Human α1,3/4 Fucosyltransferases

CHARACTERIZATION OF HIGHLY CONSERVED CYSTEINE RESIDUES AND N-LINKED GLYCOSYLATION SITES*

Received for publication, February 1, 2000, and in revised form, May 10, 2000 Published, JBC Papers in Press, May 17, 2000, DOI 10.1074/jbc.M000888200

Eric H. Holmes‡§, Ten-Yang Yen‡$, Scott Thomas§¶, Rajesh Joshi¶, Anton Nguyen¶, Tracy Long¶, Francois Gallet¶, Abderrahman Maftah**, Raymond Julien¶ and Bruce A. Macher‡‡

From the ‡Department of Cell Surface Biochemistry, Molecular Medicine, Northwest Hospital, Seattle, Washington 98134, the Department of Chemistry and Biochemistry, San Francisco State University, San Francisco, California 94132, and *Génétique Moléculaire Animale, UMR 1061, INRA, and **GlycoBio et Biotechnologie, UPRES 1074, Faculté des Sciences, Université de Limoges, 123 Avenue Albert Thomas, 87060 Limoges, France

Human α1,3 fucosyltransferases (FucTs) contain four highly conserved cysteine (Cys) residues, in addition to a free Cys residue that lies near the binding site for GDP-fucose (Holmes, E. H., Xu, Z., Sherwood, A. L., and Macher, B. A. (1995) J. Biol. Chem. 270, 8145–8151). The participation of the highly conserved Cys residues in disulfide bonds and their functional significance were characterized by mass spectrometry (MS) analyses and site-directed mutagenesis, respectively. Among the human FucTs is a subset of enzymes (FucT III, V, and VI) having highly homologous sequences, especially in the catalytic domain, and Cys residues in FucT III and V were characterized. The amino acid sequence of FucT III was characterized. Peptides containing the four conserved Cys residues were detected after reduction and alkylation, and found to be involved in disulfide bonds. The disulfide bond pattern was characterized by multiple stage MS analysis and the use of Glu-C protease and MS/MS analysis. Disulfide bonds in FucT III occur between Cys residues (Cys81 to Cys338 and Cys91 to Cys341) at the N and C termini of the catalytic domain, bringing these ends close together in space. Mutagenesis of highly conserved Cys residues to Ser in FucT V resulted in proteins lacking enzymatic activity. Three of the four mutants have molecular weights similar to wild type proteins composed of partial sequence from each of these FucTs and Cys residues are located near the N terminus and two near the C terminus of the catalytic domain. The Cys residues found at the C terminus also are conserved in FucTs from Caenorhabditis elegans (6).

Among the human FucTs, FucTs III, V, and VI share substantial sequence homology. Within the catalytic domain only about 20 out of 300 amino acids vary among the three proteins. In addition, domain swapping experiments by our group (7) and Lowe and co-workers (9) have demonstrated that chimeric proteins composed of partial sequence from each of these FucTs are active, indicating that the minor differences in their amino acid sequences does not result in major alterations in their overall structure. Therefore, we have used two (FucT III and V) of these highly homologous proteins in the present study to evaluate the structure and functional significance of the highly conserved Cys residues.

Protein chemistry experiments, coupled with mass spectrometry analyses, have been used to locate all peptides containing Cys residues in human FucT III, allowing them to be assigned either as being involved in a disulfide bond or as a free Cys residue, and identifying which Cys residues are bound to each other in disulfide linkages. The results support our (7) previously stated hypothesis that amino acids affecting acceptor substrate specificity of human FucT III and V, located near the N and C termini of the catalytic domain, are brought close together in space by disulfide bonds between these highly conserved Cys residues.

To investigate the importance of these Cys residues, we have mutated each of these Cys residues in human FucT V and evaluated the activity and other properties of each resulting protein construct. The results demonstrate that these residues affect enzyme activity, but not the interaction of the protein with GDP-Fuc in three of the four cases. In the case of one of the mutant constructs (Cys184) protein folding/stability is altered compared with the wild type protein.

The amino acid sequences of FucTs have been predicted on the basis of cDNA sequences, but none of the amino acid se-
sequences have been confirmed directly. Furthermore, little is known about sites of posttranslational modifications. Among the potential sites for posttranslational modification in FucTs, are Asn residues that may be substituted with N-linked glycosylations (Ref. 8 and references therein). Among FucTs III, V, and VI, there are two highly conserved N-linked sites, plus others that are less highly conserved (i.e. they occur in FucT V and VI, but not FucT III). From previous analyses (7, 9), it is clear that some, but not all, of these sites are posttranslationally modified. In the current study, a combination of proteolytic digestions and MS/MS analyses have been used to analyze the amino acid sequence of FucT III. This methodology in combination with PNGase F treatment has been used to locate Asn residues in FucT III that are glycosylated.

**EXPERIMENTAL PROCEDURES**

**Materials—**Unglabeled GDP-Fuc was purchased from Calbiochem (San Diego, CA), and GDP-[3H]Fuc was purchased from NEN Life Science Products. The following items were obtained from HyClone Laboratories Inc. (Logan, UT): Dulbecco’s modified Eagle’s medium, phosphate-buffered saline (PBS; 8.1 mM Na2HPO4, 1.5 mM KH2PO4, 137 mM NaCl, 2.7 mM KCl, pH 7.4), penicillin, streptomycin, and fetal calf serum. Rabbit IgG-agarose, anti-goat IgG-alkaline phosphatase conjugate, goat IgG, bovine serum albumin, 5-bromo-4-chloro-3-indoly phosphate, nitro blue tetrazolium, and formamide were obtained from Sigma. PEO-maleimide-activated biotin was purchased from Pierce. PNGase F and chymotrypsin were purchased from Roche Molecular Biochemicals, and trypsin was obtained from Worthington Biochemical Corp. (Lakewood, NJ). SP-Sepharose was from Amersham Pharmacia Biotech, and the GDP affinity resin was a gift from Dr. Ole Hindsøgaard (Department of Chemistry, University of Alberta, Alberta, Canada). All other chemicals were obtained from commercial sources and were of the highest purity available.

**FucT V Mutants—**The truncated, wild type FucT V was described previously (7) and used as a template to generate the Cys→Ser mutants at Cys94, Cys104, Cys351, and Cys354 and the double mutant Cys94Ser/Cys354Ser. The underlined segments are restriction sites that were used to amplify the Cys mutants. The flanking primers (A and D) were used in the second round of PCR mutagenesis. This new restriction site was used in PCR mutagenesis previously (7) and used as a template to generate the Cys mutant constructs expressed in COS-7 cells and adsorbed onto IgG-agarose beads. These beads were suspended in a 25% bead slurry in PBS. Aliquots (10 μl), each containing 350–540 ng of expressed protein, were dissolved in reaction mixtures, which also contained 0.5 μmol of sodium phosphate buffer, pH 7.0, 0.625 μmol of [3H]-GDP-hexanolamine-ASA (approximately 20 Ci/mmol), with or without 20 μmol of GDP-Fuc, and H2O in a total volume of 60 μl. At the end of 15 min, 0.6 μmol of GDP-hexanolamine-ASA in 35 μl of H2O was added and incubated for an additional 15 min at room temperature. The reaction was stopped by removal of the solution from the IODOBEADS and labeled GDP-hexanolamine-ASA used in photolabeling experiments.

**Peptide Mapping of Beaded Enzyme—** Recombinant PCR was used to create Cys94Ser and Cys104Ser and Cys351Ser and Cys354Ser mutants of FucT V. The truncated, wild type FucT V was described previously (7) and used as a template to generate the Cys→Ser mutants at Cys94, Cys104, Cys351, and Cys354 and the double mutant Cys94Ser/Cys354Ser. The underlined segments are restriction sites that were used to amplify the Cys mutants. The flanking primers (A and D) were used in the second round of PCR mutagenesis previously (7) and used as a template to generate the Cys mutant constructs expressed in COS-7 cells and adsorbed onto IgG-agarose beads. These beads were suspended in a 25% bead slurry in PBS. Aliquots (10 μl), each containing 350–540 ng of expressed protein, were dissolved in reaction mixtures, which also contained 0.5 μmol of sodium phosphate buffer, pH 7.0, 0.625 μmol of [3H]-GDP-hexanolamine-ASA (approximately 20 Ci/mmol), with or without 20 μmol of GDP-Fuc, and H2O in a total volume of 60 μl. At the end of 15 min, 0.6 μmol of GDP-hexanolamine-ASA in 35 μl of H2O was added and incubated for an additional 15 min at room temperature. The reaction was stopped by removal of the solution from the IODOBEADS and labeled GDP-hexanolamine-ASA used in photolabeling experiments.
Conserved Cys Residues in Human Fucosyltransferases

Amino Acid Sequence Analysis of FucT III—Purified FucT III was reacted with PEO-maleimide-activated biotin to label any free Cys residues (one was predicted to occur at Cys143; see Ref. 15), denatured and digested with trypsin under non-reducing conditions. The trypptic digest was separated by liquid chromatography and analyzed for peptides containing modified Cys residues by ESI-MS/MS. A triply charged ion (Fig. 2A) for the Cys-containing peptide (132–151) at \( m/z = 1002.3 \) and a doubly charged ion at \( m/z = 885.4 \) predicted for the peptide with Asn present, and the peptide containing amino acids 161–189 (SDS-DIFTYPGWLEPSGQPAHPL(N → D)-LSAK) gave a \( y_9 \) ion with \( m/z = 840.2 \) instead of 839.2 predicted for the peptide with Asn present. Thus, both of the highly conserved (found in FucT III, V, and VI) N-linked glycosylation sites are glycosylated when FucT III is expressed in the yeast system.

Identification of Free and Disulfide-bonded Cys Residues—Purified FucT III was reacted with PEO-maleimide-activated biotin to label any free Cys residues (one was predicted to occur at Cys143; see Ref. 15), denatured and digested with trypsin under non-reducing conditions. The trypptic digest was separated by liquid chromatography and analyzed for peptides containing modified Cys residues by ESI-MS/MS. A triply charged ion (Fig. 2A) for the Cys-containing peptide (132–151) at \( m/z = 1002.3 \) and a doubly charged ion at \( m/z = 885.4 \) predicted for the peptide with Asn present, and the peptide containing amino acids 161–189 (SDS-DIFTYPGWLEPSGQPAHPL(N → D)-LSAK) gave a \( y_9 \) ion with \( m/z = 840.2 \) instead of 839.2 predicted for the peptide with Asn present. Thus, both of the highly conserved (found in FucT III, V, and VI) N-linked glycosylation sites are glycosylated when FucT III is expressed in the yeast system.

RESULTS

Six human FucTs have been cloned (FucT III to FucT VII and FucT IX). A comparison of the distribution of Cys residues among these enzymes demonstrates that there are four highly conserved Cys residues (see Fig. 1), which are also highly conserved in FucTs from other species. The importance of these conserved Cys residues for protein structure (e.g. disulfide bond formation) and enzyme activity is unknown. Therefore, a series of protein chemistry/mass spectrometry and site-directed mutagenesis studies have been carried out to provide further information on the structure/function relationships of these highly conserved Cys residues.

Amino Acid Sequence Analysis of FucT III—The availability of a yeast expression system for FucT III provided an opportunity to obtain a sufficient quantity of enzyme for a complete characterization of the protein’s amino acid sequence, its disulfide bond pattern and the location of N-linked glycosylation. FucT III was expressed as a soluble protein in P. pastoris (14) and purified in a two-step process that yielded a single Coomassie-stained band by SDS-PAGE. Fractions containing this protein band were highly active (specific activity: 1200 nmol/min mg protein) for fucosyltransferase activity with a type I acceptor. In-gel trypptic digestion products of the band detected on the SDS-PAGE contained peptides derived from the FucT III sequence, and there was no conclusive evidence for the presence of any other protein contaminants (data not shown). A complete peptide sequence analysis of the purified protein solution (reduced with dithiothreitol and alkylated with iodoacetamide) was done using a combination of trypptic digestion and tandem mass spectrometry analysis. More than 95% of the amino acid sequence predicted from the cDNA for FucT III was confirmed by using the protein data base searching program, Sequest (data not shown). Two peptides (amino acids 152–160 and 161–189) that contain an N-linked consensus sequence were not detected. However, after treatment with PNGase F, modified (i.e. peptides with Asn converted to Asp) peptides corresponding to these sequences were detected. MS/MS analysis confirmed that the detected amino acid sequence contained Asp in place of Asn. Thus, the peptide containing amino acids 152–160 (YFPN → DILTMSYR) gave a \( y_7 \) ion with \( m/z = 885.4 \) instead of 884.4 predicted for the peptide with Asn present, and the peptide containing amino acids 161–189 (SDS-DIFTYPGWLEPSGQPAHPL(N → D)-LSAK) gave a \( y_9 \) ion with \( m/z = 840.2 \) instead of 839.2 predicted for the peptide with Asn present. Thus, both of the highly conserved (found in FucT III, V, and VI) N-linked glycosylation sites are glycosylated when FucT III is expressed in the yeast system.

Identification of Free and Disulfide-bonded Cys Residues—Purified FucT III was reacted with PEO-maleimide-activated biotin to label any free Cys residues (one was predicted to occur at Cys143; see Ref. 15), denatured and digested with trypsin under non-reducing conditions. The trypptic digest was separated by liquid chromatography and analyzed for peptides containing modified Cys residues by ESI-MS/MS. A triply charged ion (Fig. 2A) for the Cys-containing peptide (132–151) at \( m/z = 1002.3 \) and a doubly charged ion at \( m/z = 885.4 \) predicted for the peptide with Asn present, and the peptide containing amino acids 161–189 (SDS-DIFTYPGWLEPSGQPAHPL(N → D)-LSAK) gave a \( y_9 \) ion with \( m/z = 840.2 \) instead of 839.2 predicted for the peptide with Asn present. Thus, both of the highly conserved (found in FucT III, V, and VI) N-linked glycosylation sites are glycosylated when FucT III is expressed in the yeast system.

Identification of Free and Disulfide-bonded Cys Residues—Purified FucT III was reacted with PEO-maleimide-activated biotin to label any free Cys residues (one was predicted to occur at Cys143; see Ref. 15), denatured and digested with trypsin under non-reducing conditions. The trypptic digest was separated by liquid chromatography and analyzed for peptides containing modified Cys residues by ESI-MS/MS. A triply charged ion (Fig. 2A) for the Cys-containing peptide (132–151) at \( m/z = 1002.3 \) and a doubly charged ion at \( m/z = 885.4 \) predicted for the peptide with Asn present, and the peptide containing amino acids 161–189 (SDS-DIFTYPGWLEPSGQPAHPL(N → D)-LSAK) gave a \( y_9 \) ion with \( m/z = 840.2 \) instead of 839.2 predicted for the peptide with Asn present. Thus, both of the highly conserved (found in FucT III, V, and VI) N-linked glycosylation sites are glycosylated when FucT III is expressed in the yeast system.
abundant fragments generated by a loss of amino acids from N terminus of peptide 330–339. Low energy collision-induced dissociation resulted in a small amount of S-S bond dissociation, which was only observed at \([m/z] 1288.3 \) (y12)

Interestingly, the MS/MS spectrum of the triply charged ion for \([m/z] 1172.0 \) (see Fig. 4) shows that the dominant fragment ions are generated from a preferential cleavage at the amide bond of Pro86 within the peptide containing Cys 81 and Cys 91. This MS/MS fragmentation pattern is consistent with previous studies of Pro-containing peptides by Loo et al. (16), who observed a similar pattern with peptides derived from several proteins. The fragment ions, \([m/z] 1750.3 \) (b10), 881.1 (y12Y4*), correspond to singly and doubly charged ions of disulfide-containing peptides, respectively. The ion at \([m/z] 881.1 \) is predicted to result from the combination of the two peptides containing Cys81 and Cys343 (peptides containing amino acids 86–97 and 340–343, respectively), and the ion at \([m/z] 1759.1 \) is predicted to result from the combination of the two peptides containing Cys81 and Cys343 (peptides containing amino acids 81–85 and 330–339, respectively). These results suggest that the disulfide linkage patterns are Cys81-Cys343 and Cys81-Cys91. Two additional sets of experiments were used to verify the identity of these ion species as the tripeptide complex and to validate the assigned disulfide bond pattern.

Analysis of the Tripeptide Complex by MS3—An ion trap mass spectrometer is capable of performing multiple stage MS/MS analysis to obtain structural information. MS3 analysis for the MS/MS fragment ions, \([m/z] 881.1 \) and \([m/z] 1750.3 \) was employed to confirm the assignment of the Cys81-Cys343 disulfide bonding pattern. The dominant product ion in the MS3 spectrum of \([m/z] 881.1 \) (y12Y4*) is the doubly
charged ion \( (y_{12}Y_4-H_2O) \) due to the loss of 18 Da (water) from the precursor ion (Fig. 5). All prominent ions observed, as shown in Fig. 3, can be derived from the sequence of disulfide paired peptides (86–97 and 340–343), confirming the disulfide pattern of Cys\(^{91}\)-Cys\(^{343}\). The MS\(^3\) spectrum of the singly charged ion of \( m/z \) 1750.3 (b\(_5\)Y\(_{10}\)) is shown in Fig. 6. As predicted, the dominant fragments are derived from the sequence of the disulfide paired peptides 81–85 and 330–339, with sequential loss of N-terminal amino acids from peptide 330–339; confirming the disulfide pattern of Cys\(^{81}\)-Cys\(^{338}\).

**Analysis of the Tripeptide Complex by Endoproteinase Glu-C Digestion**—To further confirm the identity of the disulfide-containing tripeptide complex, fractions of the tryptic digest obtained from the biotin-modified FucT III that eluted between 34 and 38 min from the LC run were collected and digested with Glu-C. The resulting peptides were analyzed by ESI-MS/MS (Fig. 7). As would be predicted from the specificity of Glu-C, ions corresponding to two disulfide-containing peptides (Cys\(^{91}\)-Cys\(^{343}\) and Cys\(^{81}\)-Cys\(^{338}\)) were detected. These peptides produced doubly charged ions at \( m/z \) 769.7 and 995.8, corresponding to calculated values (\( m/z \) 769.9 and 995.8, due to the cleavage at the C terminus of Glu\(^{83}\) found in the peptide containing amino acids 81–97 and Cys\(^{81}\) and Cys\(^{91}\) (see Fig. 3). The MS/MS analysis (data not shown) of the doubly charged ion at \( m/z \) 995.8 for the disulfide-containing peptide Cys\(^{91}\)-Cys\(^{343}\) showed that the dominant product ions are generated from a fragmentation event at the amide bond of Pro\(^{86}\) in the peptide containing amino acids 84–97 and Cys\(^{81}\), confirming the disulfide pattern of Cys\(^{91}\)-Cys\(^{341}\). The MS/MS spectrum of the doubly charged ion at \( m/z \) 769.7 (Fig. 7) for the disulfide-containing peptide (Cys\(^{81}\)-Cys\(^{338}\)) shows that the dominant product ions are the singly charged ions (\( Y_{17}Y_n \), \( n = 2–9 \)), that result

**FIG. 3.** Scheme of (A) the disulfide-linked tripeptide from FucT III and the products observed in the MS\(^3\) (B and C) and Glu-C experiments (D and E).

**FIG. 4.** MS/MS of the disulfide-linked tripeptide for the ion at \( m/z \) 1172.0. Details of the fragments observed are shown in Fig. 3A and in the inset table.
from the sequential loss of N-terminal amino acids from peptide 330–339, confirming the disulfide pattern of Cys 81–Cys 338 between peptides 81–83 and 330–339.

**Site-directed Mutagenesis of FucT V**—To evaluate the functional significance of the highly conserved Cys residues, site-directed mutagenesis studies were carried out using a FucT (FucT V) that is highly homologous to FucT III, differing at only 21 amino acid residues within the catalytic region. A truncated form of the enzyme, containing only the catalytic domain of the enzyme, coupled to a peptide tag corresponding to the protein A, IgG binding domain was used. We have previously established that this tag does not alter the activity or substrate specificity of any of the human FucTs. Furthermore, the tag provides a means of purifying and detecting the enzyme. Finally, the tagged FucT is isolated after secretion into the cell growth medium. Therefore, any expressed form (i.e. wild type or mutant) must traverse the entire protein secretory pathway before it is isolated for analysis.

Site-directed mutagenesis was used to change each of the conserved Cys residues in FucT V independently to Ser residues, and a double mutant (Cys 351 Ser/Cys 354 Ser) was also produced (Fig. 8, lane 6). As shown in Fig. 8, each mutant construct produced a protein that could be isolated from the cell growth medium by affinity chromatography. Three of the four single mutants, and the double mutant, produced proteins with a molecular weight similar to that of the wild type FucT V (Fig. 8, lane 1). Treatment of the wild type enzyme and these mutants with PNGase F demonstrated that all of these pro-

| Fragment | Charge | Mass |
|----------|--------|------|
| Y_2S     | +      | 727.7|
| b_1Y_2   | +      | 797.2|
| b_2Y_3   | +      | 944.1|
| b_3Y_4   | +      | 1058.9|
| b_3Y_5   | +      | 1171.9|
| b_4Y_6   | +      | 1243.2|
| b_5Y_7   | +      | 1430.5|
| b_6Y_8   | +      | 1516.3|
| b_7Y_9   | +      | 1604.9|
| b_8Y_10  | +      | 1732.3|
| b_9Y_11  | +      | 1750.3|
teins are modified by N-linked oligosaccharides (data not shown). The other mutant (Cys 104 → Ser) produced a broad band on the Western blot that had a faster mobility than the wild type FucT V (Fig. 8, lane 3), and some of these bands had a higher gel mobility after treatment with PNGase F. Each protein was assayed for enzyme activity using a range of known acceptor substrates, including the preferred acceptor substrate for FucT V, H-type II. None of the mutant proteins had detectable enzyme activity, even upon prolonged incubation conditions (data not shown). The lack of enzyme activity could result from a variety of reasons (e.g. disruption of the protein structure or elimination of an essential functional group that assists in substrate binding or the catalytic mechanism). Therefore, further studies were carried out in an effort to better understand the basis for the lack of enzyme activity.

**Photolabeling Experiments**—Although the FucT V Cys mutants were enzymatically inactive, it is possible that they may still bind substrate. Since the $K_m$ for binding the nucleotide sugar donor, GDP-Fuc, is significantly lower than that of the acceptor substrate, photolabeling experiments were conducted using the photofinity probe $^{125}$I-GDP-hexanolamine-ASA (13). Wild type FucT V and FucT V Cys mutants were photolabeled with 12.5 μM $^{125}$I-GDP-hexanolamine-ASA, which is equivalent to 0.5 times the $K_m$ of GDP-hexanolamine with respect to GDP-Fuc (Ref. 13) in the absence and presence of 400 μM GDP-Fuc. After photolysis, the enzyme fractions were separated by SDS-PAGE, and the amount of labeled photoprobe incorporated into protein quantified. The results shown in Table I indicate that a significant level of specific labeling of wild type FucT V occurred (defined as the cpm incorporated into protein in the absence of GDP-Fuc minus cpm incorporated in the presence of 400 μM GDP-Fuc). The extent of protection by GDP-Fuc for wild type FucT V was 58%. The remainder of the cpm incorporated into protein most likely represents nonspecific labeling as the highly reactive nitrene generated upon photolysis reacts with protein sites at random. The level of specific labeling of the FucT V Cys mutants Cys94 → Ser, Cys351 → Ser, and Cys354 → Ser was lower (25–50% of wild type) when compared with wild type FucT V, but still significant. In addition, substantial protection from photolabeling was observed with excess GDP-Fuc, demonstrating that the photoprobe was competing with GDP-Fuc for binding to each protein. In contrast, there was very little specific labeling observed with the FucT V Cys104 → Ser mutant, suggesting this enzyme has lost the ability to bind GDP-Fuc. Since much (>90% estimated from a Coomassie-stained gel) of this protein

| Enzyme | Specific labeling | Protection |
|--------|------------------|------------|
| FucT V | 609              | 58         |
| C94S   | 220              | 36         |
| C104S  | 26               | 11         |
| C351S  | 283              | 29         |
| C354S  | 251              | 42         |

Values are derived from the difference in extent of labeling of enzyme protein in photolabeling reaction mixtures conducted in the absence and presence of 400 μM GDP-Fuc (specific labeling = cpm incorporated in the absence of GDP-Fuc minus cpm incorporated in the presence of 400 μM GDP-Fuc).

Percentage of protection of photolabeling by GDP-Fuc (cpm incorporated in the absence of GDP-Fuc – cpm incorporated in the presence of 400 μM GDP-Fuc) × 100.

After photolysis, the enzyme fractions were separated by SDS-PAGE, and the amount of labeled photoprobe incorporated into protein quantified. The results shown in Table I indicate that a significant level of specific labeling of wild type FucT V occurred (defined as the cpm incorporated into protein in the absence of GDP-Fuc minus the cpm incorporated into protein in parallel reactions in the presence of GDP-Fuc). The extent of protection by GDP-Fuc for wild type FucT V was 58%. The remainder of the cpm incorporated into protein most likely represents nonspecific labeling as the highly reactive nitrene generated upon photolysis reacts with protein sites at random.

The level of specific labeling of the FucT V mutants Cys94 → Ser, Cys351 → Ser, and Cys354 → Ser was lower (25–50% of wild type) when compared with wild type FucT V, but still significant. In addition, substantial protection from photolabeling was observed with excess GDP-Fuc, demonstrating that the photoprobe was competing with GDP-Fuc for binding to each protein. In contrast, there was very little specific labeling observed with the FucT V Cys104 → Ser mutant, suggesting this enzyme has lost the ability to bind GDP-Fuc. Since much (>90% estimated from a Coomassie-stained gel) of this protein

---

**Fig. 7.** MS/MS of the Glu-C product (see Fig. 3D), a doubly charged ion at m/z 769.9. The dominant ions observed are $y_{1-3}$ (n = 2–9) using the notation of fragments in Fig. 3A.

**Fig. 8.** Western blot of Cys mutants of human FucT V. FucT V mutants were prepared by a double PCR mutagenesis technique and expressed as chimeric proteins (catalytic domain of FucT V fused to the IgG binding domain of protein A) in COS-7 cells. The proteins were purified from the growth medium with an IgG resin and chromatographed by SDS-PAGE. Lane 1, wild type FucT V; lane 2, mutant Cys94; lane 3, mutant Cys104; lane 4, mutant Cys351; lane 5, mutant Cys354; lane 6, double mutant Cys351/Cys354.
was apparently degraded, it was not surprising that there was little incorporation of photoprobe.

**DISCUSSION**

The existence of a family of FucTs was initially proposed on the basis of differences observed in substrate specificity and sensitivity to inhibition by an amino acid modifying agent (i.e., NEM) (see Ref. 17 and references therein). However, it was not possible to establish how the enzymes differed until they were cloned. The predicted amino acid sequence information from their cloning provided several clues that have allowed us and others to begin to identify which amino acids give each enzyme its distinct properties (7, 9, 12, 15, 18–21). For example, among the human FucTs, only some (FucTs III, V, and VI) are inhibited by NEM, and we (15) have established that a free Cys residue accounts for this property in the sensitive enzymes, whereas those that are insensitive (22–24) to NEM have a corresponding Ser (FucT IV) or Thr (FucT VII) residue. More importantly, the identification of the NEM-sensitive Cys residue allowed us to pinpoint an amino acid that lies in or near the binding site for GDP-Fuc. In the current study, we have now chemically established that FucT III does contain a single, free Cys residue at this site. We have also identified a Lys residue that also lies in or near the GDP-Fuc binding site. These two residues occur at a significant distance (equivalent to amino acids Cys145 and Lys259 in FucT III) from one another in the FucT sequence and therefore, the native fold of the FucTs must bring these residues near one another in space. Furthermore, we have verified that the amino acid sequence predicted from the corresponding DNA is correct and that the only amino acids modified posttranslationally in the yeast expression system are Asn residues at the two predicted N-linked sites (Asn154 and Asn185). These are the first results available to verify the complete amino acid sequence of any fucosyltransferase, and provide the first information on the location of N-linked site, oligosaccharide occupancy for a fucosyltransferase.

We and others have also identified amino acids that affect acceptor substrate specificity and thus are most likely close to one another in the native protein (7, 9, 12, 18, 20). Our previous studies of FucTs III and V have shown that amino acids affecting acceptor substrate specificity lie at the two ends of the catalytic domain of these two enzymes (7, 20). Based on this observation, and the fact that all mammalian FucTs have two sets of conserved Cys residues at the N and C termini of their catalytic domain, we (see discussion of Ref. 7) proposed that disulfide bonds between one or both pairs of the highly conserved Cys residues are responsible for bringing the two ends of the catalytic domain close together in the native protein, and thus bringing the residues identified as being important for acceptor substrate specificity together. The results reported in this study verify that this hypothesis is true. Furthermore, the results demonstrated that both pairs of conserved Cys residues are involved in disulfide bonds and that Cys91-Cys241 and Cys81-Cys338 of FucT III are bonded together. The latter point was unequivocally established by a combination of MS experiments including MS3 analyses of the disulfide-linked tripeptide tryptic peptide isolated from nonreduced FucT III, and MS/MS analyses of the GlcUc products obtained from the disulfide-linked tryptic peptide. These results establish the maximum distance (not more than 35 Å) that amino acids we had previously demonstrated to affect acceptor substrate specificity (His73-Ile74 and Asp336 in FucT III, separated by more than 250 amino acid residues in the linear sequence) can be from one another in space when the disulfide bonds are formed.

We3 have recently demonstrated that FucT VII has a different disulfide bond pattern than that reported here for FucT III. This protein contains the four highly conserved Cys residues found in other FucTs (Fig. 1) plus two additional, closely spaced Cys residues (Cys211 and Cys214) in the middle of its catalytic domain. Our results demonstrate that all six Cys residues form disulfide bonds and that each closely spaced pair (i.e., Cys86 to Cys76, Cys211 and Cys214, and Cys318 to Cys321) of Cys residues are linked together. This results in a protein containing three short loops, in contrast to the large, single loop pattern that occurs in FucT III. Although we currently do not know whether the Cys residues in other FucTs are involved in disulfide bonds, it seems reasonable to expect that FucT V and VI would have the same disulfide bond pattern as FucT III since these enzymes have very highly conserved amino acid sequences, especially within their catalytic domain. Furthermore, we (12) and Lowe and co-workers (9) have demonstrated that it is possible to swap segments of the corresponding amino acid sequences of these proteins and create active domain swap mutants (i.e., swapping segments that do not contain all of the conserved Cys residues). Thus, it seems likely, although still unproven, that FucTs III, V, and VI would share a common disulfide bond pattern. Since human FucT IV and FucTs expressed in other species share ≤50% sequence homology with FucT III and VII, it is difficult to predict whether they will share the disulfide bond pattern of FucT III or VII, or have a distinct pattern.

The fact that both FucT III and VII bind GDP-Fuc but do not share a common disulfide bonding pattern suggests that the disulfide bonding pattern in these enzymes is not required for or involved in the formation of the GDP-Fuc binding pocket. The results obtained from our photolabeling experiments are consistent with this since the Cys mutants could still be labeled with the GDP-photo probe and labeling was blocked by GDP-Fuc.

Even though the mutant enzymes can bind GDP-Fuc, they are inactive. This may indicate that the acceptor substrate binding pocket for FucT V (and by analogy FucT III and VI) is dependent on the proper formation of disulfide bonds. Since these enzymes have a broader acceptor substrate specificity than FucT VII, one would anticipate that their acceptor substrate binding pockets would differ from that of FucT VII (see Ref. 11 and references therein; see also Refs. 22 and 23). The difference in the disulfide bond pattern of these FucTs could be critical for the formation of these distinct acceptor substrate binding sites. Further studies will be necessary to establish whether this is the case.

Another interesting observation from the mutagenesis studies is that only one of the mutants, FucT V Cys104 → Ser, gave a Western blot pattern that differed significantly from the wild type enzyme. This protein was always found to be substantially degraded, whereas all of the other mutants, including the double mutant, gave a pattern essentially identical to the wild type enzyme. It is not clear why the conversion of one of the highly conserved Cys residues in the FucTs sequence should lead to a significant alteration that allows N-linked glycosylation to occur and...
prohibits proper folding of the protein. Further studies are required to establish the actual cause of the misfolding and degradation of the FucT V Cys\textsuperscript{104} \rightarrow \text{Ser} mutant.

Comparisons of FucT amino acid sequences from several species have suggested that the α1,3 fucosyltransferase family of enzymes were derived from a common ancestral gene, and that the enzymes found today have evolved by gene duplication and divergence (1–3). Gene duplication of a common ancestral gene originated the leukocyte (FucT VII), myeloid (FucT IV), and Lewis (FucT III, V, and VI) subfamilies. Based on the results reported in this study, and those to be reported elsewhere on FucT VII’s disulfide bond pattern, not only have the amino acid sequences of the FucTs diverged during evolution but, so have the pattern of their disulfide bonds. It will be interesting to determine the disulfide bond pattern of the other members of the FucT families and incorporate the resulting information into an assessment of the evolution of this family of enzymes. It is interesting to note that the FucT reported by DeBose-Boyd \textit{et al.} \textsuperscript{(6)} from \textit{C. elegans} does not contain the N-terminal Cys residues conserved in other species and, therefore, would not be capable of forming the disulfide bond pattern found for FucT III, but could form a disulfide bond equivalent to that found in FucT VII, between the highly conserved Cys residues near the C terminus of the catalytic domain. It will be interesting to determine if this is the case or if FucTs from lower order organisms have a completely different pattern.

REFERENCES
1. Breton, C., Oriol, R., and Imberty, A. (1998) \textit{Glycobiology} 8, 87–94
2. Oriol, R., Mollicone, R., Cailleau, A., Balanzino, L., and Breton, C. (1999) \textit{Glycobiology} 9, 323–334
3. Oulmouden, A., Wierinckx, A., Petit, J. M., Costache, M., Palcic, M. M., Mollicone, R., Oriol, R., and Julien, R. (1997) \textit{J. Biol. Chem.} 272, 8764–8773
4. Leiter, H., Mucha, J., Stauffacher, E., Grimm, R., Glosol, J., and Altmann, F. (1999) \textit{J. Biol. Chem.} 274, 21830–21839
5. Kagayaama, N., Natsuka, S., and Hase, S. (1999) \textit{Biochem. (Tokyo)} 125, 838–845
6. DeBose-Boyd, R. A., Nyame, A. K., and Cummings, R. D. (1998) \textit{Glycobiology} 8, 905–917
7. Nguyen, A. T., Holmes, E. H., Whitaker, J. M., Ho, S., Shetterly, S., and Mach, B. A. (1998) \textit{J. Biol. Chem.} 273, 25244–25249
8. Weston, B. W., Smith, P. L., Kelly, R. J., and Lowe, J. B. (1992) \textit{J. Biol. Chem.} 267, 24575–24584
9. Legault, D. J., Kelly, R. J., Natsuka, S., and Lowe, J. B. (1995) \textit{J. Biol. Chem.} 270, 20967–20996
10. Higuchi, R. (1990) in \textit{PCR Protocols: A Guide to Methods and Applications}, (Innis, M. A., Gelfand, D. H., and Sninsky, J. J., eds) pp. 177–183, Academic Press, New York
11. de Vries, T., Srnka, C. A., Palcic, M. M., Swiedler, S. J., van den Eijnden, D. H., and Mach, B. A. (1995) \textit{J. Biol. Chem.} 270, 8712–8722
12. Xu, Z., Vo, L., and Mach, B. A. (1996) \textit{J. Biol. Chem.} 271, 8818–8823
13. Holmes, E. H. (1990) \textit{J. Biol. Chem.} 265, 13150–13156
14. Gallet, P. H., Vanjour, H., Petit, J. M., Maftah, A., Oulmouden, A., Oriol, R., Le Narver, C., Guilloton, M., and Julien, R. (1998) \textit{Glycobiology} 8, 919–925
15. Holmes, E. H., Xu, Z., Sherwood, A. L., and Mach, B. A. (1995) \textit{J. Biol. Chem.} 270, 8145–8151
16. Luo, J. A., Edmonds, C. G., and Smith, R. D. (1993) \textit{Anal. Chem.} 65, 425–438
17. Mach, B. A., Holmes, E. H., Swiedler, S. J. Stults, C. L. M., and Srnka, C. A. (1991) \textit{Glycobiology} 1, 577–584
18. Dupuy, F., Petit, J. M., Mollicone, R., Oriol, R., Julien, R., and Maftah, A., (1999) \textit{J. Biol. Chem.} 274, 12257–12262
19. Sherwood, A. L., Nguyen, A. T. Whitaker, J. M., Mach, B. A., and Holmes, E. H. (1998) \textit{J. Biol. Chem.} 273, 25256–25260
20. Vo, L., Lee, S., Marcinko, M. C., Holmes, E. H., and Mach, B. A. (1998) \textit{J. Biol. Chem.} 273, 25250–25255
21. Kaneko, M., Kudo, T., Iwasaki, H., Ikehara, Y., Nishihara, S., Nakagawa, S., Sasaki, R., Shina, T., Inoko, H., Saitou, N., and Narimatsu, H. (1999) \textit{FEBS Lett.} 452, 237–242
22. Lowe, J. B., Kukowska-Latala, J. F., Nair, R. P., Larsen, R. D., Marks, R. M., Mach, B. A., Kelly, R. J., and Ernst, L. K. (1991) \textit{J. Biol. Chem.} 266, 17467–17477
23. Natsuka, S., Gersten, K. M., Zenita, K., Kannagi, R., and Lowe, J. B. (1994) \textit{J. Biol. Chem.} 269, 16789–16794
24. Sasaki, K., Kurata, K., Funayama, K., Nagata, M., Watanabe, E., Ohita, S., Hana, N., and Nishi, T. (1994) \textit{J. Biol. Chem.} 269, 14730–14737