Isolation and Characterization of Human Milk Bile Salt-Activated Lipase C-Tail Fragment†

Chi-Sun Wang,*,‡,§ Azar Dashti,‡ Ken W. Jackson,|| Jiunn-Chern Yeh,§ Richard D. Cummings,§ and Jordan Tang‡,§

Protein Studies Program, Oklahoma Medical Research Foundation, Oklahoma City, Oklahoma, and Department of Biochemistry and Molecular Biology and William K. Warren Medical Research Institute, University of Oklahoma Health Sciences Center, Oklahoma City, Oklahoma

Received March 28, 1995; Revised Manuscript Received June 12, 1995®

ABSTRACT: Glycosylation positions and oligosaccharide characteristics in the proline-rich, mucin-like, C-terminal region (C-tail) of human milk bile salt-activated lipase (BAL) were studied in order to assess the possible physiological functions of this region. A large-scale purification method has been devised to purify the C-tail fragment from human milk BAL. Chymotryptic, tryptic, and cyanogen bromide cleavages of partially purified BAL and subsequent molecular sieve chromatography yielded 20−30 mg of C-tail fragment from 1 L of human milk. The N-terminal sequence and amino acid composition of the purified C-tail fragment establish that it is derived from residues 528−712 of the enzyme. The *O*-glycosylated carbohydrates of the C-tail fragment contain fucose, galactose, glucosamine, galactosamine, and neuraminic acid in a molar ratio of 1:3:2:1:0.3, respectively. β-Elimination reaction revealed that nine threonine residues and less than one serine residue were glycosylated. Edman degradation of C-tail fragment and its pronase subfragment suggest a number of glycosylation sites which are flanked by a consensus motif of PVPP. We suggest that this motif may serve as a signal for *O*-glycosylation in the C-tail region of BAL. Immunochemical studies indicated that the oligosaccharide chains in the C-tail region of BAL contain Lewis x and Lewis a antigens and, less prominently, sialyl Lewis x and sialyl Lewis a antigens. C-tail fragment was also found to bind jacalin lectin. These observations suggest the possibility that the C-tail region may contribute to adhesive activity in the physiological function of BAL.

Bile salt-activated lipase (BAL)[1] is the major lipolytic activity present in human milk (Wang & Johnson, 1983; Olivecrona & Bengtsson, 1984). The enzyme is not active in the milk but is activated in the intestine by bile salts. The function of milk BAL is the digestion of milk fats in infants (Alemi et al., 1981; Wang et al., 1989). In mammals, this enzyme is also secreted from pancreas into the intestine (Wang & Hartsuck, 1993). The cDNA sequences of human milk (Baba et al., 1991) and human pancreas (Reue et al., 1991) BAL are identical, indicating that they are expressed from the same gene. BAL shares with other pancreatic lipases the triacylglycerol hydrolase activity in the intestine, but BAL is the only intestinal lipase which hydrolyzes cholesterol ester and other fatty acid esters such as vitamin A ester. These unique activities and the large quantity of BAL in human milk and pancreatic juice suggest that this enzyme is physiologically important.

Human BAL contains 722 amino acid residues (Wang & Hartsuck, 1993). The N-terminal 538 residues comprise the catalytic domain (Downs et al., 1994). The recombinant catalytic domain itself has the full enzymic activity of BAL (Hansson et al., 1993; Downs et al., 1994). The amino acid sequence of the catalytic domain of BAL is homologous to those of acetylcholinesterase and thyroglobulin (Baba et al., 1991). However, unique in the structure of human BAL is a C-terminal region (C-tail: residues 539−722) which consists of 16 highly analogous motifs containing the consensus sequence PVPPTGDSGAP (Baba et al., 1991). BALs from different mammalian species vary considerably in the lengths of the C-tail. For example, rat BAL has four repeating motifs, and the bovine enzyme contains three repeats (Wang & Hartsuck, 1993). Since the C-tail region contains all the O-linked oligosaccharides, and since nearly one-third of the residues are prolines (Baba et al., 1991), the three-dimensional structure of the C-tail is likely in an extended conformation, resembling that of mucins. The structure and properties of C-tail are of considerable interest because they may provide insights into its physiological functions. Toward this end, we isolated the C-tail fragment from human milk BAL and studied the composition and glycosylation patterns of the C-tail region. These results are discussed below.

EXPERIMENTAL PROCEDURES

Materials. Cyanogen bromide-activated Sepharose was obtained from Pharmacia. Heparin was obtained from ICN. Heparin−Sepharose was prepared as described previously (Wang & Johnson, 1983). The Superose 12 gel-permeation

† This work was supported by Research Grants HD-23472 and CA-37627 from the National Institutes of Health, Grant HR4-026 from the Oklahoma Center for the Advancement of Science and Technology, a grant from the American Heart Association, Oklahoma Affiliate, and a grant from Astra Hässle.

* Please send correspondence and reprint requests to this author at Oklahoma Medical Research Foundation, 825 N. E. 13th Street, Oklahoma City, OK 73104. Tel. (405) 271-7284. Fax (405) 271-3980.

‡ Oklahoma Medical Research Foundation.

§ Department of Biochemistry and Molecular Biology, University of Oklahoma Health Sciences Center.

|| William K. Warren Medical Research Institute, University of Oklahoma Health Sciences Center.

® Abstract published in *Advance ACS Abstracts,* August 1, 1995.

[1] Abbreviations: BAL, bile salt-activated lipase; C-tail, C-terminal region of human milk BAL; BSA, bovine serum albumin; PBS, 67 mM potassium phosphate, 0.15 M NaCl, and 0.02% NaN_3, pH 7.4.

0006-2960/95/0434-10639$09.00/0

column was obtained from Pharmacia. All other chemicals were obtained from Sigma Chemical Co.

Isolation of C-Tail Fragment. For the isolation of C-tail, human milk BAL was first partially purified using heparin–Sepharose affinity chromatography as described previously (Wang & Johnson, 1983). Three hundred milliliters of human skim milk was first centrifuged at 20 000 rpm for 2 h. The supernate was filtered through filter paper and allowed to flow through a heparin–Sepharose column (bed volume, 1.5 × 14 cm) which had been preequilibrated with 50 mM Tris-HCl buffer, pH 8.0. The column was then washed with 200 mL of the equilibrating buffer, and the BAL was eluted with 0.3 M NaCl in the same buffer and collected in 5-mL fractions. The fractions were assayed for the esterase activity of BAL using *p*-nitrophenyl acetate as substrate (Wang, 1991). Active fractions were pooled and dialyzed overnight against distilled water and then lyophilized. The yield was about 100–150 mg of dried material from each batch. The purity of this material was about 60–70% BAL as judged by the specific activity (Wang & Johnson, 1983). For preparation of the C-tail fragment, 350 mg of the above BAL preparation was incubated in 8 M urea (10 mg/mL) for 1 h and dialyzed overnight against 50 mM Tris-HCl buffer, pH 8.0. The denatured BAL was digested with chymotrypsin and trypsin (substrate to proteases ratio, 50:1, w/w) and with 0.1 mM $CaCl_2$ for 4 h at 37 °C. The same amounts of trypsin and chymotrypsin were again added to this mixture, and the incubation was continued overnight at 37 °C. The digest was dialyzed against distilled water and lyophilized. This material was subjected to CNBr (70 mg) cleavage in 7 mL of 70% formic acid (50 mg/mL) at room temperature overnight. The solution was diluted 10-fold with distilled water and lyophilized. This material was dissolved in distilled water and dialyzed. The dialysate was centrifuged to remove insoluble material, and the supernate was lyophilized. The dried material (about 50–80 mg) was dissolved in 2 mL of 50 mM Tris-HCl buffer, pH 8.0, containing 0.15 M NaCl. Two chromatographies were performed for 1 mL each of this solution using FPLC (fast protein liquid chromatography, Pharmacia) with two tandem-linked Superose 12 columns. The eluted fractions were monitored by absorbance at 280 nm and by carbohydrate analysis (Dubois et al., 1956).

N-Terminal Amino Acid Sequence Analysis. Automated Edman degradations were performed as described by Hewick et al. (1981) in the Molecular Biology Resource Center at University of Oklahoma Health Science Center using a Model 470A gas-phase protein sequencer equipped with a Model 120A on-line phenylthiohydantoin amino acid analyzer (Perkin Elmer, Applied Biosystems Division).

Amino Acid and Hexosamine Analysis. Samples for amino acid and hexosamine analyses were hydrolyzed in 5.7 N HCl in sealed evacuated tubes at 108 °C for 20 h and in 4 N HCl at 100 °C for 4 h, respectively. All analyses were performed on a fully automated System Gold amino acid analyzer (Beckman Instruments, Inc., San Ramon, CA) by cation-exchange chromatography and postcolumn detection by reaction with ninhydrin. Amino acids were separated by a two-buffer elution system (sodium citrate, pH 3.28 and pH 7.4) as described by the manufacturer. Hexosamines were separated in the same instrument with an initial buffer of 0.2 N sodium citrate, pH 4.25. In this elution, hexosamines are eluted after all acidic and neutral amino acids and are readily quantitated. β-Thienylalanine was added to each hydrolysate as internal standard.

β-Elimination Reaction. For the release of O-glycosidically linked oligosaccharide, 1 mg of C-tail in 1 mL of 0.1 N NaOH was incubated at 4 °C for 5 days. The progress of the reaction was monitored by measuring absorbance at 242 nm (Neuberger et al., 1966). The $T_{1/2}$ of this reaction was determined to be 15 h.

Carbohydrate Analysis by Gas–Liquid Chromatography. Neutral sugars and hexosamines were determined by gas–liquid chromatography in their alditol acetate forms. The samples (0.2–0.5 mg) were hydrolyzed separately in 1, 2, and 4 N HCl in sealed evacuated tubes for 4 h at 100 °C. After the samples were cooled on ice, 25 μg of xylose was added to each of the hydrolysates. The samples were dried under vacuum with a rotatory evaporator, reduced with sodium borohydride, acetylated, and analyzed by gas–liquid chromatography as described by Griggs et al. (1971).

N-Acetylneuraminic Acid Analysis. C-tail (20 μg) was treated with 10 milliunits of *Arthrobacter ureafaciens* neuraminidase (obtained from Sigma) in 20 μL of 0.1 M sodium acetate buffer, pH 4.8, at 37 °C overnight. After the reaction, the *N*-acetylneuraminic acid released was quantified by Dionex high-pH anion-exchange chromatography (HPAEC) using a Carbopac PA 1 column (4 × 250 mm). The sample was separated with an isocratic gradient of 150 mM sodium acetate in 100 mM sodium hydroxide for 15 min. The flow rate was 1 mL/min, and the separation was monitored with a pulsed amperometric detector. Quantitation was done by calculating the peak area of the *N*-acetylneuraminic acid released by neuraminidase on the basis of the peak area of the *N*-acetylneuraminic acid standard.

Periodate Oxidation. For periodate oxidation (Tarcsay et al., 1973) the C-tail fragment (30 mg) was incubated in 5 mL of 0.05 M sodium metaperiodate for 24 h at room temperature. The excess periodate was eliminated by incubating the mixture with 0.5 mL of ethylene glycol for 1 h at room temperature. The solution was dialyzed against distilled water overnight. The dialysate was adjusted to a final concentration of 0.5 N HCl and incubated at 37 °C for 6 h in order to remove the degraded sugar. The mixture was dialyzed against distilled water and lyophilized.

Digestion of C-Tail Fragment with Pronase. Ten milligrams of the C-tail fragment was incubated with 1 mg of pronase in 1 mL of Tris-HCl buffer, pH 8.0, at 37 °C, first for 4 h first and then overnight after the addition of another 1 mg of pronase. The mixture was applied to the Superose 12 column as described above, and the fractions were analyzed for carbohydrate content. The pronase-digested C-tail fragment was found to elute with an elution peak volume of 28 mL.

Detection of Lewis x, Sialyl Lewis x, and Lewis a Epitopes on Human Milk BAL and the C-Tail by Enzyme-Linked Immunosorbant Sandwich Assay (ELISA). The wells of polystyrene microtiter plates, Immulon 4 (Dynatech Laboratories, Chantilly, VA), were coated with 100 μL of carbonate coating buffer (carbonate-buffered saline: 150 mM Na_2CO_3, 348 mM $NaHCO_3$, and 0.02% NaN_3, pH 9.6) containing different concentrations of BAL or BAL C-tail and incubated overnight at room temperature. After being washed with PBS (67 mM potassium phosphate, 0.15 M NaCl, and 0.02% NaN_3, pH 7.4), the wells were blocked

with 250 μL of 5% bovine serum albumin (BSA) in PBS at room temperature for 2 h. The blocking solutions were aspirated, and the wells were washed with PBS followed by deionized water. For neuraminidase treatment, the wells were incubated with 100 μL of 0.1 M sodium acetate buffer (pH 4.8) containing 50 milliunits/mL *Arthrobacter ureafaciens* neuraminidase (from Sigma) at 37 °C for 4 h. After treatment, the wells were washed with PBS containing 0.05% Tween 20. The wells were then incubated at room temperature with 100 μL of dilution buffer (PBS/Tween 20 with 0.1% BSA) containing 2 μg/mL of the following antibodies: anti-Lewis x antibodies (anti-human CD15) (Leinco Technologies, Inc.) and anti-Lewis a antibodies (Signet Laboratory, Inc.). The wells were washed with PBS/Tween 20 following removal of the antibody solution. The wells were then incubated at room temperature with 100 μL of goat anti-mouse IgM (alkaline phosphatase conjugate; 1:500 dilution with dilution buffer) for anti-Lewis x and anti-sialyl Lewis x antibodies and with goat anti-mouse IgG (alkaline phosphatase conjugate; 1:500 dilution with dilution buffer) for anti-Lewis a antibodies. After 1 h, the solution was removed by aspiration and the wells were washed with washing solution and deionized water. The wells were treated with 100 μL of freshly prepared alkaline phosphatase substrate solution (25 mM $MgCl_2$ and 20 mg of *p*-nitrophenyl phosphate in 20 mL of coating buffer) and incubated at 37 °C until a yellow color was developed. The optical density at 405 nm of each well was recorded by a V-Max kinetic microplate reader (Molecular Devices Corp., Menlo Park, CA). Each assay was performed in triplicate.

Binding of C-Tail with Jacalin Lectin. For examining the reactivity of C-tail with jacalin lectin, all the procedures were the same as in the ELISA assay above except that biotinylated jacalin lectin (obtained from Vector Laboratories) and streptavidin−alkaline phosphatase (1:10 000 dilution) were used. As a control for specific lectin interaction, control wells were incubated with the lectin in the presence of hapten sugar (200 mM α-methylgalactose).

RESULTS

Isolation of the C-Tail Fragment. Human milk BAL was first partially purified with heparin−Sepharose affinity chromatography. The isolated BAL was then digested with chymotrypsin and trypsin followed by cyanogen bromide reaction, and subjected to FPLC on a Superose 12 column. Carbohydrate content of the fractions indicated that the C-tail fragment was eluted in fractions 21−25 (Figure 1). Absorbancy at 280 nm showed that most peptides are much smaller than the C-tail fragment and are eluted after fraction 30. A carbohydrate-containing peak found in fractions 31−34 represents N-linked glycopeptides derived from the proteolysis of the catalytic domain. The yield of the isolated C-tail fragment was about 20−30 mg per liter of the starting human milk. The amino acid sequence in and near the C-tail region predicts that either proteolytic or CNBr cleavage would result in an intact C-tail fragment. However, we found that, in contrast to individual cleavages, the combination of the two cleavages produced a material relatively free of aggregation; therefore, it was much easier to purify the C-tail fragment. The elution position of the C-tail fragment in gel filtration (Figure 1) corresponds to a molecular weight near 220 000, considerably larger than predicted from the sequence. This overestimation is common among mucin-like

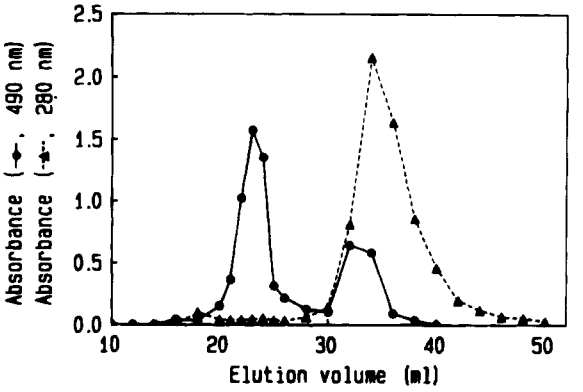

FIGURE 1: FPLC (fast protein liquid chromatography) fractionation pattern of human milk BAL after treatment with trypsin and chymotrypsin, followed by cyanogen bromide digestion. The digested BAL was dialyzed and lyophilized, and 40 mg of the sample was dissolved in 1 mL of 0.15 M NaCl and 50 mM Tris-HCl, pH 8.0, buffer solution. Fractionation was performed using two tandem-linked Superose 12 molecular sieving columns. The columns were eluted with the same buffer solution at a flow rate of 0.5 mL/min. The fractions contained 1 mL/tube. The column eluates were monitored with the UV absorbance at 280 nm. Aliquots of 0.1 mL were also taken for assay of carbohydrate content (Dubois et al., 1956) using the phenol−H_2SO_4 assay method and monitored at 490 nm.

polypeptides with high contents of proline and O-glycosylation (Timpte et al., 1988).

The N-terminal sequence of the isolated C-tail fragment indicates that the fragment starts at Ala^{528}, while the amino acid composition suggests that it ends at Lys^{712} of human BAL (Figure 2 and Table 1). Thus, the isolated C-tail fragment contains the entire 16 repeating motifs. The N-terminal Ala^{528} of the fragment is apparently the result of chymotryptic digestion of a Leu^{527}−Ala^{528} bond. Present in some preparations is also a minor (<20%) N-terminal Leu^{527} resulting from the chymotryptic cleavage of the Tyr^{526}−Leu^{527} bond. In preparations in which Ala is the only N-terminal residue, the N-terminal leucine may have been further removed by chymotryptic activity. The C-terminal Lys^{712} of the C-tail fragment is apparently the product of a tryptic cleavage.

Carbohydrate Analyses. Previous carbohydrate analyses (Wang, 1981), which were done on native human milk BAL, included both N- and O-linked oligosaccharide units. The isolation of the C-tail fragment thus provides an opportunity to analyze the carbohydrates of the O-linked oligosaccharide alone. The C-tail fragment was found to contain fucose, galactose, glucosamine, and galactosamine in a molar ratio of 1:3:2:1 (Table 2). This ratio suggests that each oligosaccharide unit contains an average of seven hexose residues. Assuming galactosamine is the usual anchoring sugar for O-glycosylation (Neuberger et al., 1966), the quantity of this amino sugar suggests the presence of about 9−11 O-glycosylation sites in the C-tail region. Also shown in Table 2, the presence of less than 1 mol of sialic acid per mole of galactosamine indicates that there is heterogeneity among the oligosaccharide chains.

Glycosylation Sites. The nature of O-glycosylated residues was first studied using the β-elimination reaction of the C-tail fragment. The losses of involved serine or threonine were analyzed by amino acid composition. Figure 3 shows that the reaction in 0.1 N NaOH was near completion at 70 h. β-Elimination of C-tail resulted in the loss of nine threonine

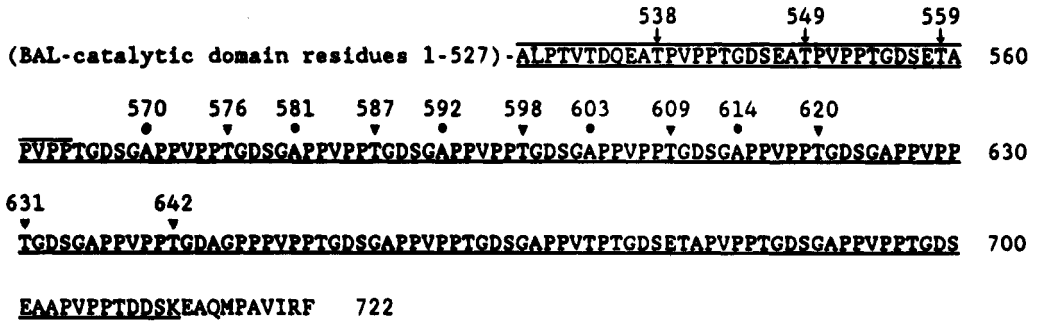

FIGURE 2: Amino acid sequence of the C-terminal domain of human milk BAL. Underlining indicates the sequence of the purified C-tail. The overline represents the sequence determined by Edman degradation of the C-tail. The O-glycosylation sites are identified as the missing residues in the sequence analysis and are marked with ↓. The pronase-digested subfragment has the sequence APPVPPXGDSGAPPVP-PXGDSGAPPVPPXGD by Edman degradation. In this sequence X is the missing residue. The possible start site (Ala) of this sequence is marked with ●. The possible identified missing Thr residues positions are marked with ▼.

Table 1: Amino Acid Composition of Isolated C-Tail of Human Milk BAL before and after Treatment with Alkali (0.1 N NaOH)

	no. of residues (set Lys = 1.0)		
residue	from known sequence	experimental	after β-elimination
Lys	1	1.0	1.0
Asp	18	18	20
Thr	23	22	14
Ser	15	14	14
GlX	6	6	8
Pro	60	57	60
Gly	26	26	28
Ala	18	18	20
Val	17	17	18
Leu	1	1	2
total	185		

Table 2: Carbohydrate Composition of C-Tail and Periodate-Treated C-Tail of Human Milk BAL

	C-tail			periodate-treated C-tail	
monosaccharide	μmol/mg	molar ratio[a]	mol/mol C-tail[b]	μmol/mg	molar ratio[a]
fucose	0.32	1	9.0	0.01	0
galactose	0.95	3	26.6	0.35	1
N-acetylglucosamine	0.69	2	19.3	0.72	2
N-acetylgalactosamine	0.36	1.0	10.1	0.34	1.0
sialic acid	0.10	0.3	2.8	nd[c]	nd[c]

[a] Relative composition upon setting *N*-acetylgalactosamine equal to 1.0. [b] Based on a C-tail molecular weight of 28 000. [c] Not determined.

residues and less than one serine residue. These results predict that O-glycosylation in the C-tail region resides, predominantly, if not exclusively, on threonine residues. They also agree well with the number of sites predicted from the galactosamine content (Table 2).

Automated Edman degradation of the isolated C-tail fragment provides evidence for some of the O-glycosylation sites in the N-terminal region of the fragment (Figure 2). Potential glycosylation sites were identified by Edman degradation on the basis of the retention of O-glycosylated PTH-amino acids in the aqueous phase, hence the absence of HPLC-identifiable PTH-amino acids at these steps. From the sequence data Thr^{538}, Thr^{549}, and Thr^{559} were identified.

In order to gain more structural information, the C-tail fragment was subjected to further proteolytic digestions. Among the proteolytic enzymes tested (elastase, collagenase, pepsin, endoproteinase Glu-C, and pronase), only pronase

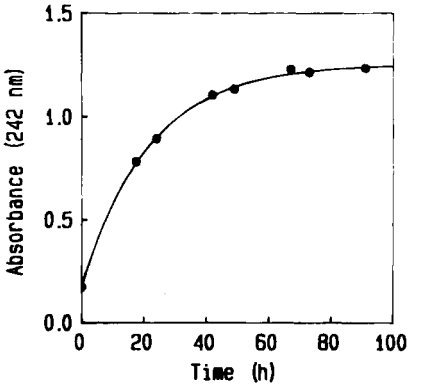

FIGURE 3: Kinetics of the β-elimination reaction of the C-tail treated with alkali. The reaction was performed using the C-tail at 1 mg/mL in 0.1 N NaOH at room temperature.

was effective and resulted in a changed elution position of the digestion product in Superose 12 column chromatography, from 23 to 28 mL. Automated Edman degradation of this pronase-generated fragment gave rise to an N-terminal amino acid sequence of APPVPPXGDSGAPPVPPXGDS-GAPPVPPXGD, where X denotes blanks in PTH-amino acid identification, which indicates that the potential start site of this peptide can be Ala^{570}, Ala^{581}, Ala^{592}, Ala^{603}, or Ala^{614}. From the known cDNA sequence other alanine residues can be eliminated as being the N-terminus of this pronase-digested C-tail. The three potential glycosylation sites identified by Edman degradation are indicated in Figure 2.

Immunochemical and Lectin Binding Properties of the O-Linked Oligosaccharides of C-Tail. The high content of fucose in BAL suggested that the glycoprotein might contain the Lewis x or Lewis a antigen. These are composed of the terminal trisaccharide sequences Gal(β1→4)[Fuc(α1→3)]-GlcNAc(β1→)R and Gal(β1→3)[Fuc(α1→4)]GlcNAc(β1→)R, respectively.

Immunochemical studies indicated that BAL is strongly recognized by anti-Lewis x and weakly recognized by anti-sialyl Lewis x antibodies (data not shown).

Subsequent studies indicated that the purified C-tail also reacts with anti-Lewis x antibodies (Figure 4A). Since the C-tail does not contain any N-linked oligosaccharide, the Lewis x antigen must be associated with the O-linked oligosaccharides in the C-tail. Further immunochemical studies using anti-Lewis a antibodies indicated that the C-tail also has the Lewis a antigen. These observations again suggest heterogeneity among the oligosaccharide chains of

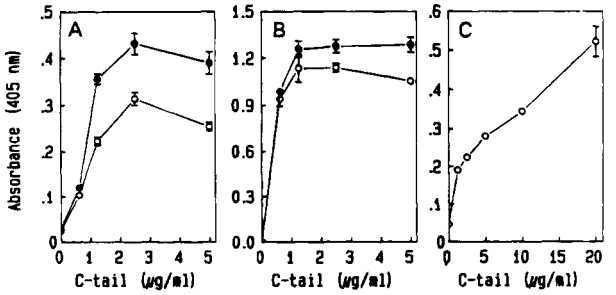

FIGURE 4: Reactivity of the C-tail with anti-Lewis x and Lewis a antibodies and jacalin lectin. (A) Effect of neuraminidase treatment on the C-tail reactivity with anti-Lewis x antibodies: (○) no treatment; (●) neuraminidase treated. (B) Effect of neuraminidase treatment on the C-tail reactivity with anti-Lewis a antibodies: (○) no treatment; (●) neuraminidase treated. (C) Reactivity of C-tail with jacalin lectin.

the C-tail. The treatment of C-tail with neuraminidase resulted in a 20−30% increase of anti-Lewis x reactivity, and also a small but detectable increased reactivity with anti-Lewis a antibodies. These results indicated that some of the Lewis x and Lewis a oligosaccharides may be sialylated. This might explain the detectable reactivity of the glycoprotein with anti-sialyl Lewis x antibody.

We have also observed that C-tail binds jacalin lectin (Figure 4C). Jacalin lectin binds with high affinity to glycoproteins containing O-linked oligosaccharides with the core 1 structure Gal(β1→3)GalNAcα1→Ser/Thr (Hortin & Trimpe, 1990). The binding of jacalin lectin to the C-tail is completely consistent with the high amount of O-linked oligosaccharide predicted by the compositional analysis and results of β-elimination.

Nonreducing End Sugars in the O-Linked Oligosaccharide Units. Periodate oxidation of the C-tail fragment resulted in the reduction of one fucose and two galactose residues, while there was no change in glucosamine and galactosamine (Table 2). This is consistent with the presence of 1 mol of fucose and 2 mol of galactose per mole of galactosamine residue as the major hexoses at the nonreducing termini of the O-linked oligosaccharides.

DISCUSSION

The method described above for the preparation of the C-tail fragment from human milk BAL is simple to perform and suitable for large-scale preparation. From combined proteolytic and chemical cleavages, a fragment encompassing the entire 16 repeating motifs and all of the O-glycosylation sites is isolated.

The discrepancy between the calculated molecular weight of the C-tail fragment (peptide portion: 17 015) and its molecular weight determined by gel filtration chromatography (220 000) apparently stems from the conformation of this polypeptide. Within the 16 repeating motifs of the C-tail region, nearly one-third of the residues are proline, and very few are hydrophobic residues. There are also about 10 units of O-linked oligosaccharides in this region. This type of structure, which resembles that of many mucins, is likely to be in an extended conformation which behaves like a much larger molecule in gel filtration. On the other hand, the recombinant BAL catalytic domain (truncated BAL without the C-tail) gave rise to a molecular weight of 60 000 in gel filtration chromatography (Downs et al., 1994), very near the true molecular size, as can be expected for the globular conformation of the catalytic domain. The overestimations of the apparent molecular weight of the native human BAL reported from different studies (Wang & Hartsuck, 1993) are attributable to the extended conformation of the C-tail region.

From the Edman degradation, potential glycosylation sites were identified at three threonine residues in the C-tail fragment, and three potential glycosylated threonine residues were located in the pronase subfragment (Figure 2). Although the identifications were made by the absence of PTH-amino acids at these positions, we feel that they are likely to be correct for two reasons. First, clear identification of PTH-threonine residues was made in other non-glycosylated threonine positions. Second, a definitive pattern in the sequences associated with the O-glycosylation has emerged. Two of the threonine residues (positions 538 and 549) are followed by a PVPP motif, and Thr^{559} is followed by a similar APVPP motif. Three other potential O-linked threonine residues are adjacent to but preceded by PVPP motifs. The precise positions of these threonine residues are not identified. The N-terminal sequence of the pronase subfragment of the C-tail produced 31 cycles which include nearly three repeating motifs of APPVPPXGDSG. Starting from residue 570, however, this motif is repeated six times without residue replacement (Figure 2). Five N-terminal positions (570, 581, 592, 603, and 614) for the pronase subfragment would produce the same result in the Edman degradation. Amino acid composition of this pronase fragment (results not shown) was uninformative because of many tandem repeats, and the molecular weight estimation by gel filtration was equally uninformative because of the gross overestimation discussed above. Since the sequences surrounding the five potential N-termini are identical, it is probable that the pronase subfragment is a mixture with five starting points. If this is the case, the fact that glycosylation takes place at the threonine residues in all three motifs following PVPP suggests that threonine residues at positions 576, 587, 598, 609, 620, 631, and 642 are all fully glycosylated. The total glycosylation sites suggested from Edman degradation (11) is near to that estimated by the other methods discussed above. These analyses also suggest that little glycosylation is associated with the threonine residues in the region from residue 643 to the C-terminus. We were unable to experimentally reach these residues because of the resistance to cleavage by this region. Our results clearly indicate that the glycosylation occurs on threonine residues. The rule of glycosylation, in the case of BAL, seems to favor the residues either before or after the PVPP motif. When threonine residues are present at both ends of the PVPP motif, as in two places between residues 538 and 554, only the residue before the PVPP motif is glycosylated. The observed O-glycosylation pattern of Thr with the neighboring PVPP motif is among the most favored sequences compiled by Elhammer et al. (1993). In the Edman degradation of the fragments, none of the serines so far sequenced were found to be glycosylated, in agreement with results from the β-elimination reaction (Table 1). This apparent O-glycosylation specificity against serine residues is probably related to their distal positions to the PVPP motifs.

The sugar composition of the O-linked oligosaccharides in the C-tail fragment produced a molar ratio of fucose:galactose:glucosamine:galactosamine of 1:3:2:1. Despite the presence of microheterogeneity, the simplest hypothesis from

these data is that each oligosaccharide unit consists of an average of seven hexoses. Since *N*-acetylgalactosamine is the anchoring sugar of the O-glycosylated oligosaccharide chain and there are only 9−11 *N*-acetylgalactosamine residues per mole of the C-tail, it seems likely that all of the *N*-acetylgalactosamine residues are involved in anchoring. The observation that BAL is positive to Lewis x and Lewis a antigenic structures suggests the presence of terminal, nonreducing end galactose and fucose. Periodate oxidation established that there are 2 mol of terminal galactose and 1 mol of terminal fucose per mole of galactosamine. This would indicate that there is one additional nonreducing end galactose of which the site of attachment remains to be determined. The agglutinin reactivity of human pancreatic BAL by Mas et al. (1993) and the jacalin lectin reactivity of C-tail of human milk BAL found in this study indicate that it is possible that one of the galactose residues could attach to the *N*-acetylgalactosamine with a $\beta 1 \rightarrow 3$ linkage. A more detailed study of the oligosaccharide structure of the C-tail is presently underway in our laboratories.

Current evidence suggests that the C-tail region of BAL has a mucin-like structure with an extended conformation. It contains threonine-linked oligosaccharides, at least some of which are related to the Lewis x structure. It is now known that many cell adhesion proteins, such as selectins, have ligands in mucin-like structures (Norgard et al., 1993; Lasky, 1992; Baumhueter et al., 1993; Sako et al, 1993). In the case of P-selectin, the ligand recognition is also thought to include sialyl Lewis x oligosaccharides (Moore et al., 1994). These comparisons suggest the possibility that the C-tail region may contribute an adhesive activity to BAL.

ACKNOWLEDGMENT

We wish to thank Dr. P. Alaupovic for providing the facility for gas−liquid chromatography analyses and Mr. Randy Whitmer for performing the carbohydrate analysis of the alditol acetate derivative of the C-tail. We wish to acknowledge the Molecular Biology Resource Facility of the University of Oklahoma Health Sciences Center for performing the N-terminal sequencing analysis and the amino acid composition analysis. We also wish to thank Mrs. Joan Pilcher for preparation of the manuscript.

REFERENCES

Alemi, B., Hamosh, M., Scanlon, J. W., Salzman-Mann, C., & Hamosh, P. (1981) *Pediatrics 68,* 484−489.

Baba, T., Downs, D., Jackson, K., Tang, J., & Wang, C.-S. (1991) *Biochemistry 30,* 500−510.

Baumhueter, S., Singer, M. S., Henzel, W., Hemmerich, S., Renz, M., Rosen, S. D., & Lasky, L. A. (1993) *Science 262,* 436−438.

Downs, D., Xu, Y.-Y., Tang, J., & Wang, C.-S. (1994) *Biochemistry 33,* 7979−7985.

Dubois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. A., & Smith, F. (1956) *Anal. Chem. 28,* 350−356.

Elhammer, Å. P., Poorman, R. A., Brown, E., Maggiora, L. L., Hoogerheide, J, G., & Kézdy, F. J. (1993) *J. Biol. Chem. 268,* 10029−10038.

Griggs, L. J., Post, A., White, E. R., Finkelstein, J. A., Moeckel, W. E., Holden, K. G., Zarembo, J. E., & Weisbach, J. A. (1971) *Anal. Biochem. 43,* 369−381.

Hansson, L., Bläckberg, L., Edlund, M., Lundberg, L., Strömqvist, M., & Hernell, O. (1993) *J. Biol. Chem. 268,* 26692−26698.

Hewick, R. M., Hunkapiller, M. W., Hood, L. E., & Dreyer, W. J. (1981) *J. Biol. Chem. 256,* 7990−7997.

Hortin, G. L., & Trimpe, B. L. (1990) *Anal. Biochem 188,* 271−277.

Lasky, L. A. (1992) *Science 258,* 964−969.

Mas, E., Abouakil, N., Roudani, S., Franc, J.-L., Montreuil, J., & Lombardo, D. (1993) *Eur. J. Biochem. 216,* 807−812.

Moore, K. L., Eaton, S. F., Lyons, D. E., Lichenstein, H. S., Cummings, R. D., & McEver, R. P. (1994) *J. Biol. Chem. 269,* 23318−23327.

Neuberger, A., Gottschalk, A., & Marshall, R. D. (1966) in *Glycoproteins* (Gottschalk, A., Ed.) pp 273−295, Elsevier, Amsterdam.

Norgard, K. E., Moore, K. L., Diaz, S., Stults, N. L., Ushiyama, S., McEver, R. P., Cummings, R. D., & Varki, A. (1993) *J. Biol. Chem. 268,* 12764−12774.

Olivecrona, T., & Bengtsson, G. (1984) in *Lipases* (Borgström, B., & Brockman, H. L., Eds.) pp 205−261, Elsevier, Amsterdam.

Reue, K., Zambaux, J., Wong, H., Lee, G., Leete, T. H., Ronk, M., Shively, J. E., Sternby, B., Borgström, B., Ameis, D., & Schotz, M. C. (1991) *J. Lipid Res. 32,* 267−276.

Sako, D., Chang, X.-J., Barone, K. M., Vashino, G., White, H. M., Shaw, G., Veldman, G. M., Bean, K. M., Ahern, T. J., Furie, B., Cumming, D. A., & Larsen, G. R. (1993) *Cell 75,* 1179−1186.

Tarcsay, L., Wang, C.-S., Li, S.-C., & Alaupovic, P. (1973) *Biochemistry 12,* 1948−1955.

Timpte, C. S., Eckhardt, A. E., Abernethy, J. L., & Hill, R. L. (1988) *J. Biol. Chem. 263,* 1081−1088.

Wang, C.-S. (1981) *J. Biol. Chem. 256,* 10198−10202.

Wang, C.-S. (1991) *Biochem. J. 279,* 297−302.

Wang, C.-S., & Johnson, K. (1983) *Anal. Biochem. 133,* 457−461.

Wang, C.-S., & Hartsuck, J. A. (1993) *Biochim. Biophys. Acta 1166,* 1−19.

Wang, C.-S., Martindale, M. E., King, M. M., & Tang, J. (1989) *Am. J. Clin. Nutr. 49,* 457−463.

BI950695U