Mouse submandibular gland salivary apomucin contains repeated N-glycosylation sites

Paul C.Denny1, Lily Mirels1 and Patricia A.Denny

Department of Basic Sciences, School of Dentistry, University of Southern California, Los Angeles, CA 90089–0641 and 1Biology Dept., University of California at San Diego, La Jolla, CA 92093–0347, USA

To whom correspondence should be addressed

A cDNA clone encoding mouse submandibular gland salivary mucin apoprotein was isolated and characterized. The mucin cDNA encodes a protein of 273 amino acids with a calculated molecular weight of 29,606. This apomucin is approximately 60% Thr, Ser, and Pro, and has a pI of 11.25. In addition to the signal sequence, the apomucin can be divided into four structural domains. The first of these contains over 30% Thr, Ser, and Pro, but only a few probable O-glycosylation sites. The second domain contains 10 repeats, each 9 or 13 amino acids in length, with Thr representing more than 50% of the amino acids, while Ser accounts for only 2%. Each repeat begins with a putative N-glycosylation site; hence this domain likely contains both N- and O-linked oligosaccharides. The third domain lacks a repeat motif, but is rich in both Thr and Ser, and therefore is potentially highly O-glycosylated. The final domain is composed mainly of basic and non-polar amino acids and does not contain Cys. This mucin shows considerable homology with the rat submandibular salivary mucin, but little overall homology with other mucins. In situ hybridization verifies that the mucin transcript is localized primarily in the acinar cells of the submandibular gland.

Key words: cDNA sequence/mouse submandibular gland/salivary mucin/N-glycosylation sites

Introduction

Mucins are high molecular weight glycoproteins that coat and protect mucosal surfaces from desiccation and from microbial or chemical damage (Tabak, 1990). A common feature among mucin apoproteins is that they are composed of a large proportion of Ser, Thr, Gly, Ala, and Pro, and contain greater than 50% carbohydrate by weight (Strous and Dekker, 1992). For most apomucins, a repeated amino acid sequence rich in Ser or Thr may provide the sites for attachment of the majority of O-linked oligosaccharides. Mucins may also contain a small proportion of N-linked oligosaccharides, as evidenced from the occurrence of Man, usually at 1% or less of the total carbohydrate (Tabak et al., 1985; Dekker et al., 1989; Corfield et al., 1991; Rumasubbu et al., 1991). Molecular cloning of mucin apoproteins has demonstrated that the consensus sequence(s) for N-glycosylation is usually found outside of the repeat domain, or in isolated cases as degenerate sequences within the repeat domain (Hoffman, 1988; Bhargava et al., 1990; Gendler et al., 1990; Eckhardt et al., 1991; Spicer et al., 1991; Toribara et al., 1991; Bobek et al., 1993; Albone et al., 1994).

In mouse submandibular gland salivary mucin, Man present in N-glycans accounts for more than 7% of the total carbohydrate (Denny et al., 1980; Denny and Denny, 1982a; Nieuw Amerongen et al., 1983). In order to understand the basis for the unusual carbohydrate composition of the mouse submandibular salivary mucin, and to determine whether it contains other features common to mucin apoproteins such as tandem repeats (Strous and Dekker, 1992) and a Cys-rich domain (Bhargava et al., 1990; Eckhardt et al., 1991), cDNA clones encoding the mouse submandibular salivary mucin apoprotein were isolated and characterized.

Results

Clones encoding mouse submandibular salivary mucin apoprotein were isolated from a mouse submandibular gland cDNA library, based on homology with the rat submandibular salivary mucin clone pRSM-3 (Albone et al., 1994). Southern blot analysis demonstrated that their cDNA inserts ranged in length from 0.35 to 1.1 kb. A total of 27 clones were sequenced from the 5′ end using the T3 promoter sequence of pBluescript as primer. Ten clones terminated in what was later determined to be the 5′ untranslated domain. Seventeen partial-length clones terminated at various positions within the apomucin protein coding sequence. Comparison of the sequences derived from these clones allowed a preliminary determination of the majority of the apomucin sequence. The sequence of the full length clone SM2–1R was confirmed by sequencing in the reverse direction using synthetic oligonucleotides. The SM2–1R cDNA was approximately 1.1 kb in length.

The nucleotide and deduced amino acid sequences of mouse submandibular gland salivary apomucin are shown (Figure 1). Nucleotides 1–46 were derived from clone SM3–2B, which had the longest 5′ untranslated sequence. The remainder of the sequence is contained in clone SM2–1R. The full sequence is 1042 bp with the open reading frame beginning at nucleotide 102 and ending at 923. An in-frame termination codon is present 6 nucleotides upstream of the initiating AUG. A second possible initiation codon is present 9 bp downstream from the first initiation site, in the same open reading frame. The surrounding sequences for the two potential initiation sites are both represented at low, but equal, frequencies in a survey of initiation site contexts (Kozak, 1983). The apomucin sequence contains a series of ten contiguous repeats beginning at bp 407. Repeats 1–5 and 10 are 39 bp in length, whereas repeats 6–9 are 27 bp. Despite 85% or greater
Fig. 1. Mouse submandibular gland salivary apomucin sequence (clone SM2-1R). The 5' and 3' non-translated domains are represented by continuous sequence. The protein coding sequence is shown in codon blocks, whereas the deduced amino acids are represented by their single letter symbol. The two possible initiation sites, the termination codon, and polyadenylation signal are underlined. Putative Asn-glycosylation sites are preceded by (*). Nucleotides and amino acids that differ from consensus in the 39 or 27 bp repeats are shown in bold. Nucleotides 1–46 were derived from clone SM3-2B.
Table I. Amino acid composition of conceptually translated secreted apomucin

<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Number of Residues</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thr</td>
<td>74</td>
<td>29.4</td>
</tr>
<tr>
<td>Pro</td>
<td>36</td>
<td>14.3</td>
</tr>
<tr>
<td>Ala</td>
<td>21</td>
<td>8.3</td>
</tr>
<tr>
<td>Lys</td>
<td>20</td>
<td>7.9</td>
</tr>
<tr>
<td>Ser</td>
<td>19</td>
<td>7.5</td>
</tr>
<tr>
<td>Asn</td>
<td>14</td>
<td>5.5</td>
</tr>
<tr>
<td>Arg</td>
<td>10</td>
<td>3.9</td>
</tr>
<tr>
<td>Phe</td>
<td>9</td>
<td>3.5</td>
</tr>
<tr>
<td>Ile</td>
<td>9</td>
<td>3.5</td>
</tr>
<tr>
<td>Leu</td>
<td>6</td>
<td>2.3</td>
</tr>
<tr>
<td>Tyr</td>
<td>6</td>
<td>2.3</td>
</tr>
<tr>
<td>Gln</td>
<td>6</td>
<td>2.3</td>
</tr>
<tr>
<td>Val</td>
<td>4</td>
<td>1.5</td>
</tr>
<tr>
<td>His</td>
<td>3</td>
<td>1.1</td>
</tr>
<tr>
<td>Glu</td>
<td>3</td>
<td>1.1</td>
</tr>
<tr>
<td>Asp</td>
<td>3</td>
<td>1.1</td>
</tr>
<tr>
<td>Cys</td>
<td>2</td>
<td>1.1</td>
</tr>
<tr>
<td>Trp</td>
<td>2</td>
<td>.7</td>
</tr>
<tr>
<td>Met</td>
<td>2</td>
<td>.7</td>
</tr>
<tr>
<td>Gly</td>
<td>2</td>
<td>.7</td>
</tr>
</tbody>
</table>

Overall identity, the majority of repeat sequences display degeneracy. In total, the repeat domain constitutes approximately 41% of the apomucin protein coding sequence.

The deduced protein sequence of clone SM2-1R contains 273 amino acids with a predicted molecular weight of 29,606. A hydropathy plot and search for sequences common to signal peptides suggest that the first 22 amino acids are likely to represent the signal peptide (von Heijne, 1986). The secreted apomucin would then be expected to have a molecular weight of 27,131. Earlier studies indicated that mouse submandibular salivary mucin was 19% protein and 81% carbohydrate by weight (Denny et al., 1980), yielding a predicted size for the mature mucin of 143 kDa. This is in good agreement with the apparent molecular weight of 140 kDa previously determined by SDS-PAGE for the purified, glycosylated mucin (Denny et al., 1980).

The amino acid composition of the conceptually translated secreted mucin is more than 50% Thr, Ser, and Pro (Table I), in agreement with earlier chemical analyses of the glycosylated (Denny et al., 1980) and HF-deglycosylated mouse submandibular salivary mucin (unpublished data).

The secreted portion of the apomucin sequence can be viewed as having four discrete regions, based on compositional weighting of amino acid classes. The first region (amino acids 23–102), although composed of more than 30% Pro, Ser, and Thr, also contains more than 20% basic amino acids and 24% non-polar amino acids. The largest domain (amino acids 103–216) comprises the area of repeats. Nearly 50% of the total amino acids in this region is Thr, while Ser is conspicuously low at 2%. Each repeat contains an N-glycosylation consensus sequence, Asn-X-Thr/Ser. In seven of the repeats, the consensus sequence is Asn-Ala-Thr; in two it is Asn-Val-Thr, and in one it is Asn-Thr-Thr. The third region (amino acids 217–256) is composed of unique sequence but is still rich in hydroxyamino acids, containing approximately 40% Thr and 20% Ser. The final domain (amino acids 257–273) is dominated by non-polar and basic amino acids, and does not contain Cys. This is the only region of the secreted apomucin that is likely to have a helical secondary structure. Without considering the possible effects of immature N-glycans on the apomucin structure (Wyss et al., 1995), by using the method of Garnier et al. (1978), a computer-assisted model suggests that approximately 70% of the remaining secreted apomucin exists as an extended secondary structure.

The apomucin sequence suggests that, as for other mucins, the N- and C-terminal domains are sparsely O-glycosylated, due to a paucity of high probability O-glycosylation sites (O’Connell et al., 1991; Wilson et al., 1991; Wang et al., 1993). The repeat domain is likely to be heavily glycosylated and contain both N- and O-linked oligosaccharides. The third domain is probably also heavily glycosylated, but mostly with O-linked oligosaccharides, because only one putative N-linked glycosylation site is present in this part of the apomucin.

The cDNA library used for this study was constructed from poly (A)’ RNA obtained from glands of 10 outbred mice. A total of nine clones containing both 5’ and 3’ untranslated apomucin sequences was analyzed. No heterogeneity was found among the 5’ or 3’ untranslated sequences, signal peptides, or regions encoding the first, third and fourth domains of the apomucin protein. However, heterogeneity in the number of repeats was found. Of the nine “full length” clones analyzed, the number of repeats was 10 (2 clones), 9 (4 clones), 8 (1 clone), and 6 (2 clones). Additionally, in four of these clones, repeat sequences similar to the third repeat of SM2–1R, but differing from this repeat by 2, 3, or 4 nucleotides, were found in place of repeats 1, 2, or 3. No sequence heterogeneity was detected in the remaining repeats, 4 to 10. The sequence of SM2–1R was selected for presentation (Figure 1) as it had 10 repeats, and the sequences at each repeat position were those most commonly found among all nine clones. Whether the differing clones resulted from cloning artifacts or were derived from alternate apomucin alleles in this outbred population will be the object of future studies.

The insert of cDNA clone SM2–1R was used as a probe for Northern blot analysis of total RNA prepared from submandibular glands of male and female mice, and from submandibular glands from females that were carbachol-treated (Figure 2). Figure 2A shows by ethidium bromide staining, that equivalent amounts of total RNA were run. In each sample, the probe hybridized to a discrete band at approximately 1 kb (Figure 2B), consistent with the size of clone SM2–1R. Furthermore, mucin transcripts appear to be equally abundant in the glands of male and female mice. Though carbachol stimulates mucin secretion in salivary glands and rectal mucosa, chronic exposure appears to have had little effect on the relative abundance of the mucin transcript. A small amount of size heterogeneity can be noted on the leading side of the apomucin transcript band. Whether this represents transcript degradation or repeat number heterogeneity, as suggested above, also remains to be determined. The Northern blot suggests that, if authentic, the 6 repeat variant is not a highly represented transcript. Clone SM2–1R was also used to generate probes for in situ hybridization to mouse submandibular glands. Reactivity was localized to the seromucous acinar cells (Figure 3A,B), consistent with the known cell-type specificity of submandibular salivary mucin (Denny and...
Fig. 2. Northern blots of total RNA from submandibular glands of adult mice using probe generated from mouse submandibular salivary mucin cDNA (clone SM2–1R). A, Ethidium bromide-stained 1.2% agarose gel containing 5 \( \mu \)g of total RNA per lane. Lane 1, adult females; lane 2, adult females following repeated stimulation of secretion with carbachol; lane 3, adult males. B, Autoradiogram of Northern blot hybridization of the ethidium-stained gel from A.

Denny, 1982b; Denny et al., 1988). Identical conditions yielded a negative reaction with the sense probe.

A search of GenBank and EMBL data bases revealed that only the rat submandibular salivary mucin shared significant overall DNA sequence homology with the mouse submandibular salivary mucin. The nucleotide and deduced amino acid sequences for submandibular salivary apomucin from mouse (Figure 1) and rat (Albone et al., 1994) were compared using sequence alignment software in PC Gene. Alignment of the 5' and 3' untranslated nucleotide sequences and of the amino acid sequences deduced from these clones is shown (Figure 4). The nucleotide sequences of the two mucins share 80.7% sequence identity, excluding the poly(A) tails. The areas of greatest homology are the 5' untranslated, 3' untranslated, and signal peptide-encoding regions, which together give over 90% identity. The overall similarity in amino acid sequence of the two apomucins is 51.7%. The most similar regions of these proteins (over 80% sequence identity) are their signal sequences, and the Thr- and Ser-rich domains C-terminal to the repeats.

Comparison of the two repeat domains (mouse amino acids 104–216) reveals that the rat apomucin has a greater number of 13 amino acid repeats but lacks 9 amino acid repeats (Figure 4). Furthermore, the repeats of the mouse apomucin are more heterogeneous than those of the rat apomucin (Figure 4). There are only two Asn in the rat apomucin repeat domain, and neither is a putative N-linked glycosylation site (rat amino acids 177 and 190). The rat mucin repeats each contain the sequence GAT TCA ACC (Asp-Ser-Thr). In the mouse, the homologous sequence is AAT GCA ACC (Asn-Ser-Thr) or a slight variant, creating an N-glycosylation site within each repeat. Despite these differences, the rat and mouse homologous 13 amino acid repeats share 73% sequence identity. The one putative N-glycosylation site in the rat apomucin is outside the repeat domain (rat amino acid 290), and is conserved in the mouse apomucin. The 5' domains between the putative signal peptide and the beginning of the repeats (mouse amino acid 23–102) and the C-terminal domains are least similar between mouse and rat apomucins, showing 35% and 14% sequence similarity, respectively. Neither rat (Albone et al., 1994) nor mouse apomucin contains a Cys-rich C-terminal domain that was found for porcine and bovine submandibular gland mucins (Bhargava et al., 1990; Eckhardt et al., 1991).

Discussion

The mouse submandibular gland salivary apomucin, though the smallest of the mammalian salivary gland apomucins described thus far, shares many features in common with other mucin apoproteins. Its aggregate content of Thr, Ser, and Pro of over 50% provides an abundance of sites available for O-linked glycosylation. The extended nature of the apomucin amino acid chain, in combination with the negatively charged NeuNAc-rich mouse submandibular salivary mucin oligosaccharides, can be expected to produce the “bottlebrush” structure common to mucins (Toribara et al., 1991). Like most other mucins, the mouse submandibular mucin contains a discrete domain of amino acid sequence repeats, which may serve as a region of abundant O-glycosylation. However, mouse submandibu-
Fig. 4. Sequence similarity of mouse and rat submandibular salivary mucins. The conceptually translated apomucins use the single letter symbols for amino acids. Gaps introduced to maximize homology are designated by (-). Tandem repeats are separated by (*). The one putative N-glycosylation site in the rat mucin is identified by (\text{\textbackslash \textbackslash A}). This glycosylation site is conserved in the mouse mucin.

The amino acid sequence of mouse submandibular salivary apomucin demonstrates that the relatively high proportion of N-linked carbohydrate observed on this mucin is made possible by the regular occurrence of an N-glycosylation consensus sequence within each of its repeats. Although N-glycosylation consensus sequences have been identified in the N- or C-terminal regions of most other mucin apoproteins (Hoffman, 1988; Bhargava et al., 1990; Gendler et al., 1990; Eckhardt et al., 1991; Spicer et al., 1991; Toribara et al., 1991; Bobek et al., 1993; Albone et al., 1994), their presence within the repeat domain is not common. Only in mouse Muc-1 are several consensus sequences present (Spicer et al., 1991). These appear to represent degeneracies in the repeat sequence. No mucin exhibits the regular occurrence of putative N-glycosylation consensus sequence within each of its repeats. Although N-glycosylation consensus sequences have been identified in the N- or C-terminal regions of most other mucin apoproteins (Hoffman, 1988; Bhargava et al., 1990; Gendler et al., 1990; Eckhardt et al., 1991; Spicer et al., 1991; Toribara et al., 1991; Bobek et al., 1993; Albone et al., 1994), their presence within the repeat domain is not common. Only in mouse Muc-1 are several consensus sequences present (Spicer et al., 1991). These appear to represent degeneracies in the repeat sequence. No mucin exhibits the regular occurrence of putative N-glycosylation consensus...
sequences in their repeats such as seen in the mouse submandibular mucin apoprotein. An average chain length of 9.9 residues containing an average of 5.4 Man can be calculated from the types and relative proportion of \(N\)-linked oligosaccharides isolated from the mouse submandibular salivary mucin (Denny et al., 1995). Given the molar relationship of Man to total amino acids in this mucin (Denny et al., 1980), we estimate that an average of 10 of the 12 possible \(N\)-glycosylation sites in the 10 repeat domain of the apomucin are occupied. These \(N\)-glycosylation sites are surrounded by the amino acids Pro-Thr-(Thr)-Thr-(Asn-X-Thr)-Thr-Thr-Pro. Thus, the majority of Thr residues that are potential targets for \(O\)-glycosylation in the repeat domain are likely to be in close proximity to an \(N\)-linked oligosaccharide. Since \(N\)-linked carbohydrate is added co-translationally (Kornfeld and Kornfeld, 1985), GaINAc transferase activity leading to \(O\)-glycosylation within the repeat domain occurs in the presence of previous added, but immature, \(N\)-linked chains. Whether the presence of these \(N\)-linked chains affects the efficiency of adjacent \(O\)-glycosylation remains to be seen.

Mouse and rat submandibular salivary apomucin cDNAs are most similar in their 5' and 3' untranslated regions, and in the regions encoding the signal peptides of their respective proteins. The secreted proteins themselves are less closely related. This pattern has been observed for many related salivary proteins, and is assumed to reflect a general tolerance of (or selection for) divergence among salivary constituents (Clements et al., 1985; Dickinson et al., 1987). The secreted forms of the two apomucins are most similar in their Thr- and Pro-rich repeat and 'domain 3' regions, reflecting the requirement that they be efficient substrates for glycosyltransferase activity. Despite the similarity in these regions, the carbohydrate content of mouse and rat mucins is quite different. In addition to the previously noted difference in \(N\)-linked oligosaccharides, the mouse mucin lacks fucose, which is present in rat mucin (Tabak et al., 1985). Furthermore, after adjustment for the contribution by \(N\)-glycans, the remaining O-linked oligosaccharides on the mouse mucin contain approximately 5 times more NeuNAc, 30 times more GlcNAc, and 2 times more Gal per mole of amino acid than the rat mucin (Denny et al., 1980; Tabak et al., 1985). On the other hand, the mouse mucin contains only about two-thirds of the GaINAc of the rat mucin per mole of amino acid.

Of the salivary apomucins other than rat submandibular, the mouse submandibular mucin most closely resembles human submandibular gland mucin MUC7 (Bobek et al., 1993). MUC7 shares no sequence identity with the rat (Albone et al., 1994) or mouse submandibular mucin apoproteins or transcripts. However, like the mouse and rat mucins, MUC7 is monomeric, favors Thr as the major hydroxyamino acid, contains repeat domains, and lacks a C-terminal Cys-rich domain (Bobek et al., 1993). An additional similarity between these three mucin apoproteins can be observed in their first domains, N-terminal to the apomucin repeat regions. Although there is little amino acid sequence similarity, the N-terminal domains of all three apomucins are rich in basic amino acids, with a predicted pI between 10.5 and 10.9. The significance of this basic domain is not known.

The mouse submandibular gland also synthesizes Muc-1, the homolog of human MUC1 (Braga et al., 1992). This membrane-associated mucin is thought to be involved in ductal differentiation of a variety of organs. In the mouse submandibular gland, Muc-1 homolog persists through adulthood in association with duct cells, but not in acinar secretory cells. Previously, we had shown by immunolocalization and immunoassay that the submandibular salivary mucin was found only in the acinar cells and their putative progenitors (Denny and Denny, 1982b; Denny et al., 1988). The difference in mouse Muc-1 and submandibular gland salivary mucin localization led to the suggestion that they were unrelated (Braga et al., 1992). Comparison of these two cDNA sequences confirms that there is no significant homology between mouse submandibular salivary mucin and mouse Muc-1. Thus, the two genes, both expressed in the submandibular gland, are probably independently regulated and serve different functions.

Materials and methods

Mice

Ten adult female Swiss-Webster mice (Simonsen Labs, Gilroy, CA) 8–10 weeks old were injected intraperitoneally with carbachol, in phosphate-buffered saline, at a dosage of 0.67 mg/kg body weight daily for 4 d. On the sixth day, the mice were sacrificed by cervical dislocation. The submandibular–sublingual gland complex was removed, and the two glands were separated with the aid of a dissecting microscope. The submandibular glands were rinsed in saline, blotted, and frozen immediately in liquid nitrogen.

RNA isolation

Total RNA was prepared from frozen glands (Han et al., 1987). Poly(A)+ RNA was purified by affinity chromatography using oligo(dT) cellulose (Avr and Leder, 1977).

Preparation of cDNA library

The submandibular gland poly(A)+ RNA provided starting material for cDNA synthesis and subsequent ligation into the lambda-ZAP II phagemid vector using the protocol and reagents supplied with the ZAP cDNA synthesis kit (Stratagene, La Jolla, CA). Afterpackaging with the Gigapack II Packaging Extract, approximately 109 plaques from the primary library were amplified.

Screening of cDNA library

Approximately 8000 plaques were screened using the cDNA insert of the rat submandibular gland mucin pRSM-3 (Albone et al., 1994). The probe was labeled with \(32\)PdCTP by random priming (Feinberg and Vogelstein, 1984). Plate lifts were performed using nitrocellulose membranes (Schleicher and Schuell, Keene, NH). Blots were hybridized and washed using high stringency conditions (Sambrook et al., 1989; Girard et al., 1993), and were exposed to Kodak XAR film at 80°C for 24 h using intensifying screens. A total of 39 positive clones were identified. pBluescript phagemids were excised using ExAssist helper phage and SOLR strain E. coli (Stratagene). All 39 cDNA clones were digested with EcoRI and XhoI restriction endonucleases (New England Biolabs, Beverly, MA), separated on 1.5% agarose gels, transferred to nitrocellulose and rehybridized to pRSM-3 as described previously (Mirels and Ball, 1992).

cDNA sequencing

The S’ ends of selected clones were sequenced using the T3 promoter site of the pBluescript phagemid by the dideoxy-method using the Sequenase II 7-deaza-dGTP kit (U.S. Biochemical Corporation, Cleveland, OH) and labeled with \(35\)SdATP (1280 Ci/mmole; New England Nuclear, Boston, MA). Most of the mucin cDNA sequence was assembled from overlapping clones. The sequence of two clones, SM2–1R and SM3–2B, was completed using three synthetic primers (Operon Technologies, Alameda, CA) with the following sequences: positions 334–353, 367–393, and 818–835 (Figure 1). SM2–1R was sequenced in the reverse direction.
using the T7 promoter site of the plBluescript phagemid and four additional synthetic primers complementary to positions 184–202, 372–393, 757–779, and 966–986 (Figure 1).

**Northern blot analyses**

Total RNA preparations from adult male and female submandibular glands, as well as the total RNA noted above to generate the cDNA library, were run on 1.2% agarose gels (Mirels and Ball, 1992). After transfer to nitrocellulose, the RNA was hybridized with 32P-labeled probes generated by random priming (Feinberg and Vogelstein, 1984) from the insert in clone SM2–1R.

**In situ hybridization**

High resolution in situ hybridization was performed on adult female submandibular glands by a modification of a protocol described previously (Bekhor et al., 1994). Each submandibular gland was cut into 1–2 mm square pieces and fixed in 1% formaldehyde for 2 h at 0°C. Following two 5 min rinses in 0.1 M sodium cacodylate, they were incubated in a solution which contained 30 mg crude collagenase, 15 mg hyaluronidase and 5 international units of Streptolysin O in 10 ml phosphate-buffered saline at 37°C for 15 min with shaking. The partially dissociated glands were then post-fixed in 1% formaldehyde for 15 min at room temperature with rocking, rinsed twice in sodium cacodylate for 5 min at 0°C, dehydrated through a series of ethanols, and stored in 100% ethanol at -20°C.

In situ hybridization, the tissue was subsequently embedded in LR Gold (Polysciences) and cut in 1 µm serial sections (Bekhor et al., 1994). Digoxigenin-UTP labeled sense and anti-sense cRNA probes were synthesized from linearized SM2–1R cDNA (Jiang et al., 1991).

**Sequence analyses**

Sequences were entered, edited and assembled in PG CENE (IntelliGene
tics). The following PG CENE subprograms were used: Chargepro, Garn
er, Ggbsm, Nalign, Novotny, Palign, Psignal, and Soap. A homology search was conducted on the NCBI BLAST e-mail server with the BLASTN program using the GenBank and EMBL Data Library databases.

**Acknowledgements**

We wish to thank Drs. E. Albome and L. Tabak for their generous gift of the pRSM-3 cDNA clone. We also wish to thank Dr. L. Tabak for his valuable suggestions with regard to the manuscript. This work was supported by USPHS (NIH/NIDR) Grant DE06892 (P.C.D.).

**References**


major phosphoproteins and mucins of human submandibular-

Laboratory Manual, 2nd Ed. Cold Spring Harbor Laboratory, Cold
Spring Harbor, NY.

ing and analysis of the mouse homologue of the tumor-associated
mucin, MUC1, reveals conservation of potential O-glycosylation sites,
transmembrane, and cytoplasmic domains and a loss of minisatellite-


Tabak, L.A., Mirels, L., Monte, I.D., Ridall, A.L., Levine, M.J., Loomi-
s, R.E., Lindauer, F., Reddy, M.S. and Baum, B.J. (1985) Isolation and
characterization of a mucin-glycoprotein from rat submandibular

Toribara, N.W., Gum, J.R., Jr., Culhane, P.J., Lagace, R.E., Hicks, J.W., Pet-

von Heijne, G. (1986) A new method for predicting signal sequence cleav-

The acceptor substrate specificity of porcine submaxillary UDP-Gal-
NAc-polypeptide N-acetylgalactosaminytransferase is dependent on
the amino acid sequences adjacent to serine and threonine residues.

Wilson, J.B.H., Gavel, Y. and von Heijne, G. (1991) Amino acid distribu-

Wyss, D.F., Cho, J.S., Li, J., Knoppers, M.H., Willis, K.J., Arulanan-
mation and function of the N-linked glycan in the adhesion domain

Received on August 7, 1995; revised Sept. 22, 1995; accepted Sept. 30,
1995.