Human Oct3 gene family: cDNA sequences, alternative splicing, gene organization, chromosomal location, and expression at low levels in adult tissues

Jun Takeda, Susumu Seino and Graeme I.Bell*

Howard Hughes Medical Institute and Departments of Biochemistry and Molecular Biology and Medicine, The University of Chicago, 5841 S. Maryland Avenue, MC1028, Chicago, IL 60637, USA

Received April 13, 1992; Revised and Accepted July 31, 1992 EMBL accession nos Z11898–Z11901 (incl.)

ABSTRACT

Transcription factors containing the POU-domain have been shown to be important regulators of tissue-specific gene expression in the pituitary and lymphoid cells. Using a polymerase chain reaction (PCR)-based strategy, we have searched for similar factors that may be expressed in adult human pancreatic islets. This approach resulted in the amplification of sequences encoding the octamer binding proteins Oct1 and Oct3 (also called Oct4). The isolation of cDNAs encoding Oct3 revealed the expression of two isoforms of this transcription factor termed Oct3A and Oct3B that are generated by alternative splicing. Human Oct3A and Oct3B are composed of 360 and 265 amino acids, respectively, of which the 225 amino acids at the COOH-termini are identical. The sequence of human Oct3A shows 87% amino acid identity with mouse Oct3. Reverse-transcriptase PCR showed low levels of expression of both Oct3A and Oct3B mRNA in all adult human tissues examined. We also isolated and characterized the human Oct3 gene (OTF3) and a related gene, OTF3C. The human Oct3 gene, localized to human chromosome 6 in the region of the MHC complex, spans about 7 kb and consists of five exons. Southern blotting and PCR amplification of human DNA indicated the presence of other OTF3-related genes as has been previously noted in the mouse. Two polymorphisms which can be typed using PCR were identified in OTF3 which will facilitate genetic studies of this gene.

INTRODUCTION

Transcription factors containing the POU-homeo domain (1–3) have been shown to be important regulators of tissue-specific gene expression in lymphoid (4–6) and pituitary differentiation (3,7–9) and in early mammalian development (10–14). To identify members of the POU family of transcription factors that may be involved in the tissue-specific regulation of expression of genes in insulin-secreting cells of the pancreas, we used the polymerase chain reaction and degenerate oligonucleotide primers specific for the homeo-domain to amplify POU-related sequences in human pancreatic islet mRNA. The sequences of two of the PCR products were identical to human Oct1 and two others were homologous to mouse Oct3.

Since the sequence of human Oct3 had not been previously reported, we characterized the human Oct3 cDNAs and showed that two forms of Oct3 mRNA are expressed in adult tissues as a result of alternative splicing. We also isolated and characterized the human Oct3 gene. The human Oct3 gene (designated OTF3) and a pseudogene-like sequence (OTF3C) were localized to chromosomes 6 and 8, respectively. Linkage studies showed that OTF3 is tightly linked to the genes for HLA-A, B, C and DR.

MATERIALS AND METHODS

General methods

Standard procedures were carried out as described in Sambrook et al. (15) and as described previously (16). Human pancreatic islets were provided by D.W.Scharp and P.E.Lacy (Washington University School of Medicine). RNA was isolated using the guanidinium thiocyanate/CsCl method. DNA sequencing was done by the dideoxynucleotide chain-termination procedure (17) after subcloning appropriate DNA fragments into M13mpl8 and M13mpl9. Both strands were sequenced.

Cloning of cDNAs encoding POU-domain transcription factors expressed in human islets

Five micrograms of DNase I-treated total human pancreatic islet RNA were reverse-transcribed using AMV reverse transcriptase (GIBCO BRL, Gaithersburg, MD) and 10 pmol of primer POU-1, 5’-GGGATTCCCGT(G/A)ACC(A/G)ACCCACACC-3’. Amplification was carried out using GeneAmp reagents and a
Isolation of human Oct3 cDNA clones

Human Oct3 cDNA clones were isolated using the rapid amplification of cDNA ends (RACE) procedure (19). For 5' RACE, one microgram of human pancreatic islet RNA was reverse-transcribed using 1 pmol of the antisense primer, 5'-TCGGACACATCCTTCTCGA-3'. Second-strand cDNA was prepared using 10 pmol of the (dT)17-adaptor primer and amplified with the adaptor primer and the internal Oct3-specific antisense primer, 5'-TGCAAGACAAATTTCTCCAGCTGCCTACTAATC-3'. For 3' RACE, the cDNA was prepared using the oligo (dT)17-adaptor primer and amplified with the adaptor primer and the Oct3-specific sense primer, 5'-AGAGGCAACCTGGAGAATT-3'. The PCR products were separated on a 5% polyacrylamide gel. The gel was dried and exposed to X-ray film containing a 32P-labeled probe at 42°C for 16-20 h; and washing—0.1 xSSC and 0.1% sodium dodecyl sulfate at room temperature for 1 h and then at 60°C for 30 min.

Amplification of Oct3A and Oct3B mRNA in adult tissues

cDNA was prepared from one microgram of DNase I-treated total RNA from adult human heart, kidney, liver, spleen, and pancreatic islets and from term placenta using reverse transcriptase (RT) and 200 ng of random hexadeoxynucleotide primers (Pharmacia LKB Biotechnology). The cDNA sample was divided into two tubes and Oct3A and Oct3B sequences were amplified in the presence of 1 μCi of [α-32P]-dCTP (Amersham, 3000 Ci/mmmole) and a common downstream antisense primer (from exon 1b), 5'-CCACATCAGCCTGGTTAGTATAT-3' (from the POU-specific domain and upstream sense primers specific for Oct3A, 5'-CTCTGAGGGCCAGAGAATC-3' (from exon 1a), and Oct3B, 5'-ATGCATATGTGCATGAAAAGC-3' (from exon 1b), mRNAs. The PCR was for 25 cycles of denaturation at 95°C for 1 min, annealing at 55°C for 1 min, and extension at 72°C for 1 min. The products were separated on a 5% polyacrylamide gel. The gel was dried and exposed to X-ray film for 3 h. The expected sizes of the Oct3A and Oct3B-specific products are 381 and 303 bp, respectively.

Isolation of human Oct3 gene and related sequence

The partial Alu/Hae III fetal liver genomic library of Lown et al. (20) was screened by hybridization as described above using kidney cDNA clones as probes.

Gene mapping

The chromosomal locations of the human Oct3 gene (OTF3) and a processed pseudogene-like sequence (OTF3C) were determined using Bios PCRable™ DNA from Somatic Cell Hybrids (Bios, New Haven, CT) and primer-pair specific for OTF3 [sense (exon 3) primer 5'-AGAGGCAACCTGGAGAATT-3' and antisense (exon 4) primer 5'-CTCTGAGGGCCAGAGAATC-3' and OTF3C [sense (5'-flanking region) primer 5'-TGAAGAATCTGACATCAGTATG3' and antisense (POU-specific domain region) primer 5'-AACACCTTCCAATAGAAC-3'].

Allele-specific amplification of OTF3

The T->G replacement in the second position of the codon for the initiating methionine of Oct3B mRNA was characterized by allele-specific amplification using the sense primers 5'-GGCAAGCTTGGAAAGCGAGA-3' and 5'-GGCAAGCTGGAAAGCGAGA-3' and the antisense primer (intron 2) 5'-ATGCAGTGAAAAATAGAAC-3'. Briefly, 50 ng of EcoRl-digested genomic DNA was amplified using the sense (intron 4) primer 5'-AGCTTGTGCTAATGCTAAT-3' and the antisense (3'-flanking region) primer 5'-CAGCTACATGCTGACCTG-3' for 30 cycles: denaturation at 95°C for 1 min, annealing at 63°C for 1 min, and extension at 72°C for 2 min. The 802 bp PCR product was digested with Hind III and the fragments were separated on a 1.2% agarose gel; the presence of the Hind III site gave fragments of 546 and 256 bp on digestion.

RESULTS

Identification of POU-domain transcription factors in pancreatic islets

Messenger RNAs encoding members of the POU-family of transcription factors expressed in human pancreatic islets were amplified in vitro using the primers POU-1 and -2 which correspond to conserved regions of the homeo domain (1-3). Since the expected size of POU-homeo-domain-containing PCR products obtained using these primers was 122 bp, fragments of this size were cloned into M13mp18 and sequenced. 100 clones were sequenced and of these four were homologous to the homeo-domain of the POU family. The sequences of two clones, A1 and A33, were identical to that of human Oct1 cDNA (23). The sequences of the other two clones, K27 and A11, were related...
A. Human Oct3A

B. Human Oct3B

![Composite nucleotide sequences of human Oct3A (A) and Oct3B (B) cDNAs and predicted amino acid sequences of the proteins.](image-url)

**Figure 1.** Composite nucleotide sequences of human Oct3A (A) and Oct3B (B) cDNAs and predicted amino acid sequences of the proteins. The number of the nucleotide at the beginning of each line is shown. The arrows indicate the sites, upstream of which, the sequences of Oct3A and Oct3B cDNAs diverge. The putative polyadenylation site is underlined as is the in-frame translational stop codon upstream of the putative initiating methionine in Oct3B. Sequence differences among cDNA clones are shown. They are C, C, G, T, TGCC, and CAC at nucleotides 92, 209, 780, 885, 910 and 1155, respectively, in clone AKHO-1; and A and G at nucleotides 1087 and 1136, respectively, in clone KXHO-14. The accession numbers in the EMBL/GenBank database for the human Oct3A and Oct3B sequences are Z11809 and Z11898, respectively.

except that A1 lacked a CAG triplet encoding Gin. The sequence of clone K27 showed 75% and 79% amino acid and nucleotide cDNA clones encoding the human pancreatic islet Oct3-like sequence identity, respectively, with the sequence of mouse Oct3 (11-13) suggesting that it may be the human homolog of Oct3.

**Sequence of human Oct3**

cDNA clones encoding the human pancreatic islet Oct3-like sequence were generated by direct amplification as described above and using 5'- and 3'-RACE. The composite cDNAs span
Except for A and G substitutions at nucleotides 1087 and 1136 below this sequence difference represents an RFLP. Change results in the loss of a 532 bp insert corresponding to nucleotides 624 to 1155 of the human sequence. The nucleotide sequence of the insert in XKHO-14 was identical to the corresponding region of XKHO-4 and -9 sequences. The nucleotide sequence of the insert in XKHO-14 lacked the ATG of the initiating Oct3; this sequence was designated as human Oct3A. Since the overall identity between the sequences of human Oct3A and Oct3B cDNA, a region that is common to both Oct3A and Oct3B (Fig. 2). The fourth kidney cDNA clone, XKHO-14, had a 532 amino acids with 87% identity with the sequence of mouse Oct3 and eight of the 56 differences are insertion/deletions.

Figure 2. Comparison of amino acid sequences of human Oct3A and mouse Oct3. Only differences between the two sequences are noted. Gaps introduced to generate this alignment are represented by colons.

In order to identify other Oct3-related sequences expressed in human tissues, we screened an adult human kidney cDNA library by hybridization. Of 1.2 \times 10^6 phage screened, four clones were identified. One of them \( \lambda \)KHO-4 contained a 1065 bp insert whose sequence was identical to the corresponding region of the composite pancreatic islet Oct3 sequence. The second clone, \( \lambda \)KHO-1, having an insert of 1161 bp, showed 99% nucleotide sequence identity with \( \lambda \)KHO-4 and encoded a protein whose sequence was 99.3% identical with that of mouse Oct3 (Fig. 1B). The third clone, \( \lambda \)KHO-9, had an insert of 1328 bp, excluding polyA tract, and encoded a protein of 359 amino acids with 87% identity with the sequence of mouse Oct3 including a region homologous to amino acids 2—128 of mouse Oct3; this sequence was designated as human Oct3A. Since the sequence of the insert in \( \lambda \)KHO-9 lacked the ATG of the initiating methionine, this region of the cDNA was obtained using the 5'-RACE procedure and human adult kidney RNA; a 45 bp fragment was isolated in this manner. The composite sequence of the human Oct3A cDNA is 1373 bp and encodes a protein of 360 amino acids (\( M_r = 38.6 \) kDa) (Fig. 1A). There is 87% overall identity between the sequences of human Oct3A and mouse Oct3 and eight of the 56 differences are insertion/deletions (Fig. 2). The fourth kidney cDNA clone, \( \lambda \)KHO-14, had a 532 bp insert corresponding to nucleotides 624 to 1155 of the human Oct3B cDNA, a region that is common to both Oct3A and Oct3B sequences. The nucleotide sequence of the insert in \( \lambda \)KHO-14 was identical to the corresponding region of \( \lambda \)KHO-4 and -9 except for A and G substitutions at nucleotides 1087 and 1136 of the Oct3B cDNA sequence, respectively. The latter nucleotide change results in the loss of a Hind III site (Fig. 1B); as described below this sequence difference represents an RFLP.

Figure 3. Expression of Oct3A and Oct3B mRNAs in human tissues by RT-PCR. RT-PCR was carried out as described in Material and Methods. The PCR fragments derived from Oct3A and Oct3B mRNAs are shown.

Expression of Oct3 mRNA

Since our cDNA probes did not show hybridization to RNA blots prepared using total RNA from adult heart, kidney, liver and spleen and term placenta, we used RT-PCR to screen for the presence of Oct3A and Oct3B mRNA as described in Materials and Methods. Since the primer pairs used span an intron in the Oct3 gene, PCR products derived from Oct3A and Oct3B mRNA can be readily distinguished from those generated by amplification of any contaminating DNA. PCR-products of the expected size of 381 and 303 bp derived from human Oct3A and Oct3B mRNA, respectively, were observed in all the adult tissues that we examined as well as pancreatic islets (Fig. 3). In pancreatic islets, the relative abundance of Oct3B mRNA was higher than Oct3A, whereas in the other tissues that we tested, the levels were similar.

Isolation and characterization of the human Oct3 gene

About \( 1 \times 10^6 \) phage were screened by hybridization with \(^{32}P\)-labeled Oct3B cDNA clone. Twenty-one clones were identified and of these five hybridized with probes derived from the 5'- and 3'-ends of the Oct3A and Oct3B cDNA clones. These latter clones were characterized further.

Clones \( \lambda \)HG-13, -14, -15 and -16 formed a contig of 18 kb that included the human \( OTF3 \) (Fig. 4A). The sequence of a region of 6.2 kb indicated that \( OTF3 \) was comprised of five exons (Fig. 4A, Table 1) that span \(-7 \) kb. The exon-intron organization of the human Oct3 gene is very similar to that previously determined for the mouse gene (24); the four introns located in identical positions in both genes. Of interest in this gene is the presence of an atypical splice donor site at the beginning of intron 1a. The human Oct3 gene has a GC here rather than the usual GT. As reviewed by Jackson (25), there are 26 other examples where the 5'-splice site is GC which is the only variant of the 5'-splice site that has been identified. The biological consequences of the presence of a GC with respect to expression of Oct3A and Oct3B mRNAs are unknown.

A comparison of the sequences of the gene and of the Oct3A and Oct3B cDNA clones indicated that Oct3A and Oct3B mRNA are derived from the same gene by alternative splicing (Fig. 4B). Oct3A mRNA is generated from exon 1a, the 3'-half of exon 1b, and exons 2—4 through the use of an internal splice acceptor site in exon 1b. Oct3B mRNA includes exons 1b—4.

The sequence of the exon 1b of clones \( \lambda \)HG-13 to -16 showed one interesting difference. Clones \( \lambda \)HG-13, -14 and -15, whose sequences appear to be identical, have AGG in the position of the initiating methionine of Oct3B mRNA while \( \lambda \)HG-16 has an ATG at this position. Since the nucleotide sequences of the exons,
Figure 4. Exon-intron organization of the human Oct3 gene and alternative splicing. A. Map of clones comprising the human Oct3 gene. The five exons are shown as boxes. The locations of the Hind III (H) and ATG/AGG polymorphisms described in the text are shown. The locations of sites for several restriction endonuclease are noted: S, Sal I; R, EcoRI; and X, Xba I. B. Alternative splicing of Oct3 mRNA. The arrow shows the location of the internal splice acceptor site in exon 1b.

Table 1. Exon-intron organization of OTF3

<table>
<thead>
<tr>
<th>Exon no.</th>
<th>Exon size</th>
<th>5'-Splice donor</th>
<th>3'-Splice acceptor</th>
<th>Intron size</th>
<th>Amino acid interrupted</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>447 bp</td>
<td>GAGGAG gaggagag...</td>
<td>GAAAGAGG GCAAAG</td>
<td>5 kb</td>
<td>Glu-133</td>
</tr>
<tr>
<td>1b</td>
<td>344 bp</td>
<td>TATTTG gaggagag...</td>
<td>GAAAGAGG GCAAAG</td>
<td>227 bp</td>
<td>Gln-176</td>
</tr>
<tr>
<td>2</td>
<td>122 bp</td>
<td>GAGGAG gaggagag...</td>
<td>GAAAGAGG GCAAAG</td>
<td>285 bp</td>
<td>Gln-219</td>
</tr>
<tr>
<td>3</td>
<td>159 bp</td>
<td>AAGGAT gaggagag...</td>
<td>GAAAGAGG GCAAAG</td>
<td>260 bp</td>
<td>Asp-272</td>
</tr>
<tr>
<td>4</td>
<td>513 bp</td>
<td></td>
<td></td>
<td>3'-end</td>
<td></td>
</tr>
</tbody>
</table>

The sites at which introns interrupt the mRNA and protein sequence are indicated. An atypical splice donor sequence occurs in intron 1a, GC. Exon and intron sequences are shown in capital and lower-case letters, respectively. Note that part of Oct3B mRNA is encoded by the 'intron' upstream of exon 1b. The partial sequence of human OTF3 has been deposited in the EMBL/GenBank database with accession number Z11900.

Figure 5. DNA polymorphisms in the human Oct3 gene. A. and B. Allele-specific amplification of the ATG and AGG alleles, respectively. The genotypes of the five individuals shown here are from left to right: ATG/AGG, ATG/ATG, ATG/ATG, ATG/AGG and AGG/AGG. C. Characterization of Hind III polymorphism. The 802 bp PCR product was digested with Hind III. The genotypes of the individuals shown are from left to right: +/-, -/-, +/-, +/+ and +/-.

Whereas the sequences of clones AHG-13 to -16 indicate that OTF3 is interrupted by introns, the sequence of the region of clone AKH-17 homologous to Oct3 mRNA showed 97.5% nucleotide sequence identity with the sequence of Oct3C cDNA and lacked introns (Fig. 6). The presence of a polyA tract at the 3'-end of the Oct3-like sequence in AKH-17 and of flanking direct repeats are consistent with this sequence, designated as OTF3C, being a retroposon (26). The sequence of the Oct3C-like protein encoded by AKH-17 differs from Oct3A at 14 residues and by the deletion of a triplet corresponding to Gln238; a similar deletion was noted in one of the PCR products, A11, obtained using DNase-I-treated human pancreatic islet RNA.

Hybridization of Southern blots of EcoRI and Hind III digests of human DNA with an Oct3B cDNA probe revealed multiple hybridizing fragments (Fig. 7) suggesting that there may be a family of Oct3-related sequences in humans as there are in mice (27). By contrast, the relatively simple patterns noted with
4 and 5 (Fig. 8). The comparison of the sequences suggest that clones showed that one was identical to intron-containing OTF3 pseudogenes since they contain in-frame translational stop codons.

The Oct3B probe hybridized to human DNA from a panel of somatic-cell hybrids, the human OTF3 family. The Oct3B probe hybridized to human, monkey (Macaca nemestrina), rat and guinea pig DNAs suggest that they may also be species differences in the size of this gene family. The Oct3B probe hybridized to human EcoRI fragments of 17, 12, 8.2, 6.1, 4.8 and 2.4 kb of which the 6.1 and 17 kb fragments correspond to the sizes of the EcoRI fragments family. The Oct3B probe hybridized to human OTF3C, respectively.

In order to estimate the size of OTF3 family, we amplified human DNA with the primers used for mapping OTF3 described in Material and Methods. Agarose gel electrophoresis of the PCR products showed fragments of 441 and 181/178 bp corresponding to the products from genes with and without introns in this region, respectively (data not shown). The sequences of five clones showed that one was identical to intron-containing OTF3 sequence. The other four clones lacked the intron between exons 2-3. The Oct3B probe hybridized to human DNA from a panel of somatic-cell hybrids, the human OTF3 family. The Oct3B probe hybridized to human, monkey (Macaca nemestrina), rat and guinea pig DNAs suggest that they may also be species differences in the size of this gene family. The Oct3B probe hybridized to human EcoRI fragments of 17, 12, 8.2, 6.1, 4.8 and 2.4 kb of which the 6.1 and 17 kb fragments correspond to the sizes of the EcoRI fragments family. The Oct3B probe hybridized to human OTF3C, respectively. The boundaries of the Oct3A-like sequence are noted by square brackets. The direct repeats at the ends of this sequence are underlined. The accession number in the EMBL/GenBank database for the sequence of OTF3C is Z11901.
The human Oct3 gene, has been mapped to chromosome OTF3, 6. The are similar, they are not identical. Since we expected to isolate the region of the promoter although there are related AT-rich expressed in embryonic tissues where it is involved in the data) or glucokinase (35) contain a canonical octamer motif in expressed in pancreatic islets including insulin (33), islet amyloid elements. Thus, the target(s) for Oct3A and Oct3B in pancreatic and Oct3 bind to the octamer motif ATGCAAAT to activate of adenovirus DNA replication suggesting that it may be involved POU-family, Oct3, had only been shown previously to be Our data indicate that Oct3 mRNA is present at low but detectable the region of the promoter although there are related AT-rich expressed in embryonic tissues where it is involved in the regulation of differentiation of embryonic stem cells (10—14). Our data indicate that Oct3 mRNA is present at low but detectable levels in adult human tissues. In addition, whereas only a single form of Oct3 mRNA has been identified in embryonic mouse tissues, two forms generated by alternative splicing were identified in adult human tissues. Oct3A and Oct3B mRNAs encode proteins that share a common DNA-binding POU-homeo-domain but differ in sequence at their NH2-termini. This region of Oct3A is rich in proline and glycine residues (24 and 28 of 135 amino acids, respectively) whereas the NH2-terminus of Oct3B is not. Since proline-rich regions form part of the transcriptional activation domain of other transcription factors such as Oct2 and CTF/NF1 (31,32), alternative splicing of Oct3 mRNA may generate proteins with common DNA-binding domains but different activation domains thereby endowing Oct3A and Oct3B with different functional properties. Both Oct1 and Oct3 bind to the octamer motif ATGCAAAT to activate transcription (1,2). However, none of the cell-type specific genes expressed in pancreatic islets including insulin (33), islet amyloid polypeptide (34), GLUT2/liver glucose transporter (unpublished data) or glucokinase (35) contain a canonical octamer motif in the region of the promoter although there are related AT-rich elements. Thus, the target(s) for Oct3A and Oct3B in pancreatic islets is unknown.

There are multiple Oct3-like genes in both mice and humans. The human Oct3 gene, OTF3, has been mapped to chromosome 6 and is tightly linked to the HLA loci. The cognate mouse gene, Octf-3, has been localized to chromosome 17 (14,27) in a region which is syntenic with the MHC region of human chromosome 6. The OTF3 maps to chromosome 8 and thus may be the human homolog of the locus Octf-3c on mouse chromosome 3 (27).

Although the nucleotide sequences of XKHO-1, -4, and -14 are similar, they are not identical. Since we expected to isolate only two different sequences corresponding to the products of both alleles of OTF3, these results imply that one or more of the Oct3-related genes is transcribed. Moreover, the isolation of a RT-PCR product from pancreatic islet RNA, clone All, whose sequence was identical to that of Oct3-related gene-4 is identical to OTF3C.

**DISCUSSION**

Using a PCR-based strategy, two members of the POU-family of transcription factors were identified in human pancreatic islets. One of these, Oct1 (23), is expressed in many tissues and has been implicated in transcription activation of a number of genes including histone H2B (28). Oct1 can also stimulate initiation of adenovirus DNA replication suggesting that it may be involved in the regulation of the cell cycle (29,30). The other member of the POU-family, Oct3, had only been shown previously to be expressed in embryonic tissues where it is involved in the regulation of differentiation of embryonic stem cells (10—14).

Our data indicate that Oct3 mRNA is present at low but detectable levels in adult human tissues. In addition, whereas only a single form of Oct3 mRNA has been identified in embryonic mouse tissues, two forms generated by alternative splicing were identified in adult human tissues. Oct3A and Oct3B mRNAs encode proteins that share a common DNA-binding POU-homeo-domain but differ in sequence at their NH2-termini. This region of Oct3A is rich in proline and glycine residues (24 and 28 of 135 amino acids, respectively) whereas the NH2-terminus of Oct3B is not. Since proline-rich regions form part of the transcriptional activation domain of other transcription factors such as Oct2 and CTF/NF1 (31,32), alternative splicing of Oct3 mRNA may generate proteins with common DNA-binding domains but different activation domains thereby endowing Oct3A and Oct3B with different functional properties. Both Oct1 and Oct3 bind to the octamer motif ATGCAAAT to activate transcription (1,2). However, none of the cell-type specific genes expressed in pancreatic islets including insulin (33), islet amyloid polypeptide (34), GLUT2/liver glucose transporter (unpublished data) or glucokinase (35) contain a canonical octamer motif in the region of the promoter although there are related AT-rich elements. Thus, the target(s) for Oct3A and Oct3B in pancreatic islets is unknown.

There are multiple Oct3-like genes in both mice and humans. The human Oct3 gene, OTF3, has been mapped to chromosome 6 and is tightly linked to the HLA loci. The cognate mouse gene, Octf-3, has been localized to chromosome 17 (14,27) in a region which is syntenic with the MHC region of human chromosome 6. The OTF3 maps to chromosome 8 and thus may be the human homolog of the locus Octf-3c on mouse chromosome 3 (27).

Although the nucleotide sequences of XKHO-1, -4, and -14 are similar, they are not identical. Since we expected to isolate only two different sequences corresponding to the products of both alleles of OTF3, these results imply that one or more of the Oct3-related genes is transcribed. Moreover, the isolation of a RT-PCR product from pancreatic islet RNA, clone All, whose sequence was identical to that of OTF3C suggests that this locus is transcribed. SCIP/Oct6/Tst-1, another member of the POU-domain family of transcription factors that is expressed in proliferative glial progenitor cells, also has features of an expressed retroposon (36). If OTF3C mRNA encodes a protein, this protein may have altered DNA binding properties because of a deletion of one amino acid in a region involved in DNA binding as well as in specific interactions with other proteins including other members of the POU-family (37—40).

In our analysis of OTF3, we identified two polymorphisms one of which was in the codon that we believe represents the initiating methionine of Oct3B, ATG—AGG. This polymorphism has important functional consequences since individuals having AGG should not be able to express Oct3B. We have confirmed this polymorphism using a PCR-based strategy and the frequencies of the three genotypes, ATG/ATG, ATG/AGG and AGG/AGG, among a groups of 51 healthy unrelated Caucasians was 0.41, 0.51 and 0.08, respectively. These results indicate that a
significant fraction of the population has no or reduced levels of Oct3B protein. Since these individuals are otherwise normal, the absence of Oct3B seems to have few if any physiological consequences. Moreover, we have examined the sequence of the mouse Oct3 gene (24) for a region homologous to that encoding the unique NH2-terminus of human Oct3B. The sequences of the human and mouse genes in the region upstream of exon 2 are quite divergent suggesting that the mouse gene does not encode a protein similar to human Oct3B. This result implies that Oct3B does not serve a function that is crucial for mammalian development.

Our analysis of human pancreatic islet RNA for the presence of mRNAs encoding members of the POU-family resulted in the identification of Oct1 and Oct3. Further studies will be necessary to determine whether the endocrine pancreas, like the pituitary, expresses a tissue-specific POU domain protein.

ACKNOWLEDGEMENTS

We thank Drs. H. Singh and E. B. Crenshaw III for their critical reading of this manuscript. These studies were supported by the Howard Hughes Medical Institute and NIH grants DK-20595 and 42086.

REFERENCES