General and Comparative Endocrinology 175 (2012) 135–143

Contents lists available at SciVerse ScienceDirect

General and Comparative Endocrinology

journal homepage: www.elsevier.com/locate/ygcen

The oxytocin/vasopressin receptor family has at least five members in the gnathostome lineage, inclucing two distinct V2 subtypes

Daniel Ocampo Daza *, Michalina Lewicka [1], Dan Larhammar

Department of Neuroscience, Science for Life Laboratory, Uppsala Universitet, Box 593, SE-751 24 Uppsala, Sweden

ARTICLE INFO

Article history:
Received 5 August 2011
Revised 30 September 2011
Accepted 20 October 2011
Available online 28 October 2011

Keywords:
Vasopressin receptor
Vasotocin receptor
Oxytocin receptor
Mesotocin receptor
Isotocin receptor
Evolution
GPCR

ABSTRACT

The vertebrate oxytocin and vasopressin receptors form a family of G-protein-coupled receptors (GPCRs) that mediate a large variety of functions, including social behavior and the regulation of blood pressure, water balance and reproduction. In mammals four family members have been identified, three of which respond to vasopressin (VP) named V1A, V1B and V2, and one of which is activated by oxytocin (OT), called the OT receptor. Four receptors have been identified in chicken as well, but these have received different names. Until recently only V1-type receptors have been described in several species of teleost fishes. We have identified family members in several gnathostome genomes and performed phylogenetic analyses to classify OT/VP-receptors across species and determine orthology relationships. Our phylogenetic tree identifies five distinct ancestral gnathostome receptor subtypes in the OT/VP receptor family: V1A, V1B, V2A, V2B and OT receptors. The existence of distinct V2A and V2B receptors has not been previously recognized. We have found these two subtypes in all examined teleost genomes as well as in available frog and lizard genomes and conclude that the V2A-type is orthologous to mammalian V2 receptors whereas the V2B-type is orthologous to avian V2 receptors. Some teleost fishes have acquired additional and more recent gene duplicates with up to eight receptor family members. Thus, this analysis reveals an unprecedented complexity in the gnathostome repertoire of OT/VP receptors, opening interesting research avenues regarding functions such as regulation of water balance, reproduction and behavior, particularly in reptiles, amphibians, teleost fishes and cartilaginous fishes.

1. Introduction

The classical neuroendocrine nonapeptides vasopressin and oxytocin, released from hypothalamic neurons that extend their axons into the posterior pituitary, have profound effects on salt and water balance and reproduction, respectively [6,9,32]. In addition, their widespread distribution in the brain in mammals correlates with important roles as neuropeptides influencing social behavior [11,20] including social olfactory memory in mice [12] and trust and generosity in humans [28,53]. In fish the two peptides have been found to influence aggressive behavior [4,40,49]. During the past few years, genetic variants of the peptide receptors have been found to correlate with variation in different aspects of human social interaction, see for instance [51].

Oxytocin and vasopressin were two of the first peptides that were sequenced in different classes of vertebrates. Unfortunately, each new sequence variant was given a separate name such as isotocin and mesotocin [2] for peptides that we now know are orthologs of mammalian oxytocin. It has also become clear that vasotocin in non-mammalian vertebrates is the ortholog of mammalian vasopressin. In fact, oxytocin and vasopressin arose from a common ancestral gene by a local duplication in a gnathostome ancestor, and the separate oxytocin and vasopressin lineages arose in early vertebrate evolution [16,17]. This makes other names for these peptides in the gnathostomes redundant in an evolutionary perspective, and also establishes a likely evolutionary time frame for the divergence of oxytocin and vasopressin systems, including the corresponding receptors.

The peptides exert their actions by activating a family of G-protein-coupled receptors (GPCR) belonging to the rhodopsin clan. In mammals, four receptors were cloned in the 1990s: one oxytocin receptor called OT with the gene name *OXTR* [24], and three vasopressin receptors named V_{1A} (*AVPR1A*), V_{1B} (*AVPR1B*) and V_2 (*AVPR2*) [7,44,46]. The OT receptor (we will use the abbreviation OTR to distinguish the receptor from the OT peptide) and the two V_1 receptors primarily stimulate Gq and the DAG/IP3/

Abbreviations: ICL3, third intracellular loop; NJ, neighbor joining; NNI, nearest neighbor interchange; PhyML, phylogenetic maximum likelihood; SPR, subtree pruning and regrafting; TMHMM, transmembrane helix hidden Markov model.

* Corresponding author. Fax: +46 018 511540.
E-mail address: Daniel.Ocampo-Daza@neuro.uu.se (D. Ocampo Daza).
[1] Current address: Department of Neuroscience, Karolinska Institutet, SE-171 77 Stockholm, Sweden.

0016-6480/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygcen.2011.10.011

Ca^{2+} pathway whereas the V_2 receptor inhibits adenylyl cyclase, thereby reducing the production of cAMP. The four receptors have approx. 36–46% overall identity to each other in human, and 42–57% in the region spanning transmembrane helix (TM) 1 to TM 7.

Over the past years, OT/VP receptors have been cloned and characterized from several species other than mammals, such as chicken, amphibians and ray-finned fishes. However, their orthology or paralogy relationships with the mammalian receptors and each other have not always been clear, although this issue has been addressed before. For example: the chicken vasotocin (vasopressin) receptor initially named VT1 [45] is clearly more similar to mammalian V2, although phylogenetic analyses have placed it outside of the mammalian V2 clade [10,14], and the chicken receptor called VT2 [10] is in fact a homolog of the mammalian V1B, a receptor that has also been known as V_3. The cloned chicken receptors are reviewed by Baeyens and Cornett [5]. In amphibians, vasopressin (vasotocin) receptors showing similarities to mammalian V1A, V1B and V2 as well as oxytocin (mesotocin) receptors have been cloned and characterized in several species [42]. In ray-finned fishes, all described receptors had been identified as V1A receptors and oxytocin (isotocin) receptors until recently when V2-type receptors from gray bichir and medaka were cloned [27] and a V2 receptor cDNA from the Amargosa pupfish was identified [33]. Together, these findings raise questions about the evolution of the OT/VP receptor family before and during the divergence of ray-fined fishes and lobe-finned fishes (including tetrapods), and before the origin of gnathostomes.

We initially assumed that the four mammalian receptor subtype genes might reflect the double tetraploidization early in vertebrate evolution (2R) [34,37]. In an effort to understand the evolutionary history of this receptor family, we have collected and characterized sequences from several species representing the major gnathostome classes. It emerged that the ancestral gnathostome actually must have had at least five receptor genes and that differential losses have occurred in different evolutionary lineages. This is proven by the fact that a few lineages have retained all five of the ancestral receptor family members. Our phylogenetic analyses that clarify important aspects of the orthology–paralogy relationships also lead to a simplified terminology for the OT/VP receptor family.

2. Methods

2.1. Database searches and sequence annotation

Amino acid sequences of OT/VP receptor family members were identified using the protein family prediction feature of the Ensembl genome browser available at <www.ensembl.org> [13]. The chromosome locations of the identified sequences were noted. The chromosomal locations as well as more information about the database mining, such as genome assembly versions, are available in Supplementary material 1. To account for failures in the automated protein family predictions, sequences were also sought through TBLASTN searches [3] in the Ensembl database as well as in the National Center for Biotechnology Information (NCBI) Reference Sequence and trace archive databases (<http://www.ncbi.nlm.nih.gov>) using the known human V1A, V1B, V2 and OT receptor sequences as queries. Hits were considered significant if their expect values (*E*) were lower than e^{-30}.

All sequences in the identified Ensembl protein family were collected from the following species (names in parenthesis indicate how the different species will be referred to henceforth): *Homo sapiens* (human), *Mus musculus* (mouse), *Monodelphis domestica* (opossum), *Ornithorhynchus anatinus* (platypus), *Anolis carolinensis* (anole lizard), *Silurana (Xenopus) tropicalis* (frog), *Gallus gallus* (chicken), *Danio rerio* (zebrafish), *Oryzias latipes* (medaka), *Gasterosteus aculeatus* (stickleback) and *Takifugu rubripes* (fugu). The functionally characterized *Ciona instestinalis* (tunicate) VP receptor sequence [22] was collected and used as query in TBLASTN searches in the *Ciona intestinalis* genome assembly available in the Ensembl database in order to obtain its chromosomal location and reveal possibly unidentified family members. The functionally characterized *Octopus vulgaris* (common octopus) sequences for OPR, CTR1 and CTR2 [21] were collected to be used as an outgroup. Family members were also sought in the *Callorhinchus milii* (elephant shark) genome database (<http://esharkgenome.imcb.a-star.edu.sg>). Genomic scaffold sequences representing the top hits (expect values < e^{-30}) were collected in order to predict the receptor sequences.

All identified sequences were verified against published sequences entered into the NCBI Reference Sequence database. Erroneous predictions in the Ensembl database were replaced with better reference database sequences where available or annotated manually by following consensus for gene initiation and splice donor and acceptor sites as well as sequence similarities to verified family members. Short predictions were extended by searching for missing exons in the flanking genomic sequence. All new protein predictions were controlled for the characteristic seven transmembrane helices using the TMHMM prediction tool available at <www.cbs.dtu.dk/services/TMHMM/> [29]. Detailed accounts of the editing process can be provided upon request.

2.2. Sequence alignments and phylogenetic analyses

All amino acid sequence alignments were made using the ClustalW [48] function on Jalview version 2.4.0 [52] and edited in Jalview or in the text editor TextWrangler version 3.1 (<http://www.barebones.com/products/TextWrangler>). An initial alignment was made using the identified sequences in order to create a phylogenetic tree with the neighbor joining (NJ) method [39] (1000 bootstrap replications and standard settings in ClustalX version 2.0 [31]). The main clusters of the OT/VP receptor tree were identified in the NJ-tree topology (Supplementary material 4), after which amino acid sequence alignments for each individual cluster were made and edited manually. Poorly aligned sequence stretches were removed from the amino terminal, intracellular loop 3 (ICL3) and carboxy terminal. Alignment details are provided in Supplementary material 3. The tunicate and octopus protein sequences were edited to remove corresponding sequence stretches. These main cluster alignments were then aligned to one another and to the tunicate and octopus protein sequences using ClustalX's profile alignment function. The final alignment was edited manually to adjust sequence stretches before and after the removed ICL3 sequence and to adjust and align conserved exon boundaries. A phylogenetic maximum likelihood (PhyML) tree was created from this alignment using the online execution of the PhyML 3.0 algorithm available at <http://www.atgc-montpellier.fr/phyml/> [15]. The amino acid frequency (equilibrium frequency), proportion of invariable sites and gamma-shape for the amino acid substitution rate heterogeneity parameters were estimated from the dataset. The number of substitution rate categories was set to 8. The starting tree was estimated using BIONJ and both the NNI and SPR tree improvement methods were used to estimate the best tree topology. Both the tree topology and branch length optimization options were selected. A non-parametric bootstrap analysis with 100 replicates was selected for statistical branch support. The best amino acid substitution model was estimated from the final alignment using ProtTest version 1.4 [1]. In ProtTest, models were tested with no add-ons and assuming 8 gamma rate categories, the optimization strategy was set to slow and the BIONJ strategy was selected to create a random input tree. As a result of this analysis, the JTT-model of amino acid substitution was assumed for the construction of the PhyML tree.