Asparagine-linked Oligosaccharides Present on a Non-consensus Amino Acid Sequence in the C_1 Domain of Human Antibodies

Received for publication, April 29, 2009, and in revised form, September 16, 2009. Published, JBC Papers in Press, September 18, 2009, DOI 10.1074/jbc.M109.014803

John F. Valliere-Douglass, Paul Kodama, Mirna Mujacic, Lowell J. Brady, Wes Wang, Alison Wallace, Boxu Yan, Pranhitha Reddy, Michael J. Treuheit, and Alain Ballard

From the Department of Process and Product Development, Amgen Inc., Seattle, Washington 98119-3105

We report that N-linked oligosaccharide structures can be present on an asparagine residue not adhering to the consensus site motif NX(S/T), where X is not proline, described in the literature. We have observed oligosaccharides on a non-consensus asparaginyl residue in the C_1 constant domain of IgG1 and IgG2 antibodies. The initial findings were obtained from characterization of charge variant populations evident in a recombinant human antibody of the IgG2 subclass. HPLC-MS results indicated that cation-exchange chromatography acidic variant populations were enriched in antibody with a second glycosylation site, in addition to the well documented canonical glycosylation site located in the C_1 domain. Subsequent tryptic and chymotryptic peptide map data indicated that the second glycosylation site was associated with the amino acid sequence TVSWN162SGAL in the C_1 domain of the antibody. This highly atypical modification is present at levels of 0.5–2.0% on most of the recombinant antibodies that have been tested and has also been observed in IgG1 antibodies derived from human donors. Site-directed mutagenesis of the C_1 domain sequence in a recombinant-human IgG1 antibody resulted in an increase in non-consensus glycosylation to 3.15%, a greater than 4-fold increase over the level observed in the wild type, by changing the −1 and +1 amino acids relative to the asparagine residue at position 162. We believe that further understanding of the phenomenon of non-consensus glycosylation can be used to gain fundamental insights into the fidelity of the cellular glycosylation machinery.

The attachment of glycan to a particular asparaginyl residue is governed by protein sequence (5, 6). It has been determined, from studies involving amino acid substitutions of the +2 residue following the asparagine residue, that a proximal hydrogen-bond acceptor is necessary for rendering the asparagine residue sufficiently nucleophilic to displace the GlcNAc_2-Man_2Glc_3 from the dolichol donor (7, 8). The canonical N-linked consensus sequence motif, NX(S/T), where X is not Pro, has one sequence exception, and that is replacement of the Ser or Thr residue in the +2 position with a Cys residue. This exception, although not favorable for N-glycosylation, was first demonstrated to be possible on synthetic substrates (7) before being found in nature. The motif NX(S/T) was shown to be constitutively glycosylated in bovine and later in human protein C (9, 10) and human von Willebrand factor (11).

Previous binding and kinetic studies have shown that the K_m for glycosylation of Asn residues by the OST complex is severalfold higher, and the V_max was reduced 2- to 3-fold when the +2 amino acid was a Ser versus a Thr (7, 12). It was subsequently determined, through the use of synthetic peptides containing Thr analogues in the +2 position, that introduction of charge or movement of the side-chain methyl group was not tolerated by the OST complex (13, 14). The outcomes of these studies suggest the presence of a hydrophobic binding pocket in the active site of the OST complex, which results in preferential glycosylation of NXT relative to NXS due to the presence of the side-chain methyl group of Thr (15).

Amino acids around the N-glycosylation consensus sequence also have an impact on the probability that a consensus asparagine residue will be glycosylated. Petrescu et al. (16) performed a statistical analysis of the upstream and downstream amino acids adjacent to occupied asparagines in N-glycosylated proteins, which are present in the Protein Data Bank crystallographic data base. They found that the occurrence of specific residues at discrete positions with respect to the consensus Asn residue was correlated with site occupancy. Glycosylated Asn residues have a greater probability of being flanked by large hydrophobic amino acids, particularly aromatic residues immediately preceding the occupied Asn. There was also a greater probability of Pro residues occurring in the vicinity of the occupied Asn with the exception of the complete absence of Pro in the +1 position (6) and reduced probability of Pro in the +3 position.

Although sequence has been shown to be a necessary requirement for N-glycosylation (5), the fact that, among the

---

1 To whom correspondence should be addressed: Amgen Inc., 1201 Amgen Ct, West Seattle, WA 98119-3105. Tel: 206-265-8603; E-mail: ballanda@amgen.com.

2 The abbreviations used are: OST, oligosaccharyltransferase; SNA, S. nigra agglutinin; PBS, phosphate-buffered saline; HPLC, high-performance liquid chromatography; PNGaseF, peptide N-glycosidase F; EIC, extracted ion current; CID, collision-induced dissociation; MS, mass spectrometry; LC, liquid chromatography; aa, amino acid(s); HC, heavy chain; CE, capillary electrophoresis; rCE, reduced CE.
Non-consensus N-Glycosylation

theoretical consensus sites found in the protein data base, only a minority are found actually glycosylated shows it is not sufficient. Investigations on conformational requirements have shown that glycosyl acceptor motifs adopt an Asn-turn conformation (17, 18). Structural studies using model tripeptides indicated that the active site of the GST enzyme complex orients the side chain of the X amino acid in the NX(S/T) sequon away from the side chains of the Asn and Ser or Thr amino acids thus allowing the side-chain hydroxyl of the +2 amino acid to be in close proximity to the Asn amide (19). The importance of the Asn turn motif for efficient N-glycosylation was mechanistically described by Imperiali et al. (20, 21) and confirmed using a synthetic peptide analogue, which was shown to adopt an Asn-turn motif (22). These results indicate that Pro in the X position is not tolerated by the GST enzyme complex, because the cyclic structure of Pro causes rigidity in the peptide backbone that prevents it from being oriented away from the Asn and the Ser or Thr amino acids in the +2 position. Retrospective analysis of glycoprotein surface topology indicates that there is a bias toward glycosylated asparagines being present on flat or convex surfaces or on the edge of a groove or a cleft on a flat surface (16). These studies also indicated that there was a marked preference for finding aromatic side chains within 4 Å of any glycogen atom thus supporting the concept that glycosylation may stabilize hydrophobic protein surfaces (23, 24).

We performed an in-depth characterization of a recombinant antibody and found that a low level population was glycosylated in the C\(_{\text{H1}}\) constant domain. While antibodies are always glycosylated in the C\(_{\text{H1}}\)2 domain of Fc, glycosylation can also be observed when complementarity determining regions in the variable domains contain a consensus N-glycosylation sequence motif (25, 26). Our initial results prompted further investigation, as potential glycosylation of natural CH1 constant domain has never been reported, which is in agreement with the fact that no consensus N-linked site is present on this conserved sequence. We describe in this report the evidence indicating that an asparagin residue, that is not part of an N-glycosylation consensus sequence, is, in fact, glycosylated. Low level non-consensus N-glycosylation in the C\(_{\text{H1}}\)1 domain was also observed on IgG1 antibodies isolated from human serum. We also describe initial findings regarding the sequence requirements for this modification. Our analysis raises the question of the fidelity of the glycosyl transferase reaction and provides the basis for a new understanding of a classic cotranslational modification.

EXPERIMENTAL PROCEDURES

Recombinant Antibodies—The IgG1 and IgG2 antibodies used in this study were human recombinant molecules stably expressed in Chinese hamster ovary cells or transiently expressed in human kidney (HEK293) cells. The antibodies were purified using conventional techniques (27) and stored in an acetate buffer at pH 5.

Antibodies from Human Serum—Lectin affinity chromatography was performed on the IgG component of pooled-normal human serum (Sigma). Enrichment of antibodies with terminal sialylation was accomplished by lectin capture using immobilized Sambucus nigra agglutinin (SNA-Sepharose, EY Laboratories, San Mateo, CA) according to standard procedures (28). Briefly, human intravenous immunoglobulin was dialyzed with PBS and applied to a column pre-equilibrated with PBS. Bound glycoconjugates containing the α2,6-linked sialic acid were washed with PBS and eluted with PBS containing 100 mM lactose.

Cation-exchange Chromatography—Cation-exchange isolation was carried out on an SCX-10 column (Dionex, Sunnyvale, CA) connected to an Agilent HP1200 quaternary HPLC or an HP1100 binary HPLC (Agilent, Santa Clara, CA). Buffer A consisted of 20 mM Tris-HCl, pH 8.0, and buffer B was 20 mM Tris-HCl, 250 mM NaCl, pH 8.0. After a 3-min hold at 0 mM NaCl for 3 min, the NaCl concentration was brought to 30 mM in 2 min at a flow rate of 0.8 ml/min. Following the initial ramp to 30 mM NaCl, the gradient continued at the rate of 1 mM per minute to a final NaCl concentration of 60 mM. A protein load of 2 mg per injection was used to obtain sufficient material for enrichment of low level species and subsequent characterization. Pooled fractions were concentrated, and buffer was exchanged into pH 5 acetate buffer using Centricon/Cen triprep spin filters (Millipore, Billerica, MA) with either a 30- or 50-kDa molecular weight cut-off.

Reversed-phase Separation and Mass Measurement of Reduced Heavy and Light Chains—Separation of antibody heavy and light chains was carried out according to a method based on Ren et al. (29). Recombinant antibody was reduced by heating samples to 55 °C for 30 min in the presence of 25 mM Tris-(2-carboxyethyl)phosphine, 4 mM guanidine HCl, 50 mM sodium acetate, pH 5.0. 100 μg of reduced antibody was injected onto a (2.0 × 150 mm) diphenyl XRS reversed-phase column (Varian, Palo Alto, CA) with 3-μm particles and 200-Å pore size. The column temperature was maintained at 80 °C during the separation, and the bound heavy and light chains were eluted with a gradient from 35 to 41% acetonitrile in 0.1% trifluoroacetic acid at a flow rate of 0.3 ml/min. The mass of reduced heavy and light chain fragments was determined on an Agilent LC-MSD-TOF mass spectrometer. Optimal desolvation of antibody fragments was obtained by adjusting the electrospray ionization source drying gas and nebulizer gas flows to 12 and 35 liters/h, respectively. The capillary voltage was set to 4500 V, and the fragmentor and octapole RF voltages were maintained at 350 and 300 V, respectively.

Capillary Scale Peptide Map Analysis of Trypsin- and Chymotrypsin-digested Antibody—Recombinant antibody was denatured and reduced at 55 °C for 30 min in the presence of 4 mM guanidine HCl, 57 mM Tris-HCl buffer, pH 8.3, and 10 mM dithiothreitol. Reduced cysteines were alkylated with 22 mM sodium iodoacetate for 15 min at room temperature in the dark. Prior to digestion, samples were buffer-exchanged into 50 mM Tris-HCl, pH 7.5, using Biospin 6 columns (Bio-Rad, Hercules, CA) and incubated with 1500 units of PNGase F (New England Biolabs, Ipswich, MA) or 15 milliunits of sialidase A (Glyko, Novato, CA) per 100 μg of protein for 2 h at 37 °C. Following deglycosylation, urea was added to a final concentration of 2 M, and all samples were incubated 1:10 (w/w) with trypsin or chymotrypsin (Roche Applied Science, Basel, Switzerland) for 4 h at 37 °C. Digestion was stopped with the addition of trifluoroacetic acid to a final concentration of 1%. Up to 5 μg of digested
antibody was injected onto two C18 Atlantis columns (Waters, Milford, MA) connected in series (0.3 × 150 mm) to a Waters nanoacquity HPLC. The particle diameter was 3 μm, and the pore size was 300 Å. Column temperature was maintained at 50 °C during the separation and, following a 10-min hold at the initial conditions of 0.1% formic acid, peptides were eluted with a gradient from 0.1% formic acid to 36% acetonitrile and 0.1% formic acid over 155 min at a flow rate of 6 μl/min. The masses of the eluted peptides were determined using a Waters/Micromass Q-TOF Premier mass spectrometer. Data were acquired using a two-segment method previously described (30–32) in which the first scan was acquired using low collision energy voltage (2 V) and the second scan was acquired under higher collision energy conditions (20 V). Mass data on intact glycopeptides were obtained from the first scan event, and glycan-specific product ion scans were detected from the second scan event. Extracted ion current (EIC) chromatograms were generated using Waters Masslynx software. The mass range was sufficient so that all isotopes of a given charge state were summed to generate the EIC. Separate methods were created for CID analysis of glycopeptides. Particular peptides were analyzed using single ion-monitoring MS acquired for 10 s per scan. The collision energy used for fragmentation was dependant on the mass and charge state of the peptide. Multiply charged peptides were deconvoluted to zero charge monoisoionic state using the MaxEnt 1 deconvolution algorithm.

$H_2^{18}O$ Labeling of Glycosylated Asparaginyl Residues—Protein samples were deglycosylated in the presence of $^{18}O$ water as previously described by Gonzalez et al. (33) with some minor changes. Stock solutions of 100 mm ammonium bicarbonate were made in $^{16}O$ and $^{18}O$ water. Reduced and alkylated samples were added to 0.5-ml Centricon spin filters with a 10-kDa molecular weight cutoff, and the volume of the samples was reduced to 25 μl by centrifugation at 9000 rpm. Each sample was brought to a final volume of 500 μl with the addition of 100 mm ammonium bicarbonate in $^{18}O$ or $^{16}O$ water, and this cycle was repeated until the total final dilution of samples into $H_2^{18}O$ or $H_2^{16}O$ ammonium bicarbonate was 400-fold. Samples were then incubated in the presence or absence of PNGase F as described above. Treated samples were then buffer-exchanged into 2.0 mM urea, 50 mM Tris-HCl, pH 7.5, using Centricon spin filters as described above. Removal of residual $^{18}O$ water before enzymatic digestion is crucial, otherwise $^{18}O$ may be incorporated into the C terminus of peptides as the peptide bond is hydrolyzed during endoprotease digestion (34). Subsequent digestion with trypsin or chymotrypsin was carried out as described above.

Analytical Scale Peptide Map Analysis of Chymotrypsin-digested Antibody—Prior to reduction and alkylation, 250 μg of recombinant antibody was digested with 0.6 milliunit of endoglycosidase F2 enzyme (Glyko) at 37 °C for 16 h according to manufacturer’s instructions. Antibody samples were reduced, alkylated, and digested with chymotrypsin according to procedures described above. Digested peptides were separated using a Varian Polaris Ether C18 column (2.1 × 250 mm). Peptides were injected onto the column using a Waters Acquity HPLC flowing at 0.2 ml/min. The initial buffer composition prior to the start of the gradient was 0.5% acetonitrile in 0.1% formic acid. After a 10-min hold at the initial conditions, peptides were eluted using a linear gradient to 27% acetonitrile, 0.1% formic acid over 90 min. The column was subsequently washed in 80% acetonitrile, 0.1% formic acid for 5 min and brought back to the initial conditions and re-equilibrated for 40 min prior to the next injection. Peptide masses were determined using a ThermoFinnigan LTQ mass spectrometer (Thermo Scientific, Waltham, MA) set to perform MS2 and MS3 in a data-dependent manner. The theoretical mass of the C141 chymotryptic peptide containing a fucosylated N-acetylglucosamine structure ([M+H]$^+$ = 810.3 Da) was added to a mass inclusion list to ensure that this species would undergo CID-MS/MS analysis. The unmodified NSGAL peptide and the glycosylated NSGAL peptide were observed eluting at 13 and 20 min, respectively, and were collected for further CID-MS/MS analysis by static nanospray infusion using a Triversa Nanomate (Advion Biosciences, Ithaca, NY). CID-MS4 was carried out on the glycosylated and non-glycosylated chymotryptic peptide NSGAL at a collision energy of 25–30 V. The masses selected for fragmentation are described under “Results.”

| Mutagenesis primer pair | Mutant | Forward (NSGAL) | Reverse (NSGAL) |
|-------------------------|--------|----------------|----------------|
| Wild-type C141 WNGS | Forward: 5′-GTCGTAATACGACAAGCTGACGAC-3′ | Reverse: 5′-GGTCAGGGCGCCGGCGTTCCACGACAC-3′ |

**Table 1.** List of C141 domain antibody heavy-chain mutants

Double-stranded DNA primer pairs associated with the mutations at residues 161–164 are shown in the right column, and the wild-type 5′ forward DNA sequence is shown at the bottom of the table. The primer sequence coding for the four amino acids shown in the table is underlined.
which they were combined with 54 ml of HEK293 cells. The cultures were then placed into a humidified incubator set at 37 °C and 5% CO2 with 150 rpm agitation. On day 3, cells were fed with 25 ml/liter FreeStyle 293 medium containing 20% Tryptone N1 (Organotechnie S.A., La Courneuve, France) and 0.1% Pluronic F68 (Invitrogen). Two days later, cultures were spun down at 1000 rpm for 10 min, and obtained supernatants were filter-sterilized using a 0.2-μm filter.

Quantitation of CH1 Glycosylation by Reduced Capillary Electrophoresis—Prior to capillary electrophoresis (CE) separation, 75–100 μg of recombinant antibody was reduced to free heavy and light chains with dithiothreitol and alkylated with iodoacetamide in the presence of 2% SDS. Samples were subsequently buffer-exchanged to 50 mM sodium phosphate at pH 7.5 using Biospin 6 columns (Bio–Rad) and incubated with 1500 units of PNGase F (New England Biolabs) at 37 °C for 16 h, at which time an additional 1500 units of PNGase F was added to the sample, and the reaction was continued for 8 h. Antibody fragments were labeled with the fluorophore 3-(2-furoyl)quinoline-2-carboxaldehyde and analyzed according to Michels et al. (36) by reduced capillary electrophoresis (rCE). The heavy-chain species were quantitated by laser-induced fluorescence detection on a ProteomeLab PA-800 CE (Beckman Coulter, Fullerton, CA).

RESULTS

Non-consensus N-Linked Oligosaccharides in C4μ Domain of Monoclonal Antibodies—Isolation of the acidic variants from a human monoclonal IgG2 antibody was accomplished by pooling fractions from repeat injections onto an analytical cation-exchange column. The purified acidic variants were reduced with Tris-(2-carboxyethyl)phosphine and analyzed by reversed-phase LC-MS (Fig. 1A). Masses were observed eluting in the leading edge of the main heavy-chain peak that were consistent with heavy-chain populations with two oligosaccharide structures present (Fig. 1B). Based on the populations found in the center of the reversed-phase heavy-chain peak (Fig. 1C), it was inferred that the composition of the secondary glycan structures present on the heavy chain were fucosylated biantennary complex structures with two terminal galactose residues (NA2F, Table 2) and the same structure with a terminal N-acetylated neuraminic acid (A1F).

Elution of molecules with an extra, neutral monosaccharide in the acidic region of the cation-exchange profile has been observed previously (37), and the similar behavior observed with molecules containing an extra N-linked oligosaccharide may be due to non-charge-based, mixed-mode separation of species. The acidic variant tryptic peptide map data were investigated for the presence of glycosylated peptides by product-ion monitoring. The collision energy was raised on every other scan to selectively fragment glycan structures on glycosylated peptides, and the glycans were then detected as positively charged oxonium product ions. The most abundant glycosyl product ions observed from N-linked oligosaccharides were protonated N-acetylglucosamine-galactose or -mannose disaccharides (HexNAc-Hex, [M + H]+ = 366.139 Da). EIC chromatograms were plotted around the mass corresponding to protonated
HexNAC-Hex product ions (Fig. 2), and peaks with retention times of 125.5 and 127.5 min were observed in the acidic variant sample that was not treated with glycosidases (Fig. 2A, peaks 1 and 2). Sialidase A digestion resulted in a loss of the second product ion peak at a retention time of 127.5 min (Fig. 2B) and an increase in the intensity of peak 1 at 125.5 min. These results indicated that peak 2 corresponded to a glycopeptide that contained a terminal sialic acid residue, which was cleaved by the action of sialidase A. The absence of both peaks in the PNGase F-treated sample (Fig. 2C) confirmed that the glycosyl product ions were derived from N-linked oligosaccharides. MS data from the low CID energy scans was investigated for the presence of glycopeptides. EIC of the [M+5H]^5+ ion at m/z = 1708.3 Da, which is consistent with the expected mass for HC aa sequence 151–213 with a NA2F oligosaccharide structure, resulted in a peak with the same retention time as HexNAC-Hex product ion peak 1 in the untreated and sialidase A-treated samples (Fig. 2D and E, respectively). EIC of the [M+5H]^5+ ion at m/z of 1766.5 Da, which is consistent with the expected mass for HC aa sequence 151–213 with a NA2F oligosaccharide structure, resulted in a peak with the same retention time as HexNAC-Hex product ion peak 1 in the untreated and sialidase A-treated samples (Fig. 2F, G and H, respectively). Neither [M+2H]^2+ ion was observed in the PNGase F-treated sample (data not shown).

The [M+2H]^2+ ion at m/z = 1115.9 Da in the sialidase A-treated sample (Fig. 4A) and the 1+ ion for the non-glycosylated peptide NSGAL were selected for CID MS2 fragmentation (Fig. 4, B and C, respectively). A comparison of both CID spectrums indicated that the [M+2H]^2+ ion at 1115.9 Da had the same amino acid sequence as NSGAL based on a partial b-ion series and that abundant HexNAC (m/z = 204.084 Da) and HexNAC-Hex (m/z = 366.139 Da) product ions were also present. This was clear evidence that the peptide NSGAL, corresponding to IgG2 heavy chain amino acids 162–166, Kabat numbering (38), was glycosylated. Confirmation that the asparagine residue at position 162 was glycosylated was obtained by sequencing of the glycopeptide.

**Site Determination of N-Glycosylation Site by 18O Labeling**—Reduced and alkylated antibody acidic variants were treated with PNGase F in the presence of H218O following digestion with chymotrypsin and peptide map analysis. Using this methodology, a 3-Da mass shift confirms N-linked glycosylation as the acidic variant that results from the action of PNGase F on a glycosylated asparagine is isotopically labeled with an18O atom (33). An [M+H]^+ ion at m/z = 464.225 Da was observed eluting after the unmodified NSGAL peptide (m/z = 461.226). A comparison of the CID MS2 fragments from these ions indicated that the Asn at position 162 was glycosylated based on the b-ion series, which was shifted 3 Da heavier, and the a1 ion, which was observed at 90.052 Da in the 464.225 Da peptide from the PNGase F-treated sample (Fig. 5A) versus the a1 observed for NSGAL at 87.059 Da from the untreated sample (Fig. 5B).

**Direct Determination of N-Glycosylation Site by CID-MS2**—Analysis of Endoglycosidase F2-Digested Monoclonal Anti-
Non-consensus N-Glycosylation

FIGURE 2. EIC chromatograms from tryptic peptide maps of acidic variants. HexNAc-Hex product ions at m/z = 366.139 Da are extracted from the untreated, sialidase A-treated and PNGase F-treated samples (A, B, and C, respectively). EIC of the [M+5H]5+ ion at m/z = 1708.3 Da, which is consistent with the expected mass of HC aa sequence 151–213 modified with a NA2F oligosaccharide structure, is shown for the untreated sample (D) and the sialidase A-treated sample (E). EIC of the [M+5H]5+ ion at m/z = 1766.5 Da, which is consistent with the expected mass of HC aa sequence 151–213 modified with a A1F oligosaccharide structure, is shown for the untreated sample (F) and the sialidase A-treated sample (G).

—The glycosylated and the non-glycosylated chymotryptic peptide NSGAL was collected from the endoglycosidase F2-treated acidic variants peptide map. Endoglycosidase F2 cleaves specifically after the 1st GlcNAc residues in the core structure of N-glycosylated oligosaccharides resulting in a fucosylated N-acetylglucosamine (HexNAc-Fuc) disaccharide at the amino acid site of attachment. Static nanospray CID-MS² analysis of the non-glycosylated peptide, NSGAL, at [M+H]+ = 461.3 Da revealed a simple b-ion series that terminated at a b2 ion at m/z = 202.00 Da (Fig. 6B). Fragmentation of the NSGAL plus (HexNAc-Fuc) resulting from Endo-F2 treatment at [M+H]+ = 810.3 Da resulted in the loss of the fucose from all fragment ions and a partial b-ion series with and without the attached HexNAc monosaccharide (data not shown). The major product ions observed from CID-MS² analysis of NSGAL plus HexNAc at m/z = 664.3 Da were b-type ions with the attached HexNAc. However, upon comparison with a zoomed view of the NSGAL MS² spectrum, the MS² spectrum of NSGAL plus HexNAc was found to contain a lower level ion series of masses consistent with b-type ions with the HexNAc monosaccharide maintained during the CID-MS² analysis (Fig. 6A). The identities of the unique ions in Fig. 6A were easily determined by adding the mass of HexNAc (203 Da) to a- and b-type ions and confirmed by subsequent MS³ analysis. The ions present in Fig. 6A, which were absent in Fig. 6B, corresponded to a b2 ion with HexNAc still attached at m/z = 405.08 and b3 ions at m/z = 390.17 Da and the water loss product at 372.17 Da (Fig. 6A, b3 and b3'-H2O, respectively) in which the HexNAc residue underwent a cross-ring cleavage. The assignment of these masses was verified by MS⁴ analysis of the 390.17-Da ion, which gave the 372.17-Da ion as the major product, and fragmentation of the 372.17-Da ion, which yielded b3 and b2 product ions (data not shown). The additional mass of 131 Da present on the b3 ion at m/z = 390.17 Da was consistent with a 3,5X₃ type cross-ring cleavage product occurring on a HexNAc monosaccharide (39). The ion at m/z = 348.17 Da (Fig. 6A) was determined to be a b4 ion based on the product ions observed as a result of MS³ analysis (data not shown). The additional mass, 18 Da above the theoretical value for b4, is likely due to a Y₀ type cleavage of a HexNAc residue during CID analysis (39). An ion at m/z = 290.08 Da, which was present in the modified MS³ spectrum (Fig. 6A), was consistent with the mass of a HexNAc monosaccharide attached to an a1 Asn ion. The HexNAc product ion at m/z = 204.00 Da (Fig. 6A) was also observed in the CID-MS³ spectrum of the putative NSGAL plus HexNAc ion. The MS⁴ spectrum that resulted from fragmentation of the ion at m/z = 290.08 Da (Fig. 6C) contained HexNAc-specific product ions at m/z = 204.00, 186.00, and 168.00 Da as well as an a1 Asn ion at 87.00 Da, which had lost the HexNAc monosaccharide during CID analysis. The occurrence of the putative Asn a-type ion at 87.00 Da and HexNAc product ions was strong evidence that the HexNAc monosaccharide was attached to the Asn residue in the sequence of NSGAL. Verification of the formation of an a1 ion

32498 JOURNAL OF BIOLOGICAL CHEMISTRY

VOLUME 284 • NUMBER 47 • NOVEMBER 20, 2009
Non-consensus N-Glycosylation

Non-consensus N-Linked Oligosaccharides in C1r1 Domain of IgG1 Antibodies in Human Serum—In the course of our work on the glycosylation of the C1r1 domain of monoclonal antibodies, we observed dramatically greater levels of sialylation on C1r1 glycans compared with C1r2 glycans. This finding indicated that antibodies with C1r1 glycosylation could potentially be enriched from human serum using SNA lectin, which is specific for terminal sialic acid. Analysis of the tryptic peptide maps from SNA lectin-enriched human serum antibody pools indicated that the IgG1 tryptic peptide, in C1r1 domain corresponding to HC aa sequence 151–213, was glycosylated at a low level relative to the unmodified peptide (Fig. 7A). It should be noted that the sequence of the IgG1 and IgG2 peptides from C1r1 are almost completely conserved and differ by only 4 amino acids out of 63 total residues (Table 3). The C1r1 glycans present on IgG1 from human serum contained ions of approximately equal intensity that were consistent with biantennary (A1F, Table 2) and bisected-biantennary (A1FB, Table 2) glycans based on the \([M+H]^+\) ions at m/z = 1756.846 Da and 1797.469 Da. It should be noted that bisected oligosaccharides are not typically observed in recombinant antibodies expressed in Chinese hamster ovary cells but are common in antibodies derived from human sources (40, 41). Due to the mode of enrichment, the predominant species observed were all sialylated. Confidence in the assignment of the ions in Fig. 7A was further increased by sialidase A treatment. Following sialidase A treatment, the mass of the \([M+H]^+\) ions was reduced by 58.3 Da, which was consistent with the loss of mass that would be expected from the removal of sialic acid (uncharged mass = 291.1 Da, 291.1 + 5 = 296.2) from a sialylated 5 + ion (Fig. 7B). In addition to the above NA2F and NA2FB species, a lower level \([M+5H]^+\) ion at m/z = 1665.863 Da, which is consistent with the expected mass of IgG1 HC aa sequence 151–213 plus GNA2F, was also observed. Ions corresponding to the masses of glycosylated peptides from IgG1 HC aa sequence 151–213 were absent in the sample treated with PNGase F prior to analysis (Fig. 7C). SNA lectin-enriched human serum IgG was treated with sialidase A and subsequently digested with chymotrypsin to verify that the glycans observed on the C1r1 tryptic peptide

[FIGURE 3. EIC chromatograms from chymotryptic peptide maps of acidic variants. HexNAc-Hex product ions at m/z = 366.139 Da are extracted from the untreated, sialidase A-treated and PNGase F-treated samples (A, B, and C, respectively). EIC of the \([M+2H]^+\) ion at m/z = 1115.9 Da, which is consistent with the expected mass of HC aa sequence 162–166 modified with a NA2F oligosaccharide structure, is shown for the untreated sample (D) and the sialidase A-treated sample (E). EIC of the \([M+2H]^+\) ion at m/z = 1261.5 Da, which is consistent with the expected mass of HC aa sequence 162–166 modified with a A1F oligosaccharide structure, is shown for the untreated sample (F) and the sialidase A-treated sample (G).]
were associated with the chymotryptic fragment NSGAL. The \([M+2H]^2+\) ion consistent with the expected average mass of NSGAL plus NA2F \((m/z = 1216.164)\) was selected for fragmentation, and the sequence NSGAL was confirmed in the b-ion series in the resulting CID-MS\(^2\) (data not shown). Also present in the CID-MS\(^2\) were abundant glycan-specific product ions at \(m/z\) of 204.096 and 366.143 Da. These masses, corresponding to protonated HexNAc and HexNAc-Hex mono- and disaccharides, respectively, confirmed the presence of low levels of CH1 glycosylation in IgG1 antibodies in human serum.

**Panel A**
Sialidase A treated
NSGAL + NA2F
\([M+2H]^2+ = 1115.443\)

**Panel B**
Sialidase A treated
CID-MS/MS of 1115.417

**Panel C**
CID-MS/MS of NSGAL

---

**Mutational Analysis of Non-consensus Glycosylation in Human IgG1 Molecule**—The C\(_{H1}\) domain amino acid sequence in a recombinant human monoclonal antibody of the IgG1 subclass was altered by single point mutations of amino acids in the -1, +1, and +2 positions relative to the Asn residue at aa position 162 (Table 1). We found that the extent of glycosylation occurring on an antibody can be conveniently ascertained by reduced capillary electrophoresis. The level of C\(_{H1}\) glycosylation was assessed by quantitation of heavy chain populations separated by this technique (Fig. 8). We determined, through exhaustive treatment with PNGase F, that antibody heavy chain with two N-linked oligosaccharides migrated after the main population with a single N-linked oligosaccharide (data not shown). Heavy chains modified by 0, 1, or 2 N-linked structures were eluted in this order by rCE-SDS (peaks 1, 2, and 3, Fig. 8). The extent of C\(_{H1}\) glycosylation could thus be reliably quantitated by comparing the fluorescence area of the post heavy-chain peak to total heavy-chain peak area. The mutant constructs, which contained a consensus sequence in the C\(_{H1}\) domain, were found to be constitutively glycosylated at Asn-162, and thus consisted of virtually 100% HC populations with two N-glycans (Table 4). The native C\(_{H1}\) heavy chain sequence resulted in a level of non-consensus glycosylation of 0.62%. Mutation of the Ser at position 163 to Thr reduced the level of non-consensus glycosylation to 0.55%, and mutation of Gly at position 164 to Phe further lowered CH1 glycosylation to 0.29%. Substitution of Ser at position 163 with Ala resulted in 1.18% non-consensus glycosylation, and substitution of the Trp at position 161 with Ala resulted in 3.15% non-consensus glycosylation (Table 4). Treatment of antibodies with PNGase F...
prior to analysis resulted in a shift in the retention time of the main heavy-chain peak (Fig. 8, peak 2) into the non-glycosylated heavy-chain peak (Fig. 8, peak 1) due to the removal of the Fc glycans (data not shown). The post heavy-chain peak (Fig. 8, peak 3) also collapsed into the non-glycosylated heavy-chain peak as a consequence of the removal of the CH1 and CH2 glycans. The percentage of released glycans was subsequently quantitated by comparing the fluorescent peak area of rCE-SDS HC peak 1 to total HC peak area after PNGase F treatment (data not shown). The percentage of N-glycans removed as a consequence of PNGase F treatment is shown in Table 4.

DISCUSSION

The results from analytical characterization of a human monoclonal antibody as well as the antibodies obtained from human patient serum clearly indicated that the asparagine residue in CH1 domain, position 162, was naturally glycosylated at low levels. With the exception of the sequence NXC, non-consensus N-glycosylation has not previously been described in any protein much less an antibody. However, N-glycosylation consensus sequences have previously been specifically engineered into the CH1 domain of monoclonal antibodies for the purpose of creating antibody-drug conjugates (42). The glycan structures that we have observed on the non-consensus site at Asn-162 are similar to those that have been reported as a consequence of mutating the Gly at position 164 to Ser, to create a consensus N-glycosylation site and to antibodies with N-glycosylation in Fab V-region domains (43–45). The CH1 domain amino acid sequence mutations that we have analyzed were chosen based on previous work elucidating the biochemical mechanism involved in the attachment of glycans to asparagine residues (7) as well as statistical analysis of the flanking residues found near glycosylated asparagine residues (16). The influence of the +3 amino acid on Asn occupancy in the consensus sequence X_{−1}NX_{−1}(S/T)_{+1}X_{+3} has been examined (46), and the role of the +2 S/T residue has been exhaustively documented. We hypothesized that the factors governing non-consensus Asn occupancy would be the presence of a hydroxyl amino acid in the +1 position and degrees of freedom available to the Asn side chain. Thus, we chose to confine our work primarily to the influence of the −1 and +1 residues on glycosylation of the Asn residue at position 162. Although a Thr residue in the +2 position in a consensus sequence is correlated with a greater likelihood of glycan occupancy (15, 16), our
results suggested that the steric hindrance associated with the side-chain methyl group on threonine at position 163 (+1 position) adversely affected non-consensus glycosylation of the asparagine residue at position 162. Based on the work of Bause et al. (7), it was expected that removal of the hydroxy amino acid at position 163 would eliminate non-consensus glycosylation at asparagine 162. However, we found that substitution of Ser with Ala actually doubled N-glycan occupancy versus the wild-type sequence. When the Ser residue at position 163 is eliminated, the nearest
C-terminal potential hydrogen acceptor is a threonine residue at position 167. It has previously been reported that N-glycosylation could still occur if the side-chain hydroxyl group on the Ser/Thr amino acid in the +2 position was replaced with an amide (13). Although Trp residues have a nitrogen atom in the indole ring, the Trp residue at position 161 is an unlikely candidate as a hydrogen acceptor, because the electron pair of the indole nitrogen is delocalized and participates in the aromatic system. Furthermore, its sequential position, N-terminal to the glycosylated asparagine residue, suggests that it does not participate mechanistically in glycosylation. Given the preference for bulky, hydrophobic amino acids in the −1 position (16), it was expected that the substitution of the Trp residue for Ala would result in a decrease in non-consensus glycosylation relative to the wild-type sequence. Occupancy at Asn-162 increased to 3.15% as a consequence of replacing the −1 Trp with Ala suggesting that glycan occupancy was correlated with a reduction in the relative bulk of the −1 amino acid. In general, observed glycosylation on heavy-chain CH1 Asn at position 162 increased as a consequence of reduction in side-chain bulk of amino acids in the −1 and +1 positions. The observation that the Asn residue in the sequence WNAG was glycosylated at >1% suggests that there is no requirement for a hydrogen acceptor in this non-consensus sequence. Taken together, all these facts draw a very different picture from what is mechanistically known to occur during glycan transfer on a consensus site. It is tempting to speculate that a specific enzymatic mechanism is necessary to modify these non-consensus sites.
Ruiz-Canada et al. (4) have very recently demonstrated that distinct catalytic subunits, STT3A and STT3B, within the OST enzyme complex, act sequentially and complementarily to optimize protein glycosylation. The STT3A subunit is implicated in the cotranslational glycosylation of acceptor sequences, whereas the STT3B subunit was shown to cotranslationally modify acceptor sequences adjacent to N-terminal signal sequences of secreted proteins, and post-translationally modify

![Image of a table and a graph]

**TABLE 3**

| Antibody subclass | Amino acid sequence | Glycan structure | Mass, monoisotopic/avg. Da |
|-------------------|---------------------|-----------------|--------------------------|
| IgG1,2            | NSGAL               | None            | 460.228/460.488          |
| IgG1,2            | NSGAL               | NA2F            | 2228.870/2230.125        |
| IgG1,2            | NSGAL               | A1F             | 2519.965/2521.383        |
| lgG2              | C_{1h} 151–213      | None            | 6763.234/6767.415        |
| lgG2              | C_{1h} 151–213      | NA2F            | 8531.870/8537.055        |
| lgG2              | C_{1h} 151–213      | A1F             | 8822.965/8828.313        |
| lgG1              | C_{1h} 151–213      | None            | 6713.291/6717.442        |
| lgG1              | C_{1h} 151–213      | GNA2F           | 8319.877/8324.932        |
| lgG1              | C_{1h} 151–213      | NA2F            | 8481.930/8487.075        |
| lgG1              | C_{1h} 151–213      | NA2FB           | 8685.009/8690.270        |
| lgG1              | C_{1h} 151–213      | A1F             | 8773.025/8778.333        |
| lgG1              | C_{1h} 151–213      | A1FB            | 8976.104/8981.527        |
| lgG2/IgG1 C_{1h} 1 amino acids 151–213 | DYPPEPVTVSWN162SGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSH | NA2F | 8481.930/8487.075 |
| lgG1              | DYPPEPVTVSWN162SGALTSGVHTFPAVLQSSGLYSLSSVVTVPSSH | A1F | 8773.025/8778.333 |

**FIGURE 8. Analysis of IgG1 antibody C_{1h} 1 domain mutants by rCE-SDS.** Antibody samples were derivatized with the fluorescent label 3-(2-furoyl)quinoline-2-carboxaldehyde and separated by rCE-SDS. Heavy chains were quantitated by laser-induced fluorescence. Heavy-chain populations occupied with 0, 1, and 2 N-glycans are denoted as peaks 1, 2, and 3, respectively.
TABLE 4

rCE-SDS quantitation of Cγ1 mutant antibody fragments

The percentages are calculated from the CE electropherogram peak areas shown in Fig. 7. Peaks 1, 2, and 3 represent HC populations with 0, 1, and 2 N-glycans, respectively. The percentage for 2 N-glycans is calculated as the peak 3 area (Fig. 7) relative to total HC peak area. The percentage of N-glycans removed (0 N-glycans) as a consequence of PNGase F treatment is calculated as the peak 1 area (Fig. 7) relative to total HC peak area.

| Cγ1 sequence | Untreated 2 N-glycans | PNGase F-treated 0 N-glycans |
|--------------|-----------------------|-------------------------------|
| WNSS         | 98.30                 | 99.74                         |
| WNGT         | 97.81                 | 99.47                         |
| WNGS         | 97.84                 | 98.05                         |
| ANSG         | 3.15                  | 99.34                         |
| WNA6         | 1.18                  | 99.63                         |
| WNSG         | 0.62                  | 99.75                         |
| WNTG         | 0.55                  | 99.91                         |
| WNSF         | 0.29                  | 100.00                        |

accep9tor sequences missed by STT3A or adjacent to the C ter-
minus in unfolded proteins. It has been shown that mutations
that affect folding of nascent glycoproteins may result in attach-
ment of N-glycans at acceptor sites that are not normally occu-
pied (47). This suggests a mechanism whereby relatively slower
folding rates for particular antibody domains may result in low
level glycosylation at a non-consensus acceptor site. It is not
known whether the higher activity and turnover exhibited by
STT3B might come at the expense of acceptor sequence fidelity
and thus also contribute to the phenomenon of non-consensus
glycosylation. The level of non-consensus glycosylation
observed with the IgG1 Cγ1 ANSG mutant (3.15%) suggests
that this construct may be an ideal model for assessing the
mechanism behind non-consensus glycosylation: specifically,
the identity of the catalytic subunit responsible for glycan addi-
tion to non-consensus acceptor sequences.

Since its description by Marshall et al. (5), the amino acid
sequence necessary for N-linked glycosylation, NX(S/T), where X
is not proline, has been universally established. The presence
of this modification on human antibodies from natural sources
indicates that nonspecific N-linked glycosylation is not artifac-
tially induced by recombinant expression systems. The use of
enrichment procedures and highly sensitive analytical tech-
niques allowed us to investigate molecules at a deeper level
resulting in a challenge of the N-glycosylation dogma and its
absolute requirements. Further investigations into non-con-
sensus amino acid sequences that can be N-glycosylated and
the types of protein that bear this unusual modification are
needed to understand the significance of these findings.

Although it seems likely that we have observed limitations on
the fidelity of the OST enzyme complex glycosylation machin-
ery to amino acid sequences, the larger biological significance
of this observation requires further investigations.

Acknowledgments—We are grateful to Yihong Han for assistance with
lectin affinity chromatography, Sean Steen for mass spectrometric
support, and Mei Han and Martin Wolfson for input on protein A
purification. We thank Greg Flynn and Gary Rogers for critical review
of the manuscript and helpful comments, and Dean Pettit for his
support. We appreciate the insightful suggestions from the two anon-
ymous reviewers.

REFERENCES

1. Rothman, J. E., and Lodish, H. F. (1977) Nature 269, 775–780
2. Rapoport, T. A., Jungnickel, B., and Kutay, U. (1996) Annu. Rev. Biochem.
65, 271–303
3. Bolt, G., Kristensen, C., and Steenstrup, T. D. (2005) Glycobiology 15, 541–547
4. Ruiz-Canada, C., Kelleher, D. J., and Gilmore, R. (2009) Cell 136, 272–283
5. Marshall, R. D. (1974) Biochem. Soc. Symp. 40, 17–26
6. Bause, E., and Hettick, H. (1979) FEBS Lett. 108, 341–344
7. Bause, E., and Legler, G. (1981) Biochem. J. 195, 639–644
8. Silberstein, S., and Gilmore, R. (1996) FASEB J. 10, 849–858
9. Grinnell, B. W., Walls, J. D., and Gerlitz, B. (1991) J. Biol. Chem. 266, 9778–9785
10. Stenflo, J., and Fernlund, P. (1982) J. Biol. Chem. 257, 12180–12190
11. Titani, K., Kumar, S., Takio, K., Ericsson, L. H., Wade, R. D., Ashida, K.,
Walsh, K. A., Chopek, M. W., Sadler, J. E., and Fujikawa, K. (1986) Bio-
chemistry 25, 3171–3184
12. Breuer, W., and Bause, E. (1995) Eur. J. Biochem. 228, 689–696
13. Bause, E., Breuer, W., and Peters, S. (1995) Biochem. J. 312, 979–985
14. Breuer, W., Klein, R. A., Hardt, B., Bartoschek, A., and Bause, E. (2001) FEBS Lett.
501, 106–110
15. Kasturi, L., Eshelman, J. R., Wunner, W. H., and Shakin-Eshleman, S. H.
(1995) J. Biol. Chem. 270, 14756–14761
16. Petrescu, A. J., Milac, A. L., Petrescu, S. M., Dwek, R. A., and Wormald,
M. R. (2004) Glycobiology 14, 103–114
17. Bause, E., Hettick, H., and Legler, G. (1982) Biochem. J. 203, 761–768
18. Bause, E. (1983) Biochem. J. 209, 331–336
19. Imperiali, B., and Shannon, K. L. (1991) Biochemistry 30, 4374–4380
20. Imperiali, B., Shannon, K. L., and Rickert, K. W. (1992) J. Am. Chem. Soc.
114, 7942–7944
21. Imperiali, B., Shannon, K. L., Unno, M., and Rickert, K. W. (1992) J. Am.
Chem. Soc. 114, 7944–7945
22. Imperiali, B., Spencer, J. R., and Struthers, M. D. (1994) J. Am. Chem. Soc.
116, 8424–8425
23. Toyoda, T., Itai, T., Arakawa, T., Aoki, K. H., and Yamaguchi, H. (2000)
J. Biochem. 128, 731–737
24. Toyoda, T., Arakawa, T., and Yamaguchi, H. (2002) J. Biochem. 131, 511–515
25. Wright, A., and Morrison, S. L. (1993) Springer Semin. Immunopathol. 15, 259–273
26. Wright, A., Tao, M. H., Kabat, E. A., and Morrison, S. L. (1991) EMBO J.
10, 2717–2723
27. Shukla, A. A., Hubbard, B., Tressell, T., Guhan, S., and Low, D. (2007)
J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci. 848, 28–39
28. Shibuya, N., Goldstein, I. J., Broekaert, W. F., Nsima-Lubaki, M., Peeters,
B., and Peumans, W. J. (1987) J. Biol. Chem. 262, 1596–1601
29. Ren, D., Pipes, G. D., Hambly, D. M., Bondarenko, P. V., Treueheit, M. J.,
Brems, D. N., and Gadgil, H. S. (2007) J. Chromatogr. A 1175, 63–68
30. Chalkley, R. J., and Burlingame, A. L. (2001) J. Am. Soc. Mass Spectrom. 12,
1106–1113
31. Huddleston, M. J., Bean, M. F., and Carr, S. A. (1993) Anal. Chem. 65,
877–884
32. Ivancek, M. M., Gadgil, H. S., Hambly, D. M., Pipes, G. D., Halsall, H. B.,
and Treueheit, M. J. (2007) Poster presented 55th ASMS Conference on Mass Spectrometry and Allied Topics, Indianapolis, IN
33. Gonzalez, J., Takao, T., Hori, H., Besada, V., Rodriguez, R., Padron, G., and
Shimoniishi, Y. (1992) Anal. Biochem. 205, 151–158
34. Angel, P. M., Lim, J. M., Wells, L. B., Bergmann, C., and Orlando, R. (2007)
Rapid Commun. Mass Spectrom. 21, 674–682
35. Durocher, Y., Perret, S., and Kamen, A. (2002) Nucleic Acids Res. 30, E9
36. Michels, D. A., Brady, L. J., Guo, A., and Balland, A. (2007) Anal. Chem. 79,
5969–5973
37. Valliere-Douglass, J. F., Brady, L. J., Farnsworth, C., Pace, D., Balland, A.,
Wallace, A., Wang, W., Treueheit, M. J., and Yan, B. (2009) Glycobiology 19,
144–152
38. Kabat, E. A., Wu, T. T., Perry, H. M., Gottesman, K. S., and Foeller, C.
(1991) Sequences of Proteins of Immunological Interest, 5th Ed, National
Institutes of Health, U. S. Department of Health and Human Services, Bethesda, MD

39. Domon, B., and Costello, C. E. (1988) Glycoconj. J. 5, 397–409
40. Jefferis, R. (2009) Nat. Rev. Drug Discov. 8, 226–234
41. Sheeley, D. M., Merrill, B. M., and Taylor, L. C. (1997) Anal. Biochem. 247, 102–110
42. Qu, Z., Sharkey, R. M., Hansen, H. J., Shih, L. B., Govindan, S. V., Shen, J., Goldenberg, D. M., and Leung, S. O. (1998) J. Immunol. Methods 213, 131–144
43. Holland, M., Yagi, H., Takahashi, N., Kato, K., Savage, C. O., Goodall, D. M., and Jefferis, R. (2006) Biochim. Biophys. Acta 1760, 669–677
44. Mimura, Y., Ashton, P. R., Takahashi, N., Harvey, D. J., and Jefferis, R. (2007) J. Immunol. Methods 326, 116–126
45. Qu, Z., Sharkey, R. M., Hansen, H. J., Goldenberg, D. M., and Leung, S. (1997) Glycobiology 7, 803–809
46. Mellquist, J. L., Kasturi, L., Spitalnik, S. L., and Shakin-Eshleman, S. H. (1998) Biochemistry 37, 6833–6837
47. McGinnes, L. W., and Morrison, T. G. (1994) Virology 200, 470–483