Partial Glycosylation of Antithrombin III Asparagine-135 Is Caused by the Serine in the Third Position of Its N-Glycosylation Consensus Sequence and Is Responsible for Production of the β-Antithrombin III Isoform with Enhanced Heparin Affinity†

Véronique Picard, Eva Ersdal-Badju, and Susan C. Bock*

Microbiology & Immunology Department and The Sol Sherry Thrombosis Research Center, Temple University Medical School, Philadelphia, Pennsylvania 19140

Received January 11, 1995; Revised Manuscript Received April 11, 1995®

ABSTRACT: Two antithrombin III (ATIII) isoforms occur naturally in human plasma. The α-ATIII isoform has four N-linked oligosaccharides attached to asparagines 96, 135, 155, and 192. The β-ATIII isoform lacks carbohydrate on asparagine-135 (N135), which is near the heparin binding site, and binds heparin with higher affinity than does α-ATIII. Two isoforms are also produced when the normal human ATIII cDNA sequence is expressed in baculovirus-infected insect cells, and the recombinant β′ isoform similarly binds heparin with higher affinity than the recombinant α′ isoform. Consensus sequences (CSs) of the ATIII N-glycosylation sites are N-X-S for 135 and N-X-T for 96, 155, and 192. On the basis of database and *in vitro* glycosylation studies suggesting that N-X-S CSs are utilized less efficiently than N-X-T CSs, we hypothesized that the β-ATIII isoform might result from inefficient core glycosylation of the N135 N-X-S CS due to the presence of a serine, rather than a threonine, in the third position. ATIIIs with N-X-S, N-X-T, and N-X-A consensus sequences were expressed in baculovirus-infected insect cells. In contrast to the N-X-S sequence, which expressed a mixture of α′ and β′ molecules, the N-X-T variant produced α′ exclusively, while the N-X-A variant produced β′ exclusively. Thus, serine in the third position of the N135 CS is responsible for its "partial" glycosylation and leads to production of β-ATIII. Further evidence for CS third position hydroxyamino acid modulation of N-glycosylation was obtained when two of the three other ATIII N-X-T CSs were changed to N-X-S. This work also supports a substrate-assisted catalysis mechanism for oligosaccharyltransferase, the evolutionarily conserved enzyme responsible for core oligosaccharide addition at N-glycosylation sites.

Antithrombin III (ATIII,[1] antithrombin) is a plasma glycoprotein with physiological anticoagulant activity. It belongs to the serpin family (Huber & Carrell, 1989) and inhibits thrombin and other activated coagulation factors through the formation of stable ATIII−proteinase complexes in which the enzyme is inactive. However, in contrast to other serpins, ATIII proteinase inhibitor and anticoagulant activities are dependent on activation by a specific, highly sulfated pentasaccharide sequence. This pentasaccharide is present in vascular wall heparan sulfate proteoglycans (HSPGs), which are the natural activators of ATIII, and in heparin, a substance that is not naturally present in the blood but may be administered pharmaceutically as an anticoagulant. The so-called "heparin cofactor" activity of ATIII is approximately 1000-fold its "progressive" activity, measured in the absence of heparin. [See Bjork et al. (1989) for review.]

In addition to increasing the rates of antithrombin proteinase inhibitor reactions, HPSG binding to ATIII serves to localize it in activated form on the vascular endothelium, where its activated coagulation factor targets are generated (Marcum & Rosenberg, 1984). Based on binding, clearance, and perfusion studies, there are two pools of ATIII-binding HSPGs associated with the vessel wall (de Agostini et al., 1990; Carlson et al., 1985a,b; Witmer & Hatton, 1991; Felsch & Owen, 1994). A small portion of the ATIII-binding HSPGs are present on the lumenal side of the endothelium; however, most are located ablumenally beneath it. Thus, current models of natural anticoagulant function propose that the small pool of ATIII, bound in activated form to HSPG on the lumenal side of the endothelium, neutralizes enzymes generated by continuous low-level activation of the coagulation mechanism, while the larger, ablumenal subendothelial pool of HSPG and ATIII provides additional anticoagulant activity when endothelial damage causes exposure of these reserves to blood (de Agostini et al., 1990). In contrast, circulating ATIII molecules which are not bound to HSPGs remain unactivated and are thought to play a minor role in the natural anticoagulant mechanism (Bauer et al., 1985). The importance of HSPG binding in the expression of physiological ATIII anticoagulant activity is supported by the occurrence of early thrombotic disease in patients homozygous for antithrombins with normal progressive

† This work was supported by NIH RO1-HL30712. V.P. was the recipient of a Southeast Pennsylvania American Heart Association postdoctoral fellowship, and S.C.B. was the recipient of American Heart Association Established Investigatorship 88-0261.

* Address correspondence to this author at the Microbiology Department, Temple University Medical School, 3400 N. Broad St., Philadelphia, PA 19140.

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

[1] Abbreviations: ATIII, antithrombin III; CS, consensus sequence; ER, endoplasmic reticulum; HSPG, heparan sulfate proteoglycan; pi, postinfection; PNGase F, peptide:*N*-glycosidase F; tPA, tissue plasminogen activator.

0006-2960/95/0434-8433$09.00/0

activity but defective heparin binding and heparin cofactor properties (Finazzi et al., 1987).

ATIII is synthesized by hepatocytes and secreted into the blood. The 58-kDa glycoprotein has four Asn-X-Ser/Thr N-glycosylation consensus sequences (CSs) and circulates as a mixture of two glycoforms. The major α-ATIII glycoform is glycosylated on all four consensus sequences, in contrast to the minor β-ATIII isoform, which is not glycosylated on an asparagine located near the heparin binding site. As a consequence of reduced steric interference around asparagine-135 (N135), β-ATIII binds heparin (and presumably also HSPGs) with higher affinity than α-ATIII (Brennan et al., 1987). Consistent with its higher affinity for heparin/HSPGs, β-ATIII is cleared from the circulation faster than α-ATIII (Carlson et al., 1985a), associates more rapidly with uninjured and de-endothelialized aorta than does α-ATIII, and is the major isoform observed in immunoblots of detergent extracts of initima-media from uninjured rabbit aortas (Witmer & Hatton, 1991). Thus, it has been suggested that the β-ATIII isoform, which constitutes only a minor portion of the circulating ATIII pool (~10%), may actually predominate in mediating physiologically relevant vascular wall anticoagulant reactions (Witmer & Hatton, 1993).

The origin of the β-ATIII glycoform is unknown. It was proposed that β-ATIII is derived from α-ATIII by specific removal of labile asparagine-135-linked oligosaccharide in the circulatory system (Brennan et al., 1987); however, no direct evidence for such an activity has been published. An alternative explanation for the origin of β-ATIII is that oligosaccharide is not transferred to asparagine-135 during the synthetic process, leading to "partial" glycosylation of this site and secretion of two glycoforms.

The first step of the N-glycosylation pathway is transfer of a preformed high-mannose core oligosaccharide from a dolichol-linked precursor to selected asparagine residues. This cotranslational reaction occurs as nascent polypeptides enter the endoplasmic reticulum (ER) and is catalyzed by oligosaccharyltransferase, an evolutionarily conserved membrane protein complex which faces the lumenal side of the ER (Kaplan et al., 1987). Only asparagines in the context of the tripeptide consensus sequence (CS) Asn-X-Ser/Thr/Cys (N-X-S/T/C), where X may be any amino acid except proline, are glycosylated. However, oligosaccharyltransferase does not add core oligosaccharides to all CSs of polypeptides transiting the secretory pathway (Kaplan et al., 1987; Gavel & von Heijne, 1990). Rather, comparison of amino acid sequence and carbohydrate modification data indicates that some CSs are always glycosylated, whle others are never glycosylated, and still others are only sometimes glycosylated ("partially" glycosylated).

Factors determining whether a given CS is glycosylated are not completely understood. With the exception of disallowed proline (Mononen & Karjalainen, 1984), the identity of the second position (X) amino acid does not control glycosylation outcome. Nor is there evidence for the influence of local sequence context (33-residue-long fragments, centered around the CS) or predicted secondary structure on glycosylation (Gavel & von Heijne, 1990). Instead, the kinetics of protein translation (Kaplan et al., 1987; Miletich & Broze, 1990) and protein folding (Miletich & Broze, 1990; Bullied et al., 1992), and interactions of the CS with various ER components (Gilmore, 1993; Nicchitta & Blobel, 1993) have been suggested to influence glycosylation outcome by affecting oligosaccharyltransferase access to a CS.

A correlation between the identity of the hydroxyamino acid in the CS third position and the occurrence of glycosylation has also been noted. Protein database analyses show that N-X-T sequences are glycosylated 2−3 times more frequently than are N-X-S sequences, despite the sequences occurring at the same frequency (Kaplan et al., 1987; Gavel & von Heijne, 1990). Moreover, in a cell-free glycosylation system utilizing calf liver microsomes and CS-containing synthetic peptide substrates, N-X-T-containing peptides were glycosylated 40-fold more efficiently than N-X-S-containing peptides (Bause & Legler, 1981).

The presence of oligosaccharide side chains on all four CS asparagines of α-ATIII but only the N96, N155, and N192 CSs of β-ATIII (Brennan et al., 1987) suggests that the two isoforms by be generated by partial glycosylation of N135. ATIII α and β isoforms have been observed in rabbits (Carlson & Atencio, 1982) and humans (Peterson & Blackburn, 1985), and in both species the 96, 155, and 192 sites have N-X-T CSs, while the 135 consensus seqeunce is N-X-S. Therefore, to investigate whether the serine in the third position of the ATIII asparagine-135 N-glycosylation consensus sequence might be responsible for partial glycosylation of N135 and the production of two glycoforms from a single polypeptide sequence, we changed it to a threonine and analyzed glycoform expression in baculovirus-infected insect (Sf9) cells. The influence of CS third-position hydroxyamino acid identity on glycoform production was also investigated for the N96, N155, and N192 consensus sites. We found that the presence of a serine rather than a threonine at the 137 position is responsible for the partial glycosylation of asparagine 135 and the production of β-ATIII, and that two of three other ATIII CSs are less efficiently glycosylated when the third-position residue is serine rather than threonine.

MATERIALS AND METHODS

Site-Directed Mutagenesis and Plasmid Construction. Mutagenesis was according to a PCR-based megaprimer method developed in this lab (Picard et al., 1994). Oligonucleotides were synthesized on an Applied Biosystems 380B DNA synthesizer or purchased from IDT (Coralville, IA). *Pfu* DNA polymerase was from Stratagene. The template for the S137 T and A substitutions was plasmid B1 (Peng, 1992), which is pUC19 with the following inserted between its *Eco*RI and *Bam*HI sites: GAATTCGTCGACGCTAGC, followed by nucleotides 47−1525 of the human ATIII cDNA sequence (Bock et al., 1982), followed by GCTAGCGTCGACGGATCC. The six B series mutants (Figure 3) were produced using the ATIII S137T pBlueBac2 transfer plasmid as template. Mutagenic primers were designed as pools of sequences with CS third-position substitutions and, in addition, incorporated translationally silent changes to introduce/destroy restriction sites for mutant identification. Sequences of mutagenic primers are given below. Mismatches are underlined and restriction site changes have been noted: S137T/A, 5′-TATCGAAAAGCCAACAAA (A,G) CCTCTAAGTTAGTAT (+*Dde*I at 555); T98S/A, 5′-CTGTAATGAC (T,G) CCCTGCAGCAACTGA

(+*Pst*I at 437); T157S/A, 5′-CTTCAATGAG(T,G) CCTAT-CAGGACATCAGT (−*Bst*NI at 616); and T194S/A, 5′-GTCCAATAAG(T,G)CTGAGGGCCGAATCACC (+*Dde*I at 723). Flanking primers used in the PCR mutagenesis procedure were M13#1212, 5′-GTTTTCCCAGTCACGAC; bv.F, 5′-TTTACTGTTTTCGTAACAGTTTTG; and AT3.820R, 5′-TCAGGGCTGAACTTTGAC. M13#1212 and bv.F were respectively used as upstream primers for mutagenesis of the B1 pUC19 and the ATIII S137T pBlueBac2 templates. AT3.820R was used in all mutagenesis reactions as the downstream flanking primer. Transfer plasmids were obtained by three-part ligation of (i) *Nhe*I−*Sac*I fragments from the PCR mutagenesis reactions, (ii) the *Sac*I−*Bam*HI fragment of B1 (nucleotides 767−1525), and (iii) pBlueBac2 (InVitrogen) transfer plasmid prepared with *Nhe*I and *Bam*HI cohesive ends. Preliminary screening for mutation-bearing constructs was performed by restriction digest. Selected mutant clones were sequenced between the *Nhe*I and *Sac*I sites to verify the presence of desired mutations and absence of unplanned ones.

Baculovirus-Mediated Expression of ATIII in Sf9 Cells. pBluebac2 transfer plasmids containing modified ATIII sequences were purified on Qiagen columns and then cotransfected (cationic liposomes) with linear baculovirus DNA (InVitrogen) into Sf9 cells. Recombinant baculoviruses carrying ATIII and β-galactosidase sequences were obtained after homologous recombination had occurred *in vivo*. They were isolated from the supernatant of cotransfected cells by plaque purification under X-gal agarose. Well-isolated blue plaques were used to infect 0.5-mL monolayer cultures of 0.5×10^6 Sf9 cells in 10% fetal bovine serum-containing supplemented Grace's medium (InVitrogen). Lysates were harvested 5 days postinfection (pi); 10-μL aliquots were checked for ATIII expression by Western blot and the remainder was stored as a virus stock. For most mutants, at least two different blue plaques were picked and analyzed and yielded identical results. Larger amounts of recombinant ATIIIs were produced in the absence of serum by infecting shake-flask cultures of logarithmic Sf9 cells (1.5×10^6 cells/mL) in Sf900 medium (Gibco/BRL) with 0.5% (v/v) virus stock. Conditioned media were harvested and frozen at −20 °C 4 days pi.

PNGase F Digestions. Purified human plasma ATIII (Enzyme Research Laboratories, South Bend, IN) and baculovirus-expressed wild-type or N135A ATIII were digested with peptide:*N*-glycosidase F (PNGase F) (New England Biolabs) using conditions recommended by the manufacturer. Briefly, 50 ng of plasma ATIII or 5 μL of 4-day pi culture supernatant were denatured at 100 °C in 0.5% SDS and 1% β-mercaptoethanol. Denatured samples were then treated with PNGase F for 60 min in 50 mM sodium phosphate, pH 7.5, and 1% Nonidet P-40. Limit digests (Figures 1A, 2C, and 3C) used 50 NEB units of enzyme/sample, while partial digests (Figure 1B) used 0.5 NEB unit/sample. Reactions were stopped by addition of electrophoresis buffer, and the samples were boiled for 10 min and subjected to SDS−PAGE and Western blotting.

Heparin Affinity Chromatography. Conditioned medium was passed through a 0.45-μm filter and applied to a 5-mL Hitrap-heparin column (Pharmacia) equilibrated with 0.1 M NaCl and 50 mM Tris, pH 8.0, at 1 mL/min. After extensive washing in 0.1 M NaCl buffer, ATIII was eluted by a linear gradient of 0.1−2.5 M NaCl in 50 mM Tris, pH 8.0. One-milliliter fractions were collected, and 10-μL aliquots of alternate fractions were analyzed by Western blotting.

Electrophoresis and Immunoblotting Techniques. Proteins separated by SDS−10% PAGE under reducing conditions were transferred to PVDF membranes (Immobilon P, Millipore). Membranes were blocked with 5% nonfat dry milk in TBST buffer (50 mM Tris, pH 8.0, 150 mM NaCl, and 0.5% Tween 20) for 30 min and then incubated overnight at 4 °C with polyclonal sheep anti-human ATIII Ig (The Binding Site, Birmingham, UK) (0.2% v/v in TBST with 5% nonfat dry milk). After extensive washing in TBST, blots were incubated for 1−3 h at room temperature with alkaline phosphatase-conjugated donkey anti-sheep Ig (The Binding Site, Birmingham, UK) (0.3% v/v in TBST with 5% nonfat dry milk). Following further washing with TBST, the blots were developed in alkaline phosphatase substrate solution consisting of 16 mg/L 5-bromo-4-chloro-3-indolyl phosphate (XP) and 32 mg/L nitro blue tetrazolium in 100 mM diethanolamine, pH 9.5, 100 mM NaCl, and 5 mM $MgCl_2$.

RESULTS

Characterization of Baculovirus-Expressed ATIII Isoforms. Baculovirus expression of the wild-type human ATIII cDNA sequence with an N-X-S asparagine-135 consensus sequence yields two isoforms, α′ and β′, which migrate in SDS−PAGE as a doublet of greater mobility than the human plasma α and β isoforms (Figure 1A). The β′ isoform comigrates with baculovirus-expressed bv.hat3.N135A, a β-like ATIII variant which lacks the asparagine-135-linked oligosaccharide (Ersdal-Badju et al., in press). This suggested that β′ might be an incompletely glycosylated molecule. PNGase F treatment of baculovirus expressed and plasma ATIIIs produced single bands of identical mobility, indicating (i) that the isoforms composing each doublet differ by their carbohydrate moieties and (ii) that overall, Sf9 cells add less N-linked carbohydrate to the ATIII polypeptide than do human hepatocytes. Ladders generated by partial PNGase F digestion indicate that Sf9 cells utilize all four ATIII CSs but the final sizes of attached oligosaccharides are smaller than those found on plasma ATIII (Figure 1B). To further investigate relationships between the baculovirus-expressed and plasma ATIII glycoforms, we studied their heparin binding. Asparagine-135 is near the ATIII heparin binding site, and β-ATIII binds heparin with higher affinity than α-ATIII because its interaction with the glycosaminoglycan is not sterically restricted by a 135-linked oligosaccharide chain (Brennan et al., 1987; Borg et al., 1988). Figure 1C shows that β′ binds heparin with higher affinity than α′ and elutes from immobilized heparin at increased salt concentrations, as does the β-like N135A variant (Ersdal-Badju et al., in press). These results are consistent with α′ being glycosylated at all four CSs and β′ lacking carbohydrate on asparagine-135.

ATIII Serine 137 Is Responsible for Production of α′ and β′ Glycoforms. To determine if serine-137, the third residue of the asparagine-135 CS, is responsible for glycoform production, we changed the asparagine-135 N-X-S CS to N-X-T and to the nonconsensus sequence N-X-A as a control (Figure 2). In contrast to the mixture of α′ and β′ glycoforms obtained with the wild-type N-X-S CS, the N-X-T S137T

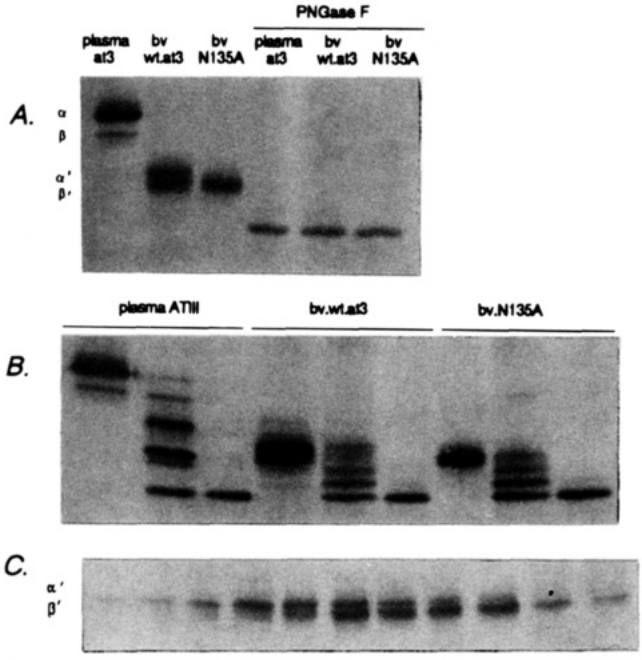

FIGURE 1: Characterization of ATIII glycoforms produced by baculovirus expression. All panels show Western blots prepared from 10% polyacrylamide Laemmli gels of reduced samples and reacted with sheep IgG to human antithrombin III. (A) Resolution of isoforms in human plasma ATIII (50 ng) and baculovirus-expressed wild-type (bv.wt.at3) and N135A ATIIIs. Baculovirus lanes contain 5-μL aliquots of supernatant harvested 4 days pi. Samples on the right were treated with 50 NEB units PNCase F as described in the Materials and Methods section. (B) Partial PNGase F digestion of plasma ATIII and baculovirus-expressed wild-type and N135A ATIIIs. Plasma ATIII (50 ng) or 4-day pi supernatants (5 μL) were treated with 0 (left lanes), 0.5 (middle lanes), or 50 (right lanes) units PNGase F as described in the Materials and Methods section. (C) Relative heparin affinities of baculovirus-expressed α' and β' glycoforms. Supernatant harvested from Sf9 cells 4 days after infection with baculovirus for wild-type ATIII expression was applied to a Pharmacia HiTrap heparin column and eluted with a linear 0.1−2.5 M NaCl gradient.

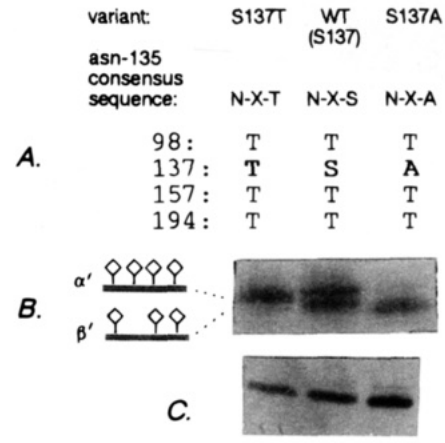

FIGURE 2: Serine 137 is responsible for synthesis of the α' and β' glycoforms. (A) CS third-position residues at the four glycosylation sites of wild-type, S137T, and S137A ATIIIs. (B) ATIII glycoforms produced by Sf9 cells infected with baculoviruses for wild-type, S137T, and S137A ATIII expression. Aliquots (5 μL) of 4-day pi supernatant were applied in each lane. Cartoons illustrate which sites are glycosylated for the α' and β' glycoforms. Western blots were prepared as in Figure 1. (C) Western blots of PNGase F treated of samples shown in panel B.

substitution directed expression of homogeneous, fully glycosylated α'. As expected, the N-X-A S137A control encoding ATIII devoid of N135-linked oligosaccharide produced a single band which comigrated with β'. Comigrating polypeptides were generated by PNGase F treatment of the N-X-S (wild type), N-X-T, and N-X-A variants, indicating that the CS third position residue affects glycosylation. The change in the pattern of protein synthesis from an α'/β' mixture with N-X-S to exclusively α' with N-X-T demonstrates that the third-position serine is directly responsible for the production of ATIII glycoforms in baculovirus-infected insect cells.

Effects of 98, 157, and 194 CS Third-Position Hydroxyamino Acids on ATIII Glycoform Production. We also investigated CS third-position hydroxyamino acid effects at the three other ATIII glycosylation sites, which are all of the N-X-T type in the wild-type protein. Threonine to serine substitutions were introduced on a S137T background, which provided homogeneous, fully glycosylated ATIII as the base molecule. Third-position alanine substitutions were also generated to produce control antithrombins not glycosylated at individual sites. Figure 3 shows the ATIII patterns obtained for these CS variants and the parent molecule (S137T = B). At the asparagine-96 CS, threonine to serine substitution lead to production of ATIII glycoforms with and without asparagine-96 glycosylation, reproducing the effect described at the 135 site. No change was observed in the pattern of glycosylation when the asparagine-155 CS was changed to N-X-S. Finally, serine substitution of the asparagine-192 CS lead mainly to ATIII which was not glycosylated at this site. Thus, at three of the four sites tested (asparagine-96, -135 and -192), glycosylation efficiency was observed to decrease when serine, rather than threonine, was present in the CS third position, and this led to the secretion of ATIIIs which were partially glycosylated or not glycosylated at the serine-containing site. However, at the fourth site (asparagine-155), full glycosylation was achieved with N-X-S as well as with N-X-T, suggesting that in this context other factors influence core glycosylation outcome.

DISCUSSION

The β-ATIII isoform represents 5−10% of circulating antithrombin and has been proposed to be the physiologically most active form of the molecule by virtue of its preferential binding to endothelial and subendothelial HSPGs (Witmer & Hatton, 1991). Using a baculovirus expression system to produce human antithrombin III, we noted that the infected insect cells secreted two ATIII glycoforms, α' and β'. On the basis of glycanase and heparin chromatography data, α' and β' corrspond to the plasma α-ATIII and β-ATIII isoforms, with β', like β-ATIII, also unmodified on a glycosylation site that is near the heparin binding site. The pattern of glycoform production was changed from a mixture of α' and β' to exclusively fully glycosylated α' by changing the hydroxyamino acid in the third position of the asparagine-135 CS from serine to threonine. This result indicates that serine-137 is responsible for inefficient glycosylation of asparagine-135 and production of the α' and β' glycoforms in the baculovirus-infected insect cells.

The process of core N-glycosylation is highly conserved in eukaryotes, based on universal use of N-X-T/S tripeptide consensus sequences as signals for N-glycosylation and recently identified homology between subunits of the

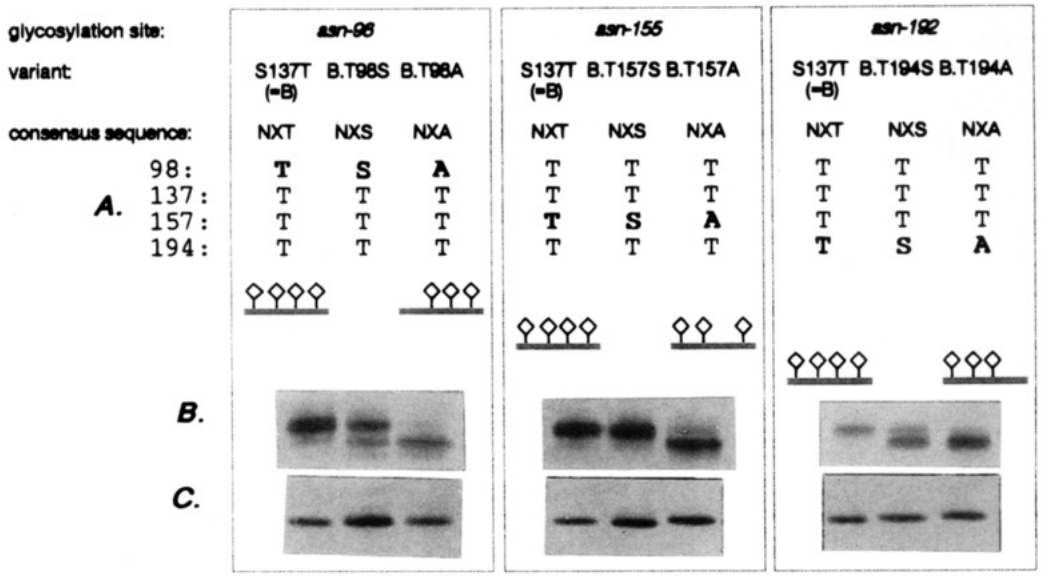

FIGURE 3: Third-position hydroxyamino acid effects on asparagine-96, -155, and -192 glycosylation. The parent molecule for all mutants was S137T, which has four N-X-T CSs and is fully glycosylated. The S137T substitution is abbreviated by the letter B in the names of variants derived from it. (A) Third-position residues at the four antithrombin III glycosylation sites of the parent S137T sequence (left column of each panel) and the asparagine-96, -155, and -192 serine and alanine CS variants (center and right columns of each panel). Cartoons under the N-X-T parent and N-X-A mutants show sites that are glycosylated for these variants. (B) ATIII glycoform production by Sf9 cells infected with baculoviruses for the expression of S137T ATIII and the indicated CS variants. Western blots were prepared as in Figure 1. (C) Western blots of PNGase F treated of samples shown in panel B.

yeast and mammalian oligosaccharyltransferase complexes (Silberstein et al., 1992; Kelleher & Gilmore, 1994). Thus, it is highly probable that serine-137 is responsible for production of the β-ATIII and α-ATIII isoforms in human hepatocytes as well as in baculovirus-infected insect cells.

In addition to addressing the role that the asparagine-135 CS third-position serine plays in ATIII glycoform production, we tested for possible third-position serine effects at the three other ATIII N-glycosylation sites, whose consensus sequences are naturally N-X-T. Glycosylation efficiency similarly decreased when serine replaced threonine in the third positions of the asparagine-96 and -192 CSs. Thus, at three of four ATIII glycosylation sites, CS third-position hydroxyamino acid usage modulated the efficiency at which the associated first-position asparagine was utilized.[2]

This finding is in accord with results from work on cell-free glycosylation of synthetic peptide substrates in which N-X-T-containing peptides were glycosylated 40-fold more efficiently than N-X-S-containing peptides (Bause & Legler, 1981). On the basis of these studies and suicide inhibition experiments with epoxypeptide derivatives (Bause, 1983), Bause formulated a mechanistic scheme for oligosaccharyltransferase which explains the tripeptide CS requirement for N-glycosylation and the catalytic component of more efficient N-X-T glycosylation.[3] This scheme is summarized in Figure 4. Its essential feature is that the CS third-position amino acid is directly involved in catalyzing glycosyl transfer. The side-chain hydroxyl (or sulfhydryl) of the CS third-position amino acid is proposed to participate in "relaying" a proton from the CS first-position asparagine to a basic group in the oligosaccharyltransferase active site, thereby activating the β-amide nitrogen for nucleophilic attack on dolichol pyrophosphate−oligosaccharide. A consequence of this scheme is that glycoslyation efficiency is dependent on the nucleophilicity of the asparagine β-amide nitrogen, which will in turn depend on the basicity of the third-position amino acid that is acting as a glycosylation site-specific component of the enzyme active site. The relative basicities of the third-position hydrogen-bonding residues which can support N-glycosylation, Thr > Ser > Cys, suggest that "threonine-containing oligosaccharyltransferases" (oligosaccharyltransferases glycosylating N-X-T CSs) will be more efficient than "serine-containing oligosaccharyltransferases", as has been observed in the *in vitro* studies and now for ATIII asparagine-135, -96, and -192 N-glycosylation in cell culture.

Interestingly, the oligosaccharyltransferase mechanism proposed by Bause more than a decade ago may represent a naturally occurring example of a subsequently described process called "substrate-assisted catalysis". As a protein engineering principle, this concept was proposed and implemented by Carter and Wells in 1987, when they showed that the activity of a disabled enzyme, lacking an essential catalytic group, could be restored if the missing functionality was supplied through an engineered substrate. Similarly, the CS requirement and sequence specificity of N-glycosylation may be explained as a consequence of a substrate-assisted catalysis mechanism in which asparagine glycosyl acceptor function is dependent on the availability of a nearby hydrogen-bonding amino acid residue to provide an essential catalytic group not endogenously present in the oligosaccharyltransferase active site.

In contrast to the partial glycosylation observed for asparagine-135, -96, and -192 when N-X-S CSs were used,

[2] Production of two glycoforms has also been observed when ATIII arginine-129 was substituted by serine, creating a new consensus sequence (Y. Zuo, A. Lu, and S. C. Bock, unpublished results).

[3] V_{max} and K_m effects observed in synthetic peptide glycosylation experiments (Bause & Legler, 1981) suggest that the CS third position amino acid affects both the catalytic efficiency of oligosaccharyltransferase as well as its binding to the CS.

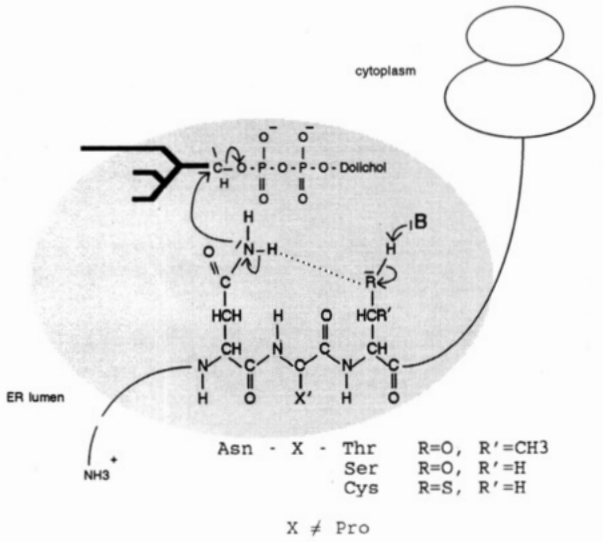

FIGURE 4: Consensus-sequence-assisted catalysis mechanism for OST. This scheme was proposed by Bause and Legler (1981) and represents an example of "substrate-assisted catalysis" as defined by Carter and Wells (1987). N-Glycosylation occurs when preformed core oligosaccharide is transferred from a carrier molecule to the CS first-position asparagine. The asparagine β-amide nitrogen is rendered nucleophilic enough to attack dolichol pyrophosphate-bound oligosaccharide by transfer of one hydrogen to a basic group in the active site of oligosaccharyltransferase. Proton relay is mediated by a hydroxyl or sulfhydryl group in the side chain of the CS third-position residue. Thus, different elements within a single CS play substrate and catalytic roles in the N-glycosylation reaction, and use of a "substrate-assisted catalysis" mechanism allows glycosylation of the first-position asparagine to be modulated in a site-specific manner by properties of the third-position amino acid. Shaded area represents OST; B is a basic group in the active site of OST. The branched structure attached to dolichol pyrophosphate is $Glc_3Man_9GlcNAc_2$; C indicates carbon 1 of the first *N*-acetylglucosamine.

CS third-position serine/threonine effects were not noted at asparagine-155. Lack of third-position effect in some contexts is consistent with instances of N-glycosylated proteins with (i) partially glycosylated N-X-T CSs[4] and (ii) fully glycosylated N-X-S CSs,[5] which indicate that the identity of the CS third-position hydroxyamino acid is not the sole determinant of whether core oligosaccharide addition occurs. Translation, translocation, protein folding, and association of a CS with various endoplasmic reticulum components are thought to influence spatial and temporal accessibility of oligosaccharyltransferase to a CS and ultimate core glycosylation outcome. For example: protein folding and the rate of protein synthesis have been proposed to limit the N-glycosylation of protein C asparagine-329 (Miletich & Broze, 1990); the efficiency of glycosylation at tPA asparagine-184 displays sensitivity to protein folding in a cell-free synthesis system (Bulleid et al., 1992); the glycosylation of integral membrane protein CSs is related to the spacing between the CS and the ends of transmembrane segments (Nilsson & von Heijne, 1993). Finally, on the basis of the cotranslational nature of N-glycosylation (Rothman & Lodish, 1977; Glabe et al., 1980), discontinuous translation rates resulting from mRNA secondary structure and codon usage patterns (Guisez et al., 1993) may affect availability of a CS to oligosaccharyltransferase. Therefore, the failure of an N-X-T to N-X-S substitution to alter the pattern of ATIII asparagine-155 glycosylation may indicate that translation, translocation, folding, and association of this polypeptide region with other ER components are all slow enough to permit full glycoslyation, even when a catalytically less favorable third-position serine is present.

In addition to ATIII, it appears likely that N-X-S and N-X-C CSs may be responsible for partial glycosylation of other glycoproteins. For example, the asparagine-184 CS of human tissue plasminogen activator (tPA) is N-X-S, and it is glycosylated in the type I isoform but not the type II isoform (Bennett, 1983; Pohl et al., 1984). The heavy chain of human protein C has three N-glycosylation CSs (N-X-S at 248, N-X-T at 313, and N-X-C at 329) and circulates as three isoforms: a major α isoform, a β isoform comprising approximately 30% of the total, and a minor γ isoform. The α, β, and γ isoforms appear to represent tri-, di-, and monoglycosylated heavy chains. β-protein C is not glycosylated on asparagine-329, which has an N-X-C CS (Miletich & Broze, 1990), and it is believed that γ-protein C additionally lacks carbohydrate on asparagine-248, which has an N-X-S CS (Grinnell et al., 1991). Thus, the rank abundance of the protein C α, β, and γ isoforms is in agreement with the efficiency of N-X-T, N-X-S, and N-X-C glycosylation predicted by the Bause model. In addition, participation of the N-X-C cysteine in disulfide bonding may further reduce its availability to oligosaccharyltransferase (Miletich & Broze, 1990). In addition to the aforementioned plasma proteins, there is also good evidence for partial glycosylation of bovine α-lactalbumin at its N(45)-X-S CS (Hill & Brew, 1975). We believe it is probable that other cases of N-X-S/C partial glycosylation exist but have not been reported as such in the literature for technical reasons. These would include separation of glycoforms prior to sequencing and the assumption of full and no glycosylation, respectively, for sequencing cycles where trace and nonstoichiometric asparagine were recovered.

Consensus-sequence third-position modulation of N-glycosylation may have functional and evolutionary consequences in cases where the partially glycosylated N-X-S or N-X-C CS is located adjacent to a ligand binding site on the surface of a protein. Inherently inefficient glycosylation of such CSs will allow one polypeptide sequence to direct the production of two different protein species, which may then have different binding (and functional) properties due to steric effects of the variably present oligosaccharide on access of ligand to the binding site. This is true for ATIII, where the α and β isoforms have different affinities for heparin (Carlson & Atencio, 1982; Peterson & Blackburn, 1985); for tPA, where the type I and type II isoforms differ in their fibrin-dependent plasminogen activating activities

[4] Examples of partially glycosylated N-X-T CSs are the human plasminogen 1 and 2 isoforms, which are respectively glycosylated and not glycosylated on asparagine-288 (Hayes & Castellino, 1979), and the bovine pancreatic ribonuclease A and B isoforms, which are respectively not glycosylated and glycosylated on asparagine-34 (Plummer & Hirs, 1964). Partial glycosylation of these N-X-T sequences may occur despite the enhanced catalytic activity associated with N-X-T CSs because the polypeptide regions in which they are located become quickly unavailable to oligosaccharyltransferase due to rapid translation, folding, or other competing interactions. Alternatively, it is also possible that the isoforms lacking carbohydrate on these CSs result from removal of these oligosaccharides from fully glycosylated molecules.

[5] Human tissue plasminogen activator glycoforms have been extensively studied and a large body of evidence indicates that the N-X-S asparagine 117 CS is fully glycosylated (Bennett, 1983; Pohl et al., 1984; Spellman et al., 1989; Parekh et al., 1989).

(Bennett, 1983; Wittwer et al., 1989; Parekh et al., 1989a); and for protein C, where the α and β isoforms have different amidolytic and anticoagulant activities (Grinnell et al., 1991). The invariant third-position serine in the asparagine-135 CS of six different antithrombins (human, Bock et al., 1982; bovine, Mejdoub et al., 1991; sheep, Niessen et al., 1992; mouse, Wu et al., 1992; rabbit, Sheffield et al., 1992; and pig, Tokunaga et al., 1994) suggests that production of β-ATIII has been evolutionarily conserved because it has an important functional role, which is distinct from that of α-ATIII.

In recent years, issues of recombinant glycoprotein heterogeneity have become important in biotechnology. This work shows that there are at least two origins for such heterogeneity. The first relates to the earliest step of N-glycosylation and the occupancy of a CS. The initial core oligosaccharide attachment step of N-glycosylation is a function of oligosaccharyltransferase activity in the endoplasmic reticulum, which we now show can be influenced by identity of the CS third-position amino acid *in vivo*. Patterns of occupancy at a given CS are well conserved in different cell types, consistent with the highly conserved nature of core oligosaccharide addition. For example, both human hepatocyte and lepidopteran insect cells produce two ATIII glycoforms (Figure 1A). The tPA type I and II isoforms have been observed in plasma and for protein expressed in Bowles melanoma, human colon fibroblast, Chinese hamster ovary, and murine C127 cells (Bennett, 1983; Parekh et al., 1989a,b; Spellman et al., 1989). Similarly, chicken oviduct and mouse L cells both N-glycosylate the N-X-T asparagine-293 of ovalbumin but not its N-X-S asparagine-312 (Sheares & Robbins, 1985).

The second source of glycoprotein heterogeneity occurs at the level of oligosaccharide processing and maturation in the Golgi (Kornfeld & Kornfeld, 1985) and is highly variable in different cell types (Parekh et al., 1989a,b), leading to the observation that different oligosaccharides are present on the same polypeptide expressed in different cells, as illustrated in Figure 1B. To date, the issue of glycosylation heterogeneity has been mainly addressed at the second level, that is, by utilizing different host cells for expression of recombinant proteins and by manipulation of processing enzymes. This work demonstrates that modification of the CS third-position amino acid may also be useful for eliminating first-level partial core glycosylation heterogeneity when it is due to an N-X-S or N-X-C CS.

In summary, we have shown that N-glycosylation of three consensus sites in ATIII is modulated in a site-specific manner by the identity of the hydroxyamino acid in the consensus sequence third position and suggest that this effect is due to oligosaccharyltransferase use of a "substrate-assisted catalysis" mechanism in which the CS third-position amino acid plays a central catalytic role. Evidence is presented to show that the serine in the third position of the antithrombin III asparagine-135 CS is responsible for the production of functionally distinct antithrombin III α and β isoforms from a single polypeptide sequence. It is likely that N-X-S and N-X-C consensus sequences contribute to structural and functional diversity of other natural and expressed glycoproteins. Thus, the finding of consensus sequence modulation of N-glycosylation may also have practical applications in biotechnology.

ACKNOWLEDGMENT

We thank Drs. I. Bjork, R. F. Colman, R. W. Colman, C. Grubmeyer, K. McCrae, and S. T. Olson for their comments on this work.

REFERENCES

Bauer, K. A., Goodman, T. L., Kass, B. L., & Rosenberg, R. D. (1985) *J. Clin. Invest. 76*, 826−836.
Bause, E. (1983) *Biochem. J. 209*, 323−330.
Bause, E., & Legler, G. (1981) *Biochem. J. 195*, 639−644.
Bennett, W. F. (1983) *Thromb. Haemostasis 50*, 106A.
Bjork, I., Olson, S. T., Shore, J. D. (1989) in *Heparin Chemical and Biological Properties, Clinical Applications* (Lane, D. A., & Lindahl, U., Eds) pp 229−255, Edward Arnold, London.
Bock, S. C., Wion, K., Vehar, G., & Lawn, R. M. (1982) *Nucleic Acids Res. 10*, 8113−8125.
Borg, J. Y., Owen, M. C., Soria, C., Soria, J., Caen, J., & Carrell, R. W. (1988) *J. Clin. Invest. 81*, 1292−1296.
Brennan, S. O., George, P. M., & Jordan, R. E. (1987) *FEBS Lett. 219*, 431−436.
Bulleid, N. J., Bassel-Duby, R. S., Freedman, R. B., Sambrook, J. F., & Gething, M.-J. H. (1992) *Biochem. J. 286*, 275−280
Carlson, T. H., & Atencio, A. C. (1982) *Thromb. Res. 27*, 23−34.
Carlson, T. H., Atencio, A. C., & Simon, T. L. (1985a) *Biochem. J. 225*, 557−564.
Carlson, T. H., Simon, T. L., & Atencio, A. C. (1985b) *Blood 66*, 13−19.
Carter, P., & Wells, J. A. (1987) *Science 237*, 394−399.
de Agostini, A. I., Watkins, S. C., Slayter, H. S., Youssoufian, H., & Rosenberg, R. D. (1990) *J. Cell Biol. 111*, 1293−1304.
Ersdal-Badju, E., Lu, A., Peny, X., Picard, V., Zendehrouh, P., Turk, B., Björk, I., Olson, S. T., & Bock, S. C. (1995) *Biochem. J.* (in press).
Felsch, J. S., & Owen, W. G. (1994) *Biochemistry 33*, 818−822.
Finazzi, G., Caccia, R., & Barbui, T. (1987) *Thromb. Haemostasis 58*, 1094.
Gavel, Y., & von Heijne, G. (1990) *Protein Eng. 3*, 433−442.
Gilmore, R. (1993) *Cell 75*, 589−592.
Glabe, C. G., Hanover, J. A., & Lennarz, W. J. (1980) *J. Biol. Chem. 255*, 9236−9242.
Grinnell, B. W., Walls, J. D., & Gerlitz, B. (1991) *J. Biol. Chem. 266*, 9778−9785.
Guisez, Y., Robbens, J., Remaut, E., & Fiers, W. (1993) *J. Theor. Biol. 162*, 243−252.
Hayes, M. L., & Castellino, F. J. (1979) *J. Biol. Chem. 254*, 8772−8776.
Hill, R. L., & Brew, K. (1975) *Adv. Enzymol. Relat. Areas Mol. Biol. 43*, 419−425.
Huber, R., & Carrell, R. W. (1980) *Biochemistry 28*, 8951−8966.
Kaplan, H. A., Welply, J. K., & Lennarz, W. J. (1987) *Biochim. Biophys. Acta 906*, 161−173.
Kelleher, D. J., & Gilmore, R. (1994) *J. Biol. Chem. 269*, 12908−12917.
Kornfeld, R., & Kornfeld, S. (1985) *Annu. Rev. Biochem. 54* 631−634.
Marcum, J. A., & Rosenberg, R. D. (1984) *Biochemistry 23*, 1730−1737.
Mejdoub, H., LeRet, M., Boulanger, Y., Maman, M., Choay, J., & Reinbolt, J. (1991) *J. Protein Chem. 10*, 205−212.
Miletich, J. P., & Broze, G. J. (1990) *J. Biol. Chem. 265*, 11397−11404.
Mononen, I., & Karjalainen, E. (1984) *Biochim. Biophys. Acta 788*, 364−367.
Nicchitta, C. V., & Blobel, G. (1993) *Cell 73*, 989−998.
Niessen, R. W. L. M., Sturk, A., Hordijk, P. L., Michiels, F., & Peters, M. (1992) *Biochim. Biophys. Acta 1171*, 207−210.
Nilsson, I., & von Heijne, G. (1993) *J. Biol. Chem. 268*, 5798−5801.
Parekh, R. B., Dwek, R. A., Rudd, P. M., Thomas, J. R., Rademacher, T. W., Warren, T., Wun, T. C., Herbert, B., Reitz,

B., Palmier, M., Ramabhadran, T., & Tiemeier, D. C. (1989a) *Biochemistry 28*, 7670−7679.
Parekh, R. B., Dwek, R. A., Thomas, J. R., Opdenakker, G., Rademacher, T. W., Witwer, A. J., Howard, S. C., Nelson, R., Siegel, N. R., Jennings, M. G., Harakas, N. K., & Feder, J. (1989b) *Biochemistry 28*, 7644−7662.
Peng, X. (1982) M.S. Thesis, Temple University, Philadelphia, PA.
Peterson, C. B., & Blackburn, M. N. (1985) *J. Biol. Chem. 260*, 610−615.
Picard, V., Ersdal-Badju, E., Lu, A., & Bock, S. C. (1994) *Nucleic Acids Res. 22*, 2587−2591.
Plummer, T. H., & Hirs, C. H. W. (1964) *J. Biol. Chem. 239*, 2530−2538.
Pohl, G., Kallstrom, M., Bergsdorf, N., Wallen, P., & Jornvall, H. (1984) *Biochemistry 23*, 3701−3707.
Rothman, J. E., & Lodish, H. F. (1977) *Nature 269*, 775−780.
Sheares, B. T., & Robbins, P. W. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 1993−1997.
Sheffield, W. P., Brothers, A. B., Wells, M. J., Hatton, M. W. C., Clarke, B. J., & Blajchman, M. A. (1992) *Blood 79*, 2330−2339.
Silberstein, S., Kelleher, D. J., & Gilmore, R. (1992) *J. Biol. Chem. 267*, 23658−23663.
Spellman, M. W., Basa, L. J., Leonard, C. K., Chakel, J. A., O'Connor, J. V., Wilson, S., & van Halbeek, H. (1989) *J. Biol. Chem. 264*, 14100−14111.
Tokunaga, F., Goto, T., Wakabayashi, S., & Koide, T. (1994) *J. Biochem. (Tokyo) 116*, 1164−1170.
Witmer, M. R., & Hatton, M. W. C. (1991) *Arterioscler. Thromb. 11*, 530−539.
Wittwer, A. J., Howard, S. C., Carr, L. S., Harakas, N. K., Feder, J., Parekh, R. B., Rudd, P. M., Dwek, R. A., & Rademacher, T. W. (1989) *Biochemistry 28*, 7662−7669.
Wu, J. K., Sheffield, W. P., & Blajchman, M. A. (1992) *Thromb. Haemostas. 68*, 291−296.

BI9500685