Available online at www.sciencedirect.com

SCIENCE DIRECT®

Genomics 85 (2005) 238–244

GENOMICS

www.elsevier.com/locate/ygeno

Variable promoter usage and alternative splicing in five mouse connexin genes

Curtis L. Anderson, Michael A. Zundel, Rudolf Werner*

Department of Biochemistry and Molecular Biology, University of Miami School of Medicine, P.O. Box 016129, Miami, FL 33136, USA

Received 19 August 2004; accepted 10 November 2004

Abstract

Recent work from our lab has demonstrated the importance of alternative promoters and variable 5′ UTRs in the regulation of two connexin genes. To see whether other connexins also utilize multiple promoters to produce different mRNA isoforms, we screened the mouse EST database for variations in the 5′ ends of each connexin EST in UniGene. 5′-RACE analysis of mouse embryo cDNA targeting five candidate genes, *Cx31, Cx40, Cx45, Cx46,* and *Cx47* (approved gene symbols *Gjb3, Gja5, Gja7, Gja3,* and *Gja12,* respectively), revealed the existence of multiple previously unknown exons upstream of the coding region that result in variations in the 5′ UTR of the mRNA. RT-PCR from 17 different mouse tissues revealed that many isoforms are expressed in a tissue-specific manner, with some being the predominant exons found in the tissues tested. Many of the novel 5′ UTRs include upstream open reading frames, suggesting varying translational efficiencies. The expression of alternative 5′ UTRs suggests that connexins, like many genes involved in development, require complex regulation at both transcriptional and translational levels.

The family of genes that code for connexins has a functional, structural, and regulatory complexity that rivals that of the genes of hormone receptors, protein kinases, and transcription factors. Connexins and their functional end-products, gap junctions, have been implicated in countless developmental processes. They have been shown to be responsible for the propagation of action potentials and apoptotic signals and appear to be required for numerous other physiologic responses. Targeted gene deletion experiments have highlighted the importance of the connexin gene family in normal vertebrate development and physiology [1,2].

The widespread use of gap junctions in different tissues and their response to varying stimuli suggest that their expression is highly regulated. Indeed, gap junctions and their subunits, connexins, have been shown to be regulated at the transcriptional, translational, and posttranslational levels in a manner that samples most of the regulatory mechanisms known. The mechanisms by which connexins are able to maintain tight regulation in numerous varied cell types is not well understood. This report describes the transcriptional variability in five connexin genes and highlights the potential for complex posttranscriptional regulation afforded by this variability.

Several examples of transcriptional and translational regulation of connexin genes have already been described. For example, the use of alternative promoters to produce different 5′ untranslated regions (5′ UTRs) of connexin mRNAs in a tissue-specific manner was first shown to occur for connexin32 (*CX32,* approved gene symbol *GJB1*) [3,4]. In addition, one of the CX32 5′ UTRs has been shown to contain an essential internal ribosomal entry site (IRES), which was found mutated in a family with Charcot-Marie-Tooth disease [5,6]. More recently, a tissue-specific alternative promoter was also found for the human connexin40

Abbreviations: 5′-RACE, 5′ random amplification of cDNA ends; Cx, connexin; 5′ UTR, 5′ untranslated region; uORF, upstream open reading frame; IRES, internal ribosome entry site; RT-PCR, reverse transcription-polymerase chain reaction; ESTs, expressed sequence tags.

* Corresponding author. Fax: +1 (305) 243 2058.

E-mail address: rwerner@molbio.med.miami.edu (R. Werner).

0888-7543/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.ygeno.2004.11.007

(approved gene symbol *GJA5*) gene [7]. Recent work in our lab has demonstrated extensive alternative promoter usage and alternative splicing in the mouse connexin43 (approved gene symbol *Gja1*) gene [8]. A highly active IRES element has also been shown to exist within the canonical connexin43 5′ UTR [9].

Based on the evolutionary ontology of the connexin family and the multifarious nature of their involvement in biological processes, it was proposed that translational regulation is likely a common and necessary component of connexin gene regulation requiring alternative 5′ UTRs expressed in a tissue-specific manner [10]. Indeed, there are ample publications citing the discrepancy between the amount of connexin mRNA found in certain tissues and the observed protein levels [11–15]. In addition, single-promoter-based expression analysis has failed to reproduce, in vitro, the complicated expression patterns seen during development and adulthood [16]. This work demonstrates that the usage of alternative promoters is not unique to only a few of the connexin genes but is likely a more common feature. The existence of such variable 5′ UTRs in five additional connexins supports this hypothesis. Data presented here are essential for a complete understanding of connexin gene regulation.

Results

EST and BLAST data mining

Our search of the mouse EST database yielded five candidate connexin genes that had varying gene structure when the ESTs were mapped to the mouse genome. The connexin genes *Cx31, Cx40, Cx45, Cx46,* and *Cx47* (approved gene symbols *Gjb3, Gja5, Gja7, Gja3,* and *Gja12,* respectively) yielded several different mRNA species for each connexin exhibiting variations in their 5′ UTRs. Mapping each EST to its genomic sequence allowed the location of the exons that coded for these 5′ UTRs.

5′-RACE

To obtain more direct sequence information and verify the existence of alternative promoters for these five genes, 5′-RACE was performed on 10- to 12-day mouse embryo cDNA. The commercial library used in our experiments contains cDNAs with adaptors ligated to their 5′ ends. The method of creating this library ensures that cDNAs that contain this adaptor were made only from full-length capped mRNAs. The PCR included forward primers corresponding to the 5′ linker sequence and reverse primers complementary to the coding region. Therefore, all products of the PCR will give both sequence and exact transcription start site information. Fig. 1 shows the agarose gel electrophoresis results for each of the five connexin cDNAs. Each gene yielded several bands, indicating considerable size heterogeneity and suggesting that the mRNA contained varying 5′ UTRs. The cDNAs contained in the individual bands of the electrophoresis gel were reamplified and cloned into pPCR-Script. In addition, the PCR products were cloned en masse into the pZero-Blunt TOPO vector for additional screening. Over 40 clones were selected from each method. Each clone was then mapped to the mouse genome. The resulting new gene structure for each of the connexins is depicted in Fig. 2.

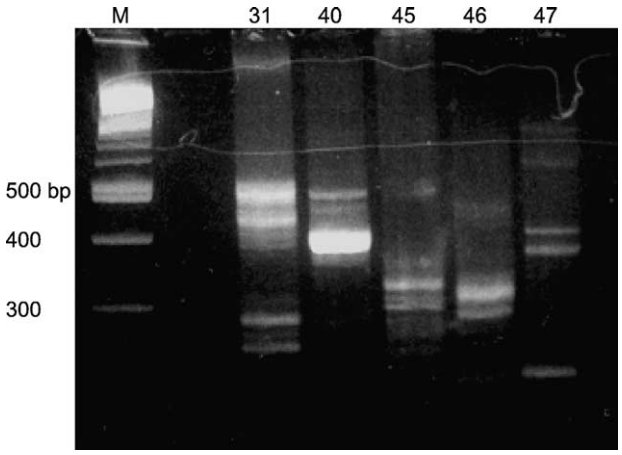

Fig. 1. Analysis of the 5′-RACE products by agarose gel electrophoresis. 10 μl of the PCR reaction was loaded on a 4% low-melting agarose gel. The presence of multiple bands reflects the heterogeneity in the 5′ UTR of five connexin genes. M, marker; 31, connexin31; 40, connexin40; 45, connexin45; 46, connexin46; 47, connexin47.

All clones isolated from connexin31 included a novel exon (1a) that appears to be transcribed from a previously unknown promoter. Exon 1a lies ~3.7 kb upstream of exon 2. Multiple transcription start sites produce transcripts of varying lengths containing from 63 to 343 bp of this exon. All transcripts containing this exon are spliced through a consensus splice acceptor site to exon 2. Two upstream start codons that are in frame with each other lie within exon 1a although neither uAUG contains a Kozak consensus sequence for translation initiation. Only transcripts with the most upstream start sites contain this open reading frame. Exon 1b, previously cloned and described by Hennemann et al. [18], could not be obtained by the 5′-RACE protocol used.

Three different connexin40 transcripts are expressed in mouse embryonic tissues. The previously described transcript, containing exons 1b and 2, is the most prevalent one [19,20]. However, an additional relatively abundant transcript containing a third exon (AS) of only 57 bp located between exon 1b and exon 2 is also found. This alternatively spliced exon AS is located 5 kb downstream of exon 1b in the genomic sequence and was present in some exon 1b-containing transcripts. A third mRNA isoform containing exon 1a is transcribed from a promoter located ~80 kb upstream of exon 2. None of the exon 1a-containing mRNAs appears to include the AS exon.

Cx45 also exhibits a considerable degree of transcript variability. In addition to the previously described con-