Gene 568 (2015) 155–164

Contents lists available at ScienceDirect

Gene

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

Research paper

Multiple non-coding exons and alternative splicing in the mouse *Mas* protooncogene

Natalia Alenina [a,b], Ilka Böhme [c], Michael Bader [a,b,d], Thomas Walther [c,e,*]

[a] Max-Delbrück-Center for Molecular Medicine (MDC), Robert-Rössle-Straße 10, 13092 Berlin-Buch, Germany
[b] Federal University of Minas Gerais (UFMG), ICB, 6627 Belo Horizonte, MG, Brasil
[c] Centre for Perinatal Medicine, University Medical Centre Leipzig, Liebigstraße 20a, 04103 Leipzig, Germany
[d] Charité University Medicine Berlin, Charitéplatz 1, 10117 Berlin, Germany
[e] Department of Pharmacology and Therapeutics, 2nd Floor, Western Road, University College Cork, Cork, Ireland

ARTICLE INFO

Article history:
Received 9 March 2015
Received in revised form 23 April 2015
Accepted 16 May 2015
Available online 21 May 2015

Keywords:
Mas gene structure
G protein-coupled receptor
Tissue-specific promoter

ABSTRACT

The *Mas* protooncogene encodes a G protein-coupled receptor with the common seven transmembrane domains, expressed mainly in the testis and brain. We provided evidence that Mas is a functional angiotensin-(1–7) receptor and can interact with the angiotensin II type1 (AT1) receptor. The gene is transcriptionally regulated during development in the brain and testis, but its structure was unresolved. In this study we used 5′- and 3′-RACE, RT-PCR, and RNase-protection assays to elucidate the complete *Mas* gene structure and organization.
We identified 12 exons in the mouse *Mas* gene with 11 in the 5′ untranslated mRNA, which can be alternatively spliced. We also showed that *Mas* transcription can start from 4 tissue-specific promoters, whereby testis-specific *Mas* mRNA is transcribed from two upstream promoters, and the expression of *Mas* in the brain starts from two downstream promoters. Alternative splicing and multiple promoter usage result in at least 12 *Mas* transcripts in which different 5′ untranslated regions are fused to a common coding sequence. Moreover, termination of *Mas* mRNA is regulated by two different polyadenylation signals. The gene spans approximately 27 kb, and the longest detected mRNA contains 2451 bp.
Thus, our results characterize the *Mas* protooncogene as the gene with the most complex gene structure of all described members of the gene family coding for G protein-coupled receptors.

1. Introduction

The *Mas* protooncogene was originally detected by its ability to transform NIH 3T3 cells (Young et al., 1986). It was isolated from DNA of a human epidermoid carcinoma cell line using the co-transfection and tumorigenicity assay in nude mice. The sequence of the Mas protein indicates that it belongs to the class of receptors that are coupled to G-proteins and share a conserved structural motif, forming 7 membrane-spanning alpha-helices. Some years ago, a family of approximately 50 G protein-coupled receptors related to *Mas*, called *Mrg* (*Mas*-related genes), has been discovered, a subset of which is expressed in specific subpopulations of sensory neurons that detect painful stimuli (Dong et al., 2001; Lembo et al., 2002).

First functional studies suggested that the *Mas* gene codes for an angiotensin II (AngII)-sensitive receptor (Jackson et al., 1988). However, the increase of the intracellular Ca^{2+}-concentrations in Mas-transfected cells after AngII treatment could only be confirmed in cells expressing the AngII receptor AT1 endogeneously (Ambroz et al., 1991; Von Bohlen und Halbach et al., 2000). We could identify this postulated AT1/Mas interaction as a constitutive hetero-oligomeric complex (Kostenis et al., 2005).

We generated mice deficient for the *Mas* protooncogene that are characterized by a sustained long-term potentiation in hippocampal neurons, sex-specific alterations in exploratory behavior (Walther et al., 1998, 2000a) and alterations in heart rate and blood pressure variability (Walther et al., 2000b). Recent studies in these mice and cell transfection experiments gave evidence that *Mas* codes for a functional receptor for angiotensin-(1–7) (Santos et al., 2003), a biologically active angiotensin metabolite, and both are critical for endothelial functionality (Peiro et al., 2007, 2013; Xu et al., 2008).

We and others identified the *Mas* gene to be expressed predominantly in the testis and distinct areas of the forebrain like hippocampus and amygdala (Bunnemann et al., 1990; Martin et al., 1992; Alenina et al., 2002b) and less strongly, but detectable in the kidney and heart (Villar and Pedersen, 1994; Metzger et al., 1995; Alenina et al., 2002a). The *Mas*-coding region was isolated and sequenced for human (Young et al., 1986), rat (Young et al., 1988), and mouse (Metzger et al., 1995). In all

Abbreviations: Ang II, angiotensin II; AT1, angiotensin II receptor subtype 1; E, exon; I, intron; IRES, internal ribosomal entry site; Mrg, Mas-related genes; ORF, open reading frame; RACE, rapid amplification of cDNA ends; RPA, RNase-protection assay; UTR, untranslated region; Y, yeast.

* Corresponding author at: Centre for Perinatal Medicine, University Medical Centre Leipzig, Liebigstraße 20a, 04103 Leipzig, Germany.
E-mail address: t.walther@ucc.ie (T. Walther).

http://dx.doi.org/10.1016/j.gene.2015.05.043
0378-1119/

three species the entire coding sequence is localized on a single exon. The gene is located on mouse chromosome 17 (Cebra-Thomas et al., 1992), rat chromosome 1 (unpublished data), and human chromosome 6 (Rabin et al., 1987) in close proximity to the paternally imprinted *Igf2r* gene. Furthermore, maternally imprinted antisense RNA (*AS* RNA) was described that starts in the neighboring *Igf2r* gene and overlaps the *Mas* coding region (Wutz et al., 1997; Lyle et al., 2000; Latos et al., 2012). Although imprinting of *Mas* itself has been postulated by using RT-PCR (Villar and Pedersen, 1994), we and others could show that not *Mas*, but exclusively its *AS* RNA, is maternally imprinted in both embryos and adult organs (Lyle et al., 2000; Alenina et al., 2002a).

Mas has been shown to be regulated on the level of its mRNA during development of the brain (Martin et al., 1992) and testis (Metzger et al., 1995; Alenina et al., 2002b) and by neuronal activity (Martin and Hockfield, 1993). Furthermore, cardiac Mas expression in rats was reduced after AT1 receptor blockade (Ferrario et al., 2005). In order to study the DNA elements relevant for these regulatory processes, the structure of the *Mas* gene and the localization of transcriptional start sites need identification.

Until recently, the structure of the *Mas* protooncogene has only partially been characterized (Metzger et al., 1995; Schweifer et al., 1997). In this study we mapped the 5′ and 3′ ends of *Mas* mRNAs, identified several promoters and polyadenylation sites, and described the full exon/intron structure of the *Mas* gene revealing multiple alternative splice variants.

2. Material and methods

2.1. Animals

Three-month old C57BL/6 mice were used in the experiments. The animals were maintained under standardized conditions with an artificial 12-h dark-light cycle, with free access to food and water. Animals were killed by cervical dislocation; tissue samples were collected and immediately snap frozen in liquid nitrogen. This research was performed in compliance with the German law and Guide for the Care and Use of Laboratory Animals published by the OPRR (Office for Protection against Research Risks) of the US National Institutes of Health, Washington, D.C. (NIH Publication No. 85-23, revised 1985).

2.2. RNA isolation

RNA of frozen organs was isolated using TRIzol reagent (Life technologies Inc., Gaithersburg, MD, USA) following the manufacturer's protocol and stored at −70 °C.

2.3. Rapid amplification of cDNA ends

The 5′- and 3′-RACE (Rapid Amplification of cDNA Ends) experiments were conducted to identify transcriptional starts and polyadenylation sites of *Mas* cDNA, respectively. One microgram of adult mouse testis and brain total RNA was used. Experiments were carried out using FirstChoice RLM-RACE (AMS Biotechnology, Witney, United Kingdom) according to the manufacturer's instruction manual. Gene-specific primers used in 5′-RACE were: MASG1: 5′-GGA GGC ATT TCT GCT GGA GG-3′ and MASG2: 5′-GGC TTT CTC TTC AGC AAG GG-3′ for the 'testis-RACE' and MRA3: 5′-GGC AGG TCT TGG TGG GG-3′ and MRA93: 5′-GAG CGT GGA CTC CCT GGT GT-3′ for the 'brain-RACE'. For 3′-RACE the following gene-specific primers were used: MAS11: 5′-CTG GTT CCT CTG CTT CCG GAT GAG G-3′; and MAS18: 5′-TTG GTG GTG AAG ATA CGG-3′. A total of 33 cycles of the first and second amplifications were carried out using a thermal cycle program (94 °C for 30 s, 60 °C for 30 s, and 72 °C for 1 min with 1 s extension time for each next cycle) and Taq polymerase (Invitrogen, Karlsruhe, Germany). The RACE products were analyzed by 1% agarose gel electrophoresis and subcloned into the vector pGEM-T (Promega GmbH, Mannheim, Germany) for sequence analysis.

2.4. Cloning of cDNA and chromosomal clones

cDNA cloning was performed using reverse transcription (RT), followed by polymerase chain reaction (PCR). First strand cDNA was synthesized from 2 μg of the total brain or testis RNA using an M-MLV Reverse Transcriptase (Invitrogen), primed by random hexamers. In total, 33 cycles of amplification were carried out using a thermal cycle program of 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 20–60 s, followed by a final 5 min extension at 72 °C. Chromosomal clones were obtained using the Expand Long Template PCR System (Roche, Mannheim, Germany) with the following parameters: 10 cycles: 92 °C for 10 s, 59 °C for 30 s, 68 °C for 6 min; 20 cycles 92 °C for 10 s, 59 °C for 30 s, 68 °C for 1 min extended by 20 s for each cycle.

Primers were designed from the published mouse *Mas* gene sequence (GenBank Acc.Nr AJ249895) and from our sequences identified by 5′- and 3′-RACE. Primer sequences were: OMA51: 5′-GGT TAC ATC CCT GGC AAA GG-3′; MRA58: 5′-CAA GCC TCT AGC CCT CTG TCC; and MAS12: 5′-GCC GTT GCC CTC CTG GCG CCT GGG-3′. Primers for the cloning of new exons using promoter prediction program were PS1–PS9 (see Section 3 for detailed primer sequences).

All PCR products were subcloned into the vector pGEM-T (Promega) and sequenced in both directions using m13-reverse, m13-forward, or fragment-specific primers.

2.5. Sequencing

The DNA samples were submitted to automatic sequencing using the thermo sequenase fluorescent-labeled primer reaction by Invitek (Berlin-Buch, Germany).

2.6. RNase-protection assay (RPA)

Mas expression was analyzed by RPA using commercially available Ambion RPA II kit (AMS Biotechnology), according to the protocol of the manufacturer. To prepare probes for RPA, PCR fragments, containing *Mas* cDNA in a range of 200–450 bp (probe sizes and locations are indicated in Figs. 3–7) were subcloned into T-vector (Promega), or pBluescript II KS(+) (Stratagene, La Jolla, CA, USA). Plasmids were linearized with the appropriate enzymes, gel purified and used for in vitro transcription (Promega). The labeled antisense RNA probe was synthesized by 5U of T7, T3, or SP6 RNA polymerase as described (Alenina et al., 2002a). Fifty micrograms of total RNA of the testis and forebrain, and 2 μg of yeast RNA (Y) as a negative control were hybridized with 20,000 cpm of the radio labeled *Mas*-specific probes. The hybridized fragments protected from RNase A + T1 digestion, were separated by electrophoresis on a denaturing gel (5% polyacrylamide, 8 M urea). Dried gels were exposed on imaging plates for 12–48 h and scanned by a FUJIX BAS 2000 Phospho-Imager system (Raytest GmbH, Straubenhardt, Germany). Spots were quantified using TINA 2.0c software. For quantification RPA experiments were run in triplicate. The specificity of protected fragments was determined by comparison with yeast RNA (Y), which represents the background and with yeast RNA, not digested with Rnase (Y^{-}), which represents the undigested probe.

2.7. Software

GenBank was searched using BLAST program (http://www.ncbi.nlm.nih.gov) (Altschul et al., 1990). Promoter regions were analyzed with PROMOTER 2.0 program for the recognition of PolII promoter sequences (Center for Biological Sequence Analysis, http://www.cbs.dtu.dk/services/Promoter/).