

Characterization of Glycosylation Sites of the Epidermal Growth Factor Receptor<sup>†</sup>Yuejun Zhen,<sup>‡,§,||</sup> Richard M. Caprioli,<sup>\*,§,⊥</sup> and James V. Staros<sup>\*,‡,#,▽</sup>

Department of Biological Sciences, Vanderbilt University, Nashville, Tennessee 37235, Mass Spectrometry Research Center, Vanderbilt University, Nashville, Tennessee 37232, Department of Biochemistry, Vanderbilt University School of Medicine, Nashville, Tennessee 37232, and Department of Biochemistry and Cell Biology, SUNY–Stony Brook, Stony Brook, New York 11794

Received November 1, 2002; Revised Manuscript Received February 4, 2003

**ABSTRACT:** The epidermal growth factor receptor is a transmembrane glycoprotein that mediates the cellular responses to epidermal growth factor (EGF) and transforming growth factor- $\alpha$  (TGF- $\alpha$ ). In this study of the human EGF receptor naturally expressed in A431 cells, the glycosylation sites of the full-length, membrane-bound receptor and of a secreted form of the receptor were characterized by mass spectrometry. Our data show that the naturally expressed human EGF receptor is fully glycosylated on eight of the 11 canonical sites; two of the sites are not glycosylated, and one is partially glycosylated, a pattern of site-usage similar but not identical to those reported for the recombinant human EGF receptor heterologously expressed in Chinese hamster ovary cells. We also confirm the partial glycosylation of an atypical NNC site first identified in the receptor expressed in Chinese hamster ovary cells. We show that an additional canonical site in the secreted form of the receptor is fully glycosylated. While the pattern of glycosylation is the same for the sites shared by the full-length and the secreted forms of the receptor, the oligosaccharides of the full-length receptor are more extensively processed. Finally, we provide evidence that in addition to the known secreted form of the receptor, a proteolytic cleavage product of the receptor corresponding to the full extracytoplasmic, ligand-binding domain is present in the conditioned medium.

Glycosylation is a ubiquitous feature of multicellular eukaryotic cells, with diversified oligosaccharides present in both the membrane-bound and the secreted proteins that are synthesized in endoplasmic reticulum (ER) and processed in Golgi complex. Oligosaccharides have been shown to play pivotal roles in protein folding and degradation during ER processing (for a review, see ref 1). They are also implicated in maintaining protein stability, modulating protein–protein interactions, etc. (2). Alteration of the structures of oligosaccharides has been observed in a number of diseases, including autoimmune diseases and cancers (3, 4).

The human epidermal growth factor receptor (EGF receptor)<sup>1</sup> is a 170 kDa transmembrane glycoprotein that mediates the mitogenic response of cells to EGF and transforming growth factor- $\alpha$  (TGF- $\alpha$ ). It plays an important role in cell growth and differentiation. The amino acid sequence of EGF receptor has been deduced from the cDNA sequence, and the full-length human EGF receptor has 1186 amino acids (5). The receptor consists of an extracellular EGF-

binding domain, a membrane-spanning segment, an intracellular protein–tyrosine kinase domain, and a carboxyl terminal tail (5). Ligand binding to the extracellular domain causes the receptor to dimerize and activate the protein–tyrosine kinase activity of the receptor (6). The activated tyrosine kinase autophosphorylates five tyrosine residues located at the carboxyl terminal tail of the receptor (7, 8). The phosphorylation provides binding sites for downstream signal transducers, which activate the mitogen-activated protein kinase (MAPK) signal transduction pathway (9).

The human EGF receptor has been found to be highly N-glycosylated but not O-glycosylated (10). Sequence analysis showed that in the extracellular domain, there are 11 potential N-glycosylation sites (NXS/T, X can be any amino acid except proline) (5). Early studies have shown that the oligosaccharides in the EGF receptor contribute about 40 kDa of the mass to the 170 kDa mature protein (11). Biosynthesis of the oligosaccharides of the EGF receptor has been thoroughly investigated using a variety of glycan processing inhibitors and radioactively labeled monosaccharides (10, 12). These studies found that the oligosaccharides of the EGF receptor are important for its translocation and maturation (12, 13). However, the roles of the oligosaccharides in the mature EGF receptor in ligand binding and signal transduction mechanisms are still not clear. One recent study

<sup>†</sup> This work was supported by grants from the National Institutes of Health: R01 GM58008 (R.M.C.), R01 GM55056 (J.V.S.), and R01 DK25489 (J.V.S.).

\* Correspondence may be addressed to the following authors. (R.M.C.) E-mail: r.caprioli@vanderbilt.edu. Telephone: (615) 322-4336. (J.V.S.) E-mail: james.staros@stonybrook.edu. Telephone: (631) 632-6976.

<sup>‡</sup> Department of Biological Sciences.

<sup>§</sup> Mass Spectrometry Research Center, Vanderbilt University.

<sup>||</sup> Current address: Myriad Proteomics, Inc., 2150 West Dauntless Ave., Salt Lake City, UT 84116.

<sup>⊥</sup> Vanderbilt University School of Medicine.

<sup>#</sup> SUNY–Stony Brook.

<sup>▽</sup> Current address: Department of Biochemistry and Cell Biology, SUNY–Stony Brook, Stony Brook, NY 11794.

<sup>1</sup> Abbreviations: EGF, epidermal growth factor; MALDI, matrix-assisted laser desorption/ionization; ACN, acetonitrile; HCCA,  $\alpha$ -cyano-4-hydroxy-cinnamic acid; PNGase F, peptide-N-Glycosidase F; S-EGFR, the secreted 105 kDa form of the EGF receptor; nanoESI, nano-electrospray ionization; MS/MS, tandem mass spectrometry; TFA, trifluoroacetic acid; Hex, hexose; HexNAc, N-acetylhexose; NeuAc, neuraminic acid; Man, mannose; Fuc, fucose;

```

LEEKVKCQGTSNKLTQLGTFEDHFLSLQRMFN32NCEVVLGNLEITYVQRNYDLSFLKTIQEVAGYVLIANTVERI
PLENLQIIRGNMYEENSALAVLSNYDAN104KTGLKELPMRNLQEILHGAVRFSNNPALCNVESIQWRDIVSSDFL
SN151MSMDFQNLGSCQKCDPSCP172GSCWGAGEENCQKLTKIICAQCCSGRCRGKSPSDCCNNQCAAGCTGPRE
SDCLVCRKFRDEATCKDTCPLMLYNPTTYQMDVNPGEKYSFGATCVKKCPRNYVVDHGSVCVRACGADSYEMEEED
GVRCKCKCEGPCRKVCNGIGIGEFKDSLSIN328ATNIKHFN337CTISIGDLHLPLVAFRGDSFTHTPLDPELQELD
ILKTVKEITGFLLIQAWPEN389RTLHAFENLEIIRGRKQHQGQFSLAVVSLN420ITSLGLRSLKEISDGDVIIISG
NKNLCYANTINWKKLFGTSGQKTKIISNRGENSKATGQVCHALCSPEGCWGPEPRDCVSCRN504VSARGRECVDKC
NLLEGEPRFVENSECIQCHPECLPQAMN544ITCTGRGPDNCIQCAHYIDGPHCVKTCPCAGVMGEN579NTLVWKY
ADAGHVCHLCHPN599CTYGCTGPGLEGCP615
      SYIVSHFPRSFYKMSVH  S-EGFR
      PK . . . . .      EC-EGFR

```

FIGURE 1: Amino acid sequences of the secreted (S-EGFR) and the extracellular domain (EC-EGFR) of the full-length EGF receptor with the potential N-glycosylation sites highlighted. The difference between the two sequences lies in the C-termini, as indicated by the split. N<sup>615</sup> is a potential site of glycosylation in S-EGFR but not in EC-EGFR.

found that elimination of the site of N-linked glycosylation at N<sup>420</sup> causes the EGF receptor to dimerize spontaneously (14), suggesting that oligosaccharides play important roles in the function of the EGF receptor. To understand the functional significance of the oligosaccharides, it is necessary to understand fully the glycosylation state of the receptor.

The glycosylation pattern of the extracellular domain of the EGF receptor expressed in Chinese hamster ovary (CHO) cells has been characterized independently by two research groups (15, 16). Though neither of these studies provided data for every potential site of glycosylation, and their data differed for some sites, the results of the two studies taken together show that most of the 11 sites are glycosylated (15, 16). Moreover, a noncanonical N-glycosylation site (N<sup>32</sup>NC) was found to be partially glycosylated in the CHO cell expressed receptor (16). Glycosylation is generally cell and tissue specific, so the glycosylation pattern of the EGF receptor expressed in CHO cells may not apply to the receptor expressed in human cells. The structures of the oligosaccharide moieties from the human EGF receptor have been thoroughly characterized using radioactively labeled sugars (10) and recently using NMR and mass spectrometry techniques (17). These studies showed that the oligosaccharides in the human EGF receptor include both high-mannose-type and complex-type structures. However, the glycosylation sites in the EGF receptor expressed in human cells have not been directly identified.

A human epidermoid carcinoma cell line, A431, has been used widely in studies of the EGF receptor because it expresses about 20 times more EGF receptor than most other cell lines (18). Moreover, in addition to the full-length membrane-bound receptor, it also secretes a truncated 105 kDa soluble form of the receptor (S-EGFR) (5, 19). S-EGFR is identical to the extracellular domain of the full-length receptor except for an additional 17 amino acids at the C-terminus (Figure 1). The extracellular domain from A431 cells has been crystallized (20, 21); however, no high-resolution structure has been reported, likely because of structural and/or conformational heterogeneity of the oligosaccharide moieties of the protein.

The expression of a truncated version of the receptor is not unique to the EGF receptor or to A431 cells. In the EGF receptor family, secreted forms of ErbB2, ErbB3, and ErbB4 have been observed and are attributed to the alternative RNA splicing and metalloprotease cleavage (22–24). This phenomenon has also been reported in a variety of other systems, including the receptors for nerve growth factor (25), growth hormone (26), interleukin 2 (IL-2) (27), and insulin-like growth factor II (28). The widespread presence of the soluble

domains of the receptors suggests that these proteins may have important physiological functions. Studies done on S-EGFR suggest that it may play an important growth-regulatory function (29). Although the secreted and the full-length EGF receptors have an identical amino acid sequence, with the exception of the 17 residue addition at the carboxyl terminus of the soluble form, whether the glycosylation patterns and the structures of oligosaccharides at specific sites are the same is not known.

In this study, the glycosylation patterns of both the secreted form and the full-length receptor expressed in A431 cells were characterized using a combination of MALDI and nanoESI mass spectrometry techniques. This is the first report identifying the sites of glycosylation of human EGF receptor expressed naturally in human cells. We show that in A431 cells, the EGF receptor is fully glycosylated on eight of the 11 canonical sites, two of the sites are not glycosylated, and one is partially glycosylated. This pattern of site-usage is similar but not identical to those reported for the recombinant human EGF receptor expressed in CHO cells (15, 16). Further, we show that the same pattern of site utilization found in the full-length receptor is found in the S-EGFR, although the S-EGFR has an additional site not shared with the full-length receptor that is fully glycosylated. Further, the oligosaccharides of the full-length receptor appear more highly processed than those in the secreted form.

## EXPERIMENTAL PROCEDURES

**Chemicals.** Endoproteinase Lys-C, modified sequencing grade Trypsin, and Glu-C were purchased from Roche (Indianapolis, IN); PNGase F was from Calbiochem (San Diego, CA); HCCA, iodoacetamide, ammonium bicarbonate, acetic acid, TFA, transferrin, hydrocortisone, and sodium selenite were from Sigma (St. Louis, MO); agarose-conjugated anti-EGF receptor antibody (528) was obtained from Santa Cruz (Santa Cruz, CA). Cell culture medium was purchased from the core facility at Vanderbilt University. All solvents used were HPLC grade, and the water was purified using MilliQ water system (Millipore, Bedford, MA).

**Purification of the EGF Receptor.** For the preparation of S-EGFR, A431 cells were grown in the serum-free Dulbecco's modified Eagle's/F-12 medium supplemented with 0.5 mg/L transferrin, 50 nM hydrocortisone, 0.025 mg/L sodium selenite, and 0.1 g/L bovine serum albumin. Two liters of the conditioned medium was subjected to centrifugation at 20 000g to remove cellular debris and was concentrated to about 30 mL using an Amicon stirred cell (Model 8200) with a YM3 membrane (Amicon, Beverly, MA). The

total protein concentration of the concentrated medium was approximately 12 mg/mL. The concentrated medium, after diluting 4-fold with 20 mM HEPES, pH 7.4, was added to agarose-conjugated anti-EGF receptor antibody, and the suspension was rocked at 4 °C for at least 4 h. The mixture was then subjected to centrifugation at 1000g, and the supernatant was discarded. The agarose pellet was washed by resuspension in 20 mM HEPES, pH 7.4, 1 M NaCl and centrifugation as above, and S-EGFR was eluted with 100 mM glycine, pH 2.5, 150 mM NaCl. The eluate, after being dialyzed in MilliQ water overnight, was concentrated using a Speed Vac concentrator (Savant SC100).

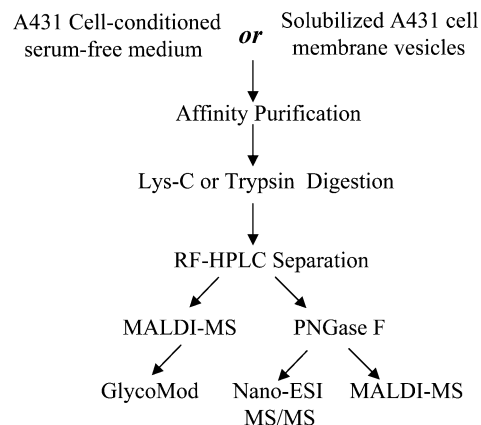
Full-length EGF receptor was purified from A431 membrane vesicles prepared as described in Cohen et al. (30, 31). Two different purification methods were employed in this study, using either agarose-conjugated anti-EGF receptor antibody (528) or EGF affi-gel. The EGF affi-gel was prepared under anhydrous conditions (31), and the purification procedure followed those described in Cohen et al. (30, 31). For the experiment using agarose-conjugated anti-EGF receptor antibody, 1 mL of membrane vesicles was solubilized in 1% Triton X-100, 0.5% Nonidet P-40, 0.1% SDS, 10% glycerol, 20 mM HEPES, pH 7.4, to a concentration of 4–6 mg/mL. The unsolubilized material was removed by centrifugation at 50 000g. The supernatant was added to the agarose beads, and the suspension was rocked at 4 °C overnight. The beads were then washed with solubilization buffer with added 1 M NaCl. The full-length EGF receptor was eluted using 100 mM glycine, 150 mM NaCl, 0.02% Triton X-100. The purified full-length EGF receptor was dialyzed in 20 mM  $\text{NH}_4\text{HCO}_3$  and was concentrated using a Speed Vac concentrator. The protein concentration was determined using BCA reagent (Pierce, Rockford, IL).

**Enzymatic Digestion.** A sample of approximately 30  $\mu\text{g}$  of S-EGFR or full-length EGF receptor was reduced by incubation in 50  $\mu\text{L}$  of 8 M urea, 0.4 M  $\text{NH}_4\text{HCO}_3$  with 10  $\mu\text{L}$  of 45 mM DTT at 60 °C for 15 min. The solution was allowed to cool to room temperature, and 10  $\mu\text{L}$  of 100 mM iodoacetamide was added to the solution, followed by 15 min incubation at room temperature in the dark. The solution was diluted with 120  $\mu\text{L}$  of  $\text{H}_2\text{O}$ , and 15  $\mu\text{L}$  of diluted Lys-C or trypsin (0.1  $\mu\text{g}/\mu\text{L}$ ) was added to the solution. The digestion was allowed to proceed at 37 °C overnight.

**HPLC Separation.** The peptide mixtures from the proteolytic digestion were separated on a Vydac C18 column (2.1  $\times$  250 mm) (Hesperia, CA) using two Waters 515 HPLC pumps and a Waters 996 photodiode array detector system (Waters, Milford MA) controlled by a workstation running Waters Millennium software. Buffer A was 0.1% TFA, and buffer B was 80% acetonitrile (ACN)/0.075% TFA. The flow rate was 0.2 mL/min, and the elution was monitored at 215 and 280 nm. The following gradient was used: 0–10 min, 5% B; 10–65 min, 5–37% B; 65–100 min, 37–75% B; 100–110 min, 75–100% B. HPLC fractions were concentrated with a Speed Vac concentrator.

**Deglycosylation of Peptides.** Five  $\mu\text{L}$  of HPLC fractions containing glycopeptides were diluted in 100 mM  $\text{NH}_4\text{HCO}_3$ . The solution was adjusted to pH 8.0 using 1 N NaOH. PNGase F, 0.2 units, was added to the solution, followed by an overnight incubation at 37 °C. The deglycosylated peptides were purified using a C18 Millipore Ziptip (Millipore, Bedford, MA), and the peptides were eluted directly onto a

Scheme 1: Procedures for the Characterization of EGF Receptor Glycosylation Sites and Oligosaccharide Structures



MALDI plate using 60% ACN/0.1% TFA with 10 mg/mL of HCCA. For nanoESI analysis, the samples were eluted using 60% methanol/1% formic acid. Some deglycosylated peptides were further digested with Glu-C. In these experiments, 0.1  $\mu\text{g}$  of Glu-C (0.1  $\mu\text{g}/\mu\text{L}$ ) was added directly to 5  $\mu\text{L}$  of the deglycosylated solution in  $\text{NH}_4\text{HCO}_3$  after the treatment with PNGase F, and the reaction was allowed to proceed at 37 °C for at least 4 h.

**Mass Spectrometric Analysis.** All MALDI spectra were acquired using an Applied Biosystems Voyager-DE STR (Framingham, MA). The spectra of glycopeptides were recorded in linear mode, and those for deglycosylated peptides were recorded in reflector mode with an accelerating voltage of 25 kV. HCCA was used as the MALDI matrix at a concentration of 10 mg/mL in 60% ACN/0.1% TFA. All nanoESI spectra and MS/MS spectra were recorded using either a Finnigan LCQ (ThermoFinnigan, San Jose, CA) or a SCIEX QSTAR (SCIEX, Ontario, Canada).

**Structural Analysis of Oligosaccharides.** Mass spectrometric data for each oligosaccharide were analyzed for possible oligosaccharide compositions using GlycoMod software (32), with constituent monosaccharide residues limited to hexose (Hex), *N*-acetylhexose (HexNAc), fucose (Fuc), and sialic acid (NeuAc) (17). Possible structures of the oligosaccharides were searched from the GlycoSuiteDB database (33).

## RESULTS

**General Strategies.** The basic strategies used in this study to characterize the glycosylation sites of the EGF receptor are summarized in Scheme 1. Purified samples were digested with Lys-C or trypsin. Since not all of the glycopeptides theoretically possible were observed from the digestions with each of the proteases, digestions with both Lys-C and trypsin provided complementary and confirmative information. The proteolytic peptides were separated by reverse-phase HPLC, each HPLC fraction was analyzed using MALDI-MS, and those fractions containing glycopeptides were identified. A distinct feature of oligosaccharides is their heterogeneity. As a result, glycopeptides tend to yield a series of low intensity peaks in a MALDI-MS spectrum, with mass differences of 146, 162, 291, and 365 Da, corresponding to the residue masses of Fuc, Hex, NeuAc, and Hex-HexNAc, respectively. Those fractions that yielded heterogeneous spectra charac-



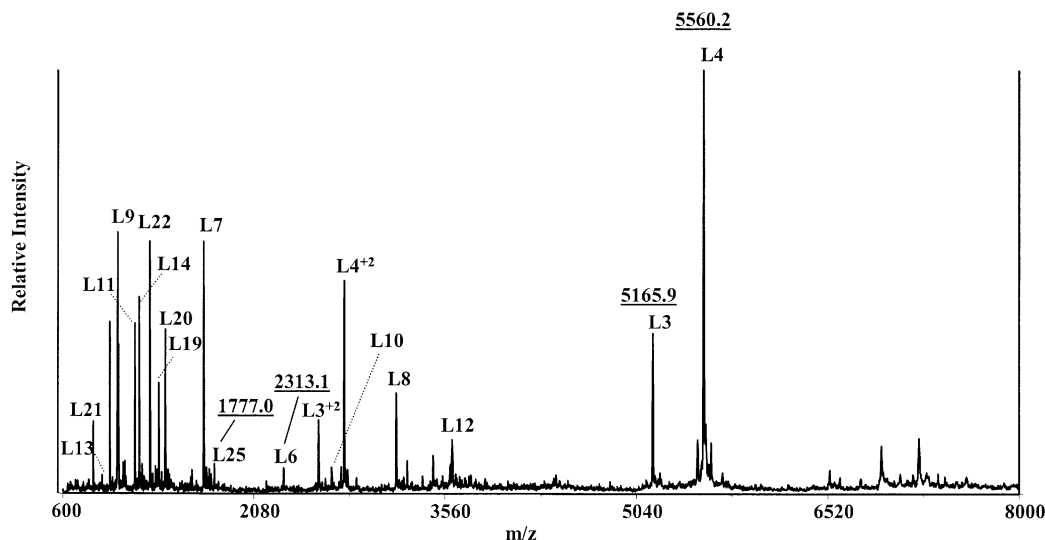


FIGURE 2: Peptide mapping by MALDI-MS of the Lys-C digested peptides from the secreted EGF receptor. The labels indicate the sequence order of peptides of mass > 500 Da in the protein. Monoisotopic masses for peptides L3, L4, L6, and L25 are shown.

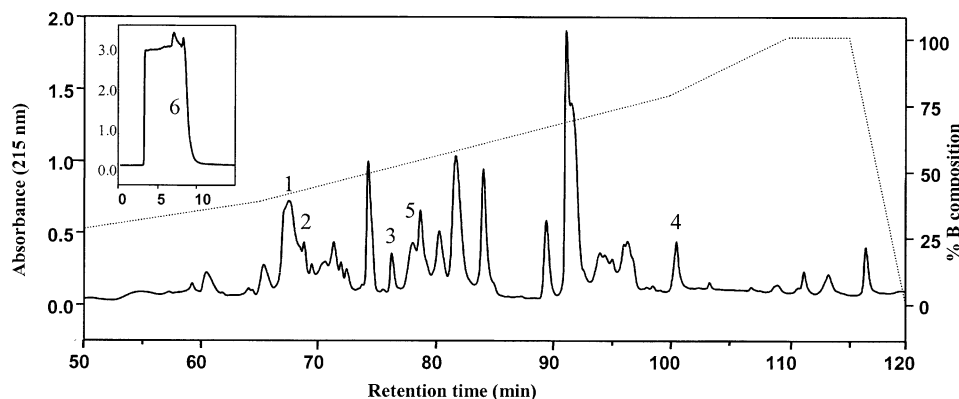


FIGURE 3: HPLC chromatograph of the Lys-C digestion of S-EGFR. The inset shows the first 15 min of the chromatograph.

teristic of glycopeptides were treated with PNGase F, and the deglycosylated samples were analyzed again by MALDI-MS. After the treatment with PNGase F, the deglycosylated peptides appear as new peaks in the mass spectrum. In addition to the removal of the oligosaccharides from the peptides, the treatment with PNGase F also converts Asn residues (114 Da) to which oligosaccharides had been attached to Asp (115 Da), resulting in a 1-Da mass shift of the deglycosylated peptide relative to the corresponding nonglycosylated peptide. The deglycosylated peptides were sequenced using nano-ESI MS/MS to confirm the conversion of Asn to Asp.

**Purification of S-EGFR and Full-Length EGF Receptor.** Several purification methods were tested in this study, and immunoprecipitation with agarose-conjugated anti-EGF receptor antibody (528) was found to give the best purification of S-EGFR. The full-length EGF receptor was purified using either agarose-conjugated anti-EGF receptor antibody (528) or EGF-affi-gel, with the latter producing the purer sample. It was found that in the EGF-affi-gel purification, a wash with 5 mM ethanolamine is important to remove tightly bound contaminants before eluting with 20 mM ethanolamine.

**Peptide Mapping.** Initially, S-EGFR was digested with trypsin, and the peptide mixture was analyzed by MALDI-MS. Only 12 peptides were observed out of the possible 42

tryptic peptides with a mass greater than 500 Da that would be obtained from a complete digest. These 12 peptides account for approximately 40% of the tryptic peptides from a complete digest without potential N-glycosylation sites. In comparison, when the Lys-C digested S-EGFR was analyzed using MALDI-MS, 16 peptides were observed, which include almost all the peptides with no potential glycosylation sites, except the first two N-terminal peptides. Among the 16 peptides observed in the Lys-C digest, four peptides had a monoisotopic mass of 5560.2, 5165.9, 2313.1, and 1777.0, respectively (Figure 2). These four correspond to peptides spanning residues 57–105 (with calculated  $m/z$  5559.9), 14–56 ( $m/z$  5165.6), 166–185 ( $m/z$  2312.8), and 570–585 ( $m/z$  1776.8), respectively. Peptides spanning residues 57–105 ( $N^{104}K(T)$ ), 166–185 ( $N^{172}GS$ ), and 570–585 ( $N^{579}NT$ ) all have the canonical N-glycosylation sites, and the peptide spanning residues 14–56 ( $N^{32}NC$ ) has been observed to be partially glycosylated in the CHO cell-expressed receptor (16). Observation of these four potential glycopeptides as nonglycosylated peptides suggests that these four glycosylation sites are either not glycosylated or are only partially glycosylated in A431 cells.

**Glycopeptide Analysis. Residue  $N^{504}$ .** The HPLC chromatogram of the Lys-C digestion of S-EGFR is shown in Figure 3. MALDI-MS analysis of the fraction containing peak 1 (Figure 3) revealed a series of peaks (from  $m/z$  5500

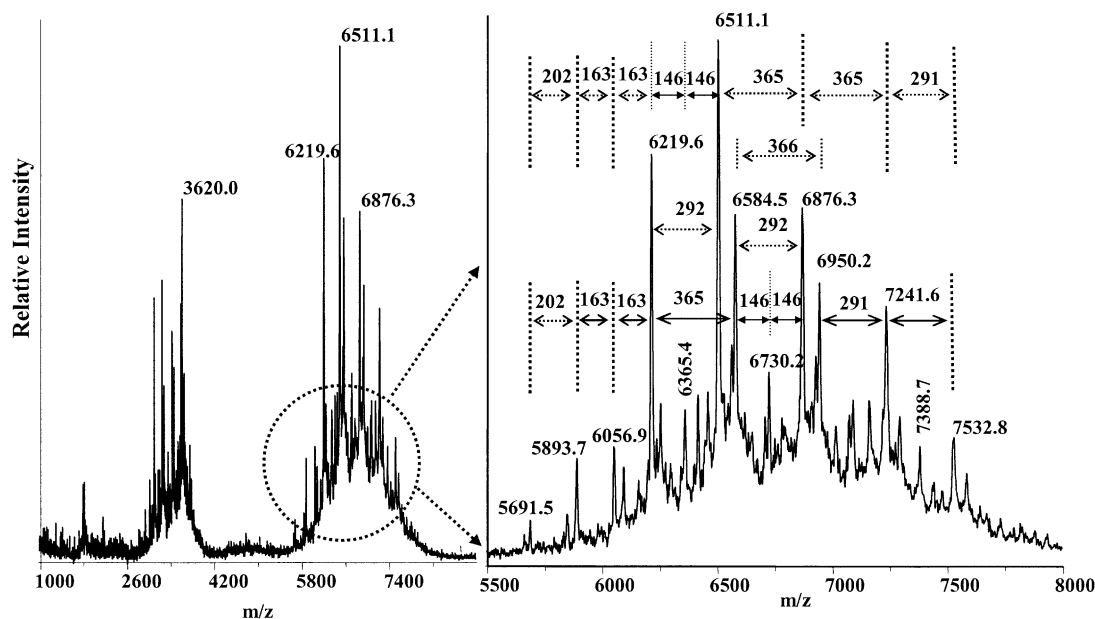


FIGURE 4: MALDI mass spectrum of glycopeptides spanning residues 477–514. The region of single-charged ions was enlarged to show the mass differences of neighboring peaks. Mass differences ( $\pm 1$  Da) corresponding to the monosaccharides are as follows: 146, fucose; 162, hexose; 203, HexNAc; 291, NeuAc; and 365, Hexose + HexNAc.

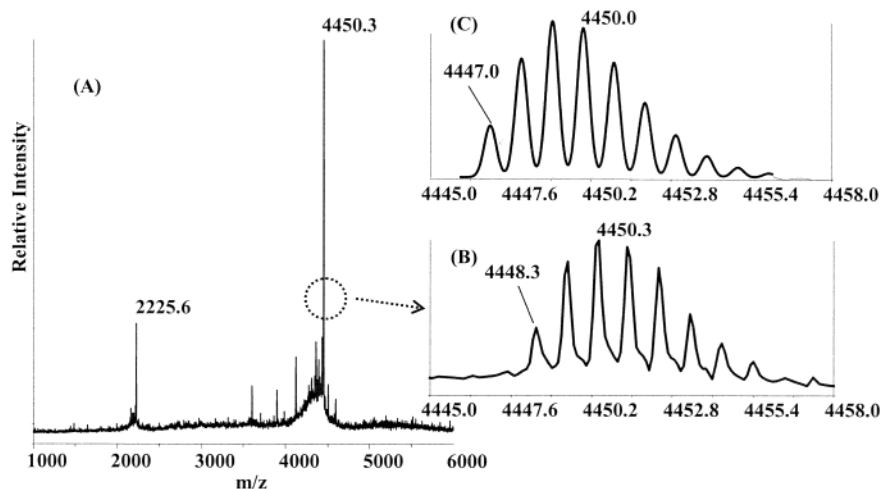


FIGURE 5: MALDI mass spectrum of the peptide spanning residues 477–514 after the treatment with PNGase F (A). The experimental isotopic pattern (B) of the peptide ( $m/z$  4450.3) is compared with the theoretical calculated pattern with Asn as residue 504 (C).

to 8000) centered at  $m/z$  6800 (Figure 4). The mass differences between these peaks were 365, 291, and 162 Da ( $\pm 1$  Da), characteristic of glycopeptides. After the treatment with PNGase F, a new peptide with a monoisotopic mass of  $m/z$  4448.3 was observed in the sample (Figure 5), attributable to the peptide spanning residues 477–514, which has a theoretical monoisotopic mass of 4447.0 Da for the non-glycosylated peptide with an Asn at position 504 (Figure 5). The observed 1-Da increase in the mass of the deglycosylated peptide corresponds to the conversion of the Asn504 to Asp by PNGase F.

The peptide was further digested with Glu-C, and the digestion products were analyzed by MALDI-MS. Two new peptides were observed with monoisotopic masses of 1429.6 and 2535.1 Da (data not shown), which correspond to peptides spanning residues 477–489 and 490–510, respectively. The peptides from the Glu-C digestion were further analyzed using nano-ESI MS/MS. In the full scan mode, the +2, +3, and +4 charged ions for the peptide containing

residues 490–510 were observed (data not shown). However, MS/MS analyses of the +2 and +3 charged ions produced few fragments. The MS/MS spectrum of the +4 charged ions ( $m/z$  634.53) is shown in Figure 6. The ions resolved in the MS/MS spectrum were compared with the theoretical values of the product ions of peptide spanning residues 490–510, and a number of  $b$  and  $y$  ions were assigned to the sequence. Among the  $b$  and  $y$  ions identified, a pair of ions,  $m/z$  409.72 and 916.89, was observed, which correspond to the doubly charged  $y_7$  ( $m/z$  409.71) and doubly charged  $b_{15}$  ( $m/z$  916.87) ions, respectively, with Asp at position 504. With Asn at position 504, the masses of the corresponding doubly charged  $y_7$  and  $b_{15}$  ions would be  $m/z$  409.22 and 916.38, respectively. On the basis of the instrumental error of  $m/z \pm 0.03$ , Asp can be assigned with confidence to position 504, resulting from conversion of Asn504 to Asp by the treatment with PNGase F.

The sugar compositions of the various glycoforms of peptide 477–514 were first analyzed manually by calculating

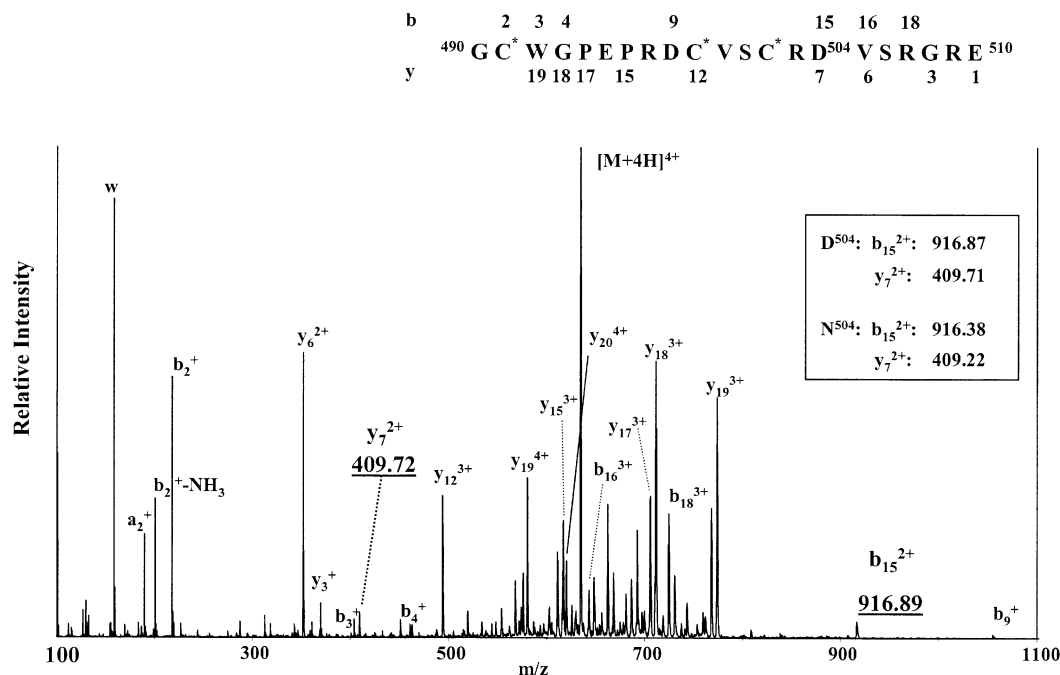
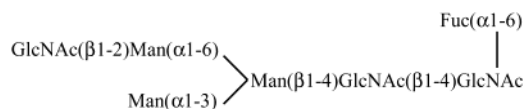


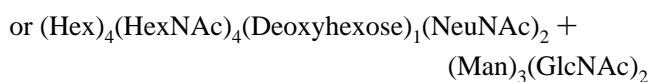
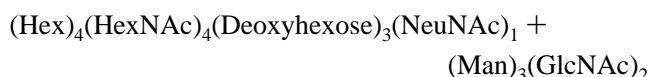
FIGURE 6: ESI/MS/MS spectrum of the peptide spanning residues 490–510. The inset shows the predicted values for  $y_7^{2+}$  and  $b_{15}^{2+}$  ions when a residue of Asp or Asn is located at position 504. A number of ions in the region from  $m/z$  500 to 800 are from internal fragmentation.

the mass differences among the peaks resolved in the spectrum (Figure 4) and then further analyzed using GlycoMod. Since in the studies of S-EGFR done by Stroop et al. (17, 34) only hexose, *N*-acetylhexose, fucose, and sialic acid were observed, these four monosaccharides were selected as the possible residues of the oligosaccharides, simplifying search results.

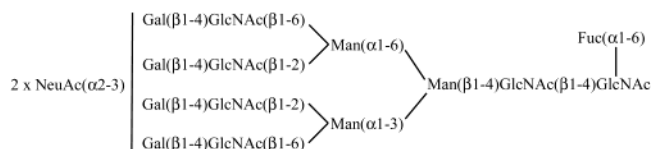
The resolved glycopeptides in Figure 4 have masses ranging from  $m/z$  5691.5 to 7532.8. Structural analysis using GlycoMod suggests that the ion at  $m/z$  5691.5 most likely has the following glycan structure:



All of the other resolved peaks from S-EGFR can be categorized as derivatives of the glycoform at  $m/z$  5691.5 since their mass differences correspond to unit masses of monosaccharides. Manual analysis of the mass difference between the largest glycoform at  $m/z$  7532.8 and the smallest at  $m/z$  5691.5 in Figure 4 suggested that the structural difference between these two glycoforms corresponds to a sugar composition of (Hex)<sub>4</sub>(HexNAc)<sub>3</sub>(NeuAc)<sub>1</sub>(Fuc)<sub>2</sub>. When analyzing the  $m/z$  7532.8 glycoform using GlycoMod, six possible sugar compositions were suggested. However, based on the observation (Figure 4) that this glycoform contains at least four Hex, three HexNAc, one NeuNAc, and one fucose residue, only two of the six possible compositions meet these criteria:



Since the mass of a sialic acid residue (291 Da) is about twice the mass of a fucose residue (146 Da), the  $m/z$  7532.8 glycoform mathematically fits both compositions. Database search using the software GlycoSuiteDB suggests that the glycan with  $m/z$  7532.8 is most likely to have the following structure:



In the structure, the two sialic acid residues are linked to the galactose residues at the nonreducing end, as observed in the study by Stroop et al. (17). However, this analysis does not reveal to which galactose residues they are linked.

**Residue N<sup>579</sup>.** In the MALDI mass spectrum of the HPLC fraction containing peak 2 (Figure 3), a series of peaks of low intensity with the characteristics of glycopeptides was observed in the range from  $m/z$  5500 to 8000. After the treatment with PNGase F, two new peptides,  $m/z$  1777.6 and 4448.4, appeared in the spectrum. The peptide with a monoisotopic mass of 4448.4 Da has been shown (above) to be the peptide with residues 477–514, which is dominant in HPLC peak 1, and the different glycoforms observed in the region from  $m/z$  5500 to 8000 are also likely to be from this peptide. The appearance of a new peptide with a mass of 1777.6 suggests the presence of another glycopeptide in peak 2, less prominent than the peptide with residues 477–514. The new peptide with  $m/z$  1777.6 corresponds to the peptide spanning residues 570–585 (Table 1). Sequence analysis of this ion confirmed that it is indeed the peptide spanning residues 570–585. Moreover, sequence analysis showed that the residue at position 579 is Asp, suggesting that the N<sup>579</sup>NT site had been glycosylated. However, when analyzing the HPLC fraction containing peak 3 (Figure 3),

Table 1: Peptide Mapping of the Potential Glycopeptides in the Secreted EGF Receptor<sup>a</sup>

potential sites	Trypsin digestion				Lys-C digestion				glyco. state
	position	calculated [M + H] <sup>+</sup>	measured [M+H] <sup>+</sup>	mass diff	position	calculated [M + H] <sup>+</sup>	measured [M + H] <sup>+</sup>	mass diff	
N <sup>32</sup> NC	30–48	2299.1	2299.2	0.1	14–56	5165.6	5165.9	0.3	partial
		2300.1 <sup>b</sup>	2299.9	−0.2		5166.6 <sup>b</sup>	5166.8	0.2	
N <sup>104</sup> KT	85–105	2400.1	2399.9	−0.2	57–105	5559.9	5559.9	0	not
N <sup>151</sup> MS	142–165	2761.2 <sup>b</sup>	2761.1	−0.1	110–165	6538.4 <sup>§</sup>	6538.3 <sup>§</sup>	−0.1	100%
N <sup>172</sup> GS	166–185	2312.8	2312.7	−0.1	166–185	2312.8	2313.1	0.3	not
N <sup>328</sup> AT	323–333	1175.6			323–333	1176.6 <sup>b</sup>	1176.4	−0.2	100%
N <sup>337</sup> CT	337–353	1901.0 <sup>b</sup>	1900.8	−0.2	337–372	4005.0 <sup>b</sup>	4004.9	−0.1	100%
N <sup>389</sup> RT	376–390	1788.0 <sup>b</sup>	1787.9	−0.1	376–407	3782.0 <sup>b</sup>	3781.7	−0.3	100%
N <sup>420</sup> IT	408–427	2141.2 <sup>b</sup>	2141.2	0	408–430	2469.4 <sup>b</sup>	2469.2	−0.2	100%
N <sup>504</sup> VS	504–507	475.5			477–514	4447.9 <sup>b</sup>	4448.3	0.4	100%
N <sup>544</sup> IT	524–550	3281.4 <sup>b</sup>	3281.5	0.1	515–569	6571.8 <sup>b</sup>	6571.5	−0.3	100%
N <sup>579</sup> NT	570–585	1776.8	1776.8	0	570–585	1777.8 <sup>b</sup>	1777.6	−0.2	80%
		1777.8 <sup>b</sup>	1777.8	0		1776.8	1777.0	0.2	
N <sup>599</sup> CT									
N <sup>615</sup> GS	586–625	4520.9 <sup>b</sup>	4520.4	−0.5	586–629	5046.2 <sup>b</sup>	5046.6	0.4	100%

<sup>a</sup> All the masses listed, except those labeled with §, are the monoisotopic masses of the peptides. <sup>b</sup> Post-deglycosylation masses of the peptides.

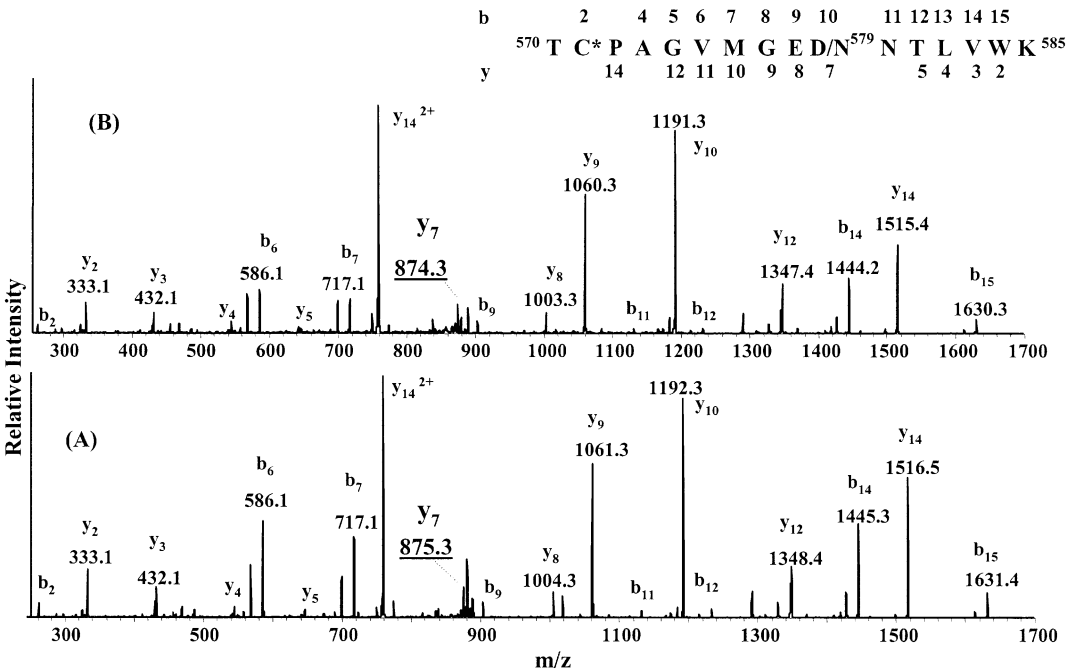


FIGURE 7: Comparison of ESI/MS/MS spectra of the deglycosylated (A) and nonglycosylated (B) peptide spanning residues 570–585. A one mass unit difference is observed between the deglycosylated peptide and the nonglycosylated form because of the conversion of Asn to Asp.

a peptide with  $m/z$  1777.0 was observed. Sequence analysis shows that this peptide also corresponds to residues 570–585 and that an Asn was present at position 579. The sample analyzed in Figure 3 had not been treated with PNGase F. Thus, the appearance of unmodified N<sup>579</sup> in this sample suggests that the fraction of the peptide in peak 3 was not glycosylated. The MS/MS spectra for both the nonglycosylated and the deglycosylated samples are shown in Figure 7. The 1-Da shift was clearly observed in all the ions containing residue 579, confirming that the glycosylation site at position 579 is partially utilized. Similar results were obtained from analyzing peptides from a tryptic digest. The glycosylated and nonglycosylated peptides from the tryptic digest were found to be present in the same HPLC fraction. In the MALDI mass spectrum of the deglycosylated sample of the tryptic digest, the monoisotopic peak (1777.6) of the deglycosylated peptide overlaps with the second isotopic

peak of the nonglycosylated peptide. The spectrum was deconvoluted using the deisotoping function of the Data Explore software supplied with the MALDI instrument, and the relative ratio of the two peptides was measured. The glycosylated fraction was found to account for about 80% of the overall peptide. No glycoforms were detected for the peptide spanning residues 570–585; therefore, the structures of its oligosaccharides could not be determined.

**Residue N<sup>32</sup>.** When analyzing the HPLC fraction containing peak 4 (Figure 3), a series of peaks having the characteristic pattern of a glycopeptide was observed to be centered at  $m/z$  ~7000 (Figure 8). After deglycosylation with PNGase F, a new peptide was observed with a monoisotopic mass of  $m/z$  5166.8, which corresponds to the mass of the deglycosylated peptide spanning residues 14–56 (Table 1). Further digestion of the deglycosylated sample with Glu-C followed by MS/MS sequence analysis confirmed that the

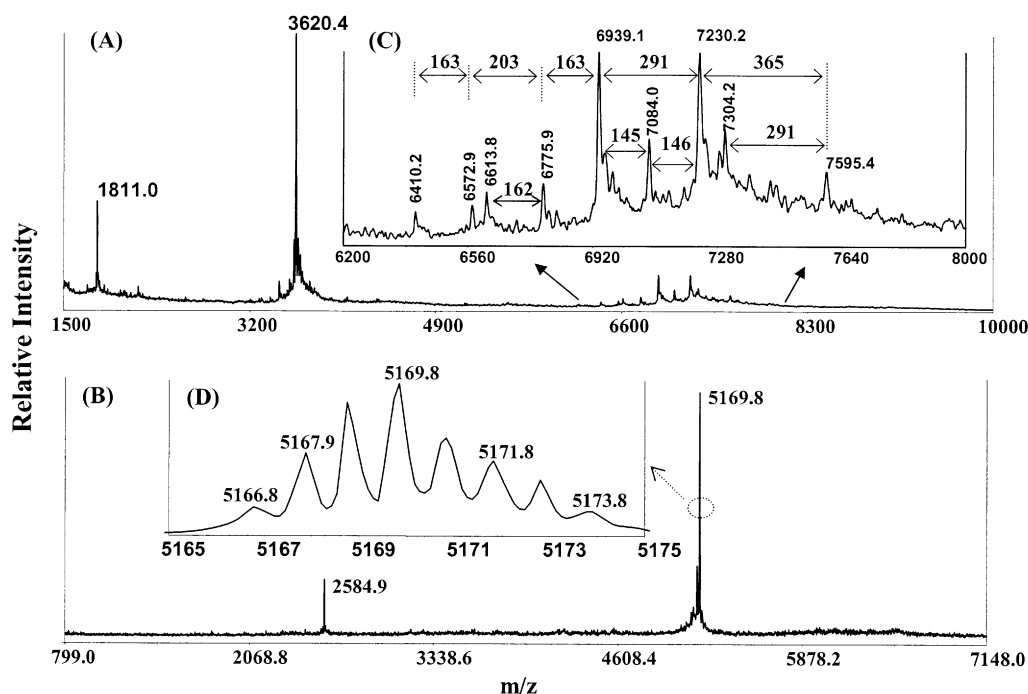
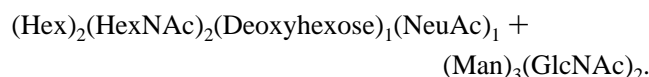


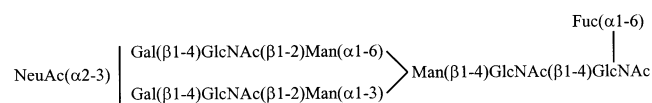
FIGURE 8: Comparison of MALDI mass spectra of an HPLC fraction containing a glycopeptide spanning residues 14–56 before (A) and after (B) the treatment with PNGase F. Inset C shows the mass differences of neighboring peaks in the region from 6200 to 8000. Inset D shows the isotopic pattern of ion with  $m/z$  5169.8.

peptide with  $m/z$  5166.8 is composed of residues 14–56 and that Asn 32 was found to be converted to Asp, confirming that  $N^{32}NC$  is a site of N-glycosylation (16). However, the nonglycosylated peptide with residues 14–56 was also observed (Figure 2, Table 1). Thus, the site at  $N^{32}NC$  is partially glycosylated in the S-EGFR from A431 cells.

Analysis of the various glycoforms of peptide 14–56 (Figure 8) suggests that the glycoform with  $m/z$  7595.4 contains at least one NeuAc, one fucose, one HexNAc, and two hexose residues. Further analysis using GlycoMod suggests only one oligosaccharide composition that matches the peptide mass:



Database search using GlycoSuiteDB suggests that the oligosaccharide is likely to have the following structure:

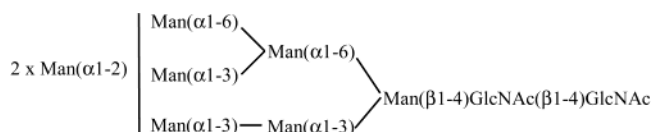


This noncanonical N-glycosylation site in the EGF receptor was first reported by Sato et al. (16). Detailed structural analysis using exoglycosidases coupled with HPLC characterization carried out in that study revealed that the oligosaccharide at the same site from CHO cell-expressed receptor has the same structure as that depicted above.

**Residue  $N^{599}$ .**  $N^{599}CT$  is the most C-terminal glycosylation site in the extracellular domain of the full-length EGF receptor; however, S-EGFR has an additional potential glycosylation site at  $N^{615}GS$ , which is not present in the full-length receptor. In the HPLC separation of the Lys-C digest of S-EGFR, peptide 586–629, containing these two sites, was found in peak 5 (Figure 3). The MALDI mass spectrum

of peak 5 displays a broad cluster of peaks from  $m/z$  8500 to 12 000 (Figure 9). After deglycosylation, a peptide with a monoisotopic mass of 5046.6 was observed (Table 1), which is the expected mass of the peptide spanning residues 586–629. The deglycosylated peptide was digested with Glu-C, and the sequences of the two peptides containing the two glycosylation sites were subjected to MS/MS analysis (data not shown). These results confirmed the sequences and showed that  $N^{599}$  and  $N^{615}$  had been converted to Asp, suggesting that both  $N^{599}$  and  $N^{615}$  were glycosylated in S-EGFR.

**Other Sites.** All of the other eight potential N-glycosylation sites were also characterized using the approach described above, and the results are summarized in Table 1. Two potential glycosylation sites,  $N^{104}KT$  and  $N^{172}GS$ , were found to be not glycosylated (i.e., the corresponding peptides were observed in tryptic and Lys-C digestions (Figure 3) but not in the PNGase F digest of glycopeptide fractions). The other six potential sites,  $N^{151}MS$ ,  $N^{328}AT$ ,  $N^{337}CT$ ,  $N^{389}RT$ ,  $N^{420}IT$ , and  $N^{544}IT$ , were found to be completely glycosylated (i.e., they were observed only after PNGase F treatment, with glycosylated Asn residues converted to Asp). Analysis of the mass spectrum of the glycosylated peptides suggests that the oligosaccharides at  $N^{337}CT$  are high-mannose type glycans with the following structure:



The oligosaccharides at  $N^{389}RT$  and  $N^{544}IT$  are likely to be complex type N-glycans with structures similar to those at  $N^{504}VS$ . Similar tetraantennary complex type N-glycans, but with four sialic acid residues at the nonreducing end, were previously found in the study of Stroop et al. (17). In their study, this tetrasialylated glycan was found to be the



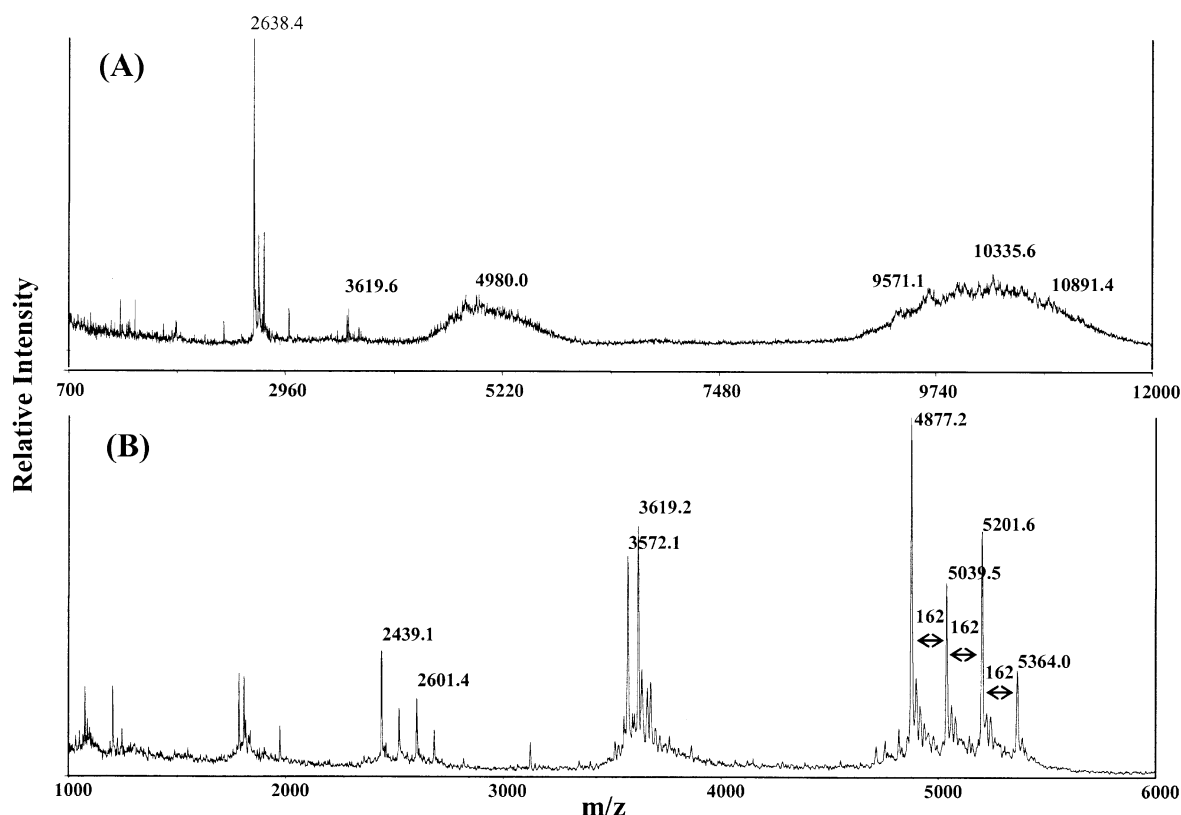
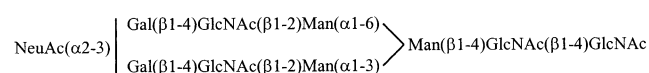


FIGURE 9: MALDI mass spectra of two Lys-C digested glycopeptides containing the glycosylated site N<sup>599</sup>CT in the EGF receptor. (A) Spectrum of HPLC peak 5 (Figure 3), which includes the peptide spanning residues 586–629 in S-EGFR. (B) Spectrum of HPLC peak 6 (Figure 3), which includes the peptide spanning residues 586–618 in the full-length EGF receptor sequence.

major glycoform of *N*-glycans of the secreted form of the EGF receptor. Sialic acid residues tend to be unstable during MALDI analysis, and losses of sialic acid may occur during analysis (35–37). It is not unlikely, therefore, that the tetra-antennary structure that we have identified could correspond to the tetrasialylated species observed by Stroop et al. (17).

The oligosaccharide at N<sup>420</sup>IT is likely to be a biantennary complex type *N*-glycans with the following structure:



The glycoforms at N<sup>151</sup>MS, N<sup>328</sup>AT, and N<sup>579</sup>NT sites were not well-resolved, and the structures of the oligosaccharides could not, therefore, be inferred.

**Presence of a Proteolytically Cleaved Form of the EGF Receptor in the Culture Medium.** When analyzing one of the HPLC run-through fractions of the Lys-C digest of the soluble form of the receptor (peak 6 in Figure 3), four ions of high intensity were observed with a mass difference of 162 Da between peaks next to each other (Figure 9B), suggesting that there is a glycopeptide in the fraction. The sample was deglycosylated, and the deglycosylated peptide was found to have a monoisotopic mass of 3658.9 Da (data not shown), which is the expected mass of the peptide spanning residues 586–618 in the full-length receptor (Figure 1 and Table 2). Sequence analysis of this peptide confirmed that it is the same as peptide 586–618 from the full-length receptor.

It is well-known that the culture medium of A431 cells contains a large amount of secreted EGF receptor (5), and the sequence of this protein has been reported (5). The current

study for the first time provides evidence suggesting that, besides the secreted EGF receptor, there is another soluble form of the EGF receptor present in the culture medium, which includes Lys618 from the full-length receptor (Figure 1). For other receptors of the ErbB family, metalloproteases have been shown to cleave the receptors at sites close to the membrane, producing soluble forms of their extracellular domains (22–24). This is the first study of which we are aware to provide evidence that the EGF receptor is subject to similar proteolytic cleavage.

Several glycoforms of the peptide 586–618 from the soluble fragment of the full-length EGF receptor were well-resolved (Figure 9B), and the structures of the oligosaccharides at this site were analyzed using the methods discussed above. The observed glycoforms were separated by a mass difference of 162 Da, suggesting that the oligosaccharides at position N<sup>599</sup> are likely high-mannose type *N*-glycans with a structure similar to those at site N<sup>337</sup>.

**Characterization of the Glycosylation Sites in the Full-Length EGF Receptor.** The full-length EGF receptor was purified using either anti-EGF receptor antibody or EGF affigel, and the purified protein was digested using either Lys-C or trypsin. Fewer glycopeptides from the full-length receptor were identified in each digestion as compared to those identified in the soluble extracellular domain studies, likely because of the presence of detergent in the samples, which suppresses some ion signals. In the Lys-C digested solution, peptides containing glycosylation sites, N<sup>32</sup>NC, N<sup>104</sup>KT, and N<sup>151</sup>MS, were found to coelute with Triton X-100 during HPLC separation, making their analysis by MALDI-MS problematic. Peptides with these glycosylation sites were

Table 2: Peptide Mapping of the Potential Glycopeptides in the Full-Length EGF Receptor<sup>a</sup>

potential sites	Trypsin digestion				Lys-C digestion				glyco. state
	position	calculated [M + H] <sup>+</sup>	measured [M + H] <sup>+</sup>	mass diff	position	calculated [M + H] <sup>+</sup>	measured [M + H] <sup>+</sup>	mass diff	
N <sup>32</sup> NC	30–48	2300.1 <sup>b</sup>	2300.2	0.1	14–56	5165.6			partial
N <sup>104</sup> KT	85–105	2400.1	2400.1	0	57–105	5559.9			not
N <sup>151</sup> MS	142–165	2761.2 <sup>b</sup>	2761.2	0	110–165	6533.1			100%
N <sup>172</sup> GS	166–185	2312.8	2313.1	0.3	166–185	2312.8	2312.8	0	not
N <sup>328</sup> AT	323–333	1175.6			312–333	2351.2 <sup>b</sup>	2351.2	0	100%
N <sup>337</sup> CT	337–353	1901.0 <sup>b</sup>	1901.0	0	337–372	4005.0 <sup>b</sup>	4004.9	–0.1	100%
N <sup>389</sup> RT	376–390	1788.0 <sup>b</sup>			376–407	3782.0 <sup>b</sup>	3781.9	–0.1	100%
N <sup>420</sup> IT	408–427	2140.2			408–430	2469.4 <sup>b</sup>	2469.3	–0.1	100%
N <sup>504</sup> VS	504–507	475.5			477–514	4447.9 <sup>b</sup>	4447.8	–0.1	100%
N <sup>544</sup> IT	524–550	3281.4 <sup>b</sup>	3281.4	0	515–569	6571.8 <sup>b</sup>	6571.7	–0.1	100%
N <sup>579</sup> NT	570–585	1777.8 <sup>b</sup>	1777.7	–0.1	570–585	1777.8 <sup>b</sup>	1778.1	0.3	partial
N <sup>599</sup> CT	586–618	3658.5 <sup>b</sup>	3658.8	0.3	586–618	1776.8	1776.8	0	
						3658.5 <sup>b</sup>	3658.7	0.2	100%

<sup>a</sup> All the masses are the monoisotopic masses of the peptides. <sup>b</sup> Post-deglycosylation masses of the peptides.

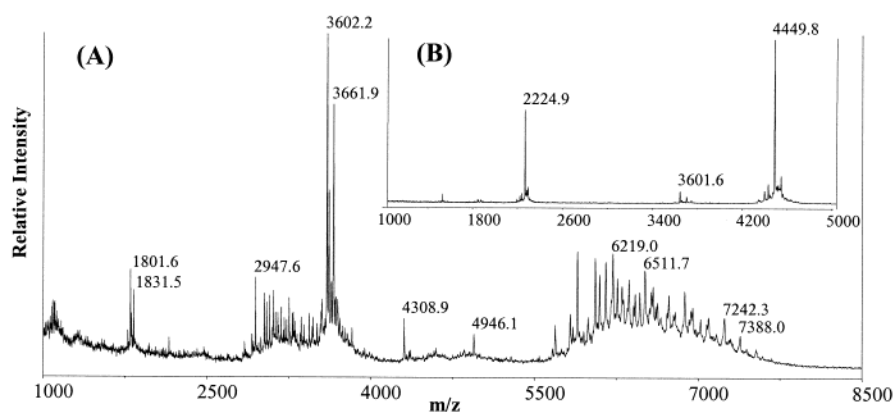


FIGURE 10: MALDI mass spectra of a glycopeptide spanning residues 477–514 from Lys-C digestion of the full-length EGF receptor before (A) and after (B) the treatment with PNGase F.

eventually resolved in the tryptic digest. The complementary data obtained from the Lys-C and trypsin digestions provide a complete picture of the glycosylation state of the full-length EGF receptor.

Isoforms of the glycopeptide spanning residues 477–514 from the full-length EGF receptor are shown in Figure 10, manifested as a broad set of peaks centered at  $m/z$  6600. The protein sequence was confirmed after deglycosylation and MS/MS analyses, which show that the glycosylation site N<sup>504</sup>VS in the full-length EGF receptor is completely glycosylated. Similar approaches were used to analyze all the other potential glycosylation sites, and the results are summarized in Table 2. As in the soluble extracellular domain, N<sup>104</sup>KT and N<sup>172</sup>GS in the full-length receptor were found to be nonglycosylated, and N<sup>579</sup>NT was partially glycosylated. All the other sites, except N<sup>32</sup>NC, were found to be completely glycosylated. The noncanonical glycosylation site, N<sup>32</sup>NC, was found to be glycosylated in the full-length receptor. However, in contrast to the results for the soluble extracellular domain study (Figure 2), the nonglycosylated fraction of this peptide spanning residues 30–48 was not observed. However, the full-length receptor was prepared in the presence of detergent, resulting in signal suppression that could have masked this peptide. On the basis of the observation that all of the other sites in the secreted and the full-length EGF receptor have identical glycosylation states, it is reasonable to suggest that N<sup>32</sup>NC may also be incompletely glycosylated in the full-length EGF receptor.

*Comparison of the Glycoforms in the Secreted and Full-Length EGF Receptor.* The different glycoforms for several glycosylation sites in both the secreted and the full-length EGF receptor have been resolved, allowing direct comparison. The data for glycopeptides containing N<sup>504</sup>VS are compared in Figure 11. The peptides from both the secreted and the full-length EGF receptor share some glycoforms, such as those with masses of  $m/z$  6219, 6511, 6876, and 7242, etc. However, the average mass of the glycoforms from the soluble domain is significantly greater than those from the full-length receptor. Furthermore, fewer glycoforms were observed in glycopeptides from the S-EGFR than in those from the full-length EGF receptor.

The glycoform from the S-EGFR with  $m/z$  7532.8 is likely to be a sialylated tetraantennary complex-type oligosaccharide with a fucose linked ( $\alpha$ 1–6) to the GlcNAc residue at the reducing end. Hydrolysis of the terminal sialic acids would produce glycans with  $m/z$  7241.6 and  $m/z$  6950.2 (Figure 12). Further hydrolysis by galactosidase and *N*-acetylhexosaminidase could result in a triantennary complex glycan with  $m/z$  6584.5 or a biantennary glycoform with  $m/z$  6219.6 (Figure 12). No intermediate oligosaccharides missing only a terminal galactose were observed. Other glycoforms present in the S-EGF receptor (Figure 11A) can be accounted for through loss of various monosaccharide residues from the above glycoforms.

Generally, more glycoforms were observed for the peptides from the full-length EGF receptor than for their counterparts from the S-EGFR, as exemplified by glycopeptides

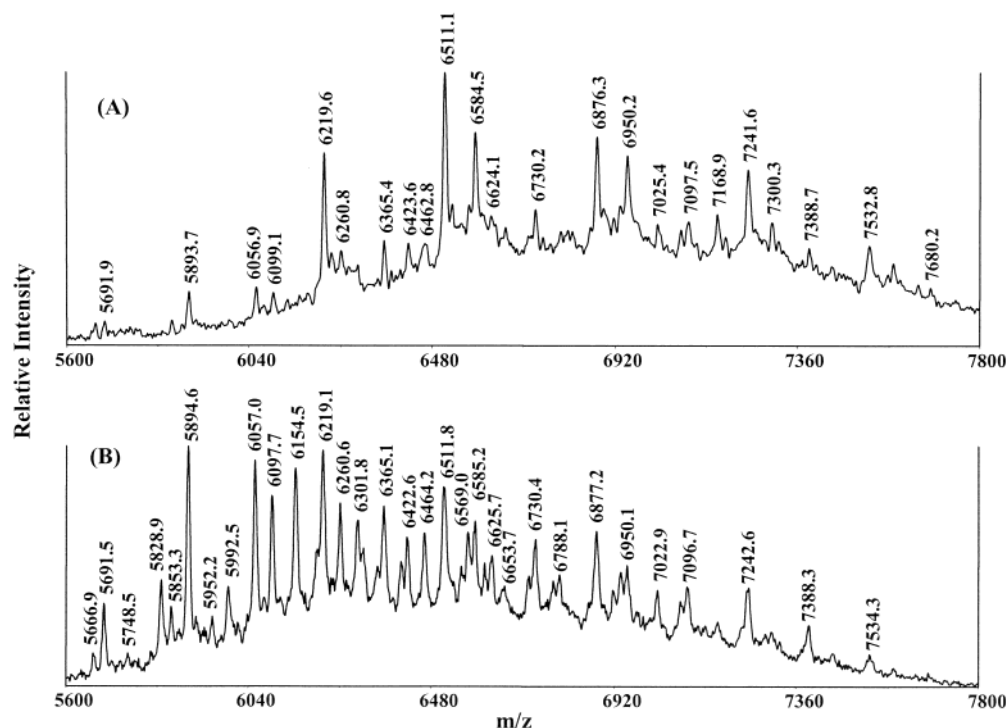
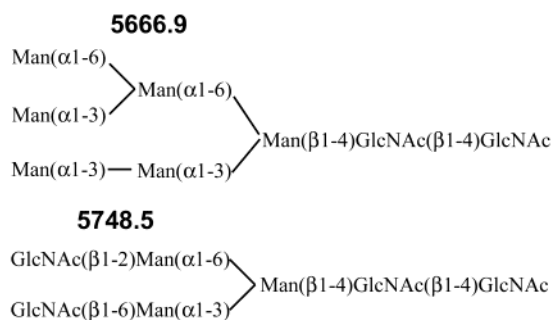


FIGURE 11: Comparison of the glycoforms in the MALDI mass spectra of a glycopeptide spanning residues 477–514 from S-EGFR (A) and the full-length EGF receptor (B).

spanning sequences 477–514 (Figure 11). Besides the glycoforms present in the secreted form (Figure 11A), the peptide with site N<sup>504</sup>VS from the full-length receptor displays other glycoforms (Figure 11B). The glycoforms observed for peptides from the protein purified using anti-EGF receptor antibody were essentially identical to those from protein purified using the EGF affi-gel method (data not shown), suggesting that the glycoforms observed were not affected by the purification methods employed. All of the glycoforms in the full-length EGF receptor can be categorized into three major groups, which are derived from the three smallest forms observed in Figure 11: *m/z* 5666.9, 5691.5, and 5748.5. The likely structure for the ion at *m/z* 5691.5 is the same as shown in Figure 12 above in the analysis of glycopeptides from the S-EGFR. A similar analysis suggests that ions at *m/z* 5666.9 and 5748.5 most likely have the following structures:



Possible compositions for all the other glycoforms present in the full-length EGF receptor were analyzed, and the results of this analysis are summarized in Table 3.

The group derived from the glycoform at *m/z* 5691.5 is present in both the secreted and the full-length receptors, suggesting that the two forms of receptor share similar glycan

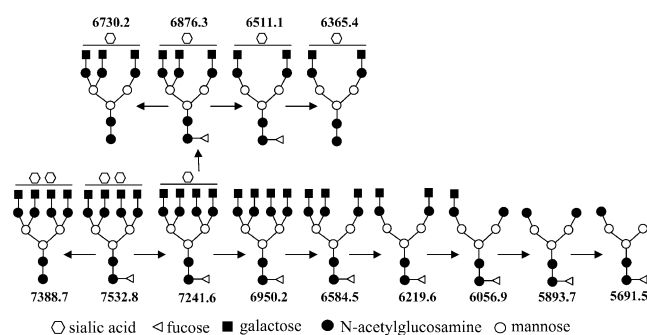


FIGURE 12: Processing of *N*-glycans at the N<sup>504</sup>VS site in S-EGFR.

structures. This group is comprised of primarily complex-type oligosaccharides with an (α1–6)-fucosylated core structure, with glycoforms in this group having di-, tri-, and tetraantennary structures. The majority of the glycans from the group derived from the glycoform with *m/z* 5748.5 are also predicted to be fucosylated complex type oligosaccharides, with three HexNAc residues in the terminal region. In contrast, the group derived from the glycoform at *m/z* 5666.9 is consistent with high-mannose type oligosaccharides, ranging from five to 12 mannose residues. The presence or absence of (α1–6)-fucose linked to the GlcNAc at the reducing end of the core structure further contributes to the structural variation of the glycans.

## DISCUSSION

In the present study, the glycosylation sites on both the secreted and the full-length, membrane-bound forms of the EGF receptor were characterized by mass spectrometry. A combination of three criteria were employed in identifying each glycosylation site: (1) the presence of peaks in the MALDI mass spectrum of an HPLC fraction with mass differences corresponding to the masses of monosaccharides;

Table 3: Suggested Compositions<sup>a</sup> for Different Glycoforms at Position Asn504 in Full-length EGF Receptor

[M + H] <sup>+</sup>	$\Delta$ mass <sup>b</sup>	composition
5748.5		(HexNAc) <sub>2</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>
5952.2	(203.7)	(HexNAc) <sub>3</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>
6097.7	(145.5)	(HexNAc) <sub>3</sub> (Fuc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>
6260.6	(162.7)	(Hex) <sub>1</sub> (HexNAc) <sub>3</sub> (Fuc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>
6422.6	(162)	(Hex) <sub>2</sub> (HexNAc) <sub>3</sub> (Fuc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>
6585.2	(162.6)	(Hex) <sub>3</sub> (HexNAc) <sub>3</sub> (Fuc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>
6730.4	(145.2)	(Hex) <sub>3</sub> (HexNAc) <sub>3</sub> (Fuc) <sub>2</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>
6877.2	(146.8)	(Hex) <sub>3</sub> (HexNAc) <sub>3</sub> (Fuc) <sub>1</sub> (NeuAc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>
7022.9	(145.7)	(Hex) <sub>3</sub> (HexNAc) <sub>3</sub> (Fuc) <sub>2</sub> (NeuAc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>
5666.9		(Hex) <sub>2</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>
5828.9	(162)	(Hex) <sub>3</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>
5992.5	(163.6)	(Hex) <sub>4</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>
6154.5	(162)	(Hex) <sub>5</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>
6301.8	(147.3)	(Hex) <sub>5</sub> (Fuc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub> <sup>#</sup>
6464.2	(162.4)	(Hex) <sub>6</sub> (Fuc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub> <sup>#</sup>
6625.7	(161.5)	(Hex) <sub>7</sub> (Fuc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub> <sup>#</sup>
6788.1	(162.4)	(Hex) <sub>8</sub> (Fuc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub> <sup>#</sup>
6950.1	(162)	(Hex) <sub>9</sub> (Fuc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub> <sup>#</sup>
7096.7	(146.6)	(Hex) <sub>9</sub> (Fuc) <sub>2</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub> <sup>#</sup>
5691.5		(Hex) <sub>1</sub> (Fuc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>
5894.6	(203.1)	(Hex) <sub>1</sub> (HexNAc) <sub>1</sub> (Fuc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>
6057.0	(162.4)	(Hex) <sub>2</sub> (HexNAc) <sub>1</sub> (Fuc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>
6219.1	(162.1)	(Hex) <sub>3</sub> (HexNAc) <sub>1</sub> (Fuc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>
6365.1	(146)	(Hex) <sub>3</sub> (HexNAc) <sub>1</sub> (Fuc) <sub>2</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>
6511.8	(146.7)	(Hex) <sub>3</sub> (HexNAc) <sub>1</sub> (Fuc) <sub>1</sub> (NeuAc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>
6877.2	(365.4)	(Hex) <sub>4</sub> (HexNAc) <sub>2</sub> (Fuc) <sub>1</sub> (NeuAc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>
7242.6	(365.4)	(Hex) <sub>5</sub> (HexNAc) <sub>3</sub> (Fuc) <sub>1</sub> (NeuAc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>
7388.3	(145.7)	(Hex) <sub>5</sub> (HexNAc) <sub>3</sub> (Fuc) <sub>2</sub> (NeuAc) <sub>1</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>
7534.3	(146)	(Hex) <sub>4</sub> (HexNAc) <sub>4</sub> (Fuc) <sub>1</sub> (NeuAc) <sub>2</sub> + (Man) <sub>3</sub> (GlcNAc) <sub>2</sub>

<sup>a</sup> Most of the glycoforms listed here may have more than one composition calculated mathematically. For those with more than one composition, only those present in GlycoSuiteDB are listed. Those compositions indicated by the # signs are not found in GlycoSuiteDB. <sup>b</sup>  $\Delta$  mass: The mass difference between this ion and the one immediately in front of it.

(2) the appearance of the deglycosylated peptide after the treatment of the fraction with PNGase F; and (3) the conversion by PNGase F of the Asn that had been the site of glycosylation to Asp. In an HPLC fraction of a larger protein, it is common for several peptides to be present in a particular fraction, and some peptides are likely to be suppressed on MALDI-MS analysis. Treatment of such an HPLC fraction with PNGase F, besides releasing the deglycosylated peptide, sometimes also resolves the suppression of other nonglycopeptides. Therefore, simply observing the appearance of the ion of a particular peptide in MALDI-MS after the treatment with PNGase F is not enough to conclude that the newly observed peptide is derived from a glycopeptide. Further, peptides and proteins are susceptible to spontaneous deamidation under physiological conditions, primarily at Asn, through a nonenzymatic reaction, especially when the Asn is followed by Gly (38). Therefore, drawing the conclusion that the Asn is an N-glycosylation site, based solely on the observation of a 1-Da mass shift of the peptide or on the replacement of Asn with Asp, is also not sufficient. Thus, fulfilling all three criteria is essential for the unambiguous identification of an N-glycosylation site.

The glycosylation pattern of the recombinant extracellular domain of the EGF receptor expressed in CHO cells has been characterized by Smith et al. (15) using a combination of chromatographic separation and protein sequence analysis. This study found that all of the 11 canonical N-glycosylation sites were glycosylated with most of them as complex-type chains, trisialylated tetraantennary oligosaccharides fucosylated on the reducing end. A second study by Sato et al. (16), also using the recombinant extracellular domain expressed

in CHO cells, showed that N<sup>328</sup>AT, N<sup>337</sup>CT, N<sup>544</sup>IT, and N<sup>579</sup>NT sites were glycosylated, but two sites, N<sup>172</sup>GS and N<sup>599</sup>CT, were not modified. Furthermore, they found an atypical N-glycosylation site at N<sup>32</sup>NC, which is partially glycosylated with a complex-type N-glycan. In the present study of the endogenously expressed EGF receptor in a human cell line, among the 11 canonical N-glycosylation sites, two sites, N<sup>104</sup>KT, N<sup>172</sup>GS, were found not to be glycosylated, and the N<sup>579</sup>NT site was partially glycosylated. The remaining eight canonical sites were completely glycosylated. This study confirmed that the N<sup>32</sup>NC site was partially glycosylated, and structural analysis of the oligosaccharides at the N<sup>32</sup>NC site reveals that the oligosaccharides are based on a core fucosylated biantennary complex-type glycan, with a structure indistinguishable from that found in the CHO cell study for the same site (16). Our results also agree with those of Sato et al. (16) for N<sup>172</sup>GS, which is not glycosylated in proteins expressed in either CHO or A431 cells. However, N<sup>104</sup>KT and N<sup>599</sup>CT appear to be modified differently in the two cell lines. In Table 4, our results are compared with those of Smith et al. (15) and Sato et al. (16).

The oligosaccharide moieties of the EGF receptor in A431 cells have previously been characterized using radioactively labeled monosaccharides, which showed that the mature full-length receptor contains both complex-type and high-mannose-type N-glycans in the approximate ratio of 2:1 (10). Many of the complex-type glycans were found to contain one or two sialic acid residues. The carbohydrate chains of the S-EGFR from A431 cells have also been thoroughly characterized using NMR and mass spectrometry (17). Over



Table 4: Comparison of EGF Receptor Glycosylation Studies

potential sites	CHO (15)	CHO (16)	human (this paper)
N <sup>32</sup> NC	nd <sup>a</sup>	partial	partial
N <sup>104</sup> KT	yes	nd	no
N <sup>151</sup> MS	yes	nd	yes
N <sup>172</sup> GS	yes	no	no
N <sup>328</sup> AT	yes	yes	yes
N <sup>337</sup> CT	yes	yes	yes
N <sup>389</sup> RT	yes	nd	yes
N <sup>420</sup> IT	yes	nd	yes
N <sup>504</sup> VS	yes	nd	yes
N <sup>544</sup> IT	yes	yes	yes
N <sup>579</sup> NT	yes	yes	partial
N <sup>599</sup> CT	nd	no	yes

<sup>a</sup> nd: Not determined.

30 structures have been revealed, and among them less than 20% are high-mannose types with di-, tri-, and tetra-antennary complex-type glycans accounting for the remainder (17). In the present study, among the 10 glycosylation sites observed, the carbohydrate structures of seven of them have been characterized. In the S-EGFR, two of these sites (N<sup>337</sup>CT and N<sup>599</sup>CT) were found to contain high-mannose-type glycans, and the sites at N<sup>32</sup>NC, N<sup>389</sup>RT, N<sup>420</sup>IT, N<sup>504</sup>VS, and N<sup>544</sup>IT were found to contain complex-type glycans. The ratio between the complex-type and the high-mannose-type glycans resolved in this study is in general agreement with the reported value. The glycan structures of the full-length EGF receptor appear more complicated. For example, we have shown that analysis of glycopeptides derived from the site at N<sup>504</sup>VS is consistent with a mixture of high-mannose and complex-type glycans.

In N-glycosylated proteins, glycans are bound to Asn in the consensus sequence Asn-X-Ser/Thr; however, the presence of this consensus sequence does not guarantee each site will be glycosylated. In S-EGFR and the full-length EGF receptor, two of the potential N-glycosylation sites (N<sup>104</sup>KT and N<sup>172</sup>GS) were not glycosylated, and the N<sup>32</sup>NC and N<sup>579</sup>NT sites were partially glycosylated. What factors determine whether a site is modified or not is still not clear. von Heijne et al. (39, 40) suggested that the glycosylation efficiency of a site depends on both the distance from the C terminus and the presence of a downstream transmembrane segment, with the glycosylation efficiency being reduced for those sites close to the C-terminus (39, 40). However, in both the S-EGFR and the full-length EGF receptor, the nonglycosylated sites (N<sup>104</sup>, N<sup>172</sup>) are quite close to the N-terminus, and the presence of the transmembrane domain in the full-length EGF receptor does not have a detectable effect on the glycosylation efficiency of N<sup>579</sup>NT site, which is partially glycosylated in both the S-EGFR and the full-length EGF receptor.

Two groups have recently reported structures determined by X-ray crystallography of recombinant ligand-binding domains of the EGF receptor in complex with EGF (41) or with TGF- $\alpha$  (42). In both structures, little of subdomain 4 was resolved, in the former case because of disorder in subdomain 4 (41) and in the latter case because only residues 1–501 were expressed (42). Both preparations had been treated with endoglycosidases prior to crystallization. Coordinates for the EGF–EGF receptor complex (Protein Data Bank accession code 1IVO) were publicly available at the

time of this writing. Examination of this structure reveals six sites of glycosylation resolved in one of the two receptor subunits of the receptor dimer. (Three of the sites are resolved in the second subunit of the dimer.) A single GlcNAc is linked to N<sup>32</sup>, N<sup>151</sup>, N<sup>172</sup>, N<sup>328</sup>, and N<sup>337</sup>, and a pair of GlcNAc's is linked to N<sup>420</sup>. All of these are consistent with our results, except for N<sup>172</sup>, which is glycosylated in the construct expressed in Lec8 cells (41), a CHO cell-derived line (43), but is not glycosylated in either the full-length receptor or the secreted form naturally expressed in A431 cells, emphasizing the cell-type dependence of glycosylation site usage. Further, the structures shed no light on why N<sup>104</sup> and N<sup>172</sup> are not glycosylated in A431 cells. Both sites are located on the surface of the receptor, and as noted above, N<sup>172</sup> is glycosylated in Lec8 cells.

In the present study, the glycoforms for the same glycosylation sites in both the secreted and the full-length receptors naturally expressed in A431 cells have been resolved. While one must be cautious about extrapolating the results from these transformed cells to normal human cells that express much lower levels of receptor, in A431 cells both the S-EGFR and the full-length EGF receptor have almost identical glycosylation patterns; however, the oligosaccharide moieties of the full-length receptor were found to be more extensively processed than those of the secreted form. A detailed analysis of the site at N<sup>504</sup>VS is reported in this study. Similar results were also observed for the site at N<sup>544</sup>IT (data not shown). Although the site at N<sup>504</sup>VS was completely glycosylated in both the secreted and the full-length receptors, this site in the two proteins displayed different glycoforms. In the full-length receptor, over 30 glycoforms were observed with roughly similar intensities in the MALDI mass spectrum; in contrast, only about 20 glycoforms were observed in the secreted protein, and among the glycoforms in the secreted form, a few displayed as the most abundant forms. The 30 glycoforms in the spectrum of glycopeptides from the full-length receptor suggest a mixture of high-mannose and complex-type glycans, while most of the glycoforms identified in the corresponding spectrum from the secreted protein are the complex type. An early statistical study by Pollack and Atkinson (44) found that transmembrane proteins tend to have both complex- and high-mannose-type glycans, while most secreted proteins have only complex-type glycans. This general rule appears to apply to the EGF receptor.

Oligosaccharide processing enzyme inhibitors have been applied to studies of EGF receptor biosynthesis and function. EGF receptor synthesized in the presence of tunicamycin, an inhibitor of N-linked glycosylation, lacks both ligand-binding and kinase activities (10). Further, when the synthesis of oligosaccharides was inhibited in A431 cells, the EGF receptor was found to be misfolded and retained in the ER lumen (12, 13). Swainsonine inhibits Golgi  $\alpha$ -mannosidase II and thus prevents the trimming of mannose residues and conversion of high-mannose-type glycans into complex-type (45). Thus, swainsonine-treated cells have an EGF receptor with mainly high-mannose-type and hybrid-type glycans (34). Functional studies on the EGF receptor derived from the swainsonine-treated cells showed that ligand binding and tyrosine kinase activities were relatively unaffected, suggesting that the complete processing of the glycans is not essential to the function of the receptor. When *N*-acetylglucosaminyltransferase III, an enzyme that inhibits the

extension of the *N*-glycan by introducing a bisecting *N*-acetylglucosamine residue, was overexpressed in cells, the EGF receptor displayed reduced binding affinity for EGF, suggesting the involvement of oligosaccharides on substrate binding (46). In a study by Tsuda and colleagues (14), a series of mutants have been generated by mutating the glycosylation sites N<sup>328</sup>, N<sup>337</sup>, N<sup>389</sup>, and N<sup>420</sup> to glutamine residues. Among the four mutants, the elimination of the oligosaccharides on N<sup>328</sup>, N<sup>337</sup>, and N<sup>389</sup> had no reported effect on EGF binding and EGF-stimulated kinase activity; on the other hand, the removal of the glycans at the N<sup>420</sup> site in the EGF receptor completely abolished EGF binding, and the receptor formed a spontaneous dimer with constitutively activated kinase activity in the absence of EGF. This study emphasizes the essential role that oligosaccharides can play in substrate binding and receptor–receptor interactions (14).

In this context, it is interesting to speculate on possible functional consequences of the partial glycosylation of N<sup>579</sup> documented here. N<sup>579</sup> is in subdomain 4, the second Cys-rich subdomain of the extracellular ligand-binding domain of the receptor (47). Saxon et al. (48) showed that deletion of residues 518–589 in subdomain 4 resulted in a receptor that was expressed on the cell surface but had greatly diminished affinity for EGF. Binding studies of the recombinant extracellular ligand-binding domain of the EGF receptor with (49) and without (50) subdomain 4 have revealed that the construct lacking subdomain 4 has much higher, more nativelike affinity for ligand than does the construct that includes subdomain 4 (49, 50). These observations suggest that subdomain 4 can regulate the affinity of the receptor for ligand. Studies are currently under way designed to test whether glycosylation at N<sup>579</sup> modulates the affinity of the EGF receptor for ligand.

In summary, among the 11 canonical *N*-glycosylation sites in human EGF receptor, two sites were not modified, one of them was partially glycosylated, and the remainder were fully glycosylated. The presence of an atypical, partially glycosylated *N*-glycosylation site at N<sup>32</sup>NC was also confirmed. The glycosylation states in both the secreted form and the full-length membrane-bound EGF receptor from A431 cells were compared, and the occupancies of the glycosylation sites in two proteins were indistinguishable. However, the oligosaccharides at the glycosylation sites in the full-length EGF receptor were more extensively processed than those at the same sites in the secreted form.

## REFERENCES

- Helenius, A., and Aebi, M. (2001) *Science* 291, 2364–2369.
- Varki, A. (1993) *Glycobiology* 3, 97–130.
- Rudd, P. M., Elliott, T., Cresswell, P., Wilson, I. A., and Dwek, R. A. (2001) *Science* 291, 2370–2376.
- Kim, Y. J., and Varki, A. (1997) *Glycoconj. J.* 14, 569–576.
- Ullrich, A., Coussens, L., Hayflick, J. S., Dull, T. J., Gray, A., Tam, A. W., Lee, J., Yarden, Y., Libermann, T. A., Schlessinger, J., Downward, J., Mayes, E. L. V., Whittle, N., Waterfield, M. D., and Seeburg, P. H. (1984) *Nature* 309, 418–425.
- Ullrich, A., and Schlessinger, J. (1990) *Cell* 61, 203–212.
- Downward, J., Parker, P., and Waterfield, M. D. (1984) *Nature* 311, 483–485.
- Margolis, B. L., Lax, I., Kris, R., Dombalagian, M., Honegger, A. M., Howk, R., Givol, D., Ullrich, A., and Schlessinger, J. (1989) *J. Biol. Chem.* 264, 10667–10671.
- Carpenter, G. (2000) *Bioessays* 22, 697–707.
- Cummings, R. D., Soderquist, A. M., Carpenter, G. (1985) *J. Biol. Chem.* 260, 11944–11952.
- Soderquist, A. M., and Carpenter, G. (1984) *J. Biol. Chem.* 259, 12586–12594.
- Gamou, S., and Shimizu, N. (1988) *J. Biochem.* 104, 388–396.
- Gamou, S., Shimagaki, M., Minoshima, S., Kobayashi, S., and Shimizu, N. (1989) *Exp. Cell Res.* 183, 197–206.
- Tsuda, T., Ikeda, Y., and Taniguchi, N. (2000) *J. Biol. Chem.* 275, 21988–21994.
- Smith, K. D., Davies, M. J., Bailey, D., Renouf, D. V., and Hounsell, E. F. (1996) *Growth Factors* 13, 121–132.
- Sato, C., Kim, J. H., Abe, Y., Saito, K., Yokoyama, S., and Kohda, D. (2000) *J. Biochem. (Tokyo)* 127, 65–72.
- Stroop, C. J., Weber, W., Gerwig, G. J., Nimtz, M., Kamerling, J. P., and Vliegthart, J. F. (2000) *Glycobiology* 10, 901–917.
- Stoscheck, C. M., and Carpenter, G. (1983) *J. Cell Biochem.* 23, 191–202.
- Reiter, J. L., Threadgill, D. W., Eley, G. D., Strunk, K. E., Danielsen, A. J., Sinclair, C. S., Pearsall, R. S., Green, P. J., Yee, D., Lampland, A. L., Balasubramaniam, S., Crossley, T. D., Magnuson, T. R., James, C. D., and Mailhe, N. J. (2001) *Genomics* 71, 1–20.
- Gunther, N., Betzel, C., and Weber, W. (1990) *J. Biol. Chem.* 265, 22082–22085.
- Weber, W., Wensch, E., Gunther, N., Marnitz, U., Betzel, C., and Righetti, P. G. (1994) *J. Chromatogr. A* 679, 181–189.
- Scott, G. K., Robles, R., Park, J. W., Montgomery, P. A., Daniel, J., Holmes, W. E., Lee, J., Keller, G. A., Li, W. L., and Fendly, B. M. (1993) *Mol. Cell. Biol.* 13, 2247–2257.
- Lee, H., and Mailhe, N. J. (1998) *Oncogene*, 16(25), 3243–3252.
- Zhou, W., and Carpenter, G. (2000) *J. Biol. Chem.* 275, 34737–34743.
- DiStefano, P. S., and Johnson, E. M., Jr. (1988) *Proc. Natl. Acad. Sci. U.S.A.* 85, 270–274.
- Leung, D. W., Spencer, S. A., Cachianes, G., Hammonds, R. G., Collins, C., Henzel, W. J., Barnard, R., Waters, M. J., and Wood, W. I. (1987) *Nature* 330, 537–543.
- Rubin, L. A., Kurman, C. C., Fritz, M. E., Biddison, W. E., Boutin, B., Yarchoan, R., and Nelson, D. L. (1985) *J. Immunol.* 135, 3172–3177.
- MacDonald, R. G., Tepper, M. A., Clairmont, K. B., Perregaux, S. B., and Czech, M. P. (1989) *J. Biol. Chem.* 264, 3256–3261.
- Flickinger, T. W., Mailhe, N. J., and Kung, H. J. (1992) *Mol. Cell Biol.* 12, 883–893.
- Cohen, S., Ushiro, H., Stoscheck, C., and Chinkers, M. (1982) *J. Biol. Chem.* 257, 1523–1531.
- Cohen, S., Fava, R. A., and Sawyer, S. T. (1982) *Proc. Natl. Acad. Sci. U.S.A.* 79, 6237–6241.
- Cooper, C. A., Gasteiger, E., and Packer, N. (2001) *Proteomics* 1, 340–349.
- Cooper, C. A., Harrison, M. J., Wilkins, M. R., and Packer, N. H. (2001) *Nucleic Acids Res.* 29, 332–335.
- Stroop, C. J., Weber, W., Nimtz, M., Gallego, R. G., Kamerling, J. P., and Vliegthart, J. F. (2000) *Arch. Biochem. Biophys.* 374, 42–51.
- Juhasz, P., and Costello, C. E., (1992) *J. Am. Soc. Mass Spectrom.* 3, 785–796.
- Sugiyama, E., Hara, A., Uemura, K., and Taketomi, T. (1997) *Glycobiology* 7, 719–724.
- Harvey, D. J. (2001) *Proteomics* 1, 311–328.
- Brennan, T. V., and Clarke, S. (1995) in *Deamidation and Isoaspartate Formation in peptides and Proteins* (Aswad, D. W., Ed.) pp 65–90, CRC Press, Boca Raton, FL.
- Gavel, Y., and von Heijne, G. (1990) *Protein Eng.* 3, 433–442.
- Nilsson, I., and von Heijne, G. (2000) *J. Biol. Chem.* 275, 17338–17343.
- Ogisso, H., Ishitani, R., Nureki, O., Fukai, S., Yamanaka, M., Kim, J.-H., Saito, K., Sakamoto, A., Inoue, M., Shirouzu, M., and Yokoyama, S. (2002) *Cell* 110, 775–787.
- Garrett, T. P. J., McKern, N. M., Lou, M., Elleman, T. C., Adams, T. E., Lovrecz, G. O., Zhu, H.-J., Walker, F., Frenkel, M. J., Hoyne, P. A., Jorissen, R. N., Nice, E. C., Burgess, A. W., and Ward, C. W. (2002) *Cell* 110, 763–773.
- Stanley, P. (1981) *Mol. Cell Biol.* 8, 687–696.

44. Pollack, L., and Atkinson, P. H. (1983) *J. Cell Biol.* 97, 293–300.
45. Kornfeld, R., and Kornfeld, S. (1985) *Annu. Rev. Biochem.* 54, 631–664.
46. Rebbaa, A., Yamamoto, H., Saito, T., Meuillet, E., Kim, P., Kersey, D. S., Bremer, E. G., Taniguchi, N., and Moskal, J. R. (1997) *J. Biol. Chem.* 272, 9275–9279.
47. Lax, I., Burgess, W., Bellot, F., Ullrich, A., and Schlessinger, J. (1988) *Mol. Cell. Biol.* 8, 1831–1834.
48. Saxon, M. L., and Lee, D. C. (1999) *J. Biol. Chem.* 274, 28356–28362.
49. Domagala, T., Konstantopoulos, N., Smyth, F., Jorissen, R. N., Fabri, L., Geleick, D., Lax, I., Schlessinger, J., Sawyer, W., Howlett, G. J., Burgess, A. W., and Nice, E. C. (2000) *Growth Factors* 18, 11–29.
50. Elleman, T. C., Domagala, T., McKern, N. M., Nerrie, M., Lönnqvist, B., Adams, T. E., Lewis, J., Lovrecz, G. O., Hoyne, P. A., Richards, K. M., Howlett, G. J., Rothacker, J., Jorissen, R. N., Lou, M., Garrett, T. P. J., Burgess, A. W., Nice, E. C., and Ward, C. W. (2001) *Biochemistry* 40, 8930–8939.

BI027101P