red) are on the solvent-exposed surface and are dispersed uniformly on the outer aspect of the folded peptide. The hydrophobic residues of conantokin G, located at residues 5, 11, and 12, are sequestered from the solvent and are the likely cause of the inherent distortion of the peptide backbone. The N-terminal cluster of two γ-carboxyglutamic acids and a glutamic acid may be a critical

component of the neurotoxin and its recognition by the receptor.

DISCUSSION

Since the discovery of γ-carboxyglutamic acid in bovine prothrombin in 1974, the presence of this amino acid in other organisms has been extensively investigated. An intensive search for γ-carboxyglutamic acid has demonstrated this amino acid in two disparate phyla: vertebrates and mollusks within the genus *Conus*. However, three indirect lines of evidence suggest a broad functional role for this amino acid and a global distribution: (1) preservation through evolution of the synthetic machinery for generating γ-carboxyglutamic acid from glutamic acid, (2) the recognition in mammalian biology that γ-carboxyglutamic acid and vitamin K are involved in processes beyond blood coagulation (Ducy et al., 1996), and (3) the presence of γ-carboxylase activity in most mammalian tissues (Vermeer et al., 1995). We have recently demonstrated that cone snails have a carboxylation system with properties similar to that of mammals. Their venom duct contains substantial amounts of γ-glutamyl carboxylase activity (Czerwiec et al., 1996). This activity has an absolute requirement for vitamin K, and substrates containing the human prothrombin propeptide bind tightly to this enzyme (Czerwiec et al., 1996). Our hypothesis is that a general carboxylation system is widely distributed and globally required in the animal kingdom for essential and diverse biological processes. By determining the structure of γ-carboxyglutamic acid-containing conotoxins, we seek to broaden our understanding of the structural and functional role that γ-carboxyglutamic acid plays in protein−membrane and protein−protein interaction.

γ-Carboxyglutamic acid is a metal binding amino acid (Sperling et al., 1978) that plays a critical role in defining the membrane binding properties of vitamin K-dependent blood-clotting and regulatory proteins. From the work of a number of laboratories [for review, see Furie and Furie (1988)], γ-carboxyglutamic acid in the blood coagulation proteins has been shown to be required for membrane binding. Prothrombin, factor VII, and factor IX are glycoprotein components of the blood coagulation system and are zymogens of serine proteases (Furie & Furie, 1988). These proteins have been prototypes for the study of the calcium-binding proteins that contain γ-carboxyglutamic acid. The N-terminal γ-carboxyglutamic acid domain of prothrombin includes 10 γ-carboxyglutamic acid residues. The X-ray crystal structure of bovine prothrombin has revealed that many of these γ-carboxyglutamic acid residues coordinate internal calcium ions (Soriano-Garcia et al., 1992). A homologous structure was evident in the X-ray crystal structure of factor VIIa (Banner et al., 1996). The N-terminal γ-carboxyglutamic acid domain of factor IX includes 12 γ-carboxyglutamic acid residues. We have recently studied the effect of metal ions, including calcium ions, in defining the stable conformation of the Gla domain of factor IX (Freedman et al., 1995a,b, 1996). Within the Gla domain, chelation of calcium by critical γ-carboxyglutamic acid residues stabilizes a unique three-dimensional structure that defines the phospholipid binding site. This site is formed within the N-terminal region of the Gla domains and includes residues 1−11. This phospholipid binding site includes a hydrophobic patch that is either within or in the immediate proximity (<3 Å) of the membrane (Freedman et al., 1996).

Although the role of γ-carboxyglutamic acid in the vitamin K-dependent blood-clotting and regulatory proteins has now been well defined, the contribution of γ-carboxyglutamic acid to the function of other vitamin K-dependent proteins is poorly understood. Specifically, the presence of γ-carboxyglutamic acid-containing proteins in mineralized tissue and in the cone snail venom peptides has been known for many years, but the role of γ-carboxyglutamic acid remains unknown. In order to shed light on γ-carboxyglutamic acid function in the neurotoxins of the cone snail, we have solved the structure of conantokin G. Conantokin G is a biologically potent constituent of the venom of *C. geographus*. This peptide induces a sleep state in young mice following intracerebral injection (McIntosh et al., 1984; Rivier et al., 1987). Conantokin G is a *N*-methyl D-aspartate (NMDA) antagonist and binds to the NMDA receptor (Haack et al., 1990; Mena et al., 1990; Skolnick et al., 1992). Unlike other classes of conotoxins, the conantokins are relatively short peptides that lack cysteine residues and disulfide bonds. Conantokin G is a metal binding peptide that interacts weakly with calcium ions (Prorok et al., 1996). The carboxyl groups of γ-carboxyglutamic acid residues are likely ligands for bound calcium ions. However, a requirement for calcium ions in the biological activity of conantokin G has not been shown, and thus, the importance of a biologically active conantokin−calcium complex remains speculative. Secondary structural calculations for conantokin G based upon circular dichroism have predicted both the presence (Chandler et al., 1993; Zhou et al., 1996) and absence (Prorok et al., 1996) of significant secondary structure in the absence of metal ions; helical structure has been observed by all groups in the presence of calcium ions. The basis of this apparent discrepancy is unclear but may be related to metal contamination of a peptide solution thought to be metal-free (Chandler et al., 1993; Zhou et al., 1996) or to the use of a form of conantokin G that lacks the C-terminal amide (Prorok et al., 1996). Asparagine 17 is amidated in the natural conantokin G and is required for receptor antagonism (Chandler et al., 1993). We report the three-dimensional structure of conantokin G in the absence of metal ions. Conantokin G is a 17-residue peptide with a well-defined stable structure, despite the absence of disulfide bonds or intramolecular interactions stabilized by metal ions. The solution structure is composed of an N-terminal loop region centered around Gla 3 and Gla 4 and a distorted type I β turn between glutamine 6 and glutamine 9, abutted to a distorted 3_{10} helix from glutamine 9 to γ-carboxyglutamic acid 14. This helical element is distorted in that the typical hydrogen bonding patterns seen in 3_{10} helices and α-helices are not observed in the structure of metal-free conantokin G. The recently published structure of conantokin G in the presence of divalent metal ions contains increased secondary structure, with elements of both 3_{10} and α-helices. The identification of a distorted 3_{10} helix between glutamine 9 and Gla 14 in conantokin G is supported by CD spectra collected on the metal-free peptide in which 25−31% of the residues were identified as being within a helical region (Chandler et al., 1993; Zhou et al., 1996). The CD spectra of conantokin G in water were characterized by negative bands blue-shifted to 204 and 222 nm and a positive band blue-shifted to 188 nm. Spectra with similar characteristics have been obtained for peptides with type I and type III β turns and 3_{10} helices (Otvos et al., 1991; Perczel et al., 1993).

The backbone structure of conantokin G has features in common with other conotoxins that contain disulfide bonds. Most notable are the distorted type I β turns or 3_{10} helices (Pardi et al., 1989; Lancelin et al., 1991; Wakamatsu et al., 1992; Guddat et al., 1996; Hill et al., 1996). The three-dimensional X-ray crystallographic structure of α-conotoxin GI possesses a region composed of two type I β turns, with similar disruption of ϕ and ψ angles (Guddat et al., 1996). Both the N termini and C termini of conantokin G are flexible under the conditions used for structural analysis. However, it is plausible that these termini are flexible in solution but stabilize upon interaction with the cell surface receptor, as is the case with some polypeptide hormones (Schwyzer, 1995). This structure of conantokin G is unrelated to the known structures of the Gla domain of the vitamin K-dependent proteins, prothrombin (Soriano-Garcia et al., 1992), factor IX (Freedman et al., 1995a,b, 1996), factor X (Sunnerhagen et al., 1995), or factor VII (Banner et al., 1996). In the absence of metal ions, NMR solution structures of factor IX and factor X show limited formal structure in the γ-carboxyglutamic acid-rich region. Similarly, in the absence of metal ions, the first 33 residues of prothrombin fragment 1 were not observed in the X-ray electron density maps due to the flexibility of this region (Park & Tulinsky, 1986). Upon binding of calcium ions to the vitamin K-dependent proteins and the induction of a conformational transition, most of the γ-carboxyglutamic acid side chains are reoriented to the interior of the folded protein where they chelate core calcium ions. In conantokin G, the five γ-carboxyglutamic acid residues are oriented toward solvent along one side of the conantokin G structure (Figure 7). It is possible that this arrangement facilitates a structural transition of the conantokin G peptide in the presence of metal ions.

Structure−function studies to date have implicated the N terminus of conantokin G in being particularly important for receptor binding. Substitution of γ-carboxyglutamic acid 3 or γ-carboxyglutamic acid 4 with alanine, serine, or phosphoserine yields a conopeptide without NMDA receptor antagonist activity (Zhou et al., 1996). In contrast, γ-carboxyglutamic acid 7, γ-carboxyglutamic acid 10, or γ-carboxyglutamic acid 14 may be replaced without loss of biological activity (Zhou et al., 1996). We have shown that the N-terminal tail is loosely structured in the absence of metal ions. The N-terminal region including γ-carboxyglutamic acid 3, γ-carboxyglutamic acid 4, and possibly glutamic acid 2 may be stabilized by metal ions, or the NMDA receptor itself may induce a stable structural element in this region.

ACKNOWLEDGMENT

We thank Drs. Johan Stenflo and Eva Czerwiec for many helpful discussions and Margaret Jacobs for the peptide synthesis.

SUPPORTING INFORMATION AVAILABLE

Table S-1 containing the proton resonance assignments for the metal-free conantokin G peptide (1 page). Ordering information is given on any current masthead page.

REFERENCES

Banner, D. W., D'Arcy, A., Chene, C., Winkler, F. K., Guha, A., Konigsberg, W. H., Nemerson, Y., & Kirchofer, D. (1996) *Nature 380*, 41−46.

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355−359.

Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521−528.

Chandler, P., Pennington, M., Maccacchini, M.-L., Nashed, N. T., & Skolnick, P. (1993) *J. Biol. Chem. 268*, 17173−17178.

Cruz, L. J., Ramilo, C. A., Corpuz, G. P., & Olivera, B. M. (1992) *Biol. Bull. 183*, 159−164.

Czerwiec, E., Carleton, J., Bronstain, L., Furie, B. C., & Furie, B. (1996) *Blood 88*, 2079 (abstract).

Ducy, P., Desbois, C., Boyce, B., Pinero, G., Story, B., Dunstan, C., Smith, E., Bonadio, J., Goldstein, S., Gundberg, C., Bradley, A., & Karsenty, G. (1996) *Nature 382*, 448−452.

Esmon, C. T., Suttie, J. W., & Jackson, C. M. (1975) *J. Biol. Chem. 250*, 4095.

Fainzilber, M., Hasson, A., Oren, R., Burlingame, A. L., Gordon, D., Spira, M. E., & Zlotkin, E. (1994) *Biochemistry 33*, 9523−9529.

Farr-Jones, S., Miljanich, G. P., Nadasdi, L., Ramachandran, J., & Basus V. J. (1995) *J. Mol. Biol. 248*, 106−124.

Freedman, S. J., Furie, B. C., Furie, B., & Baleja, J. D. (1995a) *J. Biol. Chem. 270*, 7980−7987.

Freedman, S. J., Furie, B. C., Furie, B., & Baleja, J. D. (1995b) *Biochemistry 34*, 12126−12137.

Freedman, S. J., Blostein, M. D., Baleja, J. D., Jacobs, M., Furie, B. C., & Furie, B. (1996) *J. Biol. Chem. 271*, 16227−16236.

Furie, B., & Furie, B. C. (1988) *Cell 53*, 505−518.

Furie, B., & Furie, B. C. (1990) *Blood 75*, 1753−1762.

Guddat, L. W., Martin, J. A., Shan, L., Edmundson, A. B., & Gray, W. R. (1996) *Biochemistry 35*, 11329−11335.

Haack, J. A., Rivier, J., Parks, T. N., Mena, E. E., Cruz, L. J., & Olivera, B. M. (1990) *J. Biol. Chem. 65*, 6025−6029.

Hauschka, P. V., Mullen, E. A., Hintsch, G., & Jazwinski, S. (1988) *Current Advances in Vitamin K Research*, (Suttie, J. W., Ed.) pp 237−243, Elsevier, New York.

Hauschka, P. V., Lian, J. B., Cole, D. E., & Gundberg, C. M. (1989) *Physiol. Rev. 69*, 990−1047.

Havel, T. F. (1991) *Prog. Biophys. Mol. Biol. 56*, 43−78.

Hill, J. M., Alewood, P. F., & Craik, D. J. (1996) *Biochemistry 35*, 8824−8835.

Hyberts, S. G., Goldberg, M. S., Havel, T. F., & Wagner, G. (1992) *Protein Sci. 1*, 736−751.

Jacobs, M., Freedman, S. J., Furie, B. C., & Furie, B. (1994) *J. Biol. Chem. 269*, 25494−25501.

Jeener, J., Meier, B. H., Bachman, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546−4553.

Kobayashi, Y., Ohubo, T., Kyogoku, Y., Nishiuchi, Y., Sakakibara, S., Braun, W., & Go, N. (1989) *Biochemistry 28*, 4853−4860.

Kohno, T., Kim, J. I., Kobayashi, K., Kodera, Y., Maeda, T., & Sato, K. (1995) *Biochemistry 34*, 10256−10265.

Lancelin, J.-M., Kohda, D., Tate, S., Yanagawa, Y., Abe, T., Satake, M., & Inagaki, F. (1991) *Biochemistry 30*, 6908−6916.

Ludvigsen, S., & Poulsen, F. M. (1992) *J. Biomol. NMR 2*, 227−233.

McIntosh, J. M., Olivera, B. M., Cruz, L. J., & Gray, W. R. (1984) *J. Biol. Chem. 259*, 14343−14346.

Mena, E. E., Gullak, M. F., Pagnozzi, M. J., Richter, K. E., Rivier, J., Cruz, L. J., & Olivera, B. M. (1990) *Neurosci. Lett. 118*, 241−244.

Nakamura, T., Yu, Z., Fainzilber, M., & Burlingame, A. L. (1996) *Protein Sci. 5*, 524−530.

Nelsestuen, G. L., Zytkovicz, T. H., & Howard, J. B. (1974) *J. Biol. Chem. 249*, 6347−6350.

Olivera, B. M., Rivier, J., Clark, C., Ramilo, C., Corpuz, G., Bogadie, F., Woodward, S., Hillyard, D., & Cruz, L. (1990) *Science 249*, 257−263.

Olivera, B. M., Rivier, J., Scott, J. K., Hillyard, D. R., & Cruz, L. J. (1991) *J. Biol. Chem. 33*, 22067−22070.

Otvos, L., Jr., Thurin, J., Kollat, E., Urge, L., Mantsch, H. H., & Hollosi, M. (1991) *Int. J. Pept. Protein Res. 38*, 476−482.

Pallaghy, P. K., Duggan, B. M., Pennington, M. W., & Noton, R. S. (1993) *J. Mol. Biol. 243*, 405−420.

Pardi, A., Billeter, M., & Wüthrich, K. (1984) *J. Mol. Biol. 180*, 741−751.

Pardi, A., Galdes, A., Florance, J., & Maniconte, D. (1989) *Biochemistry 28*, 5494−5501.

Park, C. H., & Tulinsky, A. (1986) *Biochemistry 25*, 3977−3982.

Perczel, A., Hollosi, M., Sandor, P., & Fasman, G. D. (1993) *Int. J. Pept. Protein Res 41,* 223−236.
Plateau, P., & Guéron, M. (1982) *J. Am. Chem. Soc. 104,* 7310−7311.
Prorok, M., Warder, S. E., Blandl, T., & Castellino, F. J. (1996) *Biochemistry 35,* 16528−16534.
Rance, M., Sørenson, O. W., Bodenhauen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117,* 479−495.
Richardson, J. S. (1981) *Adv. Protein Chem. 34,* 167−339.
Rivier, J., Galyean, R., Simon, L., Cruz, L. J., Olivera, B. M., & Gray, W. R. (1987) *Biochemistry 26,* 8508−8512.
Schwyzer, R. (1995) *Biopolymers 37,* 5−16.
Skjaerbaek, N., Nielsen, K. J., Lewis, R. J., Alewood, P., & Craik, D. J. (1997) *J. Biol. Chem. 272,* 2291−2299.
Skolnick, P., Boje, K., Miller, R., Pennington, M., & Maccecchini, M.-L. (1992) *J. Neurochem. 59,* 1516−1521.
Soriano-Garcia, M., Padmanabhan, K., de Vos, A. M., & Tulinsky, A. (1992) *Biochemistry 31,* 2554−2566.
Sperling, R., Furie, B. C., Blumenstein, M., Keyt, B., & Furie, B. (1978) *J. Biol. Chem. 253,* 3898−3905.
Stenflo, J., Fernlund, P., Egan, W., & Roepstorff, P. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71,* 2730−2733.
Sunnerhagen, M., Forsén, S., Hoffrén, A.-M., Drakenberg, T., Teleman, O., & Stenflo, J. (1995) *Nat. Struct. Biol. 2,* 504−509.
Suttie, J. W. (1993) *FASEB J. 7,* 445−452.
Vermeer, C., Jie, K. S., & Knapen, M. H. (1995) *Annu. Rev. Nutr. 15,* 1−22.
Wakamatsu, K., Kohda, D., Hatanaka, H., Lancelin, J.-M., Ishida, Y., Oya, M., Nakamura, H., Inagaki, R., & Sato, K. (1992) *Biochemistry 31,* 12577−12584.
Wishart, D. S., Sykes, B. D., & Richards, F. M. (1991) *J. Mol. Biol. 222,* 311−333.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids,* Wiley-Interscience, New York.
Zhou, L.-M., Szendrei, G. I., Fossom, L. H., Maccacchini, M.-L., Skolnick, P., & Otvos, L. (1996) *J. Neurochem. 66,* 620−628.

BI970321W