Genomic organization and promoter activity of glucosidase I gene

Farhat A.Khan, Gayatri M.Varma and Inder K.Vijay

Department of Animal and Avian Sciences, University of Maryland, College Park, MD 20742–2311, USA

Received on December 11, 1998; revised on January 26, 1999; accepted on January 26, 1999

1To whom correspondence should be addressed

Glucosidase I initiates the processing of asparagine (N-)linked glycoproteins by removing the distal α1,2-linked glucosyl residue of the tetradecasaccharide Glc3Man9GlcNAc2. The gene encoding this enzyme was isolated and its structural organization and promoter activity determined. The major transcript for glucosidase I on northern blot appeared to be 3.1 kb; Southern blotting and DNA sequencing indicated the size of the gene to be 6.8 kb, comprising four exons separated by three introns. The first exon encodes the cytoplasmic tail and transmembrane domain; the fourth encodes the putative catalytic domain of the enzyme. Exon–intron junctions are flanked by consensus splice donor and acceptor sequences. Transcription initiation sites were mapped by primer extension, ribonuclease protection assay and RT-PCR analysis. Primer extension results showed multiple initiation sites at −150, −156, and −272 bp relative to the translation initiation codon ATG. Sequence analysis of 5′ flanking region showed no canonical TATA box, a high GC content, Sp1 and ETF binding sites (typical of a housekeeping gene promoter). Also noteworthy, the promoter region contains several generic STAT factor binding sites, one nearly perfect, and two half GR binding elements. Other cis-acting elements recognized by transcription factors such as AP-2, NF-κB, estrogen receptor, and progesterone receptor (PR) were also present in the putative promoter region. To determine the promoter activity, a construct encompassing the region between −2114 to −5 bp of the putative promoter was ligated to the chloramphenicol acetyltransferase (CAT) reporter plasmid and transiently transfected into COS 7 cells. CAT assay results clearly show transcriptional activity of the promoter.

Key words: N-glycosylation/glycosidase/glycosidase/glycoprotein/processing

Introduction

The biosynthesis of the carbohydrate precursor of asparagine-linked glycoproteins is initiated by a stepwise, dolichol phosphate-linked assembly of Glc3Man9GlcNAc2 unit in the endoplasmic reticulum (ER) and its transfer en bloc to the asparagine residue in the sequon, Asn-X-Ser/Thr in the nascent polypeptide on the lumenal face of the ER. This is followed by co- and posttranslational processing reactions that begin with the excision of the glucosyl residues by ER-localized glucosidases (Kornfeld and Kornfeld, 1985). The incipient glycoprotein may not undergo any additional modification, fold, and be transported as such through the secretory pathway. More commonly, depending on the structural features of the underlying polypeptide, extensive modifications of the carbohydrate moiety of the glycosylated polypeptide occur in the rough ER and Golgi compartment; these involve the action of processing mannosidases, present in both RER and the Golgi, and glycosyltransferases within the different subcompartments of the Golgi and trans-Golgi network, culminating in the biosynthesis of high mannose, complex and hybrid glycoproteins that are transported to their intra- or extracellular sites of destination (Moremen et al., 1994).

Glucosidase I catalyzes the onset of processing reactions by clipping the terminal α1,2-linked glucosyl residue even as early as the polypeptide is still undergoing synthesis on the polysomes and translocating into the lumen of the ER (Atkinson and Lee, 1984). Later, glucosidase II sequentially removes the internal α1,3-linked glucosyl residues (Kornfeld and Kornfeld, 1985; Moremen et al., 1994). An alternative pathway utilizes an endomannosidase that can remove Glc3-1Man fragment from the newly synthesized glycosylated polypeptide (Rubas and Spiro, 1987; Fujimoto and Kornfeld, 1991). A model has been proposed according to which the action of the processing glucosidases, after the removal of two of the three glucosyl residues, allows the binding of the Glc3ManGlcNAc2-polypeptide by the lectin-chaperones calnexin and calreticulin followed by other chaperones such as BiP, for the proper folding and maturation of glycoproteins during transport in the secretory pathway (Kim and Arvan, 1995; Rodan et al., 1996; Oliver et al., 1997). For quality control purposes, this model also invokes the additional participation of an ER-localized UDP-Glc: glucoprotein glucosyltransferase that can reglucosylate a deglucosylated but partially folded or malformed glycoprotein and allow a rebinding by calnexin and calreticulin to ensure proper folding (Sousa et al., 1992; Rodan et al., 1996). The folded glycoprotein is perceived to dissociate from the lectin-chaperones, as glucosidase II cleaves the innermost α1→3Glc before the egress of the glycoprotein from the ER.

Because of the critical role in initiating the processing and maturation phase during the biosynthesis of N-linked glycoproteins, glucosidase I may be regarded as a potential candidate for the regulation of their biosynthesis. The study of protein N-glycosylation during the ontogeny of the mammary gland has been a long term focus of our laboratory. Biochemical investigations on glucosidase I in our and others’ laboratories have concentrated on the purification and characterization of the enzyme from different animal tissues and yeast (Bause et al., 1986, 1989; Shailubhai et al., 1987). In the mammary gland, it is a tetrameric N-linked glycoprotein with a subunit molecular mass of 85 kDa. Its topology was defined to have a lumenally-oriented catalytic domain, a transmembrane segment and a short, cytoplasmic tail (Shailubhai et al., 1991). The active site of the enzyme is contained within a 39 kDa segment, obtained by tryptic digestion of the purified protein. Glycosylation is not absolutely required for its catalytic activity. A tagging strategy delineated the presence of a cysteine residue at the substrate binding region within the active site of the enzyme (Pukazhenthhi et al., 1993).
Substrate protection followed by chemical modification of the side chains of amino acids on the surface of the enzyme also implicated the involvement of an arginine and a tryptophan residue at the oligosaccharide binding site (Romaniouk and Vijay, 1997). The data on the cDNA of glucosidase I, recently cloned from the human brain, is in agreement with the proposed topology, and predicts the enzyme to be a type II membrane protein (Kalz-Fuller et al., 1995). Conspicuously, of the three cysteines at positions 136, 601, and 816 in the polypeptide predicted from the cDNA sequence of the human brain enzyme, cysteine 601 is flanked by arginine and tryptophan and is within the active site-containing domain. These observations were combined with the information in the protein sequence data banks for the putative glucosidase I gene in C.elegans to propose a nine amino acid motif, ERHLDLRCW, within the catalytic cleft, as constituting the active site region of glucosidase I that binds the oligosaccharidyl moiety of the immature glycoprotein for processing (Pukazhenthhi et al., 1993; Romaniouk and Vijay, 1997). A recent comparison of the product of the yeast gene, CWH41, with the human glucosidase I indicated that the gene encodes glucosidase I (Romero et al., 1997).

To meet the demands of the secretion of large amounts of soluble and membrane glycoproteins into milk, there is an enormous increase in glycoprotein synthesis in the mammary gland during lactation. Several key enzymes of the dolichol cycle and glucosidase I are coordinately upregulated during the ontogeny of the gland in the mouse, with peak activities observed at mid-lactation (Vijay and Oka, 1986; Shailubhai et al., 1990). Using an explant culture system to mimic lactogenesis, in a previous study, we showed that glucosidase I is regulated during the hormonally modulated differentiation of the explants of the rat mammary gland (Shailubhai et al., 1990). The maximum activation was obtained by the synergistic action of the lactogenic hormones, insulin, prolactin, and hydrocortisone in the culture medium. To further elucidate the molecular mechanisms involved in the developmental and a potential hormonal regulation of the enzyme in the mammary gland, it is important to isolate its gene and procure the promoter. This study describes the structure, organization, size of the transcript, and an analysis of the promoter of the gene for the enzyme in the mouse.

Results

Genomic Southern analysis

To estimate the size and copy number of mouse glucosidase I gene a Southern blot of mouse genomic DNA digested with restriction enzymes (EcoRI, BamHI, HindIII, KpnI, and PstI), was probed with the partial cDNA of rat ovary glucosidase I under conditions of high stringency (Figure 1). The results obtained indicate the size of glucosidase I gene to be between 7 kb and 16 kb. Isolation and characterization of the genomic clone indicated the actual size of the gene to be around 6.8 kb. Multiple bands observed in some of the lanes of genomic southern were most likely due to the cleavage at restriction sites of EcoRI, BamHI, and KpnI in the exons or HindIII and PstI in intronic sequences and not because of the presence of multiple copies of glucosidase I gene (Samaha et al., 1996).

Isolation and characterization of mouse glucosidase I gene

In order to study the regulation of glucosidase I in the mouse mammary gland, it was necessary to isolate the genomic sequence with the promoter. The partial cDNA of rat ovary enzyme was used to screen a λ EMBL3 SP6/T7 mouse genomic library. A screening of 1 million plaques generated five clones: λ 4, 5, 10, 12, and 16. To determine whether one of the clones contained the complete gene, the isolated clones were further subjected to Southern analysis with 5′ and 3′ oligonucleotide probes from the rat cDNA clone as given in Materials and methods. Clone 5 showed positive fragments of 4.2 and 2.6 kb on autoradiogram for 5′ and 3′ oligonucleotide probes, respectively, indicating that it was the longest clone. These DNA inserts were subcloned in pGEM-7Zf(-) vector and designated as 5a and 3a for 5′ and 3′ region of the gene for further analysis.

Genomic organization of mouse glucosidase I gene

DNA sequencing of subclones 5a and 3a, and a comparison of the resulting sequence with human glucosidase I cDNA (Kalz-Fuller et al., 1995) indicated that the entire mouse glucosidase I gene was present in clone 5; it is ~6.8 kb in size, comprising four exons separated by three introns. A schematic diagram of the gene is

![Fig. 1. Southern analysis of mouse genomic DNA. Mouse genomic DNA (10 µg/lane) was either not digested or digested with EcoRI, BamHI, HindIII, KpnI, and PstI; transferred onto nitrocellulose membrane; and probed with the partial cDNA of rat glucosidase I.](image-url)
Mouse glucosidase I gene

Fig. 2. Schematic representation of mouse glucosidase I gene. Location of BamHI site, multiple transcription start sites (arrows), exon 1–4 (solid boxes, coding region), intron A,B,C (open boxes), translation initiation (ATG) and termination (TAC) codons, polyadenylation site and 5' and 3' untranslated regions (hatched boxes) are shown. The positive genomic clone λ5, subclones 5a and 3a, and the scale in kilobase pairs are defined by lines.

presented in Figure 2. As shown in Table I, the first exon includes the 5' noncoding region and translation initiation codon ATG, while the last exon contains the largest protein coding region (>1.7kb), translation termination codon TAC and the 3' untranslated region including the polyadenylation signal AATAAA. The introns range in size from 92 to 150 bp. The exon–intron boundaries were determined by DNA sequencing and by comparison with human and rat cDNA clones. All exon–intron junctions follow the GT-AG rule of Breathnach and Chambon (Breathnach and Chambon, 1981) and the consensus sequence proposed by Mount (Mount, 1982). Nucleotide and amino acid comparison of human glucosidase I with mouse enzyme shows 83% and 84.8% identity, respectively.

Analysis of 5' flanking region

An examination of the 5' noncoding region reveals that it lacks a canonical TATA box but contains other well characterized features of a promoter, shown underlined and in boldface letters in Figure 3. Noteworthy features of the promoter include a relatively high GC content (71%) in the proximal region (from nucleotide –1 to –176) and the presence of numerous cis-acting elements including multiple transcription initiation sites (see below). These are typical features of the promoter of housekeeping genes (Kageyama et al., 1989). The 5' noncoding region of glucosidase I contains six binding sites for the zinc finger transcription factor Sp1; four of these sites are located in the proximal region at positions –2, –84, –445, and –646, one at position –877 and one in the distal region at position –1869. Sequence element for the binding by epidermal growth factor receptor-specific transcription factor (ETF) is located at position –22, –71, –116, –1201, –1839, –1862, –1880, –1901, –1910, –1908, and –2250. ETF is known to stimulate transcription of promoters lacking a TATA box (Kageyama et al., 1989). There are three AP2 binding elements at positions –2, –204, –650, two of which overlap with Sp1 sites. The half sites of hormone responsive elements for estrogen receptor are located at position –41, for progesterone (PR)/estrogen receptor at –384 and for glucocorticoid receptor (GR) at positions –1145,-1154, –1376, and –1390 in the 5' flanking region. The generic sequence for binding by STAT (signal transducer and activator of transcription) factor, i.e., TTNNNNNAA, where N stands for any nucleotide (Darnell, 1997), occurs in the 5' flanking region of the gene at positions –680,-852, –968, –1136, –1130, –1446, –1528, –1704, and spanning a region of 2 kb from translation initiation site. A CCAAT box is located at position –874 from translation initiation codon ATG. Another noteworthy feature of this promoter region is that it contains 11 repeats of GT dinucleotide at position –495 and 10 repeats of tetranucleotide ATTT, at position –553. The (ATTT)10 contains the direct repeat motif ATT found in the promoters of histones H3 and H4 for binding by HinF-A (van Winjen et al., 1988). Polymorphic forms of such motifs, designated as variable number of tandem repeats (VNTR) or microsatellites, have been found in the DNA of different individuals (Nakamura et al., 1987).

Table I. Exon/intron organization of the mouse glucosidase I gene

<table>
<thead>
<tr>
<th>Number</th>
<th>Exon Location in cDNA (nt)</th>
<th>Size (bp)</th>
<th>Sequence at exon–intron junctiona</th>
<th>Intron Location Size (bp)</th>
<th>Amino acid interruptedb</th>
<th>Typec</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5'NC-346</td>
<td>5'NC+346</td>
<td>TCACAG gtaacc......tcaccag</td>
<td>GACTGA</td>
<td>92</td>
<td>G-116</td>
</tr>
<tr>
<td>2</td>
<td>347–577</td>
<td>230</td>
<td>CCCCAAG gtcagg......accagg</td>
<td>GCCTCA</td>
<td>479</td>
<td>Q-192</td>
</tr>
<tr>
<td>3</td>
<td>578–775</td>
<td>197</td>
<td>TGGCAG gtaa......ttcag</td>
<td>CTACAA</td>
<td>150</td>
<td>S-259</td>
</tr>
<tr>
<td>4</td>
<td>776–3'NC</td>
<td>1726+3'NC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

aUppercase and lowercase letters represent exon and intron, respectively.
bWhen splicing occurred between codons, the amino acid 5' to the splice site is indicated.
cIntron type is according to Sharp (45).

NC, Noncoding region.
Fig. 3. Nucleotide sequence of the 5′ flanking region of the mouse glucosidase I gene. The sequence of 5′ flanking region of the mouse glucosidase I gene is shown (GenBank accession number AF001719). Numbering of nucleotides is relative to adenosine (taken as +1) of translation initiation codon ATG. Primer used for primer extension is double underlined. Multiple transcription start sites are indicated by arrows. Sequences that share homology to known transcription factor binding sites are boldface and underlined. Generic STAT factor binding sites are boxed. Regions –1164 to –1101 and –1399 to –1359 are underlined.

Determination of transcription start site

The transcription start site(s) was determined by primer extension analysis using primer PR-2, positioned 10 bp upstream of the translation initiation codon ATG, with total mouse mammary RNA as the template. Three extended products of 150, 156, and 272 bp in size were obtained, suggesting the presence of multiple initiation sites for the transcription of the gene (data not shown). The same results were observed consistently in three repetitions of the experiment. No primer extension product was observed with tRNA in the control lane. RNase protection assay was performed to confirm the results of primer extension. An RNase
The genomic subclone from the coding region of the gene were selected for RT-PCR. Primers from 5′ primers, UP1 (-145 to –129) and UP2 (-352 to –335) as forward initiation sites determined by primer extension. Two sets of glucosidase I gene, a segment that would include the transcription region of 348 bp upstream to the translation initiation site of the extension may be the major site for transcription initiation. To indicates that the first transcription start site determined by primer reasonably with the first primer extended product of 150 bp. This (Figure 4). The RNase protected fragment of 140 bp matches as the major band. No protection was observed for control tRNA [32P]labeled denatured and electrophoresed on 6% polyacrylamide gel. lane 1, both samples were digested with RNase A and RNase T1, and riboprobe total RNA of mouse mammary gland and yeast tRNA as a control at 37°C under an in vitro transcription reaction and subsequently purified as described under Materials and methods. The riboprobe was hybridized overnight with total RNA of mouse mammary gland and yeast tRNA as a control at 37°C, both samples were digested with RNase A and RNase T1, and riboprobe alone was also treated in a similar fashion. The digested products were denatured and electrophoresed on 6% polyacrylamide gel. lane 1, [32P]labeled φX174/Hind I markers; lane 2, total RNA + riboprobe; lane 3, yeast tRNA + riboprobe; and lane 4, riboprobe alone.

protected fragment of 140 bp was identified on the sequencing gel as the major band. No protection was observed for control tRNA (Figure 4). The RNase protected fragment of 140 bp matches reasonably with the first primer extended product of 150 bp. This indicates that the first transcription start site determined by primer extension may be the major site for transcription initiation. To substantiate these results, RT-PCR was carried out covering a region of 348 bp upstream to the translation initiation site of the glucosidase I gene, a segment that would include the transcription initiation sites determined by primer extension. Two sets of primers, UP1 (-145 to –129) and UP2 (-352 to –335) as forward primers from 5′ UTR and UP0 (+29 to +12) as reverse primer from the coding region of the gene, were selected for RT-PCR. First strand cDNA synthesized from the total RNA of the mouse mammary gland was used as the template. The genomic subclone 5a was used as the positive template in the reaction; the DNA template was omitted in the negative control. First strand cDNA and genomic DNA gave the anticipated size of the product whereas no product was obtained in the negative control lane (Figure 5). The abundance of PCR products was similar with UP1 and UP0 primer pairs for genomic, as well as first strand cDNA, whereas UP2 and UP0 primer pair gave a lower yield with first strand cDNA. Primers selected upstream to –352 bp did not give any PCR product. These results strongly support the data from primer extension that multiple initiation sites for the transcription of the glucosidase I gene exist in the region –348 bp to –1 bp.

Northern blot analysis of glucosidase I

The levels of glucosidase I mRNA in various tissues were determined by northern hybridization (Figure 6). A major band corresponding to 3.1 kb and one or more slow moving, faint bands of higher size were detected. Several faster moving bands, especially in the lanes for liver and kidney RNA, may represent breakdown products of the larger transcript(s). Since equal amount of poly[A+ ]RNA isolated from the different tissues was used for blotting, the expression level of glucosidase I would appear to be higher in mammary gland, liver, kidney, and brain as compared to the lower level in the heart.

Analysis of promoter activity

To determine the promoter activity, the glucosidase I promoter construct, Glu I (-2114/-5) pCAT was assayed by transient transfection into COS 7 cells. β-Casein gene promoter construct was used as positive control. Both the constructs showed positive activity, whereas pCAT basic vector had no activity when assayed under similar conditions (Figure 7).

Discussion

A differential expression of a constellation of more than 100 glycosyltransferases and glycosidases is required for the elaboration of the extensive repertoire of glycoconjugate structures in mammalian cells. An ensemble of some seventeen glycosyltransferases of the dolichol cycle and two glycosidases, viz., glycosidases I and II, participating in concert, is minimally required for the biosynthesis of all N-linked glycoproteins (Kornfeld and Kornfeld, 1985; Moremen et al., 1994). Among these enzymes, glucosidase I is critically juxtapositioned to trigger the post-translational processing of the oligosaccharide moiety and set the stage for the folding and maturation phase of these glycoproteins in the secretory pathway. Therefore, a study of the structure-function and the regulation of glucosidase I would appear to be important for understanding the overall assembly of N-linked glycoproteins. The present study has focused on isolating the gene of the enzyme as a prelude to studying its regulation during the growth and differentiation of the mouse mammary gland.

The structural organization of glucosidase I gene shows a relationship with the functional domains of the enzyme. The first exon codes for the cytoplasmic tail and transmembrane domain of the enzyme, the second and third exons encode the amino acid residues 116 through 192 and 193 through 257, respectively, whereas the fourth exon encodes the putative catalytic domain containing both the active and the glycosylation sites (Shailubhai et al., 1991; Pukazhenthi et al., 1993; Romaniouk and Vijay, 1997). At present, no clear function can be assigned to second and third exons. The organization of the exons correlates well with the biochemical studies on the topology of glucosidase I (Shailubhai et al., 1991). Comparison of the mouse glucosidase I coding
region with the corresponding sequence of the human cDNA shows 84% homology (GenBank AF001719).

Analysis of the 5′ flanking region reveals that it lacks a TATA box, and possesses a perfect CCAAT box, located at position ~870. This is much farther upstream than the typical location of the CCAAT box, usually ~80 nucleotides upstream from the transcription start site. The proximal region of the promoter comprising 176 bp upstream of the translation initiation codon ATG, is 71% rich in GC content, and contains two Sp1 binding sites, and three sites for binding by ETF. These are typical features of housekeeping gene promoters. Indeed, the results of primer extension analysis and RT-PCR confirm the presence of multiple transcription start sites in the promoter. These observations are explicable, given the ubiquitous requirement for glucosidase I in the biosynthesis of N-linked glycoproteins by eukaryotic cells. Earlier, the genes for UDP-GlcNAc: dolichol-P GlcNAc-1-P transferase, the enzyme that catalyzes the first step for the assembly of the oligosaccharide unit of N-linked glycoproteins, and the Golgi-localized UDP-Gal:polypeptide acceptor galactosyl transferase involved in the biosynthesis of complex glycoproteins, were shown to have housekeeping promoters (Shaper et al., 1988; Rajput et al., 1994a). Binding elements for a number of well-known transcription factors, e.g., NFxB, AP2, half-sites for progesterone receptor (PR), and estrogen receptor are also present in the glucosidase I promoter. The size of the major transcript for glucosidase I expression was determined to be 3.1 kb by Northern analysis. These data as well as the results of ribonuclease protection assay indicate that glucosidase I gene is predominantly transcribed from a single transcription site.

The activity of glucosidase I (Shailubhai et al., 1990) and several key, ER-localized enzymes of protein N-glycosylation, viz. UDP-GlcNAc:dolichol-P GlcNAc-1-P transferase; GDP-Man:Dolichol-P mannosyltransferase; and UDP-Glc:dolichol-P glucosyltransferase (Vijay and Oka, 1986), have been shown to be stimulated markedly over the basal level in the virgin tissue during the hormonally modulated growth and lactogenic differentiation of the mammary gland during pregnancy and lactation. The levels of enzyme activities in the postlactating animal decline and reach a basal level (Vijay and Oka, 1986; Shailubhai et al., 1990). Using explant and primary epithelial cell cultures of rodent mammary glands, glucosidase I and UDP-GlcNAc:dolichol-P GlcNAc-1-phosphate transferase have been shown to be regulated by the synergistic action of insulin, prolactin and hydrocortisone (Shailubhai et al., 1990; Rajput et al., 1994a). Within this context, it is noteworthy that the promoter of glucosidase I has two regions that warrant attention. The sequence TGGATTTTAGCAGACTGGCTGTTCCTTATTATTAGAAGGATCTTTGGAGAAAGCCATTCATTTC, covering the segment ~1164 to ~1101 has two underlined hexads. While the right hand hexad has the perfect consensus for binding by the glucocorticoid receptor (GR), the left hand hexad has the critical G and C
required for binding by the receptor (the critical bases in the hexads are shown in italicized bold face); further, the hexads are separated by three nucleotides, a stringent spacing requirement for binding by the dimeric GR (Stöcklin et al., 1996). This part of the promoter sequence, while not perfect to bind the dimeric GR, is an excellent candidate for binding by this receptor.

In recent years, the Jak/Stat signal transduction pathway has been shown to induce the transcriptional activation of many genes. In this paradigm, cytokines, hormones, and growth factors bind their respective member of the transmembrane cytokine receptor superfamily, and activate an intracellular Janus protein tyrosine kinase (Jak) to tyrosine-phosphorylate an intracellular signal transducer and activator of transcription (Stat) (Darnell, 1997). An interaction of Stat protein with the SH2 domain of a dimerization partner confers on (Stat)2 the ability to bind DNA at the Stat-binding element and induce transcription. Jak2 and Stat5 have been identified for signaling by prolactin (Campbell et al., 1994; Rui et al., 1994; Waters et al., 1995). The transcription of genes during lactogenic differentiation of the mammary gland, involves the synergistic action of prolactin and glucocorticoid, operating through the Jak2/Stat5 signal transduction pathway (Darnell, 1997). Transactivation mediated by GR usually requires the binding of dimeric receptor molecules to the palindromic inverted repeat DNA-binding sites, and half-sites binding monomeric GR complexes were considered to be insufficient by themselves to confer hormone responsiveness (Stöcklin et al., 1996). However, monomeric GR molecules binding to the receptor half-palindromic sites were proposed to gain competence as transcriptional activators by interacting with other transcription factors (Slater et al., 1993). GR half-palindromic sites were found to recruit monomeric GR molecules to the β-casein gene promoter and shown to be instrumental in mediating the synergistic effect between glucocorticoids and prolactin for β-casein gene activation (Lechner et al., 1997). In the −1164 to −1101 segment of glucosidase I promoter given above, the two underlined nonads have the generic sequence, TTN5AA that is minimally required for binding by a Stat dimer. The proximity of what may be otherwise two weak but close-neighbor (Stat)2 binding sites, to a potentially strong candidate dimeric GR-binding site, and certainly a monomeric GR binding half-site raises the possibility that this region represents a portion of the promoter where a (Stat)2 or a (Stat)2-(Stat)2 tetramer may form, and interact with the bound GR in close proximity by protein–protein interaction to give a synergized activation of glucosidase I gene expression, as observed during lactogenic induction by glucocorticoid and prolactin in mammary explant culture (Shailubhai et al., 1990).

Another region within the glucosidase I promoter, viz.-1399 AAATAATACC TGTTGCTAGGTCGATAGTAAAGT –1359, could potentially represent an enhancer that binds monomeric forms of GR and interacts with Stat(s) bound to remote sites in the promoter. The segment has the direct repeat

![Fig. 6](https://academic.oup.com/glycob/article-abstract/9/8/797/823325/56979932305)
hexads of glucocorticoid receptor half-sites separated by seven nucleotides (DR7).

Steroid hormone receptors generally bind to hormone response elements composed of two half-sites organized as a palindrome with a three nucleotide spacer, while nonsteroid hormone receptors preferentially bind to hormone response elements composed of two half-sites organized as a direct repeat, with the number and composition of the nucleotides separating the half-sites serving as important determinants of receptor selectivity (Stocklin et al., 1996). There is presently no known receptor that binds a DR7 motif of the type in the promoter segment, –1399 to –1359. Thus, the above region in glucosidase I promoter might represent the binding site for a yet-to-be-discovered receptor (perhaps a new orphan receptor). Recently, an orphan nuclear receptor activated by pregnanes has been proposed to define a novel steroid signaling pathway (Kliewer et al., 1998). This receptor binds to a conserved region within the promoters of genes, CYP31A1 and CYP3A2, that encode cytochrome P450 hemoproteins involved in the hydroxylation of steroids. This region does not contain a typical glucocorticoid response element, but instead contains a DR of the nonsteroid nuclear receptor half-site sequence –1359 to –1359 segment of the glucosidase I promoter half-sites for binding by the monomeric form of the GR, and their presence has much similarity to the DR12 half-sites. TGGACCTCTTGGGATTTAA GGAAC ..., in rat β-casein gene promoter (Lechner et al., 1997). A functional analysis via cotransfection studies in COS cells after introducing mutations in these GR-binding half-sites and other GR-binding half-sites, removed as far away as –140 bp upstream from the Stat5-binding site in the β-casein promoter, dramatically reduced the synergism between prolactin and hydrocortisone for the activation of the reporter gene (Lechner et al., 1997).

Transfection of the glucosidase I promoter fused to CAT reporter gene in COS 7 cells presented in this study clearly demonstrates its transcriptional activity. Future transfection experiments using different deletion constructs of the promoter region of this gene will confirm the role of GR and STAT factors binding elements in hormonal regulation during development at the transcription level.

It is also known that an ER-localized endomannosidase can cleave Glc1-3Man oligosaccharides from the incipient glycoprotein during maturation and can provide an alternative pathway for the initiation of processing and folding during glycoprotein maturation (Lubas and Spiro, 1984; Fujimoto and Kornfeld, 1991). With the availability of the gene of glucosidase I in the mouse, as reported here, it should now be possible to knock out this gene and evaluate the relative significance of the endomannosidase-mediated maturation of N-linked glycoproteins and the essentially of glucosidase I during embryogenesis, growth, and differentiation of eukaryotes.

Materials and methods

Materials

Mouse genomic library in λ EM-BL3 SP6/T7 and human cDNA probe of glyceraldehyde 3 phosphate dehydrogenase (G3PDH) were from Clontech; Trizol reagent, restriction endonucleases, T4 DNA ligase, T4 polynucleotide kinase, deoxyribonucleotide triphosphates, MMLV reverse transcriptase, Klenow fragment of DNA polymerase I, Taq polymerase, DNA molecular weight markers, tissue culture media, and oligonucleotide primers were from Life Technologies; radioactively-labeled nucleotides and [3H] chloramphenicol were from NEN Life Sciences Products. The vector pGEM2-Zf(–), pCAT, primer extension kit, riboprobe system, Poly A Tract mRNA isolation system III, and mouse genomic DNA were from Promega Corporation. β-Casein gene promoter pβc (–282/-82) CAT and glucocorticoid receptor expression vector pSTC GR 3–795 were gifts from Dr. Wolfgang Doppler, Institut für Medizinische Chemie und Biochemie, Universität Innsbruck, Fritz-pregl-Strasse 3, A-6020 Innsbruck, Austria. STAT 5a expression vector pXM was kindly provided by Dr. Bernd Groner, Institute for Experimental Cancer Research, Tumor Biology Center Freiburg, D-79106 Freiburg, Germany. Long form prolactin receptor expression vector pR/C MV was a gift from Dr. Paul Kelly, INSERM, Unité 344, Endocrinologie Moléculaire, Faculté de Médecine, Neckler-Enfant Malades, Paris, France. Sequenase version 2.0 DNA sequencing kit was purchased from U.S. Biochemical Corp.; nitrocellulose membranes were from Schleicher & Schuell; Nu-Trap columns were from Stratagene; TA Cloning system was from Invitrogen Corp. The mouse mammary glands were obtained from C3H mice maintained at our animal colony. All other chemicals were from Sigma or Fisher Chemical Co.

Southern hybridization

Mouse genomic DNA (10 µg/lane) was digested with the restriction enzymes BamH I, HindIII, EcoRI, Kpn I, and Pst I; electrophoresed in 0.8% agarose; transferred to nitrocellulose membranes; and hybridized overnight with the partial cDNA probe of rat ovary glucosidase I (GenBank accession no. AF087431) under high stringency condition in 50% formamide, 5× SSPE (0.75 M NaCl, 0.05 M NaH2PO4, (pH 7.4) and 0.005 M EDTA), 5× Denhardt’s solution (0.1% Ficoll 400, 0.1% polyvinylpyrrolidone, 0.1% bovine serum albumin), 10 mg/ml denatured salmon sperm DNA, and 0.5% SDS at 42°C (Khan et al., 1996). cDNA probes were labeled with [32P]dCTP using Klenow fragment of DNA polymerase I and random hexamer primers (Feinberg and Vogelstein, 1983).

Screening of mouse genomic clone

One million recombinant phages of mouse genomic library were plated for each screening. Duplicate filters were lifted and hybridized with the partial cDNA probe of rat glucosidase I. Five genomic clones, λ 4, 5, 10, 12, and 16, were isolated by the screening. The clones were plaque-purified and bacteriophage DNA was purified by the method of Meese (Meese et al., 1990). Each genomic clone was analyzed by Southern hybridization with [32P] labeled oligonucleotide primer selected from 5′ and 3′ end of the rat cDNA clone used for library screening. Subclones 5a and 3a, progenitors of λ 5 clone, were found to be positive by Southern analysis and used for further experiments.

Primer extension

Primer extension was carried out using oligonucleotide primer PR-2 (5′ TGCAGCTTCCCCGCTGCCAGCTGGCGT 3′) according to manufacturer’s instructions (Promega Corp.). Position of the primer is shown in Figure 3. The primer-extended products were separated on a 6% polyacrylamide, 8 M urea gel. A sequencing reaction carried out on the genomic subclone 5a with PR-2 primer was run in parallel (Church and Gilbert, 1984).
Reverse transcription-polymerase chain reaction (RT-PCR)

First strand cDNA was synthesized with MMLV reverse transcriptase using mouse mammary gland total RNA as a template and random hexamer primers. For 5' transcription start analysis, first strand cDNA equivalent to 1 μg of total RNA was subjected to PCR with 10 pmol of sense primers UP-1 (5' ACCAGCCAC-AGGCGTCTCC 3') and UP-2 (5' TGCCACGACTTGTTAGAG3'), and the antisense primer UP-0 (5' AGGGGCGGCTCGG-TCCTG 3'). Genomic subclone 5a was used as a template for positive control; no template DNA was added in the negative control.

Ribonuclease protection assay

A PCR product of 382 bp was generated using genomic subclone 5a as a template and UP0/UP2 as forward and reverse primers respectively. This fragment was subcloned in pGEM respectively. It was sequenced completely to confirm its se-

Chloramphenicol acetyltransferase construct

Glucosidase I gene promoter construct was prepared by subcloning PstI and XbaI cut product of PCR in corresponding sites of pCAT Basic vector designated as Glu I (-2114/-5) pCAT+. Genomic subclone 5a was used as a template for PCR amplification reaction. The upstream primer was complementary to 18 bp starting at –2114 flanked by XhoI site and HindIII site as described previously (Sambrook et al., 1989). The resultant products were separated on a 6% polyacrylamide; 8 M urea gel in parallel with [32P]labeled ωX174/Hinf I DNA markers. Yeast RNA was hybridized with equal amount of riboprobe as a control.

Cell culture and transfection

COS 7 cells were maintained in Dulbecco’s modified Eagle’s medium containing 10% fetal bovine serum and 1x penicillin, streptomycin, and neomycin antibiotics. After reaching 80% confluency, cells were cotransfected with 20 μg of plasmid DNA containing 5 μg each of Glu I (-2114/-5) pCAT, PSTC GR 3–795, prolactin receptor pRC/CMV, and STAT 5a pXM expression vectors per 100 mm plate, by the calcium phosphate precipitation method (Sambrook et al., 1989). Transfection with β-casein gene promoter construct pBβ (-282/-82) CAT (5 μg) was used as a positive control under similar conditions. Cells were washed with cold PBS and harvested using a cell scraper in 0.25 M Tris–HCl, (pH 7.8) at 4°C. Cell lysate was prepared by freeze–thaw followed by centrifugation at 12,000 × g for 20 min. The supernatant was used for CAT activity assay.

CAT assay

CAT activity in the culture lysates was measured as described previously (Sambrook et al., 1989). The incubation mixture contained 75 μg of cell lysate, 50 μl of 1 M Tris–HCl (pH 7.8), 10 μl of [14C] chloramphenicol (50 mCi/mmol; 0.1 mCi/ml), and 20 μl of 4 mM acetyl CoA in a final volume of 130 μl. The reaction was carried out at 37°C overnight. Reaction products were separated on TLC in the solvent system, chloroform:meth-

Abbreviations

ER, endoplasmic reticulum; RER, rough endoplasmic reticulum; BiP, binding protein; bp, base pair(s); kb, kilobase pairs; GR, glucocorticoid receptor; PR, progestogen receptor; ETF, epider-

References


