Comparative Studies of Human UDP-glucuronosyltransferase 1A8 and 1A9 Proximal Promoters Using Single Base Substitutions

Zhang XIAO1, Kana NUNOME2, Taemi YAHARA2, Emi INOUE2, Masakazu NABESHIMA2, Shirou TSUCHIDA2, Naoya HAMAUE2 and Takashi AOKI2,*

1Department of Biochemistry and Molecular Biology, Medical College of Qingdao University, Qingdao, China
2Department of Molecular Biosciences, School of Pharmaceutical Sciences, Health Sciences University of Hokkaido, Hokkaido, Japan

Summary: The nucleotide sequences of the proximal promoters of UDP-glucuronosyltransferase (UGT) 1A8 and 1A9 genes are very similar. However, UGT1A8 and 1A9 are mainly expressed in extra-hepatic and hepatic cells, respectively. Using mutants of UGT1A8 and 1A9 proximal promoters, we revealed their critical differences in terms of promoter activity and the role of the T-repeat region (T-region) conserved in both promoters. In extra-hepatic cells, Caco2, the activity of UGT1A9 proximal promoter increased to 73.4 ± 8.5% of that of the UGT1A8 proximal promoter with only 4 base changes: ¹160C, ¹152A, ¹62T, and ¹59G. The derivatives of the T-region showed that this region is not necessary for promoter activity, but the length of T repeats influences the activity somewhat. Therefore, the cause of the low activity of the UGT1A9 proximal promoter may be not only 4 base changes, but also the truncation of T repeats. From these results, the UGT1A9 proximal promoter was assumed to change into the non-active form from the original sequence, and this might be one of the reasons for the tissue-specific expression of UGT1A9.

Keywords: UGT1A8; UGT1A9; proximal promoter; T-region; tissue-specific expression; single base substitution

Introduction

The UDP glucuronosyltransferase 1 (UGT1) locus contains 13 individual versions of exon 1 (A1 to A13) in addition to shared exons 2–5.1) Four of the 13 individual versions of exon 1 are pseudogenes, while the remaining nine encode mature proteins (UGT1A1, 3–10). Each exon 1 possesses a specific promoter at 5'-UTR and creates an individual RNA molecule. Each UGT1 form has different specificity for the substrate and tissue distribution. UGT1A7, 1A8, and 1A10 are mainly expressed in the gastrointestinal tract, but the other UGT1 forms are found in many tissues including the liver.2–4)

The UGT1A7–10 genes have high homology and comprise a cluster,5) and only UGT1A9 is found in the liver.6) Gregory et al. reported that the promoters of UGT1A8 and 1A10 were activated depending on hepatocyte nuclear factor 1α (HNF1α) and caudal-related homeodomain protein 2 (Cdx2).6) On the other hand, the promoter activity of UGT1A9 depends on HNF1α and 4α.7–10) HNF4α is expressed in the liver, kidney, pancreas, and intestine, and it controls the expression of many metabolizing enzymes in human hepatocytes.11) However, the regulatory mechanism behind the hepatic and extra-hepatic expression of UGT1A7–10 has yet to be clarified completely.

The nucleotide sequences of 5'-UTR in UGT1A8 and 1A9 genes are very similar within the region about 900 bp from the initiation codon, and the proximal promoters are located within 200 bp. Both promoters contain a conserved T-repeat region (T-region) at the same position. It was previously suggested that the T-region may function as a TATA box;1) however, Gregory et al. reported that the activity of the UGT1A8 proximal promoter was not dependent on the T-region.6) Although the role of the T-region is still unclear, Yamanaka et al. reported that a single nucleotide polymorphism (SNP) in the T-region of UGT1A9 (T9 or T10) affected the promoter activity.12) In this study, using mutation derivatives of UGT1A8 and 1A9 proximal promoters, we examined (i) the critical differences between the two in terms of promoter activity, and (ii) the role of the T-region in this activity.

Materials and Methods

Primer and enzymes: The primers used in this study are listed in Supplementary Table S1. The enzymes were purchased...
from Invitrogen (Carlsbad, CA) or Nippon Gene Co. (Toyama, Japan), and all reactions were performed as specified by the manufacturers. Plasmids were transformed into *Escherichia coli* JM109 with CaCl₂, and the transformants were then selected on LB agar plates containing 50 μg of ampicillin/ml. The DNA was sequenced with an ABI PRISM 310 Genetic Analyzer (Applied Biosystems Japan, Tokyo, Japan).

**Cloning of the UGT1A8 and 1A9 proximal promoters:**

The proximal promoter region of UGT1A8 (−170/−5, A of the initiation codon corresponds to a position of +1) was amplified by PCR using the primers 1A8F and 1A8R. One hundred nanograms of human genomic DNA prepared from peripheral blood was used as the template. The amplified DNA was digested with *Kpn*I and *Hind*III, and cloned into a luciferase reporter vector, pGL3-basic (Promega, Madison, WI). The resulting plasmid was designated pGL1A8. The proximal promoter region of UGT1A9 (−171/−5) was amplified using the primers 1A9F and 1A9R. The amplified DNA was cloned into pGL3-basic by the same procedure. The resulting plasmid was designated pGL1A9.

**Construction of pGL1A8 derivatives:**

Plasmids pGL1A8-T₀ to T₉₀ were constructed by sequentially increasing or decreasing T by 2 nucleotides. Site-directed mutagenesis was performed by the circular mutagenesis method. One microgram of pGL1A8-T₁ was used as the template, and the primers 1A8Tₓ₁₋₃F and 1A8Tₓ₋₃R were used for mutagenesis.

The T-region in pGL1A8-T₁₀ was also changed to a C₁₀ repeat using the following steps. The PCR products amplified using primers 1A8Cₓ₁₀F and 1A8R and using 1A8F and 1A8Cₓ₁₀R were mixed, and the fragment extension was repeated five times under conditions of 95°C for 30 s, 61°C for 30 s, and 72°C for 120 s. The extension product was then amplified using primers 1A8F and 1A8R, and cloned into pGL3-basic. The resulting plasmid was designated pGL1A8-C₁₀.

**Construction of pGL1A9 derivatives:**

The sequence upstream of the T-region in pGL1A9 was amplified using the primers 1A9F and 1A9r2R. The sequence downstream of it was also amplified using the primers 1A9R and 1A9r2F. The PCR products (upstream and downstream of the T-region) were mixed, and fragment extension and second PCR using the primers 1A9F and 1A9R were performed as described above. The amplified DNA was cloned into pGL3-basic, and the resulting plasmid was designated pGL1A9-r2. Parts of the UGT1A9 promoter (−171/−44) in pGL1A9 and pGL1A9-r2 were amplified using the primers 1A9F and 1A9r3R. Another part of the UGT1A9 promoter (−82/−5) in pGL1A9 was also amplified using the primers 1A9R and 1A9r3F. The fragment extension, second PCR, and the cloning were performed by the same procedure. The resulting plasmids were designated pGL1A9-r3 and pGL1A9-r23.

Next, the UGT1A9 promoters (−157/−5) in pGL1A9, pGL1A9-r2, pGL1A9-r3, and pGL1A9-r23 were amplified using the primers 1A9r1F and 1A9R. The PCR products were then re-amplified using the primers 1A8F and 1A9R. These fragments were cloned into pGL3-basic, and the resulting plasmids were designated pGL1A9-r1, pGL1A9-r12, pGL1A9-r13, and pGL1A9-r123, respectively. Base substitutions in regions 1 and 3 of pGL1A9 were performed by circular mutagenesis method using the primers shown in Supplementary Table S1.

**Cell culture, transfection, and luciferase assay:**

All results about the promoter activity were provided using Caco2 cells. Caco2 cells 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, and 80 μg/ml gentamicin at 37°C in 5% CO₂. Seven-tenths μg of target plasmids, 0.3 μg of *Renilla* luciferase vector phRL-TK (Promega), and 2 μl of Lipofectamine (Invitrogen) were mixed, and the total volume was adjusted to 100 μl by adding the medium. The plasmid solutions were incubated for 20 min at room temperature, and were added to the cells cultured in 24-well plates at a density of about 1.0 × 10⁵ cells in 500 μl of medium/well. After incubation for 24 h, the medium was removed and the cells were washed with PBS (8.1 mM Na₂HPO₄, 1.5 mM KH₂PO₄, 137 mM NaCl, 2.7 mM KCl, pH 7.4).

The cells were lysed using 100 μl of 100 mM K-phosphate buffer (pH 7.8) containing 0.5% Triton X-100, and the firefly luciferase and *Renilla* luciferase activities of 50 μl of lysate were assayed using the Dual-Luciferase Reporter Assay System (Promega). To verify the validity of this luciferase assay system, pGL3-control containing SV40 promoter and pGL3-basic containing no promoter were used as the positive and negative controls, respectively. Statistical analyses were performed using Student’s t-test.

**Results and Discussion**

**Activation of UGT1A9 proximal promoter by conversion into UGT1A8 sequence:**

The activities of UGT1A8 and 1A9 proximal promoters were examined in Caco2 cells, which is an *in vitro* model of the gastrointestinal tract. The results obtained from the cells might not necessarily correspond to the *in vivo* results, but at least we are able to understand the tendency of both promoters. Despite the high level of sequence similarity between the proximal UGT1A8 and 1A9 promoters (Fig. 1A), the activity of the UGT1A8 promoter was shown to be about 5-fold higher than that of the equivalent UGT1A9 promoter in Caco2 cells (Fig. 1B).

In the proximal promoter of UGT1A9, as shown in Figure 1A, the nucleotide sequence differs from that of UGT1A8 in three main regions, region 1 (upstream of the T-region), region 2 (around the T-region), and region 3 (Sp1/initiator-like region). To determine which differences are more critical for the promoter activity, the UGT1A9 promoter was converted into the UGT1A8 sequence. In the conversion of region 1, 4 bases (−160T, −152T, −149G, and −137T) were replaced by the UGT1A8 sequence (C, A, A, and G, respectively). In the conversion of region 2, although the T-region was conserved as T₉₀, 3 bases (−130C, −113A, and −104A) were replaced by G, C, and T, and 2 nucleotides (T and G) were added between −128 and −127. In the conversion of region 3, 5 bases (−72T, −71G, −67T, −62C, and −59A) were replaced by C, T, G, T, and G.

As shown in Figure 1C, the conversion of region 3 had the greatest effect on the promoter activity. Gregory *et al.* reported that the mutation of this region decreased UGT1A8 promoter activity to background levels. Our results showed that the UGT1A9 promoter activity increased up to about 60% of the UGT1A8 promoter with the conversion of this region to the UGT1A8 sequence. The difference in region 3 is, therefore, the main reason for the low activity of the UGT1A9 proximal promoter, but as shown in Figure 1C, the difference in region 1 is also likely to influence the activity somewhat. The conversions of both region 1 and region 3 brought increased activity compared with that solely by the conversion of region 3. Because the conversion of region 2 did not influence the activity, it is thought that the importance of the differences between the promoters is as follows: regions 3 >
Then, to determine which differences in nucleotides in regions 1 and 3 of the UGT1A9 promoter lead to the low activity, many mutants converted to the UGT1A8 sequence were constructed by one base change. As shown in Table 1, in region 3, the most important base changes were $¹^6^2^C$ and $¹^5^9^A$, while three differences, $¹^7^2^T$, $¹^7^1^G$, and $¹^6^7^T$, were not critical for the activity. The activity of the mutant having $¹^6^2^T$ and $¹^5^9^G$ was 49.9 ± 5.0% compared with that with the UGT1A8 promoter, which was almost equal to the effect of the entire conversion of region 3 (Fig. 1C, r3). In region 1, the most important base changes were $¹^1^6^0^T$ and $¹^1^5^2^T$, and the relative activity of the mutant having $¹^6^2^T$ and $¹^5^9^G$ was 49.9 ± 5.0% compared with that with the UGT1A8 promoter, which was almost equal to the effect of the entire conversion of region 3 (Fig. 1C, r3). In region 1, the most important base changes were $¹^1^6^0^T$ and $¹^1^5^2^T$, and the relative activity of the mutant having $¹^6^2^T$ and $¹^5^9^G$ was 49.9 ± 5.0% compared with that with the UGT1A8 promoter, which was almost equal to the effect of the entire conversion of region 3 (Fig. 1C, r3). In region 1, the most important base changes were $¹^1^6^0^T$ and $¹^1^5^2^T$, and the relative activity of the mutant having $¹^6^2^T$ and $¹^5^9^G$ was 49.9 ± 5.0% compared with that with the UGT1A8 promoter, which was almost equal to the effect of the entire conversion of region 3 (Fig. 1C, r3). In region 1, the most important base changes were $¹^1^6^0^T$ and $¹^1^5^2^T$, and the relative activity of the mutant having $¹^6^2^T$ and $¹^5^9^G$ was 49.9 ± 5.0% compared with that with the UGT1A8 promoter, which was almost equal to the effect of the entire conversion of region 3 (Fig. 1C, r3).
The expression of UGT1A10 is regulated by HNF1α and 4α which are expressed at the highest levels in the liver. Therefore, HNF1α and HNF4α might function as activation factors for the distal promoter in hepatic cells. The suppression of the UGT1A8 proximal promoter in hepatic cells is unclear. Since this promoter is activated by HNF1α and Cdx2, this activation process might be interfered with by other regulatory factors. Although more studies are necessary to clarify the tissue-specific expression of UGT1A7–10, T repeats may influence the promoter activity to some extent.

**References**


