A common motif in proparts of Cnidarian toxins and nematocyst collagens and its putative role

Gregor Anderluh *, Zdravko Podlesek, Peter Maček

Department of Biology, Biotechnical Faculty, University of Ljubljana, Večna pot 111, 1000 Ljubljana, Slovenia

Received 3 September 1999; received in revised form 15 October 1999; accepted 10 November 1999

Abstract

In Cnidarians, cnidoblast cells contain organelles called cnidocysts, which are believed to be the product of an extremely complex regulated secretory pathway. When matured, these stinging organelles are capable of storing and delivering toxins. We hypothesized that translated nematocyst proteins might comprise specific sequences serving as signals in sorting to the organelle. A sodium channel neurotoxin from the sea anemone Actinia equina was cloned and the toxin precursor sequence was compared to those of nematocyst collagens, pore-forming toxins and ion channel neurotoxins. It was found that all the analyzed sequences possess a highly conserved stretch of nine amino acid residues ending with Lys-Arg N-terminally of the mature region. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Cnidaria; Pore-forming toxin; Collagen; Regulated sorting pathway; Sorting signal; Sodium channel neurotoxin

Cnidarians (Coelenterates), evolutionarily the most primitive metazoans, evolved a special weaponry called cnidocysts (nematocysts). These organelles are produced in cnidocytes (nematocytes), and are considered to be one of the most complex cellular secretory products. A number of different toxins have been isolated from them (for examples see [1–4]) providing evidence for a widely accepted belief that cnidocysts serve for storage and active delivery of venom [2,3,5]. The mature organelle consists of a collagen capsule, a long tubule connected with a stylet, with a cover which seals the cyst. Upon proper external stimulus, the cover opens, and the tubule evacinates at extreme speed, releasing the cyst contents into the victim [6,7]. In many respects, the formation of cnidocyst constituents and trafficking of their proteins is considered as an example of a very complex secretory apparatus. It is reasonable to anticipate that translated cnidocyst constituents and toxins could have a common signal directing them to a maturing cnidocyst. It is notable that all cloned cnidarian toxins [8–10] and nematocyst collagens [11,12] have a similar structural organization (Fig. 1A). Between the signal peptide and mature region they contain a propart 9–17 residues long, always terminating with Lys-Arg. The toxin proparp is composed of mainly polar (16.7%) and charged amino acid residues (57.5%). Collagen sequences have a slightly shorter propart of 8–12 amino acid residues. The composition favoring polar and charged residues, however, is preserved. The aim of this study was to search for an eventual motif in common to peptides and proteins of nematocyst origin.
From the sea anemone *Actinia equina* L. three different types of nematocyst toxins have already been isolated: very potent cytolysins, equinatoxins [5], and potassium [13] and sodium channel neurotoxins [14], but only equinatoxins were studied at cDNA level [9,15,16]. For the purpose of present study, we cloned an additional cDNA from *A. equina* cDNA library coding for the nematocyst derived precursor polypeptide, namely, sodium channel neurotoxin. The cDNA library from a single specimen of sea anemone *A. equina* [9] prepared in bacteriophage Vgt11 was used for PCR amplification. The library was first amplified with the oligonucleotides λgt11F (5’-GGTGCCAGCATCGCTGAGCC-3’) and λgt11R (5’-GACACCAGCAACTGGAATG-3’) in order to reduce false positive results after secondary PCR amplification. The mixture for amplification contained appropriate buffer, 50 pmol of each primer, and 2.5 U of Taq polymerase (New England Biolabs). Twenty-five cycles of amplification were performed at the following conditions: denaturation, 92°C for 1 min; annealing, 55°C for 1 min; extension at 72°C for 1 min. The smear between 300 and 1000 bp was excised, purified and used in the secondary PCR using Vgt11R as a sense oligonucleotide and degenerate AeNa (5’-CGGAATTCAYTGYTTRCARC-3’) as an antisense oligonucleotide. AeNa is complementary to the last five amino acid residues of an inhibitor of the sodium channel from *A. equina* [14] having a degeneracy of 32. DNA was amplified under the same conditions. The propart is doubly underlined. The protein sequence is numbered according to the presumed initiation methionine residue, the numbering of the mature toxin is in parentheses.

Fig. 1. The structural organization of cnidocyst proteins. (A) The structural organization of cnidarian toxins and nematocyst collagens. Cnidarian toxins are presented schematically. Black, the signal sequence; gray, the propart; and white, the mature region. A bar denotes a length of 10 amino acid residues. Clx1, calitoxin 1 from *Calliactis parasitica* [10]; EqtII, equinatoxin II from *Actinia equina* [9]; HmK, inhibitor of the potassium channel from *Heteractis magnifica* [8]; AeNa, inhibitor of sodium channel from *A. equina* (this work); N-Col, collagens from *Hydra magnipappilata* [11]; AdC1, collagen from *Acropora donei* [12]. (B) Nucleotide and amino acid sequences of *A. equina* sodium channel neurotoxin (AeNa). The signal peptide is underlined. The propart is doubly underlined. The protein sequence is numbered according to the presumed initiation methionine residue, the numbering of the mature toxin is in parentheses.

From the sea anemone *Actinia equina* L. three different types of nematocyst toxins have already been isolated: very potent cytolysins, equinatoxins [5], and potassium [13] and sodium channel neurotoxins [14], but only equinatoxins were studied at cDNA level [9,15,16]. For the purpose of present study, we cloned an additional cDNA from *A. equina* cDNA library coding for the nematocyst derived precursor polypeptide, namely, sodium channel neurotoxin. The cDNA library from a single specimen of sea anemone *A. equina* [9] prepared in bacteriophage λgt11 was used for PCR amplification. The library was first amplified with the oligonucleotides λgt11F (5’-GGTGCCAGCATCGCTGAGCC-3’) and λgt11R (5’-GACACCAGCAACTGGAATG-3’) in order to reduce false positive results after secondary PCR amplification. The mixture for amplification contained appropriate buffer, 50 pmol of each primer, and 2.5 U of Taq polymerase (New England Biolabs). Twenty-five cycles of amplification were performed at the following conditions: denaturation, 92°C for 1 min; annealing, 55°C for 1 min; extension at 72°C for 1 min. The smear between 300 and 1000 bp was excised, purified and used in the secondary PCR using λgt11R as a sense oligonucleotide and degenerate AeNa (5’-CGGAATTCAYTGYTTR-CARC-3’) as an antisense oligonucleotide. AeNa is complementary to the last five amino acid residues of an inhibitor of the sodium channel from the sea anemone *A. equina* [14] having a degeneracy of 32. DNA was amplified under the same conditions. The fragment of approximately 350 bp was excised from the gel, restricted with EcoRI, and cloned into the pUC19. Three independent clones
were sequenced in both directions using T7 DNA polymerase (Pharmacia) and $[^{35}S]$dATP (Amersham). Notably, we could not amplify the inhibitor for potassium channels [13] by the same procedure. Antisense AeK oligonucleotide (5'-CGGAATTCA-R-CAYTTNCCRCANGT-3') was used in the combination with $\lambda gt11F$ or $\lambda gt11R$ for the secondary PCR. AeK has a degeneracy of 128. This means that inhibitor was either not expressed in this particular specimen, or was expressed in such low amounts that it was not possible to amplify it with PCR.

The nucleotide and derived amino acid sequences of the 316-bp fragment obtained by PCR are presented in Fig. 1B. An open reading frame of 246-bp codes for a precursor protein of 82 amino acid residues. A potential translation start site is found at positions 68–70, with highly conserved adenine at positions −3 and +4 for the initiation of translation in vertebrates [17]. The initiation start provides translation of a typical signal peptide of 19 residues [18]. A further insertion of nine mostly negatively charged amino acid residues is found between the signal peptide and the mature peptide. It ends with a pair of basic residues Lys-Arg at positions 27–28, a potential proteolytic cleavage site. The deduced sequence of a 54 amino acid residues long mature protein matches the already published one obtained by protein sequencing [14].

In order to search for a motif within the precursors of nematocyst toxins and collagens, a MEME tool [19] was used. Twelve precursor sequences of pore-forming proteins, equinatoxins, inhibitors of ionic channels, calitoxins, HmK and AeNa and five collagens were used as an input data set. The most conserved sequence of 9 residues DEDEDIEKR was recognized in all proteins located N-terminally to the mature region (Fig. 2). The nine amino acid residues long motif ends with the highly conserved sequence Glu-Lys-Arg (similar to the same part around the dibasic recognition sequence discovered by Block Maker [20]). The $E$-values for the alignment of motifs and cnidarian sequences were from 6.7e-3 to 9.7e-9, indicating statistically significant results (usually $E<0.001$). This was the only motif common to all 12 proteins. None was recognized in 12 sequences if only mature parts were used as an input, further indicating that the motif must reside in the propart. Two sets of randomized sequences of the same length
were used for the input as a control. In both cases, no motif common to all 12 sequences was recognized. Similarly, no motif was discovered if shuffled sequences of our set were used as an input for MEME. Terminal Lys-Arg are conserved in all sequences. Dibasic sites are usually present before mature part of prohormones as a cleavage point for subtilisin like proteases, most commonly PC1 and PC2, in secretory granules. Such processing enzymes also exist in cnidarians: i.e. a subtilisin-like prohormone convertase has been cloned from Hydra attenuata [21].

In general, proparts have various functions in the precursors of proteins. Most important are assistance in folding, prevention of activity, and trafficking a precursor to different locations within the cell. The role of the propart in folding is not likely for cnidarian toxins since it was possible to obtain fully active toxins by expressing them in bacterial hosts as mature proteins without a propart [8,9,22]. Furthermore, fully active cnidarian potassium channel neurotoxins were chemically synthesized and properly folded [23,24]. The second possibility, that the propart prevents activity of toxins within cnidocytes is unlikely. Cnidarian toxins exhibit positive net charge and proparts might neutralize such a positive charge, disabling their interaction with lipid membranes. This was suggested for an acidic propart present in antibacterial peptides [25]. However, once the toxins arrive to the cnidocyst, this propart is no longer needed, since the interaction of toxins with the lipid membrane is prevented by a collagen wall. Furthermore, it has been proposed by some authors [8,10] that toxins are stored as inactive precursors, being processed prior to release after the proper stimulus. However, as toxins are released in less than 3 ms [7], it is unlikely that any modification could occur at this stage. They are more likely processed during cyst maturation. In addition, Meinardi et al. [26] proved that cytolsins from the sea anemone Phy- nactis clematis, belonging to the same family of lysins as equinatoxins, are inactive against the membranes of sea anemones. It was shown that the P. clematis membrane contains phosphonosphingolipids but not sphingomyelin which is considered a specific acceptor for that type of toxins [5]. Therefore, we rather suggest, that the propart may have a role in the regulated secretory pathway (RSP), i.e. the targeting of toxin precursors to secretory membranous structures involved in cnidocyst formation.

In RSP, proteins are stored in secretory vesicles and released from the cell after a proper stimulus [27]. Typical examples of proteins following the RSP include hormones and neuropeptides [28,29]. All are synthesized as long precursors, usually composed of a signal peptide, propart, and mature protein. They are processed during transportation to the secretory vesicles, first by signal peptidases and later with subtilisin-like endopeptidases. How proteins are sorted to RSP is not well understood. In general, signals for trafficking proteins to different cell compartments are diverse and include short, conserved sequences of amino acid residues, amphiphilic helices or surface patches. A specific signal common to all types of proteins which follow the RSP is not known.

Our suggestion on the putative biological role of the nematocyst motif is contrasted by the fact that no signal composed of linear amino acid sequence has yet been determined for proteins following RSP despite attempts to identify such a motif from aligned sequences of various prohormones [30]. Instead, a surface patch within the proregion has been proposed to be the signal for binding to carboxipeptidase E (CPE), a sorting receptor for POMC, proinsulin, and proenkephalin [31]. This 24-residue surface patch is located after the signal sequence, and consists of highly conserved two charged and two hydrophobic amino acid residues [32]. It is folded into the right conformation with two disulfide bonds. A similar structure of a N-terminally disulfide bridged loop of 22 amino acid residues has been recently proposed to be responsible for the sorting of chromogranin B to secretory vesicles [33]. Considering the proteins analyzed here, only N-Col4 possesses one Cys residue within the proregion, thus the needed bridge cannot be formed. Furthermore, none of the analyzed proparts is long enough to form an appropriate patch.

Similar precursor organization was also found in frog antimicrobial peptides of the magainin and dermaseptin family. They form a large heterogeneous family produced as precursors with a highly conserved signal peptide and a negatively charged propart [34]. They are stored in secretory granules within poisonous glands and are released after mechanical stimulus [35]. An acidic propart is usually longer, but
exhibits similar sequence around dibasic site (not shown).

To conclude, we determined the nucleotide sequence of sodium channel neurotoxin from the sea anemone *A. equina*. The analysis of its propart and those from other nematocyst toxins and collagens revealed a conserved stretch of nine residues similar to that found previously in precursors of antimicrobial peptides. We suggest that the motif may have a role in trafficking to RSP, which, however, calls for experimental confirmation.

This work was supported by a grant from the Ministry of Science and Technology of Slovenia. The authors would like to thank to Prof. Dr. Franc Gubenšek, Jožef Stefan Institute, Ljubljana, Slovenia, for enabling us to determine the DNA sequence of sodium channel inhibitor.

References