FEBS Letters 410 (1997) 490–492 FEBS 18794

cDNA cloning of Clavanins: antimicrobial peptides of tunicate hemocytes

Chengquan Zhao, Lilian Liaw, In Hee Lee, Robert I. Lehrer*

Department of Medicine, Section of Molecular Host Defense, UCLA School of Medicine, Box 951690, Los Angeles, CA 90095-1690, USA

Received 7 May 1997

Abstract Clavanins are a family of α-helical antimicrobial peptides found in hemocytes of the tunicate, *Styela clava*. We examined a cDNA library prepared from pharyngeal tissues of *S. clava* and sequenced 24 clones that encoded prepropeptides of Clavanins A, C, D or E. These sequences indicated that Clavanins are synthesized as 9.2 kDa prepropeptides which contain a 19-residue signal peptide, followed in turn by a highly polar 'pro' region (LEERKSEEEK) with five glutamic acid residues, the 23 residues of the mature Clavanin peptide, the glycine residue needed for its amidation and a 27-residue polar C-terminal extension that is removed in later processing. Although the signal sequence and anionic propiece of Clavanin precursors share features with corresponding regions in precursors of the certain frog peptides, including ranalexin, gaegurins, dermaseptins and deltorphins, their unique multipartite structure suggests that they are not actually homologues of these amphibian peptides.

Key words: cDNA; Antimicrobial peptide; Clavanin; *Styela clava*; Tunicate

1. Introduction

'Tadpole' larvae of protochordates such as the solitary tunicate *Styela clava*, display a constellation of features that reveals their ancestry to vertebrates [1]. These include the presence of pharyngeal gill slits, postanal tail, notochord and a dorsal tubular nerve cord. The body cavity and tissues of protochordates contains phagocytic cells ('hemocytes') that resemble the granulocytes and macrophages found in vertebrates in appearance and function [2]. The hemocytes of *S. clava* contain at least two families of antimicrobial peptides, Clavanins and Styelins [3,4]. Clavanins are histidine-rich peptides that contain 23-amino-acid residues, are α-helical and have an amidated C-terminus [3]. Styelins are somewhat larger peptides (≈3.6 kDa) that contain hydroxylysines and other modified residues [4]. Although other tunicate antimicrobial peptides have been recognized [5] including the unusual antimicrobial tetrapeptide (halocyamine) [6], no precursors of tunicate antimicrobial peptides have been previously described. We describe cloning studies that defined the structures of four Clavanin precursors, including a newly recognized member of this family, Clavanin E.

2. Materials and methods

2.1. RNA isolation and 3′-RACE analysis

Tunicate pharyngeal and hemocyte total RNAs were purified from live *Styela clava* (Marinus, Long Beach, CA) using a RNA separator kit (Clontech, Palo Alto, CA). Rapid 3′ amplification of cDNA end (3′-RACE) was performed with a kit (Gibco BRL, Gaithersburg, MD), using a degenerate 30-base primer for PCR amplification. This primer (5′-GTCGACTAGTCAYCAYTIGGIAAYTTYGT-3′) corresponded to seven invariant amino acids, His–His–Val–Gly–Asn–Phe–Val, which constituted residues 10–16 in Clavanins A–D. In the notation used above for the primer, Y represents either T or C, I signifies inosine and underlining shows the *Spe*I restriction site. The resulting ≈250 bp PCR product was subcloned into pCRScript SK vector (Stratagene, La Jolla, CA) and sequenced.

2.2. Styela clava cDNA library

Tunicate pharyngeal tissues were harvested from living *S. clava*, quickly frozen and stored at −70°C. A custom cDNA library was constructed in λ TripIEx® by Clontech Laboratories, using *E. coli* strain XL1-Blue as a host. Phage particle DNA was transferred to nylon membrane filters (Dupont, Boston, MA) which were hybridized overnight at 50°C with the ^{32}P-labeled ≈250 bp PCR product mentioned above in Rapid-hyb buffer (Amersham, Arlington Heights, IL). After several washes, and finally at 60°C in 0.1×SSC and 0.1% SDS, the filters were exposed to X-ray films at −70°C with an intensifying screen. After subjecting positive clones to additional rounds of plaque screening at low density, 80 positive clones were identified from approximately 1.2×10^5 plaques.

2.3. DNA purification, amplification and sequencing

λ phage DNA was purified using a Lambda kit (QIAGEN, Chatsworth, CA). Purified DNA or directly picked plaques were subjected to long-distance PCR using LD-Insert Screening Amplimers (Clontech Lab., Palo Alto, CA). PCR amplification was performed with Pfu DNA polymerase in a GeneAmp PCR system 2400 (Perkin-Elmer, Palo Alto, CA). PCR products of the inserts were purified from low-melting agarose gel, sequenced directly by the fluorescein-labeled, dideoxynucleotide terminator method, and analyzed on an Applied Biosystems 373A DNA Sequencer (Perkin-Elmer, Palo Alto, CA) at the UCLA DNA Sequencing Facility.

2.4. Northern blot analysis

Total RNA (10 μg) from *S. clava* hemocytes and pharyngeal tissues was denatured in 50% formamide, separated on a 1% formaldehyde gel, and transferred to Gene Screen Plus nylon membranes (DuPont, Boston, MA). A ≈250 bp PCR product containing amplified Clavanin 5′ side cDNA (from start codon to stop codon) was labeled with ^{32}P by random primer synthesis and hybridized to the immobilized RNA in rapid hybridization buffer, with the final wash in 0.1×SSC with 0.1% SDS at 60°C.

3. Results and discussion

We recently purified Clavanins A, B, C and D from hemocytes of the solitary tunicate, *S. clava*, and described their primary structures [3]. To characterize their precursors, we examined a cDNA library prepared from pharyngeal tissue, a repository of hemocytes as well as of digestive cells. By

*Corresponding author. Fax: (310) 206-8766.
E-mail: rlehrer@med1.medsch.ucla.edu

PII S0014-5793(97)00646-7

```
1   CAAACTCAGACAAACAACAGGAAAGATGAAAACAACAATTTTGATTCTTCTCATACTGGG
                                  M  K  T  T  I  L  I  L  L  I  L  G

                                                          ade  edd de
                                                           G   TCT A
61  ACTTGGCATCAATGCAAAATCTCTGGAGGAAAGAAAATCGGAGGAAGAAAAAGTATTCCA
      L  G  I  N  A  K  S  L  E  E  R  K  S  E  E  E  K  V  F  H
                                                        Ad/Le Qa/Kde

    ade a
    A T          Gd                                  Cae
121 T CTCCTTGGCAAAATTATTCATCATGTTGGCAATTTTGTATATGGTTTTAGCCACGTGTT
      L  L  G  K  I  I  H  H  V  G  N  F  V  Y  G  F  S  H  V  F
      Fa       Rd                                  Hae

                                                 Te
181 CGGCGACGACCAACAAGATAATGGAAAGTTTTATGGCCACTACGCAGAAGACAATGGCAA
      G  D  D  Q  Q  D  N  G  K  F  Y  G  H  Y  A  E  D  N  G  K
                                                 Ye

241 GCATTGGTATGATACCGGGGATCAATAAAAAAGTTTTAAACAGCTACGCGACTTGAAGAC
      H  W  Y  D  T  G  D  Q  ***

301 GGACGGACCCGGCAGAACATTGATATTTCTTGTTTTCTTTGATTAAAGGCTAGCCTTATT

361 ACTCAGAATATAACACTACATTGCATTC(A)n
          αααααα
```

Fig. 1. cDNA sequences of Clavanins A, C, D, and E. The main sequence is Clavanin C. Other Clavanin sequences are shown only where they differ from Clavanin C, and are identified in superscript ([a]Clv = Clavanin A, [d]Clv = Clavanin D, [e]Clv = Clavanin E). The mature peptide sequence is shown in boldface type. The signal sequence, propiece and post-piece are underlined with single and double and broken lines, respectively. The stop codon is indicated by asterisks and polyadenylation signal (underlined) by αααααα.

performing 3′-RACE analyses on pharyngeal tissue RNA with a degenerate primer based on a region of invariant sequence, we obtained a PCR product of ≈250 bp in length. When subcloned and sequenced, this was found to encode the 14 C-terminal amino acids found in both Clavanin C or D [3] plus additional 3′ side sequence.

We used this fragment to probe a *S. clava* pharyngeal cell cDNA library in λ TripIEx and identified 80 positives from approximately 1.2×10^5 clones. We determined the insert size of each positive clone by PCR, using a vector-specific insert screening primer, and sequenced 32 of these PCR products directly by the fluorescein-labelled dideosynucleotide terminator method. Twenty-six of these 32 clones encoded Clavanin isoforms: three Clavanin A; 12 Clavanin C; seven Clavanin D and two Clavanin E. Clavanins A, C and D were previously described [3] and Clavanin E was subsequently identified, purified and sequenced at the peptide level (unpublished). Two of the clones encoded a more distantly related isoform that we

Signal Sequence and Propiece

```
CLAVANIN A. MKTTILILLILGLGINAKSLEERKSEEEK
CLAVANIN C. MKTTILILLILGLGINAKSLEERKSEEEK
CLAVANIN D. MKTTILILLILGLGINAKSLEERKSEEEK
CLAVANIN E. MKTTILILLILGLGINAKSLEERKSEEEK
```

Mature Peptide and Anionic C-terminal Extension

```
A. VFQFLGKIIHHVGNFVHGFSHVFGDDQQDNGKFYGHYAEDNGKHWYDTGDQ
C. VFHLLGKIIHHVGNFVYGFSHVFGDDQQDNGKFYGHYAEDNGKHWYDTGDQ
D. AFKLLGRIIHHVGNFVYGFSHVFGDDQQDNGKFYGHYAEDNGKHWYDTGDQ
E. LFKLLGKIIHHVGNFVHGFSHVFGDDQQDNGKFYGYYAEDNGKHWYDTGDQ
```

Fig. 2. Primary structures of Clavanin precursors. The domains have been underlined as follows: signal sequence residues, no underline; propiece residues, double underline, mature peptide, single underline; post-piece residues, dotted underline. Anionic propiece and post-piece residues are bolded.

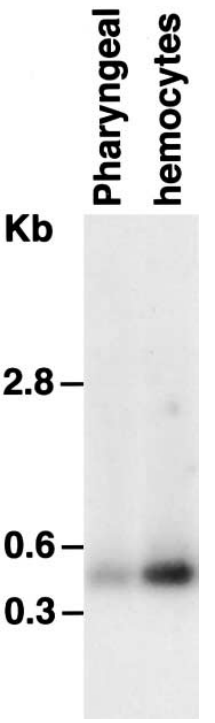

Fig. 3. Northern blot analysis. Total RNA was prepared from the pharyngeal tissues and hemocytes of *S. clava* and probed with ^{32}P-labelled Clavanin A 5′-side cDNA (from start to stop codon). Note that both tissues contained an mRNA species of ≈0.5 kb. Size markers are shown on the left.

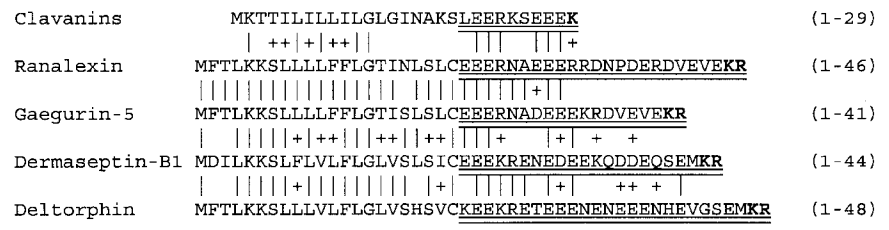

Fig. 4. Comparison of tunicate and frog peptides. The Clavanin signal sequence and anionic propiece are aligned with the corresponding regions of four frog prepropeptides: ranalexin, an antimicrobial peptide from *Rana catesbiana*[8], gaegurin-5, an antimicrobial peptide from *Rana rugosa* [9], dermaseptin B1, an antimicrobial peptide from *Phyllomedusa bicolor* [10] and deltorphin, a bioactive peptide precursor from *Phyllomedusa bicolor* [11]. Identical residues are connected by vertical lines and conservative substitutions are indicated by +. The anionic propieces have been doubly underlined and their monobasic (Clavanin) or dibasic cleavage-site residues are bolded. Residue numbers are shown in parentheses.

have named 'Clavaspirin'. This peptide is currently under study and will be described in a future report. We did not identify any clones encoding Clavanin B, whose primary differs from Clavanin A only by containing an arginine instead of a lysine at residue 7.

The cDNA sequences of Clavanins A, C, D and E are shown in Fig. 1. Each contained 388 bp including the 25-nucleotide-long 5′ untranslated end. Their 240-bp-long reading frame encoded an 80-amino-acid residue prepropeptide, with a calculated mass of ≈9.2 kDa. In each Clavanin precursor, the typical signal peptide (residues 1–19) was followed by a highly polar pro sequence (residues 20–29). Then came, in order, the mature peptide (residues 30–52), a glycine (residue 53) and a 27-residue polyanionic C-terminal extension (residues 54–80). Given the known structure of the mature peptides. [3], it is apparent that Gly^{53} contributes to the formation of the C-terminal amide [7] and the C-terminal extension is removed during processing.

As shown in Fig. 1, the cDNA sequences of Clavanins were highly conserved. Clavanin A and C transcripts differed by only 1 bp in the pro-sequence and by 3 bp in the region encoding the mature peptide, accounting for the three amino acid differences between the corresponding mature peptides (Fig. 2). The 6 bp differences between Clavanin C and D resulted in three amino acid differences between the corresponding peptides, of which two were in the mature peptide and one in the C-terminal extension. The 6 bp differences between Clavanins C and E were responsible for three amino acid differences between the mature peptides and single difference in the post-piece. The structural organization of these Clavanin precursors is also shown in Fig. 2. By Northern blot analysis, a ≈500 bp message was found in both pharyngeal tissue and hemocytes (Fig. 3) indicating that both tissues contained Clavanin-producing cells.

The signal sequences and polar propieces of Clavanin precursors resemble those of the precursors of certain peptides found in amphibia, including the antimicrobial peptides ranalexin [8], gaegurins [9] and dermaseptins [10] found in the skin of *Rana catesbiana, Rana rugosa* and *Phyllomedusa bicolor*, respectively (Fig. 4). However, whereas the anionic propieces of the frog peptides ranged from 19 (gaegurin-5) to 26 residues (deltorphin) in length and ended with a dibasic KR cleavage site, the Clavanin propiece was shorter (10 residues) and ended with a monobasic residue, lysine. Taken together with the long C-terminal extension (residues 54–80) that is found in Clavanins but not in the amphibian peptide precursors, these differences suggest that the precursors of tunicate Clavanins and amphibian skin peptides are not truly homologous. This inference, of course, should be re-visited once structural information about the respective tunicate and amphibian genes becomes available.

Acknowledgements: Dr. Lee received a Frederick B. Bang Scholarship from the American Association of Immunologists and a postdoctoral fellowship award from KOSEF, the Korea Science and Engineering Foundation. We thank Lorenzo Menzel for providing valuable technical assistance. Nucleotide sequences have been deposited with EMBL with the following accession numbers: Clavanin A, Y11020; Clavanin C, Y11019; Clavanin D, Y11018 Clavanin E, Y11017. Supported, in part, by a grant from the National Institutes of Health AI 22839.

References

[1] Bone, Q., The Origin of Chordates, 2nd edn., Carolina Biol. Supply Co., Burlington, NC, 1979.
[2] Sawada, T., Zhang, J. and Cooper, E.L. (1994) Biol. Bull. 184, 87–96.
[3] Lee, I.H., Zhao, C., Cho, Y., Harwig, S.S.L., Cooper, E.L. and Lehrer, R.I. (1997) FEBS Lett. 400, 158–162.
[4] I.H. Lee, Y. Cho, R.I. Lehrer, 1997, in press.
[5] Findlay, C. and Smith, V.J. (1995) J. Exp. Zool. 273, 434–444.
[6] Azumi, K., Yokosawa, H. and Ishii, S. (1990) Biochemistry 29, 159–165.
[7] Kulathila, R., Consalvo, A.P., Fitzpatrick, P.F., Freeman, J.C., Snyder, L.M., Villafranca, J.J. and Merkler, D.J. (1994) Arch. Biochem. Biophys. 311, 191–195.
[8] Clark, D.P., Durell, S., Maloy, W.L. and Zasloff, M. (1994) J. Biol. Chem. 269, 10849–10855.
[9] Park, J.M., Lee, J.Y., Moon, H.M. and Lee, B.J. (1995) Biochim. Biophys. Acta 1264, 23–25.
[10] Amiche, M., Ducancel, F., Mor, A., Boulain, J.C., Menez, A. and Nicolas, P. (1994) J. Biol. Chem. 269, 17847–17852.
[11] Richter, K., Egger, R., Negri, L., Corsi, R., Severini, C. and Kreil, G. (1990) Proc. NatL. Acad. Sci. USA 87, 4836–4839.