Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Genomics 87 (2006) 366 – 381
GENOMICS
www.elsevier.com/locate/ygeno

Comparative genomic and expression analysis of the conserved NTPDase gene family in *Xenopus* ☆

Karine Massé, Robert Eason, Surinder Bhamra, Nicholas Dale [1], Elizabeth A. Jones *,[1]

Department of Biological Sciences, Warwick University, Coventry, CV47AL, UK

Received 8 July 2005; accepted 9 November 2005
Available online 27 December 2005

Abstract

The purines, ATP and adenosine, are important signaling molecules in the nervous system. ATP is sequentially degraded to adenosine by the ectonucleotidase proteins. The NTPDase (or CD39) family is a subfamily of these enzymes, which consists of nine members in mammals. In *Xenopus* embryos, we have shown that ATP, and its antagonist adenosine, regulate the rundown of swimming and we therefore proposed that ectonucleotidase proteins are key regulators of locomotor activity. Here, we report the cloning of all nine members of the NTPDase family in *Xenopus laevis* and *Xenopus tropicalis*. Our phylogenetic analysis shows that this family is highly conserved between the frog species and also during vertebrate evolution. In the adult frog, NTPDase genes are broadly expressed. During development, all NTPDase genes, except for NTPDase8, are expressed and display a distinct specific expression pattern, suggesting potentially different functions of these proteins during embryogenesis of *X. laevis*.

Keywords: Ectonucleotidases; E-NTPDase; CD39; Purine metabolism; Purinergic; *Xenopus laevis*; Embryogenesis; Gene expression

Introduction

Nucleotides are involved in many cellular and extracellular processes such as cardiac function, muscle contraction and relaxation, secretion of hormones, and cell growth (for review see [1]). In the peripheral and central nervous systems, ATP and adenosine are important neurotransmitter and modulator molecules that have been implicated in the control of sleep [2], regulation of pain pathways [3,4], and mediation of neuroprotection in the brain during transient ischemia [5]. In many cases, adenosine is not directly released by neurons but is produced from the breakdown of ATP by the ectonucleotidase proteins [6]. These proteins therefore control both the termination of ATP signaling and the accumulation of adenosine. As both ATP and adenosine can act at several distinct types of receptors, the ectonucleotidase proteins are a key site of regulation for purinergic signaling.

In *Xenopus* embryos, these purines are important regulators of spinal motor pattern generation. We have shown that ATP and adenosine regulate the fatigue or run-down of swimming in tadpoles. ATP, which is released from spinal neurons [7], inhibits K^+ currents and therefore increases excitability of the locomotor network, whereas adenosine has the opposite effect by inhibiting Ca^{2+} currents [8]. The production of adenosine can be measured in spinal cord during locomotion [9,10]. Moreover, we have proposed that the feed-forward inhibition by ADP of the 5′-ectonucleotidase is a key determinant that delays the generation of adenosine and thus regulates the rate of run-down in *Xenopus* embryos [8,11]. By terminating the actions of ATP and by initiating the effects of adenosine, members of the ectonucleotidase families are therefore the major players in the control of locomotor activity in frog embryos.

Nucleoside triphosphate diphosphohydrolases are members of the ectonucleotidase superfamily. They are membrane-bound, highly glycosylated proteins. Previously known as E-type ATPases, ATPDases, ecto-ATPases, or ecto-apyrases, the members of this family are now named NTPDases (or E-NTPDases)

☆ Sequence data from this article have been deposited with the GenBank/EMBL Data Libraries under Accession Nos. DQ117604-8 and DQ118409-13.

* Corresponding author. Fax: +44 2476523701.
E-mail address: Elizabeth.Jones@warwick.ac.uk (E.A. Jones).
[1] Joint PIs.

0888-7543/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2005.11.003

after the nomenclature of Zimmermann [6]. These proteins can all hydrolyze nucleoside 5′-triphosphates and nucleoside 5′-diphosphates. Each member has specific substrate specificity and subcellular localization. *NTPDase* genes have been cloned in vertebrates, and homologues have also been identified in invertebrates, plants, yeast, and protozoans. The family members are characterized by the presence of five highly conserved domains [apyrase conserved regions (ACR)], by their activity dependence on divalent cations (Ca^{2+} and/or Mg^{2+}), and their insensitively to inhibitors of P-type, F-type, and V-type ATPases.

In mammals, nine *NTPDases* have been identified thus far. In adults, these genes display a broad tissue distribution with overlapping expression domains. NTPDases 1–3 and NTPDase8 are bound to the plasma membrane by two transmembrane domains and their catalytic site is in the extracellular milieu, classifying these proteins as ecto-NTPDases. The other members (NTPDase4, NTPDases 5–7) are endo-NTPDase proteins, bound either to the ER/Golgi or to lysosomal membranes. However, NTPDases 5 and 6 have also been found on the plasma membrane and can be secreted. *NTPDase1* was the first member cloned and was identified as *CD39* [12]. *NTPDases 2* and *3* [13,14] and the recently cloned *NTPDase8* [15] are highly homologous to *NTPDase1* but their substrate specificity differs. NTPDase2 has a strong preference for triphosphate nucleotides (ATP:ADP rate of hydrolysis of ~30:1) [16], whereas NTPDases 1, 3, and 8 catalyze the hydrolysis of triphospho- and diphosphonucleosides efficiently with ATP:ADP rates of hydrolysis of ~1:1, ~3:1, and ~2:1, respectively [15–17]. Two splice variants of NTPDase4 have been isolated, LALP70 and UDPase (LALP70v), and are now referred as NTPDase4 and NTPDase4a [18–20]. They are localized to the lysosomal/autophagic vacuole membranes and to the Golgi, where they catalyze the hydrolysis of UTP and UDP, respectively. LALP1 is now referred as NTPDase7 even though its catalytic characteristic, protein sequence, and cellular sublocalization suggest that this protein is more related to the NTPDase4 proteins [21]. The last two proteins, NTPDases 5 and 6, are the only family members with one transmembrane domain and they hydrolyze preferentially diphosphonucleosides [22,23].

The role of the NTPDases is to regulate the concentration of extracellular and intracellular nucleotides and therefore they are involved in many functions mediated by NTPs. However, specific functions have been ascribed to some of these proteins. The knock-out mouse for *NTPDase1* identified a role for this protein in hemostasis, thromboregulation, and angiogenesis [24,25] and also in the modulation of inflammation and immune responses mediated by extracellular nucleotides [26]. These studies are the first to identify the functional role of a member of this family in vivo in a vertebrate. However, in vitro, NTPDase2 has been shown to promote platelet aggregation, an opposite role to the NTPDase1 function [27]. Finally, the identification of *NTPDase5* as the PCPH proto-oncogene links the NTPDase family to carcinogenesis [28]. The functions of these nine NTPDases are likely to be numerous and, therefore, still need to be elucidated in vivo.

X. laevis provides a very powerful model organism for the establishment of gene function. Misexpression of genes can be achieved by microinjection of the mRNA of interest and knock-down models can now be obtained by using morpholinos ([29] or for review see [30]). This, coupled with the availability of physiological methodologies and computational models [11] by which to test the roles of these enzyme families in physiological processes, makes *Xenopus* an ideal organism in which to study this important gene family.

Here, we report the cloning and characterization of nine members of the *NTPDase* family in *X. laevis*, the first and necessary step before analyzing their functions *in vivo*. We also present and compare their temporal and spatial expression patterns in frog embryos and then their distribution in adult tissues. We have also identified the members of this family in *Xenopus tropicalis* and conducted a phylogenetic analysis of these genes in both of these two *Xenopus* species. This work is the first study that describes the complete embryonic expression pattern of all the members of this family during the development of a vertebrate model organism.

Results

Cloning of Xenopus NTPDase genes

To clone the *NTPDase* family of genes in *Xenopus laevis* and *X. tropicalis*, we opted to search the NCBI Internet site and JGI Internet site to identify EST clones. Where no ESTs were available, we searched the *X. tropicalis* genomic database on the JGI website. A combination of approaches was used, detailed below for each gene. All the positive EST clones (Table 1) were ordered from the UK MRC HGMP Resource Centre, Cambridge, and checked by sequencing. Further rounds of sequencing were carried out to obtain the full-length sequences.

Isolation and sequence characterization of Xenopus NTPDase1 gene

Interrogation of the Unigene database on the NCBI internet site identified the clone CF282534 (now BC073267), as being 2933 bp long. The open reading frame encodes a protein of 508 aa, with a calculated molecular weight of 57,012 kDa. The ATG at nucleotide position 24 was identified by homology with the human and mouse *NTPDase1* sequences, since the sequence of the translation initiation GAACTGAUGG showed a weak homology to the consensus Kozak sequence [31]. The deduced amino acid sequence of *X. laevis* NTDase1 (Accession No. AAH73267) is shown in Fig. 1. The five apyrase conserved domains, characteristic of the NTPDase family, are present. Prediction of transmembrane helices using the TMBase software allowed the identification of two transmembrane domains, one at the N-terminal region (aa positions 17 to 38) and the other at the C-terminal region of the protein (aa positions 477 to 495).

X. tropicalis NTPDase1 protein sequence is available on the NCBI Internet site under Accession No. AAH76662. This protein is 508 aa long with a calculated molecular weight of