Secretory IgA (SIgA) is a multi-polypeptide complex consisting of a secretory component (SC) covalently attached to dimeric IgA containing one joining (J) chain. We present the analysis of both the N- and O-glycans on the individual peptides from this complex. Based on these data, we have constructed a molecular model of SlgA1 with all its glycans, in which the Fab arms form a T shape and the SC is wrapped around the heavy chains. The O-glycan regions on the heavy (H) chains and the SC N-glycans have adhesin-binding glycan epitopes including galactose-linked β1–4 and β1–3 to GlcNAc, fucose-linked α1–3 and α1–4 to GlcNAc and α1–2 to galactose, and α2–3 and α2–6-linked sialic acids. These glycan epitopes provide SIgA with further bacteria-binding sites in addition to the four Fab-binding sites, thus enabling SIgA to participate in both innate and adaptive immunity. We also show that the N-glycans on the H chains of both SlgA1 and SlgA2 present terminal GlcNAc and mannose residues that are normally masked by SC, but that can be unmasked and recognized by mannose-binding lectin, by disrupting the SC-H chain noncovalent interactions.

Secretory IgA (SIgA) is the major immunoglobulin responsible for protecting the mucosal surfaces against invasion by pathogens. In humans, mucosa covers a vast surface area (~400 m²), and the body produces more SIgA each day than all pathogens. In humans, mucosa covers a vast surface area (~400 m²), and the body produces more SIgA each day than all pathogens. In humans, mucosa covers a vast surface area (~400 m²), and the body produces more SIgA each day than all pathogens.

SIgA with attached J chain is produced in plasma cells close to the epithelium. The epithelial cells express the polymeric immunoglobulin receptor (pIgR) that binds to dimeric IgA; this complex is then translocated across the epithelial cell. During translocation disulfide bonding occurs between the pIgR and one H chain. On reaching the mucosal surface, the (50–90-kDa) secretory component (SC) is cleaved from the pIgR transmembrane tail, and the whole IgA/J chain/SC (SIgA) complex is secreted. Thus, SIgA is a multimeric glycoprotein complex originating from two cell types (3). This is in contrast to serum IgA, which is predominantly monomeric and lacks the J chain and SC.

There are two isotypic forms of IgA: IgA1 and IgA2. Both forms contain two conserved N-glycan sites per H chain, one at Asn263 in the Co2 domain and one on the terminal amino acid (Asn459) of its 18-amino acid tail piece (compared with IgG, IgA has an extra 18 amino acids on the C-terminal of the H chain). There are two allotypes of IgA2, IgA2m(1) and IgA2m(2), that have an extra 18 amino acids on the C-terminal of the H chain. There are two allotypes of IgA2, IgA2m(1) and IgA2m(2), that contain further conserved N-glycan sites: one on the Ca2 domain and one or two on the Ca1 domain respectively (1) (Fig. 1). SC is highly glycosylated, it has seven N-glycan sites, and sugars contribute up to 25% of its molecular mass, whereas the J chain has only one N-glycan site (3). In addition, IgA1 has a 23-amino acid, proline-rich hinge region with nine potential O-glycosylation sites (serine and threonine residues) of which three to five sites have been shown to be occupied in serum IgA1 (4–7). IgA2 lacks this 13-amino acid hinge region and is not O-glycosylated.

A newborn relies on passive immunity from the SIgA in its mother’s milk until its own immune system has matured. The SIgA in colostrum and milk binds to microorganisms, their metabolic products and toxins, preventing their attachment to the gut epithelium and facilitating their expulsion in the feces, a process known as immune exclusion (8). The adhesion of many pathogenic organisms to mucosal membrane cells is mediated by adhesins on their surface. These are lectin-like receptors that can bind to complementary carbohydrate constituents expressed by the host tissues (9). For example, S-fimbriated Escherichia coli, which causes sepsis and meningitis in newborn, can be prevented from binding to epithelial cells, independently of the antigen-binding sites, by siaIyated glycans on SIgA that bind to the pathogen (10). Fucose, linked α1–2 to galactose as in Lewisα and Lewisβ epitopes, on SC N-glycans compete with Helicobacter pylori for binding to gastrotic receptors (11, 12). Free SC also binds to E. coli (13, 14) and toxin A from Clostridium difficile (15), and both free and SIgA-bound SC interact specifically with a surface protein of Strep-tonococcus pneumoniae (16, 17). Type 1-fimbriated E. coli express a mannose-specific lectin that binds to SIgA (13). This has led...
to an increased interest in the use of orally administered recombinant SlgA for passive immunization against virulent pathogens such as *C. difficile* and *Neisseria meningitidis* (18, 19).

However, accurately defining glycan structures is a prerequisite to understanding their binding properties, particularly because proteins made in different cell types are differently glycosylated, and this can have major implications when producing engineered antibodies (20).

In this paper, we present the first total N- and O-glycan analysis of each of the different peptide chains (H, J, and SC) from the same sample of normal pooled human SLgA. Sensitive analytical procedures have enabled us to identify minor components of the glycan pool in addition to confirming the major glycan structures found by previous investigators (21–28). The O-glycans on the H chain and the N-glycans on SC presented a wide range of epitopes for adhesin binding. Over 75% of the N-glycans on the J chain were sialylated, whereas over 66% of the N-glycans on the H chain were truncated complex structures with free terminal GlcNAcs.

On the basis of these data, we have constructed a molecular model of SlgA1 with its glycans attached, in which the Fab arms form a T shape and the SC is wrapped around the H chains. This model shows that glycans cover most of the SlgA1 complex with the exception of the Fab regions. Each SlgA1 molecule has several sites for binding to pathogens. In addition to the four Fab antigen-binding sites (adaptive immunity), there are two O-glycosylated regions containing up to 10 glycans per region and SC with seven N-glycans, which present a wide range of sugar epitopes capable of binding to adhesins (innate immunity).

Our model of SlgA1 shows the SC wrapped around the H chains, masking the H chain N-glycans, which are truncated complex structures with free terminal GlcNAc residues. These GlcNAc residues are potential ligands for lectins such as mannos-binding lectin (MBL). MBL is a calcium-dependent serum lectin that is able to bind to α-mannose, L-fucose, GlcNAc, and N-acetyllactosamine, recognizing two common equatorial hydroxyl groups (29). Binding of MBL to microorganisms, including viruses, bacteria, and yeast species, can induce activation of the complement system via the lectin pathway, thus leading to target opsonization. MBL represents a key component of the innate immune system, as illustrated by the increased susceptibility to infections occurring in MBL-deficient individuals (30). Recently, binding of the lectin domain of MBL to polymeric serum IgA has been reported (31). We show that MBL does not bind to SlgA at neutral pH, supporting the proposal that SC is masking the H chain N-glycans, and that disruption of the SC-IgA noncovalent interactions, by preincubation at pH 3, unmasks the H chain N-glycans, allowing MBL to bind.

**EXPERIMENTAL PROCEDURES**

**Secretory IgA**—Purified human SlgA, purified from colostrum, was obtained from Sigma. The secretory IgA complex was reduced, alkylated, and then separated into SC, J chain, H chain, and light chain by reducing SDS-PAGE (80 × 80 × 1 mm, 10% BisTris NuPAGE gel, MES SDS running buffer [Invitrogen]). The protein bands were visualized by Coomassie staining. MultiMark molecular mass standards were used (Invitrogen). Relative amounts of IgA1 and IgA2 were determined by ELISA according to Ref. 32 using normal human serum with known concentrations of IgA1 (2.1 mg/ml) and IgA2 (0.2 mg/ml) as standards.

Identification of Gel Bands by Mass Spectrometry—Coomassie-stained bands were excised and in-gel digested with trypsin (sequencing grade; Roche Applied Science) as described in Ref. 33. The mixtures of recovered trypptic peptides were desalted by loading onto a PepMap C18 0.3 × 5 mm cartridge (LC packings; Presequence Ltd., Hitchin, UK) in water with 0.1% formic acid and then eluting with 80% acetonitrile, 0.1% formic acid at a flow of 0.2 µl min⁻¹ directly into a hybrid quadrupole time-of-flight mass spectrometer fitted with a nanospray source (Waters-Micromass Ltd., Manchester, UK). The peptides were sequenced from fragmentation data as described (34). A BLAST search of the NCBI data base was performed using Mascot software.

**Release and Fluorescent Labeling of Glycans—N-Glycans** were released from excised gel bands by in-gel digestion of the protein with peptide: N-glycosidase F (Roche Applied Science) (35). Manual hydrazinolysis (36, 37) was used to release O-glycans (60°C for 6 h) from the whole SlgA complex. Released glycans were fluorescently labeled with 2-aminobenzamide (2AB) by reductive amination according to the method of Bigge et al. (38) using an Oxford GlycoSciences Signal™ labeling kit (Oxford GlycoSciences, Abingdon, UK).

**Analysis of Glycans by High Performance Liquid Chromatography (HPLC)—Normal phase** (NP) HPLC was performed according to the low-salt buffer system as previously described (39) using a 4.6 × 250-mm GlycSep-X column (Oxford GlycoSciences). The system was calibrated using an external standard of hydrolyzed and 2AB-labeled glucose oligomers to create a dextran ladder. Weak anion exchange HPLC (40) was performed using a Vydac 301VHPS75 7.5 × 50-mm column (Anachem Ltd., Luton, Bedfordshire, UK) according to the modified methodology (41). These HPLC methods are described in detail in Ref. 37.

**Exoglycosidase Digestions—** Arrays of exoglycosidases were used in combination with HPLC to determine the sequence, monosaccharide type, and linkage of sugar residues as described in Ref. 37. The enzymes used were: *Arthrobacter ureafaciens* sialidase (EC 3.2.1.18), 1–2 units/ml; Newcastle disease virus (Hitcher B1 Strain) sialidase (EC 3.2.1.18), 1–2 units/ml; *Clostridium perfringens* α-N-acetyl glucosaminidase (EC 3.2.1.31), 10 units/ml; *Clostridium perfringens* β-galactosidase (EC 3.2.1.23), 10 units/ml; *Arthrobacter ureafaciens* α-L-fucosidase (EC 3.2.1.72), 10 units/ml; and *Aspergillus niger* α-1,6-glucosidase (EC 3.2.1.72), 10 units/ml.
0.2 unit/ml; S. pneumoniae sialidase recombiant in E. coli (EC 3.2.1.18) 1 unit/ml; bovine kidney a-fucosidase (EC 3.2.1.51), 1 unit/ml; almond meal a-fucosidase (EC 3.2.1.111), 3 milliunits/ml; bovine testes b-galactosidase (EC 3.2.1.23), 2 units/ml; S. pneumoniae b-galactosidase (EC 3.2.1.23), 80 milliunits/ml; S. pneumoniae b-N-acetylgalactosaminidase (EC 3.2.1.30), 120 milliunits/ml; Jack bean b-N-acetylhexosaminidase (EC 3.2.1.30), 10 milliunits/ml; and Jack bean a-mannosidase (EC 3.2.1.24), 50 units/ml (with a second aliquot added after 12 h). All of the enzymes were from Glyko Inc. (Novato, CA).

Molecular Modeling—Molecular modeling was performed on a Silicon Graphics Fuel work station using InsightII and Discover software (Accelrys, San Diego, CA). The figures were produced using the program Molscripts (43). Crystal structures used as the basis for modeling were obtained from the Brookhaven data base (44). N- and O-glycan structures were generated using the data base of glycosidic linkage conformations (45) and in vacuo energy minimization to relieve unfavorable steric interactions between the glycans and the protein surface.

RESULTS
Determinaton of the Relative Amounts of IgA1 and IgA2 in SIgA—The percentages of IgA1 (39%) and IgA2 (61%) in the total SIgA from pooled human colostrum were determined by ELISA (data not shown).

Separation and Identification of SIgA Component Proteins—Pooled normal human SIgA from colostrum was resolved into SC, H, L, and J chains by SDS-PAGE (Fig. 2). Tryptic peptides were sequenced to confirm the identity of the proteins in the gel bands (Table I). The J chain migrated with a higher apparent molecular mass than 16 kDa, in agreement with results by Chuang and Morrison (47).

N-Glycan Analysis—Fig. 2 shows the NP HPLC profiles of the pools of glycans released from the different gel bands. The glucose unit (GU) scale indicates the relative size of the glycans (the larger the GU, the larger the glycan). The N-glycans from the H chains ranged between GU 5 and 7; SC had the largest glycans (GU 6–12), whereas the J chain glycans were GU 6–9. Because there is only one glycosylation site per J chain and one J chain per SIgA, the abundance of glycans is lower; the light chain had negligible levels of N-glycosylation.

The preliminary assignment of structures was made by comparing GU values with standards and confirmed by following the elution positions (measured in GU) of peaks through the different exoglycosidase arrays and by mass spectrometry (37, 48). In Tables II–VI, the relative amounts of glycans in the pool released from each gel band can be found in the column headed “Undig.” To detect the presence of a bisecting GlcNAc residue, both Jack bean b-N-acetylgalactosaminidase and S. pneumonia b-N-acetylgalactosaminidase digestions were performed. Jack bean b-N-acetylgalactosaminidase does not digest glycans with bisecting GlcNAc, whereas S. pneumonia b-N-acetylgalactosaminidase does digest bisection biantenary glycans but is inefficient at digesting tri- or tetra-antennary structures with GlcNAc β1→4 or 1–6 linked to mannose under the conditions used in this paper (data not shown). The presence of bisecting GlcNAc was also confirmed by tandem mass spectrometry (MS/MS).

H Chain N-Glycans—Over 75% of the H chain N-glycans contain a bisecting GlcNAc, 66% contain a free terminal GlcNAc on one antenna, less than 20% of structures were fully galactosylated, less than 15% were sialylated (only n2–6 sialic acids detected), no glycans contained outer arm fucose residues, all galactose residues were β1–4-linked, about half the structures contained core fucose, and there were about 12% of oligomannose structures (Figs. 3 and 4 and Table II). The major structures were FcA2B (30%), A2B (21%), and FcA2BG1 (8%) (the notation is explained in footnote 1 of Table II). The glycans were of the biantenary complex type with a bisecting GlcNAc residue rather than triantenary, because they did not digest beyond A2B (GU 5.8) with Jack bean b-N-acetylgalactosaminidase but were digested with S. pneumonia b-N-acetylgalactosaminidase, and during LC-ESI-MS/MS, fragments were found that had lost the two GlcNAc-Man antennas but not the third bisecting GlcNAc (data not shown).

Fig. 5 shows the profiles of N-glycans from the H chains of SIgA1 and SIgA2, purified from colostrum from a single human donor. The N-glycan profiles from SIgA1 and SIgA2 H chains were almost identical, even though SIgA2 has two or three more N-glycan sites, depending on the isotype, than SIgA1.

J Chain N-Glycans—Over 75% of the structures at the single N-glycan site on the J chain were sialylated (Figs. 3 and 6 and...
Table III. Five major structures were detected in approximately equal proportions: biantennary complex structures FcA2G2S2, A2G2S2, FeA2G2S, A2G2S, and the hybrid structure FcMan4A1G1S. The di-sialylated biantennary structures contained one sialic acid α-2-3-linked and one α-2-6-linked (because all of the di-sialylated structures were digested to the mono-sialylated forms by Newcastle disease virus, which is specific for α-2-3 sialic acids). The mono-sialylated structures contained either α-2-3- or α-2-6-linked sialic acid. Bisecting GlcNAc was found in about half of the neutral structures but not in the sialylated structures. About half of all of the structures (neutral or charged) were core-fucosylated. No outer arm fucosylation was seen, and all galactose residues were 1-4-linked.

Secretory Component N-Glycans—A much greater range of structures was found on SC than on the H or J chain (Figs. 3, 7, and 8 and Table IV). The majority of structures were fully galactosylated, nonbisected biantennary structures. Triantennary (11.7%) structures were also present, and a small (≈1%) amount of tetraantennary (or polylactosamine) structures were detected by matrix-assisted laser desorption-ionization-MS. Over 70% of the glycans were sialylated, and the majority were mono-sialylated. The glycan pool was separated into neutral and charged forms. The composition of the glycans was determined by matrix-assisted laser desorption-ionization and liquid chromatography-electrospray ionization mass spectrometry. The results are shown in Table IV.
and mono- and di-sialylated fractions by weak anion exchange, and then these fractions were run again on NP HPLC (Fig. 8). The two di-sialylated peaks digested to A2G2 and FcA2G2 with A. ureafaciens sialidase, and, when subjected to Newcastle disease virus, about 20% lost one sialic acid and therefore originally contained one /\(^2\)H\(^9\)2–3 and one /\(^2\)H\(^9\)2–6 sialic acid; the remaining 80% had two /\(^2\)H\(^9\)2–6 sialic acids. The mono-sialylated fraction consisted of a mixture of glycans carrying either /\(^2\)H\(^9\)2–3 or /\(^2\)H\(^9\)2–6 sialic acid. Over 65% of structures contained core fucose. Digestion with glycosidases of differing specificity (bovine testes /\(^1\)H\(^9\)1–4 and S. pneumoniae /\(^1\)H\(^9\)1–3 galactosidase) gave different results, indicating that galactose was linked both /\(^1\)H\(^9\)1–4 and /\(^1\)H\(^9\)1–3 to GlcNAc. Digestion with almond meal /\(^2\)H\(^9\)1–3 and /\(^2\)H\(^9\)1–4 fucosidase showed the presence of outer arm fucoses linked both /\(^2\)H\(^9\)1–3 and /\(^2\)H\(^9\)1–4 to GlcNAc (because fucose was removed from GlcNAc, which also had galactose either /\(^1\)H\(^9\)1–3 and /\(^1\)H\(^9\)1–4 linked, the fucose was assumed to be in the unoccupied linkage position). After digestion with all of the exoglycosidases listed in Table IV, except for bovine kidney fucosidase, 1.8% of the glycan pool was left at GU 7.2. This material digested with bovine kidney fucosidase, indicating the presence of fucose /\(^2\)H\(^9\)1–2 linked to a galactose. LC-ESI-MS/MS analysis of the major mass peaks gave results consistent with the majority of structures with outer arm fucosylation and sialylation having fucose on one arm and sialic acid on the other. There was, however, some fucosylation and sialylation on the same arm (i.e. sialylated Lewis epitopes) because the fragment Hex-HexNAc-Fuc-Neu5Ac (m/z 803.2) was detected from fragmentation of FeA2G2FS (Hex\(_5\)-HexNAc\(_4\)-Fuc\(_2\)-Neu5Ac-2AB) (data not shown). Notably there were a number of unusual structures identified with outer arm fucosylation but no core fucose. Thus, there is a wide range of different epitopes present on the SC N-glycans, including all of the different Lewis and sialylated Lewis epitopes.

**O-Glycosylation**—Fig. 9 shows the NP HPLC profile from the undigested glycan pool followed by a range of whole pool digestions. Over 50 peaks, from glycans ranging in size from two- (GU 1.8) to fifteen-monosaccharide residues (GU 12), were found, and many of the peaks contained more than one structure. We used a combination of different HPLC and MS strategies to identify these structures as described in Ref. 37. The HPLC profiles were very complex; therefore individual peaks were collected, digested separately, and then run on NP and reverse phase HPLC. Table V gives the results of a selection of digestions on the whole pool of O-glycans, with 33 structures fully characterized in the whole O-glycan pool. The major structures (present at over 3% and labeled in the top panel of Fig. 9)
| 1Structure | 2Molecular Mass | 1Composition | NPHPLC GU | 1Exoglycosidase added | NPHPLC % Area |
|------------|----------------|-------------|----------|----------------------|----------------|
|            |                |             | Abs | Abs | Abs | Abs | Abs | Abs | Abs | Abs | Abs | Abs | Abs | Abs | Abs | Abs |
|            |                |             | Undig | Amf | Bkf | Btg | Btg | Spg | Spg | Spg | Spg | Btg | Btg | Spg | Spg | Spg |
| Man3       | -              | -           | 4.49  | -   | -   | -   | -   | -   | -   | -   | 219 | 215 | -   | -   | -   | -   |
| fucMan3    | -              | -           | 4.90  | -   | -   | -   | -   | -   | 3.9 | 64.7| 63.7| 1.7 | 6.3 | 2.0 | 7.4 | -   |
| A1         | -              | -           | 2.20  | -   | -   | -   | -   | -   | -   | -   | 2.5 | 3.1 | -   | -   | -   | -   |
| A2 from A3 | -              | -           | 5.40  | -   | -   | -   | -   | -   | 7.7 | -   | -   | 10.5| 9.8 | -   | -   | -   |
| A2G1       | 2(1421.5)      | 4, 3, 1   | 5.63  | -   | 2.4 | 4.7 | -   | -   | -   | -   | 3.2 | 2.9 | 3.0 | 2.6 | 2.4 | 1.0 |
| A2G1F      | 2(1421.5)      | 4, 3, 1   | 5.91  | 0.5 | 1.7 | 4.8 | -   | -   | -   | -   | 1.9 | 1.0 | 1.0 | -   | -   | -   |
| CeA2       | -              | -           | 4.90  | -   | -   | -   | -   | -   | -   | 9.9 | 10.6| 9.8 | 1.0 | 1.0 | 1.0 | -   |
| Man5       | 1244.4         | 6, 5, 4   | 6.20  | -   | 2.1 | 2.1 | 1.6 | 1.2 | 2.4 | 1.2 | 1.2 | 1.2 | 2.6 | 2.0 | 1.6 | -   |
| FeA3       | 1665.6         | 3, 5, 1   | 5.91  | -   | 1.7 | 4.8 | -   | 1.2 | 2.4 | 1.2 | 1.2 | 1.2 | 2.6 | 2.0 | 1.6 | -   |
| A2G2       | 4(1421.5)      | 4, 3, 1   | 4.98  | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   |
| A2G2F      | -              | -           | 4.52  | -   | 1.2 | 1.6 | 1.6 | 1.6 | 1.2 | 1.2 | 1.2 | 1.2 | 2.6 | 2.0 | 1.6 | -   |
| A2G2FS     | -              | -           | 4.93  | 0.4 | 2.8 | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   |
| A2G2F1     | -              | -           | 6.97  | 0.7 | 2.0 | -   | -   | -   | -   | 12.9| 2.2 | 1.2 | 1.2 | 12.9| 11.8| 1.4 |
| A2G2F2     | 4(1421.5)      | 4, 3, 1   | 7.17  | 1.8 | 2.0 | 2.0 | 2.0 | 2.0 | 2.0 | 2.0 | 2.0 | 2.0 | 2.0 | 2.0 | 2.0 | 2.0 |
| A2G2F3     | 4(1421.5)      | 4, 3, 1   | 7.37  | -   | -   | 12.9| 2.0 | 2.0 | 2.0 | 2.0 | 2.0 | 2.0 | 2.0 | 2.0 | 2.0 | 2.0 |
| A2G2F4     | -              | -           | 5.46  | 0.7 | 2.0 | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   |
| A2G2F5     | 4(1421.5)      | 4, 3, 1   | 7.37  | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   |
| A2G2F6     | 4(1421.5)      | 4, 3, 1   | 7.17  | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   |
| A2G2F7     | 4(1421.5)      | 4, 3, 1   | 7.37  | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   |
| A2G2F8     | 4(1421.5)      | 4, 3, 1   | 7.17  | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   |
| A2G2F9     | 4(1421.5)      | 4, 3, 1   | 7.37  | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   |
| A2G2F10    | 4(1421.5)      | 4, 3, 1   | 7.17  | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   |
| A2G2F11    | 4(1421.5)      | 4, 3, 1   | 7.37  | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   |
| A2G2F12    | 4(1421.5)      | 4, 3, 1   | 7.17  | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   |
| A2G2F13    | 4(1421.5)      | 4, 3, 1   | 7.37  | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   |
| A2G2F14    | 4(1421.5)      | 4, 3, 1   | 7.17  | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   |
| A2G2F15    | 4(1421.5)      | 4, 3, 1   | 7.37  | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   |
| A2G2F16    | 4(1421.5)      | 4, 3, 1   | 7.17  | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   | -   |

1 See Table II for abbreviations and footnotes.
2 Fucose linked α1–2 to Gal.
3 Composition Hex8HexNAc8 = 0,2,3, and 4 Fuc also detected as minor components by matrix-assisted laser desorption-ionization. These could be tetranennary or polylactosamine structures.
were: unsialylated and mono- and di-sialylated core I; Galβ1–4 core II; Galβ1–4 core II with an α1–3 sialic acid, with and without an α1–4 fucose; lactosamine (with galactose as either α1–3 or α1–4) extensions of Galβ1–4 core II with an α1–3 sialic acid, with and without an α1–4 fucose; and the largest structure that was fully sequenced was a furosylated-lac-
tosamine on Gal β1–4 core II with an α1–3 sialic acid, with an α1–4 fucose (Table VI, GU 7.51). In addition there were also at least 50 structures that were unidentified or not fully characterized, including some containing α1–2 fucose, as bovine kidney α-fucosidase digestions produced different results to almond meal α-fucosidase digestions.

Molecular Model of SIgA1—The molecular model (Fig. 10) was constructed from models of the IgA1 monomer, J chain, and secretory component. The IgA1 monomer model was based on the model of serum IgA1 previously published by us (6), with the orientation of the Fab domains adjusted to fit with the recent low angle x-ray and neutron scattering data obtained on serum IgA1 (49). It was also necessary to modify the conformations of the C-terminal tail pieces on building SIgA1 (see below). The glycan structures were also modified according to the glycan sequence data for SIgA presented here, with the most

| Structure | Molecular Mass | Composition | NPHPLC GU |
|-----------|----------------|-------------|------------|
| GlcNAc    | -              | -           | 6.24       |
| GlcNAc    | -              | -           | 6.39       |
| Galβ1–4   | 1478.5         | 4 4 0 0     | 6.50       |
| Galβ1–4   | 1478.5         | 4 4 0 0     |            |
| Galβ1–3   | 1550.6         | 3 3 1 1     | 6.59       |
| Galβ1–3   | 1550.6         | 3 3 1 1     |            |
| nd        | -              | -           | 6.67       |
| nd        | -              | -           | 6.73       |
| Galβ1–3   | 1769.6         | 4 4 0 1     | 7.11       |
| Galβ1–3   | 1769.6         | 4 4 0 1     |            |
| Galβ1–3   | 1696.6         | 3 3 2 1     | 7.51       |

**Table VI**

Analysis of the O-glycans from SIgA1

| Exoglycosidase added | NPHPLC % Area |
|----------------------|---------------|
| Undig               | Btg           |
| Abs                 | Btg           |
| Abs Amf             | Btf           |
| Abs Amf Bkf         | Btf Sph       |
| Abs Amf Bkf         | Sph           |
| Abs Amf Bkf         | Sph           |

| Structure | Symbol representation of glycan structure is: GlcNAc, black square; galactose, white diamond; fucose, diamond with a dot inside; sialic acid, black star; β linkage, solid line; α linkage; dotted line; 1–6 linkage, \; 1–4 linkage, /; 1–3 linkage, \; 1–2 linkage, ; nd, not determined. |
|-----------|---------------------------------------------------------------------------------------------------|
| 1         | See Table II for abbreviations.                                                                   |

Human Secretory IgA: N- and O-Glycosylation
abundant glycans distributed between the attachment sites. Although the site specificity of the glycans is not known, the overall conclusions made from the model would not be significantly affected, because the major glycans on each peptide are of similar size. An A2 glycan was attached to residue 263 (C\textsubscript{H} domain), and an FcA2B glycan was attached to residue 459 (tail piece) on each heavy chain; 4 of the 5 possible O-glycan sites on each H chain are shown occupied; on one H chain, sites 225, 228, 230, and 232 have O-glycans with GU values 1.80 (core 1), 3.58, 5.87 (the top structure), and 4.57, whereas the second H chain has sites 228, 230, 232, and 236 occupied by glycans of GU 2.93, 4.57, 7.51, and 3.58, respectively (see Tables V and VI for glycan structures). No Fab glycosylation was found. The J chain was modeled based on the three-dimensional fold of superoxide dismutase (1cbl.pdb) (50), as suggested by secondary structure predictions and circular dichroism spectroscopy (51). The sequence homology between the J chain and superoxide dismutase is not high, so the model is intended only to indicate the general size and shape of the J chain. An FcA2G2S2 glycan was attached to residue 49 (SWISSPROT numbering, P01591).

Fig. 3. Mass spectrometry data for SC, H, and J chain N-glycans. LC-ESI-MS combined data from N-glycan region of chromatogram. The structures are labeled for the main masses [M+H]\textsuperscript{+} (see Table II for abbreviations).

Fig. 4. NP HPLC profiles of N-glycans from the H chain following arrays of exoglycosidase digestions (see Table II for percentage areas). Aliquots of the total 2AB-labeled glycan pool were incubated with different exoglycosidases, as shown in each panel. Following digestion, the products were analyzed by NP HPLC. The major peaks have been annotated. The structures were allocated by their elution position measured in GU before and after digestion with exoglycosidases and reference to known GU values (37, 39).

Fig. 5. NP HPLC profiles of N-glycans from H chains of SIgA1 and SIgA2. Purified total SIgA from colostrum from a single human donor was separated into SIgA1 and SIgA2 by Jacalin-agarose affinity chromatography (SIgA1 binds). ELISA tests were done to check the purity of the samples, the IgA1, was 98% IgA1, and the IgA2 was 97% IgA2 (ELISA plates were coated with SIgA samples at a concentration of 5 g/ml (250 ng/well) in carbonate buffer pH 9.6; monoclonal anti-human IgA1- and IgA2-peroxidase conjugates (Nordic Immunological Labs, Tilburg, The Netherlands) were used at 1:1000 dilution; incubation with primary and secondary antibody was 1 h; 12 wells were run per sample). The samples were reduced, alkylated, and then run on a 10% BisTris gel. The N-glycans were released from the H chains by in-gel N-glycosidase F digestion, 2AB-labeled, and run on NP HPLC.
Fig. 6. NP HPLC profiles of N-glycans from J chain following digestions with arrays of exoglycosidases of known specificity. Aliquots of the total 2AB-labeled glycan pool were incubated with different exoglycosidases, as shown in each panel. Following digestion, the products were analyzed by NP HPLC. The major peaks have been annotated. The structures were allocated by their elution position in GU before and after digestion with exoglycosidases and reference to known GU values (37, 39). For example, the A2G2S peak (GU 7.90) moves to GU 7.20 (A2G2) after digestion with sialidase (A. ureafaciens sialidase (Abs)), whereas further digestion with galactosidase (bovine testes β-galactosidase (Btg)) moves the peak to GU 5.47 (A2), and finally digestion with S. pneumoniae β-N-acetylhexosaminidase moves the peak to GU 4.40 (Man3) (see Table III for percentage areas).

Fig. 7. NP HPLC profiles of N-glycans from SC following arrays of exoglycosidase digestions (see Table IV for percentage areas).

Ig-like domains connected by variable length peptides (52). Each individual domain was modeled separately, based on the crystal structure of the Ig-like domain with the highest sequence homology. The sequence homologies were generally in the region of 25–30% identity and 40% homology (using BLAST), so again the model provides only the overall size and shape of the domains. No attempt was made to model the linkers between the domains, except to put a limit on the possible distances between the domains. Glycans were added as follows: A2G2S2 glycans on residues 83, 186, and 499; an FcA2G2S2 glycan on residue 90; and FcA2G2S glycans on residues 135 and 469 (SWISSPROT numbering, P01833).

Dimeric SlgA1 was modeled from two IgA1 monomers. The relative positions of the IgA1 monomers were based on electron microscopy studies (53) showing that the dimer is linear with the four Fab domains in roughly the same plane and a separation between the two hinge regions of about 125 Å. This places the two Fe domains very close together, with insufficient room between them to accommodate the J chain. The J chain was, therefore, positioned asymmetrically to one side of the dimer with the glycans pointing away from the IgA1 monomers. The tail pieces were then modeled as random coil peptides, with the requirements that the C-terminal Cys on one tail of each monomer has to form a disulfide bond with Cys\textsuperscript{311} on the Ca2 domain of one of the IgA1 monomers (55), which fixes its position. In addition, residues on domain I of the SC have been shown to interact with the J chain (56) and one of the Ca3 domains (57, 58). This has been modeled as the secretory component wrapping around the IgA1/J chain complex, so that domain V is bonded to one IgA1 monomer and domain I interacts with the other monomer.

**MBL Binding Assays**—Although the SlgA H chains contain terminal sugars known to be ligands for MBL, no binding of MBL to SlgA-coated microtiter plates was detected (Fig. 11A). In contrast a sample of polymeric serum IgA, with known MBL affinity, did bind MBL (31). This suggests that the SC in the SlgA complex may be preventing access of MBL to these H chain N-glycans. To determine whether pH changes could disrupt the association of SC with SlgA-H chains and thereby expose the GlcNac and mannose bearing N-glycans on the H chains, the SlgA-coated plates were pretreated for 60 min with buffers of different pH. Preincubation at pH 2.0 or 3.0 resulted in MBL binding to SlgA but not to albumin (Fig. 11B). A time course of 5–60 min of preincubation with a pH 3.0 buffer showed increasing MBL binding to SlgA with time (Fig. 11C), indicating that the disruption of the noncovalent interactions between SC and IgA-H chains progressed over this period. Fig. 11D shows that this MBL binding to SlgA was a C-type lectin-sugar interaction because preincubation of MBL with either EDTA or D-mannose prevented MBL-SlgA binding.

**DISCUSSION**

This analysis of all of the major N- and O-glycans on each of the constituent peptides of SlgA indicates that there are significant differences in the processing of the glycan structures on the different peptides. The N-glycans on the SC and the O-glycans on the hinge region of the H chain of SlgA1 present a wide range of glycan epitopes, including Lewis type structures that are potentially able to bind to bacterial adhesins,
whereas the majority of the N-glycans on the H chains of both SIgA1 and SIgA2 carry terminal GlcNAc residues. We show that these GlcNAc residues are masked from binding to MBL in native SIgA by SC but that they can be unmasked, allowing interaction with MBL by disrupting the SC-H chain noncovalent interactions by incubation at low pH. We hypothesize that unmasking of the H chain N-glycans may take place in vivo upon the interaction of SC with bacterial adhesins or cellular lectins. Subsequent lectin binding to the exposed GlcNAc residues would promote opsonization by complement and/or direct phagocytosis, leading to presentation of pathogen to the adaptive immune system.

N-Glycans of SC Present Many Epitopes for Bacterial Adhesin and Lectin Binding—Our results show that SC presents a wide range of glycan structures, including all of the different Lewis and sialyl-Lewis epitopes that can potentially bind lectins and bacterial adhesins. We found galactose linked both β1–4 and β1–3 to GlcNAc; fucose linked α1–3 and α1–4 to GlcNAc and α1–2 to galactose, as well as both α2–3- and α2–6-linked sialic acids. This is a larger range of epitopes than reported by previous workers (21, 24–26). As well as having a general role of protecting both SC and the SIgA from proteases (59), these glycans on SC can specifically interact with adhesins and lectins. SC has been shown to bind to a range of bacteria via its glycans (e.g. H. pylori (11, 12), E. coli (10, 13, 14), C. difficile toxin A (15), and S. pneumoniae (16, 17)), thereby inhibiting attachment and the subsequent infection of epithelial surfaces. The SC glycans are also involved in the localization of SIgA by anchoring the SIgA to mucus lining the epithelial surface through its carbohydrate residues. Phalipon et al. (60) showed that glycosylated SC was required to locate the SIgA to specific areas of the epithelium and that antibody specificity was required for binding to Shigella flexneri. To protect the mice from S. flexneri infection, the whole glycosylated SC-IgA complex was required. Burns et al. (61) also found that SC was needed with IgA antibodies to protect mice from rotavirus, because the IgA was only effective if given systemically, not when presented at the luminal side of the intestinal tract (without SC), indicating that IgA transcytosis (with the addition of SC) was required for viral inactivation in vivo. A further role for SC glycans is in binding to the lectin-binding domain of Mac-1 (CR3, CD11b/CD18) (62), which is involved in inducing SIgA signaling via the IgA receptor FcR (CD89), leading to respiratory burst, phagocytosis, and cytokine secretion (reviewed in Refs. 63 and 64). Mac-1 has a broad specificity to sugars, including β-glucans (65), and has been shown to bind to recombinant SC, either as free SC or in the SIgA complex (62). The physiological roles of serum and SIgA are quite different, and these data show that the presence of the highly glycosylated SC has major effects on the biological functions of SIgA.

J Chain Conformation Requires an N-glycan—The J chain on dimeric IgA is essential for binding to pIgR and translocation across the epithelial cells (2, 56). Deletion of the N-glycan site, by substitution of Asn with alanine, prevents IgA dimer formation (54), indicating that the N-glycan plays an important role in this process. We found that over 75% of the structures on this single N-glycan site were sialylated with a mixture of both α2–3- and α2–6-linked sialic acids. Baenziger (28) also reported 85% sialylation (although all α2–6-linked sialic acids), and these data suggest that the charge on this glycan is important in maintaining the correct conformation of the J chain for presentation to pIgR.

The N-Glycans on SIgA H Chain Present Terminal GlcNAc Residues—Our analysis shows that the majority of N-glycans on the H chains of SIgA are bisected biantennary glycans.
without (~51%) or with one (~15%) galactose and that less than 15% of structures are sialylated. These truncated glycans are very different from the mainly fully galactosylated, sialylated structures that have been found on normal serum IgA (6, 66). Serum IgA is produced in plasma cells in the spleen or lymph and then distributed around the whole body. In contrast, SIgA is made in plasma cells local to the site of secretion. The difference in glycosylation between serum and secretory IgA is likely to be a reflection of the different glycosylation machinery operating in these different plasma cells. It has been shown that H chain N-glycans are needed to maintain proper conformation, because selective removal of these N-linked sites by mutagenesis results in degradation and reduced secretion of IgA (67) and markedly reduced dimer assembly (68). The SIgA H chain N-glycans on both SIgA1 and SIgA2 have 66% of glycans terminating in GlcNAc residues and 12% oligomannose structures. There are four N-glycans on SIgA1 and up to 20 on SIgA2; therefore these GlcNAc and mannose residues are multiply presented and may be expected to bind to lectins such as MBL.

**MBL Does Not Bind to Native SIgA**—When we tested the ability of MBL to bind to SIgA coated onto microtiter plates, no binding was detected. We therefore concluded that the SC is masking the N-glycans on the H chains. SC is covalently bound at one end (domain V) to the Co2 domain of one of the IgA1 H chains (55), and the other end (domain I) interacts noncovalently with both the J-chain (56) and one of the Co3 domains on a H chain (57, 58). The model in Fig. 10 shows these interactions with the SC wound around the dimeric IgA. The SC with its large N-glycans (shown in orange) thus obstructs the MBL from accessing the H chain N-glycans (shown in yellow). To find out whether disruption of the noncovalent interactions between SC and SIgA-H chains could unmask the underlying H chain N-glycans, the SIgA-coated plates were preincubated with low pH buffers. Indeed, preincubation at pH 3 allowed binding of MBL to SIgA, and longer preincubation times (0–60 min) directly correlated with increasing binding, suggesting that the H chain N-glycans were being unmasked. We therefore hypothesize that such unmasking may take place upon the interaction of SC glycans with bacterial adhesins or cellular lectins. The binding of lectins to the SC glycans (up to seven glycans) may be stronger than the noncovalent interactions between SC and the IgA H chains and therefore pull one end of the SC away from the H chains. Subsequent MBL binding to

![Molecular model of SIgA1 with N- and O-glycans](https://example.com/molecule.png)

**Fig. 10. Molecular model of SIgA1 with N- and O-glycans.** The two IgA monomers are shown in light blue and green, the J chain is in pink, and the secretory component is in purple. The O-glycans (on the hinge region of the H chains) are shown in red, the N-glycans on the SC and J chain are in orange, and the N-glycans on the H chains are in yellow. This model illustrates that most of the molecule is covered with glycans with the exception of the Fab antibody-binding sites. This leaves the Fab regions clear for interaction with antigens (adaptive immunity), whereas the O-glycans on the hinge region of the H chains and the N-glycans on SC cloak the rest of the surface, free for interaction with bacterial adhesins (innate immunity). In this conformation, with the SC wrapped around the H chains, the N-glycans on the H chain (which bear terminal GlcNAc and mannose residues) are masked from binding with lectins such as MBL.
the exposed GlcNAcs and mannose residues would promote opsonization by complement and/or direct phagocytosis, leading to presentation of pathogen to the adaptive immune system. In addition to the potential interaction with soluble lectins such as MBL, sugars present on SIgA may also bind to lectin receptors present on phagocytic cells. In this respect it has been shown that SIgA, but not serum IgA, could be internalized by dendritic cells via mannose receptor (69), which is an important receptor for antigen uptake, and that this binding and uptake could be blocked by specific sugars (mannose, fucose, GlcNAc, but not galactose) or partially by antibody reactive with mannose receptor. Recognition and presentation of antigen by SIgA may involve modulation of the immune response via its binding to host lectins. In this respect, it is important that uptake of SIgA by dendritic cells, which was shown to be partially mediated by the mannose receptor, did not induce DC maturation (69). No data are currently available about the potential effect of antigen opsonization by MBL on phagocytosis and maturation of dendritic cells. Because a recent study showed that MBL can be produced on the apical side of intestinal epithelium in the mouse (70), an interaction between SIgA and MBL in the gut lumen is quite feasible. In view of the well known roles of both SIgA and MBL in host defense, we hypothesize that such an interaction will contribute to the protection against mucosal infection.

A Wide Range of O-Glycan Epitopes Are Available for Interaction with Bacterial Adhesins—We found over 50 different O-glycan structures, up to 15 sugars in size, of which we have fully sequenced 33 structures (13 of which were previously identified by Pierce-Cretel et al. (22, 23)). As with the SC N-glycans, a wide range of structures were identified with many different sugar epitopes including galactose β1–4 and β1–3 linked; sialic acid α2–3 and α2–6 linked; and fucose α1–4, 3 and 2 linked, showing that these O-glycans (present on IgA1 but not IgA2) also present adhesin-binding sites. The T shape of IgA, revealed by x-ray resolution (49), indicates that these O-glycans would be exposed for interaction with bacterial adhesins. These SIgA O-glycans are much larger and more elaborate structures than the simple sialylated core 1 structures (6, 71) found on serum IgA and can be expected to have other functions in addition to stabilizing the hinge region. Evidence to support the proposal that these O-glycans can interact with bacterial adhesins in a similar way to those found on SC comes from Bos et al. (72). They established eight IgA-producing hybridomas from the mesenteric lymph nodes of mice and showed that these hybridomas recognize different but partially overlapping fecal bacterial populations. They suggested that this overlapping recognition was caused by interactions with conserved regions of the antibodies such as the glycans. The presence of the types of O-glycans we have shown could readily account for these overlapping interactions.

Differences in O-glycan structures between serum and secretory IgA are likely to be due to different glycosylation enzymes in the plasma cells located in the bone marrow/spleen/liver to those at mucosal surfaces. It is also plausible that plasma cells at different mucosal surfaces (e.g. lungs and upper or lower
...}  

**Acknowledgments**—We thank Brian Matthews for assistance with hyalurondiase procedures, Alex von Kriegsheim for advice and assistance with peptide sequencing, and Dr. T. Fujikata (Fukushima, Japan) for providing monoclonal antibody 3E7.

**REFERENCES**

1. Kerr, M. A. (1990) Biochem. J. 271, 285–296
2. Johansen, F. E., Braathen, R., and Brandtzaeg, P. (2001) J. Immunol. 167, 5165–5169
3. Norderhaug, I. N., Johansen, F. E., Schjerven, H., and Brandtzaeg, P. (1999) Crit. Rev. Immunol. 19, 481–508
4. Baenziger, J., and Kornfeld, S. (1974) J. Biol. Chem. 249, 60–68
5. Iwase, H., Tanaka, A., Hiki, Y., Kokubo, T., Ishii-Karakasa, I., Kobayashi, Y., et al. (1996) Mol. Immunol. 32, 55–63
6. Mattu, T. S., Pleass, R. J., Willis, A. C., Kilian, M., Wormald, M. R., Lellouch, J. M., and Hotta, K. (1996) J. Immunol. 157, 61–69
7. Novak, J., Tomana, M., Kilian, M., Coward, L., Kulhavy, R., Barnes, S., and Mestecky, J. (2000) Mol. Immunol. 37, 1047–1056
8. Lamm, M. E. (1997) Annu. Rev. Microbiol. 51, 311–340
9. Sharon, N., and Ofek, I. (2000) Glycocon. J. 17, 659–664
10. Schotten, H., Stupper, C., Plagmann, R., Kohler, H., Hacker, J., and Hanisch, F. G. (1998) Infect. Immun. 66, 3791–3797
11. Fahl, P., Roth, K. A., Boren, T., Westblom, T. U., Gordon, J. I., and Normark, S. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 2035–2039
12. Boren, T., Falk, P., Roth, K. A., Larson, G., and Normark, S. (1993) Science 262, 1892–1895
13. Weld, A. E., Mestecky, J., Tomana, M., Kobata, A., Ohashiyashi, H., Endo, T., and Mestecky, J. (1990) Infect. Immun. 58, 3073–3077
14. de Oliveira, I. R., de Araujo, A. N., Bao, S. N., and Giugliano, L. G. (2001) FEMS Microbiol. Lett. 203, 29–33
15. Dallas, S. D., and Rolfe, R. D. (1998) J. Mol. Microbiol. 47, 879–888
16. Hammerschmidt, S., Talay, S. R., Brandtzaeg, P., and Chahatwal, G. S. (1995) Mol. Microbiol. 25, 1113–1124
17. Zhang, J. R., Mostov, K. E., Lamm, M. E., Nanno, N., Shimida, S., Ohwaki, M., and Tum posing, E. (2000) Cell 102, 827–837
18. Cortesby, B. (2002) Trends. Biotechnol. 20, 65–71
19. Cortesby, B., and Svertin, F. (1999) Biochem. Biophys. Res. Commun. 260, 1251–1256
20. Wright, A. L., and Slonim, S. (1997) Trends. Biotechnol. 15, 32–36
21. Pierce-Cretel, A., Debray, H., Montreuil, J., Spik, G., Van Halbeek, H., Mutsaers, J. H., and Vliegenthart, J. F. (1984) Eur. J. Biochem. 139, 337–349
22. Pierce-Cretel, A., Decottignies, J. P., Wiersuzski, J. M., Strecker, G., Montreuil, J., and Spik, G. (1989) Eur. J. Biochem. 182, 457–476
23. Pierce-Cretel, A., Pambianco, M., Strecker, G., Montreuil, J., and Spik, G. (1981) Eur. J. Biochem. 114, 169–178
24. Pierce-Cretel, A., Pambianco, M., Strecker, G., Montreuil, J., Spik, G., Dorland, L., Van Halbeek, H., and Vliegenthart, J. F. (1982) Eur. J. Biochem.
Secretory IgA N- and O-Glycans Provide a Link between the Innate and Adaptive Immune Systems

Louise Royle, Anja Roos, David J. Harvey, Mark R. Wormald, Daniëlle Van Gijlswijk-Janssen, El-Rashdy M. Redwan, Ian A. Wilson, Mohamed R. Daha, Raymond A. Dwek and Pauline M. Rudd

J. Biol. Chem. 2003, 278:20140-20153.
doi: 10.1074/jbc.M301436200 originally published online March 10, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M301436200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 72 references, 26 of which can be accessed free at http://www.jbc.org/content/278/22/20140.full.html#ref-list-1