A Single γ-Carboxyglutamic Acid Residue in a Novel Cysteine-Rich Secretory Protein without Propeptide[†,‡]

Karin Hansson,[*,§] Ann-Marie Thämlitz,[§] Bruce Furie,[||,⊥] Barbara C. Furie,[||,⊥] and Johan Stenflo[§]

Department of Clinical Chemistry, Lund University, University Hospital, Malmö, S-20502 Malmö, Sweden, Marine Biological Laboratory, Woods Hole, Massachusetts 02543, and Center for Hemostasis, Thrombosis and Vascular Biology, Beth Israel Deaconess Medical Center and Department of Medicine, Harvard Medical School, Boston, Massachusetts 02215

Received June 29, 2006; Revised Manuscript Received August 15, 2006

ABSTRACT: γ-Glutamyl carboxylase catalyzes the modification of specific glutamyl residues to γ-carboxyglutamyl (Gla) residues in precursor proteins that possess the appropriate γ-carboxylation recognition signal within the propeptide region. We describe the immunopurification and first biochemical characterization of an invertebrate high molecular weight Gla-containing protein with homologues in mammals. The protein, named GlaCrisp, was isolated from the venom of the marine cone snail *Conus marmoreus*. GlaCrisp gave intense signals in Western blot experiments employing the Gla-specific antibody M3B, and the presence of Gla was chemically confirmed by amino acid analysis after alkaline hydrolysis. Characterization of a full-length cDNA clone encoding GlaCrisp deduced a precursor containing an N-terminal signal peptide but, unlike other Gla-containing proteins, no apparent propeptide. The predicted mature protein of 265 amino acid residues showed considerable sequence similarity to the widely distributed cysteine-rich secretory protein family and closest similarity (65% identity) to the recently described substrate-specific protease Tex31. In addition, two cDNA clones encoding the precursors of two isoforms of GlaCrisp were identified. The predicted precursor isoforms differed at three amino acid positions (−6, 9, and 25). Analysis by Edman degradation and nanoelectrospray ionization mass spectrometry, before and after methyl esterfication, identified a Gla residue at amino acid position 9 in GlaCrisp. This is the first example of a Gla-containing protein without an obvious γ-carboxylation recognition site. The results define a new class of Gla proteins and support the notion that γ-carboxylation of glutamyl residues is phylogenetically older than blood coagulation and the vertebrate lineage.

A unique feature of the vitamin K-dependent proteins is the presence of γ-carboxyglutamyl (Gla)[1] residues. They are formed by posttranslational carboxylation of glutamyl (Glu) residues in a reaction that is catalyzed by the γ-glutamyl carboxylase, an integral endoplasmatic reticulum (ER) membrane protein (*1*−*4*). In the presence of the reduced form of vitamin K and molecular oxygen, the enzyme abstracts a proton with subsequent incorporation of carbon dioxide at the γ-carbon of specific glutamyl residues in appropriate proteins. These proteins carry a carboxylation recognition site (γ-CRS) in an N-terminal cleavable propeptide that not only mediates binding of the substrate to the enzyme but also activates the enzyme (*5*, *6*). Historically, Gla was first discovered and characterized in the N-terminal so-called Gla domain of certain mammalian proteins involved in the hemostatic process. The modification enables these proteins to bind Ca^{2+}, which promotes binding to phospholipid surfaces. Various mammalian Gla-containing proteins have since been discovered, implicating a role of Gla in bone morphogenesis (bone Gla protein, BGP; matrix Gla protein, MGP) and growth control (Gas6) (*7*−*9*).

The identification of Gla in certain neuropeptides (conotoxins) isolated from the venom duct of predatory marine cone snails of the genus *Conus* provided evidence that this posttranslational modification is of ancient evolutionary origin and not confined to the vertebrate lineage (*10*, *11*). Cloning and expression of the *Conus* γ-glutamyl carboxylase revealed an enzyme with the same general requirements as has been established for its vertebrate counterpart (*12*−*14*). Like the mammalian enzyme, the *Conus* carboxylase requires reduced vitamin K and the presence of a γ-CRS located in the propeptide region on a precursor form of the conotoxin substrate (*15*, *16*). However, the propeptide of the conotoxin precursors bears no sequence similarity to the known

[†] This work was supported by Grant K2001-03X-04487-27A from the Swedish Medical Research Council, the Swedish Foundation of Strategic Research, the Kock Foundations, and the Foundation of University Hospital, Malmö.

[‡] The nucleotide sequences reported in this paper have been submitted to the GenBank/EBI Data Bank with accession numbers DQ647193, DQ647194, and DQ647195.

* To whom correspondence should be addressed. Phone: 46-40-337226. Fax: 46-40-929023. E-mail: Karin_M.Hansson@med.lu.se.

[§] Lund University.

[||] Marine Biological Laboratory.

[⊥] Harvard Medical School.

[1] Abbreviations: Gla, γ-carboxyglutamyl; Glu, glutamyl; ER, endoplasmatic reticulum; γ-CRS, γ-carboxylation recognition site; BGP, bone Gla protein; MGP, matrix Gla protein; PRGP, proline-rich Gla protein; TMG, transmembrane Gla protein; PR, pathogenesis-related; CRISP, cysteine-rich secretory protein; nanoESI, nanoelectrospray ionization; PCR, polymerase chain reaction; Abu, 2-aminobutyric acid; SCP, sperm-coating protein; CRD, cysteine-rich domain; VA5, venom antigen 5.

10.1021/bi061311a CCC: $33.50
Published on Web 09/30/2006

mammalian γ-CRS, and a consensus sequence of the *Conus* carboxylase binding site has been difficult to identify. Moreover, a C-terminal "postpeptide" was recently demonstrated to contain the recognition signal for carboxylation in two novel conotoxins (*17*). A functional vitamin K-dependent carboxylase has also been demonstrated in *Drosophila melanogaster*, although a substrate has yet to be identified (*18*).

The cone snail venom is typically loaded with at least 100 different pharmacologically active components of varying molecular weights (*11*, *19*). Most are small peptides (<30 amino acids) specifically targeted to receptors and ion channels to paralyze the prey by antagonizing neurotransmission. To date many Gla-containing conotoxins have been described and characterized from the venoms of various cone snails (*20*−*27*). Even though the importance of Gla for the biological activity of some of these peptides has been demonstrated, the function of the posttranslational modification in the cone snails is not well established.

Although larger protein components are present in the cone snail venoms, only a few have been biochemically characterized and none have been reported to contain Gla (*28*−*32*). Recently, a 30 kDa calcium-dependent substrate-specific protease, Tex31, was identified in the venom of *Conus textile* (*33*). The protease was reported to be a member of the widely distributed pathogenesis-related (PR) protein superfamily and most closely related to the cysteine-rich secretory protein (CRISP) family of mammalian proteins (*34*, *35*). The mammalian CRISPs have been found in secretions of exocrine glands, first of all in male reproductive tracts and saliva, and in human granulocytes and plasma (*36*, *37*). The function of these proteins is not well defined at present, but various functions related to sperm−egg fusion and the innate immune system have been suggested (*34*, *38*, *39*).

The broad distribution of the γ-glutamyl carboxylase gene suggests critical function(s) for Gla that has been retained throughout phylogeny. Isolation of novel Gla-containing proteins from invertebrate species with homologues in vertebrates should provide insight into these functions. Recently, we developed mouse monoclonal antibodies that permit the sensitive identification of Gla in complex biological protein extracts (*40*). In the present study, we used one of these antibodies to detect and purify a novel Gla-containing protein belonging to the PR superfamily from the venom of *Conus marmoreus*. A full-length cDNA clone encoding the protein, here referred to as GlaCrisp, predicted a sequence with closest similarity to the substrate-specific protease Tex31. Unlike all previously reported Gla proteins, GlaCrisp contains a single Gla residue in a context without propeptide. These results indicate that vitamin K-dependent γ-carboxylation is phylogenetically older than blood coagulation and tissue mineralization. They also raise the possibility that Gla residue(s), with yet unknown function(s), remain to be identified in proteins, perhaps also in humans.

EXPERIMENTAL PROCEDURES

Purification of GlaCrisp and Its Two Isoforms from the Venom of C. marmoreus. Crude venom from *C. marmoreus* snails obtained from Vietnam was extracted and size-fractionated on a Sephadex G-50 Superfine column (25 mm × 920 mm) equilibrated with 0.2 M ammonium acetate buffer (pH 7.5) and eluted with a flow rate of 10.3 mL/h (Figure 1A) (*24*). The Gla content of every third collected fraction was determined after alkaline hydrolysis. Under reducing conditions an immunoreactive polypeptide band of ~40 kDa (migrating at ~30 kDa under nonreducing conditions) was observed in the size-fractionated material by Western blot experiments employing the Gla-specific mouse monoclonal antibody M3B (pool 1, denoted by horizontal bar 1 in Figure 1A). The M3B antibody had been purified and prepared as described previously (*40*). The Gla-specific antibody (2.3 mg) was coupled to 2 mL (bed volume) of Sulfolink coupling gel (Pierce Chemical Co., Rockford, IL) according to the manufacturer's instructions. The M3B-coupled resin was used to prepare an immunoaffinity column that was equilibrated with loading buffer (20 mM Tris−HCl, 50 mM NaCl, 10 mM EDTA, pH 7.4). The pooled material (Figure 1A) was dialyzed against the loading buffer using Spectra/Por 3 (MW cutoff 3500, Spectrum Laboratories, Inc., California) and loaded onto the immunoaffinity column at a flow rate of 3 mL/h. After A_{280} decreased to near the baseline the column was eluted at 10 mL/h with 20 mM Tris−HCl, 300 mM NaCl, 10 mM EDTA, pH 7.4. The eluted material (horizontal bar 2, Figure 1B) was concentrated using a Vivaspin centrifugal filter (MW cutoff 5000, Vivascience, Hannover, Germany) and applied to a Superdex 75 10/300GL column (10 mm × 300 mm, Amersham Biosciences AB, Uppsala, Sweden) equilibrated with 50 mM Tris−HCl, 150 mM NaCl, pH 8.0 (Figure 1C). A flow rate of 0.02 mL/min was employed, and 140 μL fractions were collected while A_{280} was monitored. The purity of the isolated 30 kDa protein in the major peak (denoted by horizontal bar 3 in Figure 1C) was >95% as assessed by SDS−PAGE analysis of the

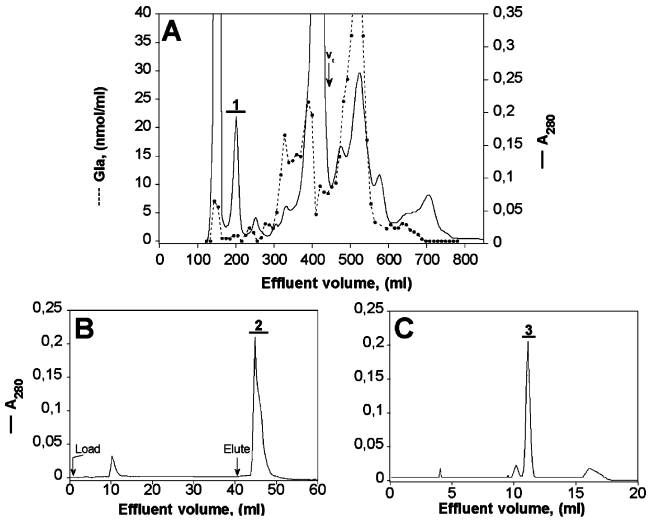

FIGURE 1: Purification of 30 kDa CRISP-like proteins from the venom of *C. marmoreus*. (A) An extract prepared from the venom of *C. marmoreus* was chromatographed on a Sephadex G-50 Superfine gel filtration column. The Gla content (---) of every third collected fraction was determined after alkaline hydrolysis. The vertical arrow indicates one column volume. (B) The fractions (pool 1, shown by horizontal bar 1 in (A)) containing the 30 kDa GlaCrisp protein and its two isoforms were subjected to immunoaffinity chromatography on a column of monoclonal antibody M3B-coupled resin. The column was eluted employing 300 mM NaCl. (C) The NaCl-eluted fraction (shown by horizontal bar 2 in (B)) was further purified on a Superdex 75 gel filtration column. The peak marked with horizontal bar 3 contained the purified 30 kDa CRISP protein isoforms including GlaCrisp.

reduced and denaturated material. A total of 30 nmol of purified protein was obtained from five venom ducts. M3B gave a strong response against the purified protein under Western blot conditions.

Partial Purification of Tex31 from the Venom of C. textile. Crude *C. textile* venom was extracted according to the protocol described by Milne et al. (*33*). The extract was separated by reversed-phase HPLC on a Genesis C4 column (4 μm, 10 mm × 250 mm, Grace Vydac, California) using an acetonitrile gradient in 0.1% TFA (v/v). Fractions interacting with M3B in Western blot analysis were pooled and applied to a HiTrap Q-HP column (1 mL, Amersham Pharmacia Biotech AB, Uppsala, Sweden) eluted with a linear 0−1 M NaCl gradient in 50 mM Tris−HCl, pH 8.0. M3B-immunoreactive fractions eluting at ∼350 mM NaCl were pooled and concentrated by ultrafiltration (as described above) and loaded onto a Superdex 75 10/300GL column at a flow rate of 0.02 mL/min in 50 mM Tris−HCl, 150 mM NaCl, pH 8.0. The partially purified protein binding to M3B was identified as Tex31 by enzymatic in gel digestion followed by mass spectrometry (see below).

SDS−PAGE and Western Blot Analysis. Protein samples were reduced, alkylated, and separated in 12% (w/v) SDS−polyacrylamide gels and silver stained using the silver staining kit PlusOne (Amersham Biosciences AB). For Western blot analysis proteins were electrotransferred to Immobilon-P transfer membrane (Millipore, Bedford, MA) using wet electrophoretic transfer. The membranes were treated as described previously, and Gla proteins were visualized using the Gla-specific monoclonal antibody M3B followed by alkaline-phosphatase-conjugated rabbit anti-mouse IgG (DAKO A/S, Glostrup, Denmark) (*40*).

Amino Acid Sequence and Composition Analyses. N-terminal sequences were determined on an ABI Procise 494 automatic sequencer (Foster City, CA) according to the protocol from the manufacturer. To allow the identification of the Gla residue in the sequence, samples were sequenced after methyl esterfication with methanolic HCl (*41*). The amino acid composition was determined on a Beckman 3000 amino acid analyzer after acid hydrolysis, except for Gla, which was measured after alkaline hydrolysis (*42*, *43*).

Reduction and Alkylation of Disulfide Bonds. For reduction and alkylation of cysteine residues with iodoacetic acid (BDH Chemicals Ltd., Poole, England), the protein sample was dissolved in 2 mL of 6 M guanidine hydrochloride in 1 M Tris−HCl buffer, 10 mM EDTA, pH 8.6. Dithiothreitol (DTT) was added to 20 mM followed by incubation at 37 °C for 2 h, after which iodoacetic acid was added to a final concentration of 50 mM. After incubation at room temperature for 30 min, β-mercaptoethanol was added to a final concentration of 1% to quench the reaction. The protein sample was then dialyzed against 20% (v/v) acetic acid and lyophilized.

Endoproteinase Digestions and Separation of Peptides. For endoproteinase digestions the reduced and alkylated protein sample was dissolved in 100 mM NH_4HCO_3 buffer (pH 7.9) and digested with either trypsin (sequencing grade modified, Promega, Madison, WI) or trypsin followed by Asp-N (sequencing grade, Roche Diagnostics GmbH, Penzberg, Germany). Digestions were performed in substrate to enzyme ratios of 20:1. The reaction mixtures were incubated at 37 °C for 18 h and terminated by freezing at −20 °C. For mass spectrometry analysis the generated peptides were micropurified and concentrated on $ZipTip_{C18}$ (Millipore) columns and eluted with aqueous 50% (v/v) methanol containing 1% (v/v) formic acid. The tryptic digest was separated by reversed-phase HPLC on a Genesis C4 column (4 μm, 4.5 mm × 100 mm, Grace Vydac) in 0.1% (v/v) trifluoroacetic acid and developed with an acetonitrile gradient. The Gla content of the collected fractions was determined after alkaline hydrolysis. The material in the Gla-containing fraction (eluting at 44% acetonitrile) was subjected to mass spectrometry analysis and Edman degradation before and after methyl esterfication with methanolic HCl.

Enzymatic in Gel Digestions. In gel enzymatic digestion of Tex31 was performed following a protocol slightly modified from that of Shevchenko et al. (*44*). Briefly, the excised gel pieces were rinsed twice with 50% acetonitrile (v/v) and shrunk in 100% acetonitrile, followed by reduction of the cysteines with 10 mM DTT/0.1 M NH_4HCO_3 and subsequent S-carbamidomethylation of free cysteines with 55 mM iodoacetamide in 0.1 M NH_4HCO_3. For enzymatic digestion of the reduced and alkylated protein the sample was dissolved in 50 mM NH_4HCO_3 buffer (pH 7.9) and digested with trypsin followed by Asp-N (substrate to enzyme ratio of 20:1). The reaction mixture was incubated on ice for 45 min followed by 37 °C for 18 h and terminated by freezing at −20 °C. For mass spectrometry analysis the peptides generated were micropurified and concentrated on $ZipTip_{C18}$ columns as described above.

Mass Spectrometry. Nanoelectrospray ionization (nanoESI) experiments were carried out on an API QSTAR Pulsar-i quadrapole/time-of-flight mass spectrometer (Applied Biosystems/MDS Sciex, Toronto, Canada) equipped with a nanoESI source (MDS Protana, Odense, Denmark). The samples were sprayed from a silver-coated glass capillary supplied from New Objective, Woburn, MA. Spectra were obtained in positive ion mode with an ion spray voltage of 800−1000 V. The acquisition and the deconvolution were performed on an AnalystQS Windows PC data system. Bioanalyst version 1.0 software (Applied Biosystems/MDS Sciex) was used to analyze the mass spectra. Parent ions (identified in a time-of-flight MS survey scan) for MS/MS analysis were selected in Q1, and product ions were generated in Q2 using N_2 as the collision gas and collision energies of 14−36 eV.

cDNA Cloning and Sequencing. The cDNA sequences of GlaCrisp and its two isoforms were determined by polymerase chain reaction (PCR) amplification followed by subcloning and sequence analysis by standard techniques. Oligonucleotide PCR primers were designed on the basis of the reported cDNA sequence of Tex31 (*33*). The primer sequences were 5′-ACG ATC TTC CGT CGG ACA GTG-3′ (KHmex-13), 5′-GCT TGA AGT CTG TCA GGA C-3′ (KHmex-14), and 5′-GCC ACA AAT CAA CAG AGA C-3′ (KHmex-15). The primers were used in PCR reactions using a λ ZAPII custom phage library of *C. marmoreus* venom duct cDNA (Stratagene, La Jolla, CA) as the template. The conditions for PCR were 95 °C for 8 min, followed by 35 cycles of 94 °C for 1 min, 46 °C for 1 min, and 72 °C for 1 min. After the completion of 35 cycles, the reaction was maintained at 72 °C for 10 min. The PCR products were purified and TA-cloned into the pGEM-T vector (Promega), and the resulting plasmid DNA was sequenced using the

primers M13 forward, 5′-GTT TTC CCA GTC ACG AC-3′, M13 reversed, 5′-CAG GAA ACA GCT ATG AC-3′, KHmex-4, 5′-TAC TTT GTC TGT AAC TAT TAC AA G-3′, KHmex-10, 5′-CTT TTC CTG GCA GTC AGC TCC-3′, KHmex-11, 5′-TGC GAG TGC CAG TGT CCT AG-3′, and KHmex-12, 5′-TGG CAT TAT CCG TCA GGC AC-3′. The oligonucleotides were synthesized by DNA Technology A/S, Aarhus, Denmark.

Amino Acid Sequence Alignment. For sequence comparisons, the multiple sequence alignment program CLUSTAL W was employed using the default settings (http://www.ch.embnet.org/software/ClustalW.html).

Synthetic Substrate Synthesis and Proteolytic Activity Assay. The peptide substrate for the protease activity assay ∧KLNKRWAbuKQSG* (∧, acetylation; Abu, 2-aminobutyric acid; *, amidation) was synthesized with Dpfd L-Fmoc-amino acids (Perseptive Biosystems, Framingham, MA) on a Milligen 9050 Plus peptide synthesizer (PerkinElmer Life Sciences, Boston, MA). The peptide was deprotected and cleaved from the resin by treatment with 95% anhydrous trifluoroacetic acid containing relevant scavengers. The peptide was then applied to a Kromasil C8 preparative reversed-phase HPLC column (5 μm, 21.2 mm × 250 mm) and eluted with an acetonitrile gradient in 0.1% (v/v) trifluoroacetic acid. The major peak corresponding to the peptide fraction was confirmed by mass spectrometry. The synthetic peptide substrate was used in protease activity assays performed in accordance with those of Tex31 (*33*). A typical reaction mixture was made up of 100 μM peptide substrate, 1.5 μM Superdex 75-purified *C. marmoreus* protein in 100 μL of 100 mM bis-Tris propane, and 10 mM $CaCl_2$, pH 8.0. The reactions were incubated for 3 h at 37 °C, and the potential cleavage products were monitored by C18 reversed-phase HPLC (3 μm, 2.1 mm × 100 mm, Amersham Biosciences AB) using an acetonitrile gradient in 0.1% (v/v) trifluoracetic acid and nanoESI mass spectrometry.

RESULTS

Identification and Purification of a 30 kDa Gla Protein. To search for a high molecular weight Gla component in the venom of *C. marmoreus*, a crude venom duct extract was size-fractionated on a Sephadex G-50 Superfine gel filtration column and the effluent was pooled on the basis of Gla content (Figure 1A). Early-eluting high molecular weight polypeptides were screened for Gla components by Western blot analysis employing the Gla-specific mouse monoclonal antibody M3B. Under reducing conditions an immunoreactive polypeptide with an apparent molecular mass of ∼40 kDa (migrating at ∼30 kDa under nonreducing conditions) was detected in pool 1 (Figures 1A and 2). The immunoreactive protein was purified by affinity chromatography using a column of M3B-coupled resin followed by gel filtration on a Superdex 75 column (Figure 1B,C). The purified protein, which was more than 95% homogeneous as judged by SDS−polyacrylamide gel electrophoresis, gave an intense signal in Western blot experiments under reducing conditions using M3B (Figure 2).

The native 30 kDa protein gave a single N-terminal sequence (His-Ala-Xxx-Asp-Ser-Lys-Tyr-Ser-Asp-Val-Thr-Pro-Thr-His-Thr) by automated Edman degradation. The N-terminal amino acid sequence together with sequence tags

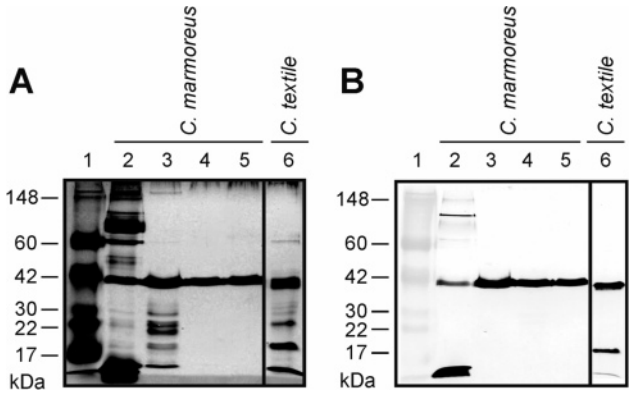

FIGURE 2: Gel electrophoresis and Western blot analysis of the purification steps of GlaCrisp and Tex31. Lanes 2−5 show purification steps of GlaCrisp from *C. marmoreus*, and lane 6 shows the purification step of Tex31 from *C. textile*. (A) Protein samples from purification steps of GlaCrisp (Figure 1) and Tex31 were reduced, alkylated, and resolved by 12% SDS−PAGE and silver stained. Under reducing conditions GlaCrisp and Tex31 migrated with an apparent molecular weight of ∼40 kDa (migrating at 30 kDa under nonreducing conditions). (B) The Gla-specific monoclonal antibody M3B was used to detect the Gla-containing components in the Western blot experiments. Key: lane 1, molecular weight marker; lane 2, crude *C. marmoreus* venom extract; lane 3, pool 1 from the Sephadex G50 Superfine gel filtration column (Figure 1A); lane 4, NaCl-eluted fraction from the M3B immunoaffinity column (Figure 1B); lane 5, purified GlaCrisp and isoforms from the Superdex 75 gel filtration column (Figure 1C); lane 6, partially purified Tex31 from the Superdex 75 gel filtration column.

obtained by nanoESI mass spectrometry of internal tryptic peptide fragments (denoted in Figure 3) indicated a novel protein that is a homologue of the recently described substrate-specific protease Tex31 isolated from *C. textile* (*33*). The amino acid composition determined after acid and alkaline hydrolysis revealed that the purified protein contained 0.4 mol of Gla/mol of protein (three separate measurements). As will be discussed below the protein was a mixture of three isoforms, one carrying a Gla residue.

Amino Acid Sequence Prediction of GlaCrisp and Its Two Isoforms from Cloned cDNAs. The entire amino acid sequence of the three precursor isoforms of the purified Tex31 homologue was predicted by sequencing isolated cDNA clones obtained from a *C. marmoreus* cDNA venom duct library (Figure 3). The deduced sequences consist of a putative signal peptide sequence of 22 residues (amino acids −24 to −3) and a canonical dibasic signal for proteolytic cleavage (amino acids −2 to −1), followed by a 265-residue-long mature translation product. No apparent propeptide sequence was observed between the signal sequence and the mature protein. The predicted precursor sequences are identical except for three differences at amino acid positions −6 ($Ala^1/Ala^2/Val^3$), 9 ($Asp^1/Asp^2/Glu^3$), and 25 ($Met^1/Val^2/Val^3$) due to variations in the cDNA sequences at nucleotides 57 ($C^{1,2}/T^3$), 100 ($T^{1,2}/G^3$), and 146 ($A^1/G^{2,3}$), respectively (Figure 3). (The superscripts refer to the three isoforms; isoform 3 is also referred to as GlaCrisp.) On the basis of the isolated cDNA clones, the predicted mature protein sequences contain either 14 (isoforms 1 and 2) or 15 (GlaCrisp) Glu residues that could be posttranslationally carboxylated to Gla. The N-terminal amino acid sequence obtained by Edman degradation of the purified Tex31 homologue corresponds to the sequences of isoforms 1

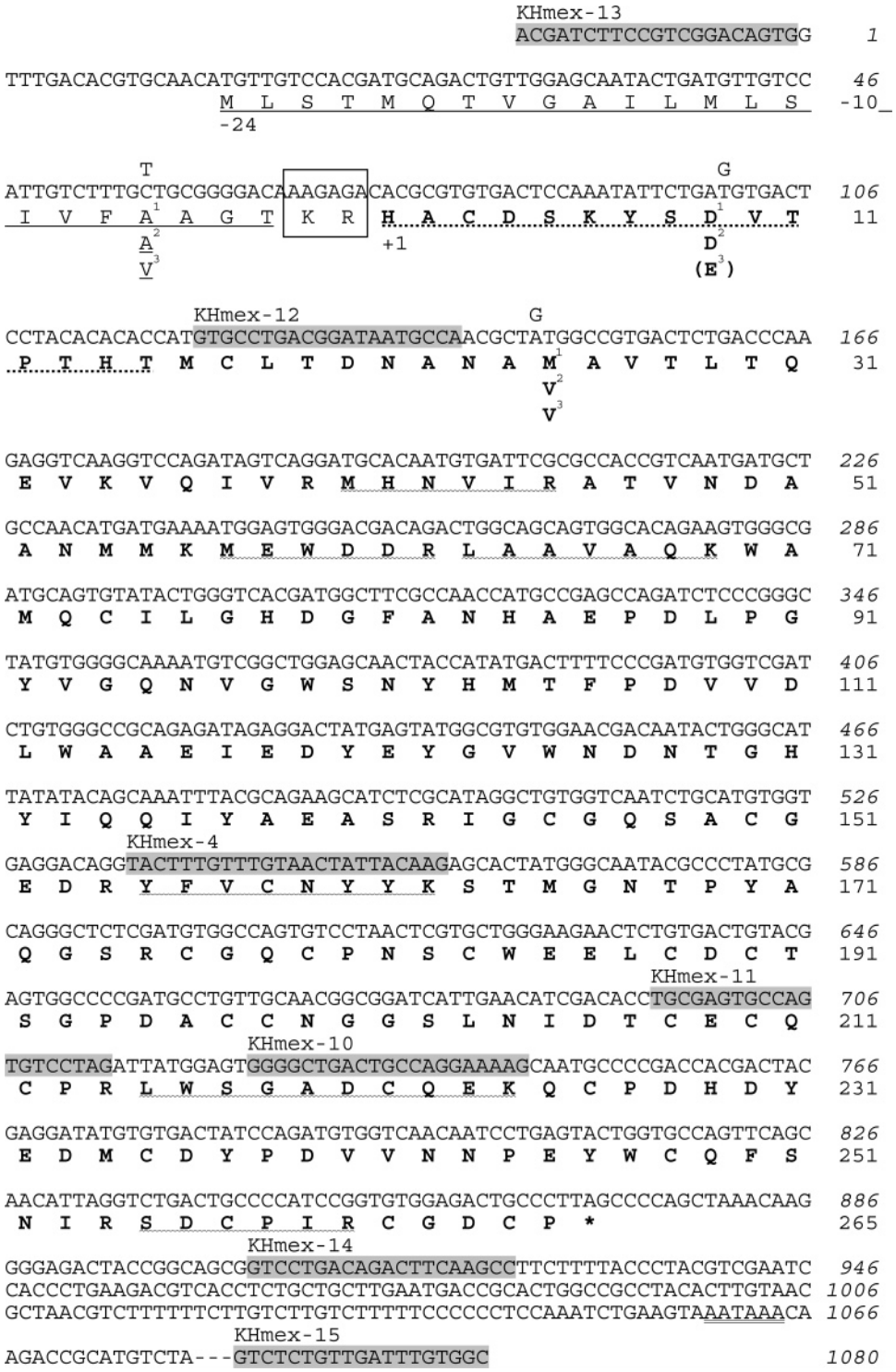

FIGURE 3: cDNA sequences encoding the precursors of GlaCrisp and its two isoforms. The deduced amino acid sequence is shown in single letter code below the nucleotide sequence. The Glu residue in GlaCrisp that is posttranslationally carboxylated to Gla is shown in parentheses. The putative signal peptide predicted by PSORT analysis is underlined (from amino acid positions −24 to −3), and the putative dibasic proteolytic cleavage site (amino acid positions −2 to −1) is boxed. Nucleotide (in italics) and amino acid numbers are shown to the right. The polyadenylation signal is doubly underlined. Superscript numbers 1, 2, and 3 refer to isoforms 1, 2, 3 (here also named GlaCrisp), respectively. The amino acid sequences determined by (underlined by a hatched line) Edman degradation or (underlined by a jagged line) mass spectrometry used to identify the protein as a homologue of the *C. textile* protein Tex31 are marked. The positions for the oligonucleotide primers used to determine the cDNA sequences are shaded.

and 2, having an Asp residue at position 9. However, the identification of a potential Gla residue at position 9 in GlaCrisp needs methyl esterfication of the protein prior to sequence analysis. Thus, the determined cDNA sequences support the notion that the purified protein is a mixture of three isoforms.

A BLAST analysis of the SwissProt database showed that GlaCrisp and its two isoforms are closely related to members of the CRISP family and most closely resemble Tex31 (amino acid sequence identity 65%). Like the earlier reported

CRISP proteins, the *C. marmoreus* proteins possess an SCP (sperm-coating protein) domain and a characteristic cysteine-rich C-terminal region.

Identification of a Single Gla Residue in GlaCrisp. To identify the Gla residue in the purified 30 kDa *C. marmoreus* protein, it was digested with trypsin after reduction and S-carboxymethylation of cysteine residues. The resulting tryptic peptide fragments were separated by reversed-phase HPLC using a linear acetonitrile gradient in 0.1% trifluoroacetic acid. The collected fractions were analyzed for the content of Gla after alkaline hydrolysis, and the material in the single-Gla-containing fraction, eluting at ∼44% acetonitrile, was subjected to nanoESI mass spectrometry analysis. Three dominating components were identified in a mass spectrum with monoisotopic molecular masses of 3111.42 Da (calculated from the 3+ ion at *m*/*z* 1038.12 and the 4+ ion at *m*/*z* 778.87), 3079.48 Da (calculated from the 3+ ion at *m*/*z* 1027.48 and the 4+ ion at *m*/*z* 770.88), and 3137.50 Da (calculated from the 3+ ion at *m*/*z* 1046.79 and the 4+ ion at *m*/*z* 785.41) (Figure 4A and Table 1). The observed masses most likely correspond to the tryptic peptides covering amino acids 7−34 in the three predicted isoforms, assuming that GlaCrisp contains a carboxylated Glu residue (Gla) at position 9. The 3+ and 4+ ion peaks at *m*/*z* 1046.79 and 785.41 were accompanied by losses of 44 Da (corresponding to the mass of CO_2), which further suggested the presence of a Gla residue in the tryptic peptide derived from GlaCrisp. The monoisotopic mass of the decarboxylated peptide was determined from the mass spectrum to be 3093.49 Da. A number of ion signals corresponding to oxidized species were probably produced during sample preparation.

To verify the primary structure of the internal tryptic peptides covering amino acids 7−34, each of the triply charged ions at *m*/*z* 1038.12 (isoform 1), *m*/*z* 1027.48 (isoform 2), and *m*/*z* 1046.79 (GlaCrisp) were selected for MS/MS fragmentation (Figure 4B). The spectra obtained revealed y ions (y1−y10) according to the C-terminal part of the proposed sequences. The masses of the y10 ions confirmed the presence of a Met residue at position 25 in isoform 1 and a Val residue at the corresponding positions in isoform 2 and GlaCrisp.

NanoESI mass spectrometry after methyl esterfication of the tryptic peptide mixture (residues 7−34 of the three isoforms) showed a 70 Da increase in the monoisotopic mass for the peptides derived from isoform 1 (mass calculated as 3181.52 Da from the 3+ ion at *m*/*z* 1061.49 and the 4+ ion at *m*/*z* 796.39) and isoform 2 (mass calculated as 3149.54 Da from the 3+ ion at *m*/*z* 1050.83 and the 4+ ion at *m*/*z* 788.40) (Figure 4C and Table 1). This is consistent with methyl esterfication of the Asp residues at positions 9 and 20, the S-carboxymethylated Cys residue at position 17, the Glu residue at position 32, and the C-terminal carboxyl group. A mass increase of 84 Da was observed for the peptide derived from GlaCrisp (mass calculated as 3221.56 Da from the 3+ ion at *m*/*z* 1074.83 and the 4+ ion at *m*/*z* 806.41), corresponding to the introduction of six methyl groups including two on the presumed Gla residue at position 9. The ion peaks at *m*/*z* 1055.51 and *m*/*z* 791.90 most likely correspond to the decarboxylated 3+- and 4+-charged species, respectively, derived from GlaCrisp. The fragment ion spectrum obtained for the selected trapped ion at *m*/*z* 1074.83 revealed y ions according to the proposed C-terminal sequence of the methyl-esterfied tryptic peptide (amino acids 7−34) derived from GlaCrisp (Figure 4B).

To confirm the presence of a Gla residue at position 9 in GlaCrisp, the methyl-esterfied tryptic peptide mixture (residues 7−34 of the three protein isoforms) was subjected to Edman degradation. According to the proposed sequences, a combination of Asp and Gla residues was obtained at position 9 in the mature proteins (Figure 3). In agreement, Edman degradation analysis after methyl esterfication of the intact purified 30 kDa protein revealed two N-terminal sequences having a mixture of Asp and Gla residues at position 9 (His-Ala-Xxx-Asp-Ser-Lys-Tyr-Ser-Asp/Gla-Val-Thr-Pro-Thr-His-Thr).

The Gla residue in GlaCrisp was further verified by digestion of the S-carboxymethylated protein with trypsin followed by Asp-N and subsequent analysis of the derived peptide fragments with nano-ESI mass spectrometry. The internal peptide fragment with an apparent monoisotopic molecular mass of 1583.68 Da (calculated from the +2 ion at *m*/*z* 792.84) covering residues 7−19 of GlaCrisp was subjected to MS/MS fragmentation (Figure 5A,B, and Table 1). The y and b ion series obtained for the selected trapped ion at *m*/*z* 792.84 readily confirmed the peptide sequence to be Tyr-Ser-Gla-Val-Thr-Pro-Thr-His-Thr-Met-Cys-Leu-Thr. The *m*/*z* region from 230 to 630 has been expanded in Figure 5B to show the b2−b5 ions from which the N-terminal peptide sequence, including the Gla residue, was determined. These results clearly demonstrate the presence of a Gla residue at position 9 in the GlaCrisp isoform.

Negative Results of Proteolytic Activity Assays of GlaCrisp. Proteolytic propeptide cleavage activity was recently demonstrated for the *C. textile* Tex31 protein. To examine the potential proteolytic activity of GlaCrisp and its isoforms, a peptide substrate was synthesized based upon the proposed Tex31 propeptide substrate (∧KLNKRWAbuKQSG*). The substrate was incubated with the purified protein isoforms in the presence of 10 mM $CaCl_2$, and the possible cleavage products were assessed by reversed-phase chromatography and nanoESI mass spectrometry (data not shown). No proteolytic activity was detected. However, it is possible that the *C. textile* and the *C. marmoreus* homologues possess different substrate sequence specificities.

Identification of Gla in the Substrate-Specific Protease Tex31. Under reducing conditions Tex31 that had been partially purified from a venom extract of *C. textile* interacted strongly in Western blot experiments with the Gla-specific antibody M3B (Figure 2). These results suggested that the GlaCrisp homologue also is a Gla-containing protein. Amino acid sequence alignment of the two homologues revealed that Tex31 carries a Glu residue at position 9 corresponding to the Gla residue in GlaCrisp. To verify a Gla residue at this position in Tex31, the protein was subjected to in gel digestion with trypsin followed by Asp-N, and the derived peptides were analyzed by mass spectrometry. MS/MS fragmentation of the selected trapped ion at *m*/*z* 822.35 generated b and y ions in agreement with the peptide sequence Tyr-Tyr-Gla-Leu-Thr-Pro-Ala-His-Thr-Met-Cys-Leu-Thr covering residues 7−21 of Tex31 (Figures 5C and 6). Thus, the conservation of Gla in GlaCrisp and Tex31

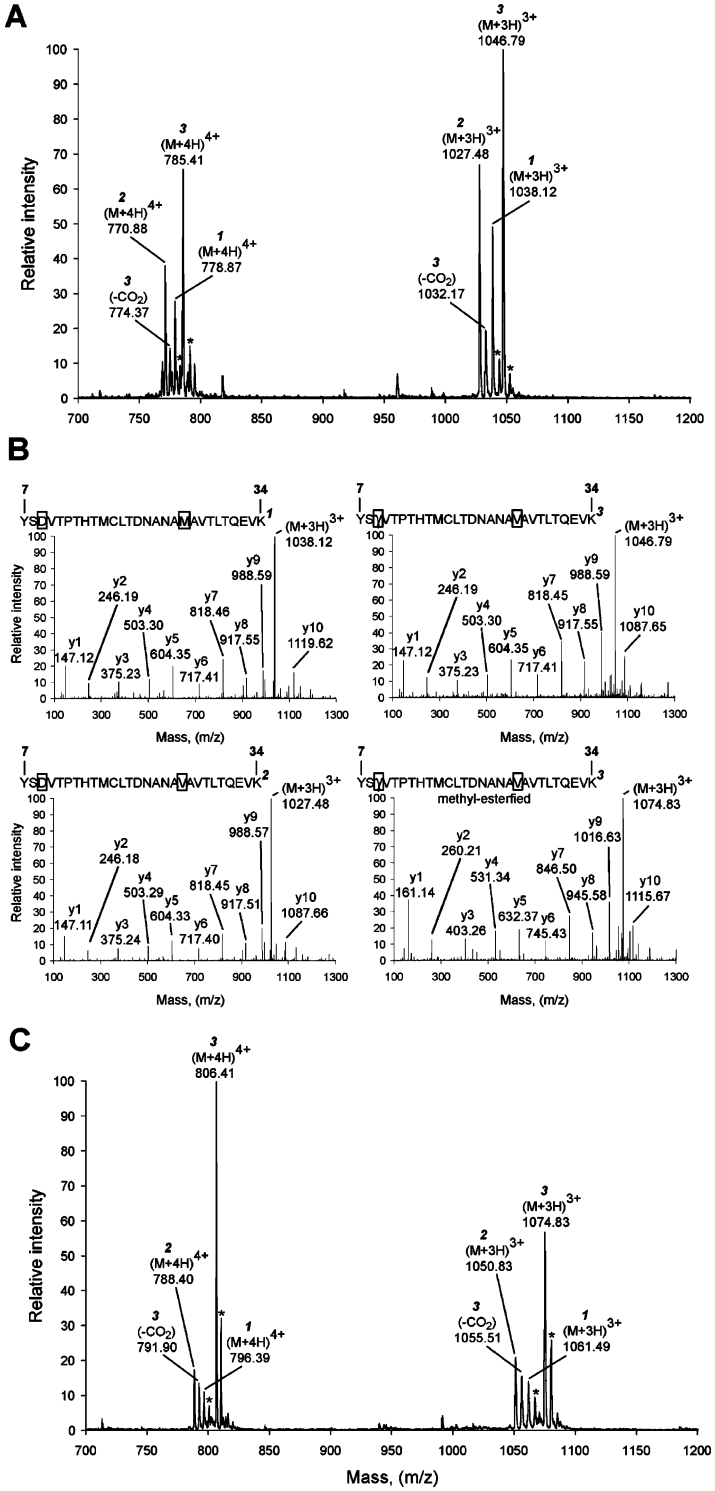

FIGURE 4: Mass spectrometric analysis of the internal tryptic peptide fragments covering amino acids 7−34 of GlaCrisp and its two isoforms. Italic numbers 1, 2, and 3 refer to the three CRISP-like protein isoforms from which the peptides were derived after trypsin digestion. Mass spectra were acquired in the positive ion mode. Ion peaks labeled with an asterisk represent oxidized products. (A) The nanoESI-MS spectrum of the S-carboxymethylated internal tryptic peptide fragments (amino acids 7−34) is shown. The monoisotopic molecular masses of the tryptic peptides were determined from the mass spectrum to be 3111.42 Da (isoform 1; from the 3+ ion at *m*/*z* 1038.12 and the 4+ ion at 778.87), 3079.48 Da (isoform 2; from the 3+ ion at *m*/*z* 1027.48 and the 4+ ion at 770.88), and 3137.50 Da (isoform 3; from the 3+ ion at *m*/*z* 1046.79 and the 4+ ion at 785.41). The ions at *m*/*z* 1046.79 (3+) and 785.41 (4+) are accompanied by losses of 44 Da (corresponding to the mass of CO_2), which is consistent with the presence of a Gla residue in the peptide. (B) The nanoESI-MS/MS spectra of the triply charged parent ions at *m*/*z* 1038.12, 1027.48, and 1046.79 (acquired before methyl esterfication) and at *m*/*z* 1074.83 (acquired after methyl esterfication) are shown. The amino acid residues that differ between GlaCrisp and its isoforms are boxed. The y ions according to the C-terminal part of the proposed sequences (y1−y10) of the tryptic peptides are labeled. (C) The nanoESI-MS spectrum of the S-carboxymethylated internal tryptic peptide fragments (amino acids 7−34) after methyl esterfication is shown. After methyl esterfication a 70 Da increase (corresponding to the methylation of all five carboxyl groups including the one at the C-terminus) in the monoisotopic molecular masses was observed for the peptide fragments derived from isoforms 1 and 2, whereas an 84 Da increase (corresponding to the methylation of all six carboxyl groups including the two at the Gla residue and the one at the C-terminus) was observed for the peptide derived from isoform 3 (GlaCrisp).

Table 1: Observed and Calculated Monoisotopic Masses of Peptide Fragments Derived from Endoproteinase Digestions of the *C. marmoreus* 30 kDa CRISP Protein Isoforms

residue[a]	sequence	enzyme[b]	MW_{obsd}	MW_{calcd}
$7-34^1$	YS*D*VTPTHTMCLTDNANA*M*AVTLTQEVK	trypsin	3111.42	3111.42
$7-34^2$	YS*D*VTPTHTMCLTDNANA*V*AVTLTQEVK	trypsin	3079.48	3079.45
$7-34^3$	YSγVTPTHTMCLTDNANA*V*AVTLTQEVK	trypsin	3137.50	3137.45
$-CO_2$	YS*E*VTPTHTMCLTDNANA*V*AVTLTQEVK		3093.49	3093.45
$7-34^1$	YS*D*VTPTHTMCLTDNANA*M*AVTLTQEVK	trypsin[c]	3181.52	3181.50
$7-34^2$	YS*D*VTPTHTMCLTDNANA*V*AVTLTQEVK	trypsin[c]	3149.54	3149.53
$7-34^3$	YSγVTPTHTMCLTDNANA*V*AVTLTQEVK	trypsin[c]	3221.56	3221.54
$-CO_2$	YS*E*VTPTHTMCLTDNANA*V*AVTLTQEVK		3163.54	3163.53
$7-8^{1,2}$	YS	trypsin + Asp-N	ND	268.11
$9-19^{1,2}$	*D*VTPTHTMCLT	trypsin + Asp-N	1275.50	1275.55
$7-19^3$	YSγVTPTHTMCLT	trypsin + Asp-N	1583.68	1583.65
$-CO_2$	YS*E*VTPTHTMCLT	trypsin + Asp-N	1539.71	1539.66

[a] The superscript numbers 1, 2, and 3, indicate isoform 1, isoform 2, and isoform 3 (GlaCrisp), respectively. [b] The proteins were reduced with DTT and alkylated with iodoacetic acid prior to endoproteinase digestion. [c] Peptide fragments were methyl-esterfied prior to nanoESI-MS analysis.

might indicate a functional role of this unusual amino acid in these proteins, and it is possible that also other PR proteins contain this modification.

DISCUSSION

This work identified a novel invertebrate 30 kDa Gla protein with homologues in species from mammals to insects and plants. The protein, here referred to as GlaCrisp, was purified from the venom of the mollusc-hunting cone snail *C. marmoreus*. cDNA cloning of GlaCrisp deduced a primary translation product with a signal peptide sequence of 22 amino acids immediately followed by a dibasic proteolytic cleavage site and a mature protein sequence of 265 amino acids. Thus, unlike all previously identified Gla proteins, the predicted precursor polypeptide contained no apparent propeptide sequence. In the case of the conotoxin precursors that are substrates for the *Conus* carboxylase, the propeptide N-terminal to the mature toxin region to be modified has been shown to contain the γ-CRS that is required for carboxylation (*15*, *16*). Similarly, γ-carboxylation of the mammalian Gla proteins requires a γ-CRS within the propeptide sequence, N-terminal to the Glu residues to be converted to Gla (*3*, *5*, *45*−*48*). In one case, MGP, the propeptide-like sequence containing the γ-CRS is part of the mature form of the protein (*49*). There are sequence similarities among the propeptides of all known mammalian Gla proteins, and a γ-CRS motif has been defined by the sequence Z-F-Z-X-X-X-X-A (where X is any amino acid, and Z is Ile, Val, or Leu) (*3*). However, the propeptides of the conotoxin precursors bear no sequence similarity to the mammalian carboxylase binding sites. Recently, a C-terminal postpeptide was demonstrated to contain the γ-CRS for carboxylation of two novel conotoxins (*17*). On the basis of the post- and propeptide sequences from all known Gla-containing conotoxins, a possible consensus sequence was suggested including two basic residues and one hydrophobic residue, K/R-X-X-J-X-X-X-X-K/R (where X is any amino acid and J is a hydrophobic residue). Attempts in this work to identify a recognition element within the mature region of GlaCrisp on the basis of the defined carboxylase recognition site in either the mammalian Gla proteins or the *Conus* conotoxins failed. Presumably GlaCrisp and the previously identified Gla-containing proteins/peptides have a separate mechanism for substrate recognition. Identification of which sequence within the mature region of GlaCrisp plays a critical role for recognition in the carboxylation reaction will require further investigation.

GlaCrisp is to our knowledge the first Gla-containing high molecular weight protein reported to contain a single Gla residue. With the exception of BGP (containing 3 Gla residues) and MGP (containing 4−5 Gla residues depending on the species), all known vertebrate Gla proteins contain 9−13 Gla residues within a structurally conserved Gla domain of ∼45 amino acids (*3*). In these proteins the Gla residues bind Ca^{2+}, which generates a structural transition that is crucial for the biological function of the proteins (*4*). Similarly, the Ca^{2+}-binding properties and the three-dimensional structure of some of the Gla-containing *Conus* peptides suggest a specific structural role for Gla (*50*−*54*). Binding of Ca^{2+} to the Gla residues in conantokin-G (containing 5 Gla residues) introduces α-helicity to the peptide, which is important for binding of the peptide to NMDA receptors. In Glacontryphan-M, the first Gla peptide isolated from *C. marmoreus*, the two Gla residues provide the peptide with Ca^{2+}-binding properties that are crucial for the interaction of the conotoxin with L-type Ca^{2+} channels (*26*). In GlaCrisp it is conceivable that the single Gla residue in conjunction with the carboxyl group of an Asp residue (for instance, Asp 4 and Asp 20) can participate in Ca^{2+}-binding. Given that the γ-glutamyl carboxylase activity is phylogenetically older than blood coagulation and tissue mineralization, Gla most likely plays a more critical role in biology than has so far been discovered (*12*−*14*, *18*, *55*). GlaCrisp with its single Gla residue may be a particularly favorable protein for investigating functions of this post-translational modification, other than Ca^{2+}-binding.

The predicted product of the GlaCrisp full-length cDNA clearly is a member of the widely distributed CRISP family found in a variety of mammalian tissues and venoms of lizards and snakes (Figure 6) (*56*, *57*). The CRISP proteins are modular, each containing two potentially functional domains, an N-terminal SCP domain and a C-terminal cysteine-rich domain (CRD), that are connected by a hinge region. The N-terminal region is a homologue to members of the pathogenesis-related protein superfamily 1 (PR-1) and the insect venom antigen 5 proteins (VA5) that are involved in plant-stress responses and venom allergens, respectively (*35*, *58*). The precise biological functions of the CRISP family proteins are not well defined at present. Among these

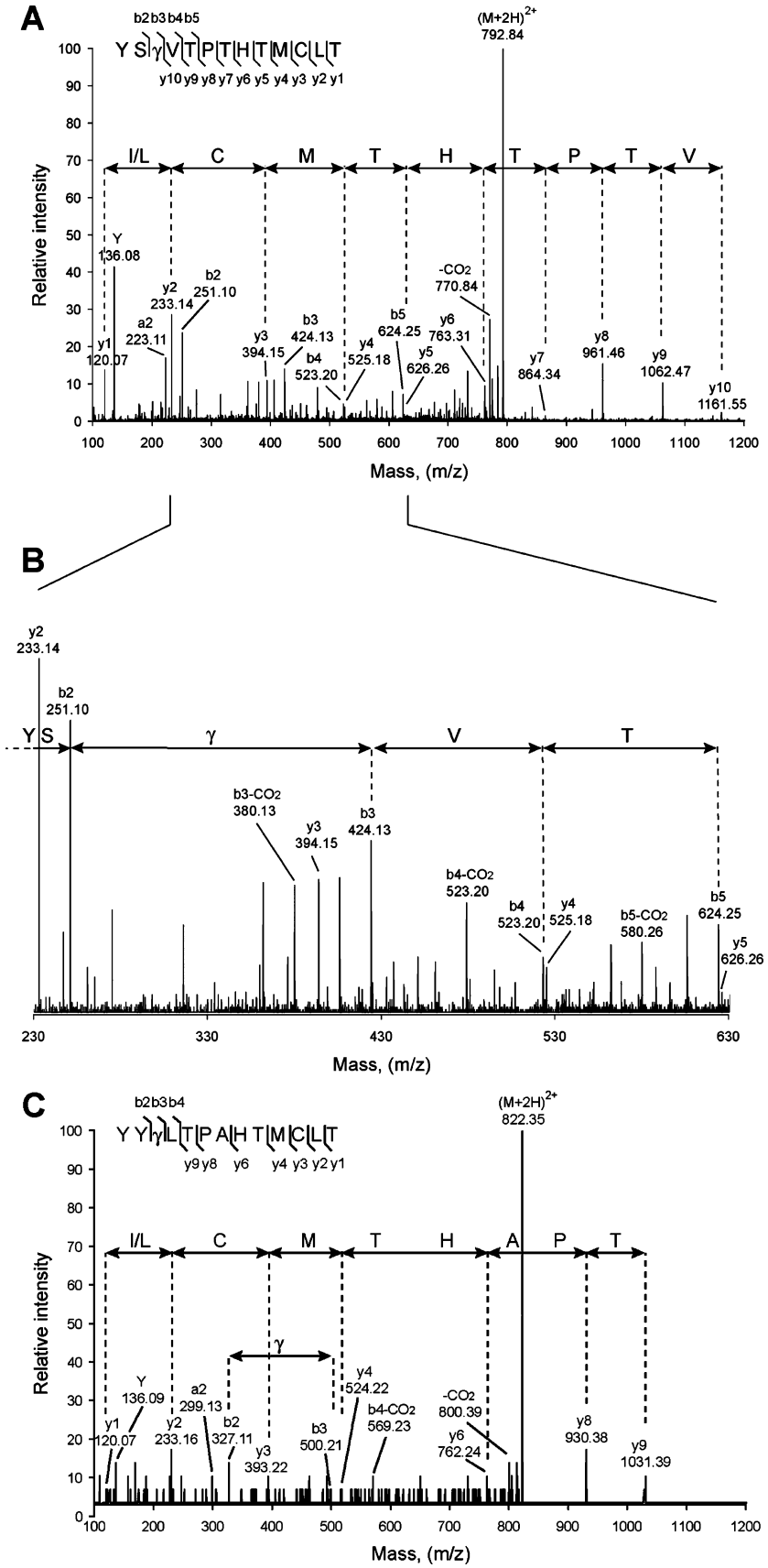

FIGURE 5: Positive ion nanoESI-MS/MS analysis of the doubly charged peptides at *m*/*z* 792.84 and 822.35 derived from digestions of GlaCrisp and Tex31, respectively, with trypsin and Asp-N. (A) The b and y ions according to the proposed sequence of the S-carboxymethylated internal peptide that covers amino acids 7−19 of GlaCrisp are labeled. (B) The *m*/*z* region from 230 to 630 has been expanded to show the b2, b3, b4, and b5 ions from which the N-terminal sequence of the peptide of GlaCrisp, including the Gla residue, could be determined. (C) Some of the b and y ions of the S-carbamidomethylated internal peptide fragment containing Gla of Tex31 are labeled. The presence of Gla in the peptides of GlaCrisp and Tex31 is notable in the spectra via decarboxylation of the peptides and product ion fragments, resulting in a loss of 44 Da.

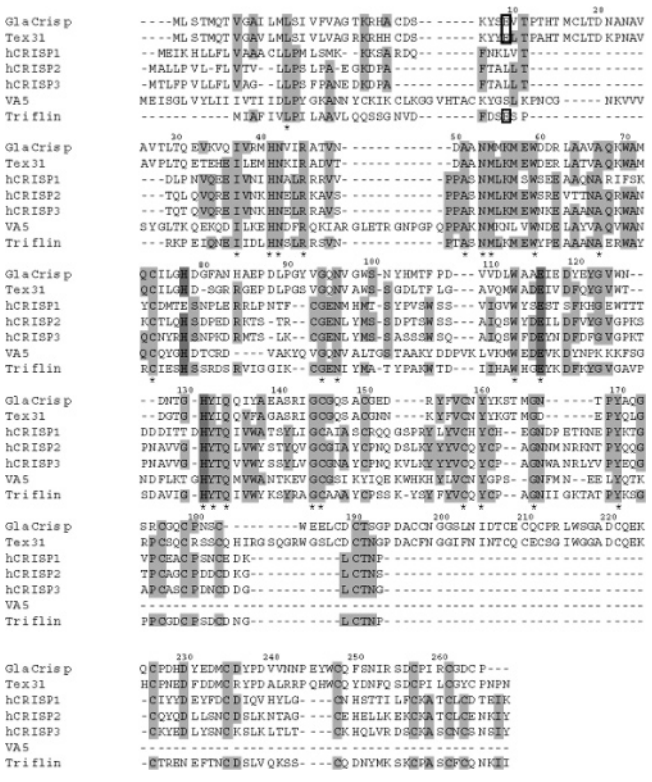

FIGURE 6: Amino acid sequence alignment of GlaCrisp with PR-1-type proteins. Sequences were aligned using Clustal W. Amino acid residues that are identical in at least four of the proteins are marked in light gray. The positions of amino acids that were suggested as potential active site residues are indicated in dark gray. Asterisks indicate conserved amino acid residues. The Glu residues in GlaCrisp and Tex31 that are posttranslationally carboxylated to Gla and the Glu residues in the corresponding position in Triflin are boxed. Key: Tex31, *C. textile* (SwissProt entry Q7YT83); hCRISP1, human cysteine-rich secretory protein 1 (SwissProt entry P54107); hCRISP2, human cysteine-rich secretory protein 2 (SwissProt entry P16562); hCRISP3, human cysteine-rich secretory protein 3 (SwissProt entry P54108); VA5, *Vespula vulgaris* venom allergen 5 (SwissProt entry Q05110), Triflin, *Trimeresurus flavoviridis* (SwissProt entry Q8JI39).

proteins GlaCrisp displays the highest sequence similarity to Tex31, which was shown to provide proteolytic activity. At this time, the molecular target of GlaCrisp remains unknown. Preliminary tests to determine whether the protein possesses a proteolytic activity similar to that of Tex31 were negative. In these experiments a peptide substrate based on the proposed Tex31 propeptide substrate was used.

Comparison of GlaCrisp to representative members of the CRISP family revealed that the Gla protein, similarly to Tex31, contains an extended CRD domain and the three highly conserved residues at His 78, Glu 116, and His 131 that have been suggested as potential active site residues (Figure 6) (*59*−*61*). Milne et al. suggested that Tex31 might exhibit serine protease activity and proposed Ser 80, Glu 115, and His 130 (corresponding to Phe 81, Glu 116, and His 131 in GlaCrisp) as a protease-like triad on the basis of their proximity and placement (*33*). A serine nucleophile residue, corresponding to that in Tex31, is however not present in GlaCrisp, which might explain the negative results observed in the proteolytic assays. The identified Gla residue in GlaCrisp is located near the N-terminal end of the mature protein in a region that is not highly conserved. Therefore, the presence of a potential Gla residue in the homologues cannot be predicted. We also showed that a partially purified sample of Tex31 contains a Gla residue at the position corresponding to Gla 9 in GlaCrisp. Interestingly, the protease activity of Tex31 was reported to depend on Ca^{2+} (*33*). Whether the Gla residue is involved in Ca^{2+}-binding and is crucial for the biological function of the protein remains to be established.

The major implication of this work is the identification and purification of a novel Gla-containing protein that may provide insight into a new class of Gla-containing proteins with homologues broadly distributed in animal phyla. Moreover, the absence of an apparent propeptide sequence suggests a separate recognition mechanism for γ-carboxylation in the novel Gla protein.

ACKNOWLEDGMENT

We thank Ingrid Dahlqvist for performing sequence analyses and synthesis of peptides and Dr. Eva Czerwiec for providing the *C. marmoreus* cDNA library.

REFERENCES

1. Stenflo, J., and Suttie, J. W. (1977) Vitamin K-dependent formation of gamma-carboxyglutamic acid, *Annu. Rev. Biochem. 46*, 157−172.
2. Suttie, J. W. (1985) Vitamin K-dependent carboxylase, *Annu. Rev. Biochem. 54*, 459−477.
3. Furie, B., Bouchard, B. A., and Furie, B. A. (1999) Vitamin K-dependent biosynthesis of gamma-carboxyglutamic acid, *Blood 93*, 1798−1808.
4. Stenflo, J. (1999) Contributions of Gla and EGF-like domains to the function of vitamin K-dependent coagulation factors, *Crit. Rev. Eukaryotic Gene Expression 9*, 59−88.
5. Knobloch, J. E., and Suttie, J. W. (1987) Vitamin K-dependent carboxylase. Control of enzyme activity by the "propeptide" region of factor X, *J. Biol. Chem. 262*, 15334−15337.
6. Stanley, T. B., Wu, S.-M., Houben, R. J. T. J., Mutucumarana, V. P., and Stafford, D. W. (1998) Role of the propeptide and gamma-glutamic acid domain of factor IX for in vitro carboxylation by the vitamin K-dependent carboxylase, *Biochemistry 37*, 13262−13268.
7. Luo, G., Ducy P., McKee, M. D., Pinero, G. J., Loyer, E., Behringer, R. R., and Karsenty, G. (1997) Spontaneous calcification of arteries and cartilage in mice lacking matrix GLA protein, *Nature 386*, 78−81.
8. Melaragno, M. G., Fridell, Y. W., and Berk, B. C. (1999) The Gas6/Axl system: a novel regulator of vascular cell function, *Trends Cardiovasc. Med. 9*, 250−253.
9. Hoang, Q. Q., Sicheri, F., Howard, A. J., and Yang, D. S. (2003) Bone recognition mechanism of porcine osteocalcin from crystal structure, *Nature 425*, 977−980.
10. McIntosh, J. M., Olivera, B. M., Cruz, L. J., and Gray, W. R. (1984) Gamma-carboxyglutamate in a neuroactive toxin, *J. Biol. Chem. 259*, 14343−14346.
11. Olivera, B. M., Rivier, J., Clark, C., Ramilo, C. A., Corpuz, G. P., Abogadie, F. C., Mena, E. E., Woodward, S. R., Hillyard, D. R., and Cruz, L. J. (1990) Diversity of Conus neuropeptides, *Science 20*, 257−263.
12. Begley, G. S., Furie, B. C., Czerwiec, E., Taylor, K. L., Furie, G. L., Bronstein, L., Stenflo, J., and Furie, B. (2000) A conserved motif within the vitamin K-dependent carboxylase gene is widely distributed across animal phyla, *J. Biol. Chem. 275*, 356245−36249.
13. Bandyopadhyay, P. K., Garrett, J. E., Shetty, R. P., Keate, T., Walker, C. S., and Olivera, B. M. (2002) Gamma-glutamyl carboxylation: An extracellular posttranslational modification that antedates the divergence of molluscs, arthropods, and chordates, *Proc. Natl. Acad. Sci. U.S.A. 99*, 1264−1269.
14. Czerwiec, E., Begley, G. S., Bronstein, M., Stenflo, J., Taylor, K., Furie, B. C., and Furie, B. (2002) Expression and characterization of recombinant vitamin K-dependent gamma-glutamyl carboxylase from an invertebrate, Conus textile, *Eur. J. Biochem. 269*, 6162−6172.

15. Bandyopadhyay, P. K., Colledge, C. J., Walker, C. S., Zhou, L., Hillyard, D. R., and Olivera, B. M. (1998) Conantokin-G precursor and its role in gamma-carboxylation by a vitamin K-dependent carboxylase from a Conus snail, *J. Biol. Chem. 273*, 5447−5450.
16. Bush, K., Stenflo, J., Roth, D. A., Czerwiec, E., Harrist, A., Begley, G. S., Furie, B. C., and Furie, B. (1999) Hydrophobic amino acids define the carboxylation recognition site in the precursor of the gamma-carboxyglutamic-acid-containing conotoxin epsilon-TxIX from the marine cone snail Conus textile, *Biochemistry 38*, 14660−14666.
17. Brown, M. A., Begley, G. S., Czerwiec, E., Stenberg, L. M., Jacobs, M., Kalume, D. E., Roepstorff, P., Stenflo, J., Furie, B. C., and Furie, B. (2005) Precursors of novel Gla-containing conotoxins contain a carboxy-terminal recognition site that directs gamma-carboxylation, *Biochemistry 44*, 9150−9159.
18. Li, T., Yang, C. T., Jin, D., and Stafford, D. W. (2000) Identification of a Drosophila vitamin K-dependent gamma-glutamyl carboxylase, *J. Biol. Chem. 275*, 18291−18296.
19. Terlau, H, and Olivera, B. M. (2004) Conus venoms: a rich source of novel ion channel-targeted peptides, *Physiol Rev. 84*, 41−68.
20. Haack, J. A., Rivier, J., Parks, T. N., Mena, E. E., Cruz, L. J., and Olivera, B. M. (1990) Conantokin-T. A gamma-carboxyglutamate containing peptide with N-methyl-d-aspartate antagonist activity, *J. Biol. Chem. 265*, 6025−6029.
21. Fainzilber, M., Gordon, D., Hasson, A., Spira, M. E., and Zlotkin, E. (1991) Mollusc-specific toxins from the venom of Conus textile neovicarius, *Eur. J. Biochem. 202*, 589−595.
22. Nakamura, T., Yu, Z., Fainzilber, M., and Burlingame, A. L. (1996) Mass spectrometric-based revision of the structure of a cysteine-rich peptide toxin with gamma-carboxyglutamic acid, TxVIIA, from the sea snail, Conus textile, *Protein Sci. 5*, 524−530.
23. Fainzilber, M., Nakamura, T., Lodder, J. C., Zlotkin, E., Kits, K. S., and Burlingame, A. L. (1998) gamma-Conotoxin-PnVIIA, a gamma-carboxyglutamate-containing peptide agonist of neuronal pacemaker cation currents, *Biochemistry 37*, 1470−1477.
24. Rigby, A. C., Lucas-Meunier, E., Kalume, D. E., Czerwiec, E., Hambe, B., Dahlqvist, I., Fossier, P., Baux, G., Roepstorff, P., Baleja, J. D., Furie, B. C., Furie, B., and Stenflo, J. (1999) A conotoxin from Conus textile with unusual posttranslational modifications reduces presynaptic Ca2+ influx, *Proc. Natl. Acad. Sci. U.S.A. 96*, 5758−5763.
25. Kalume, D. E., Stenflo, J., Czerwiec, E., Hambe, B., Furie, B. C., Furie, B., and Roepstorff, P. (2000) Structure determination of two conotoxins from Conus textile by a combination of matrix-assisted laser desorption/ionization time-of-flight and electrospray ionization mass spectrometry and biochemical methods, *J. Mass Spectrom. 35*, 145−156.
26. Hansson, K., Ma, X., Eliasson, L., Czerwiec, E., Furie, B., Furie B. C., Rorsman, P., and Stenflo J. (2004) The first gamma-carboxyglutamic acid-containing contryphan. A selective L-type calcium ion channel blocker isolated from the venom of Conus marmoreus, *J. Biol. Chem. 279*, 32453−32463.
27. Hansson K, Furie B, Furie BC, Stenflo J. (2004) Isolation and characterization of three novel Gla-containing Conus marmoreus venom peptides, one with a novel cysteine pattern, *Biochem. Biophys. Res. Commun. 319*, 1081−1087.
28. Kobayashi, J., Nakamura, H., and Ohizumi, Y. (1985) Potent excitatory effects of a cardiotonic protein from the venom of the marine snail Conus magus on the guinea-pig left atria, vas deferens and ileum, *Toxicon 23*, 783−789.
29. Gonoi, T., Ohizumi, Y., Kobayashi, J., Nakamura, H., and Catterall, W. A. (1987) Actions of a polypeptide toxin from the marine snail Conus striatus on voltage-sensitive sodium channels, *Mol. Pharmacol. 32*, 691−698.
30. Saxena, V. K., Partoens, P., de Block, J., Coen, E. P., Vauquelin, G., and De Potter, W. P. (1992) Inhibition of evoked neurotransmitter release from rat hippocampus by a polypeptide toxin isolated from the marine snail Conus distans, *Neurochem. Int. 20*, 69−74.
31. McIntosh, J. M., Ghomashchi, F., Gelb, M. H., Dooley, D. J., Stoehr, S. J., Giordani, A. B., Naisbitt, S. R., and Olivera, B. M. (1995) Conodipine-M, a novel phospholipase A2 isolated from the venom of the marine snail Conus magus, *J. Biol. Chem. 270*, 3518−3526.
32. Bulaj, G., Buczek, O., Goodsell, I., Jimenez, E. C., Kranski, J., Nielsen, J. S., Garrett, J. E., and Olivera, B. M. (2003) Efficient oxidative folding of conotoxins and the radiation of venomous cone snails, *Proc. Natl. Acad. Sci. U.S.A. 100*, 14562−14568.
33. Milne, T. J., Abbenante, G., Tyndall, J. D., Halliday, J., and Lewis, R. J. (2003) Isolation and characterization of a cone snail protease with homology to CRISP proteins of the pathogenesis-related protein superfamily, *J. Biol. Chem. 278*, 31105−31110.
34. Eberspaecher, U., Roosterman, D., Kratzschmar, J., Haendler, B., Habenicht, U. F., Becker, A., Quensel, C., Petri, T., Schleuning, W. D., and Donner, P. (1995) Mouse androgen-dependent epididymal glycoprotein CRISP-1 (DE/AEG): isolation, biochemical characterization, and expression in recombinant form, *Mol. Reprod. Dev. 42*, 157−172.
35. Van Loon, L. C. and Van Strien, E. A. (1999) The families of pathogenesis-related proteins, their activities, and comparative analysis of PR-1 type proteins, *Physiol. Mol. Plant Pathol. 55*, 85−97.
36. Kratzschmar, J., Haendler, B., Eberspaecher, U., Roosterman, D., Donner, P., and Schleuning, W. D. (1996) The human cysteine-rich secretory protein (CRISP) family. Primary structure and tissue distribution of CRISP-1, CRISP-2 and CRISP-3, *Eur. J. Biochem. 236*, 827−836.
37. Udby, L., Cowland, J. B., Johnsen, A. H., Sorensen, O. E., Borregaard, N., and Kjeldsen, L. J. (2002) An ELISA for SGP28/CRISP-3, a cysteine-rich secretory protein in human neutrophils, plasma, and exocrine secretions, *Immunol. Methods 263*, 43−55.
38. Maeda, T., Nishida, J., and Nakanishi, Y. (1999) Expression pattern, subcellular localization and structure-function relationship of rat Tpx-1, a spermatogenic cell adhesion molecule responsible for association with Sertoli cells, *Dev., Growth Differ. 41*, 715−722.
39. Cohen, D. J., Ellerman, D. A., Busso, D., Morgenfeld, M. M., Piazza, A. D., Hayashi, M., Young, E. T., Kasahara, M., and Cuasnicu, P. S. (2001) Evidence that human epididymal protein ARP plays a role in gamete fusion through complementary sites on the surface of the human egg, *Biol. Reprod. 65*, 1000−1005.
40. Brown, M. A., Stenberg, L. M., Persson, U., and Stenflo, J. (2000) Identification and purification of vitamin K-dependent proteins and peptides with monoclonal antibodies specific for gamma-carboxyglutamyl (Gla) residues, *J. Biol. Chem. 275*, 19795−19802.
41. Cairns, J. R., Williamson, M. K., and Price, P. A. (1991) Direct identification of gamma-carboxyglutamic acid in the sequencing of vitamin K-dependent proteins, *Anal. Biochem. 199*, 93−97.
42. Stenflo, J., Fernlund, P., Egan, W., and Roepstorff, P. (1974) Vitamin K dependent modifications of glutamic acid residues in prothrombin, *Proc. Natl. Acad. Sci. U.S.A. 71*, 2730−2733.
43. Fernlund, P., Stenflo, J., Roepstorff, P., and Thomsen, J. (1975) Vitamin K and the biosynthesis of prothrombin. V. Gamma-carboxyglutamic acids, the vitamin K-dependent structures in prothrombin, *J. Biol. Chem. 250*, 6125−6133.
44. Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels, *Anal. Chem. 68*, 850−858.
45. Pan, L. C., and Price, P. A. (1985) The propeptide of rat bone gamma-carboxyglutamic acid protein shares homology with other vitamin K-dependent protein precursors, *Proc. Natl. Acad. Sci. U.S.A. 82*, 6109−6113.
46. Jorgensen, M. J., Cantor, A. B., Furie, B. C., Brown, C. L., Shoemaker, C. B., and Furie, B. (1987) Recognition site directing vitamin K-dependent gamma-carboxylation resides on the propeptide of factor IX, *Cell 48*, 185−191.
47. Foster, D. C., Rudinski, M. S., Schach, B. G., Berkner, K. L., Kumar, A. A., Hagen, F. S., Sprecher, C. A., Insley, M. Y., and Davie, E. W. (1987) Propeptide of human protein C is necessary for gamma-carboxylation, *Biochemistry 26*, 7003−7011.
48. Hubbard, B. R., Jacobs, M., Ulrich, M. M., Walsh, C., Furie, B., and Furie, B. C. (1989) Vitamin K-dependent carboxylation. In vitro modification of synthetic peptides containing the gamma-carboxylation recognition site, *J. Biol. Chem. 264*, 14145−14150.
49. Price, P. A., Fraser, J. D., Metz-Virca, G. (1987) Molecular cloning of matrix Gla protein: implications for substrate recognition by the vitamin K-dependent gamma-carboxylase, *Proc. Natl. Acad. Sci. U.S.A. 84*, 8335−8339.
50. Prorok, M., Warder, S. E., Blandl, T., and Castellino, F. J. (1996) Calcium binding properties of synthetic gamma-carboxyglutamic acid-containing marine cone snail "sleeper" peptides, conantokin-G and conantokin-T, *Biochemistry 35*, 16528−16534.
51. Rigby, A. C., Baleja, J. D., Furie, B. C., and Furie, B. (1997) Three-dimensional structure of a gamma-carboxyglutamic acid-

containing conotoxin, conantokin G, from the marine snail Conus geographus: the metal-free conformer, *Biochemistry 36*, 6906−6914.
52. Rigby, A. C., Baleja, J. D., Li, L., Pedersen, L. G., Furie, B. C., and Furie B. (1997) Role of gamma-carboxyglutamic acid in the calcium-induced structural transition of conantokin G, a conotoxin from the marine snail Conus geographus, *Biochemistry 36*, 15677−15684.
53. Blandl, T., Warder, S. E., Prorok, M., and Castellino, F. J. (2000) Structure-function relationships of the NMDA receptor antagonist peptide, conantokin-R, *FEBS Lett. 470*, 139−146.
54. Grant, M. A., Hansson, K., Furie, B. C., Furie, B., Stenflo, J., and Rigby, A. C. (2004) The metal-free and calcium-bound structures of a gamma-carboxyglutamic acid-containing contryphan from Conus marmoreus, glacontryphan-M, *J. Biol. Chem. 279*, 32464−32473.
55. Walker, C. S., Shetty, R. P., Clark, K., Kazuko, S. G., Letsou, A., Olivera, B. M., Bandyopadhyay, P. K. (2001) On a potential global role for vitamin K-dependent gamma-carboxylation in animal systems. Evidence for a gamma-glutamyl carboxylase in Drosophila, *J. Biol. Chem. 276*, 7769−7774.
56. Morrissette, J., Kratzschmar, J., Haendler, B., el-Hayek, R., Mochca-Morales, J., Martin, B. M., Patel, J. R., Moss, R. L., Schleuning, W. D., and Coronado, R. (1995) Primary structure and properties of helothermine, a peptide toxin that blocks ryanodine receptors, *Biophys. J. 68*, 2280−2288.
57. Yamazaki, Y., Hyodo, F., and Morita, T. (2003) Wide distribution of cysteine-rich secretory proteins in snake venoms: isolation and cloning of novel snake venom cysteine-rich secretory proteins, *Arch. Biochem. Biophys. 412*, 133−141.
58. Lu, G., Villalba, M., Coscia, M. R., Hoffman, D. R., and King, T. P. (1993) Sequence analysis and antigenic cross-reactivity of a venom allergen, antigen 5, from hornets, wasps, and yellow jackets, *J. Immunol. 150*, 2823−2830.
59. Fernandez, C., Szyperski, T., Bruyere, T., Ramage, P., Mosinger, E., and Wuthrich, K. (1997) NMR solution structure of the pathogenesis-related protein P14a, *J. Mol. Biol. 266*, 576−593.
60. Szyperski, T., Fernandez, C., Mumenthaler, C., and Wuthrich, K. (1998) Structure comparison of human glioma pathogenesis-related protein GliPR and the plant pathogenesis-related protein P14a indicates a functional link between the human immune system and a plant defense system, *Proc. Natl. Acad. Sci. U.S.A. 95*, 2262−2266.
61. Henriksen, A., King, T. P., Mirzam O., Monsalve, R. I., Meno, K., Ipsen, H., Larsen, J. N., Gajhede, M., and Spangfort, M. D. (2001) Major venom allergen of yellow jackets, Ves v 5: structural characterization of a pathogenesis-related protein superfamily, *Proteins 45*, 438−448.

BI061311A