Characterization of the 5'-flanking fragment of the human GM3-synthase gene

Guichao Zeng*, Luoyi Gao, Tian Xia, Tewin Tencomnao, Robert K. Yu

Institute of Molecular Medicine and Genetics, Medical College of Georgia, 1120 15th Street CB-2083, Augusta, GA 30912, USA

Received 28 August 2002; received in revised form 10 October 2002; accepted 10 October 2002

Abstract

To investigate the transcriptional regulation of human GM3-synthase, a 5'-flanking fragment of 1379 bp was cloned by a PCR-based procedure. Analysis of the human genomic sequence showed that the gene consists of seven exons, locates at chromosome 2, and spans over 62 kb. There are a number of potential consensus binding sites in the cloned promoter region, but TATA and CCAAT boxes were not found in the promoter. Primer extension analysis identified two transcription start sites approximately 11 and 57 bp upstream of the exon 1. The transcription activity of the promoter was assessed in human HeLa cells by transient transfection. Of the fragments assayed, the proximal 409 bp fragment exhibits the highest transcription activity. Transcription factors that bound to the 409 bp fragment were pulled down by DNA-coupled magnetic beads. Identities of the pull-down proteins were determined by array analysis. Eight transcription factors were identified, which might either bind to the proximal region or be recruited as co-activators of the transcription factor complexes.

Gangliosides are glycosphingolipids that contain sialic acid as ubiquitous membrane components in mammalian cells. Ganglioside GM3 is a common precursor for nearly all of the naturally occurring gangliosides and is suggested to play important roles in various cellular functions. It induces monocytic differentiation of human leukemia cell lines HL-60 and U937 [1–3]. GM3 is known as a modulator of transmembrane signaling by regulating various receptor functions. For example, GM3 inhibits phosphorylation of the receptors of epidermal growth factor (EGF) and platelet-derived growth factor (PDGF) [4,5], and thus suppresses cell proliferation. GM3 may also participate in cell adhesion/recognition through glycosphingolipid microdomains [6]. GM3 is synthesized by GM3-synthase (CMP-NeuAc:lactosylceramide α2-3 sialyltransferase, EC 2.4.99.9), a key regulatory enzyme for initiation of the ganglioside biosynthesis. The cDNA for the human and mouse GM3-synthase has been cloned [7–10]. However, transcription regulation of GM3-synthase has not been examined. On the other hand, while many of the components of the basal transcription machinery have been identified, cloned and studied in varying detail, much less is known about regulatory transcription factor complexes [11]. Therefore, in the present study, we characterized the 5'-flanking fragment of the gene and demonstrated the basal promoter of the human GM3-synthase gene. We also identified some transcription factors that bound to the proximal promoter region directly or were recruited as components of the regulatory complexes.

A human genome BLAST homology search of the public database (http://www.ncbi.nlm.nih.gov/genome/guide/human) was performed using the human GM3-synthase cDNA (GenBank Accession numbers AB018356 and AF119415) reported by Ishii et al. [7] and us [9]. The result showed that the gene is located at chromosome 2 between p24.1 and p24.3, consists of seven exons, and spans over 62 kb (Fig. 1). The ATG codon is located in exon 2, which is 25 kb downstream of exon 1. Our analysis revealed that the sizes of the introns, and thus the overall size of the gene, were larger than those reported by Kim et al. [12]. The smaller size of the gene probably resulted from the use of incomplete human genomic contig sequences from the database. The authors also found that there were four isoforms of GM3-synthase mRNA in fetal brain, which differ in the 5' untranslated region and only one of these isoforms was detected in adult brain [12], indicating that regulation of GM3-synthase expression may be development-dependent.
The published cDNA sequences [7,9] showing two sizes of the exon 1 (Fig. 1) may represent two of such isoforms [12].

In order to understand the regulatory mechanism for the human GM3 synthesis gene expression, the 5′ flanking untranscribed fragments of the human GM3-synthase gene were cloned by PCR amplification using the nested primers designed from the genomic sequence (Fig. 2). The forward primers were derived from the upstream region of the exon 1. They were F1: 5′-ACATCTGCCCATATATACAAGC (bp 1344 to 1323); F2: 5′-GCAGTGGCGATCTCGGCTC (bp 1256 to 1238); F3: 5′-CACCTTTTCCTTTGCAACGC (bp 658 to 639); and F4: 5′-CTTAAGCTTTGTTTCCTTCC (bp 286 to 267) (bp 1 represents the first nucleotide upstream of the exon 1). The nested reverse primers were derived from the exon 1. They were R1: 5′-GCCTCGGTCCGCGGCTGCAGG (+148 to +168) and R2: 5′-GCCTTCGTCCGCATACTAATG (+103 to +123) (+1 represents the first nucleotide of the exon 1). Human genomic DNA was prepared from human HeLa cells as described [13]. The first round of PCR amplification was carried out using primers F1 and R1, and human genomic DNA as template. The PCR products were used as templates in the second round of PCR with the nest primers. Three PCR fragments of 1379, 781 and 409 bp were obtained using the primer pairs of F2 + R2, F3 + R2, and F4 + R2, respectively, and cloned into the pGL-3 Basic reporter, a promoterless luciferase expression vector (Promega, Madison, WI). DNA sequencing of the cloned promoter fragments showed that the sequences of the fragments were identical to the human genomic sequence (Fig. 2). Further computer analysis of the promoter fragments was conducted to detect putative cis-acting elements using a transcription factor database (Wisconsin Package Version 10.2, Genetics Computer Group (GCG), Madison, WI). A number of putative consensus binding sites for known transcription factors were identified including, but not limited to SP1, CREB, C/EBP, CNBP, and NFI, while both TATA and CCAAT boxes were not found in the promoter (Fig. 2).

Fig. 1. Structure of the human GM3-synthase gene. The structure of the human GM3-synthase gene (not drawn to scale) is shown illustrating the seven exons (1–7) and six introns (I–VI) (top). Number of base pairs of each exon or intron is presented in the table (bottom). Two sizes (191 and 107 bp) of the exon 1 were reported [7,9].

![Fig. 2. Nucleotide sequence of the 5′-flanking region of the human GM3-synthase gene. This sequence was originally downloaded from the public database as described in the text, to which the sequence of the cloned promoter fragment was identical. The first exon was presented as reported by Ishii et al. [7]. A smaller exon 1 was identical to +85 to +191 as reported by us [9]. Putative recognition sites for known transcriptional regulatory factors are underlined. +1 represents the first nucleotide of the exon 1. Human genomic DNA was prepared from human HeLa cells as described [13]. The first round of PCR amplification was carried out using primers F1 and R1, and human genomic DNA as template. The PCR products were used as templates in the second round of PCR with the nest primers. Three PCR fragments of 1379, 781 and 409 bp were obtained using the primer pairs of F2 + R2, F3 + R2, and F4 + R2, respectively, and cloned into the pGL-3 Basic reporter, a promoterless luciferase expression vector (Promega, Madison, WI). DNA sequencing of the cloned promoter fragments showed that the sequences of the fragments were identical to the human genomic sequence (Fig. 2). Further computer analysis of the promoter fragments was conducted to detect putative cis-acting elements using a transcription factor database (Wisconsin Package Version 10.2, Genetics Computer Group (GCG), Madison, WI). A number of putative consensus binding sites for known transcription factors were identified including, but not limited to SP1, CREB, C/EBP, CNBP, and NFI, while both TATA and CCAAT boxes were not found in the promoter (Fig. 2).]
The promoter activity of the cloned fragments was assayed for luciferase expression by transfection of the plasmids into human HeLa cells. Cells were grown in Dulbecco's Modified Eagle Medium supplemented with 10% fetal bovine serum (Invitrogen, Carlsbad, CA). Five hundred thousand cells cultured in a 35 mm Petri dish were transfected with 2 μg of each plasmid, along with 0.5 μg of either the pβ-gal Control (Clontech, Palo Alto, CA) or the pRL-CMV, a Renilla luciferase control reporter vector (Promega). After an overnight transfection in serum-free medium, cells were allowed to recover for 24 h in serum-containing medium and lysed for luciferase activity, which was normalized by co-transfected β-galactosidase activity detected by the Luminescent β-galactosidase Detection Kit II (Clontech) or Renilla luciferase activity by the Dual Luciferase Reporter Assays kit (Promega). Results are shown in Fig. 3. All of the three fragments showed significant luciferase activities in HeLa cells. The highest promoter activity was present in the 409 bp fragment (−286 to +123).

The transcription initiation site was determined by primer extension analysis. Primer extension was carried out using a primer extension system (Promega) as previously described [13]. An antisense oligonucleotide (5'-ATTCAAGCTGGG-GGCGCGCCGCTC), corresponding to positions +11 to +32 (Fig. 2) was 5’ end-labeled with [γ-32P]ATP using T4 polynucleotide kinase. Total cellular RNA was isolated from HeLa cells using the RNeasy Mini Kit (Qiagen, Chatsworth, CA). The end-labeled primer (~10^6 cpm) was added to 20 μg of total RNA, incubated for 30 min at 60 °C and then extended with AMV reverse transcriptase for 1 h at 42 °C. The extension products were analyzed on 6% polyacrylamide/7 M urea gels. Two bands were identified (Fig. 4, lane B), indicating that there are two transcription initiation sites of approximately 11 and 57 bp upstream of the exon 1 (Fig. 2).

Table 1

<table>
<thead>
<tr>
<th>Ap-1</th>
<th>AP-1</th>
<th>AP-2</th>
<th>AP-2</th>
<th>ARE</th>
<th>ARE</th>
<th>ARE</th>
<th>ARE</th>
<th>Brn-3</th>
<th>Brn-3</th>
<th>C/EBP</th>
<th>C/EBP</th>
<th>C/EBP</th>
<th>C/EBP</th>
<th>MEF-1</th>
<th>MEF-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>CREB</td>
<td>CREB</td>
<td>E2F1</td>
<td>E2F1</td>
<td>EGR</td>
<td>EGR</td>
<td>EGR</td>
<td>EGR</td>
<td>ERE</td>
<td>ERE</td>
<td>Ets</td>
<td>Ets</td>
<td>Ets</td>
<td>Ets</td>
<td>MEF-1</td>
<td>MEF-1</td>
</tr>
<tr>
<td>GATA</td>
<td>GATA</td>
<td>GRE</td>
<td>GRE</td>
<td>HNF-4</td>
<td>HNF-4</td>
<td>HNF-4</td>
<td>HNF-4</td>
<td>IFN-1</td>
<td>IFN-1</td>
<td>IFN-1</td>
<td>IFN-1</td>
<td>IFN-1</td>
<td>IFN-1</td>
<td>MEF-1</td>
<td>MEF-1</td>
</tr>
<tr>
<td>NFATc</td>
<td>NFATc</td>
<td>NF-E1</td>
<td>NF-E1</td>
<td>NF-E2</td>
<td>NF-E2</td>
<td>NF-E2</td>
<td>NF-E2</td>
<td>NFk-B</td>
<td>NFk-B</td>
<td>NFk-B</td>
<td>Oct-1</td>
<td>Oct-1</td>
<td>Oct-1</td>
<td>Oct-1</td>
<td>Oct-1</td>
</tr>
<tr>
<td>Pπ1</td>
<td>Pπ1</td>
<td>PPAR</td>
<td>PPAR</td>
<td>PRE</td>
<td>PRE</td>
<td>PRE</td>
<td>PRE</td>
<td>RAR(DR3)</td>
<td>RAR(DR3)</td>
<td>RAR(DR5)</td>
<td>RAR(DR5)</td>
<td>RXR(DR1)</td>
<td>RXR(DR1)</td>
<td>RXR(DR1)</td>
<td>RXR(DR1)</td>
</tr>
<tr>
<td>Sp1</td>
<td>Sp1</td>
<td>SRE</td>
<td>SRE</td>
<td>Stat1</td>
<td>Stat1</td>
<td>Stat1</td>
<td>Stat1</td>
<td>Stat3</td>
<td>Stat3</td>
<td>Stat3</td>
<td>Stat3</td>
<td>Stat4</td>
<td>Stat4</td>
<td>Stat4</td>
<td></td>
</tr>
</tbody>
</table>

* The 54 transcription factors spotted on the array membrane were listed. Each factor was present in four spots in two rows and two columns. The first row (either A, C, E, G, I, K, and M) was DNA spotted normally. The second row (either B, D, F, H, J, L, and N) was DNA diluted 1:10. Columns 17 and 18 were spotted with the second consensus sequences of the transcription factors. The transcription factors with strong hybridization signals were indicated in bold in clear areas while those with slight or no hybridization signals were shown in shadowed areas.
Then, we attempted to define the transcription factors that were actually involved in activation of the 409 bp basal promoter fragment. The proteins bound to the 409 bp fragment were purified by the DNA pull-down assay according to the procedure described previously [14]. The 409 bp fragment of the pGL-3 Basic vector was cut out with restriction enzymes XhoI and NcoI. The reaction mixture was further incubated with restriction enzyme SacI to eliminate the XhoI site at the 3’ end of the vector fragment. The 409 bp fragment with the XhoI site at the 5’ end was then labeled with biotin using Bio-16-dUTP (Roche, Mannheim, Germany) and the Klenow fragment of DNA polymerase I (Invitrogen). Briefly, 0.2 nmol of the digested DNA fragment was labeled with 3 nmol Bio-16-dUTP, 125 units of Klenow 1 in labeling buffer (10 mM Tris–HCl, 10 mM MgCl2, 50 mM NaCl and 1 mM DTT, pH 8.0) and incubated at 15 °C for 2 h. Unincorporated Bio-16-dUTP was removed by a spin column and the labeling efficiency was determined by gel electrophoresis of the biotin-labeled fragment compared to the nonlabeled fragment. The biotin-labeled DNA fragment was then coupled to the M-280 Streptavidin magnetic beads (Dynal Biotech, Oslo, Norway) under the conditions described by the manufacturer. Nuclear extracts were prepared from HeLa cells using a slightly modified procedure [15], initially described by Ausubel et al. [16]. Protein concentrations of the extracts were measured by the Lowry method [17] using bovine serum albumin as a standard. The DNA coupled magnetic beads were incubated with 500 µg HeLa nuclear extracts for 2 h at 4 °C in the presence of an excess amount of dC-dI for competition of nonspecific binding. After several washes, binding proteins were disassociated from the DNA-coupled beads to a buffer containing 2 M NaCl.

The identities of the DNA pull-down proteins were determined by a novel procedure using the TranSignal Protein/DNA Array kit (Panomics, Redwood, CA). This array-based technology is a significant improvement over gel mobility-shift assays, and allows the functional analysis of dozens of eukaryotic transcription factors at a time. According to the manufacturer’s instruction, the binding proteins pulled down from the 409 bp promoter region were incubated with the TranSignal probe mix (a set of 54 biotin-labeled DNA binding oligonucleotides corresponding to the consensus sequences of 54 transcription factors, respectively) to allow the formation of DNA/protein complexes, which were then separated from the free probes by agarose gel electrophoresis. The probes in the complexes were

---

**Fig. 5.** Hybridization signals of the transcription factors bound to the 409 bp promoter fragment. The proteins from nuclear extracts of HeLa cells, which bound to the 409 bp (+123 to −286) fragment of the human GM3-synthase promoter were pulled down and their identities were determined using the TranSignal Protein/DNA Array (Panomics) as described in the text. The identities of the transcription factors with strong hybridization signals are indicated in Table 1.

**Protein/DNA Array kit** (Panomics, Redwood, CA). This array-based technology is a significant improvement over gel mobility-shift assays, and allows the functional analysis of dozens of eukaryotic transcription factors at a time. According to the manufacturer’s instruction, the binding proteins pulled down from the 409 bp promoter region were incubated with the TranSignal probe mix (a set of 54 biotin-labeled DNA binding oligonucleotides corresponding to the consensus sequences of 54 transcription factors, respectively) to allow the formation of DNA/protein complexes, which were then separated from the free probes by agarose gel electrophoresis. The probes in the complexes were
protein, is one of the best characterized stimulus-region (Fig. 2). CREB, the cAMP response element-binding of many target genes [21]. Although phosphorylation of 54 known transcription factors (Fig. 5 and Table 1). As a preliminary attempt, we identified eight binding/C0 proteins in this region, which were determined by the arrays /C0

For the past decade, most of the genes coding for glycosyltransferases involved in the biosynthesis of the major gangliosides have been cloned [18]. However, promoters for sialyltransferases involved in ganglioside synthases have not yet been examined except that of the rat and mouse GD3-synthase genes [13,19]. In the present study, we report the promoter sequence for the human GM3-synthase gene. We found that the proximal promoter region of −1 to −286 was sufficient for transcription of the gene (Fig. 3). As a preliminary attempt, we identified eight binding proteins in this region, which were determined by the arrays of 54 known transcription factors (Fig. 5 and Table 1). However, only CREB has the consensus sequence in this region (Fig. 2). CREB, the cAMP response element-binding protein [20], is one of the best characterized stimulus-induced transcription factors, which activates transcription of many target genes [21]. Although phosphorylation of CREB is enough to induce cellular gene expression in response to cAMP, additional promoter-bound factors are required for target gene activation [21]. Many transcription activators have been suggested to associate with CREB and CREB-binding protein (CBP) [21–23]. The other seven factors without consensus binding sites could be recruited as co-factors in the regulatory complexes, although we do not know whether they are associate with CREB. On the other hand, there were three Sp1 binding sites on the fragment (Fig. 2) but Sp1 was not pulled-down by the 409 bp fragment. Sp1 is a ubiquitous transcription factor that has been implicated in the activation of a large number of genes. Sp1 associates directly with members of the basal transcription machinery, involving the TATA box binding protein [24]. We showed here that Sp1 might not be involved in this TATA-less promoter. It should also be noted that there are consensus sites for transcription factors NFI and CNBP in the 409 bp region (Fig. 2). Because the TranSignal Array (Panomics) used in this analysis did not include these factors, the presence of the two factors in the pull-down samples could not be concluded in this report. Therefore, whether transcription factors Sp1, NFI and CNBP are involved in activation of the basal transcription of human GM3-synthase gene, as well as how the identified factors/co-factors function in the complex of processes during transcription activation, remain a major topic for future studies.

Acknowledgements

This study was supported by NIH Grant RO1NS11853.

References


