A complete $\alpha_{1,3}$-galactosyltransferase gene is present in the human genome and partially transcribed

Marion Lantéri$^{2}$, Valérie Giordanengo$^{2}$, Frédérique Vidal$^{3}$, Patrick Gaudray$^{4}$, and Jean-Claude Lefebvre$^{1,2}$

$^{1}$INSERM U526, IFR50, Faculté de Médecine, 06107 Nice Cedex 2, France; $^{2}$INSERM U470, Centre de Biochimie, Parc Valrose, 06108 Nice Cedex 2, France; and $^{3}$CNRS-UNS-UMR 6549, IFR50, Faculté de Médecine, 06107 Nice Cedex 2, France

Received on December 20, 2001; revised on May 29, 2002; accepted on July 2, 2002

The synthesis of Gal$\alpha_{1}$-$\beta$-Gal-terminated oligosaccharides ($\alpha$-Gal) epitopes has been interrupted during the course of evolution, starting with Old World primates. Partial sequences similar to the $\alpha_{1,3}$-galactosyltransferase ($\alpha_{1,3}$GalT) gene, which governs the synthesis of $\alpha$-Gal epitopes, have been detected in the human genome and were found to correspond to pseudogenes. We completed the sequence of the human $\alpha_{1,3}$GalT pseudogene present on chromosome 9 and found it to be organized like the murine $\alpha_{1,3}$GalT gene. In human cell lines and several normal and tumor tissues we detected truncated transcripts corresponding to this pseudogene. Considering these mRNAs, translation of an open reading frame containing the first four translated exons but missing the two catalytic exons could predict a truncated $\alpha_{1,3}$GalT polypeptide that should be enzymatically inactive. We show that transcription of human $\alpha_{1,3}$GalT is prematurely terminated at the level of a strong transcriptional stop signal in the middle of intron VII. We were able to reproduce this effect in vitro by subcloning the implicated DNA region upstream from a reporter cDNA. The premature transcriptional arrest of human $\alpha_{1,3}$-GalT gene leads to an ectopic splicing event and to the connection of a short intronic sequence downstream from translated exons. Finally, we show that these truncated transcripts are over-expressed in cell lines with modifications of $O$-glycans.

Key words: Alu element/gene expression/transcription stop/xenograft

Introduction

Most of the genes expressed in humans have counterparts in other mammals. The $\alpha_{1,3}$-galactosyltransferase ($\alpha_{1,3}$GalT) gene constitutes a noteworthy exception (Galili, 1999). Only partial sequences, similar to $\alpha_{1,3}$GalT genes of nonprimate mammals, have been detected in the human genome and are considered to correspond to pseudogenes because of multiple base deletions leading to premature translation stops (Joziasse et al., 1991; Larsen et al., 1990). Attempts to detect human $\alpha_{1,3}$GalT transcripts have remained unsuccessful as of today (Galili, 1999).

The $\alpha_{1,3}$GalT gene is responsible for the synthesis of $\alpha$-Gal epitopes in all mammals, except Old World primates. The expression of this gene has been completely lost in Old World monkeys, apes, and humans, thus enabling the production of significant amounts of natural antibodies to $\alpha$-Gal (Galili et al., 1984). These antibodies are strongly active, in the presence of complement (Good et al., 1992). This explains hyperacute rejection of xenografts, and in particular, these antibodies prevent projects of transplantation of pig organs to humans (Kobayashi and Cooper, 1999).

The $\alpha_{1,3}$GalT gene encodes for the UDP-Gal:Gal$\beta_{1,4}$GlcNAc $\alpha_{1,3}$-galactosyltransferase enzyme that catalyzes the transfer of a Gal residue, with an $\alpha_{1,3}$ linkage, on terminal lactosamine (Gal$\beta_{1,4}$GlcNAc-R) disaccharide beared by a glycoprotein or a glycolipid (Blanken and Van den Eijnden, 1985). The enzyme is not able to graft $\alpha$-Gal on Fuc$\alpha_{1,2}$Gal$\beta_{1,3,4}$GlcNAc-R that is a Fuc$\alpha_{1,2}$-substituted lactosamine (i.e., blood group H structure), as does the histo-blood group B transferase (B-transferase). Therefore the $\alpha_{1,3}$GalT gene has to be clearly distinguished from the B-transferase, which is strongly expressed in blood group B individuals.

The murine $\alpha_{1,3}$GalT gene is composed of nine exonic sequences, six of which (exons 4 to 9) being translated (Joziasse, 1992). Its structure diverges from that of the blood group transferase genes that make up seven translated exons with extensive similarities between the two last exons that correspond to the catalytic domain of this group.

Two $\alpha_{1,3}$GalT homologs have been described in the human genome. Both contained several frame-shift mutations and internal stop codons (Joziasse et al., 1991; Larsen et al., 1990). The first one is located on chromosome 12 and corresponds to a copy of an $\alpha_{1,3}$GalT cDNA, without intronic sequences (Joziasse et al., 1991; Larsen et al., 1990). This pseudogene has also been retrieved in apes and Old World monkeys (Galili and Swanson, 1991). The other one has been localized on chromosome 9 and could correspond to an authentic $\alpha_{1,3}$GalT gene, because intronic sequences (Joziasse et al., 1991) as well as one exonic sequence corresponding to the largest part of the catalytic domain of the enzyme (Larsen et al., 1990) have been detected in human genomic libraries. However, two frame-shift mutations present in the exonic sequence rendered this pseudogene nonfunctional (Larsen et al., 1990). Finally,
natural transcripts of α1,3GalT genes have widely been shown in nonprimate but never in human cells (Galili, 1999).

In the course of our study of the alterations of glycosylation patterns (hyposialylation and poly-N-acetylactosaminyl extensions) in human HIV-1-infected lymphoblasts (Lefebvre et al., 1994a,b), we have found transcripts similar to some of the coding exons of marmoset α1,3GalT, a New World monkey gene, overexpressed in latently HIV-1-infected lymphoblasts and also retrieved in human Expressed Sequences Tag (EST) databases. In the present article, we report the characterization of these transcripts and of the genomic sequences corresponding to these messengers. Moreover, we propose that Alu-like sequences may act as transcriptional terminator in this gene, accounting for part of the regulation of α1,3GalT gene expression in humans.

Results and discussion

We have previously reported modifications of O-glycans that appear on major lymphocytic surface glycoproteins CD45 and CD43 after HIV-1 infection. The modified O-glycans side chains, harbored by the latently HIV-1-infected CEM<sub>LAi</sub>/NP cells, are hyposialylated and elongated with poly-N-acetylactosamine chains (Lefebvre et al., 1994a,b).

Among various mechanisms, it was conceivable that an α1,3GalT enzymatic activity could be responsible for these changes. Actually, it has been demonstrated that expression of a transgene α1,3GalT could compete with endogenous α2,3-sialyltransferases without limiting the elongation of poly-N-acetylactosamine chains (Smith et al., 1990).

HIV-1 latently infected CEM cells are reactive with the plant lectin Griffonia simplicifolia specific for α-Gal epitopes

Although the hypothesis of expression of α1,3GalT in human cells was unlikely, we were led to explore it after the discovery of a weak differential reactivity of CEM<sub>LAi</sub>/NP versus parental CEM cells with Griffonia simplicifolia GS-1 isoelectin B4 specific for terminal α-Gal (Murphy and Goldstein, 1977) (data not shown). This surprising result has already been reported by others (Castronovo et al., 1989; Galli et al., 1985; Kagawa et al., 1988).

Detection of human α1,3GalT mRNA

To validate this observation, we decided to search for α1,3GalT mRNA by reverse transcription polymerase chain reaction (RT-PCR) on CEM<sub>LAi</sub>/NP cells lysates. Several pairs of degenerate primers were designed on the basis of conserved sequences between α1,3GalT of marmoset (Henion et al., 1994), ox (Joziasse et al., 1989), and mouse (Larsen et al., 1989). One of them, Galfive, allowed the production of a 200-bp fragment. The sequence of this fragment was nearly identical (94.5%) to the 5′ region of marmoset α1,3GalT cDNA. The Galfive amplified DNA fragment was then used to probe a northern blot carried out with poly(A)-rich RNA isolated from parental CEM and CEM<sub>LAi</sub>/NP cells. As shown in Figure 1, a ~1.6-kb species was detected in both cell lines but was much stronger in CEM<sub>LAi</sub>/NP cells. It is noteworthy that the size of these mRNA is much smaller than the size of α1,3GalT transcripts (3.6–3.9 kb) detected in mammals with a functional α1,3GalT gene (Joziasse et al., 1989; Smith et al., 1990).

Hence, the expression of truncated mRNA similar to α1,3GalT seemed to be correlated with changes in glycosylation in CEM<sub>LAi</sub>/NP cells.

An entire human α1,3GalT-inactivated gene (pseudogene) was retrieved in the GenBank database

Cloning the entire human α1,3GalT cDNA was performed by the screening of a cDNA library synthesized by retrotranscription of poly(A)-mRNA isolated from CEM<sub>LAi</sub>/NP cells, using a 5′-end 32P-radiolabeled oligonucleotide designed on the basis of the sequence of Galfive PCR fragments (see Materials and methods). Three positive clones were isolated and the longest (α1,3GalT-FRAG, 1.1 kb) was sequenced. An unique open reading frame of only 303 nt was identified, corresponding to the first four translated exons 4–7 of the marmoset α1,3GalT cDNA (Figure 2A), linked to an unknown untranslated region of 0.8 kb (unk3′) (Figure 2C). The unk3′ region is followed by a poly-A sequence. The first coding exon of this truncated cDNA matched exactly with the clone HGT-10 (Figure 2B) described by Joziasse et al. (1991)). The size discrepancy between the transcripts seen after northern blot (1.6 kb) and revealed after cloning (1.1 kb) can be explained by an unusually long 5′-untranslated region (UTR) of these truncated α1,3GalT as it has already been reported for the bovine α1,3GalT mRNA with a quite long 5′ UTR, totaling 468 bp (Joziasse et al., 1989). It is also known that the three exons composing the α1,3GalT 5′ UTR are differentially spliced depending on the tissues among various vertebrates, very likely leading to differential expression (Galili, 1999).

The recent availability of the human genome allowed us to retrieve a 172-kb-long clone (GenBank accession no. AL359644) from chromosome 9 containing all the six translated exons that matched almost exactly with those of marmoset (Figure 3). The size of each coding exon (89, 36, 66, 116, 138, 692 nt for exons IV–IX, respectively) was remarkably similar to the corresponding exon of mouse α1,3GalT (Joziasse et al., 1992), apart from exon VII (116 versus 102 nt). (Roman numerals are used to number human α1,3GalT exons by analogy with mouse exons [Arabic numerals], to take into account that the first three exons of human gene are hypothetical.) It is also noteworthy that the mouse coding exons IV–VI and VIII–IX are highly similar (82%) to the corresponding potentially coding human exons, whereas exons VII are <50% similar. This gene contains the partial α1,3GalT genomic sequences previously described as clones HGT-10 (Joziasse et al., 1991).
Human α1,3-galactosyltransferase gene

**A** human α1,3GalT mRNA (Genbank no. AF305838)

<table>
<thead>
<tr>
<th>Exon</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV</td>
<td>ATG AAT GTC AAA GGA AAA GTA ATT CTG TCA ATG CTG GTC TCA ACT GTG ACT ATT GTG</td>
</tr>
<tr>
<td>V</td>
<td>ACA GAA GGC TCT TCG TGG ATA TAT CAC TCA AAA AAC</td>
</tr>
<tr>
<td>VI</td>
<td>GAA</td>
</tr>
</tbody>
</table>

**B** clone HGT-10 (Joziasse et al., 1991)

<table>
<thead>
<tr>
<th>Exon</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV</td>
<td>AGAATCCGAATATACAACGTTGTTTCTCTGACTCGTCTCTCTGTAAGAGATCAAGAGAAGAAAAAGACCACTGGCTTGGTTGTTTCTCTGACTCGTCTCTCTGTAAGAGATCAAGAGAAGAAAAAGACCACTGGCTTGGTTG</td>
</tr>
</tbody>
</table>

**C** human α1,3GalT gene

Fig. 2. Sequence and localization of α1,3GalT-FRAG cDNA. (A) The four potentially translated exons IV–VII of α1,3GalT-FRAG and the corresponding sequence of marmoset α1,3GalT were aligned; only divergent bases were denoted. The position of the base 240 deletion is indicated by a vertical arrow, and the poly(A) signal is underscored. (B) The sequence of exon IV, underscored on panel A, matched exactly with the exon described in the clone HGT-10 flanked by intronic sequences (Joziasse et al., 1991). (C) The unknown 3′ UTR sequence (unk3′) of α1,3GalT-FRAG cDNA is localized on the intron VII of the human α1,3GalT gene. The sequence of this gene was extracted from clone RP11-162D16 (Genbank accession no. AL359644). The unk3′ sequence is linked to the potentially translated exon VII after an intermediate splicing event that excises a segment of 4361 bases after a premature transcriptional arrest in the middle of intron VII. The positions of stopGal and AluGal sequences are denoted by hooks. Splice donor and acceptors are framed. The splice acceptor at position 7920 corresponds to the expected end of intron VII in the case of a complete transcription of the human α1,3GalT gene.

1991) and HGT-1 (Larsen et al., 1990) that correspond to exons IV and IX, respectively.

These potentially coding sequences are interrupted by five successive introns of 4.6, 2, 1, 7.9, and 4.5 kb (Figure 3). The unknown 3′ region (unk3′) of 0.8 kb of the α1,3GalT-FRAG cDNA was retrieved in the middle of intron VII (Figure 2C). Using the BLAST program, we searched for alignments between α1,3GalT-FRAG and Expressed Sequence Tag databases. This
search allowed the discovery of several clones that matched almost exactly (95–100% identities) with the 3’ end of α1,3GalT-FRAG (Table I). Through the IMAGE Consortium, it has been possible to receive one of these clones (Genbank accession no. R24770), originated from an infant brain cDNA library constructed by Bento Soares and M. Fatima Bonaldo (University of Iowa). The insert sequence (1.2 kb) matched exactly (100% identities) with our α1,3GalT-FRAG and
allowed us to confirm the presence in human cells of a particular transcript of α1,3GalT containing the first four translated exons elongated with the unk3′ region is polyadenylated downstream from the 3′ end exon that encodes for a transmembrane domain, and thus to the production of a functional secreted form of the receptor.

The DNA structure downstream from the intronic poly(A) signal is propitious to a transcriptional stop

Comparison between human α1,3GalT-FRAG cDNA and α1,3GalT gene organization led to the conclusion that the splicing of exon VII with the unk3′ sequence in intron VII is the result of the excision of a 4.3-kb intron. However, according to the splicing events that generate a complete α1,3GalT cDNA in nonprimate mammals, the size of the corresponding intron might be 7.8 kb. No obvious mutations within the donor or the acceptor consensus sites of splicing could be revealed in the genomic sequence in these regions (Figures 2C and 3). Therefore, one way to explain this observation was that the transcription of α1,3GalT gene was terminated using a transcriptional-stop signal and/or a polyadenylation signal present within intron VII.

Little is known about either the signals for termination or the process involved for most eukaryotic RNA polymerases. However, experiments performed in Xenopus oocytes suggest that formation of secondary structures is more important in that process than the exact sequence (Nishikura et al., 1982). To check for the presence of a sequence that could be involved in forming secondary structures, we looked at the genomic sequence located downstream from the 3′ sequence of α1,3GalT-FRAG. By computer analysis, we found significant similarities between this sequence and AluSx sequences belonging to the Alu family (Batzer et al., 1996), (AluGal, Figures 2C and 4).

The Alu sequence corresponds to an human family of repetitive sequences related to 7SL RNA. These sequences were composed of a tandem duplication of 130 bp, with an unrelated insertion of a 31-bp sequence in the right half of the dimer. The individual members of the Alu family are rather related than identical (Rowold and Herrera, 2000). Some of them appeared to be transcribed in vivo by the RNA polymerase III, thus leading to transcriptional interferences. A variety of properties have been found for this family, and its ubiquity has led to many suggestions concerning its function. In fact, repeated sequences seem to be modular components of regulatory elements whose effects depend on the influence of flanking sequences. One of the proposed functions is that due to their ability to form hairpin structures, some of these sequences could act as transcriptional terminators and in that way could be involved in the control of gene expression (Marais and Sarrowa, 1995). To test this hypothesis in the case of α1,3GalT

<table>
<thead>
<tr>
<th>GenBank accession no.</th>
<th>Tissue</th>
<th>Sequence length</th>
<th>Identity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AW973637</td>
<td>Colon tumor metastasis</td>
<td>448</td>
<td>99</td>
</tr>
<tr>
<td>AF966270</td>
<td>Pooled germ cell tumors</td>
<td>342</td>
<td>99</td>
</tr>
<tr>
<td>AA489740</td>
<td>Pooled melanocyte, fetal heart, pregnant uterus</td>
<td>432</td>
<td>99</td>
</tr>
<tr>
<td>AI588750</td>
<td>Prostate</td>
<td>313</td>
<td>98</td>
</tr>
<tr>
<td>AW292272</td>
<td>Brain</td>
<td>212</td>
<td>100</td>
</tr>
<tr>
<td>AI468141</td>
<td>Glioblastoma</td>
<td>221</td>
<td>100</td>
</tr>
<tr>
<td>AA554133</td>
<td>Colon</td>
<td>244</td>
<td>100</td>
</tr>
<tr>
<td>AA435778</td>
<td>Testis</td>
<td>207</td>
<td>100</td>
</tr>
<tr>
<td>AW075560</td>
<td>Wilm’s tumors (pooled)</td>
<td>202</td>
<td>99</td>
</tr>
<tr>
<td>AW080178</td>
<td>Endometrial adenocarcinoma</td>
<td>232</td>
<td>95</td>
</tr>
<tr>
<td>R24770</td>
<td>Infant brain</td>
<td>349</td>
<td>100</td>
</tr>
<tr>
<td>AW204146</td>
<td>Germ cell</td>
<td>188</td>
<td>100</td>
</tr>
<tr>
<td>AI634808</td>
<td>Kidney</td>
<td>272</td>
<td>99</td>
</tr>
<tr>
<td>N67359</td>
<td>Fetal cochlea</td>
<td>293</td>
<td>96</td>
</tr>
<tr>
<td>BF930379</td>
<td>Nervous tumor</td>
<td>420</td>
<td>99</td>
</tr>
<tr>
<td>BG700213</td>
<td>Hippocampe</td>
<td>197</td>
<td>96</td>
</tr>
<tr>
<td>BF914104</td>
<td>Uterus tumor</td>
<td>155</td>
<td>100</td>
</tr>
</tbody>
</table>

*Aligned by the BLAST program.

*Percentage of identities.
gene, we cloned the genomic fragment containing the Alu-like sequence (stopGal) in a reporter plasmid between the \( \beta \)-galactosidase gene and its promoter.

The 865-bp sequence around the poly(A) signal of intron VII functions as a transcriptional stop

A fragment of 865 bp (stopGal) was synthesized from total DNA extracted from CEM cells using PCR and a pair of primers flanking the AAATAAAAA signal from the –83 to the +782 nucleotide (Figure 2C). The stopGal fragment was then ligated into the pCMV\( \beta \) plasmid between the promoter CMV IE and the \( \beta \)-galactosidase reporter gene (see Materials and methods). Stop-sense and Stop-antisense constructs were obtained. These recombinants and control plasmids were transfected in human embryonic kidney (HEK)-293 cells. As shown in Figure 5, the activity of \( \beta \)-galactosidase was quantified from lysates of transfected cells, using ortho-nitrophenyl \( \beta \)-D-galactopyranoside (ONPG) substrate. When compared to corresponding controls, the stopGal sequence was able to block almost completely the activity of the reporter gene when it was constructed in the sense orientation. Moreover, although significant, its blocking efficiency was much more moderate when it was constructed in the antisense orientation (Figure 5).

These results could be explained by the orientation of the Alu sequence inserted in the constructs. In the sense construction, in addition to secondary hairpin structures, transcriptional interferences could arise due to the presence of potential RNA polymerase III promoter retrieved in Alu sequence. This observation is reminiscent of what is found in natural transcriptional stop of some viruses (Lee et al., 1981).

An important question arising in view of the widespread tissue distribution of \( \alpha \)1,3GalT-FRAG transcripts concerns the role that these mRNA could play, especially in the case of their overexpression discovered in CEM LAI/NP cells (Figure 1). Indeed, the CD43/CD45 molecules of CEM LAI/NP cells are severely hyposialylated and this status is associated with a (apparently paradoxical) normal expression of concerned sialyltransferases (Giordanengo et al., 1997).

Even if any glycosyltransferase activity cannot be expected from \( \alpha \)1,3GalT-FRAG, a short protein (100 amino acids) containing the signal peptide and the almost entire stem region of \( \alpha \)1,3GalT can be translated from the corresponding mRNA. This putative protein could thus be addressed to the Golgi apparatus just like an authentic glycosyltransferase. It is generally admitted that the concentration of these enzymes is well regulated, implicating sorting, anterograde and retrograde transport vesicles (Bannykh et al., 1998; Klumperman, 2000). On this basis, it
could be hypothesized that α1,3GalIT-FRAG proteins interfere in the trans-Golgi network compartment with functional enzymes, such as sialytransferases, by way of their conserved stem domain in the manner of negative transdominants. This could account for results indicating that expression of a transgene α1,3GalIT could compete with endogenous α2,3-sialyltransferases (Smith et al., 1990). To pursue the hypothesis, the hyposialylation status of CD43/CD45 molecules of CEM_{LA/NP} could be explained by the overexpression of nonfunctional but competitive α1,3GalIT-FRAG peptides.

**Materials and methods**

**Cells and cultures**

The cell line HEK-293 (ATCC CRL-1573) was maintained in Dulbecco’s modified Eagle’s medium (Life Technologies, Rockville, MD) containing 7% (v/v) fetal calf serum (FCS) (Whittaker Bioproducts, Walkersville, MD), and transfected with the Transfection MBS Mammalian Transfection Kit (Stratagene, La Jolla, CA). The HIV-1 latently infected CEM_{LA/NP} cell line bearing hyposialylated surface glycoproteins was described elsewhere (Lefebvre et al., 1994a). The parental CCRF-CEM (ATCC CCL-119) and CEM LAI/NP cell lines were propagated in RPMI 1640 medium (Life Technologies) containing 5% (v/v) FCS.

**PCR**

All the PCR reactions were carried out in buffer added with 10% (v/v) dimethyl sulfoxide, and the amplicons were subcloned into the plasmid pCR II-TOPO (TOPO TA Cloning Kit, Invitrogen, Faraday, CA). PCR primers pairs were: Galfive: 5′-ATGAAATTGCTAARGAAAAATRAT and 5′-CTTSTTGGATAATYTGGRTCC (R = A + G; S = G + C; Y = C + T) for detection of α1,3GalIT sequences (Genbank S71333, nt 1–23 and nt 181–202, respectively) by RT-PCR; and stopGal: 5′-CTAAAGACCACTGTAAGCTACTAA and 5′-CTCAT-TTAGTCTCTGCAAGACG for the cloning of the transcriptional stop sequence from total genomic DNA (Genbank AL359644, nt 102,539–102,560 and 103,380–103,404, respectively). The probe was 5′-end radiolabeled using T4 Polynucleotide Kinase from New England Nuclear and hybridized with the double strand oligonucleotide generated by hybridization of the 5′-AGCCTGTAACCGGAGCTCG and 5′-GATCCGACGTC-CCGGTACCA (the Sma I linker underscored) oligonucleotides. It was thereafter possible to ligate the Xho I–stopGal–Sma I fragment into the Sma I unique downstream from the SV40 intron. This probe was used in the orientation expected to obtain the constructs called STOP-antisense were nevertheless tested as indicated.

To insert stopGal in sense orientation, the SV40 intron of pCMVβ flanked by the unique restriction sites Xho I and Sma I was excised and replaced by appropriated stopGal. For that purpose, a Sma I site was first created between the Hind III and Bam HI sites located in the multiple cloning site of the intermediate stopGal/pCR II-TOPO construct, by insertion of a double strand oligonucleotide generated by hybridization of the 5′-AGCCTGTAACCGGAGCTCG and 5′-GATCCGACGTC-CCGGTACCA (the Sma I linker underscored) oligonucleotides. This probe was used during hybridization of the double strand oligonucleotide generated by hybridization of the 5′-AGCCTGTAACCGGAGCTCG and 5′-GATCCGACGTC-CCGGTACCA (the Sma I linker underscored) oligonucleotides. It was thereafter possible to ligate the Xho I–stopGal–Sma I fragments into pCMVβ in the orientation expected to obtain the constructs called STOP-sense. The control pCMVβ without the SV40 intron (pCMVβ delta SV40) was also prepared.

HEK-293 cells were transiently transfected, on six-well plates using MBS mammalian transfection kit (Stratagene). To control transfection efficiency, the expression vector GL2-promoter (Promega) containing the luciferase gene was cotransfected (3 µg/well) with each of the β-galactosidase gene constructs (3 µg/well).

**Northern analysis**

Total RNA were extracted according to Guanidine Isothiocyanate-Technic, and Poly(A)-rich RNA were selected as described elsewhere (Giordanengo et al., 1997). Poly(A)-rich RNA were electrophoresed on denaturing 1.2% agarose gel and transferred in 20× NaCl/Cit to a hybrid N nylon (Amersharm Corp). The probe was [α-32P]-random-labeled Galfive amplicons. The detection of the glyceraldehyde 3-phosphate-dehydrogenase mRNA were used as internal control.

**β-Galactosidase and luciferase assays**

Colorimetric assays were conducted with soluble ONPG (Sigma) on HEK-293 cells transfected in 24-well plates (5 × 10^5 cells/well) transfected with STOP-sense, STOP-antisense, and control plasmids (2 µg/well). A colorimetric assay conducted with soluble ONPG (Sigma) on HEK-293 cells transfected in 24-well plates (5 × 10^5 cells/well) transfected with STOP-sense, STOP-antisense, and control plasmids (2 µg/well) was performed. A 48-h period of culture, cells were disrupted in 100 µl lysis Tris buffer (0.25 M, pH 7.8; 0.5% Nonidet P-40; 1% Triton) for 15 min. From supernatant of each well, a 0.5 µl aliquot was taken off to protein quantification, and 0.007 µl were transferred in a new plate. Then, successively and in the same order, to each well was added 1 ml of 2-ME buffer (0.1 M NaHPO4, pH 7.5; 10 mM KCl; 1 mM MgSO4; 0.1% 2-ME) for 15 s, 200 µl ONPG solution (0.1 M NaHPO4, pH 7.5; 4 mg/ml ONPG) for 15 s, and finally 500 µl 1 M Na2CO3 to stop the reaction. Two days after transfection, β-galactosidase...
activity was evaluated by measuring the release of o-nitrophenol at 420 nm. Luciferase assay was performed according to Brasier et al. (1989). Colorimetric and luminescent points were normalized according to the amount of protein evaluated in each sample with the BCA Protein Assay Reagent (Pierce, Rockford, IL). All transfections were performed in triplicate for each point and repeated at least three times.

Acknowledgments

This work was supported by institutional grants from the Institut National de la Santé et de la Recherche Médicale (INSERM) and the Association pour le Développement du Diagnostic des Maladies Virales (ADDMV). The sequence reported in this article has been deposited in the GenBank database (accession no. AF305838).

Abbreviations

EST, Expressed Sequence Tag; FCS, fetal calf serum; HEK, human embryonic kidney; ONPG, ortho-nitrophenyl β-D-galactopyranoside; RT-PCR, reverse transcription polymerase chain reaction; SDS, sodium dodecyl sulfate; UTR, untranslated region.

References


