Identification of N-Linked Glycosylation Sites in Human Testis Angiotensin-converting Enzyme and Expression of an Active Deglycosylated Form*

(Received for publication, September 16, 1996, and in revised form, December 3, 1996)

X. Christopher Yu‡, Edward D. Sturrock‡‡, Zucchini Wu‡, Klaus Biemann§, Mario R. W. Ehlers¶, and James F. Riordan‡‡‡

From the ‡Center for Biochemical and Biophysical Sciences and Medicine, Harvard Medical School, Boston, Massachusetts 02115, the ¶Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, and the §Department of Medical Biochemistry, University of Cape Town Medical School, Observatory 7925, South Africa

© 1997 by The American Society for Biochemistry and Molecular Biology, Inc.

The sites of glycosylation of Chinese hamster ovary cell expressed testicular angiotensin-converting enzyme (tACE) have been determined by matrix-assisted laser desorption ionization/time of flight/mass spectrometry of peptides generated by proteolytic and cyanogen bromide digestion. Two of the seven potential N-linked glycosylation sites, Asn650 and Asn1099, were found to be fully glycosylated by analysis of peptides before and after treatment with a series of glycosidases and with endoproteinase Asp-N. The mass spectra of the glycopeptides exhibit characteristic clusters of peaks which indicate the N-linked glycans in tACE to be mostly of the biantennary, fucosylated complex type. This structural information was used to demonstrate that three other sites, Asn155, Asn337, and Asn586, are partially glycosylated, whereas Asn72 appears to be fully glycosylated. The only potential site that was not modified is Asn620. Sequence analysis of tryptic peptides obtained from somatic ACE (human kidney) identified six glycosylated and one unglycosylated Asn. Only one of these glycosylation sites had a counterpart in tACE. Comparison of the two proteins reveals a pattern in which amino-terminal N-linked sites are preferred. The functional significance of glycosylation was examined with a TACE mutant lacking the O-glycan-rich first amino-terminal 36 residues and truncated at Ser625. When expressed in the presence of the a-glucosidase I inhibitor N-butyldeoxynojirimycin and treated with endoglycosidase H to remove all but the terminal N-acetylgalactosamine residues, it retained full enzymatic activity, was electrophoretically homogeneous, and is a good candidate for crystallographic studies.

Both forms of angiotensin-converting enzyme (ACE1; EC 3.4.15.1 peptidyl-dipeptidase A) are class I transmembrane ectoenzymes (1) that have N- and O-linked oligosaccharides attached to their polypeptide chains (2, 3). Expression of ACE in human HeLa cells in the presence of tunicamycin resulted in complete inhibition of glycosylation, rapidly degraded intracellular ACE, and no enzyme released in the medium (4). An enzymatically active ACE was produced with partial glycosylation in a mutant Chinese hamster ovary (CHO) cell line (ddID), although it was released to a lesser extent (4). Similarly, it was reported (5) that inhibitors of glucosidase I and II in the endoplasmic reticulum (ER) and mannose 1 in the cis-Golgi reduced the amount of oligosaccharide attached to human intestinal ACE and delayed protein release significantly. These data strongly suggest that glycosylation plays an important role in the membrane targeting and release of ACE, possibly by affecting the folding of the polypeptide and its recognition by a variety of enzymes in the folding and transport machineries. Recently, Sadhukhan and Sen (6) reported that mutations at individual N-linked glycosylation sites (sequons) in rabbit testis ACE (tACE) resulted in varied efficiencies in enzyme release, which suggests that N-linked glycans at each site may make different contributions to ACE transport and release.

Comparison of the cDNA sequences of ACE in human, rabbit, and mouse further supports such a role for glycosylation in ACE. Five of the seven potential N-linked glycosylation sequons in human tACE have counterparts in rabbit, and a sixth sequon is also present in mouse. It is not known whether the two additional N-linked glycosylation sequons in human tACE are utilized, although studies have shown heterogeneity across different species, both in the sites of oligosaccharide attachment and the types of carbohydrate components (3). TACE from all three species contains a serine/threonine-rich NH2-terminal motif that is heavily glycosylated, although no apparent function for the O-glycosylation has been demonstrated (4, 7).

Information on ACE active site residues and structures is based largely on homology between ACE and other zinc metalloenzymes, and attempts to crystallize ACE have not been successful. It is thought that partial or complete removal of the carbohydrate might facilitate the crystallization and structural studies of ACE. Expression of rabbit TACE in Escherichia coli resulted in a carbohydrate-free form of the protein, but it was devoid of any enzyme activity (6). Partially glycosylated ACE proteins generated by transient expression in human HeLa cells and in yeast were found to be enzymatically active but left open the question of whether glycosylation affects in any way...
the specific activities of ACE in vitro.

In this study, we have identified the N-linked glycosylation sites in human tACE expressed in CHO cells by a combination of enzymatic digestion and chemical cleavage of the protein followed by mass spectrometry. Four of the five conserved sequons are glycosylated, and a fifth is likely glycosylated as well. A sequon that is present in human and mouse but not in rabbit ACE is partially glycosylated. In addition, we provide evidence that a chemically homogeneous form of tACE can be prepared by inhibition of complex N-linked glycosylation and enzymatic removal of the high mannose oligosaccharides. Kinetic analysis indicates that the enzyme is fully active in vitro, suggesting that it is a good candidate for crystallographic studies.

Our results further support the hypothesis that glycosylation plays a critical role in the folding of ACE and that the effects on transport and enzyme release may be site-dependent.

EXPERIMENTAL PROCEDURES

Materials—Endoproteinase Lys-C and Asp-N, peptide N-glycosidase F, endoglycosidase H, neuraminidase, and O-glycosidase were purchased from Boehringer Mannheim. Cytochrome c (Sigma), trypsin, and chymotrypsin (Calbiochem-Behring). Bovine serum albumin (Fraction V, Baker). Dialysis tubing (1000–5000 molecular weight cutoff), dialysis bags (6000 molecular weight cutoff), and polyethyleneimine (Sigma). A nitrogen laser (337 nm) was used for desorption ionization. Measurements were carried out either in the linear or reflector mode with mass accuracies of 0.1 and 0.01%, respectively. Spectra were collected on 100 laser shots. Typical matrices used in these experiments were 3,5-dimethoxy-4-hydroxyacetophenone (Sigma, Aldrich). About 1 μl of sample solution was mixed with 2 μl of the matrix solution (10 mg/ml in 50% v/v CH 3CN and H 2O). A 0.5-μl volume (containing 1–10 pmol of peptide or peptide mixture) of the above solution was loaded on the sample plate and allowed to dry. All mass values reported are isotopically averaged masses.

RESULTS

Determination of N-Linked Glycosylation Sites—Purified human tACE (wild-type minus COOH-terminal residues 628–701) was digested with endoproteinase Lys-C, and the resulting peptide fragments were resolved by HPLC (Fig. 1) and analyzed by MALDI/TOF/MS. A separate digestion was carried out with CNBr. As shown in Table I, about 75% of the entire sequence of the glycosylated protein could be mapped from the two sets of peptides, and of the seven potential sites, two sets of peptides, and of the seven potential

FIG. 1. HPLC chromatogram of a Lys-C digest of wild-type tACE. The peptide mixture was injected onto a C18 Delta-Pak column (Waters) and developed with a linear gradient of 8.5–51% acetonitrile in 0.1% trifluoroacetic acid over 90 min. Fractions designated by an asterisk provided no assignable MALDI/TOF/MS signals (did not match the mass of any peptide) and were thus digested further with glycosidases and analyzed as described under “Experimental Procedures.”

Protein Digestions—Generally 200 μl of endoproteinase Lys-C (0.1 mg/ml in H 2O) was added to 0.5 mg of tACE in 100 mM ammonium bicarbonate, pH 8.5, and the digestion was allowed to proceed for 16 h at 37°C. For CNBr digestion, purified tACE (4 nmol) was lyophilized and dissolved in 70% trifluoroacetic acid (1 ml). CNBr (40 mg) was added and the reaction mixture incubated at room temperature for 4 h. The digestion was stopped by the addition of ice-cold water (1 ml) and kept on ice for 1 h before lyophilization. The dried sample was then dissolved in 0.1 M ammonium bicarbonate and subjected to reversed phase HPLC on a C18 column as described below.

Purified human kidney ACE was digested with trypsin, fractionated by HPLC, and the peptides were sequenced by automated Edman degradation as described previously (13).

ACE Peptide Deglycosylation—Lyophilized fractions containing glycosylated peptides (2–3 nmol) were dissolved in 200 μl of 20 mM sodium phosphate, pH 7.2. One aliquot was designated as a control, and 150 μl was digested with 50 μl of a glycosidase mixture containing neuraminidase (5 milliunits), O-glycosidase (2.5 milliunits), and peptide N-glyco-
sidase F (0.4 milliunit) at 25°C for 24 h.

Peptide Separation—Mixtures of peptides (200 μg in 50–100 μl) were resolved by reversed phase HPLC using either a C18 Delta-Pak column, 5 μm, 4.6 × 250 mm (Waters), or a C3, Microsorb-MV column, 5 μm, 4.6 × 250 mm (Rainin), and eluted with a 10–60% gradient of acetonitrile in a flow rate of 1 ml/min. The UV absorbance was monitored at 214 nm.

Purified ACE—Purified ACE was obtained on a MALDI/TOF/MS instrument (Voyager-Elite Biowavepectra System, PerSeptive Biosystems, Inc.). A nitrogen laser (337 nm) was used for desorption ionization. Measurements were carried out either in the linear or reflector mode with mass accuracies of 0.1 and 0.01%, respectively. Spectra were collected on 100 laser shots. Typical matrices used in these experiments were 3,5-dimethoxy-4-hydroxyacetophenone (Sigma, Aldrich). About 1 μl of sample solution was mixed with 2 μl of the matrix solution (10 mg/ml in 50% v/v CH 3CN and H 2O). A 0.5-μl volume (containing 1–10 pmol of peptide or peptide mixture) of the above solution was loaded on the sample plate and allowed to dry. All mass values reported are isotopically averaged masses.

RESULTS

Determination of N-Linked Glycosylation Sites—Purified human tACE (wild-type minus COOH-terminal residues 628–701) was digested with endoproteinase Lys-C, and the resulting peptide fragments were resolved by HPLC (Fig. 1) and analyzed by MALDI/TOF/MS. A separate digestion was carried out with CNBr. As shown in Table I, about 75% of the entire sequence of the glycosylated protein could be mapped from the two sets of peptides, and of the seven potential N-linked glycosylation sites, four (Asn155, Asn327, Asn386, and Asn620) were found unglycosylated. The CNBr fragment Leu593–Arg627 is thought to be the COOH terminus of the soluble form of ACE.
from CHO cells, as discussed in more detail elsewhere (8).

All of the peptides that make up the remaining 25% of the protein sequence not observed prior to deglycosylation contain potential N- and/or O-linked glycosylation sites. To determine unambiguously the glycosylation states of Asn residues in all seven sequons, the oligosaccharides were removed by treatment with a series of glycosidases. HPLC fractions from the CNBr and Lys-C digests that did not contain identifiable peptides (Fig. 1) were treated with a mixture of peptide CNBr and Lys-C digests that did not contain identifiable peptides, Asn90

Table I

| Peptide | Fraction | Calculated m/z | Observed m/z |
|---------|----------|----------------|--------------|
| Lys-C   |          |                |              |
| 80–84   | 14       | 614.8          | 614.5        |
| 115–137 | 10       | 2,345.6        | 2,344.5      |
| 175–188 | 22       | 1,688.1        | 1,688.7      |
| 189–307 | 43       | 13,716.1       | 13,713.8     |
| 308–317 | 8        | 1,307.6        | 1,307.1      |
| 318–338 | 28       | 2,423.7        | 2,424.8      |
| 359–363 | 12       | 2,842.2        | 2,841.8      |
| 364–368 | 5        | 678.8          | 678.7        |
| 369–395 | 18       | 3,236.7        | 3,235.7      |
| 396–425 | 24       | 3,075.5        | 3,074.8      |
| 426–449 | 34       | 2,695.0        | 2,693.3      |
| 455–478 | 30       | 2,946.4        | 2,946.4      |
| 479–491 | 15       | 1,766.9        | 1,766.6      |
| 492–511 | 9        | 2,147.4        | 2,146.7      |
| 512–556 | 36       | 5,224.0        | 5,223.4      |
| 557–567 | 2        | 1,176.4        | 1,176.3      |
| 568–613 | 37       | 5,293.1        | 5,293.2      |
| 614–627 | 14       | 1,691.0        | 1,690.3      |

| CNBr |          |                |              |
|------|----------|----------------|--------------|
| 143–169 |         | 2,874.2        | 2,874.0      |
| 170–223 |         | 6,320.1        | 6,319.5      |
| 224–278 |         | 6,430.3        | 6,430.5      |
| 300–315 |         | 1,674.1        | 1,675.3      |
| 393–450 |         | 6,254.0        | 6,254.7      |
| 341–385 |         | 5,154.8        | 5,154.8      |
| 386–399 |         | 849.0          | 849.0        |
| 451–566 |         | 13,437.4       | 13,450       |
| 567–578 |         | 1,371.6        | 1,372.5      |
| 593–627 |         | 4,262.8        | 4,262.0      |

a Numbered according to the sequence of the wild-type protein (13).
b Bold characters indicate peptides that contain potential N-glycosylation sites.
c Calculated for COOH-terminal homoserine lactone.

To confirm these findings, the original, glycosylated CNBr-peptide Gln87-Hse (Met142) was treated with Asp-N, the digests were separated by HPLC, and the resulting fractions were analyzed by MALDI/TOF/MS. The spectra of two fractions (Figs. 4A and 5A, respectively) exhibited typical glyco-patterns (clusters of peaks separated by 162, 203, and 291 Da, corresponding to the addition of hexose, N-acetylhexosamine, and sialic acid residues, respectively) indicating the presence of heterogeneous N-linked glycans. After treatment with glycosidases, these multiplets converged to a single dominant peak (Figs. 4B and Fig. 5B). They were identified as peptide Asp103-Gln120 (Figs. 4B, expected m/z 2141.5 and Gln87-Phe102 (Fig. 5B, expected m/z 1878.2). Thus both Asn90 and Asn109 are glycosylated in the wild-type tACE.

As shown in Fig. 4A, the different molecular ions observed in the glycosylated peptide Asp103-Gln120 demonstrate the heterogeneous nature of the sugars attached to Asn109. The mass difference between the molecular ion at m/z 3910.1 and m/z 2141.5 (calculated for the deglycosylated peptide) is 1768.6, consistent with the calculated value of 1769.6 for the increment in molecular mass due to the addition of Hex5HexNAc4Deoxy-Hex1 to Asn109. These data indicate that the N-linked glycans in tACE are mostly of the biantennary, fucosylated complex biantennary, fucosylated complex.
type (Fig. 6). The addition of either one or two sialic acids at the two termini would produce molecular ions at \( m/z \) 4201.4 (observed at \( m/z \) 4201.4) and 4492.7 (observed at \( m/z \) 4492.6). The presence of shorter glycans was also observed in the spectrum shown in Fig. 4A. The likely compositions of the various carbohydrate moieties attached to Asn109 are listed in the inset.

The glycan structures on Asn90 as deduced from the mass spectrum are listed in Fig. 5A. The mass difference between the molecular ion at \( m/z \) 3646.9 and the deglycosylated calculated peptide mass at 1878.2 is 1768.7, again consistent with the calculated mass of 1769.6 for the above oligosaccharide increment. Masses corresponding to the addition of one and two sialic acids, respectively, as well as shorter glycans are apparent in the spectrum (Fig. 5A).

Three residues, Asn155, Asn337, and Asn586, are present in both glycosylated and unglycosylated forms. As mentioned earlier, peptides with these three sequons were observed in their unglycosylated forms (see Table I). Based on the molecular mass of the observed major oligosaccharide structure (1769.6), it was possible to identify them in their glycosylated forms as well (Table II). For Asn155, the same molecular ions were observed before and after incubation with the glycosidases, which may be due to the presence of a disulfide bond between Cys152 and Cys158 (see discussion below) which may prevent enzymatic deglycosylation. For Asn586, a molecular ion at \( m/z \) 5309.1 was observed after glycosidase treatment, which is 16 Da higher than that expected for deglycosylated peptide Leu568-Lys613 (calculated \( m/z \) 5294.1). This is most likely due to oxidation of one or more of the three Met residues present in that peptide (peaks 32 and 48 Da higher and of decreasing signal intensity are also present).

As listed in Table II, molecular ions were also observed which suggest the glycosylation of Asn72. This is based on the assumption that the same type of oligosaccharide structure is attached to this site but lacks the fucose moiety. However, the amount of peptide recovered after treatment with deglycosidases was insufficient for positive identification.

Identification of N-Linked Glycosylation Sites in Purified Human Kidney ACE—Automated peptide sequencing was applied to tryptic fragments of somatic ACE isolated from human
A total of 28 tryptic peptides was identified (Table III), among which six contained glycosylated Asn residues (residues 9, 25, 82, 117, 480, and 913), and one contained an unglycosylated Asn (residue 1196). There are 17 N-linked glycosylation sequons in human somatic ACE, and the states of glycosylation of the remaining 10 potential sites were not determined in this study.

Effects of NB-DNJ and Treatment of ACEΔ36NJ with Endo H—The deletion mutant tACEΔ36NJ lacks the first 36 NH2-terminal residues of the mature protein (7); it is also truncated after Ser125 to encode a soluble protein that lacks part of the juxtamembrane stalk and the transmembrane and cytoplasmic domains (these modifications were introduced to facilitate later crystallization attempts). The glucosidase I inhibitor NB-DNJ prevents maturation of N-linked oligosaccharides of recombinant protein expressed in CHO cells (14). These sugars remain as oligomannose forms that are cleaved with endoglycosidase H

**Fig. 4.** Panel A, MALDI/TOF/MS of an HPLC fraction from the Asp-N digest of CNBr-glycopeptide Gln89-Hse (Met145). The multiple peaks from m/z 3000 to 4800 show the glycoform distribution. Panel B, MALDI/TOF/MS of the above fraction after treatment with a mixture of glycosidases. The peak at m/z 2142.1 corresponds to the protonated peptide Asp103-Gln120.
under nondenaturing conditions to leave single N-acetylglucosamine residues (12).

After expression in the presence of NB-DNJ, tACEΔ36NJ migrated as a sharp band on SDS-polyacrylamide gel electrophoresis (Fig. 7) consistent with increased homogeneity of its N-linked oligosaccharides. Digestion of this protein with endoglycosidase H produced an electrophoretically homogeneous product at 68 kDa, in agreement with mass spectrometric analysis that gave a [M+H]^+ ion at m/z 68,924 (expected m/z 69,008, with five GlcNAc/mol of protein). The glycosylated protein (tACEΔ36NJ) from cells not treated with NB-DNJ was found to be present as a multiplet at m/z 74,136. The different

![Image](image-url)
glycoforms of the mutant protein as well as those of the wild-type tACE migrated as broad diffuse bands on SDS-polyacrylamide gel electrophoresis with molecular masses of approximately 75 and 84 kDa, respectively (Fig. 7).

Catalytic Properties—tACE36NJ retained its enzyme activity after deglycosylation with endoglycosidase H. The specific activity of the glycosylated protein, in terms of the hydrolysis of furanacryloyl-Phe-Gly-Gly, was somewhat higher than that of the deglycosylated form as reflected in the values for $k_{cat}$ (22,876 min$^{-1}$ for the glycosylated form and 20,500 for the deglycosylated form). However, its $K_m$ was unchanged: 2.00 and 2.03 $\times 10^{-4}$ M for the glycosylated and deglycosylated forms, respectively. The $k_{cat}$ and $K_m$ values of the deglycosylated protein were in agreement with those reported for the wild-type tACE (11).

**DISCUSSION**

Glycosylation is an essential feature of the biosynthesis of ACE. Rabbit tACE, transiently expressed in human HeLa cells in the presence of tunicamycin, a chemical inhibitor of N-linked glycosylation, appeared only in the cytosol and was degraded rapidly (4). Studies of partial glycosylation of tACE by mutation of potential N-linked glycosylation sites indicated that the oligosaccharides at each site make different contributions to in vivo stability and localization (6). Such selective effects are not unprecedented. For example, in the case of the human vasoactive intestinal peptide receptor it was reported that Asn$^{58}$ or Asn$^{69}$ is critical for its correct delivery to the plasma membrane (15), and glycans at Asn$^{25}$ in human intestinal peptide receptors were found to confer protease resistance (16). We have identified three fully glycosylated and three partially glycosylated Asn residues among the seven N-glycosylation sequons in human tACE when the enzyme is expressed in CHO cells. Our results are similar to those obtained with rabbit tACE (6) in that all five conserved potential glycosylation sites are glycosylated. Three of these five are located near the NH$_2$ terminus of the enzyme, and either of the first two was found to be sufficient for the release of rabbit tACE (6). Our observation is consistent with the suggestion that glycosylation at these sites is probably involved in recognition and intracellular transport during later stages of enzyme biosynthesis.

The potential glycosylation site Asn$^{155}$ was found to be present in both glycosylated and unglycosylated forms. It is interesting to note that Asn$^{155}$ is located between cysteines 152 and 158, which form a disulfide bond (17), and it is “homologous” to Asn$^{131}$ in somatic ACE. We have suggested that specific disulfide formations in ACE could be critical in its folding process (17). This would be consistent with the report that ACE expressed in E. coli did not have a conformation that generated enzymatic activity (6). It appears that glycosylation processes in either the ER or Golgi may be affected by disulfide formation. Specific constraints imposed by the disulfide linkage between Cys$^{152}$ and Cys$^{158}$ in human tACE may prevent or reduce oligosaccharide attachment at Asn$^{155}$ much as it seems to prevent enzymatic deglycosylation (see above). In fact, Asn$^{155}$ is replaced by an Asp in rabbit tACE, whereas the positions of the two cysteines are conserved. It was demonstrated that conditions that prevent disulfide bond formation in tissue-type plasminogen activator allow complete glycosylation of a sequon that is otherwise variably glycosylated (18). Thus it may well be that mutations of cysteines that form the three disulfide bonds in tACE would not only affect folding but also affect glycosylation and consequently its transport and release.

A summary of the N-glycosylation state of the seven sequons is presented in (Fig. 6). It is interesting that the fully glycosylated sequons all end with threonine, whereas those that are only partially glycosylated and with serine. This agrees with the recent report that when serine is in the third position, glycosylation is less complete than in the case of threonine (19). For Asn$^{155}$ the effect of the disulfide bond discussed above may be an additional factor in reducing the level of glycosylation. Indeed, the particular order in which folding, disulfide formation, and other post-translational modifications occur within the ER will likely determine the accessibility of sequons to core glycosylation.

**Table II**

| Sequon   | Peptide | Peptide + carbohydrate | Before deglycosylation | After deglycosylation |
|----------|---------|------------------------|-----------------------|-----------------------|
|          |         |                        | Calc. m/z             | Obsd. m/z             | Calc. m/z | Obsd. m/z |
| Asn$^{72}$ | 47–79a  | 4004.2 + 1769.6-Fuc    | 5627.7                | 5626.3                |           |           |
|          |         | +NeuAC                 | 5919.0                | 2920.4                | 6210.3    | 6211.6    |
|          |         | +NeuAC                 | 3646.7                | 3646.6                | 4228.9    | 4228.9    |
|          |         | +NeuAC                 | 3544.8                | 3545.0                | 3748.0    | 3748.1    |
|          |         | +NeuAC                 | 4039.3                | 4039.2                | 4201.4    | 4201.6    |
|          |         | +NeuAC                 | 7964.7                | 7960.3                | 8256.0    | 8252.2    |
| Asn$^{90}$ | 87–102b | 1877.2 + 1769.6        | 4192.3                | 4216.0\textsuperscript{c} | 4484.6    | 4496.8\textsuperscript{f} |
|          |         | +NeuAC                 | 4775.9                | 4781.6                | 4775.9    | 4781.6    |
|          |         | +NeuAC                 | 7062.7                | 7062.2                | 7354.0    | 7353.2    |
|          |         | +NeuAC                 | 7645.3                | 7644.8                | 5294.1    | 5309.1\textsuperscript{d} |
|          |         |                        | 5293.1                | 1769.6                | 5294.1    | 5309.1\textsuperscript{d} |

\textsuperscript{a} Peptides were isolated from Lys-C digestion.

\textsuperscript{b} Peptides were isolated from CNBr and Asp-N digestions.

\textsuperscript{c} Large error due to low signal-to-noise ratio.

\textsuperscript{d} Mass increase due to oxidation of Met.
### Table III

| Tryptic peptides from human kidney ACE identified by Edman sequencing |
|---------------------------------------------------------------|
|Bold characters indicate potential glycosylation sites.|
| 1–30   | LDPLGQRGNFS"ADEAQGQLFAQSYNSS"AEQ− |
| 53–67  | QREEAALSLQFEEAFAEAA |
| 74–89  | ELYEFPINEQFDPQR |
| 91–96  | IIGAVYR |
| 97–107 | TLQGANLPLAK |
| 109–120 | QQYNALLSNSM3 |
| 121–126 | IYSTAK |
| 185–199 | QDFGTDTGAYWR |
| 296–316 | VAEFFTSLELSMPPFEFWEGL |
| 327–344 | EVCWNASDFYRNKD |
| 447–453 | WGVFSGR |
| 480–489 | NETNFDAK |
| 543–557 | VLGQASSRFQVEVL |
| 573–585 | YFQVQVQVLQQN |
| 691–693 | IIK |
| 701–713 | AALPAQELEENYK |
| 776–785 | YVEILNQAAR |
| 786–797 | LNGYVADGSMR |
| 798–811 | SMYETPSLEQDLER |
| 830–833 | ALHR |
| 884–889 | QGWTPR |
| 915–924 | SMLKPFEDG |
| 979–1001 | EGANPFPHEAIDGVLAIVSTPK |
| 1047–1054 | VFQGSTK |
| 1055–1065 | ENYQWNSLR |
| 1078–1087 | TQGDFDPAG |
| 1190–1203 | LGWQPYQWTNTPSAR |

a Identified as glycosylated by the absence of Asn peak. Sample was not treated with peptide N-glycosidase F.

b Identified as glycosylated by the presence of Asp peak. Sample was treated with peptide N-glycosidase F.

c Identified as not glycosylated by the presence of Asn peak. Sample was treated with peptide N-glycosidase F.

The only unglycosylated site in tACE was identified at Asn928. It has been shown that ACE is released from the plasma membrane of the CHO cells into the medium by a cleavage between Arg627 and Ser628 (8, 20). The exact nature of the protease(s) involved in the cleavage reaction is not clear, but it is possible that if the nearby Asn620 was glycosylated, this proteolytic modification might be sterically hindered. Indeed, this potential glycosylation site is absent in rabbit and mouse. From our initial analyses of human somatic ACE, human immunoglobulin E receptor-α-subunit caused misfolding and retention of the protein in the ER (21). Glycosylation may also affect protein secretion in a more direct way (22). It has been documented that N-glycans in properly folded lysosomal enzymes are recognized by a specific mechanism in the Golgi which generates a mannose 6-phosphate marker critical

### Fig. 7

SDS-polyacrylamide gel electrophoresis of wild-type and mutant tACE on 4–20% gradient gel stained with Coomassie Brilliant Blue. Lane 1, wild-type tACE; lane 2, tACEΔ36NJ; lane 3, endoglycosidase H-treated tACEΔ36NJ; lane 4, molecular mass markers.

### Fig. 8

Schematic representations of testis and somatic ACE showing the relative positions of the N-linked sites identified as being glycosylated only (closed box), glycosylated and nonglycosylated (dotted box), and nonglycosylated only (open box).

would seem to be another example of this general but not well understood phenomenon.

Our sequencing analysis of somatic ACE, although not yet complete, provides interesting information on N-linked glycosylation sites. Human seminal plasma ACE has been reported to consist of 14% carbohydrate by weight and to have approximately seven N-linked glycosylation sites (3). Whereas human kidney ACE may have slightly more carbohydrate, it is possible that we have identified most if not all of its glycosylated Asn sites and that the general pattern observed is, perhaps not surprisingly, similar to that seen in tACE (Fig. 7). Both isozymes are heavily glycosylated at their NH2 termini, with the Asn in sequon Asn-Lys-Ser glycosylated in both isozymes (Asn337 in tACE and Asn913 in somatic ACE). Further to the discussion above, it is likely that carbohydrates attached at the NH2 termini of the ACE proteins are used as general signals in trafficking to the plasma membrane. This concept is consistent with the observation by Sadhukhan and Sen and (6) that only one sequon at the NH2 terminus of the rabbit tACE was necessary and sufficient for TACE release, but the exact site of that sequon was not critical. On the other hand, it is tempting to speculate that the oligosaccharide chain on Asn-Lys-Ser in both isozymes may play a more specific functional role in ACE, as the nearby zinc binding site and the disulfide-linked cysteines are located in a region where sequence homology is significantly higher. It should be noted, however, that the sequence of residues 36–701 of TACE is identical to that of residues 613–1277 (the COOH-terminal domain) of somatic ACE. Yet although six of the seven sequons in tACE are glycosylated, the limited evidence (Table III) suggests that only one is glycosylated in the corresponding segment of somatic ACE. Moreover, while the N- and C- domains of somatic ACE are homologous only one of the 10 sequons in the former has a counterpart in the latter. The five that are glycosylated have no equivalent sequons. Thus, the similarity of glycosylation patterns seen in somatic and tACE (Fig. 8) seems to be determined by the order in which sequons enter the ER rather than by overall sequence.

It remains unclear whether and which glycosylation sites influence the transport, release, and stability of ACE. Naim (5) reported that selective inhibition of glucosidases I and II in the ER and mannosidase I in the Golgi causes a significant delay in intestinal ACE secretion. In addition, rabbit tACE appeared to be trapped intracellularly and undergo rapid degradation in tunicamycin-treated human HeLa cells. These observations would be consistent with the concept that a specific interaction between oligosaccharides and the folding machinery in the ER is required for proper protein processing in eukaryotic cells. It was demonstrated that mutations in glycosylation sites in the human immunoglobulin E receptor-α-subunit caused misfolding and retention of the protein in the ER (21). Glycosylation may also affect protein secretion in a more direct way (22). It has been documented that N-glycans in properly folded lysosomal enzymes are recognized by a specific mechanism in the Golgi which generates a mannose 6-phosphate marker critical

The only unglycosylated site in tACE was identified at Asn620. It has been shown that ACE is released from the plasma membrane of the CHO cells into the medium by a cleavage between Arg627 and Ser628 (8, 20). The exact nature of the protease(s) involved in the cleavage reaction is not clear, but it is possible that if the nearby Asn620 was glycosylated, this proteolytic modification might be sterically hindered. Indeed, this potential glycosylation site is absent in rabbit and mouse. From our initial analyses of human somatic ACE, human immunoglobulin E receptor-α-subunit caused misfolding and retention of the protein in the ER (21). Glycosylation may also affect protein secretion in a more direct way (22). It has been documented that N-glycans in properly folded lysosomal enzymes are recognized by a specific mechanism in the Golgi which generates a mannose 6-phosphate marker critical
to their delivery via endocytic pathways (23). Scheiffele et al. (24) demonstrated that nonglycosylated growth hormone is secreted both apically and basolaterally but only apically when glycosylated.

Structural studies of ACE have long been hampered by an inability to crystallize the enzyme. It is thought that the removal of the carbohydrates on ACE may help alleviate this problem. As tACE expressed in Escherichia coli is catalytically inactive (6), enzymatic removal of oligosaccharides appears to be a promising alternative strategy. Complete enzymatic deglycosylation can only occur when ACE is denatured. To obtain structurally meaningful information on the active site of ACE, it is essential that the deglycosylated ACE retains its native conformation and is enzymatically active. Our results show that the NB-DNJ-treated ACE mutant digested with endoglycosidase H is deglycosylated yet fully active. (It actually retains a single N-acetylglucosamine residue at each glycosylation site but should be devoid of oligosaccharide-based heterogeneity.) This form of ACE is thus considered a good candidate for crystallographic studies. Our results further support the notion that in vivo, glycosylation confers higher stability and plays a role in the transport and release of ACE.

Acknowledgments—We are indebted to Dr. D. J. Strydom and J. Brito for the sequence and amino acid analyses. We thank Dr. F. Platt for helpful discussions and R. J. Marks from Searle for providing the NB-DNJ.

REFERENCES
1. Ehlers, M. R. W., and Riordan, J. F. (1991) Biochemistry 30, 7118–7126
2. Conroy, J. M., Hartley, J. L., and Soffer, R. L. (1978) Biochim. Biophys. Acta 524, 403–412
3. Ripka, J. E., Ryan, J. W., Valido, F. A., Chung, A. Y. K., Peterson, C. M., and Urry, R. L. (1993) Biochem. Biophys. Res. Commun. 196, 503–508
4. Kasturi, S., Jabbar, M. A., Sen, G. C., and Sen, I. (1994) Biochem. 33, 6228–6234
5. Naim, H. Y. (1993) Biochem. J. 296, 607–615
6. Sadhu, R., and Sen, I. (1996) J. Biol. Chem. 271, 6429–6434
7. Ehlers, M. R. W., Chen, Y.-N. P., and Riordan, J. F. (1992) Biochem. Biophys. Res. Commun. 183, 199–205
8. Ehlers, M. R. W., Schwartz, S. L. U., Scholle, R. R., Manji, G. A., Brandt, W. F., and Riordan, J. F. (1996) Biochemistry 35, 9549–9559
9. Ehlers, M. R. W., Chen, Y.-N. P., and Riordan, J. F. (1991) Protein Expression Purif. 2, 1–9
10. Bebbington, C. R., and Hentschel, C. C. G. (1987) Cloning: A Practical Approach (Glover, D. M., ed) vol. 3, pp. 163–188, IRL Press, Oxford
11. Ehlers, M. R. W., Chen, Y.-N. P., and Riordan, J. F. (1991) Proc. Natl. Acad. Sci. U.S.A. 88, 1009–1013
12. Davis, S. J., Davies, E. A., Barclay, A. N., Daenke, S., Bodian, D. L., Jones, E. Y., Stuart, D. I., Butters, T. D., Dwek, R. A., and van der Merwe, P. A. (1995) J. Biol. Chem. 270, 369–375
13. Ehlers, M. R. W., Fox, E. A., Strydom, D. J., and Riordan, J. F. (1989) Proc. Natl. Acad. Sci. U.S.A. 86, 7741–7745
14. Karlsson, G. B., Butters, T. D., Dwek, R. A., and Platt, F. M. (1995) J. Biol. Chem. 266, 570–576
15. Covieneau, A., Fabre, C., Gaudin, P., Maoret, J.-J., and Laburthe, M. (1996) Biochemistry 35, 1745–1752
16. Sarvella, T., Pirhonen, J., Cantell, K., and Jukunen, I. (1995) Biochem. J. 308, 9–14
17. Storroock, E. D., Yu, X. C., Wu, Z., Biemann, K., and Riordan, J. F. (1996) Biochemistry 35, 9560–9566
18. Allen, S., Naim, H. Y., and Balleseid, J. N. (1995) J. Biol. Chem. 270, 4797–4804
19. Picard, V., Kersal-Bach, E., and Bock, S. C. (1995) Biochemistry 34, 8433–8440
20. Ramachandran, R., and Sen, I. (1995) Biochemistry 34, 12645–12652
21. Latourneur, O., Sica, S., Willette-Brown, J., Robertson, M. W., and Kinet, J.-P. (1995) J. Biol. Chem. 270, 8249–8256
22. Fiedler, K., and Simons, K. (1995) Cell 81, 309–312
23. Kornfeld, S., and Mellman, I. (1989) Annu. Rev. Cell Biol. 5, 483–525
24. Scheiffele, P., Peränen, J., and Simons, K. (1995) Nature 378, 96–98