Cloning and Characterization of the Human UDP-glucuronosyltransferase 1A8, 1A9, and 1A10 Gene Promoters

DIFFERENTIAL REGULATION THROUGH AN INITIATOR-LIKE REGION*

Received for publication, May 28, 2003, and in revised form, July 2, 2003
Published, JBC Papers in Press, July 7, 2003, DOI 10.1074/jbc.M305565200

Philip A. Gregory‡, Dione A. Gardner-Stephen‡, Rikke H. Lewinsky, Kym N. Duncliffe, and Peter I. Mackenzie‡

From the Department of Clinical Pharmacology, Flinders University School of Medicine, Flinders Medical Centre, Bedford Park, South Australia 5042, Australia

The human UDP-glucuronosyltransferases, UGT1A8, 1A9, and 1A10, are closely related in sequence and have a major role in the elimination of lipophilic chemicals by glucuronidation. UGT1A8 and 1A10 are expressed exclusively in the gastrointestinal tract, whereas UGT1A9 is expressed mainly in the liver and kidneys. To determine the factors contributing to the extrahepatic expression of these UDP-glucuronosyltransferases, we have cloned and characterized the promoters of the UGT1A8, 1A9, and 1A10 genes and studied their regulation in the colon cell line, Caco2. Their transcription start sites were mapped, and a functional overlapping Sp1/initiator-like site was identified which strongly contributed to UGT1A8 and 1A10 promoter activity. The high promoter activity of UGT1A8 and 1A10 correlated with the binding of nuclear proteins (complex B) to this region. Two-bp differences in the corresponding site in the UGT1A9 promoter prevented the binding of complex B and reduced promoter activity. Although Sp1 was able to bind to the Sp1/initiator-like site, its binding was dispensable for promoter activity. However, the binding of Sp1 to a second Sp1 site 30 bp 5' to the Sp1/initiator-like site greatly enhanced the activity of the UGT1A8 and 1A10 promoters. These results provide evidence that the UGT1A8, 1A9, and 1A10 genes are differentially regulated through an initiator element in their 5'-flanking regions.

Glucuronidation represents one of the major pathways through which the body is able to deactivate and excrete a wide range of lipophilic compounds including drugs, environmental pollutants, dietary chemicals, and endogenous compounds such as bilirubin, steroids, and bile acids (1). Accommodation of such a structurally diverse range of chemicals is provided by the superfamily of UDP-glucuronosyltransferase (UGT) enzymes, of which there are 16 known functional enzymes in humans to date. UGTS are classified into two families, UGT1 and UGT2, based on their amino acid sequence similarity (2). The UGT2 family is subdivided further into UGT2A and UGT2B subfamilies, each encompassing genes encoded by six unique exons located on chromosome 4 (3, 4). In contrast, the UGT1A subfamily is encoded by a single gene locus on chromosome 2, recently shown to span >200 kb (5). Eight functional UGT1A enzymes have been shown to be expressed in humans (UGT1A1, 1A3, 1A4, 1A6, 1A7, 1A8, 1A9, and 1A10), and four pseudogenes (1A2, 1A11, 1A12, and 1A13) have been identified (6). Each UGT1A form contains a unique exon 1, which is joined to a shared set of exons 2–5 through RNA splicing. Characterization of the UGT1A locus has shown that the first exon of each individual form is organized in series upstream from exon 2 (5, 7). The first exons are separated from each other by 5'-flanking regions of 5–23 kb in length. Much accumulated evidence suggests that the 5'-flanking region of each UGT1A exon 1 contains promoter elements capable of regulating their expression. Analysis of the promoters of the human UGT1A1 and 1A6 genes showed that they were functional in liver and colon cells, respectively (8). Studies on the rat UGT1A6 and UGT1A7 genes have also shown that their promoter regions are capable of stimulating transcription in liver cells (9, 10). In addition, UGT1A1 genes have recently been demonstrated to be induced by the nuclear receptors constitutive active receptor (CAR), aryl hydrocarbon receptor (AhR), pregnane X receptor (PXR) and peroxisome proliferator-activated receptor (PPAR) through binding to their promoter regions (9, 11–14). The tissue specificity of UGT1A expression has been an area of intense study recently. Although the liver is considered to be the major site of glucuronidation, the importance of the gastrointestinal tract in the glucuronidation of ingested chemicals has been shown by the expression of multiple UGT1A enzymes in these tissues (15). The liver is known to express the UGT1A forms 1A1, 1A3, 1A4, 1A6, and 1A9 (16–19) but not the 1A7, 1A8, and 1A10 genes. These three isoforms are expressed exclusively in extrahepatic tissues, particularly in the tissues of the gastrointestinal tract (20–23). The functional UGT1A genes are divided into two clusters, the UGT1A7–10 cluster of genes being >70% similar in their first exon sequence and <60% similar to the other UGT1A genes (including the UGT1A3–5 cluster) (5). Unlike the UGT1A1–6 genes, which have interindividual variation in expression in the gastrointestinal tract, UGT1A7–10 appear to be expressed consistently in these tissues (24, 25). This unique pattern of expression of the UGT1A7–10 gene cluster suggests that coordinated, tissue-specific mechanisms of regulation may be important for their expression.

Both the human UGT1A1 and UGT1A6 genes utilize TATA

* This work was supported by grants from the Cancer Council of South Australia and the National Health and Medical Research Council of Australia. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ These authors contributed equally to this work.

© 2003 by The American Society for Biochemistry and Molecular Biology, Inc.

Published, JBC Papers in Press, July 7, 2003, DOI 10.1074/jbc.M305565200

Published, JBC Papers in Press, July 7, 2003, DOI 10.1074/jbc.M305565200


This paper is available on line at http://www.jbc.org
boxes to initiate transcription (12, 26). However, initiation of gene transcription by RNA polymerase II may also occur through initiator regions in the promoters of genes. TATA box-dependent mechanisms of transcription are well characterized, the first step involving the binding of transcription factor IID (TFIID) to AT-rich sequences generally located 25–30 bp upstream from the transcription start site, and subsequent formation of a preinitiation complex containing RNA polymerase II (27). In contrast, the mechanism of transcription initiation through initiator regions is less clearly understood. Initiator regions are found in both TATA-containing and TATA-less promoters and generally encompass the transcription start site (28). The initiator site is sufficient for direct transcription alone, but transcription is often enhanced through upstream TATA box or Sp1 binding sites (29–31). Although the sequence requirements of initiator sites vary, they adhere loosely to a Py N T/A Py Py consensus, where Py represents a pyrimidine (32). The relative strength of the initiator site appears to be enhanced by the presence of further pyrimidine residues surrounding the core consensus site (32, 33). A wide range of proteins has been demonstrated to bind to initiator-like sites including TFIID (34–36), RNA polymerase II (29), ying-yang 1 (YY1) (37, 38), upstream stimulating factor (USF) (39), YY1 (40), and a nearby overlapping Sp1/initiator-like site was identified which strongly contributed to UGT1A8 and 1A10 promoter activity. Initiation of transcription through TATA box-dependent mechanisms was not observed for these UGT genes.

**EXPERIMENTAL PROCEDURES**

**Isolation of the UGT1A8, 1A9, and 1A10 Genes and 5'-Flanking Regions**—The UGT1A8, 1A9, and 1A10 promoter regions were isolated from a human placenta lambda genomic library (Clontech) after several rounds of screening for their adjacent exons. Briefly, a 480-bp EcoRI/NcoI fragment was prepared from exon 2 of the UGT1A8 (corresponding to bases 56–576) and labeled with [32P]dATP by random priming using the DECAprime II DNA labeling kit (Ambion). Hybrid-N membranes (Amer sham Biosciences) from plaque lifts were denatured, neutralized, and UV cross-linked as in the manufacturer’s protocol (PT1010-1). Prehybridization was carried out in a solution of 5 × SSC, 5 × Denhardt’s and 0.5% SDS for 4 h at 42 °C before hybridization with 1 × 10⁶ cpm/μl probe overnight. Membranes were then washed serially twice in 2 × SSC, 0.1% SDS for 5 min at room temperature and twice in 0.1 × SSC, 0.1% SDS for 30 min at 65 °C followed by exposure to autoradiographic film with intensification at −70 °C overnight. Positive plaques were picked and screened by restriction digestion and sequencing for appropriate inserts from the UGT1A locus.

**Construction of Luciferase Reporter Constructs**—The proximal 1 kb of the UGT1A9 and 1A10 promoters were amplified from the lambda Clontech library isolated by following the manufacturer’s protocol. Resulting full-length RNA was ligated into the pGEM-T Easy plasmid using the QIAquick PCR purification kit (Qiagen) and cloned into the plasmid pBlunt vector (Invitrogen) according to the manufacturer’s protocol. Several clones were selected and sequenced to confirm the position of each transcription start site.

**Transcription Start Site Mapping**—Transcription start sites were mapped using the First Choice RNA ligase-mediated rapid amplification of cDNA ends (RACE) kit designed to amplify cDNA only from directionally into a 5-kb genomic region. The upstream Sp1 site was mutated by PCR using the Sp1rev primers. The proximal 5′ bp of the UGT1A8 and 1A10 promoters were amplified from the lambda Clontech library isolated by following the manufacturer’s protocol. Resulting full-length RNA was ligated into the pGEM-T Easy plasmid using the QIAquick PCR purification kit (Qiagen) and cloned into the plasmid pBlunt vector (Invitrogen) according to the manufacturer’s protocol. Several clones were selected and sequenced to confirm the position of each transcription start site.

**Cell Culture and Transfection**—Caco2 cells obtained from the American Type Culture Collection were cultured in Dulbecco’s modified Eagle’s medium supplemented with 10% fetal calf serum, 1 mM sodium pyruvate, 0.1 mM mixture of nonessential amino acids (all from Invitrogen), and 80 μg/ml gentamycin at 37 °C in 5% CO₂. Cells were plated...
Differential Regulation of the UGT1A8, 1A9, and 1A10 Genes

RESULTS

Characterization of the UGT1A8, 1A9, and 1A10 Promoters— The promoter regions of the UGT1A8, 1A9, and 1A10 genes were isolated from a human placenta genomic library by screening with a UGT1A8 exon 1 probe. Several clones were selected and identified as containing the UGT1A8, 1A9, or 1A10 exon 1 by sequencing. Comparison of the proximal ~1 kb of the UGT1A8, 1A9, and 1A10 promoters showed a high degree of sequence similarity (>75%) with several putative regulatory elements. Fig. 1 shows an alignment of the proximal 149 bases (UGT1A8 promoter relative to the ATG start codon) of each promoter, highlighting a conserved T-repeat region, two putative Sp1 binding sites (GC box), as well as several differences in nucleotide sequence. It has been suggested previously by Gong et al. (5) that the T-repeat region may function as a TATA box for the UGT1A8, 1A9, and 1A10 genes. To investigate this possibility, the transcription start sites of the genes were mapped using RNA ligase-mediated RACE to amplify the transcription start sites of the genes marked with a bold base. The putative Sp1 sites, initiator binding sites, and the T-region are boxed. Several base differences near the overlapping Sp1/initiator-like site have been highlighted (numbered 1–4).

![Fig. 1. Alignment of the UGT1A8, 1A9, and 1A10 proximal promoters and location of putative transcription factor binding sites.](image)

Approximately 150 bp of the proximal UGT1A8, 1A9, and 1A10 promoters were aligned with the methionine (ATG) initiation codon (shown in bold italics). Identical bases in the UGT1A9 and 1A10 genes compared with the UGT1A8 gene are indicated with a period. Nonidentical bases with the base change, and missing bases with a dash. The transcription start sites of each promoter, highlighting a conserved T-repeat region, two putative Sp1 sites, initiator-like sites, and the T-region are boxed. Several base differences near the overlapping Sp1/initiator-like site have been highlighted (numbered 1–4).

<table>
<thead>
<tr>
<th>Gene</th>
<th>Sp1</th>
<th>T-region</th>
</tr>
</thead>
<tbody>
<tr>
<td>UGT1A8</td>
<td>5'-TCATGCGAGTGGTAGCTGGATATTCATAGACAGTAGACCATCACCTCTTA-3'</td>
<td></td>
</tr>
<tr>
<td>UGT1A9</td>
<td>5'-TCATGCGAGTGGTAGCTGGATATTCATAGACAGTAGACCATCACCTCTTA-3'</td>
<td></td>
</tr>
<tr>
<td>UGT1A10</td>
<td>5'-TCATGCGAGTGGTAGCTGGATATTCATAGACAGTAGACCATCACCTCTTA-3'</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 1. Alignment of the UGT1A8, 1A9, and 1A10 proximal promoters and location of putative transcription factor binding sites. Approximately 150 bp of the proximal UGT1A8, 1A9, and 1A10 promoters were aligned with the methionine (ATG) initiation codon (shown in bold italics). Identical bases in the UGT1A9 and 1A10 genes compared with the UGT1A8 gene are indicated with a period. Nonidentical bases with the base change, and missing bases with a dash. The transcription start sites of each promoter, highlighting a conserved T-repeat region, two putative Sp1 sites, initiator-like sites, and the T-region are boxed. Several base differences near the overlapping Sp1/initiator-like site have been highlighted (numbered 1–4).

Nuclear Proteins Bind Differentially to the UGT1A8, 1A9, and 1A10 Initiator-like Region—Double-stranded oligonucleotide probes corresponding to nucleotides ~34 to +5 of the UGT1A8 promoter (Fig. 2B, bases referenced from the 5’-transcription start site) were synthesized and incubated with nuclear extracts from the colon-derived cell line Caco2 in EMSAs. Probes were also synthesized to the identical region in the UGT1A9 and 1A10 genes (Fig. 2B) and compared with the UGT1A8 probe. As shown in Fig. 2A, two major complexes (labeled A and B) were observed using the UGT1A8 wild-type (wt) probe, and several minor complexes also formed (lane 1). The UGT1A10wt probe appeared to form the same complexes as the UGT1A8wt probe (lane 10), whereas the UGT1A9wt probe only formed the top complex A and the minor complexes (lane 7). There are four differences in sequence between the UGT1A8 and 1A9 genes in this region (Fig. 1, labeled I–4) but only one difference at base 2 between UGT1A8 and UGT1A10. To determine which of these differences influences the protein binding patterns observed, probes were synthesized incorporating the nucleotide changes as shown in Fig. 2B (labeled m1–4). Mutation of base 1 in the UGT1A8 gene (UGT1A8m1) had no impact on the binding of nuclear proteins (lane 2). Mutation of base 2 (UGT1A8m2) appeared to diminish the formation of complex B slightly but did not affect complex A (lane 3). Mutation of base 3 (UGT1A8m3) completely abolished the formation of complex B and increased the formation of complex A, indicating this base is crucial for complex B formation (lane 4). Mutation of base 4 in the UGT1A8 gene (UGT1A8m4, lane 5) altered the intensity of complex A and the mobility of complex B. Mutation of both bases 3 and 4 in the UGT1A8 gene (UGT1A8m3/4, lane 6) resulted in the formation of complexes
Differential Regulation of the UGT1A8, 1A9, and 1A10 Genes

Initiator-like Region Influence Promoter Activity—

The higher promoter activity of the UGT1A8 and 1A10 genes correlates directly with their promoter activity. Sp1/initiator-like region is important for promoter activity of the UGT1A8, 1A9, and 1A10 genes and that differences in these nucleotides sharply decreased UGT1A8 promoter activity to 6-fold. These results indicate that the Sp1/initiator-like region is important for promoter activity of the UGT1A8, 1A9, and 1A10 genes and that differences in protein binding to this region between the UGT1A8/1A10 and UGT1A9 genes correlate directly with their promoter activity. The higher promoter activity of the UGT1A8 and 1A10 genes appears to be, in part, caused by the complex B formation on the Sp1/initiator-like region because the loss of this complex appears to be, in part, caused by the complex B formation on UGT1A9 genes (UGT1A9m3/4, lane 8). The results shown are a representative experiment of three individual experiments performed in triplicate (±S.D.).

2-fold relative luciferase activity over pGL3-basic. To determine whether the differences at nucleotides 3 and 4 in the Sp1/initiator-like region influenced promoter activity, mutant constructs were prepared containing the same mutations used in the EMSA probes (Fig. 2B). Mutation of base 3 in the context of the 1-kb UGT1A8 and 1A9 promoters decreased UGT1A8 promoter activity from 14- to 9-fold above pGL3-basic but slightly increased UGT1A9 promoter activity from 2- to 3-fold (Fig. 3). Similarly, mutation of base 4 decreased UGT1A8 promoter activity quite sharply to 5-fold and slightly increased UGT1A9 promoter activity to 3-fold. Mutation of both nucleotides sharply decreased UGT1A8 and UGT1A9 promoter activity to 5-fold and 2-fold and significantly increased UGT1A9 promoter activity to 6-fold. These results indicate that the Sp1/initiator-like region is important for promoter activity of the UGT1A8, 1A9, and 1A10 genes and that differences in protein binding to this region between the UGT1A8/1A10 and UGT1A9 genes correlate directly with their promoter activity. The higher promoter activity of the UGT1A8 and 1A10 genes appears to be, in part, caused by the complex B formation on the Sp1/initiator-like region because the loss of this complex (Fig. 2A, lanes 4–6) results in lower promoter activity. Similarly, the lower promoter activity of UGT1A9 appears to be, in part, caused by the lack of complex B formation, because when this complex is able to form (Fig. 2A, lane 10) the resulting promoter activity is higher. Therefore, we investigated the Sp1/initiator-like region further, examining the proteins present in complexes A and B.

Sp1 Is Present in Complex A of the UGT1A8, 1A9, and 1A10 Sp1/initiator-like Region—To determine whether Sp1 was able to bind to the UGT1A8, 1A9, and 1A10 Sp1/initiator-like region, supershift assays were performed with the UGT1A8, 1A9, and 1A10 wild-type probes (Fig. 2B) using a monoclonal Sp1-specific antibody. As shown in Fig. 4, addition of the Sp1 antibody significantly reduced the formation of complex A but not complex B with the UGT1A8wt and the UGT1A9wt probes (comparing lanes 1 and 5, and 6 and 7). The same reduction in

![Fig. 2. EMSA of the UGT1A8, 1A9, and 1A10 Sp1/initiator-like region.](image)

Fig. 2. EMSA of the UGT1A8, 1A9, and 1A10 Sp1/initiator-like region. A, double-stranded oligonucleotide probes (50,000 cpm) corresponding to wild-type and mutant UGT1A8, 1A9, and 1A10 Sp1/initiator-like regions (−34/+8 of UGT1A8) were incubated with 5 μg of Caco2 nuclear extracts and resolved on a 4% nondenaturing polyacrylamide gel. The two major complexes, A and B, are indicated with arrows, and the minor complexes are bracketed (labeled with an asterisk). B, sequences of the sense strand of the oligonucleotide probes used in A are shown. Identical bases are marked with a dash, and mutations at positions 1–4 are indicated by the appropriate nucleotide change. The overlapping Sp1/initiator-like site is boxed.

![Fig. 3. Expression of the UGT1A8, 1A9, and 1A10 wild-type and Sp1/initiator-like mutant proximal promoters in Caco2 cells.](image)
complex A formation was also observed with the UGT1A10wt probe (data not shown). Furthermore, the mobility of complex A was similar to that of the complex formed on the Sp1 consensus (Sp1 con) probe (lane 8). This indicates that complex A in the UGT1A8, 1A9, and 1A10 genes contains Sp1 and that complex B does not contain this protein. These observations were confirmed by competition studies using unlabeled UGT1A8wt, UGT1A8mut3/4, and Sp1 consensus probes. Addition of a 50-fold molar excess of the UGT1A8wt probe decreased the formation of both complexes A and B (lane 2), whereas addition of an equal amount of the UGT1A8mut3/4 or Sp1 con probe decreased the formation of complex A only (lanes 3 and 4). In each case, the minor complexes were not competed out and therefore represent nonspecific binding to the probe. These results implied that complex A formation with the UGT1A8wt and UGT1A8mut3/4 probes are identical and dependent on the presence of Sp1.

Sp1 and Complex B Bind to Overlapping Sites in the UGT1A8 Sp1/initiator-like Region—Having established that bases 3 and 4 are important for both complex A and B formation (Fig. 2) and that complex A contains Sp1 (Fig. 4), it appeared likely that Sp1 and complex B were binding to overlapping regions of the Sp1/initiator-like site and that complex B consisted of initiator-like proteins. To test this hypothesis, probes were synthesized containing mutations in the regions 5’-(UGT1A8 5’ mut) and 3’-(UGT1A8 3’ mut) to the Sp1/initiator site to prevent the binding of complex B and Sp1, respectively, in EMSA (Fig. 5). Mutation of 4 bases in the 3’-region inhibited Sp1 binding and increased complex B formation (lane 2) compared with the wild-type probe (lane 1). Similarly, mutation of 3 bases in the 5’-region inhibited complex B formation and increased binding of the Sp1 complex (lane 3). Taken together, these results suggest that Sp1 binding is dependent on bases 3’ of the Sp1/initiator-like site, complex B formation is dependent on bases within the 5’-region of the Sp1/initiator-like site, and that binding to their respective, overlapping sites is competitive. To determine the identity of the proteins composing complex B, competition studies and supershift assays were carried out using competitor oligonucleotides and antibodies to known initiator binding proteins. In particular, the UGT1A8/1A10 initiator-like site appeared to resemble closely YY1 and GABP sites found in the murine cytochrome c oxidase subunit Vb and cytochrome c oxidase subunit IV promoters (38, 50). However, despite extensive testing using competitor oligonucleotides and/or antibodies to these factors and the initiator-like binding proteins USF1, USF2, TFII-I, TafII250, c-Myc and E2F-1, the identity of the proteins in complex B could not be resolved (data not shown).

Binding of Sp1 to the Overlapping Sp1/initiator-like Site Is Not Essential for UGT1A8 Promoter Activity—Transfection of the UGT1A8, 1A9, and 1A10 promoters and their base 3 and 4 mutants showed that the binding of complex B was needed for maximal promoter activity (Fig. 3). To determine whether the binding of Sp1 to the Sp1/initiator-like site was essential for UGT1A8 promoter activity, an Sp1 mutant construct corresponding to the UGT1A8 3’-mutant probe in Fig. 5 was generated and tested in transfection assay. Compared with the wild-type UGT1A8 wild-type 1A8–1036/+45 construct (100%), the 1A8–1306/+45 Sp1-1036 construct had slightly elevated promoter activity (Fig. 6), which most likely reflects the increase in complex B binding to this mutant promoter (see Fig. 5). These data indicate that complex B formation, in the absence of Sp1 binding, is sufficient for maximal UGT1A8 promoter activity.
Fig. 6. Sp1 binding to the UGT1A8 Sp1/initiator-like region is not essential for promoter activity. An Sp1 mutant promoter construct with the same sequence as the UGT1A8 3'–mutant probe (Fig. 5B) was generated as described under “Experimental Procedures.” 0.5 μg of the wild-type and mutant promoter constructs was transfected into Caco2 cells with 0.025 μg of the pRL-null vector as an internal control of transfection efficiency. Cells were harvested 48 h post-transfection and assayed for firefly and Renilla luciferase activities. The relative luciferase activities of the mutant construct are expressed as a percentage change in promoter activity relative to the wild-type construct (set at a value of 100%). The results shown are a representative experiment of three individual experiments performed in triplicate (±S.D.).

The Activity of the UGT1A8 Promoter Is Not Dependent on the T-region but Is Enhanced by an Sp1 Binding Site 30 Base Pairs Upstream from the Sp1/initiator-like Site—It has been suggested previously that the UGT1A8, 1A9, and 1A10 T-region (see Fig. 1) may function as a TATA box in initiating their transcription (5). However, our data demonstrate that UGT1A8 and 1A10 are strongly regulated through the Sp1/initiator-like element, which is more indicative of TATA-independent transcription. In addition, the UGT1A8, 1A9, and 1A10 promoters have a conserved Sp1 site located 30 bp upstream from the functional Sp1/initiator-like region (Fig. 1). Sp1 sites are frequently found to enhance the activity of TATA-less, initiator-containing promoters (30, 31, 51). Therefore, to determine the importance of the T-region and the upstream Sp1 site in UGT1A8 and 1A10 regulation, −50 bp of the proximal promoters was cloned into the pGL3-basic vector and tested in transfection assays. The proximal 50-bp promoter fragments contain the upstream Sp1 site and Sp1/initiator-like region but not the upstream T-region. As shown in Fig. 7A, the proximal 50-bp UGT1A8 construct (1A8−55/+45) had a moderate activity of −2.5-fold over the pGL3-basic vector. A similar result was also obtained with the proximal 50-bp UGT1A10 promoter (data not shown). This suggests that the Sp1/initiator-like region is capable of stimulating UGT1A8 and 1A10 promoter activity in the absence of the T-region. Mutation of 3 bases in the upstream Sp1 site (1A8−55/+45 upstream Sp1mut) significantly decreased the activity of the proximal 50-bp UGT1A8 (Fig. 7A) and 1A10 (data not shown) promoters to levels similar to background (Fig. 7A), indicating that this site is important in enhancing UGT1A8 and 1A10 transcription. The importance of the upstream Sp1 site was also confirmed by experiments showing that mutation of this site in the context of the 1-kb UGT1A8 promoter construct resulted in an ~80% decrease in promoter activity (data not shown). The binding of Sp1 to the upstream Sp1 site was confirmed by gel shift assay with a probe corresponding to the upstream UGT1A8/10 upstream Sp1 site (Fig. 7C). The formation of two major complexes with the UGT1A8/10 upstream Sp1 wt probe (Fig. 7B, labeled C and D, lane 1) was observed. The formation of both complexes was reduced by a 50-fold molar excess of UGT1A8/10 upstream Sp1 wt (lane 2) and Sp1 consensus (Sp1 for; Table I). 0.5 μg of the wild-type and 5′-Sp1mut promoter constructs was transfected into Caco2 cells with 0.025 μg of the pRL-null vector as an internal control of transfection efficiency. Cells were harvested 48 h post-transfection and assayed for firefly and Renilla luciferase activities. Relative luciferase activities are expressed as a -fold induction over the promoterless pGL3-basic vector (set at a value of 1). The results shown are a representative experiment of two individual experiments performed in triplicate (±S.D.). B, the UGT1A8/10 5′-Sp1 wild-type probe (50,000 cpm) was incubated with 5 μg of Caco2 nuclear extracts and resolved on a 4% nondenaturing polyacrylamide gel. Unlabeled competitor oligonucleotides were added at a 50-fold molar excess before the addition of labeled probe. The Sp1 monoclonal antibody was added after the addition of labeled probe and incubated for 30 min before loading. The two major complexes, C and D, are labeled with arrows. C, sequences of the sense strand of the UGT1A8/10 5′-Sp1 wild-type and mutant (mut) probes are shown. Identical bases are marked with a dash, and the Sp1 mutation is indicated by the base changes. The 5′-Sp1 site is boxed.

Fig. 7. An Sp1 site upstream from the Sp1/initiator-like site enhances UGT1A8 promoter activity. A, a promoter construct containing the proximal 55 bp of the UGT1A8 promoter was generated as described under “Experimental Procedures.” A promoter construct of equal length containing a mutated upstream (5′) Sp1 site was generated as above using a forward primer with a 3-base mutation in the Sp1 site (1A8−55mut5′Sp1for; Table 1). 0.5 μg of the wild-type and 5′-Sp1mut promoter constructs was transfected into Caco2 cells with 0.025 μg of the pRL-null vector as an internal control of transfection efficiency. Cells were harvested 48 h post-transfection and assayed for firefly and Renilla luciferase activities. The proximal promoter fragment was cloned into the pGL3-control vector as an internal control of transfection efficiency. Cells were harvested 48 h post-transfection and assayed for firefly and Renilla luciferase activities. Relative luciferase activities are expressed as a _fold induction over the promoterless pGL3-basic vector (set at a value of 1). The results shown are a representative experiment of two individual experiments performed in triplicate (±S.D.). B, the UGT1A8/10 5′-Sp1 wild-type probe (50,000 cpm) was incubated with 5 μg of Caco2 nuclear extracts and resolved on a 4% nondenaturing polyacrylamide gel. Unlabeled competitor oligonucleotides were added at a 50-fold molar excess before the addition of labeled probe. The Sp1 monoclonal antibody was added after the addition of labeled probe and incubated for 30 min before loading. The two major complexes, C and D, are labeled with arrows. C, sequences of the sense strand of the UGT1A8/10 5′-Sp1 wild-type and mutant (mut) probes are shown. Identical bases are marked with a dash, and the Sp1 mutation is indicated by the base changes. The 5′-Sp1 site is boxed.
Differential Regulation of the UGT1A8, 1A9, and 1A10 Genes

formulation was dependent on the intact Sp1 site (lane 4). The presence of Sp1 in complex C was confirmed by the supershifting of complex C, but not complex D, with a monoclonal Sp1 antibody (lane 5). Taken together, these results suggest that the UGT1A8 and 1A10 promoters are regulated through an overlapping Sp1/initiator-like site, which is highly dependent on an upstream Sp1 site, and appears to be TATA-independent.

**DISCUSSION**

The UGT1A gene locus represents a unique structure in which eight functional UGT1A enzymes are encoded by a separate exon 1 alternatively spliced to a shared set of exons 2–5. The complex, tissue-specific expression of the UGT1A genes appears to be regulated by promoter elements in the 5′-region flanking their respective first exon. To date, only the promoter region of the human UGT1A1 has been characterized extensively. The UGT1A1 promoter contains a functional TATA box consisting of a polymorphic TA repeat region. The TA repeat region has been shown to vary between 5 and 8 repeats, the length inversely correlating with the activity of the UGT1A1 promoter (52). Comparison of the UGT1A2p, UGT1A3, UGT1A4, UGT1A5, and UGT1A6 promoters suggests that these genes have a TATA box in a position similar to that of the UGT1A1 gene (7), but their function is still to be tested. In this study, we have characterized the promoters of the UGT1A8, UGT1A9, and UGT1A10 genes and have shown that an overlapping Sp1/initiator-like site regulates their activity differentially. A putative TATA box (T-region) located upstream from the overlapping Sp1/initiator-like site was not essential for UGT1A8 and 1A10 promoter activity, indicating that these promoters are likely to be TATA-less in Caco2 cells. In support of this hypothesis, an Sp1 binding site upstream from the overlapping Sp1/initiator-like site, a common feature of TATA-less promoters, was demonstrated to be important for UGT1A8 and 1A10 promoter activity. However, it is possible that the T-region may function as a TATA box to regulate the UGT1A8, 1A9, and 1A10 promoters in other cell types.

Despite the high level of sequence similarity between the proximal UGT1A8, 1A9, and 1A10 promoters (>75%), the activity of the UGT1A8 and 1A10 proximal ~1-kb promoters were shown to be 7–8-fold higher than the equivalent UGT1A9 promoter in the Caco2 colon cell line (Fig. 3). The difference in promoter activity strongly, but not entirely, correlated with the binding of complex B to the overlapping Sp1/initiator-like site was shown to be important for activity and that the affinity of each site for Sp1 was sufficient for UGT1A8 promoter activity (Figs. 5 and 6). However, despite their apparent similarity in sequence, gel shift assays with a YY1 consensus oligonucleotide probe failed to demonstrate competition for complex B, indicating that YY1 was not present in this complex (data not shown). The UGT1A8/1A10 initiator-like site was similar in sequence to the tandemly repeated GABP binding sites in the cytochrome c oxidase subunit IV (COXIV) promoter (50). The COXIV GABP sites were sufficient to activate and direct transcription initiation of the COXIV gene in the absence of a TATA box or other initiator elements (53). However, extensive testing using competitor oligonucleotides and/or antibodies to GABP and the initiator-like binding proteins USF1, USF2, TFI-I, TAFII250, c-Myc, and E2F-1 demonstrated that these proteins were not present in complex B (data not shown). Initiator regions have also been shown to interact directly with the multisubunit TFIID complex, but these interactions have only been observed in the presence of an upstream TATA box (35, 36, 54). Therefore, it is likely that the protein(s) composing complex B are initiator-like-binding proteins not yet tested or identified.

Transfection and gel shift assays demonstrated that an upstream Sp1 site was important for activation of the UGT1A8 and 1A10 promoters (Fig. 7). It is intriguing that the binding of Sp1 to this site enhances promoter activity, whereas the binding of Sp1 to the Sp1/initiator-like site is not necessary for activity. This suggests that the context in which Sp1 binds is important for activity and that the affinity of each site for Sp1 may determine the effect of this transcription factor on rates of transcription. Sp1 sites are frequently found in TATA-less, initiator-containing promoters, where they enhance the activity and accuracy of transcription initiation. Although the mechanism of Sp1-mediated activation is not fully elucidated, Sp1 in the right context has been demonstrated to interact with components of the TFIID complex and stabilize their binding to initiator elements (34, 55–57). Sp1 has also been shown to interact with the initiator-like protein YY1 and synergistically enhance transcription through the adeno-associated virus P5 initiator site (37). Our data suggest that the UGT1A8/1A10 Sp1/initiator-like site was highly dependent on the upstream Sp1 site to enhance promoter activity because mutation of this Sp1 site decreased UGT1A8 promoter activity to background levels (Fig. 7A). The upstream Sp1 site was also present in the UGT1A9 promoter, but the lack of the functional Sp1/initiator-like site appeared to negate the effect of the Sp1 site. Interest-
ingly, UGT1A9 transcription was accurately initiated at 37 bp from the ATG, in the absence of the functional Sp1/initiator-like site (Fig. 1). The position of UGT1A9 transcription initiation varied from UGT1A8 and 1A10, being located 10 bp downstream of their 3′-transcription start sites. These findings suggest that the Sp1/initiator-like site may play an important role in directing the position of transcription initiation in UGT1A8 and 1A10 and that UGT1A9 may be regulated through a different initiator site. The sequences surrounding the UGT1A8, 1A9, and 1A10 transcription start sites also closely resembled the initiator consensus sequence, and the importance of these sites in transcription initiation needs to be assessed.

In summary, we have characterized the UGT1A8, 1A9, and 1A10 gene promoters and demonstrated that these genes are differentially regulated through an overlapping Sp1/initiator-like element in their proximal promoters. The UGT1A8/1A10 initiator-like site has several properties that are characteristic of initiator sites including its sequence similarity to the consensus initiator site, its proximity to the transcription start site, and its dependence on an upstream Sp1 site to promote transcription.
Cloning and Characterization of the Human UDP-glucuronosyltransferase 1A8, 1A9, and 1A10 Gene Promoters: DIFFERENTIAL REGULATION THROUGH AN INITIATOR-LIKE REGION

Philip A. Gregory, Dione A. Gardner-Stephen, Rikke H. Lewinsky, Kym N. Duncliffe and Peter I. Mackenzie


Access the most updated version of this article at doi: 10.1074/jbc.M305565200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 56 references, 32 of which can be accessed free at http://www.jbc.org/content/278/38/36107.full.html#ref-list-1