Introduction

A number of mammalian developmental control genes have recently been isolated and characterized by reversed genetics (for review see Kessel and Gruss, 1990). This approach has greatly enhanced our understanding of the molecular processes that take place during mammalian development. Most of these developmental control genes, such as homeobox (Hox), paired box (Pax) and POU genes, encode transcription factors (Gehring et al., 1990; Treisman et al., 1991; Chalepakis et al., 1991; Rosenfeld, 1991). It is believed that these genes are capable of modulating the expression of corresponding dependent downstream target genes, thus providing positional information and/or controlling organogenesis (Kessel and Gruss, 1990).

The paired box was first identified in the Drosophila segmentation genes paired (Frigerio et al., 1986), gooseberry-distal and gooseberry-proximal (Baumgartner et al., 1987; Côté et al., 1987), as a conserved domain of 128 amino acids (Bopp et al., 1986). To date, eight paired box genes from mouse (Walther et al., 1991) and three such genes from man (Burri et al., 1989) have been isolated on the basis of their homology to the paired box sequences of Drosophila. The restricted temporal and spatial pattern of the expression of Pax genes during mouse embryogenesis suggests a regulatory role in mammalian development.

For example, the Pax1 gene is expressed in the sclero-
tome cells which form the intervertebral disks of the vertebral column (Deutsch et al., 1988). Pax2 is expressed in particular cells along the entire developing spinal cord, in the optic and otic vesicles, in the Wolffian duct and nephric cord and subsequently in the ureter and condensed metanephric mesenchyme (Dressler et al., 1990; Nornes et al., 1990; Dressler and Douglass, 1992). Another murine paired box gene, Pax3, contains a paired type homeodomain in addition to the paired domain and has been shown to bind DNA in a sequence-specific manner (Goulding et al., 1991). The Pax3 gene is expressed exclusively during embryogenesis, and its expression is limited to the dorsal part of the developing spinal cord as well as to distinct regions in the brain (Goulding et al., 1991). Preliminary characterization of a further paired box gene with an additional homeobox, Pax7, reveals its expression in the developing nervous system and in myoblasts (Jostes et al. 1990).

The role of Pax genes in mammalian development is further substantiated by a recent finding that certain congenital malformations in mice and man are associated with mutations in Pax genes. Mutations in the Pax3 gene are associated with splotch phenotype in mice (Epstein et al., 1991) and with the Waardenburg’s syndrome in man (Tassabehji et al., 1992; Baldwin et al., 1992). Mutations in the Pax6 gene cause a semidominant mutation small eye (Sey) in mice (Hill et al., 1991; Walther and Gruss, 1991) and aniridia disorder in humans (Ton et al., 1991).

The last isolated member of the murine paired box gene family, Pax8, is highly homologous to Pax2, and both genes share an overlapping, but distinct, pattern of expression during murine embryogenesis (Plachov et al., 1990). Pax8 is expressed during initial stages in induced parts of the developing excretory system, i.e. in the nephric cord and metanephric condensations (Plachov et al., 1990); regions undergoing intense differentiation during mesenchyme–epithelium transition (Saxén, 1987). In the developing spinal cord, the expression of Pax8 is transient, coinciding with that of Pax2 during days 11-12 of embryogenesis. Pax8 is also expressed during very early stages of thyroid differentiation (Plachov et al. 1990) and in adult thyroid (Zannini et al., 1991). Binding sites for Pax8 within the promoters of two thyroid-specific genes have been extensively characterized leading to the suggestion that Pax8 could function as a transcription control factor (Zannini et al., 1991).

In this report, we describe the isolation and sequence of the human PAX8 gene, with the aim of investigating its expression during human embryogenesis and its possible involvement in congenital malformations. In addition, the possible involvement of PAX8 in Wilms’ tumor, an embryonic tumor of the kidney (Mierau et al., 1987), was evaluated. We have found a high degree of conservation between the amino acid sequences of the human and mouse Pax8 proteins. Furthermore, we describe evidence for differential splicing of the human PAX8 transcript that results in removal of a 63 amino acid serine-rich region at the carboxy end of the PAX8 protein. These spliced PAX8 transcripts are homologous to the sequence of the Pax2 gene in this domain. We demonstrate the presence of such transcripts in normal human kidney, five Wilms’ tumors and human thyroid by RNAase protection assays. Interestingly, these alternatively spliced transcripts are not detectable in the mouse.

Using somatic cell hybrids and comparative mapping data, we mapped the human PAX8 gene to the distal part of human chromosome 2q.

### Materials and methods

**cDNA library screening**

The mouse 2A cDNA (Plachov et al., 1990) fragment encompassing most of the paired box was labeled by random priming (Feinberg and Vogelstein, 1983) and used as a probe to screen a human adult kidney cortex cDNA library in λgt10 (Bell et al., 1986) under low-stringency hybridization conditions (42% formamide, 5×SSPE, 5×Denhardt’s, 0.5% SDS at 37°C; washing in 2×SSPE, 0.1%SDS at 37°C). In total 1.2×10^6 plaques were screened.

**DNA sequencing**

Overlapping restriction fragments of the human PAX8 cDNA were cloned into Bluescript KS (Stratagene) plasmid and sequenced by the dideoxy termination method (Sanger et al., 1977) using a T7 sequencing kit (Pharmacia).

**RNA isolation and RNAase protection analysis**

Total cellular RNA was prepared by the acid guanidinium thiocyanate-phenol-chloroform extraction method (Chomczynski and Sacchi, 1987) from human tissues (kidney, thyroid and Wilms’ tumors), mouse tissues (kidney, brain and liver) and from the human cervix carcinoma cell line HeLa. RNAase protection experiments were carried out according to Stamminger et al. (1990). Total cellular RNA (40 µg) was hybridized with 2.5×10^5 counts/minute of 32P-labelled riboprobe at 37°C and subjected to digestion by RNAase A (6 µg/ml) and RNAase T1 (12 U/ml). Fragments were separated on a 6% denaturing polyacrylamide gel and visualized after exposure to X-ray film (Kodak).

**Tissue in situ hybridization**

Human fetuses of 9-12 weeks gestational age were obtained after therapeutic abortions using prostaglandin pessaries or suction termination. The tissue was fixed in 4% formalin in PBS for 12-24 hours, placed in 0.5 M sucrose for 6-12 hours to prevent freezing artifacts and then snap frozen in liquid nitrogen. Processing of the tissues was as described earlier (Mundlos et al., 1991). In brief, cross-sections of the developing kidney were cut on a Leitz cryostat and mounted on triethoxyethanolamine-treated microscopic slides. After prehybridization, 30-50 µl of the labelled RNA-probes (10^5 counts/minute/µl) in a buffer containing 50% formamide, 10% dextran sulfate, 0.3 M NaCl, 10 mM Tris pH 7, 10 mM NaHPO4 pH 5, 5.5 mM EDTA, 0.02% Denhardt’s, and t-RNA 0.5 mg/ml were applied to the slides. Hybridization was carried out at 47°C overnight. Slides were washed at 50°C for 2 hours in 50% formamide/SSC, followed by incubation with 20 µg/ml RNAase (Sigma, Munich, FRG.) for 15 minutes and washed with 1×SSC three times for 30 minutes at room temperature. After dehydration, slides were dipped in Kodak NTB2 photoemulsion diluted 1:1 with water and air dried. Autoradiography was performed at 4°C in a dry chamber. Exposure time varied from 10 to 20 days. Sense probes served as negative controls.

The following cDNA-fragments were used as probes: the PvuII/StuI fragment of human PAX8 and a 550 bp cDNA-fragment of human PAX2. These two clones, which correspond to the sequences outside of the paired box, show little homology to each
and to other Pax clones. To study the expression pattern of the WT1 gene, a 610 bp HincII/EcoRI-fragment was used. The probe is derived from the 3′ region of the WT33 cDNA (Call et al., 1990) outside the zinc finger domain to avoid cross-hybridization with corresponding sequences in other genes. The cDNAs were subcloned in Bluescript transcription vectors and transcribed with high efficiency from the T3 or T7 promoters in the presence of 35S-UTP, 800 µCi/mmol (Amersham, Poole, UK), to generate labelled sense and antisense RNA probes.

Mapping studies
The human-mouse somatic cell hybrid lines used for chromosomal localization of the human PAX8 gene have been isolated and described previously (Zabel et al., 1985). The mapping panel consisted of these somatic cell hybrids derived from cells from several unrelated individuals and different rodent strains. The hybrids were characterized by PCR amplification of a 329 bp fragment from the PAX8 cDNA (nt 1622-1950). The PCR reaction was performed using AmpliTaq (Perkin Elmer Cetus) as described by the manufacturers (94 °C, 10 seconds; 60 °C, 10 seconds; 72 °C, 10 seconds; 35 cycles). Oligonucleotides for PCR screening were primer 1 = 5′-AAGTCCAGGATGCCGACA-3′ and primer 2 = 5′-GAGGGAAATGCTTATGGTC-3′.

Results
Isolation and structure of the human PAX8 cDNAs
Using the mouse Pax8 cDNA c2A, which includes most of the paired box sequence (Plachov et al., 1990), as a probe, we isolated twenty clones from a human adult kidney cortex cDNA library (Bell et al., 1986) in a low-stringency screen. All eleven clones characterized so far represented the human PAX8 gene. Expression of another paired box gene, Pax2, that is highly homologous to the Pax8 gene, can be detected in mouse adult kidney upon RNAase protection analysis (H. Fickenscher, unpublished data). However, in our search, we were not able to identify any human Pax2 cDNAs. Four representative PAX8 cDNA clones are depicted in Fig. 1. Two of them, H8 and H29, contain the entire coding sequence of Pax8. All the clones contain at their 3′ ends the polyadenylation sequence AATAAA (Proudfoot and Brownlee, 1976), followed by a long stretch of adenosine residues, twelve nucleotides downstream, in the case of clones H2 and H12 (Fig. 2). The region around the polyadenylation sequence is highly conserved in mouse and human Pax8 genes (data not shown) and probably represents the authentic 3′ end of the Pax8 mRNA. One PAX8 clone, H14, contains five additional nucleotides, ATCTC, in front of the poly (A) tail, which are absent in the other Pax8 clones (data not shown).

Two cDNA plasmids, H12 and H19, display a deletion of a 189 bp fragment between positions 1059 and 1247 (Figs 1 and 2). This deletion probably represents a differential splicing event in the PAX8 primary transcript, rather than being a cloning artifact, because the clones H12 and H19 were isolated independently. Furthermore, we identified one mouse Pax8 cDNA clone, Z8, presenting the same deletion (data not shown).

A comparison of the mouse and human Pax8 coding sequences reveals two small deletions in the human PAX8 gene that do not change the reading frame. Three nucleotides, AAC, at the murine positions 738-740 and 18 nucleotides at the murine positions 1173-1190 (Plachov et al., 1990) are absent from the human transcript. Apart from these deletions the coding regions from both species show a high degree of homology: of the 1350 compared nucleotides 91.3% are conserved at their corresponding positions. The mouse and human Pax8 cDNAs are similar in size (2528 and 2712 nt, respectively).

The open reading frame of the human PAX8 gene
Two PAX8 cDNA clones contain a large open reading frame predicting a human PAX8 protein of 450 amino acids (aa) (Fig. 2). The first AUG at position 161 is not surrounded by any Kozak’s consensus sequence (Kozak, 1986). However, 16 nucleotides upstream and 11 nucleotides downstream of this AUG are identical in mouse and man. This sequence conservation around the first AUG suggests similar mechanisms of Pax8 translation in both species. The absence of strong homology between the murine and human sequences 5′ upstream of the human position 140 indicates that the AUG at position 161 is the start codon of the open reading frame. This open reading frame contains the 128 aa paired domain implicated in sequence-specific DNA binding (Zannini et al., 1992) at its amino terminus and, 41 aa further downstream, an octapeptide conserved in several other genes containing a paired box, YSINGLLG, the function of which is still unknown (Burri et al., 1989; Plachov et al., 1990). There are four

Fig. 1. The structure of human PAX8 cDNA. The coding region is indicated by an open box. The paired box domain and the adjacent octapeptide are indicated by a dotted and filled box, respectively. The PAX8 cDNA clones are shown underneath as black bars. The presence of the poly(A) tails and the deletion of a sequence encoding a serine-rich region are indicated by corresponding cDNA clones.
leucine residues at aa positions 264, 271, 278, 285 that could contribute to a 'zipper' structure (Landschulz et al., 1988). Leucines are conserved at these positions in mouse Pax8 and Pax2 conceptual proteins.

The deletion of 189 bp in plasmids H12 and H19 would result in loss of a 63 aa domain. This region is rich in hydroxylated amino acids, there are 14 serine and 3 threonine residues. Being a multiple of 3, the removal of this 189 bp region retains the reading frame. One clone, H26, shows a deviation in the nucleotide sequence of this region

---

**Fig. 2.** Sequence of human *PAX8*. The paired box domain and the polyadenylation signal are in bold type. The octapeptide is underlined. The borders of a deletion in H12 and H19 cDNA clones are indicated by triangles.
Analysis of human PAX8 gene

(C at position 1145 compared to T in other clones) leading to an amino acid change from Phe to Leu. This difference could either represent a polymorphism in the PAX8 gene or, alternatively, be a consequence of a mistake by reverse transcriptase.

Comparison of the conceptual murine and human Pax8 proteins

A comparison of the predicted murine and human Pax8 amino acid sequences reveals a very high degree of conservation (Fig. 3). Both proteins are identical from their amino terminus down to the octapeptide. There is a deletion of an Asn residue in the human PAX8 gene corresponding to the mouse aa position 195. Furthermore, there is a deletion of six amino acid residues in the human PAX8 gene corresponding to the mouse Pax8 aa positions 340-346 (Fig. 3). We do not yet know whether this last deletion is due to the differential splicing or a genomic deletion. Of the remaining 450 aa of the human PAX8 protein, there are only 10 aa substitutions, three of them conservative, as compared to the corresponding positions of the mouse Pax8 protein (Fig. 3).

The chromosomal localization

The PAX8 probe (H29) hybridized to a 2 kb fragment in HindIII digested human DNA, not present in mouse DNA. Analysis of 12 hybrid cell lines demonstrated consistent PAX8 gene segregation only with human chromosome 2 (Table 1).

Expression of the PAX8 gene in human tissues and Wilms' tumors

In order to investigate the distribution of the differentially spliced human PAX8 transcripts in tissues known to express Pax8 in mouse (Plachov et al., 1990), we used an RNAase protection analysis. The 505 bp Stul fragment from the H26
clone (positions 899-1404), designated H26Stu4, was chosen as a probe. Its sequence is identical to that of H8 and H29 in this region. This fragment encompasses a region in the human *PAX8* gene where the differential splicing occurs (Fig. 1). The sequencing of the cDNAs depicted in Fig. 1 predicts that the antisense riboprobe corresponding to the H26Stu4 fragment would protect a 505 base fragment corresponding to the full *PAX8* sequence and two additional fragments each 157 bases long arising from hybridization to an mRNA corresponding to the sequence represented by the H12 and H19 clones. Indeed, the RNAse protection analysis shows the protected fragments of the expected lengths with total RNA from human kidney, and five different Wilms’ tumors. No protection is observed with RNA from human HeLa cells as negative control (fragments fl and B, Fig. 4). Interestingly, the analysis reveals an additional specifically protected fragment of approximately 260 bases that cannot be accounted for by the presented sequencing data (fragment A, Fig. 4). These data might indicate the presence of a third form of *PAX8* transcripts generated by differential splicing in this region of the *PAX8* gene. We are currently investigating this possibility. Furthermore, this experiment shows a particularly strong *PAX8* expression in Wilms’ tumors, when compared with the transcript level from normal adult kidney.

<table>
<thead>
<tr>
<th>Human chromosome</th>
<th>(+/+)</th>
<th>(+/-)</th>
<th>(-/-)</th>
<th>(-/+)</th>
<th>% Discordancy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>6</td>
<td>36</td>
</tr>
<tr>
<td>2</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>6</td>
<td>36</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>43</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>6</td>
<td>29</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>6</td>
<td>29</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>6</td>
<td>36</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>43</td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>6</td>
<td>43</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>5</td>
<td>29</td>
</tr>
<tr>
<td>11</td>
<td>2</td>
<td>5</td>
<td>0</td>
<td>7</td>
<td>36</td>
</tr>
<tr>
<td>12</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>50</td>
</tr>
<tr>
<td>13</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>6</td>
<td>43</td>
</tr>
<tr>
<td>14</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>6</td>
<td>29</td>
</tr>
<tr>
<td>15</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>43</td>
</tr>
<tr>
<td>16</td>
<td>1</td>
<td>6</td>
<td>2</td>
<td>5</td>
<td>57</td>
</tr>
<tr>
<td>17</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>64</td>
</tr>
<tr>
<td>18</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>36</td>
</tr>
<tr>
<td>19</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>6</td>
<td>43</td>
</tr>
<tr>
<td>20</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>50</td>
</tr>
<tr>
<td>21</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>64</td>
</tr>
<tr>
<td>22</td>
<td>1</td>
<td>6</td>
<td>2</td>
<td>5</td>
<td>57</td>
</tr>
<tr>
<td>X</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>4</td>
<td>50</td>
</tr>
</tbody>
</table>

Tabulation of the presence or absence of the human *PAX8* gene in cell hybrids. The detection of *PAX8* is correlated with the presence of each human chromosome in the somatic cell hybrids. Discordancy represents the presence of the gene in the absence of the chromosome (+/-) or the absence of the gene despite the presence of the chromosome (-/+); the sum of these numbers divided by total hybrids examined (×100) represents percentage discordancy.

Fig. 4. *PAX8* expression in human kidney and Wilms’ tumors. Total cellular RNA from HeLa cells, human kidney or from Wilms’ tumors (WT) were analyzed by RNAase protection experiments. A riboprobe of high specific activity was prepared from the cDNA plasmid H26Stu4 after linearizing with restriction enzyme HindIII (probe H26Stu4). After RNAase digestion, specific protected fragments were observed from kidney and Wilms’ tumor RNA. The largest band (fl; 505 bases) corresponds to the full-length protected internal fragment, while two predominant smaller fragments (A, B: 260 and 157 bases, respectively) seem to result from alternate spliced messages. 40 µg of total cellular RNA were used from each tissue, except from WT2, where only 10 µg of RNA were used, supplemented with 30 µg of yeast tRNA. Two different exposures are shown, the upper one represents a four-day exposure, while the lower panel shows activities after 12 hours. On the right a DNA marker and the undigested riboprobe are shown.
were visualized from human fetal thyroid tissue using the H26Stu4 riboprobe. RNA from Wilms’ tumor number 4 was used as a strong positive, kidney RNA as a weak positive, and HeLa RNA as a negative control. After RNAase digestion, the same specific fragments appeared as in Fig. 4. RNAase protection analysis with this probe showed that only one fragment corresponding to the full-length H26PS3 probe was present at low levels. Parallel to the folding of the induced mesenchyme to comma-shaped and S-shaped bodies, PAX2 expression declined and PAX8 transcripts were detected at high levels (Fig. 6A,B). PAX8 expression was found in all cells of the S-shaped bodies. The ingrowth of blood vessels and the formation of a podocyte layer ultimately leads to the development of the glomerulus. During these developmental stages, neither PAX2 nor PAX8 expression was demonstrable. No expression was found in the collecting tubules or any other structures. Thus, during human kidney development, PAX2 expression is found mainly in induced, condensing mesenchyme, whereas PAX8 expression is present mainly in the intermediate stages of development corresponding to the comma-shaped and S-shaped bodies.

In contrast, expression of the Wilms’ tumor gene WT1 is present throughout these stages of development (Fig. 7A,B). A low level of expression was present in condensed, induced mesenchyme but not in the inducing ureteric bud. Parallel to the formation of S-bodies, WT1 expression increases. Within the S-shaped structures, the WT1-mRNA positive cells are restricted to the precursors of the podocyte layer. This pattern of expression is maintained in the more mature glomerulus, where a ring of positive cells (podocytes) surrounds the glomerulus.

Only one form of Pax8 transcripts can be detected in mouse kidney

Evidence that differential splicing occurs in a serine-rich region of the human PAX8 gene prompted us to investigate this region of the mouse Pax8 transcripts by RNAase protection analysis. The probes designated c27B/x22/1 (a 297 bp PstI/DdeI fragment of the mouse Pax8 cDNA, positions 855 and 1152, respectively; Plachov et al., 1990) and c27BDde3 (a 107 bp PstI/DdeI fragment of the mouse Pax8 cDNA, positions 1152 and 1259, respectively; Plachov et al., 1990) were used. If a differential splicing resulting in the removal of a serine-rich region occurred in the mouse Pax8 gene, then the c27B/x22/1 probe would protect two fragments corresponding to the two types of transcripts. A 297 base protected fragment would correspond to the full-length probe, indicating the transcripts coding for the serine-rich region, while a 201 base protected fragment would indicate differential splicing of this region. Only one fragment was protected specifically in the RNA samples containing PAX8 transcripts. This is demonstrated in Fig. 6C,D. The strongly positive structures in the cortical layers have a pairwise appearance caused by the separation through the ureteric bud. The ureteric bud itself showed significantly less expression when compared to the condensed mesenchyme. At this stage, PAX8 transcripts were present at low levels. Parallel to the folding of the induced mesenchyme to comma-shaped and S-shaped bodies, PAX2 expression declined and PAX8 transcripts were detected at high levels (Fig. 6A,B). PAX8 expression was found in all cells of the S-shaped bodies. The ingrowth of blood vessels and the formation of a podocyte layer ultimately leads to the development of the glomerulus. During these developmental stages, neither PAX2 nor PAX8 expression was demonstrable. No expression was found in the collecting tubules or any other structures. Thus, during human kidney development, PAX2 expression is found mainly in induced, condensing mesenchyme, whereas PAX8 expression is present mainly in the intermediate stages of development corresponding to the comma-shaped and S-shaped bodies.

In contrast, expression of the Wilms’ tumor gene WT1 is present throughout these stages of development (Fig. 7A,B). A low level of expression was present in condensed, induced mesenchyme but not in the inducing ureteric bud. Parallel to the formation of S-bodies, WT1 expression increases. Within the S-shaped structures, the WT1-mRNA positive cells are restricted to the precursors of the podocyte layer. This pattern of expression is maintained in the more mature glomerulus, where a ring of positive cells (podocytes) surrounds the glomerulus.

Only one form of Pax8 transcripts can be detected in mouse kidney

Evidence that differential splicing occurs in a serine-rich region of the human PAX8 gene prompted us to investigate this region of the mouse Pax8 transcripts by RNAase protection analysis. The probes designated c27B/x22/1 (a 297 bp PstI/DdeI fragment of the mouse Pax8 cDNA, positions 855 and 1152, respectively; Plachov et al., 1990) and c27BDde3 (a 107 bp PstI/DdeI fragment of the mouse Pax8 cDNA, positions 1152 and 1259, respectively; Plachov et al., 1990) were used. If a differential splicing resulting in the removal of a serine-rich region occurred in the mouse Pax8 gene, then the c27B/x22/1 probe would protect two fragments corresponding to the two types of transcripts. A 297 base protected fragment would correspond to the full-length probe, indicating the transcripts coding for the serine-rich region, while a 201 base protected fragment would indicate differential splicing of this region. Only one fragment was protected specifically in the RNA samples containing PAX8 transcripts. This is demonstrated in Fig. 6C,D. The strongly positive structures in the cortical layers have a pairwise appearance caused by the separation through the ureteric bud. The ureteric bud itself showed significantly less expression when compared to the condensed mesenchyme. At this stage, PAX8 transcripts were present at low levels. Parallel to the folding of the induced mesenchyme to comma-shaped and S-shaped bodies, PAX2 expression declined and PAX8 transcripts were detected at high levels (Fig. 6A,B). PAX8 expression was found in all cells of the S-shaped bodies. The ingrowth of blood vessels and the formation of a podocyte layer ultimately leads to the development of the glomerulus. During these developmental stages, neither PAX2 nor PAX8 expression was demonstrable. No expression was found in the collecting tubules or any other structures. Thus, during human kidney development, PAX2 expression is found mainly in induced, condensing mesenchyme, whereas PAX8 expression is present mainly in the intermediate stages of development corresponding to the comma-shaped and S-shaped bodies.

In contrast, expression of the Wilms’ tumor gene WT1 is present throughout these stages of development (Fig. 7A,B). A low level of expression was present in condensed, induced mesenchyme but not in the inducing ureteric bud. Parallel to the formation of S-bodies, WT1 expression increases. Within the S-shaped structures, the WT1-mRNA positive cells are restricted to the precursors of the podocyte layer. This pattern of expression is maintained in the more mature glomerulus, where a ring of positive cells (podocytes) surrounds the glomerulus.

Only one form of Pax8 transcripts can be detected in mouse kidney

Evidence that differential splicing occurs in a serine-rich region of the human PAX8 gene prompted us to investigate this region of the mouse Pax8 transcripts by RNAase protection analysis. The probes designated c27B/x22/1 (a 297 bp PstI/DdeI fragment of the mouse Pax8 cDNA, positions 855 and 1152, respectively; Plachov et al., 1990) and c27BDde3 (a 107 bp PstI/DdeI fragment of the mouse Pax8 cDNA, positions 1152 and 1259, respectively; Plachov et al., 1990) were used. If a differential splicing resulting in the removal of a serine-rich region occurred in the mouse Pax8 gene, then the c27B/x22/1 probe would protect two fragments corresponding to the two types of transcripts. A 297 base protected fragment would correspond to the full-length probe, indicating the transcripts coding for the serine-rich region, while a 201 base protected fragment would indicate differential splicing of this region. Only one fragment was protected specifically in the RNA samples containing PAX8 transcripts. This is demonstrated in Fig. 6C,D. The strongly positive structures in the cortical layers have a pairwise appearance caused by the separation through the ureteric bud. The ureteric bud itself showed significantly less expression when compared to the condensed mesenchyme. At this stage, PAX8 transcripts were present at low levels. Parallel to the folding of the induced mesenchyme to comma-shaped and S-shaped bodies, PAX2 expression declined and PAX8 transcripts were detected at high levels (Fig. 6A,B). PAX8 expression was found in all cells of the S-shaped bodies. The ingrowth of blood vessels and the formation of a podocyte layer ultimately leads to the development of the glomerulus. During these developmental stages, neither PAX2 nor PAX8 expression was demonstrable. No expression was found in the collecting tubules or any other structures. Thus, during human kidney development, PAX2 expression is found mainly in induced, condensing mesenchyme, whereas PAX8 expression is present mainly in the intermediate stages of development corresponding to the comma-shaped and S-shaped bodies.

In contrast, expression of the Wilms’ tumor gene WT1 is present throughout these stages of development (Fig. 7A,B). A low level of expression was present in condensed, induced mesenchyme but not in the inducing ureteric bud. Parallel to the formation of S-bodies, WT1 expression increases. Within the S-shaped structures, the WT1-mRNA positive cells are restricted to the precursors of the podocyte layer. This pattern of expression is maintained in the more mature glomerulus, where a ring of positive cells (podocytes) surrounds the glomerulus.

Only one form of Pax8 transcripts can be detected in mouse kidney

Evidence that differential splicing occurs in a serine-rich region of the human PAX8 gene prompted us to investigate this region of the mouse Pax8 transcripts by RNAase protection analysis. The probes designated c27B/x22/1 (a 297 bp PstI/DdeI fragment of the mouse Pax8 cDNA, positions 855 and 1152, respectively; Plachov et al., 1990) and c27BDde3 (a 107 bp PstI/DdeI fragment of the mouse Pax8 cDNA, positions 1152 and 1259, respectively; Plachov et al., 1990) were used. If a differential splicing resulting in the removal of a serine-rich region occurred in the mouse Pax8 gene, then the c27B/x22/1 probe would protect two fragments corresponding to the two types of transcripts. A 297 base protected fragment would correspond to the full-length probe, indicating the transcripts coding for the serine-rich region, while a 201 base protected fragment would indicate differential splicing of this region. Only one fragment was protected specifically in the RNA samples containing PAX8 transcripts. This is demonstrated in Fig. 6C,D. The strongly positive structures in the cortical layers have a pairwise appearance caused by the separation through the ureteric bud. The ureteric bud itself showed significantly less expression when compared to the condensed mesenchyme. At this stage, PAX8 transcripts were present at low levels. Parallel to the folding of the induced mesenchyme to comma-shaped and S-shaped bodies, PAX2 expression declined and PAX8 transcripts were detected at high levels (Fig. 6A,B). PAX8 expression was found in all cells of the S-shaped bodies. The ingrowth of blood vessels and the formation of a podocyte layer ultimately leads to the development of the glomerulus. During these developmental stages, neither PAX2 nor PAX8 expression was demonstrable. No expression was found in the collecting tubules or any other structures. Thus, during human kidney development, PAX2 expression is found mainly in induced, condensing mesenchyme, whereas PAX8 expression is present mainly in the intermediate stages of development corresponding to the comma-shaped and S-shaped bodies.

In contrast, expression of the Wilms’ tumor gene WT1 is present throughout these stages of development (Fig. 7A,B). A low level of expression was present in condensed, induced mesenchyme but not in the inducing ureteric bud. Parallel to the formation of S-bodies, WT1 expression increases. Within the S-shaped structures, the WT1-mRNA positive cells are restricted to the precursors of the podocyte layer. This pattern of expression is maintained in the more mature glomerulus, where a ring of positive cells (podocytes) surrounds the glomerulus.
protected fragment in the RNA sample from mouse kidney corresponding to the 297 bp of the full-length c27B/xx22/1 probe can be detected in our analysis (Fig. 8). These data show that only Pax8 transcripts including the serine-rich region can be detected in the RNA from mouse kidney. The presence of the sequence coding for the serine-rich region

Fig. 6. Gene expression of PAX8 (B) and PAX2 (D) in section through a 10-week fetal kidney with bright-field images given in (A) and (C). Note expression in condensed mesenchyme (PAX2, PAX8) and S-bodies (PAX8) but not in glomeruli or collecting tubules. G, glomerulus; C, condensed mesenchyme; S, S-shaped bodies; magnification ×70 (A,B) and ×52 (C,D).

Fig. 7. Gene expression of WT1 in a section through a 10-week fetal kidney; (A) bright field, (B) dark field images. Note expression in condensed mesenchyme, S-bodies and in the outer cell layer of glomeruli. G, glomerulus; magnification ×70.
Analysis of human PAX8 gene

Discussion

In this report, we describe the isolation and sequence of cDNA clones representing the human PAX8 gene. A comparison of the predicted human and mouse Pax8 proteins reveals a very high degree of conservation: there are only 10 amino acid substitutions, three of them conserved, out of 450 amino acids. Such evolutionary conservation may reflect an important function for Pax8 during mammalian development. However, the longest open reading frame, represented by human cDNAs H8 and H29 is seven amino acids shorter than its mouse counterpart. This seven aa truncation results from deletions at two different positions. First, the human PAX8 protein lacks an asparagine that is present at position 195 of the mouse Pax8 protein (Plachov et al., 1990). This gap is probably caused by a deletion of three nucleotides at the corresponding position of the human PAX8 gene because this region of the protein is encoded by a distinct exon in the mouse gene. Second, the PAX8 protein lacks six amino acids corresponding to the positions 340-345 aa of the mouse protein. Both proteins are identical in their paired domains. This fact suggests that DNA sequences that bind the Pax8 protein in the promoters of putative Pax8 target genes are conserved in mouse and man. Two thyroid-specific candidate target genes, thyroglobulin and thyroperoxidase, have been proposed (Zannini et al., 1992).

It has been shown that the murine Pax8 protein can function as a transcriptional activator (Zannini et al., 1992). The serine-rich region encompassing positions 302-369 aa of the mouse and 301-362 aa of the human Pax8 protein could be a candidate for an activation region (Ptashne, 1988). It has been shown in functional assays that CTF/NF-1 (Mermod et al., 1989), bicoid (Struhl et al., 1989), GHF-1/Pit-1 (Theill et al., 1989), OTF-2/oct-2 (Gerster et al., 1990; Tanaka and Herr, 1990) and LFB1 (Nicosia et al., 1990) contain a serine-rich activation region. Also a paired type homeobox gene, S8, contains such a region (Opstelten et al., 1991). The serine-rich region of the Pax8 protein might be a target for phosphorylation, which plays a crucial functional role for many transcription factors (Lee et al., 1990; Sheng et al., 1991; Ghosh and Baltimore, 1990; Kapiloff et al., 1991; Binetrui et al., 1991; Smeal et al., 1991; Pulverer et al., 1991). Currently we are investigating phosphorylation status of the Pax8 protein. The function of this serine-rich region in the Pax8 protein remains to be determined. Interestingly, this domain is removed in two independent PAX8 cDNA clones presumably due to differential splicing. Indeed, both transcript variants can be found by RNase protection experiments in human thyroid, kidneys and five investigated Wilms’ tumors. A preliminary characterization of the genomic organisation of PAX8 indicates that the boundaries of the spliced region correspond to exon acceptor/donor sites (not shown). Differential splicing producing different forms of a transcription factor is a well-documented event (e.g. a homeotic gene Ultrabithorax in Drosophila; Lopez and Hogness, 1991). Interestingly, a protein form with a truncated serine-rich domain arises due to a differential splicing in the Oct2 gene (Hatzopoulos et al., 1990). In some cases, differential splicing produces a dominant negative form of a transcription factor (Nakabeppu and Nathans, 1991; Mumberg et al., 1991; Roman et al., 1991).

Similarities between murine Pax8 and Pax2 genes have been observed previously (Plachov et al., 1990). Comparison of the published conceptual amino acid sequences of

Fig. 8. Murine Pax8 transcripts. Using mouse cDNA plasmids c27B/xx22/1 and c27B/DdeI/3, riboprobes were synthesized and hybridized to total cellular RNA from mouse kidney, brain or liver. In both cases only single protected fragments could be found. Their sizes suggest the absence of alternate splicing for the mouse Pax8 gene in contrast to human PAX8.
the murine Pax8 and Pax2 proteins indicates that these proteins are highly homologous not only in their paired domains but also in their downstream sequences (Fig. 9). However, the homology is interrupted in two regions. Amino acid residues between positions 302 and 369 of the Pax8 protein, encompassing a serine-rich region, are absent in Pax2. Interestingly, exactly this region is altered by alternative splicing in the human Pax2 protein in structure and possibly in function. It would be interesting to investigate the amino acid sequence of the human PAX2 protein in this domain. Conversely, amino acid residues between positions 203 and 229 of the Pax2 protein are present in Pax8. Of two described Pax2 cDNA clones, one shows a deletion of 23 aa in the predicted protein that coincides in size with this insertion and is shifted by two amino acids only (Dressler et al., 1990). The genes on 11p possibly represent ‘tumor suppressor’ genes (Weissman et al., 1987; Mundlos et al., 1991). Being putative transcription factors, PAX2 and PAX8 may be involved in the transcriptional regulation of these matrix genes. One important objective behind the isolation of the human PAX8 gene was to investigate its function in Wilms’ tumors, embryonal kidney tumors (Mierau et al., 1987). At least three genomic regions have been implicated in the genesis of Wilms’ tumors (van Heyningen and Hastie, 1992). A deleted region on chromosome 11p, linked to 11p, has been implicated by genetic linkage analysis (Roman et al., 1991). In this respect, it would be interesting to determine whether Pax8 and Pax2 proteins are capable of binding to common DNA targets modulating transcription from downstream genes in a manner similar to the WT1 (Wilms’ tumor gene 1) and EGR-1 proteins (Madden et al., 1991). The targeted inactivation of both Pax8 and Pax2 genes in transgenic mice would clarify these questions.

During the process of induction, mesenchymal condensations and the formation of S-bodies, profound changes in the expression of extracellular matrix genes take place. Induced cells no longer produce collagens type I and III but switch to collagen IV, osteonectin, laminin, fibronectin and proteoglycans (Ekblom et al., 1981; Mundlos et al., 1992; Laurie et al., 1989). Being putative transcription factors, PAX2 and PAX8 may be involved in the transcriptional regulation of these matrix genes.
et al., 1990; Gessler et al., 1990) and mutations were found in some Wilms’ tumors (Haber et al., 1990; Pelletier et al. 1991a,b). In agreement with these findings it has been shown that big DNA deletions at the 11p13 locus are rare in Wilms’ tumors (Royer-Pokora et al., 1991). It has been suggested that the WT1 gene would not be involved in the familial cases of Wilms’ tumors due to infertility of the carriers of mutated WT1 (Pritchard-Jones et al., 1990; van Heyningen et al., 1990). As demonstrated in Fig. 7, WT1 expression in the human metanephros was present throughout the different stages of glomeruli maturation. Low levels of expression were found in condensed mesenchyme followed by strong expression in the cells of the podocyte layer of S-bodies and glomeruli. These results are in agreement with the pattern of expression published earlier (Pritchard-Jones et al., 1990). Thus, WT1 expression is comparable to PAX2 and PAX8 in the early stages of development, but persists in the glomeruli where PAX2 and PAX8 expression is not present anymore (Table 2).

Because Wilms’ tumors are believed to arise as a consequence of the failure of the metanephrogenic mesenchyme to differentiate, any gene potentially involved in controlling mesenchymal differentiation is also a potential ‘Wilms’ tumor gene’. Somatic mutation of both alleles of such a developmental gene could result in failure of the metanephrogenic mesenchyme to differentiate, the possible result being a Wilms’ tumor. In this sense, the PAX8 gene could potentially be involved in the genesis of at least some Wilms’ tumors. As a first step in defining possible PAX8 functions in Wilms’ tumors, we have investigated its expression in five different tumors by RNAaase protection. PAX8 is strongly expressed in all five tumors as compared to its expression in adult human kidneys. This fact is probably a reflection of an embryonic origin of the Wilms’ tumor. The pattern of differential splicing in the PAX8 gene described above is similar in all investigated tumors. We are currently characterizing the genomic organization of PAX8 to perform a mutation analysis of DNA from tumor and normal tissues of patients with Wilms’ tumors.

It is interesting to note that a mouse mutant, Dickie’s Small-eye (Sey<sup>Dey</sup>), heterozygously deleted for one allele of the WT1 gene does not develop nephroblastomas as would be expected by analogy with the human WAGR syndrome (Buckler et al., 1991; Glaser et al., 1990). This suggests that there may be differences in murine and human kidney development that are relevant to the initiation of nephroblastoma development. In this respect, it is noteworthy that differential splicing of the serine-rich region in murine Pax8 transcripts could not be detected by RNAaase protection although this splicing was easily detectable among human PAX8 transcripts. This fact could indicate that the ratio between two variants of Pax8 transcripts would be quite different in mouse and man, respectively. However, we did also manage to isolate one mouse cDNA clone with a removed serine-rich domain (data not shown).

Thyroid is another site of PAX8 expression. As in kidney, both types of PAX8 transcripts can be detected in human thyroid where their corresponding protein products could act in concert with another tissue-specific transcription factor, TTF-1 (Guazzi et al., 1990; Zannini et al., 1992). It has been reported that proto-oncogenes could be activated in thyroid tumors by gene rearrangement to thyroid-specific genes (Grieco et al., 1990). PAX8 would be a candidate to mediate such thyroid-specific expression. Furthermore, the reports of association of thyroid malignancies with kidney cancers (Gault et al., 1974) should be reevaluated in the light of our results.

Using a somatic cell hybrids system, we have mapped the human PAX8 gene to chromosome 2q. This finding is in agreement with the mapping of the murine Pax8 gene to the centromeric region of chromosome 2 (Plachov et al., 1990; Walther et al., 1991), in close vicinity to the surfet gene cluster (Surf), loci marked by defects like the Danforth’s short tail mutation (Sd), the HOX-5 gene cluster and the proto-oncogene Abl. Some of these genes have been identified as part of a region of synteny with the human chromosome 2q (von et al. 1989; Stubbs et al., 1990).

The present data make it possible to search for mutations in the human PAX8 gene. The expected phenotype of such a mutation would be similar to a developmental mutation of the mouse called Danforth’s short tail mutation (Sd; Dunn et al., 1940; Gluecksohn-Waelsch and Rota, 1962) which was proposed to be a candidate for murine Pax8 mutations (Plachov et al., 1990; Hastie, 1991). The Sd mutation is characterized by abnormalities in the axial skeleton and the reduction or absence of kidneys. A comparable pattern of human malformation can be found in subtypes of Klippel-Feil syndrome, which is characterized by vertebral fusion and/or hemivertebrae. More than 59% of patients show urogenital malformations, mostly unilateral or bilateral renal agenesis (Helmi and Pruzansky, 1980).

We thank G. Bell for the human kidney cDNA library and R. Bleackley for the human genomic library. We thank S. Naylor for providing hybrid cell DNAs and D. Housman for the WT 33 cDNA probe. We thank U. Deutsch for the initial Pax8 and Pax2 sequence comparison. We thank R. Di Lauro, P. van der Saag and A. Reeve for transmission of unpublished results. Many thanks go to Andreas Weber and Ralf Altschäffel for the excellent photographic work. The valuable suggestions of R. Di Lauro on the manuscript are gratefully acknowledged. We thank S. Naylor for providing hybrid cell DNAs. We thank D. Housman for providing the WT33 cDNA probe. We thank C. Abbott for providing sequences for PAX8 primers. Special thanks to F. Wendler, E. Gurtner, and K. Kratochwil. A.P. was supported by a fellowship from the Austrian Academy of Sciences. D.P. was supported by the EMBO long-term fellowship and the austrian FWF grant (P 8682 MOB). H.F. was supported by the Deutsche Forschungsgemeinschaft (Fi 424/1-1).

Table 2. Expression of PAX2, PAX8 and WT1 genes in human metanephros

<table>
<thead>
<tr>
<th>gene</th>
<th>condensed mesenchyme</th>
<th>S-bodies</th>
<th>glomeruli</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAX2</td>
<td>+++</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>PAX8</td>
<td>++</td>
<td>+++</td>
<td>−</td>
</tr>
<tr>
<td>WT1</td>
<td>+</td>
<td>+++</td>
<td>+++</td>
</tr>
</tbody>
</table>

References


chromosome and histogenesis of the renal tumors of childhood: an overview. Ultradept. Pathol. 11, 313-333.


