ELSEVIER

Toxicon 51 (2008) 218–229

TOXICON

www.elsevier.com/locate/toxicon

# $I_1$-superfamily conotoxins and prediction of single D-amino acid occurrence ☆, ☆☆

Olga Buczek[a,*], Elsie C. Jimenez[a,c], Doju Yoshikami[a], Julita S. Imperial[a], Maren Watkins[b], Alex Morrison[a], Baldomero M. Olivera[a]

[a]*Department of Biology, University of Utah, 257 South 1400 East, Salt Lake City, UT 84112, USA*
[b]*Department of Pathology, University of Utah, Salt Lake City, UT 84112, USA*
[c]*Department of Physical Sciences, University of the Philippines Baguio, Baguio City 2600, Philippines*

Received 7 June 2007; received in revised form 6 September 2007; accepted 25 September 2007
Available online 29 September 2007

**Abstract**

The considerable diversity of *Conus* peptides in the $I_1$-superfamily provides a rare opportunity to define parameters important for the post-translational L- to D-isomerization of amino acids. This subtlest of post-translational modifications is not readily detectable by most techniques, and it would be a considerable advance if one could predict its potential occurrence purely from gene sequences. We previously described three $I_1$-conotoxins, *ι*-RXIA (formerly designated r11a), r11b and r11c, each containing a D-amino acid at the third position from the C-terminus. In this work, we investigated two novel $I_1$-superfamily members, r11d and ar11a, which we show have only L-amino acids. Based on these observations and an analysis of cDNA sequences of other group members, we suggest that there is a rule to predict D-amino acids in $I_1$-superfamily peptides. Two factors are important: the residue to be modified should be three amino acids from the C-terminus of the precursor sequence, and it should be in a suitable sequence context. We apply the rule to other members of the $I_1$-superfamily, to determine *a priori* which are probably modified.


*Keywords:* Conotoxin; Epimerization; D-amino acid; Post-translational modifications; Excitatory peptide

## 1. Introduction

Venom peptides from cone snails (conotoxins) comprise a complex library of ~100,000 peptides containing the most highly post-translationally modified gene products known (for review see Buczek et al., 2005a). These modifications, in conjunction with hypermutation of amino acids between conserved cysteine residues, are responsible for the great molecular diversity of *Conus* sequences. The variety of modified conotoxins and an easy access to genomic information provide a

☆This work was supported by program Project Grant GM 48677 (to B.M.O.) from the National Institute of General Medical Sciences.

☆☆*Ethical statement*: All experiments conducted for this paper were completed according to high standards of scientific ethics. The results are true and accurate to the best of our knowledge. All use of animals in experiments were reviewed and approved by the University of Utah's Institutional Animal Care and Use Committee.

*Corresponding author. Tel.: +1 801 581 8370; fax: +1 801 585 5010.
*E-mail address:* buczek@biology.utah.edu (O. Buczek).

0041-0101/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.toxicon.2007.09.006

rare opportunity to identify general 'rules' for a number of post-translational modifications that allow an accurate prediction of their occurrences from the gene sequence. This is particularly important for modifications such as the conversion of an L- to a D-amino acid, which is a most subtle post-translational modification, very often not detectable by most modern bioanalytical techniques. Even if very advanced analyses are employed, such as the recently developed tandem mass spectrometry (MS/MS) utilizing both electron capture and collisionally activated dissociations (Adams and Zubarev, 2005), there are drawbacks such as limitations on the size of the polypeptide tested, the amount of material needed, requirement for internal standards, not to mention the involvement of costly instrumentation. Thus, we examined whether the occurrence of epimerization might be predicted purely from gene sequence information.

Although not many D-amino-acid-containing secreted polypeptides have been identified to date (very likely because of the above mentioned limitations), their potential significance for the biological function of the gene products in which they occur and the fact that the potency of these gene products depends on the chirality of specific amino acids cannot be neglected. Although epimerization is rather rare, it is surprisingly widespread in different phyla, including chordates (from amphibia to mammals), arthropods and mollusks (Monteccuchi et al., 1981; Kreil et al., 1989; Mor et al., 1989; Kreil, 1997; de Plater et al., 1998; Torres et al., 2002). In the last group, modified gene products are very well represented in the venom peptides of marine cone snails.

We previously described three related *Conus* peptides belonging to the $I_1$-superfamily, *ι*-RXIA, r11b and r11c, isolated from venom of fish-hunting snail *Conus radiatus* (Jimenez et al., 2003). Each peptide contained a single D-amino acid at the homologous position, which was always the third amino acid from the C-terminus of the precursor sequence. These were D-Phe$^{44}$ in *ι*-RXIA, D-Phe$^{44}$ in r11b and D-Leu$^{42}$ in r11c. This modification of a single residue to the D configuration was found to be critical for the biological activity of these peptides (Buczek et al., 2005b, c, 2007).

The key insight from our previous study of epimerization was that neither the nature of the side chain of the modified residue nor the vicinal sequence around it seemed to be important. However, there may be favored loci for isomerization of an amino acid to its D configuration (Buczek et al., 2005b, c). In this report we test our hypothesis and expand the database for predicting whether an amino acid is likely to be isomerized from the L to D configuration by examining two novel $I_1$-conotoxins: r11d from *C. radiatus* and ar11a from the worm-hunting *Conus arenatus* (for sequence comparison see Table 1). We have determined experimentally whether these peptides are $I_1$-superfamily members containing D-amino acid. Our new results,

Table 1
Comparison of $I_1$-conotoxin sequences determined from cDNA clones and from the sequencing of peptides purified from venom

| | cDNA/peptide sequence | Reference |
|---|---|---|
| R11.6 | GPSF**C**KADEKP**C**EYHAD**CC**N**CC**LSGI**C**APSTNWILPG**C**STSSFFKI | Jimenez et al. (2003) |
| *ι*-RXIA | *GOSF**C**KADEKO**C**EYHAD**CC**N**CC**LSGI**C**AOSTNWILPG**C**STSSFFKI* | Buczek et al. (2005b) |
| R11.14 | GPSF**C**KANGKP**C**SYHAD**CC**N**CC**LSGI**C**KPSTNVILPG**C**STSSFFRI | Jimenez et al. (2003) |
| r11b | *GOSF**C**KANGKO**C**SYHAD**CC**N**CC**LSGIC**C**KOSTNVILPG**C**STSSFFRI* | Buczek et al. (2005c) |
| R11.4 | GPSF**C**KADEKP**C**KYHAD**CC**N**CC**LGGI**C**KPSTSWI—G**C**STNVFLTR | Jimenez et al. (2003) |
| r11c | *GOSF**C**KADEKO**C**KYHAD**CC**N**CC**LGGI**C**KOSTSWI—G**C**STNVFLT* | Buczek et al. (2005c) |
| R11.8 | G—**C**KKDRKP**C**SYHAD**CC**N**CC**LSGI**C**APSTNWILPG**C**STSTFT | Jimenez et al. (2003) |
| r11d | G—*C**C**KKDRKO**C**SYHAD**CC**N**CC**LSGI**C**AOSTNWILPG**C**STSTFT* | This work |
| Ar11.1 | RT**C**SRRGHR**C**IRDSQ**CC**GGM**CC**QGNR**C**FVAIRR**C**FHLPF | This work |
| ar11a | RT**C**SRRGHR**C**IRDSQ**CC**GGM**CC**QGNR**C**FVAIRR**C**FH**L**PF | |

Peptides *ι*-RXIA (formerly designated r11a), r11b and r11c (encoded by R11.6, R11.14 and R11.4, respectively) contain a D-amino acid (underlined) at the third position from the C-terminus of precursor sequences. Peptide r11c is naturally synthesized with a C-terminal Arg as determined from cDNA clone, but the Arg is possibly cleaved off by carboxypeptidase during maturation. The *italic* residues show the sequence homology to *ι*-RXIA; ar11a has no extended sequence homology to the other peptides. Bold underline indicates possible epimerization. O, 4-*trans*-hydroxyproline.