Characterization of the Mouse TFF1 (pS2) Gene Promoter Region

Tomoyuki Terada, Reina Sakagami, Yoshiaki Tabuchi, and Masatomo Maeda

Laboratory of Biochemistry and Molecular Biology, Graduate School of Pharmaceutical Sciences, Osaka University, Osaka, Japan. Received August 7, 2000; accepted October 23, 2000

Trefoil peptides (TFFs) with a unique trefoil domain(s) are presumed to function in protection and repair of the gastrointestinal epithelial layer. Three peptide family members are differentially distributed in the mouse gastrointestinal tract: TFF1/pS2 specifically in stomach, TFF2/SP mainly in stomach, pancreas and duodenum, and TFF3/ITF in intestine. We cloned and sequenced the mouse TFF1 gene 5'-upstream region by means of the genomic walking procedure. The cloned region was ligated to the luciferase reporter gene and then introduced into mouse gastric surface mucous GSM10 cells which express TFF1 and TFF2. The minimal promoter was located in the region containing the TATA-box between −39 and the transcriptional start site. Further upstream regions stimulated (−2192 to −1630 bp, −641 to −243 bp, −137 to −39 bp) and inhibited (−1630 to −641 bp, −243 to −137 bp) luciferase gene expression. These regions as well as short segments conserved in the mouse and human 5'-upstream sequences may be important for modulation of the mRNA level of the TFF1 gene.

Key words trefoil factor; TFF1; transcription; GSM10, pS2; stomach

Trefoil peptides (TFFs) are small and stable molecules secreted by the mammalian gastrointestinal tract. The term “trefoil”, meaning three leaves, is derived from their three intrachain loops (trefoil domain) that are maintained by disulfide bonds. Three family members are found in mammals: TFF1/pS2 and TFF3/ITF contain one trefoil domain, whereas TFF2/SP has two domains. It is suggested that the TFFs could be involved in protection and repair of the gastrointestinal tract through stimulation of cell migration and mucus polymerization. Actually, mice lacking TFF3 exhibited impaired mucosal healing and died from extensive colitis after polymerization. Actually, mice lacking TFF3 exhibited impaired mucosal healing and died from extensive colitis after oral administration of dextran sulfate. Furthermore, loss of paired mucosal healing and died from extensive colitis after polymerization. Actually, mice lacking TFF3 exhibited impaired mucosal healing and died from extensive colitis after oral administration of dextran sulfate. Furthermore, loss of TFF1 induced severe hyperplasia and dysplasia of the gastric mucosa, and the development of adenomas and carcinomas.

The results suggest that TFF1 is essential for normal differentiation of the gastric mucosa and that it functions as a tumor suppressor.

The three TFFs exhibit strict tissue-specific expression patterns: TFF1 in stomach, TFF2 in stomach and duodenum, and TFF3 in small intestine and colon. We are interested in the gene regulation of TFF1 from the viewpoint of stomach-specific transcription, as our previous studies focused on the sequence motifs in the upstream regions are discussed in comparison with those in the human gene.

MATERIALS AND METHODS

Reverse Transcriptase Polymerase Chain Reaction (RT-PCR) Total RNA was prepared from tissues of a 6-week-old male ddY mouse by the guanidine thiocyanate-CsCl method. cDNA was synthesized from total RNA (8 μg) with a First-Strand cDNA Synthesis Kit (Amersham-Pharmacia-Biotech) and an oligo d(T)12–18 primer, and then subjected to PCR to amplify specific nucleotide sequences for the three TFFs. The pairs of PCR primers used for TFF1 (S29 and S30), TFF2 (S26 and S27) and TFF3 (S3 and S4) are shown in Table 1. The PCR conditions were: preheating (94 °C, 3 min), followed by 35 cycles of denaturation (94 °C, 0.5 min), annealing (54 °C for TFF1, 58 °C for TFF2, or 62 °C for TFF3, 0.5 min), and extension (72 °C, 0.5 min). The 2nd PCR was carried out for TFF3 (54 °C annealing). The products were analyzed by agarose gel electrophoresis (1% (w/v) agarose (TaKaRa L03) in TAE-buffer), and visualized by ethidium bromide staining.

Total cellular RNA was prepared from GSM10 cells as above and then PCR analysis was carried out. The PCR conditions were essentially the same as for tissues except that the annealing temperature and cycle number for the three TFFs were 61 °C and 40 cycles, respectively. The GSM10 cells, the generous gift of Daiichi Pharmaceutical Co., Ltd., were cultured in Dulbecco’s modified Eagle medium/F12 medium (GIBCO BRL) containing 10% (v/v) fetal bovine serum (GIBCO BRL), 1% (v/v) insulin-transferrin-ethanolamine-selenium mixture (Wako) and 10 ng/ml epidermal growth factor (Wako) at 33 °C or 39 °C.

Genomic Walking The BamH I-digested mouse genomic DNA fragments (1 μg) attached to a Sau3AI cassette (1 μg) were subjected to PCR with Ampli Taq (Perkin-Elmer Cetus). The primer pairs for the 1st and 2nd PCR were S9 and C1, and S10 and C2, respectively (Table 1). The PCR conditions in both cases were preheating (94 °C, 2 min), followed by 25 cycles of denaturation (94 °C, 0.5 min), annealing (55 °C, 2 min), and extension (72 °C, 3 min). The PCR products were size-separated on an agarose gel, eluted with GeneClean III (Bio 101), and ligated into the pGEM T-Easy vector (Promega) with a ligation kit Ver. 2 (TaKaRa), and then sequenced by the dideoxy chain-termination method (Shimadzu DNA Sequencer Model DSO1000L).

Determination of the Transcription Start Site for TFF1 mRNA Single-stranded cDNA was synthesized from total RNA of GSM10 cells with phosphorylated primer S20 (Table 1) and circularized according to the manual for a 5’-Full Race Core Set (TaKaRa). Primer pairs S9 and S17, and S10 and S18 were used for the 1st and 2nd PCR with Ampli Taq, respectively. The coding region for TFF1 was also amplified...
was not only found in stomach but also in duodenum and rather broad distributions in the gastrointestinal tract; TFF2 scripts for the three TFFs in mouse by RT-PCR. TFF1 was...}

**RESULTS**

All other chemicals used were of the highest grade commercially available.

**RESULTS**

First, we determined the tissue distributions of the transcripts for the three TFFs in mouse by RT-PCR. TFF1 was only detected in stomach, while TFF2 and TFF3 showed rather broad distributions in the gastrointestinal tract; TFF2 was not only found in stomach but also in duodenum and pancreas (Fig. 1). Ileum and colon expressed TFF2 in very small amounts. TFF3 was distributed throughout the intestine (duodenum, ileum and colon). Such expression patterns of the TFFs were in good agreement with the previous observations. The S9 and S10 primers were also used for 5'-RACE together with S17–S20.

**Table 1. PCR Primers Used in This Study**

<table>
<thead>
<tr>
<th>Name</th>
<th>Amplification</th>
<th>Sequence</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>S3</td>
<td>TFF3 (antisense)</td>
<td>5’-TGTTGAGCTGCCCTTCTCCTGGA3’</td>
<td>Residues 32–13 from end codon</td>
</tr>
<tr>
<td>S4</td>
<td>TFF3 (sense)</td>
<td>5’-TTGCGTGTGCACTAGTAGGACACGAG3’</td>
<td>12 and 10 residues on 5'- and 3'-sides, respectively, from 1st codon</td>
</tr>
<tr>
<td>S9</td>
<td>Genomic walking of TFF1</td>
<td>5’GAAAGCTCGCGAAGCGACAG3’</td>
<td>Residues 59–40 from 1st codon</td>
</tr>
<tr>
<td>S10</td>
<td>Genomic walking of TFF1</td>
<td>5’ACCAACGAGGACACGATCA3’</td>
<td>Residues 32–11 from 1st codon</td>
</tr>
<tr>
<td>S11</td>
<td>TFF1 (sense)</td>
<td>5’-actCTAGAAAGCGACATTCTCTTCT3’</td>
<td>9 residues on both sides from 1st codon</td>
</tr>
<tr>
<td>S12</td>
<td>TFF1 (antisense)</td>
<td>5’-GAAAGCTCGCGAAGCGACAG3’</td>
<td>4 and 19 residues on 3'- and 5'-sides, respectively, from end codon</td>
</tr>
<tr>
<td>S17</td>
<td>5'-RACE of TFF1 (sense)</td>
<td>5’-CCCCGCGGAGAGGATATAATTG3’</td>
<td>Codon 37 (3rd letter) – codon 44</td>
</tr>
<tr>
<td>S18</td>
<td>5'-RACE of TFF1 (sense)</td>
<td>5’GCTGCCGCGGTGACCCG3’</td>
<td>Codon 45 (2nd letter) – codon 51 (2nd letter)</td>
</tr>
<tr>
<td>S20</td>
<td>5'-RACE of TFF1 (antisense)</td>
<td>5’-TTTGCTCTGATGGCCATGGGG3’</td>
<td>Codon 80 (1st letter) – codon 73 (3rd letter)</td>
</tr>
<tr>
<td>S26</td>
<td>TFF2 (sense)</td>
<td>5’-GAAAGACCTCCTCCTCCTG3’</td>
<td>Codon 24 – codon 30 (2nd letter)</td>
</tr>
<tr>
<td>S27</td>
<td>TFF2 (antisense)</td>
<td>5’-GGATGAAACACCAAGGGAC3’</td>
<td>Codon 121 (2nd letter) – codon 115</td>
</tr>
<tr>
<td>S29</td>
<td>TFF1 (sense)</td>
<td>5’-AATTTGCTCCTCCCGGTCA3’</td>
<td>Codon 43 – codon 50 (1st letter)</td>
</tr>
<tr>
<td>S30</td>
<td>TFF1 (antisense)</td>
<td>5’-ATTCTACCTCTTTTTATTCCTCGACGCC3’</td>
<td>Residues 178–154 from end codon</td>
</tr>
</tbody>
</table>

Primer sequences were amplified by PCR using primers S11 and S12. The PCR conditions were as follows: preheating (94 °C, 3 min), followed by 40 cycles of denaturation (94 °C, 0.5 min), annealing (61 °C, 0.5 min), and extension (72 °C, 0.5 min). The amplified fragments were ligated into the pGEM T-Easy vector and both strands of the cloned DNA were sequenced with the Silver Sequence DNA Sequencing System (Promega).

**Reporter Gene Assay**

The reporter plasmids were constructed by inserting the 5'-upstream region into the pGVB2 vector without a promoter or enhancer (Toyo Ink): pGV2192 (–2192–NcoI (+28)), pGV1630 (NheI (–1630)–NcoI (+28)), pGV641 (SacI (–641)–NcoI (+28)), pGV496 (–496–NcoI (+28)), pGV243 (PvuII (–243)–NcoI (+28)), pGV137 (ApalI (–137)–NcoI (+28)), and pGV39 (AgeI (–39)–NcoI (+28)). GSM10 cells (5×10^5 cells/collagen-coated i.d. 35 mm dish) were prepared 20 h before transfection. One of the reporter plasmids (4 μg) was introduced with pSV-GAL (0.05 μg) into the cells by the Ca⁺²-phosphate method, and the cells were cultured at 33 °C for 48 h. A cell extract was prepared with Reporter Lysis Buffer (Promega) and then luciferase activity was measured (Lumat LB9501, Berthold). The activity was normalized based on the β-galactosidase activity determined with an AURORA Gal-XE (ICN Pharmaceuticals, Inc.).

**Chemicals**

Restriction enzymes were obtained from New England Biolab, Takara Shuzo, Toyobo, or Nippon Gene. The PCR primers were purchased from Gibco BRL. All other chemicals used were of the highest grade commercially available.

**RESULTS**

First, we determined the tissue distributions of the transcripts for the three TFFs in mouse by RT-PCR. TFF1 was only detected in stomach, while TFF2 and TFF3 showed rather broad distributions in the gastrointestinal tract; TFF2 was not only found in stomach but also in duodenum and pancreas (Fig. 1). Ileum and colon expressed TFF2 in very small amounts. TFF3 was distributed throughout the intestine (duodenum, ileum and colon). Such expression patterns of the TFFs were in good agreement with the previous observations. The S9 and S10 primers were also used for 5'-RACE together with S17–S20.

![Fig. 1. Tissue Distributions of mRNAs for TFF1, TFF2 and TFF3](image-url)
region of the mouse TFF1 gene by genomic walking. The 700 bp sequence from the initiation codon of the amplified 2200 base pairs (bp) fragment is shown (Fig. 3), compared with the reported human sequence. The transcriptional start site was determined by the 5′-rapid amplification of cDNA ends (5′-RACE) technique with RNA prepared from GSM10 cells. The most frequent residue following the 5′-terminus of the phosphorylated primer (indicated by X in Fig. 4), i.e. the “adenine” residue 29 bases upstream from the initiation codon (ATG), was assigned as the transcriptional start site. This position is about 30 bp downstream from the potential TATA-box (Fig. 3). Such a location conformed with the tissue-specific eukaryotic promoter.

Harr plot analysis of the mouse and human upstream sequences demonstrated that there are a few conserved regions (Regions I, II and III) between −350 and the TATA-box (Fig. 5). These conserved regions could be important for the transcription of TFF1 gene. It must be noted that among common sequence motifs of human TFF genes (Motifs I—IV), the phosphorylated primer was assigned to a major transcriptional start site (open boxed adenine residue numbered +1). The initiation codon is indicated by a boxed dot.

Fig. 2. Expression of mRNAs for TFFs in GSM10 Cells
cDNAs from total RNA of GSM10 cells grown at 33°C (upper) and those from mouse stomach as a control (lower) were used. The primer pairs and other conditions were essentially the same as in the legend to Fig. 1. cDNAs of GSM10 gave slightly stronger background staining when primer pairs S3 and S4 were used to amplify the fragment for TFF3 (upper right). However, 290 bp TFF3 fragment, the size of which was indistinguishable from that for TFF2, could not be amplified.

Fig. 3. Comparison of the Nucleotide Sequences of the 5′-Upstream Regions of the Mouse and Human TFF1 Genes
The 5′-upstream region of the mouse TFF1 gene was amplified by the genomic walking procedure [this study], and the nucleotide sequence is shown together with the corresponding portion of the human gene. TATA-box sequences, transcriptional start sites and conserved motifs (Motifs I—IV) are indicated by bold letters on a dotted background, arrows and bold letters, respectively. The GATA sites [GAT(A or T)] are underlined. The underlines with asterisks are the GATA motifs pointed out recently. ERE and AP-1 sites are also shown as boxes (dotted and open, respectively). The chief cell-specific motif is double-overlined. The 5′-ends of the reporter genes are indicated by the vertical bars with open arrows and the residue numbers. Initiation codons are shown by bold letters. The nucleotide sequence data of mouse TFF1 gene have been submitted to DDBJ/EMBL/GenBank under accession number AB043538. Three conserved regions (Regions I—III) shown in Fig. 5 are indicated by underlines with double-arrowheads. Twelve residues are included on both sides of each region considering the analytical condition (span length).

Fig. 4. Determination of the 5′-end of TFF1 mRNA
The 5′-end of TFF1 mRNA was determined by the 5′-RACE technique as described in the text. The most frequent residue following the 5′-terminus of the phosphorylated primer was assigned to a major transcriptional start site (open boxed adenine residue numbered +1). The initiation codon is indicated by a boxed dot.
only Motifs III and IV are located in the conserved region (Region III) (Fig. 3). An estrogen responsive element found in the human gene was not conserved in the mouse gene. Furthermore, the AP-1 sites of mouse and human genes are located in Region I and its upstream, respectively.

To determine which part of the cloned 5' region affects the promoter activity, the luciferase reporter gene assay was carried out. pGV39 carrying the shortest fragment was introduced into GSM10 cells and this construct exhibited a significant level of luciferase activity (Fig. 6). Since a TATA-box was located at the 5'-end of the pGV39, the observed activity could be ascribed to a minimum promoter. pGV137 showed further enhanced luciferase activity, suggesting that the region between −137—−39 bp contains a proximal enhancer(s). pGV243 carrying the region between −243—−137 bp showed lower activity, possibly due to the presence of a negative regulatory element(s). The region of −496—−243 bp may contain a positive regulatory element(s), since pGV496 exhibited increased luciferase activity compared with pGV243. The further upstream −641—−496 bp region of the mouse gene showed a strong positive effect. The upstream −2192—−1630 bp and −1630—−641 bp regions could contain positive and negative elements, respectively, for transcription of the reporter gene.

**DISCUSSION**

TFF1 is found in the surface epithelium and the pits in mouse stomach, whereas TFF2 is found in the neck cells in the same organ. Although the GSM10 cells have properties of surface mucous cells, they express both TFF1 and TFF2, suggesting that they could not undergo complete differentiation or that the culture conditions would reflect some repair process during which the TFFs are co-expressed. Thus the studies on transcriptional regulation of TFF1 and TFF2 in this cell line could give valuable information for cell specific and context dependent gene expression.

Critical and common DNA sequences for transcriptional regulation of TFF1 genes could be located proximal to the TATA-box. Actually, conserved regions (Regions I—III)
were found by Harr plot analysis of 5'-upstream sequences of human and mouse TFF1 genes (Fig. 5). These three regions could be important for stomach specific transcription of TFF1 gene. Region I was included in the enhancer region of pGV496. Region II and Region III were mainly located in the silencer region of pGV243 and in the proximal enhancer of pGV137, respectively (Fig. 3). There are short sequence motifs (Motifs I—IV) in the human TFF genes. Among them, only Motifs III and IV are located in the conserved Region III. A site-directed mutagenesis study demonstrated that the Motif III and Motif IV contributed positively to transcriptional regulation of human TFF1 gene. In this regard, the roles of HNF3/FKHL proteins which recognized human Motif IV and triggered transcription of the TFF1 reporter gene are of interest from the viewpoints of TFF1 induction upon injury, inflammation and tumor development as well as tissue-specific TFF1 expression. However, it should also be mentioned that DNA methylation of the promoter region of TFF1 gene is another important factor for gene regulation associated with carcinogenesis.

The human gene corresponding to the mouse −496—−243 bp region contains a strong enhancer(s) responsive to estrogens, the epidermal growth factor, a tumor promoter, the c-Ha-ras oncoprotein and the c-jun protein. Motif I and Motif II, ERE and the AP-1 site could play important roles in the enhancer function in the human gene. However, the mouse gene does not have Motif I, Motif II or ERE (Fig. 3). We were consistently unable to detect estrogen responsiveness in our reporter gene assay (not shown), although the −496—−243 bp region had an enhancer function. It is also reported that the TFF1 mRNA level was unchanged upon administration of estrogen to ovariectomized rat. Transcription of TFF1 gene is upregulated in both chronic pancreatitis patients and TGFα overproduced mouse. Although the AP1 sites are differentially located between the mouse and human genes, these sites may have a potential role in such upregulation induced by growth factors through c-jun protein.

Other proteins important for stomach-specific transcription could be GATA DNA-binding proteins. Recently, it has been reported that GATA-6 has a positive effect on the TFF1 gene expression in human gastric cancer cells. Actually, many GAT(A/T) sites as potential GATA protein binding sites are distributed in the human upstream sequence. However, such sites are found less frequently in the mouse gene and their positions are apparently not conserved (Fig. 3). Thus, the mode of gene regulation of TFF1 by GATA proteins could have diverged between mouse and human. It would also be interesting to know whether or not the non-conserved region of mouse gene can participate in species-specific gene regulation.

Acknowledgements This research was supported in part by grants from the Ministry of Education, Science, Sports and Culture of Japan, and the Yakult Bio-Science Foundation.

REFERENCES