Cloning and deorphanization of three inotocin (insect oxytocin/vasopressin-like) receptors and their ligand from the tick Ixodes scapularis

Hauser, Frank; Al-Ribaty, Tara; Stebegg, Marisa; Thygesen, Gedske; Grimmelikhuijzen, Cornelis J.P.

Published in:
Biochemical and Biophysical Research Communications

DOI:
10.1016/j.bbrc.2023.09.009

Publication date:
2023

Document version
Publisher's PDF, also known as Version of record

Document license:
CC BY

Citation for published version (APA):
Cloning and deorphanization of three inotocin (insect oxytocin/vasopressin-like) receptors and their ligand from the tick *Ixodes scapularis*

Frank Hauser, Tara Al-Ribaty, Marisa Stebegg, Gedske Thygesen, Cornelis J.P. Grimmelmikhuijzen

Section for Cell and Neurobiology, Department of Biology, University of Copenhagen, Universitetsparken 15, DK-2100, Copenhagen, Denmark

ARTICLE INFO

**Keywords:** Vasopressin Oxytocin Neurophysin Cystine-knot GPCR Evolution

**ABSTRACT**

Many insects produce the cyclic neuropeptide inotocin (CLITNCPRGamide), which is the insect orthologue of the mammalian neuropeptides oxytocin and vasopressin. These insects also have one inotocin G protein-coupled receptor (GPCR), which is the orthologue of the mammalian oxytocin and vasopressin receptors. The tick *Ixodes scapularis* belongs to the subphylum Chelicerata, an arthropod taxon different from insects, to which also spiders, scorpions, and mites belong. *I. scapularis* is an ectoparasite and a health risk for humans, because it transfers pathogenic microorganisms to its human host during a blood meal, thereby causing serious neurological diseases, among them Lyme disease and tick-borne encephalitis (TBE). By annotating the genomic sequence of *I. scapularis*, we previously found one presumed tick inotocin preprohormone gene and, in contrast to insects, three genes coding for presumed inotocin GPCRs. We now find that these GPCR genes cluster on one genomic contig, suggesting that they originated by recent gene duplications. Closely located on the same contig are also four adipokinetic hormone/corazonin-related peptide (ACP) GPCR genes, and one crustacean cardioactive peptide (CCAP) GPCR gene, suggesting evolutionary relationships. These evolutionary relationships are confirmed by phylogenetic tree analyses of their gene products. We also cloned the tick inotocin preprohormone, which has a structural organization closely resembling mammalian oxytocin and vasopressin preprohormones, including the presence of a conserved neurophysin sequence, having seven cystine bridges. This neurophysin sequence has two cystine-knot domains, but in contrast to mammalian neurophysins, the tick neurophysin contains a canonical prohormone convertase cleavage signal and a peptide C-terminal amidation sequence (GKR), suggesting cleavage into two biologically active cystine-knot peptides. This cleavage/amidation sequence occurs in neurophysins from most hard tick species, but not in other chelicerates. Mature tick inotocin is different from insect inotocin and has the sequence CFITNCPPGamide. Finally, we cloned and stably expressed the three tick inotocin receptors in Chinese Hamster Ovary cells and found that each of them was activated by nanomolar concentrations of tick inotocin (EC$_{50}$ for ITR1 = 1.6 × 10$^{-8}$ M; EC$_{50}$ for ITR2 = 5.8 × 10$^{-9}$ M; EC$_{50}$ for ITR3 = 9.3 × 10$^{-9}$ M), thereby establishing that they are genuine tick inotocin receptors.

1. Introduction

Most multicellular animals belong to two evolutionary lineages, the Deuterostomia (such as mammals, and other chordates), and the Protostomia (such as insects, molluscs and other invertebrates). Nearly all multicellular animals use neuropeptides and neuropeptide G protein-coupled receptors (GPCRs) for the control of reproduction, development, feeding, homeostasis, growth, and many other physiological processes. The first neuropeptides to be isolated from mammals were oxytocin and vasopressin, which were sequenced by the groups of Du Vigneau and Acher in 1953 [1-3]. Other neuropeptides followed in the 1960ies, such as the hypothalamic releasing hormones [4-6] and in the 1970ies, when additional gut neuroendocrine peptides were discovered [7]. In the 1980ies, it became clear that orthologues of some of the neuropeptides that first were discovered in mammals, also existed in protostomes. This was a surprise, as it was often believed that protostomes and deuterostomes had their own taxon-specific neuropeptides. The first two deuterostome neuropeptide orthologues to be discovered in protostomes were insect sulfakinin, which was the orthologue of the mammalian neuroendocrine peptides gastrin and cholecystokinin [8], and conus snail vasopressin (called conopressin),
which was the orthologue of mammalian vasopressin and oxytocin [9]. The discovery of orthologous neuropeptides in deutero- and protostomes was later followed by the identifications of orthologous neuropeptide GPCRs in both animal taxa [10–14], indicating that these neuropeptide/receptor couples were already present in the common ancestor of deutero- and protostomes, which lived 630–760 million years ago [14–17].

After the initial finding of a vasopressin-like peptide in gastropod molluscs [9], vasopressin-like peptides were also isolated from various cephalopod molluscs, annelids, and other lophotrochozoans, followed by the cloning of their receptors [18–21].

Insects are the largest animal taxon with over 1.3 million identified species, while some researchers estimate that a more realistic species number might be more than 6 million [22]. The occurrence of a vasopressin-like peptide in insects has initially been somewhat uncertain. In 1987 Proux and co-workers published an oxytocin/vasopressin peptide in insects (now called inotocin) and its receptor, which was later followed by the identifications of orthologous neuropeptide GPCRs in the evolutionary lineages, leading to these insects [27].

The subphylum Chelicerata, to which ticks, scorpions, spiders and mites belong, is an arthropod taxon different from insects. Ticks are medically important, because they are vectors for pathogens that cause serious diseases in humans. Ticks belonging to the genus *Ixodes*, such as *I. scapularis* and *I. ricinus*, often contain the spirochete bacterium *Borrelia burgdorferi*, which they transfer to their human host during a blood meal. If not treated by antibiotics, *B. burgdorferi* will cause Lyme disease (borreliosis), which is a serious, neurological disease.

During our annotation of the sequenced genome from *I. scapularis*, we found that the tick had three presumed inotocin GPCR genes, while insects had only one or none [29]. We hypothesized that these additional inotocin GPCR genes might have something to do with a specific aspect in the life cycle of the tick and that these GPCRs, therefore, might be promising drug targets [29]. The purpose of the research project described in the current paper was to clone the three presumed tick inotocin GPCRs, express them in cells in cell culture and functionally characterize them to establish that they are genuine tick inotocin receptors.

### 2. Materials and methods

No experiments with live animals have been carried out and our investigations comply with the ARRIVE guidelines for animal experiments. RNA purification and cDNA synthesis were done as described previously [30]. The coding sequences of the tick inotocin receptors were amplified by PCR using HotStarTaq Plus DNA Polymerase (Qiagen). The inotocin preprohormone was amplified by 3′-RACE using the First Choice RLM-RACE kit (Thermo Fisher Scientific) in the presence of 8% DMSO. All PCR primers are given in Supplementary file 1. The PCR...
Fig. 1. Alignment of the three tick inotocin receptors Isca-ITR1 (*I. scapularis* inotocin receptor1), Isca-ITR2, Isca-ITR3, the beetle inotocin receptor Tcas-ITR, and the wasp inotocin receptor Nvit-ITR. Amino acid residue positions common to at least three aligned receptors are highlighted in yellow. The position of a common intron shared by the five receptor genes is marked by a blue star. For GenBank accession numbers see Supplementary file 2. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)
products were ligated into the vector pCR4-TOPO using the TOPO TA Cloning kit (Thermo Fisher Scientific), sequenced by Eurofins Genomics and analyzed with the CLC Main Workbench software package. The coding sequences of the receptors were inserted into the expression vector pIRE62-EGFP (Takara Bio Europe), using the EcoRI restriction sites and the Rapid DNA Dephos and Ligation kit (Merck).

Transfection of Chinese Hamster Ovary (CHO) cells was done as in Ref. [30]. The bioluminescence assay was performed on a Wallac Victor® multilabel counter (PerkinElmer Life Sciences) [31]. Tick inotocin was from Genemed Synthesis.

MUSCLE and MEGA 11 software were used for protein alignments and phylogenetic tree analyses. Neighbor-Joining trees were bootstrapped using 1000 replicates. Accession numbers are given in Supplementary file 2. SignalP 6.0 software [32] was used to predict the signal peptide. Deep TMHMM software [33] was used to predict transmembrane helices. Dose-response curves and EC50 values were calculated using Prism (GraphPad Software).

3. Results and discussion

3.1. Cloning of the three tick inotocin receptors

We searched the sequenced genome from I. scapularis [29] with a query corresponding to the beetle inotocin GPCR, which resulted in the identification of three presumed tick inotocin receptors (tick ITR1, ITR2, ITR3). We cloned these GPCRs, using standard molecular cloning techniques and PCR primers based on the genomic sequences of the three receptors (Fig. 1, Supplementary file 1). The protein sequences of both ITR1 and ITR2 were 471 amino acid residues long, while the length of ITR3 was 465 residues (Fig. 1). Each receptor had seven transmembrane helices and the canonical DRY sequence at the border of TM3 and the second intracellular loop (Fig. 1).

3.2. Phylogenetic tree analyses of arthropod oxytocin/vasopressin-like receptors

Fig. 2 (left side, with yellow background) shows a phylogenetic tree of various oxytocin/vasopressin-like receptors and the canonical DRY sequence at the border of TM3 and the second intracellular loop (Fig. 1).

3.2. Phylogenetic tree analyses of arthropod oxytocin/vasopressin-like receptors

Fig. 2 (left side, with yellow background) shows a phylogenetic tree analysis of the newly cloned I. scapularis inotocin receptors (Isca-ITR1, -ITR2, -ITR3), the previously cloned T. castaneum inotocin receptor (Tcas-ITR) [27] and the human oxytocin (Hsap-OTR) and vasopressin receptors (Hsap-V1aR, -V1bR, -V2R) [10–13]. Also included in this analysis are annotated oxytocin/vasopressin-like receptors from tardi-grades, insects, and chelicerates with a focus on the last group. The phylogenetic tree shows that we could annotate several new inotocin receptors from chelicerates. These findings, however, do not mean that all chelicerate species have inotocin receptors. For example, the spiders Argoipe bruennichi (abbreviated: Abru), Nephila pilipes (Npil), and Stegodyphus dumicola (Sdum) (presented at the right side of Fig. 2, because they have CCAP receptors) lack inotocin receptors (also noticed in Ref. [34]). Other chelicerates have high numbers of inotocin receptors, such as the soil mite Oppia nitenis (Onit), which has five of these receptors (Fig. 2). The tree clearly shows that the arthropod inotocin receptors and human oxytocin/vasopressin receptors are evolutionarily related.
3.3. Evolutionary relationships of the tick inotocin receptors with other neuropeptide receptors

The three tick inotocin receptor genes are lying on one genomic contig (NW_024609837), which has a size of 92 Mb (Fig. 3A). ITR2 and ITR3 are separated by only 0.7 Mb, which suggests that they originated by recent gene duplications. Also located on the same contig is one crustacean cardioactive peptide (CCAP) receptor gene and four adipokinetic hormone/corazonin-related peptide (ACP) receptor genes (ACPR2, ACPR3, ACPR4, and ACPR5) (Fig. 3A). A second contig contains a fifth ACP receptor gene (ACPR1) and a third contig contains a corazonin receptor gene (CRZR) (Fig. 3A).

The presence of three inotocin receptor genes, four ACP receptor genes, and one CCAP receptor gene on one contig (NW_024609837), which has a size of 92 Mb (Fig. 3A), suggests that these genes might be evolutionarily related. This conclusion is confirmed, when we look at other arthropods, for example at the malaria mosquito Anopheles gambiae, where one genomic contig (NT_078266) houses one adipokinetic hormone receptor (AKHR) gene, one CCAP receptor gene, AKH, and ACP receptor proteins, showing that the phylogenetic tree has two major branches of related receptors: One branch, containing the inotocin and CCAP receptors, and one branch, containing the corazonin, AKH, and ACP receptors. Abbreviations: ACPR1, ACPR2, ACPR3, ACPR4, ACPR5, adipokinetic hormone/corazonin-related peptide receptor1, -2, -3, -4, -5; AKHR, adipokinetic hormone receptor; CCAPR, crustacean cardioactive peptide receptor; CRZR, corazonin receptor; Dmel-ASTAR1, Drosophila melanogaster allatostatin-A receptor 1; ITR1, ITR2, ITR3, inotocin receptor –1, –2, –3. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

The evolutionary relationships between the inotocin, CCAP, AKH, corazonin, and ACP receptor genes are also confirmed, when we compare their gene products in a phylogenetic tree analysis. Fig. 3C shows such an analysis, where we used receptor protein sequences from both I. scapularis and A. gambiae. It is interesting that this tree shows two major branches: One branch, containing the inotocin and CCAP receptor proteins; and one branch containing corazonin, AKH, and ACP receptor proteins. Thus, inotocin receptors appear to be more related to CCAP receptors than to corazonin, AKH, and ACP receptor proteins.

3.4. Intron/exon organizations of the inotocin receptor genes

All three tick inotocin receptor genes share one common intron that is located at a position corresponding to an amino acid residue in the receptor proteins between transmembrane helix 6 (TM6) and TM7 (marked by a blue star in Fig. 1).

An alignment of the intron/exon organizations of the three tick inotocin receptors with those from the human oxytocin receptor and three vasopressin receptor genes shows a complete conservation of their gene structures (Supplementary file 3). Other chelicerates, myriapods, and insects have either identical or similar gene structures, but always one common intron with the human oxytocin and vasopressin receptor.
genes (Supplementary file 3). These results support the close evolutionary relationships between human and arthropod oxytocin/vasopressin receptor genes.

3.5. Cloning of the tick inotocin preprohormone

We have previously annotated the inotocin preprohormone gene from *I. scapularis* [29]. In the current paper, we have cloned and further analyzed this preprohormone. Fig. 4A and B show that the preprohormone gene consists of three exons (exon 1, 2, and 3). Exon 1 codes for the signal peptide and the immature inotocin sequence. Exon 2 codes for the neurophysin region of the preprohormone. Exon 3 codes for the C-terminal part of the protein.

The tick preprohormone has the same structural features as the mammalian oxytocin and vasopressin preprohormones: The immature tick inotocin sequence is located immediately after the signal sequence and is followed by a neurophysin part that contains fourteen cysteine residues, which form seven cystine bridges (Fig. 4C). After the signal peptide is removed by signal peptidase during RER membrane transport (left arrow), immature tick inotocin is liberated from its prohormone by cleavage at the canonical KR prohormone convertase (PC1/3) cleavage signal (middle arrow), followed by the conversion of the C-terminal G residue into a C-terminal amide group by peptidylglycine α-hydroxylating monooxygenase [37]. Mature tick inotocin has the sequence CFIITNCPPGamide and differs from insect inotocin (CLITNCPRGamide) by two amino acid residues. Mature tick inotocin is cyclic after oxidation of its cysteine residues and formation of a cystine bridge.

The neurophysin part of the tick inotocin preprohormone has two domains, one containing four cystine bridges and one containing three bridges (Fig. 4C). In between the two domains is a second canonical prohormone convertase cleavage site (PC1/3) cleavage signal (middle arrow), followed by the conversion of the C-terminal G residue into a C-terminal amide group by peptidylglycine α-hydroxylating monooxygenase [37]. Mature tick inotocin has the sequence CFIITNCPPGamide and differs from insect inotocin (CLITNCPRGamide) by two amino acid residues. Mature tick inotocin is cyclic after oxidation of its cysteine residues and formation of a cystine bridge.

The neurophysin part of the tick inotocin preprohormone has two domains, one containing four cystine bridges and one containing three bridges (Fig. 4C). In between the two domains is a second canonical prohormone convertase cleavage site (PC1/3) cleavage signal (middle arrow), followed by the conversion of the C-terminal G residue into a C-terminal amide group by peptidylglycine α-hydroxylating monooxygenase [37]. Mature tick inotocin has the sequence CFIITNCPPGamide and differs from insect inotocin (CLITNCPRGamide) by two amino acid residues. Mature tick inotocin is cyclic after oxidation of its cysteine residues and formation of a cystine bridge.

The neurophysin part of the tick inotocin preprohormone has two domains, one containing four cystine bridges and one containing three bridges (Fig. 4C). In between the two domains is a second canonical prohormone convertase cleavage site (PC1/3) cleavage signal (middle arrow), followed by the conversion of the C-terminal G residue into a C-terminal amide group by peptidylglycine α-hydroxylating monooxygenase [37]. Mature tick inotocin has the sequence CFIITNCPPGamide and differs from insect inotocin (CLITNCPRGamide) by two amino acid residues. Mature tick inotocin is cyclic after oxidation of its cysteine residues and formation of a cystine bridge.

The neurophysin part of the tick inotocin preprohormone has two domains, one containing four cystine bridges and one containing three bridges (Fig. 4C). In between the two domains is a second canonical prohormone convertase cleavage site (PC1/3) cleavage signal (middle arrow), followed by the conversion of the C-terminal G residue into a C-terminal amide group by peptidylglycine α-hydroxylating monooxygenase [37]. Mature tick inotocin has the sequence CFIITNCPPGamide and differs from insect inotocin (CLITNCPRGamide) by two amino acid residues. Mature tick inotocin is cyclic after oxidation of its cysteine residues and formation of a cystine bridge.

The neurophysin part of the tick inotocin preprohormone has two domains, one containing four cystine bridges and one containing three bridges (Fig. 4C). In between the two domains is a second canonical prohormone convertase cleavage site (PC1/3) cleavage signal (middle arrow), followed by the conversion of the C-terminal G residue into a C-terminal amide group by peptidylglycine α-hydroxylating monooxygenase [37]. Mature tick inotocin has the sequence CFIITNCPPGamide and differs from insect inotocin (CLITNCPRGamide) by two amino acid residues. Mature tick inotocin is cyclic after oxidation of its cysteine residues and formation of a cystine bridge.

The neurophysin part of the tick inotocin preprohormone has two domains, one containing four cystine bridges and one containing three bridges (Fig. 4C). In between the two domains is a second canonical prohormone convertase cleavage site (PC1/3) cleavage signal (middle arrow), followed by the conversion of the C-terminal G residue into a C-terminal amide group by peptidylglycine α-hydroxylating monooxygenase [37]. Mature tick inotocin has the sequence CFIITNCPPGamide and differs from insect inotocin (CLITNCPRGamide) by two amino acid residues. Mature tick inotocin is cyclic after oxidation of its cysteine residues and formation of a cystine bridge.

The neurophysin part of the tick inotocin preprohormone has two domains, one containing four cystine bridges and one containing three bridges (Fig. 4C). In between the two domains is a second canonical prohormone convertase cleavage site (PC1/3) cleavage signal (middle arrow), followed by the conversion of the C-terminal G residue into a C-terminal amide group by peptidylglycine α-hydroxylating monooxygenase [37]. Mature tick inotocin has the sequence CFIITNCPPGamide and differs from insect inotocin (CLITNCPRGamide) by two amino acid residues. Mature tick inotocin is cyclic after oxidation of its cysteine residues and formation of a cystine bridge.

The neurophysin part of the tick inotocin preprohormone has two domains, one containing four cystine bridges and one containing three bridges (Fig. 4C). In between the two domains is a second canonical prohormone convertase cleavage site (PC1/3) cleavage signal (middle arrow), followed by the conversion of the C-terminal G residue into a C-terminal amide group by peptidylglycine α-hydroxylating monooxygenase [37]. Mature tick inotocin has the sequence CFIITNCPPGamide and differs from insect inotocin (CLITNCPRGamide) by two amino acid residues. Mature tick inotocin is cyclic after oxidation of its cysteine residues and formation of a cystine bridge.
(the one with four cystine bridges) will be amidated [37]. This would make the two cystine-knot peptides potential signal molecules for intercellular communication. The neuropsin cleavage site occurs in I. scapularis and fifteen other hard tick species, but not in other chelicerates (Fig. 4D, Supplementary file 4).

Normally, chelicerates have only one gene for an inotocin preprohormone. However, we found that the genome from the citrus mite and pest, Panonychus citri, contains as many as nine genes for this preprohormone, of which five genes might possibly be pseudogenes (Supplementary file 5). The genome from the cattle tick, Haemaphysalis longicornis, contains two genes for the inotocin preprohormone (Supplementary file 5).

3.6. Deorphanization of the three tick inotocin receptors

We transfected Chinese Hamster Ovary (CHO) cells with cDNA, coding for each of the three presumed tick inotocin receptors and selected cell clones that stably expressed one of the receptors. These cells also stably expressed the promiscuous G protein G\( \alpha \)G\( \beta \)(which couples the receptors to the \( \text{IP}_3/Ca^{2+} \) second messenger pathway) and transiently expressed the cnidarian protein apoaequorin. Shortly before the assay, the cnidarian compound coelenterazine was added to the cells. Binding of the ligand to the GPCRs expressed in these pre-treated cells initiates a second messenger cascade, leading to an intracellular Ca\(^{2+} \) pulse and a Ca\(^{2+} \)-mediated bioluminescence response, which can be quantified on a plate reader [27,31].

Fig. 4E shows that tick inotocin induced a strong bioluminescence response in CHO cells stably expressing the tick ITR1 receptor (EC\(_{50} = 1.6 \times 10^{-8} \) M). Also CHO cells stably expressing ITR2 (EC\(_{50} = 5.8 \times 10^{-9} \) M) and CHO cells stably expressing ITR3 (EC\(_{50} = 9.2 \times 10^{-9} \) M) showed strong bioluminescence responses after addition of nanomolar concentrations of tick inotocin (Fig. 4F and G). These results deorphanized the three receptors.

CRediT author statement

FH and CJPG: conceptualization, supervision, resources; FH, TAR, MS, and GT: investigation, formal analysis, validation; FH: visualization; CJPG: writing- Original Draft; CJPG and FH: writing- Review & Editing, project administration.

Funding

This work was supported by Carlsberg Foundation (grant number CF18-0101 to CJPG). This funding source had no role in the study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Cornelis Grimmelikhuijzen reports financial support was provided by Carlsberg Foundation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbrc.2023.09.009,

References

[29] C. deorphanized the three receptors.


