Genomic organization of the rat connexin40 gene: identical transcription start sites in heart and lung

W. Antoinette Groenewegen *, Toon A.B. van Veen, Huub M.W. van der Velden, Habo J. Jongsma

Utrecht University, Department of Medical Physiology and Sports Medicine, P.O. Box 80043, 3508 TA Utrecht, The Netherlands

Received 22 July 1997; accepted 10 December 1997

Abstract

Objectives: The gap junction protein connexin(Cx)40 is developmentally and tissue-specifically expressed. How Cx40 expression is regulated is unknown. We therefore set out to characterize the 5'-untranslated end of both the Cx40 gene and mRNA from different tissues and ages and to identify the Cx40 promoter region. Methods: The PCR method 5'-RACE was used to amplify the 5'-end of rat Cx40 mRNAs. Genomic rat Cx40 clones were isolated from a EMBL3 library. The promoter sequence was isolated by long distance PCR. The transcription start site was identified by primer extension and RNase protection assays. Results: Comparison of Cx40 genomic DNA and mRNA sequences revealed that the Cx40 gene contains a small untranslated exon, exon 1, which is separated from the coding sequences by an intron of at least 5.5 Kb. The untranslated 5'-end of Cx40 mRNA sequences from adult rat lung, neonatal and adult rat heart and the rat aortic smooth muscle cell line A7r5 were identical. While the same transcription start site was found for the Cx40 mRNAs from different tissues and ages, the amount of Cx40 mRNA differed between tissues as follows: A7r5 cells > neonatal lung > adult lung > neonatal atrium > neonatal ventricle; Cx40 mRNA from adult atrium and ventricle was not readily detected by primer extension and RNase protection analyses. The genomic sequence upstream of the transcription start site contains multiple consensus binding sites for transcription factors putatively responsible for spatio-temporal control of Cx40 gene expression. Conclusions: Similar to other connexin genes, the Cx40 gene contains two exons. The same exon 1 sequence is present in all tissues and developmental stages examined and the relative amounts of Cx40 mRNA in these compare well with published data. Together our data suggest that tissue-specific and developmentally regulated expression of the Cx40 gene is controlled within the same promoter region by mechanisms that have yet to be detailed. © 1998 Elsevier Science B.V. All rights reserved.

Keywords: Rat; Connexin40; Gap junctions; Expression; Promoter; Gene; Regulation; Tissue-specific; Developmental

1. Introduction

Gap junction channels facilitate the intercellular passage of both electrical and metabolic signals. Each channel consists of two hemi-channels, one in each opposing cell membrane, and in turn each hemi-channel comprises 6 subunit proteins, connexins. At least 13 different, closely related connexin genes have been identified in mammalian cells. They differ in their electrical properties as well as their spatial and temporal patterns of expression [1]. Some connexins show widespread distribution (e.g. Cx43 and Cx32) while others have a much more restricted pattern of expression. In adult rat myocardium, expression of the gap junction protein Cx40 is restricted to the ventricular conduction system, particularly the atrioventricular bundle and bundle branches [2], while in most other species Cx40 is also expressed throughout the atrium [1]. Besides expression in the heart, Cx40 is also expressed in a variety of other adult tissues, including cells of the vascular system, and is highly expressed in lung [3,4]. During rat embryonic
development, Cx40 expression in the working myocardium becomes detectable at embryonic day (ED)13, peaks at ED16 and subsequently declines towards birth [5]. Under hypertensive conditions, i.e. in spontaneously hypertensive rats and in transgenic rats that exhibit hypertension, a 3.1-fold increase in Cx40 protein was detected [2]. Together these studies indicate that Cx40 expression is under precise spatio-temporal control.

Connexin genes appear to have a similar overall genomic structure. The complete coding region is contained in one exon, exon 2. In addition, a small untranslated exon, exon 1, of between 65–250 bp, is separated from exon 2 by a large intron that varies in size from approximately 3–10 Kb, depending on the connexin gene. However, recently, Neuhaus et al. [6] and Sohl et al. [7] demonstrated that the Cx32 gene has two additional exon 1 sequences that can be spliced to exon 2 in a tissue-specific manner. These studies suggest separate promoter sequences upstream of each of the alternative exon 1 sequences.

In order to begin to understand the regulation of the gap junction protein Cx40 at the transcriptional level, the aim of the present study was to identify the upstream regulatory sequences for the rat Cx40 gene. We therefore investigated the structure at the 5' end of the Cx40 gene and its mRNA from different sources.

2. Materials and methods

2.1. RNA isolation from rat tissues and cells

RNA was isolated from adult male Wistar rats and 1–2-day-old neonatal rats as described by Chomczynski and Sacchi [8]. The embryonic rat smooth muscle thoracic aorta cell line A7r5 was obtained from the American Type Culture Collection (ATCC), Maryland, USA, cultured in DMEM supplemented with 10% fetal bovine serum and penicillin/streptomycin, and RNA was isolated by the same method [8]. The investigation conforms with the Guide for the Care and Use of Laboratory Animals published by the National Institutes of Health (NIH Publication No. 85-23, revised 1985).

2.2. 5'RACE

The 5'end of rat Cx40 mRNA was isolated by 5'RACE (Rapid amplification of cDNA ends, Gibco/BRL) as per protocol. Briefly, cDNA was synthesized by Superscript reverse transcriptase using total RNA (1 μg) and antisense primer A (1 μM) (Cx40 bp #303-322 [4]: 5'TTC-CTGCATGGCGACAGTTGT-3'). The RNA template was removed by RNaseH and the resulting cDNA product was purified on Glassmax resin supplied by the manufacturer (Gibco/BRL). The cDNA was tailed with dCTP using terminal transferase. For each tailed cDNA, a control sample was included to which no terminal transferase was added, and this sample served as a negative control in subsequent PCR reactions. Both tailed cDNA templates and their untailed controls were subjected to two consecutive rounds of PCR. In the first PCR round, the Anchor Primer provided with the 5'RACE system (Gibco/BRL), which consisted mainly of dGTP residues to allow annealing with the dCTP tail, was used together with the Cx40 specific antisense primer B derived from the coding sequence (Cx40 bp # 173-192 [4], 5'CAACCAG-GCTGAATGGTATC-3'). The second round of PCR was performed with the Universal Amplification Primer, also provided with the 5'RACE system and which is designed to anneal to restriction sites incorporated into the PCR product by the Anchor Primer, together with antisense primer C (see Fig. 1), situated at the translation start site. PCR products were separated on 1.5% agarose gels, purified by Wizard PCR prep (Promega) and cloned into the pGEM-T vector (Promega). For each PCR product, multiple clones were sequenced using T7 DNA polymerase according to the standard double-stranded DNA sequencing protocol provided by the manufacturer (Pharmacia).

2.3. Isolation of genomic clones

Genomic clones were isolated from a λEMBL3 rat genomic library using a BglII/XhoI fragment of the Cx40 cDNA (plasmid kindly provided by Dr. D.L. Paul, Harvard Medical School, Boston, USA) which contains the first 600 bp of the Cx40 coding sequence as a probe (Fig. 3A). Positive clones were purified by several rounds of re-screening and characterized by restriction enzyme digests and Southern blotting. Suitable fragments were subcloned into pBluescriptSK vector and partially sequenced.

2.4. Long distance PCR-mediated genomic walking

Genomic DNA upstream of the 5' untranslated end of the Cx40 gene was amplified by long-distance PCR as

![Fig. 1. 5'end DNA sequences from Cx40 cDNA isolated by 5'RACE and from genomic Cx40 DNA from a λEMBL3 clone. Capital letters indicate exon 1 and 2 sequences, lower case letters indicate intron sequence, italics indicate bases of the tail sequence. The translation initiation signal ATG is in bold type. The arrows indicate the primer sequences. For the antisense primers C, E and F, the complement was synthesized. These sequences can be accessed in the Genbank database with accession numbers AF025765 (cDNA sequence) and AF025767 (genomic sequence).](image-url)
described by Siebert et al. [9] using the PromoterFinder DNA Walking Libraries from Clontech. This method uses rat genomic DNA that had been digested separately with five different blunt-cutting restriction enzymes (EcoRV, Scal, DraI, PvuII or SspI) after which special, partially double-stranded DNA adaptors had been added to the digested blunt-ended genomic DNA. Specific PCR products were obtained after two consecutive rounds of PCR using nested gene specific primers (see below) together with primers that hybridize to the DNA adaptors. The size of the PCR product depends on the distance between the gene-specific primer and the restriction site with which the genomic DNA had been digested. Two separate ‘PCR-walks’ were performed. In the antisense PCR walk, antisense primers E and F were used (Fig. 1). PCR products were cloned into pGEM-T and for each cloned fragment, multiple clones were sequenced. Two sense primers G and H (Fig. 4) based on the DNA-sequence from the antisense walk were used to perform a sense walk as above. The PCR conditions were exactly as described in the manufacturers protocol (Clontech).

2.5. Primer extension analysis

Primer extension analysis was performed as described [10]. Antisense primer I (Fig. 4) was 5'-end labelled with γ-32P-ATP and gel-purified. Before addition of 8.5 \times 10^5 cpm primer, the RNA was denatured at 70°C for 5 min. Primer and RNA were hybridized at 65°C for 90 min in 15 mM KCl, 10 mM Tris, pH 8.3, 1 mM EDTA. The primer was extended with 5 Units AMV-RT at 42°C for 1 h in 20 mM Tris pH 8.3, 10 mM MgCl2, 4.4 mM DTT, 0.15 mg/ml actinomycin D and 150 μM of each of the 4 dNTP’s. The products were treated with RNase A (14 μg/ml) in the presence of 70 μg/ml salmon sperm DNA, extracted with phenol/chloroform (1:1), precipitated and separated by 6% denaturing acrylamide gel electrophoresis. A sequencing ladder was used as size marker.

2.6. RNase protection analysis

RNase protection was carried out using the RPA II system (Ambion). High specific activity (~8 \times 10^5 cpm/μg) antisense RNA probes were synthesized using the Riboprobe II in vitro transcription system (Promega). Excess probe (~2 \times 10^7 cpm) and total RNA were added together, heated at 95°C for 4 min and hybridized overnight at 42°C in 80% formamide, 100 mM sodium citrate pH 6.4, 300 mM sodium acetate pH 6.4 and 1 mM EDTA. At the end of the hybridization, the samples were treated with 0.05 U RNaseA/0.2 U RNaseT1, for 30 min at 37°C, unless stated otherwise, and the protected fragments were analysed on a 6% sequencing gel.

3. Results

3.1. Isolation and sequencing of the 5'-end of Cx40 mRNA in different tissues

Until recently, ten base pairs of sequence upstream of the rat Cx40 translation start signal ATG were known [4]. Since other connexin genes were known to have a 5'-untranslated exon of 65 bp to 250 bp, situated 3–10 Kb upstream of exon 2, we speculated that this 5'-untranslated exon also exists for the Cx40 gene. This putative DNA was amplified by the PCR method 5'-RACE, initially from adult lung RNA, since this tissue had been shown to express high levels of Cx40 mRNA [4]. A PCR product migrating between DNA marker fragments of 154 bp and 200 bp on an agarose gel was detected after two rounds of PCR with nested primers and no product was detected in the negative control PCR reaction (cDNA not tailed with dCTP, see Section 2.2). The PCR product was cloned and the cDNA sequence upstream of the translation start ATG is given in Fig. 1. To determine whether the same 5'-end sequences were present in Cx40 mRNA from adult or neonatal rat heart, 5'-RACE was also performed on RNA from these tissues. After two rounds of PCR, a shorter band (~120 bp) than in the lung 5'-RACE was amplified from adult heart cDNA while from neonatal heart cDNA, no product could be amplified (not shown). In both cases the control reactions were negative. The ~120 bp PCR product from adult heart was cloned and the sequence corresponded with that found for lung Cx40 cDNA. In order to check if adult heart Cx40 mRNA extends as far upstream as lung Cx40 mRNA, and whether Cx40 mRNA with the same 5'-end is present in neonatal heart RNA, PCR reactions were carried out with a forward primer D (Fig. 1) and reverse primer B (see Section 2.2) using the cDNAs generated for the 5'-RACE as templates. A PCR product of the expected size (~300 bp) could be amplified from both adult and neonatal cDNA (Fig. 2 lanes 1–3), confirming that the 120 bp product amplified from adult heart RNA by 5'-RACE was an incomplete reverse transcriptase product. The fact that the ~300 bp PCR product was also detected in neonatal rat heart cDNA, indicated that Cx40 mRNA from this tissue extends at least as far upstream as in adult heart and lung (see also Fig. 5 below). It is unclear why no PCR product was obtained in the 5'-RACE experiment described above using neonatal RNA. Fig. 2 also shows that more PCR product could be amplified from neonatal than from adult heart cDNA. Furthermore, we tested RNA isolated from A7r5 cells, known to express high levels of Cx40 mRNA [3], for the presence of the same 5'-end sequence. cDNA was synthesized as before (see Methods), after which a 300 bp product was amplified with the same primers D and B (Fig. 1 and Section 2.2; Fig. 2 lanes 4 and 5). The control reaction was negative. This confirmed that A7r5 Cx40 mRNA has the same 5'-end sequence as that given in Fig. 1.
3.2. Sequence of genomic DNA upstream of the Cx40 coding region

We investigated whether an intron is present in rat genomic DNA upstream of the Cx40 coding sequences as suggested by the presence of a putative splice acceptor site upstream of the translation start site in the mouse Cx40 gene [11]. A rat genomic library was screened with a DNA probe derived from the coding region and several positive clones were identified. Two overlapping clones were analysed in more detail and a partial restriction enzyme map was constructed (Fig. 3A). The ~9 Kb HindIII fragment was subcloned and sequenced using primer B (section 2.2). Part of the sequence is shown in Fig. 1 and comparison of this genomic sequence with that obtained by 5'-RACE revealed a sequence divergence approximately 34 bp upstream of the translation start site (see also Section 3.3), indicating the presence of an intron. We therefore conclude that the additional upstream sequence found by 5'-RACE constitutes a separate exon, which we will refer to as exon 1 from hereon. Southern blotting of the genomic clone resulting from the 5'-end of Cx40 cDNAs from different sources with sense primer D (see Fig. 1) and antisense primer B, derived from the coding region (see also Section 2.2). Lane M: DNA size marker (1 Kb DNA ladder from Gibco/BRL; lane 1: adult heart; lane 2: neonatal heart; lane 3: control reaction for lanes 1 and 2; lane 4: A7r5 cells; lane 5: control reaction for lane 4. In both control reactions the cDNA template was replaced by water. The lower band in lanes 4 and 5 is smaller than 75 bp (lowest band of DNA size marker) and is probably a PCR primer artifact.

3.3. Isolation of DNA sequence upstream of exon 1

Long distance PCR-mediated genomic walking was used to amplify DNA either upstream or downstream of exon 1. In the antisense walk, two antisense primers E and F that anneal near the end of exon 1 (Fig. 1) were used to PCR amplify DNA upstream of exon 1. From all five library-digests (libraries), specific products were obtained with the following approximate sizes: 0.35 Kb (PvuII), 3.0 Kb (SspI), 4.0 Kb (ScaI), 4.5 Kb (EcoRI) or 6 Kb (DraI) (not shown). The 0.35 Kb and 3 Kb products were subcloned and partial sequencing immediately upstream of exon 1, i.e., at the 3'-end of the PCR products, showed the same sequence in both cases. Two sense primers G and H (Fig. 4) were synthesized and a subsequent sense PCR walk was performed to confirm that the upstream sequence obtained in the antisense walk does indeed belong to the rat Cx40 gene. Specific PCR products were obtained in the DraI library (~200 bp), the PvuII library (~1.5 Kb) and the SspI library (~0.7 Kb). The latter two products were cloned and the sequence included that of exon 1 as given in Fig. 1. By comparing the (amplified) genomic sequence with the cDNA sequence in Fig. 1, the exact position of the 5'-splice site of the intron can be determined (see also Fig. 4). These results establish that the upstream DNA sequence found in the antisense walk is upstream of the Cx40 gene and confirm the sequence of exon 1 in genomic DNA. The 3 Kb-product from the antisense PCR walk was subjected to restriction enzyme mapping and these data together with those deduced from the PCR walks are given as a restriction map in Fig. 3B. Multiple clones from both the 0.35 Kb and 3 Kb antisense products were sequenced. Some sequence variation was found between these clones, which are presumed to be PCR-generated. Fig. 4 gives the consensus sequence of approximately 1100 bp of upstream DNA for the Cx40 gene. Except for the first base of the API site at −819 bp (T→A), none of the sequence variations alter any of the putative transcription factor binding sites indicated in this figure.

Fig. 3. A) Partial restriction enzyme map of genomic DNA flanking the Cx40 coding sequences (hatched) obtained from two overlapping genomic clones. Probe indicates the part of the coding region DNA that was used as a probe to screen the EMBL3 library. S: SmaI; H: HindIII; Xh: XhoI; Xb: XbaI; E: EcoRI; the lines underneath the map indicate the position of the two clones. Dotted lines indicate that the clone extends beyond this part of the map. B) Partial restriction enzyme map deduced from PCR amplification of genomic DNA fragments (see also Section 2.4) and further restriction enzyme digests of the 3 Kb and 1.5 Kb PCR products from the antisense and sense PCR walks respectively: D: DraI; EV: EcoRV; Sc: ScaI; Ss: SspI; H: HindIII; El: EcoRI; X: XhoI; P: PvuII; N: Ncol. There is an additional PvuII site either 1 or 2 kb upstream of exon 1.
3.4. Transcription start site mapping in different tissues

Fine mapping of the transcription start site for rat Cx40 mRNA was performed by both primer extension and RNase protection assays. RNAs isolated from A7r5 cells, adult and neonatal rat lung, atrium and ventricle were assayed as described in the methods section. By primer extension analysis, in each tissue the same two major bands, 84 and 85 bp long, with about the same intensity were seen (Fig. 5A). The lower band possibly represents incompletely extended cDNA molecules. RNase protection analysis also showed two bands but these were 85 and 87 bp long (Fig. 5B). However, on prolonged incubation with RNase A/T1 mixture (Fig. 5B, lanes 1–3) and with a higher concentration of RNase A/T1 (not shown), the 87 bp band became less intense suggesting that this band represents an incomplete RNase digestion product. We therefore conclude that Cx40 mRNA transcription starts at the C-residue, indicated by the arrowhead in Fig. 4, although possible minor transcription start sites cannot be ruled out completely. The size of exon 1 is therefore 85 bp.

With respect to the level of Cx40 mRNA detected, both primer extension and RNase protection (Fig. 5A and 5B) assays gave comparable results. However, the relative levels of Cx40 mRNA in the various tissues examined were clearly different. A7r5 cells express the highest levels of all tissues and cells examined, followed in order of decreasing abundance by neonatal lung and adult lung, which expresses a level about equal to neonatal atrium. Cx40 mRNA levels in neonatal ventricle were only just detectable. In adult atrium and ventricle, Cx40 mRNA was not detected by these methods using 20 μg total RNA. However, while we could not determine the exact position of the transcription start site in these tissues, a PCR product was obtained from adult heart cDNA (Fig. 2).
using primer D, which is derived from the extreme 5′-end of exon 1. This suggests that the cDNA was sufficiently extended to allow primer D to anneal and indicates that the transcription start site in adult heart is the same or close to that determined in the other tissues.

3.5. Analysis of the DNA sequence upstream of the transcription start site

The genomic DNA upstream of the Cx40 transcription start site contains many consensus sequences for transcription factor binding sites and some are indicated in Fig. 4. At 29 bp upstream of the transcription start site, the TATA-like element TTTAAAAA is found. In addition, a consensus TATA-box sequence is present at −148 bp (not indicated in Fig. 4), but this seems too far upstream to directly transcribe from the transcription start site indicated in Fig. 4. Sites for the basal transcription factors Sp1, AP1 and AP2 are present in the immediate upstream sequence. Around position −16 bp, two overlapping recognition sites for Sp1 are found and these are also overlapped by an AP2 site on the reverse strand. Interestingly, a number of E-box sequences (CAnNTG), which have been implicated in the transcriptional regulation of muscle-specific gene expression, are present throughout this region. Some E-box sequences are more specific for muscle-specific gene expression, are present throughout the myocardium of atria and ventricles, but the left half of the SRF binding site as described by Rivera et al. [12] is present on the reverse strand at −868 bp. The MEF2 binding site CTA(A/T)TAG, related to the CArG box elements, was also not detected in this 5′-upstream region. A GArC element, identified for the human cardiac myosin heavy chain gene, was detected on the reverse strand at −398 bp [13]. A binding site for the factor Nkx-2.5 [14], thought to play an important role in differentiation and maintenance of myocardial cells, is present at −705 bp. No binding sites were detected for the GATA family of proteins which are thought to be important in cardiogenesis and cardiac-specific gene regulation. A consensus binding site for nuclear factor-kappa B (NF-κB) is present at −678 bp on the reverse strand [15]. A number of other putative binding sites are present in 5′-upstream sequence but are not indicated in Fig. 4 for clarity. These include nuclear factor-interleukin-6 (NF-IL6) binding sites at −729 bp, −838 bp and −926 bp (all on reverse strand), the first two sites are also C/EBP binding sites, and CCAAT-boxes at −969 bp and −1053 bp also on the reverse strand. Thus a number of potentially important transcription factor binding sites for Cx40 expression in cardiac and vascular tissue are present in the sequence immediately upstream of the Cx40 gene.

4. Discussion

In the present study we have investigated the 5′-end structure of the rat Cx40 gene and identified the putative upstream regulatory DNA sequences. We have demonstrated that the Cx40 gene structure is similar to that of Cx32, Cx26 and Cx43 consisting of a small untranslated exon 1, located at least 5.5 Kb upstream of the coding sequences in exon 2. The genomic rat Cx40 sequence shown in Fig. 1 is highly homologous to the same sequence region in the mouse Cx40 gene [11], but shows less homology with that for dog Cx40 [16] or the human Cx40 gene [17]. It can be seen (Figs. 1 and 4) that the intron/exon boundary sequences conform to the GT-AG consensus splice site rule.

In each tissue under study, the same, singular transcription start site for Cx40 mRNA was determined by two different methods, primer extension and RNase protection. In addition, the data obtained by 5′-RACE also suggested approximately the same transcription start site. In contrast, for some other connexin genes multiple start sites have been found. For both rat and mouse liver Cx32 mRNA, multiple major and minor transcription start sites have been reported [18–20]. Rat heart Cx43 mRNA was reported to have multiple closely spaced start sites in a stretch of 5 A-residues [21], but only one transcription start site was reported for mouse heart Cx43 mRNA [22]. In addition to this transcription start site, another minor start site may be present in mouse ovary Cx43 mRNA [22]. Mouse liver Cx26 mRNA has one major transcription start site but a minor site may be present, especially in skin tissue [20].

We detected clear differences in the abundance of Cx40 mRNA in the various tissues examined by both primer extension and RNase protection analysis. These methods do not rely on DNA amplification to detect mRNA and therefore are more representative of the relative abundance of Cx40 mRNA in different tissues. The highest level of Cx40 mRNA was detected in RNA from the smooth muscle cell line A7r5, followed in order of decreasing abundance by neonatal lung, adult lung, neonatal atrium and neonatal ventricle. Cx40 mRNA levels in adult atrium and ventricle were undetectable by either of these methods using 20 μg of total RNA per assay but could be detected by 5′-RACE and PCR (Fig. 2, see also below). Cx40 expression in lung tissue is mainly due to the endothelial cells of the vascular system. Our data compare well with published findings. By Northern blotting, Beyer et al. [3] demonstrated high levels of rat Cx40 mRNA in A7r5 cells and Haeffiger et al. [4] detected high levels of Cx40 mRNA in rat lung and a lower level of expression in adult rat heart. By in situ hybridization techniques using cRNA probes, van Kempen et al. [5] have shown that in rat heart Cx40 mRNA expression is developmentally regulated. mRNA levels peak at ED16, Cx40 being expressed throughout the myocardium of atria and ventricles, but
expression subsequently declines in a region-specific manner during foetal development through the neonatal period towards adulthood. In our experiments we used neonatal rats of between one and two days old or adult rats. Our data showed that in adult rat heart Cx40 mRNA levels are extremely low, as may be expected from its restricted expression pattern, such that it could not be detected by primer extension or RNase protection analysis but could be detected by PCR. We detected slightly higher levels of Cx40 mRNA in neonatal atria than in neonatal ventricle which is also in agreement with the study by van Kempen et al. [5] as they described that Cx40 expression first is reduced in the working myocardium of ventricles and later in the atria.

We did not find evidence for the existence of alternative exon 1 sequences for Cx40 in the tissues examined as has been described for Cx32 [6,7]. This does not rule out the possibility that an alternative exon 1 exists for Cx40 since the PCR method 5’-RACE may not have detected a minor alternative mRNA species in our experiments. However, using probes specifically directed at exon 1, our primer extension and RNase protection assays detected Cx40 mRNA levels from different sources that are in agreement with published data (see above). This suggests that, if an alternative exon 1 exists for the Cx40 gene, the regulatory sequences upstream of that alternative exon likely only play a minor contribution to the overall level of expression of Cx40 in the tissues and developmental stages examined in our study.

The DNA sequence upstream of the rat Cx40 gene presented in this study should be the correct upstream sequence for two reasons. First, the upstream sequences found in two independently generated PCR products (0.35 Kb and 3 Kb) by nested gene-specific primers E and F that hybridized to exon 1 were the same. Second, when PCR was performed in the opposite direction (sense walk) with primers G and H from the new sequence, again the sequences found in two independently generated PCR products were the same. Furthermore, these last sequences contained the Cx40 exon 1 sequence generated by 5’-RACE from RNA, thus confirming exon 1 in genomic DNA. Exon 1, as identified by 5’-RACE, belongs to Cx40 mRNA since the cDNA for the 5’-RACE was generated from Cx40 mRNA by antisense primers that start their cDNA synthesis in the Cx40 coding sequence.

For the rat, the promoter sequences for the Cx43 and Cx32 genes have been reported, while for Cx26, the promoter sequence from mouse is known [18–21]. The Cx43 promoter contains the sequence TTTTAAAAA as a degenerate TATA-box sequence [21] while the Cx26 and Cx32 genes were said to have TATA-less promoters [20]. Our results show that the Cx40 promoter contains the TATA-like element, TTTAAAAA, at −29 bp. Interestingly, on closer examination of the mouse Cx26 promoter sequence, the same TA-rich element (TTAAAAA) is found 30 bp upstream of the Cx26 transcription start site. Whether these sequences function as consensus TATA-box elements would, remains to be established. In addition, analogous to the consensus TATA-box we identified for Cx40 at −148 bp, the mouse Cx26 promoter also contains a consensus TATA-box sequence (at −450 bp) [20]. Both these TATA-boxes are presumed to be located too far upstream to direct the initiation of transcription from their known transcription start sites. Thus there are some unexpected similarities between Cx26 and Cx40, in spite of the fact that these two connexins traditionally are assigned to different classes of connexin genes. Both are tissue-specifically expressed, although their expression patterns are not the same.

The Cx40 promoter contains a SP1 binding site immediately upstream of the transcription start site at −16 bp. SP1 has been shown to play an activating role in promoters that lack TATA-boxes [23] and since Cx40 does not have a consensus TATA-box sequence, SP1 could perhaps play a similar role in Cx40 gene transcription. In addition, the SP1 binding site overlaps with an AP2 binding site. AP2 can potentiate the binding of SP1 to GC-box sequence in the human metallothionine IIA gene [24]. We therefore speculate that the SP1/AP2 binding sites may play a role in enhancing the initiation of transcription of the Cx40 gene.

The sequence upstream of the Cx40 gene contains a number of E-box elements. These sequences are involved in the regulation of muscle-specific gene expression [25]. The cardiac-specific expression of myosin light chain-2 and α-myosin heavy chain genes involves E-boxes as part of composite DNA elements in their promoters. In the Cx40 promoter, some E-box sequences are also close to or overlap other known transcription factor binding sites e.g. AP1 at bp −310. E-box elements are binding sites for the bHLH family of transcription factors. These proteins form heterodimers that bind to the E-box sequence. The heterodimer often consists of a ubiquitous bHLH protein such as E2A together with a tissue-specific bHLH protein. Two cardiac bHLH proteins, eHAND and dHAND, have recently been cloned and shown to be essential for cardiac morphogenesis [26]. They are expressed early in mouse embryonic heart development (around ED8.5) and coincides with the time at which Cx40 expression can be detected in the developing mouse heart [27]. The time course of expression of the HAND proteins in the rat cardiac development is unknown. Whether eHAND and/or dHAND have a role in the regulation of embryonic expression of rat Cx40 will need to be determined. It is also interesting to note the consensus binding site for the homeobox protein Nkx-2.5. In the mouse, this protein is expressed in adult heart muscle but again also very early in mouse development (by ED8.5) coinciding with the earliest Cx40 expression (see above). In addition, the promoter for the region-specific and developmentally regulated gene Atrial Natriuretic Factor (ANF) contains binding sites for Nkx-2.5 [28]. These sites were shown to contribute to
region-specific expression of the ANF gene: they were essential for ANF expression in 4-day-old rat atrial cells but not essential in 4-day-old rat ventricular cells. By contrast, the Nkx-2.5 protein itself does not show this regional expression pattern but its role in the regulation of ANF promoter activity was shown to be affected by interaction with other factors such as SRF [29]. Whether Nkx-2.5 plays a similar role in controlling region-specific and/or time-dependent expression of the Cx40 gene warrants further investigation.

The presence of the NF-κB binding site is potentially interesting. The cytokine tumor necrosis factor α (TNFα) can stimulate binding of NF-κB to DNA in a number of different cell types [30]. TNFα has been shown to cause a reduction in gap junctional communication in human smooth muscle cells [31], which express Cx40. In addition, recent results from our laboratory suggest that treatment of human vascular endothelial cells for 48 h with TNFα leads to a down-regulation of the expression of Cx40 [32]. Further, in addition to the similarities between Cx40 and Cx26 mentioned above, the Cx26 promoter also contains a potential binding site for NF-κB at around −150 bp, but whether this site is transcriptionally active is unknown.

In conclusion, our results show that Cx40 mRNA transcription is initiated at the same point in the Cx40 gene in the different tissues and developmental stages examined in this study. This suggests that in these tissues the same region of upstream sequence regulates Cx40 transcription. Therefore, tissue-specific and time-dependent regulation of Cx40 gene expression may be due to differences in the level or activity of different transcription factors. A number of potentially important transcription factor binding sites are present within the first 1100 bp of Cx40 upstream sequence. In future investigations, additional DNA elements involved in spatio-temporal regulation of the Cx40 gene may be identified within this region or upstream of this sequence. Reporter gene constructs containing different parts of the DNA upstream of the Cx40 gene will need to be analysed for their ability to drive transcription in cardiac or other cell types. Subsequent DNA binding studies using nuclear extracts from these cells should begin to elucidate the protein factors involved in the tissue- and time-specific transcriptional regulation of the Cx40 gene.

Acknowledgements

This study was supported by the Netherlands Heart Foundation, grant numbers M93.002 and M96.001.

References


