Analysis of the 5′-upstream regions of the human relaxin H1 and H2 genes and their chromosomal localization on chromosome 9p24·1 by radiation hybrid and breakpoint mapping

J L Garibay-Tupas, K Csiszár1, M Fox2, S Povey2 and G D Bryant-Greenwood

Molecular Endocrinology Laboratory, Pacific Biomedical Research Center, University of Hawaii at Manoa, Honolulu, Hawaii, USA
1Laboratory of Matrix Pathobiology, Pacific Biomedical Research Center, University of Hawaii at Manoa, Honolulu, Hawaii, USA
2MRC Human Biochemical Genetics Unit, Galton Laboratory, University College, London, UK

(Requests for offprints should be addressed to J L Garibay-Tupas, Department of Anatomy and Reproductive Biology, 1960 East–West Road, Honolulu, Hawaii 96822, USA; Email: jerme@pbrc.hawaii.edu)

ABSTRACT

Relaxins are known endocrine and autocrine/paracrine hormones that play a major role in reproduction. In the human there are two relaxin genes, H1 and H2 which share 90% sequence homology within their coding region. The biological and evolutionary significance of two highly homologous and biologically active human relaxins is unknown. In order to achieve a better understanding of the regulatory mechanisms involved in the differential expression of these two genes and to gain insight into their role(s) in the preterm premature rupture of the membranes, we have investigated the properties of their 5′-upstream regions and mapped them both by radiation hybrid and breakpoint mapping into the same chromosome 9p24·1 locus. The 5′ ends of these relaxin genes could be divided into a proximal highly homologous segment and a distal non-homologous region. Within the proximal region are contained several putative regulatory elements common to both genes, suggesting a similar regulatory mechanism. The clustering of the relaxin genes within the same chromosomal locus suggests that these genes may be under a common regulation. On the other hand, a distinct gene-specific regulation may also exist for the individual relaxin genes since cis elements specific to each gene were identified at their 5′ ends. Moreover, the observed divergence at the distal region of their 5′-upstream sequences may provide the structural features that act as gene-specific transcription regulators. Since the two genes are highly homologous in both their coding and flanking regions, the divergence at the distal region of their 5′ ends may be important in the regulation of these genes and in their involvement in the pathology of preterm birth.

Journal of Molecular Endocrinology (1999) 23, 355–365

INTRODUCTION

Relaxin is a 6 kDa polypeptide hormone and a member of the insulin superfamily. In the human there are two non-allelic genes for relaxin, termed H1 and H2 that encode two biologically active relaxin peptides (Crawford et al. 1984). Relaxin is best known as an endocrine hormone produced by the ovary and acting on the mammalian reproductive system to cause uterine quiescence and to remodel connective tissues, thereby facilitating the birth process (Bryant-Greenwood & Schwabe 1994). The systemic relaxin acting in this way is the H2 form and the only relaxin produced by the human corpus luteum (Hudson et al. 1984). It has been shown, however, that relaxins are also autocrine/paracrine hormones in a number of other tissues such as prostate (Hansell et al. 1991), placenta (Sak bun et al. 1987).
An overexpression of the relaxins in the decidua and placenta has been shown to be associated with the preterm premature rupture of the membranes (Bogic et al. 1997), but it is not yet known whether one or both relaxins are upregulated in this condition. Relaxin overproduction in turn upregulates the expression of the matrix metalloproteinases (MMPs), which then increase the degradation of the extracellular matrix in this tissue, leading to tissue weakening, and premature rupture and preterm birth (Qin et al. 1997a, b).

There has been no in-depth study of the genetic etiology of preterm birth, although it has been shown that a previous history of preterm birth is a strong risk factor for another preterm delivery (Carr-Hill & Hall 1985, Porter et al. 1997). Moreover, there are well-described racial differences in the incidence of preterm deliveries (Shiono & Klebanoff 1986, Migone et al. 1991). A single study suggests the association of a rare allele of the tumor necrosis factor-α (TNF-α) gene arising from genetic polymorphism in the promoter region, with a 2.7-fold increased risk of African-Americans delivering preterm (Roberts et al. 1999). There have been no similar studies on the relaxin genes in the ethnic groups in which preterm delivery is markedly increased. Thus, the determination of the precise localization of the two human relaxin genes was undertaken in order to further understand the evolutionary significance of the two relaxins in the human, and to gain insight into the regulation of their expression in the pathology of the preterm premature rupture of the fetal membranes.

### Table 1. PCR primers used for the promoter walking experiments

<table>
<thead>
<tr>
<th>Primer</th>
<th>Outer Primer</th>
<th>Nested Primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relaxin H1 Walk 1</td>
<td>5’-ATAACATCGCTCCATCCATTGG-3’</td>
<td>5’-TTCTAGCAGGTGAACAAAGACAG-3’</td>
</tr>
<tr>
<td>Relaxin H2 Walk 1</td>
<td>5’-CTTCCCTCATTGGATGTC-3’</td>
<td>5’-GTAGTAAACAGACCTCTAGGAGGGG-3’</td>
</tr>
<tr>
<td>Relaxin H1 Walk 2</td>
<td>5’-CTAATCTACATCTCTGGAGCTGG-3’</td>
<td>5’-GATCCACACTGTTCAGGTCATAGAT-3’</td>
</tr>
<tr>
<td>Relaxin H2 Walk 2</td>
<td>5’-TGGTGCTGATTTGCACTGCT-3’</td>
<td>5’-CAGATTGAGGTTGTTCATGCAC-3’</td>
</tr>
<tr>
<td>Relaxin H1 Walk 3</td>
<td>5’-GTAATACGACCTATAGGGA-3’</td>
<td>(Adaptor 2: AP2): nested</td>
</tr>
<tr>
<td>Relaxin H2 Walk 3</td>
<td>(Adaptor 1: AP1): outer</td>
<td>5’-ACTATAGGCGACCTGCTG-3’</td>
</tr>
</tbody>
</table>

### Materials and Methods

#### Isolation of the 5’-flanking regions of the relaxin H1 and H2 genes

Genomic clones containing the 5’-flanking regions of the human relaxin H1 and H2 genes were isolated using a PCR-based chromosome walking method based on the Promoter Walking Kit (Clontech, Palo Alto, CA, USA). The initial outer and nested 3’-gene specific primers were designed for relaxins H1 and H2 based on their published sequences (Gunnersen et al. 1995). Subsequent primers were then designed from the cloned fragments as shown in Table 1. The PCR amplification was performed as per the manufacturer’s instructions using seven cycles of denaturation/annealing at 94 °C, 2 s; 70 °C, 3 min, followed by 37 cycles of denaturation/annealing/extension at 94 °C, 2 s; 65 °C, 3 min, and a final extension at 65 °C for 4 min for the primary PCR reaction. The PCR products generated were then used as templates for the secondary PCR which was carried out under the following conditions: five cycles at 94 °C, 2 s, 70 °C, 3 min; then 20 cycles at 94 °C, 2 s, 65 °C, 3 min; and a final extension at 65 °C for 4 min. The amplified genomic DNAs were purified after agarose gel electrophoresis and subcloned into the TA cloning vector pCR2.1 (Invitrogen, San Diego, CA, USA). Clones were sequenced in both directions by the Biotechnology Molecular Biology Instrumentation Facility at the University of Hawaii, using the double-stranded dideoxychain termination method (Sanger et al. 1977). The 5’ end of relaxin H2 isolated from the promoter library was derived from three overlapping clones, therefore, a single non-overlapping clone was subsequently isolated by genomic PCR using undigested placental genomic DNA obtained from Clontech. A genomic clone was similarly isolated for the relaxin H1 using the same placental genomic DNA, and its identity confirmed.
The primers used for the genomic PCR amplification are shown in Table 2, this was performed with Clontech’s Advantage Genomic PCR kit, using the following conditions: initial denaturation at 95 °C, 1 min; followed by 25 cycles of 95 °C, 10 s; 55 °C, 30 s; 65 °C, 3 min and a final extension of 70 °C for 2 min. The two DNAs were sequenced in both directions in order to confirm the identities of the genomic clones obtained.

**Primer extension analysis of transcription initiation**

Poly(A)+ RNAs were prepared from the prostate gland of a young male obtained from the National Disease Research Interchange (Philadelphia), and from the prostate adenocarcinoma cell line LNCaP. FGC. The prostate was selected because it has been shown to express both relaxin H1 and H2 genes at a relatively high level. The LNCaP cells were grown from the prostate adenocarcinoma cell line LNCaP (Disease Research Interchange (Philadelphia), and the gland of a young male obtained from the National Institutes of Health, Bethesda, MD, USA) in RPMI 1640 medium containing 10% fetal calf serum (Gibco BRL, Grand Island, NY, USA), at 37 °C and 5% CO₂. The poly(A)+ RNAs were extracted using Invitrogen’s (Carlsbad, CA, USA) fast track mRNA kit. The two sets of oligonucleotide primers for each relaxin gene used were the outer and nested primers for promoter walk 1 shown in Table 1. These primers (100 µM) were 5’-end labeled with γ²P-dATP (NEN Research Products, Boston, MA, USA) using T4 polynucleotide kinase (Gibco BRL) then purified by passing the labeled products in a Sephadex G-25 column. Five picomoles of the labeled primers were annealed to 5 µg poly(A)+ RNA using MuLV reverse transcriptase and reagents from PE Applied Biosystems (Foster City, CA, USA), then incubated at 23 °C for 10 min, 42 °C for 15 min and 94 °C for 5 min. The reverse transcribed cDNAs were analyzed by electrophoresis using 8% polyacrylamide gel containing 6 M urea followed by autoradiography of the fixed/dried gel. Plasmid DNAs encoding the 5’ ends of the two relaxin genes were sequenced using the nested primers (Table 1) with reagents from Gibco BRL’s dsDNA cycle sequencing system and following the manufacturer’s instructions. These sequence ladders were used as size markers for the primer extension products.

**Radiation hybrid mapping**

To map the precise chromosomal localization of the relaxin H1 and H2 genes by radiation hybrid mapping, gene-specific primers were made to a specific sequence of each of the relaxins’ 5’ ends with low (37%) homology (Fig. 1). These gene-specific primers were used for the PCR-based screening of the GeneBridge 4 radiation hybrid panel (Walter et al. 1994, Gyapay et al. 1996), obtained from Research Genetics (Huntville, AL, USA). A 579 bp segment of genomic DNA was amplified with the relaxin H1 primers (forward: 5’-TGGAGGTTGCACTGAGCTGAGCTTG-3’ and reverse: 5’-GTGCACATTCCCGGATAGA CCATTG-3’) while those for relaxin H2 (forward: 5’-TCAAGTGCCACTAAAGTATGCC-3’ and reverse: 5’-TGAAGCAGACATAGAAAGG GAG-3’) amplified a 419 bp genomic DNA. Polymerase chain reactions were performed as described by Szabo et al. (1997), with minor modifications. The reaction volume (15 µl) contained 40 mM Tricine–KOH (pH 9.2 at 25 °C), 15 mM KCl, 3.5 mM Mg(OAc)₂, 75 µg/ml bovine serum albumin, 200 µM each dNTP, 1 µM primers, 37.5 ng hybrid DNA and 0.3 µl of 50 × advantage KlenTaq polymerase mix (Clontech). For amplification of the relaxin H1 DNA, the initial denaturation was carried out at 95 °C for 1 min followed by 30 cycles of 95 °C, 10 s; 64 °C, 30 s; 72 °C, 1 min; and a final extension of 72 °C for 2 min. For relaxin H2 these were: initial denaturation at 95 °C for 1 min followed by 30 cycles of 95 °C, 10 s; 55 °C, 30 s; 65 °C, 1 min; and a final extension of 70 °C for 2 min. Two positive controls were used and these were genomic DNA from two human cell lines: cervical carcinoma HeLa cells and human lung fibroblasts (HFLs). The negative control was genomic DNA from hamster. All three controls were included in all amplifications. The PCR products were separated in single four-tiered 1.5% agarose gels in 1 × Tris–acetate (TAE) buffer for 90 min at 100 V and visualized by ethidium bromide staining. Photographs of gel images were taken and used to score the presence of PCR products of the expected size. The 93 radiation hybrid panels were scored following the instructions

**Table 2. PCR primers used for genomic PCR amplification of the 5’-upstream regions of relaxin H1 and H2 genes**

<table>
<thead>
<tr>
<th>Relaxin</th>
<th>5’ Primer</th>
<th>3’ Primer</th>
</tr>
</thead>
<tbody>
<tr>
<td>H1</td>
<td>5’-TCTTAGCTAGTTAAGGGTTTGTCAG-3’</td>
<td>5’-TTCTAGCAGGTGGAACAAGAACAG-3’</td>
</tr>
<tr>
<td>H2</td>
<td>5’-TTCTGGTGAACCTACATAGTCTAGGTAATATG-3’</td>
<td>5’-GATCCACACTGTTTAGCAGTGAATATG-3’</td>
</tr>
</tbody>
</table>

Journal of Molecular Endocrinology (1999) 23, 355–365
available at the website: (http://www-genome.wi.mit.edu/ftp/distribution/human_STS_releases/oct96.INTRO.html#RH) and the results were submitted for linkage analysis to a server at the Whitehead Institute/MIT Center for Genome Research (http://www-genome.wi.mit.edu/cgi-bin/contig.rhmapper.pl). Linkage analysis was carried out at logarithm of the odds (LOD) scores 15, 17, 19 and 21 using RHMAPPER (Stein et al. 1995).

Isolation of a P1 clone

Using a PCR-based strategy the P1 library was screened by Genome Systems Inc. (St Louis, MO, USA), with the same relaxin H2 gene-specific primers used for the radiation hybrid mapping. A single positive clone P1# 16225 was isolated. Before extraction of the P1, plasmid DNA clone 16225, which was maintained in E. coli strain NS3529, was transferred to strain NS3516 by transduction in order to increase the yield of plasmid DNA (Sternberg et al. 1994). The transformed E. coli cells were then grown to stationary phase overnight at 37 °C in LB medium containing 25 µg/ml kanamycin and 1 mM isopropyl β-d-thiogalactopyranoside (IPTG), which induces the lytic replicon to amplify the plasmid copy number (Sternberg et al. 1994). The P1 plasmid DNA was isolated by a modified alkaline lysis procedure, extracted with phenol/chloroform (1:1 v/v), precipitated in isopropanol and washed in 70% ethanol (Sternberg et al. 1994). The size

**FIGURE 1.** (A) The 5’-upstream regions of the human relaxin H1 and H2 genes. The nucleotide sequences of the relaxin H1 (GenBank accession no. AF104934) and H2 (GenBank accession no. AF104935) 5’-upstream regions were compared using the GCG-GAP program. The numbering of the nucleotide residues is relative to the position of the major transcription start site which was designated as (+1). To obtain maximal alignment, gaps were introduced and shown as dots while the vertical bars between sequences show identical nucleotides. The proximal region of these 5’-upstream sequences show 91% homology while the distal region has only 37% homology. The gene-specific primers used for the radiation hybrid mapping were designed from the distal region and are enclosed by solid rectangles for the relaxin H1 primers or dashed/dotted rectangles for the relaxin H2 primers. The CT/GT microsatellite repeats are underlined, the putative TATA boxes are dotted, the ATG translation initiation codon is enclosed in dotted rectangle. Three transcription starts sites were detected through primer extension, the minor upstream (T at −71nt for RLN1 or −76 for RLN2) and downstream (A at +44nt for RLN1; C at +48nt for RLN2) sites are indicated by four-pointed stars, while the major transcription initiation sites (A for both relaxin genes) are indicated by double-sided arrowheads and +1. The nested primers used for the primer extension and for the sequencing ladders are underlined (RLN1) or overlined (RLN2) with arrows showing the direction of cDNA synthesis. (B) A schematic comparison of the relaxin H1 and H2 5’-upstream regions indicate the positions of putative transcription regulatory elements. The major transcription start site (arrow) was used as the reference point for the numbering of nucleotides. Non-canonical TATA boxes for RLN1 and RLN2 and the various putative cis elements are indicated. Most of these elements are concentrated within the proximal region. A dashed vertical line separates the proximal homologous region from the distal non-homologous portion. A nucleotide scale above the diagram shows the length of these regions.
of the P1 clone was estimated by restriction digestion of the plasmid DNA with several endonucleases, following the methods of MacLaren & Clarke (1996). The DNA fragments were analyzed using a Field Inversion Gel Electrophoresis (FIGE) Mapper system (BioRad, Hercules, CA, USA), blotted onto nylon membranes (MSI, Westborough, MA, USA), then hybridized with ³²P-labeled probes specific to either side of the insert cloning site of the vector (MacLaren & Clarke 1996). To verify that both genes were cloned into the P1 plasmid the blots were deprobed then reprobed with ³²P-labeled full-length (558 bp) relaxin H2 cDNA which recognizes both relaxin genes. This was verified further by PCR amplification of the H1 and H2 DNA fragments, using the same gene-specific primers and conditions as described in the radiation hybrid mapping section.

**Breakpoint mapping and fluorescent in situ hybridization analysis**

Fluorescent in situ hybridization (FISH) was carried out on metaphase chromosome spreads obtained from normal male lymphocyte cultures which served as control and from Epstein–Barr virus transformed lymphoblastoid lines obtained from individuals with balanced 9p translocations. There were three lines forming the panel: D1151 with karyotype 46,XY, t(7;9)(q34+1; p24+1); HW1069 with karyotype 46,XX, t(9;11)(p24+1;q22) and D118 with karyotype 46,XX, t(9;10) (p22-3; q11-2) as described by M T Rebello (unpublished observations). Metaphase chromosome spreads were prepared using standard cytogenetic methods. P1 plasmid DNA containing the relaxin H1 and H2 genes was biotin-labeled by nick translation and hybridized to metaphase chromosomes. The signal was detected by fluorescein isothiocyanate (FITC) using techniques described in Gillett et al. (1993). Chromosomes were counterstained with 4',6-diamidino-2-phenylindole (DAPI), examined using a Zeiss Axioskop fluorescent microscope, and images were captured by cooled charge-coupled device (Photometrics) using Smartcapture software (Digital Scientific, Ltd, Essex, UK).

**RESULTS**

**Isolation and analysis of the relaxin H1 and H2 5’-flanking regions**

A 3.0 kb relaxin H1 genomic DNA was isolated after a single round of PCR using the Promoter Walking kit, and was shown to contain the 5’-upstream region. On the other hand, the 2.9 kb 5’-upstream region for relaxin H2 was obtained after a series of three promoter walks. A comparison of the two 5’-end sequences (Fig. 1A) showed an overall homology of 76% (GCG-GAP program, University of Wisconsin), however, two distinct regions with differing sequence homologies were observed. The proximal region spanning ~2.0 kb immediately upstream of the ATG translation initiation codon showed a 91% homology between the two genes, while the ~1.0 kb distal segment showed only 37% homology (Fig. 1B). To confirm that both relaxin 5’-end segments were indeed present in the human genome, genomic PCR was performed using undigested placental DNA to amplify the ~3.0 kb fragments. These genomic DNAs were sequenced and were found to be identical with those obtained through the promoter walking strategy, confirming that the observed drop in the sequence homology at the distal region of both genes was correct.

Analysis of the 5’-upstream regions of relaxin H1 and H2 showed that a typical TATA box is lacking from both genes, instead relaxin H1 has the variant TACTAA, while in relaxin H2 the TACAA which is found in the same position may act as a TATA box (Fig. 1A). A schematic comparison of the relaxin H1 and H2 5’ ends and the presence of several putative regulatory elements are shown in Fig. 1B. Three Sp1 binding sites were identified in relaxin H1 in close proximity to the TACTAA sequence and corresponding two Sp1 binding sites were observed in the relaxin H2 5’ end. Both sequences showed long stretches of CT-GT repeats within −750 to −650 nucleotides. Although the CT-GT repeats appear in the same region of both genes, the sequences are not identical and shared very limited similarities. Both genes have putative mineralocorticoid (MRE) and zinc responsive elements (ZRE), although they were found at different positions. A unique feature for the relaxin H1 5’-end is the GC-rich CArG box, while sites for AP1 and NFkB binding and a glucocorticoid response element (GRE) were found only in the 5’ end of relaxin H2.

Through primer extension analysis using mRNA from the prostate and LNCaP cells, the transcription initiation sites were determined for the relaxin H1 and H2 genes. Each gene has three transcription initiation sites found to be in similar locations within their 5’-upstream regions (Fig. 1A). The major transcription initiation site in both relaxin H1 and H2 genes was found to be in a similar position as the transcription initiation site for the porcine relaxin gene (Haley et al. 1987). Thus, the size of the human relaxin H1 and H2 5’-untranslated
region of \( \sim 109 \) nucleotides corresponds to the reported size of \( \sim 106 \) nucleotides for the porcine relaxin gene (Haley et al. 1987).

**Chromosomal localization by radiation hybrid mapping**

The presence of a region with marked sequence difference (37% homology) between the two relaxin 5'-upstream regions was useful in designing gene-specific primers which allowed the separate chromosomal mapping of each relaxin gene. Initial determination of the chromosomal localization of the relaxin genes was done by radiation hybrid mapping using the GeneBridge 4 panel. Triplicate mapping experiments were performed for both genes and the results were analyzed after agarose gel electrophoresis. The scoring and retention pattern for both genes in all the experiments were identical at all the LOD scores used and comparable with the sequence-tagged site (STS) marker WI-5527, positioned at 30.1 centirays (cR) from the telomeric end of chromosome 9. The distance between WI-5527 and the two relaxin genes was 0.0 cR, thus both relaxin genes are clustered in the same locus as this STS marker which lies on the p24.1 interval of the chromosome 9 short arm (Whitehead Institute map). The two relaxin genes were also located 4.1 cR or 1.35 megabases (Mb) telomeric of STS marker FB2G7 and 5.9 cR or 1.59 Mb centromeric of STS marker D9S286, which covers approximately 8.1 Mb (Fig. 2). In comparison with the genetic map, this position lies within the 9–15 cM (centimorgan) interval (Fig. 2).

**Breakpoint mapping by FISH**

To confirm the radiation hybrid mapping results, breakpoint mapping by FISH was carried out using a P1 clone containing both relaxin H1 and H2 genes, as a probe. This P1 clone has a genomic DNA insert of approximately 80 kb estimated by restriction mapping and Southern blotting, large enough to contain both full-length relaxin genes, each estimated to have a size of approximately 7–9 kb (Hudson et al. 1983). The presence of the relaxin H1 and H2 genes in the P1 clone was confirmed through PCR amplification and Southern blotting using a full-length (558 bp) relaxin H2 cDNA as a probe (data not shown).

FISH was carried out on metaphase chromosomes from normal male lymphocyte cultures and from Epstein–Barr virus transformed lymphoblastoid lines obtained from individuals with constitutional 9p translocations. The control metaphase chromosomes showed intense signals at the telomeric end of the short arm of chromosome (chr) 9 indicated by arrows in Fig. 3a. Finer mapping was achieved using three transformed cell lines bearing various translocation breakpoints. Beginning at the most distal breakpoint on chromosome 9, the cell line D1151 with a karyotype 46,XY, t(7;9)(q34;1;p24.1) gave strong signals nearly at the end of the short arm of the normal (N) chr 9 and its translocation derivative (der), but not in either normal chr 7 or its derivative (Fig. 3b). Using the HW1069 cell line which has a balanced translocation between chr 9 short arm and chr 11 long arm and a karyotype of 46,XX, t(9;11)(p24.1;q22), fluorescent signals were observed in both the normal and chr 9 derivative but not in N11 or in der 11 (Fig. 3c). In the cell line D118 which contains the most proximal balanced translocation with a
46,XX, t(9;10)(p22·3;q11·2) karyotype, signals were detected in those chromosomes (N9 and der 10) where the 9p22·3–9p24·1 interval was retained or translocated, but not in those where these regions were deleted or replaced (der 9 and N10; Fig. 3d).

These data indicate that the relaxin genes are within the p22·3–p24·1 intervals of chromosome 9, and exclude them from the more distal interval 9p24·1–9pter. The combined results of the FISH and the radiation hybrid mapping show that the

Journal of Molecular Endocrinology (1999) 23, 355–365
relaxin genes are clustered in the proximal portion of this interval, namely 9p24·1.

**DISCUSSION**

The 5′-upstream regions of the two human relaxin H1 and H2 genes were isolated and putative regulatory elements that may affect their gene expression were identified. Gene-specific primers were designed, based on a region at their 5′ ends which showed a low percentage homology. These were used to map their individual chromosomal localization using a panel of 93 radiation hybrids. The same primers were also used to isolate a P1 clone containing the full-length relaxin H1 and H2 genes. This P1 clone was subsequently used to further confirm the chromosomal localization of these genes by breakpoint mapping using FISH analysis.

The human relaxin H1 and H2 genes share a very high degree of sequence homology, not only over the coding region (90%) but also within their 5′- and 3′-untranslated regions (UTRs). Both relaxin genes are actively transcribed and translated in several reproductive tissues (Bryant-Greenwood & Schwabe 1994), however, only the relaxin H2 is expressed in the corpus luteum (Hudson et al. 1984). The reason for this differential expression of the relaxin genes is unknown. Likewise, the biological significance of two relaxins in the human is an enigma, since any separate function(s) for either hormone has not been established. In an effort to provide some insights into the mechanisms involved in the differential expression of these genes, we have isolated and analyzed their 5′-upstream sequences. We have also determined their precise chromosomal localization in order to provide insight into their functional significance.

The detailed analysis of the 5′-upstream regions of the relaxin H1 and H2 genes showed that these genes share very similar features. Both genes lack a typical TATA box, but were GC-rich at their proximal regions where several Sp1 binding sites were noted. Moreover, three transcription initiation sites were found in both genes, however no consensus initiator sequences were found flanking these transcription cap sites. These features are very similar with TATA-less and GC-rich promoters of housekeeping genes that are constitutively expressed in low amounts (Ye et al. 1993, Bohm et al. 1995). In many of the tissues producing relaxins, the mRNA level is constantly low and they therefore appear to be constitutively expressed. However, in certain pathological conditions, such as in the preterm premature rupture of membranes, an increased level of total relaxin gene expression was observed in both the decidua and placenta (Bogic et al. 1997), implying an inducible control mechanism. The presence of cis elements such as the MREs, ZREs, and the enhancer core element in both genes suggests that the expression of these genes may be influenced by other metabolites or hormones.

An interesting feature observed in both genes are the CT-GT repeats. These repeats were positioned in the same region of both relaxin genes, but the sequences only had limited homology. These repeats may act as another transcription regulatory element (Meloni et al. 1998) and may influence their tissue-specific expression. The observed putative cis elements were found in both genes and may be involved in a common regulatory mechanism. However, the presence of a CArG box only in relaxin H1, and binding sites for AP1 and NFkB and a GRE only in relaxin H2, suggest that a gene-specific control mechanism may also be involved.

At the more distal region of the relaxin 5′ends, a major decrease in the sequence homology was noted, but its significance is unknown. A similar trend was also noted in the 3′-UTRs of both genes (91% overall homology) where at the distal region, a drop to 44% homology was noted (J L Garibay-Tupas et al. 2000). It is possible that the DNA sequences at the distal 5′-upstream regions of the relaxin genes may act as gene-specific transcription regulators by providing a distinct DNA strand conformation. Since the coding and flanking regions of both relaxin genes are highly conserved, this divergence at the distal 5′ ends may have evolved recently to further define the control mechanisms of each gene. The significance of these regions in the transcriptional control of the relaxin H1 and H2 gene expression is currently being investigated.

The positioning of relaxin H1 on chromosome 9p24·1 confirms the earlier report by Bouzyk et al. (1997), and extends it to show that relaxin H2 is clustered within the same locus. The physical distance between these two genes and their relative orientation within chromosome 9 are still unknown, however, finding both genes within a P1 clone with a genomic insert of 80 kb suggests that they are relatively close together. We do not exclude the possibility that there may be other genes in-between the two relaxin genes or that they may be separated by a short intergenic region, which may even include regulatory elements shared by the two genes. In such a situation, it is likely that the relaxin genes are oriented in a head-to-head direction. Further sequencing information on the genes flanking the two relaxin genes and the putative intergenic region is needed to confirm this.
One of the STS markers that was identified from the Whitehead Institute database to map at the same locus as the two relaxin genes was SGC34067, also known as the INSL4 gene. The chromosomal localization of this gene was earlier reported by Chassin et al. (1995) to the 9p24 locus by FISH analysis, while the Whitehead Institute maps it to a more precise position to 9p24-1. The localization of these three insulin-related genes within the same chromosomal locus probably reflects their evolutionary relationship. These genes are thought to have evolved from a common insulin-like gene through a series of duplication and chromosomal translocations (McRory & Sherwood 1997). The first duplication of this ancestral gene and the subsequent translocation event may have given rise to the modern day insulin which maps at chromosome 11p15 (Harper et al. 1981). A second duplication event followed by translocation to chromosome 9p may have given rise to the INSL4 and relaxin genes, which further diverged in time. While the INSL4 gene did not duplicate further, a primitive relaxin gene may have given rise to two highly homologous relaxin H1 and H2 genes. This event is likely to have been very recent, since the two relaxin genes share a very high sequence homology. Moreover, the only other primate shown to have a second relaxin gene that is actively transcribed is the chimpanzee (Evans et al. 1994).

However, it is not yet known whether the chimpanzee Ch1 relaxin, which is the counterpart of relaxin H1, is functional (Evans et al. 1994). Perhaps, because of the very recent duplication of relaxins, a separate function for relaxin H1 has not totally evolved. Relaxin H1 may act in synergy with and augment the functions of relaxin H2. The identification of a single type of relaxin receptor in the decidua which is believed to mediate the biological actions of both relaxins (Garibay-Tupas et al. 1995) supports the hypothesis of synergism between relaxins H1 and H2. However, we do not exclude the possibility that relaxin H1 may act to regulate or modulate the activity of relaxin H2 by competitively binding to the same receptor.

The results presented in this study will enable us to further investigate the relative roles of relaxin H1 and H2 in the pathology of the preterm premature rupture of the membrane and preterm birth. We are now in a position to further analyze the mechanisms involved in the transcriptional control of the human relaxin genes.

ACKNOWLEDGEMENTS

We acknowledge the technical help and advice of Zoltan Szabo, Shunzhong Bao, and the helpful discussions with Dr Frederick C Greenwood. This work was supported by the Research Centers in Minority Institutions Programs of NIH (RR-03061).

REFERENCES


REVISED MANUSCRIPT RECEIVED 29 June 1999