Identification of Functional Promoter Haplotypes of Human Concentrative Nucleoside Transporter 2, hCNT2 (SLC28A2)

Linghui Li, Seok Hwee Koo, Ivan Kean Hong and Edmund Jon Deoon Lee*
Department of Pharmacology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore

Summary: The human concentrative nucleoside transporter 2 (hCNT2) plays a major role in the intestinal absorption of naturally occurring nucleosides as well as some nucleoside analog drugs. To determine if single nucleotide polymorphisms (SNPs) in the promoter region of hCNT2 affect gene expression, we examined approximately 1 kb upstream the hCNT2 transcription start site. Ninety Chinese samples were screened and seven SNPs were identified: -115T>G, -146T>A, -264A>G, -564G>A, -861A>C, -880T>C and -906C>T. Based on these seven variants and their relative positions, eight haplotypes were identified using PHASE v2.1.1. Three naturally occurring haplotypes were cloned into the pGL3-Basic vector and transfected into HEK293 cells. Dual luciferase assay revealed that haplotype 4 (GTAGACC) and 7 (GAGAACT) exhibited significantly lower expression levels compared to the published haplotype 1 (TTAGATC). Results from our in-vitro study showed that the hCNT2 promoter region haplotype may modulate gene expression and cause different drug responses.

Keywords: hCNT2; promoter haplotype; expression

Introduction

Synthetic analogs of nucleosides are widely used in anticancer and antiviral chemotherapies. Like naturally occurring nucleosides, these hydrophilic nucleoside derivatives require nucleoside transporters (NTs) to traverse the cell membrane. Therefore, NTs are critical determinants likely to influence the pharmacokinetics of nucleoside analog drugs and consequent response to therapy. Currently, two major families of NTs in humans have been identified by molecular cloning and functional expression: human equilibrative nucleoside transporters (hENTs) and human concentrative nucleoside transporters (hCNTs).

The human concentrative nucleoside transporter 2 (hCNT2), also known as hSPNT1 or SLC28A2, was first cloned from human kidney.1) The hCNT2 gene is located on chromosome 15q13–14. The hCNT2 protein exhibits the functional characteristics of a Na+–dependent nucleoside transporter with selectivity for adenosine, other purine nucleosides and uridine.2) hCNT2 has a wide tissue distribution and appears to have higher mRNA level in human duodenum compared to other NTs such as hCNT1, hENT1 and hENT2, implying a highly significant role in the uptake of nucleoside drugs in a variety of human tissues especially the intestine.1,3) hCNT2 has been reported to facilitate the uptake of antiviral compounds such as didanosine and ribavirin, used in the treatment of HIV2) and hepatitis C,9) respectively. More recently, intestinal hCNT2 was also found to interact with some novel anticancer nucleoside analogs and amino acid ester prodrugs of nucleoside drugs.1,5)

Inter-individual variation in response to anticancer and antiviral nucleoside analogs remains a great challenge to physicians. Genetic variations altering the functions or expression of NTs in various tissues among individuals may result in differences in pharmacokinetics and toxicity of nucleoside analogs, which, in turn, may lead to inter-individual variation in drug response. Population studies in the United States and Singapore both suggest that genetic variations in the coding region of hCNT2 are unlikely to contribute much to variation in response to ribavirin.6,7) However, the hCNT2 gene promoter region has never been cloned and studied to in-
vestigate the possible influence of promoter polymorphisms on hCNT2 gene expression. This study has two major goals: (1) identify polymorphisms occurring in the proximal promoter region of hCNT2; (2) characterize their potential influence on the expression of hCNT2 in vitro.

Materials and methods

Genetic analysis of the upstream region of hCNT2: The 1072 bp sequence from −1 to −1072 upstream the transcription start site of hCNT2 (accession number: NT_010194.16) was examined for the presence of single nucleotide polymorphisms (SNPs) by sequencing the genetic material from 90 fully anonymized white cell lines developed from 90 unrelated healthy Chinese volunteers. All consented donors had been recruited in accordance with local ethics requirements and study protocol approved by the National University of Singapore Ethics Committee. 1 kb of the proximal promoter region was divided into three overlapping fragments of approximately 500 bp each. They were amplified by polymerase chain reaction (PCR) using the primers pairs presented in Table 1. PCR products were subjected to direct sequencing using BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, USA). Sequencing data were analyzed with Mutation Surveyor™ version 2.61 (Softgenetics LLC, State College, USA). After identification of SNPs in this region, putative transcription factor binding sites affected by positions of SNPs were screened using the TRANSFAC database available at a public website (http://motif.genome.jp). Finally, haplotypes were reconstructed from variant positions by PHASE v2.1.1.

Subcloning: Three naturally occurring common haplotypes (1, 4 and 7) were selected for functional analysis. A representative sample for each of the three haplotypes was selected and the 1088 bp fragment from 16 to −1072 containing seven variants was amplified using the cloning primer pairs shown in Table 1. Restriction enzyme sites KpnI and NheI were incorporated at the ends of the forward and reverse primers, respectively. The amplification was performed under the following conditions: 1 cycle at 94°C for 4 min, 30 cycles at 94°C for 1 min, 64°C for 45 sec, 72°C for 1 min and 1 final cycle at 72°C for 10 min. Amplified products were cloned into the TA vector pCRII-TOPO (Invitrogen, Carlsbad, USA). Positive clones were screened by colony-PCR and subjected to direct sequencing to verify the sequence of each haplotype. The TOPO-haplotype constructs were doubly digested with enzymes: KpnI and NheI (Promega, Madison, USA). The digested fragments were then subcloned into pGL3-Basic vector (Promega, Madison, USA). The sequence of each pGL3-Basic-haplotype was reconfirmed by direct sequencing.

Cell culture: HEK293 cells were maintained in EMEM (Essential Minimum Eagle Medium) supplemented with 2 mM L-glutamine, 1 mM Sodium Pyruvate, 1.5 g/L NaHCO3 and 10% fetal bovine serum (Sigma, St. Louis, Missouri, USA). All cells were cultured in an atmosphere of 5% CO2 and 90% relative humidity at 37°C.

Transfection: Cells were seeded in a 96-well plate at a density of 30,000 cells per well the day before transfection. Cells were cotransfected with 150 ng pGL3-Basic-haplotype and 0.5 ng pRL-CMV as an internal control and 0.5 μl Lipofectamine 2000 (Invitrogen, Carlsbad, USA) was used for the cotransfection. The negative control (pGL3-Basic vector without a promoter) and positive control (pGL3-Control vector with SV40 promoter) were carried out in parallel. The cells were harvested 24 hours and 48 hours after transfection and assayed with the Dual luciferase system (Promega, Madison, Wisconsin, USA).

Dual luciferase reporter assay: The experimental Firefly luciferase and internal control Renilla luciferase activity assayed according to manufacturer’s instructions, with minor modification. Each sample was lysed in 20 μl 1 × Passive Lysis Buffer for 30 min after washing with phosphate buffered saline (PBS). The lysate was transferred to an opaque white bottomed NUNC plate and assayed for firefly and renilla activity by adding 40 μl Luciferase assay reagent II (LARII) and 40 μl Stop & Glo reagent in tandem. The luminescence activity of each firefly and Renilla was read on Tecan GENios Pro (Tecan, Switzerland).

<table>
<thead>
<tr>
<th>Primer</th>
<th>Primer sequences</th>
<th>PCR product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fragment 1</td>
<td>Forward 5’-CAGGGTCTCAGTGCAGTCG-3’</td>
<td>61–448</td>
</tr>
<tr>
<td></td>
<td>Reverse 5’-CCAAACACCTTATAGATACTAC-3’</td>
<td>−327–859</td>
</tr>
<tr>
<td>Fragment 2</td>
<td>Forward 5’-TACAAAGGCTAGTACCGC-3’</td>
<td>−61–1258</td>
</tr>
<tr>
<td></td>
<td>Reverse 5’-AGACAGGGCTGCTCGGC-3’</td>
<td>−61–1258</td>
</tr>
<tr>
<td>Fragment 3</td>
<td>Forward 5’-TTGTCGACATGTCATGTGCTGTCAGTCG-3’</td>
<td>16–1072</td>
</tr>
<tr>
<td></td>
<td>Reverse 5’-AGCGTAGCTCTCCAGTAAGGACT-3’</td>
<td></td>
</tr>
</tbody>
</table>

* Underlined nucleotide sequences in the primer pairs for cloning represent specially designed restriction sites KpnI and NheI respectively.
Fig. 1. hCNT2 putative promoter
The nucleotide sequence (NT_010194.16) is numbered from the transcription start site as +1. Putative transcriptional binding sites are shown in boldface and named below the sequence. Newly identified SNPs are highlighted in red. Putative transcription factor binding sites affected by SNPs are underlined with lines and named behind. +: predicted gain of a transcription factor binding site while −: predicted loss of a transcription factor binding site.
**Statistical analysis:** The ratio of experimental firefly luciferase activity to control Renilla luciferase activity was calculated and compared for different haplotypes using ANOVA analysis (SPSS 13.0, Chicago, USA)

**Results**

**Genetic analysis of promoter region of hCNT2:** About 1 kb of the 5′-flanking sequence upstream the transcription start site was analyzed to identify putative transcription factors and other motifs. A number of putative transcription factor binding sites was identified including binding sites for CCAAT-binding transcription factors (C/EBPα, NF-1 and NF-Y) and HNF-3 transcription factors (HNF-3β). Genetic analysis of this region in 90 unrelated Chinese samples identified seven polymorphisms: −115T>G, −146T>A, −264A>G, −564G>A, −861A>C, −880T>C and −906G>T (Fig. 1). The allele frequencies of each SNP are presented in Table 2. The allele frequencies are all greater than 20% in the Chinese population. Interestingly, SNPs at positions −115 and −880 are represented at frequencies greater than 75%.

Following the identification of SNPs in this region, search for putative transcription factor binding sites affected by positions of SNPs predicted three of the seven SNPs to gain or loss of a consensus eukaryotic transcription site (Fig. 1). The −146T>A led to the gain of binding site for Dfd (Deformed), S8, HNF-1 (Hepatic nuclear factor 1) and Oct-1 (Octamer factor 1); the −264A>G led to loss of GATA-1 (GATA-binding factor 1) and MATα2 (Mating factor α2); the −861A>C led to loss of Bcd (Bicoid), Deformed and Ik-2 (Ikaros 2) sites.

Eight haplotypes were reconstructed from variant positions by PHASE v2.1.1 and their frequencies are listed in Table 3. Seven of these occurred more than once and haplotype 4 (GAGAACT) was the commonest haplotype in the Chinese population.

**Functional study of promoter haplotype:** To determine if promoter region haplotype influences hCNT2 expression, three common haplotypes (1, 4 and 7) out of the eight haplotypes were selected for functional analysis. These haplotypes represent 76% of the total haplotypes and also naturally occurring haplotypes from the sequencing results. The HEK-293 cell line was chosen for reporter gene assay because: (1) hCNT2 is known to have a high expression level in human kidney; 2) the HEK-293 cell line is commonly used in transfection studies because of its high efficiency. The average expression level and standard deviation (SD) of three replicate assays at 24 and 48 hours for the three haplotypes and the control (negative & positive) are plotted in Figure 2. As expected, the pGL-3 Basic vector (negative control) exhibited very low activity. The vector/insert constructs showed much higher activity comparable to the pGL3-Control vector. ANOVA was run to compare the expression level of different haplotypes using SPSS 13.0. There was no statistical difference between haplotypes 4 and 7 (P > 0.05). However, when compared to haplotype 1, the expression level of hCNT2 was significantly lower in haplotype 4 and haplotype 7, both at 24 hours and 48

**Table 2.** Location, allele frequencies and predicted transcription factor binding site changes of single nucleotide polymorphisms in the proximal promoter region of hCNT2

<table>
<thead>
<tr>
<th>Location</th>
<th>g. −115T&gt;G</th>
<th>g. −146T&gt;A</th>
<th>g. −264A&gt;G</th>
<th>g. −564G&gt;A</th>
<th>g. −861A&gt;C</th>
<th>g. −880T&gt;C</th>
<th>g. −906C&gt;T</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allele frequencies</td>
<td>f(T)=0.2</td>
<td>f(T)=0.75</td>
<td>f(A)=0.75</td>
<td>f(G)=0.74</td>
<td>f(A)=0.78</td>
<td>f(T)=0.22</td>
<td>f(C)=0.74</td>
</tr>
<tr>
<td>TRANSFAC prediction</td>
<td>None</td>
<td>(+)Kd (−)AG</td>
<td>(+)S8 (−)GATA-1</td>
<td>(+)MATα2 (−)Bcd</td>
<td>None</td>
<td>(−)Dfd (−)Ik-2</td>
<td></td>
</tr>
</tbody>
</table>

![Location relative to the transcription start site (NT_010194.16)](https://example.com/Location.jpg)

**Table 3.** Promoter haplotypes of hCNT2

<table>
<thead>
<tr>
<th>Haplotype name</th>
<th>g. −115T&gt;G</th>
<th>g. −146T&gt;A</th>
<th>g. −264A&gt;G</th>
<th>g. −564G&gt;A</th>
<th>g. −861A&gt;C</th>
<th>g. −880T&gt;C</th>
<th>g. −906C&gt;T</th>
<th>Frequency (% of total haplotype)</th>
</tr>
</thead>
</table>
Fig. 2. Expression activity (ratio of experimental firefly luciferase activity to the control renilla luciferase activity) of negative & positive control and experimental haplotypes at 24 hrs and 48 hrs. Data are representative of two transfection experiments with six replicates for each haplotype (three replicates for lysis at 24 hrs and three replicates for lysis at 48 hrs), presented as mean ± SD.

hours (P < 0.05). Both haplotypes 4 and 7 showed average expression approximately 30% lower than haplotype 1 at 48 hours.

Discussion

This study characterizes seven common SNPs (−115T > G, −146T > A, −264A > G, −564G > A, −861A > C, −880T > C and −906C > T) in the proximal promoter region of hCNT2. All are common variants with a frequency rate greater than or equal to 20%. Transcription factor binding motif search revealed that the putative hCNT2 gene promoter contains various potential binding sites for CCAAT-binding transcription factors (C/EBP, NF-1/CTF, NF-Y/CBF) and HNF3 transcription factors which may potentially explain the recent finding that transcription factors such as C/EBPα and HNF3γ modulate the expression of hCNT2. Among the seven SNPs identified in this study, only three (g. −146T > A, g. −264A > T and g. −861T > C) led to gain or loss of a consensus eukaryotic transcription site predicted by TRANSFAC. Through mutation in these transcription sites, the affinity of transcription factors to DNA may be changed. This is the most common way in which SNPs may alter gene expression.

Among the eight haplotypes inferred from the positions of these seven variants, haplotype 4 (GTAGACC) is the commonest in the Chinese population, while haplotype 1 (TTAGATC) with the same sequence as that reported in the online database accounts for 13% of all haplotypes. Haplotype 4 differs from haplotype 1 in nucleotides at positions −115 and −880. G and C are predominant variants at −115 and −880 in the Chinese population, taking up 80% and 78% of all alleles respectively. It is uncertain if haplotype 1 represents the commonest haplotype globally since this has never been reported. In addition to being high-frequency variants, half of the eight haplotypes has a frequency rate exceeding 10%, suggesting that the haplotype distribution pattern may be quite different for various populations.

Dual luciferase assay results revealed that haplotypes 4 and 7 had a 30% lower mean relative expression level compared to haplotype 1. A comparison of variants among these three haplotypes showed that differences between haplotypes 1 and 4 are the variants at positions −115 and −880, while the only similarity between haplotypes 1 and 7 exists at position −861. In general, mutations affecting functional transcription factor binding sites are more likely to have impact on expression activity directly through regulatory sequences. However, in our study, we found that mutations which have no influence on transcription factor binding sites such as T−115G and T−880C can also contribute to differences in expression activity. For the three SNPs located at the predicted transcription factor binding sites, haplotype 7 differs from both haplotypes 1 and 4 in T−146A and A−264G. However, there is no significant difference in expression activity between haplotypes 7 and 4 in the in vitro assay. This implies that the overall effect of these two transcription binding site SNPs is not significant and they probably do not disrupt the transcription binding site. Given the complexity of gene transcription involving many transcription factors, a single position change in any binding site may not be the only determinant of expression. These findings indicate that the multisite haplotype rather than one SNP may ultimately be responsible for hCNT2 regulation.

Quantitative northern analysis of the differential expression of NTs in normal and tumor tissues showed that there is considerable variation in all NT expression profiles from different individuals. This variable expression of NTs in normal and tumor tissues showed that there is considerable variation in all NT expression profiles from different individuals. It has been indicated that a number of transcription factors can influence the expression of NTs. However, the molecular mechanism of their transcriptional activation of NT genes is still poorly understood. Promoter sequences may
harbor functionally relevant polymorphisms affecting gene expression since they are involved in initiating transcription. Screening, the first of 500 bp of the 5' flanking region in 170 genes revealed that 35% of these proximal promoters is polymorphic and one third of these promoter variants may modify gene expression. However, there have been few studies on the promoter region of NTs and only the promoter for hENT1 has been characterized so far.

From our in-vitro functional study, different haplotypes in the promoter region of hCNT2 carried by each individual could have important effects on the variable expression level of hCNT2 and thus affect the transport activity at the intestinal mucosa. However, more studies are needed in order to apply our in-vitro findings in-vivo. Together, genetic variation in the promoter region of hCNT2 may be associated with in-vitro variations in hCNT2 expression but the clinical significance of this potential variation with respect to the clinical response to anti hepatitis C treatment with a purine analog nucleoside such as ribavirin needs to be clarified.

Acknowledgments: This study was supported by a grant from the National Medical Research Council NMRC/0952/2005.

References


