THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 272, No. 6, Issue of February 7, pp. 3511–3519, 1997
Printed in U.S.A.

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‡§, Zhuchun 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

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, Asn^{90} and Asn^{109}, 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, Asn^{155}, Asn^{337}, and Asn^{586}, are partially glycosylated, whereas Asn^{72} appears to be fully glycosylated. The only potential site that was not modified is Asn^{620}. 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 Ser^{625}. When expressed in the presence of the α-glucosidase I inhibitor *N*-butyldeoxynojirimycin and treated with endoglycosidase H to remove all but the terminal *N*-acetylglucosamine residues, it retained full enzymatic activity, was electrophoretically homogeneous, and is a good candidate for crystallographic studies.

Both forms of angiotensin-converting enzyme (ACE[1]; 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 (*ldlD*), although it was released to a lesser extent (4). Similarly, it was reported (5) that inhibitors of glucosidases I and II in the endoplasmic reticulum (ER) and mannosidase I 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 NH_2-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

* This study was supported in part by National Institutes of Health Grants HL34704 (to J. F. R.) and GM05472 (to K. B.) and by a Roche Research Foundation grant (to M. R. W. E.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Associated with the Medical Research Council Liver Research Centre, University of Cape Town Medical School, Observatory 7925, South Africa.

** To whom correspondence should be addressed: Center for Biochemical and Biophysical Sciences and Medicine, Harvard Medical School, 250 Longwood Ave., Boston, MA 02115. Tel.: 617-432-1367; Fax: 617-566-3137.

[1] The abbreviations used are: ACE, angiotensin-converting enzyme; CHO, Chinese hamster ovary; ER, endoplasmic reticulum; tACE, testis ACE; NB-DNJ, *N*-butyldeoxynojirimycin; MALDI/TOF/MS, matrix-assisted laser desorption ionization/time-of-flight/mass spectrometry; HPLC, high performance liquid chromatography.

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. Cyanogen bromide (CNBr), trifluoroacetic acid, and calibration standards (angiotensin, insulin, myoglobin, oxidized insulin B-chain, and tosylphenylalanyl chloromethyl ketone-treated trypsin) were from Sigma. Glycosylation inhibitor *N*-butyldeoxynojirimycin (NB-DNJ) was kindly provided by Searle Co.

Construction of Expression Vector—pEE-ACEΔ36NJ encodes human tACE that lacks the heavily *O*-glycosylated, 36-residue NH_2-terminal sequence (7) and is truncated after Ser^{625}, thereby lacking most of the juxtamembrane stalk as well as the transmembrane and cytoplasmic domains (8), and was constructed as follows. The 5′ half of the ACE cDNA in the plasmid pLEN-ACE-JMΔ24 (8) was excised by digestion with *Bam*HI and *Nhe*I (9) and replaced with the similarly digested fragment from plasmid pLEN-ACEΔ36N (7). pLEN-ACE-JMΔ24 has an engineered *Eco*RI site at nucleotide 1984 in the ACE cDNA (8). The sequence between nucleotide 1854 (the start of the unique *Bcl*I site) and nucleotide 1990 (the end of the codon for Ser^{625}) in the native ACE cDNA was amplified by the polymerase chain reaction, using a 3′ primer that contained two stop codons (TAA and TAG) after the Ser^{625} codon, followed by an *Eco*RI site. The recombinant sequence was inserted into the pLEN-ACEΔ36N/JMΔ24 hybrid cut with *Bcl*I and *Eco*RI, to generate pLEN-ACEΔ36NJ. The ACEΔ36NJ coding sequence was excised by digestion of unique *Xba*I (generated after first subcloning in pBluescript) and *Eco*RI sites and inserted into the polylinker of the expression vector pEE14 (10), to generate pEE-ACEΔ36NJ.

Cell Culture, Transfections, and Enzyme Purification—CHO-K1 cells stably transfected with pLEN-ACEVII and expressing recombinant, wild-type human tACE were grown and maintained as described (9, 11). In addition, native CHO-K1 cells were cotransfected with pEE-ACEΔ36NJ (10 μg) and pSV2NEO (1 μg) by the calcium phosphate precipitate method, and clones stably resistant to G418 (Geneticin, Life Technologies, Inc.) were selected and assayed for ACE activity, by procedures detailed previously (9, 11). Clones stably expressing pEE-ACEΔ36NJ were further selected for resistance to methionine sulfoximine and then amplified, as described (10, 12). Methionine sulfoximine-amplified cells were grown first in GMEM-10 (Life Technologies, Inc.) containing 10% dialyzed fetal bovine serum (Life Technologies, Inc.) and 1.5 mM NB-DNJ for 3 days and then refed with GMEM-10, 5% dialyzed fetal bovine serum, 2 mM NB-DNJ. This medium was changed twice over a period of 9 days before harvesting. Soluble, recombinant tACE (wild-type and ACEΔ36NJ), purified from conditioned media by lisinopril affinity chromatography, was quantitated by amino acid analysis and assayed for activity, as described (9).

Deglycosylation of ACE—tACEΔ36NJ (12.5 nmol) purified from cultures treated with NB-DNJ was digested with endoglycosidase H (30 milliunits) in 100 mM sodium phosphate, 0.1 mM $ZnCl_2$, 1% bovine serum albumin, pH 6.0, for 16 h at 37 °C. The endoglycosidase H-treated ACE was passed through a lectin affinity column consisting of equal parts of concanavalin A, wheat germ, and lentil lectin, after equilibration with 20 mM Tris-HCl, 0.5 M NaCl at pH 7.4. The deglycosylated ACE was collected in the break-through. Free oligosaccharides and any other impurities were removed from the break-through fraction by a final lisinopril-Sepharose affinity chromatography step. The homogeneity of the tACEΔ36NJ after deglycosylation was confirmed by SDS-polyacrylamide gel electrophoresis on a 4–20% acrylamide gel and MALDI/TOF/MS.

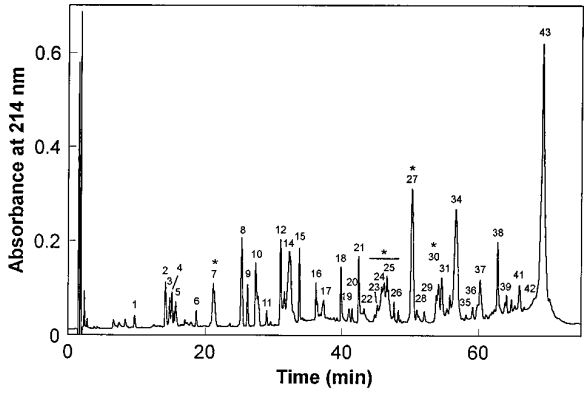

FIG. 1. **HPLC chromatogram of a Lys-C digest of wild-type tACE.** The peptide mixture was injected onto a C_{18} Delta-Pak column (Waters) and developed with a linear gradient of 8.5–51% acetonitrile in 0.1% trifluoroacetic acid over 90 min. Fractions designated by an *asterisk* provided no assignable 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 C_8 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*-glycanase (2.5 milliunits), and peptide *N*-glycosidase 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 C_{18} Delta-Pak column, 5 μm, 3.9 × 150 mm (Waters), or a C_8 Microsorb-MV column, 5 μm, 4.6 × 250 mm (Rainin), and eluted with a 10–60% gradient of 0.08% (v/v) trifluoroacetic acid in acetonitrile at a flow rate of 1 ml/min. The UV absorbance was monitored at 214 nm.

Mass Spectrometry—All mass spectra were obtained on a MALDI/TOF/MS instrument (Voyager-Elite Biospectrometry Workstation, PerSeptive Biosystems, Inc.). A nitrogen laser (337 nm) was used for desorption ionization. Measurements were carried out either in the linear or reflectron mode with mass accuracies of 0.1 and 0.01%, respectively. Spectra were collected over 100 laser shots.

Typical matrices used in these experiments were 3,5-dimethoxy-4-hydroxycinnamic acid (sinapinic acid) and α-cyano-4-hydroxycinnamic acid (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 *m/z* 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 (Asn^{155}, Asn^{337}, Asn^{586}, and Asn^{620}) were found unglycosylated. The CNBr fragment Leu^{593}-Arg^{627} 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 *N*-glycosidase F, neuraminidase, and *O*-glycosidase. Subsequent molecular mass determination by MALDI/TOF/MS identified four peptides, each of which contains at least one of the potentially glycosylated Asn residues.

TABLE I
Observed $[M+H]^+$ ions of unglycosylated peptides in human testis ACE

Peptide	Fraction	Calculated *m/z*	Observed *m/z*
Lys-C			
80–84[a]	1	614.8	614.5
118–137	10	2,345.6	2,344.5
175–188	22	1,868.1	1,868.7
189–307	43	13,716.6	13,713.8
308–317	8	1,307.6	1,307.1
318–338[b]	28	**2,423.7**	**2,424.8**
339–363	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	21	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[c]			
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		1,874.1	1,875.3
393–450		6,254.0	6,254.7
341–385		5,154.8	5,154.8
386–392		848.0	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.

The mass spectrum of the first CNBr-peptide identified after the removal of oligosaccharides is shown in Fig. 2. The observed peak at *m/z* 6506.2 agrees well with 6509.5 calculated for peptide Gln^{87}-Hse (Met^{142}), assuming both potential sites, Asn^{90} and Asn^{109}, were originally glycosylated and converted to Asp by peptide *N*-glycosidase F. To determine whether one or both residues are indeed glycosylated in the native ACE, the peptide was treated with endoproteinase Asp-N, which would cleave at Asp residues (including those newly generated) but not at Asn residues. After HPLC separation of the digest and MALDI/TOF/MS analysis of the fractions, both expected peptides, Asp^{90}-Phe^{102} and Asp^{109}-Gln^{120} were observed (Fig. 3) as molecular ions at *m/z* 1566.3 (calculated 1565.8) and 1444.5 (calculated 1443.8), respectively. These results are consistent with glycosylation at both Asn^{90} and Asn^{109}.

To confirm these findings, the original, glycosylated CNBr-peptide Gln^{87}-Hse (Met^{142}) 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. 4*A* and 5*A*, 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. 4*B* and Fig. 5*B*). They were identified as peptide Asp^{103}-Gln^{120} (Fig. 4*B*, expected *m/z* 2141.5) and Gln^{87}-Phe^{102} (Fig. 5*B*, expected *m/z* 1878.2). Thus both Asn^{90} and Asn^{109} are glycosylated in the wild-type tACE.

As shown in Fig. 4*A*, the different molecular ions observed in the glycosylated peptide Asp^{103}-Gln^{120} demonstrate the heterogeneous nature of the sugars attached to Asn^{109}. 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 $Hex_5HexNAc_4Deoxy\text{-}Hex_1$ to Asn^{109}. These data indicate that the *N*-linked glycans in tACE are mostly of the biantennary, fucosylated complex

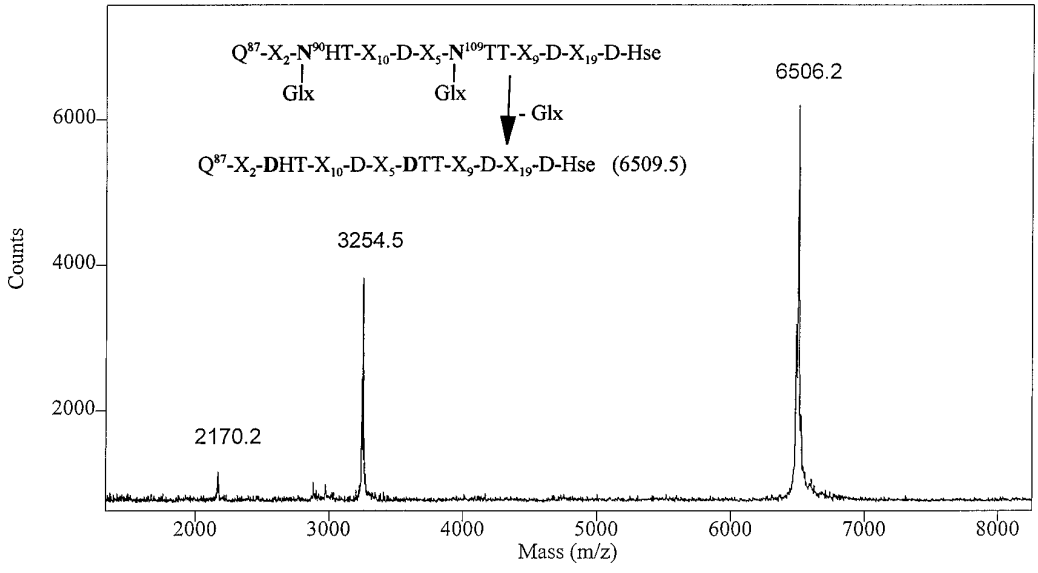

FIG. 2. **MALDI/TOF/MS of a peptide *N*-glycosidase F deglycosylated CNBr fragment from tACE.** The peaks at *m/z* 6506.2, 3254.5, and 2170.2 correspond to the singly, doubly, and triply protonated peptide Gln^{87}-Hse (Met^{142}), which contains two potential glycosylation sites, Asn^{90} and Asn^{109}. X_n refers to spacing by *n* amino acids.

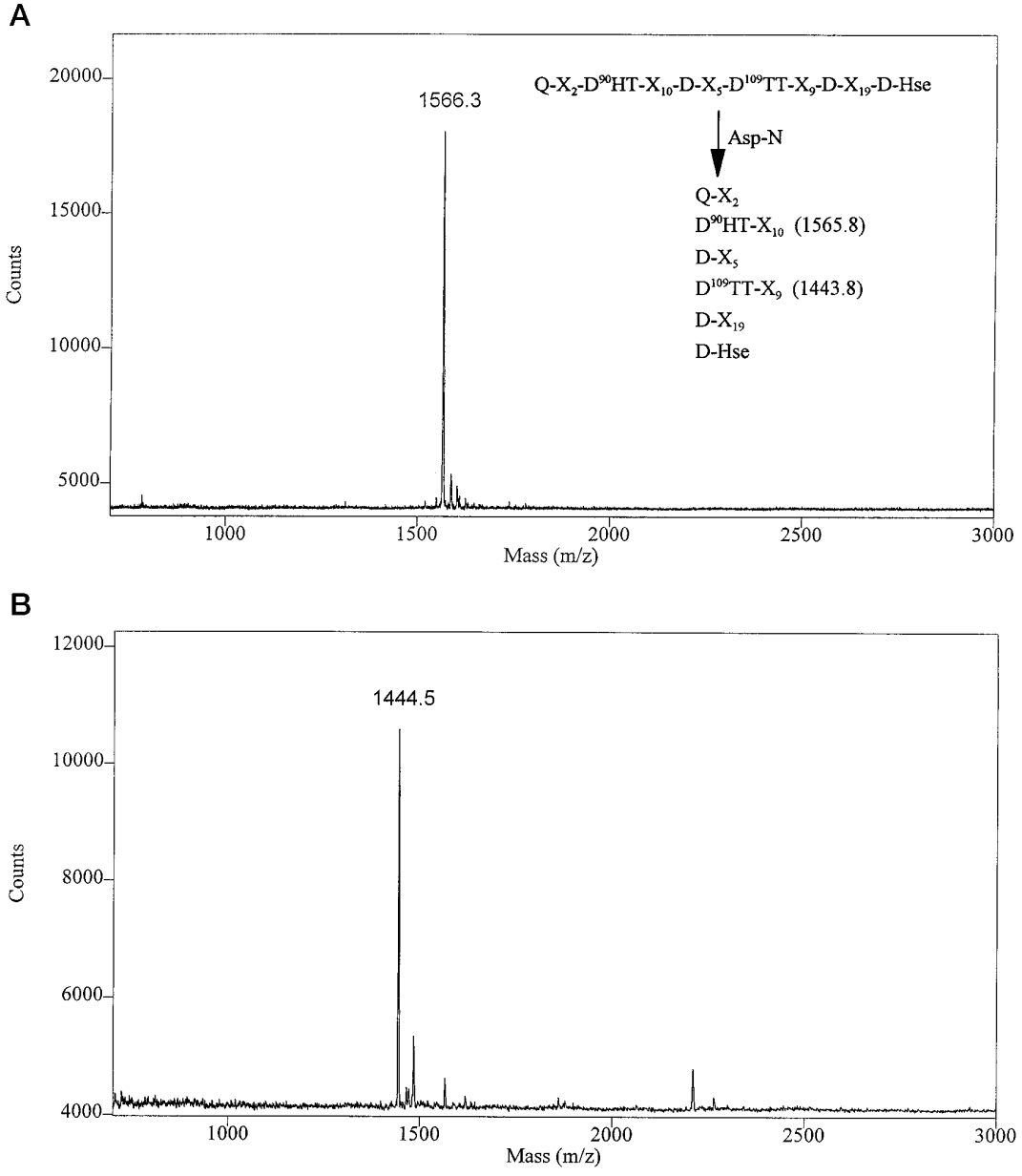

FIG. 3. **MALDI/TOF/MS of peptides obtained by Asp-N digestion of peptide *N*-glycosidase F deglycosylated peptide Gln^{87}-Hse (Met^{142}).** *Panel A* corresponds to the protonated peptide Asp^{90}-Phe^{102} and *panel B* to the protonated peptide Asp^{109}-Gln^{120}.

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. 4*A*. The likely compositions of the various carbohydrate moieties attached to Asn^{109} are listed in the *inset*.

The glycan structures on Asn^{90} as deduced from the mass spectrum are listed in Fig. 5*A*. 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. 5*A*).

Three residues, Asn^{155}, Asn^{337}, and Asn^{586}, 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 Asn^{155}, 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 Cys^{152} and Cys^{158} (see discussion below) which may prevent enzymatic deglycosylation. For Asn^{586}, a molecular ion at *m/z* 5309.1 was observed after glycosidase treatment, which is 16 Da higher than that expected for deglycosylated peptide Leu^{568}-Lys^{613} (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 Asn^{72}. 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

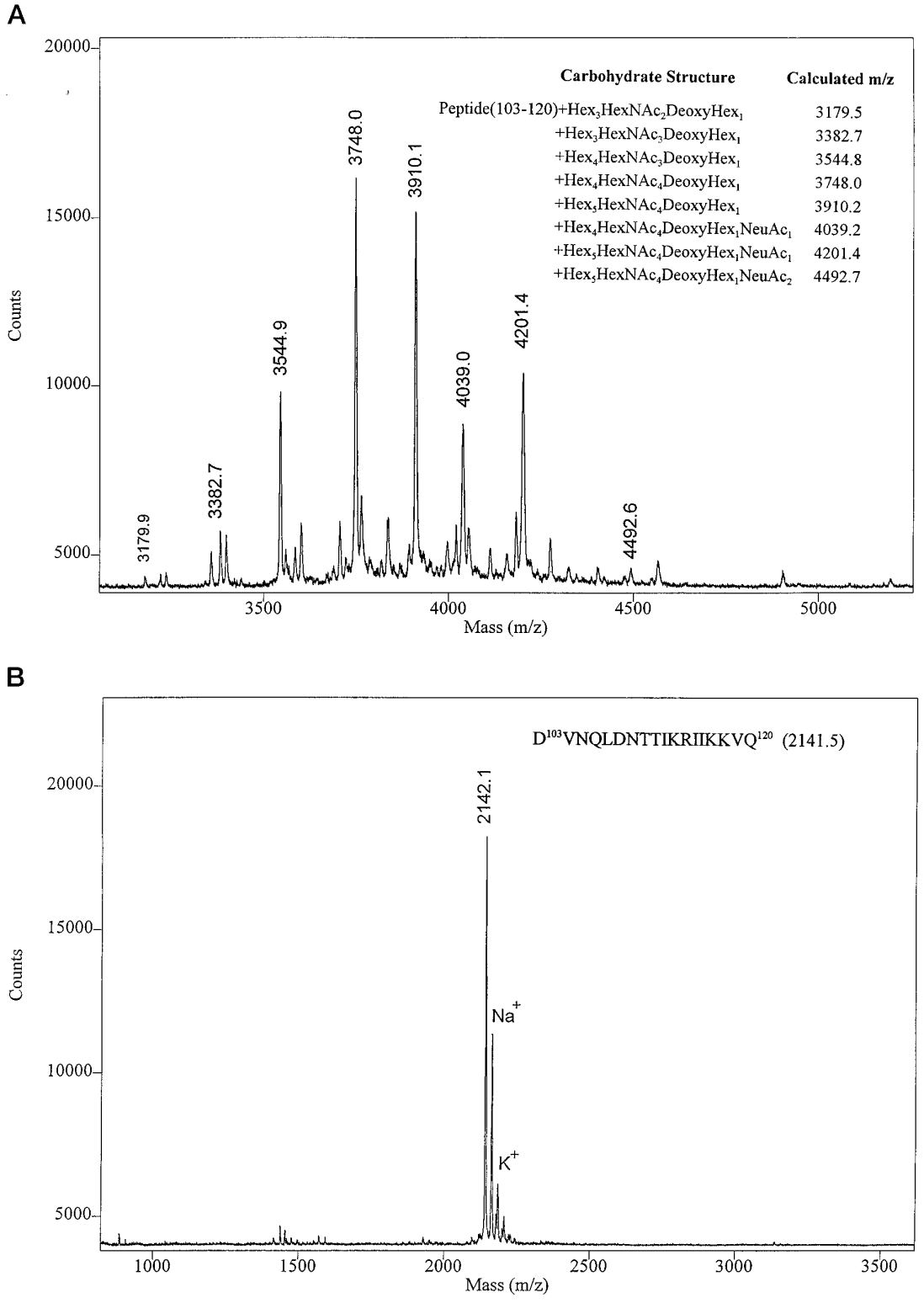

FIG. 4. *Panel A*, MALDI/TOF/MS of an HPLC fraction from the Asp-N digest of CNBr-glycopeptide Gln^{87}-Hse (Met^{142}). The multiple peaks from *m/z* 3000 to 4800 show the glycoform distribution. *Panel B*, MALDI/TOF/MS of the above fraction after treatment with a mixture of glycosidases. The peak at *m/z* 2142.1 corresponds to the protonated peptide Asp^{103}-Gln^{120}.

kidney. 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*-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 NH_2-terminal residues of the mature protein (7); it is also truncated after Ser^{625} 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

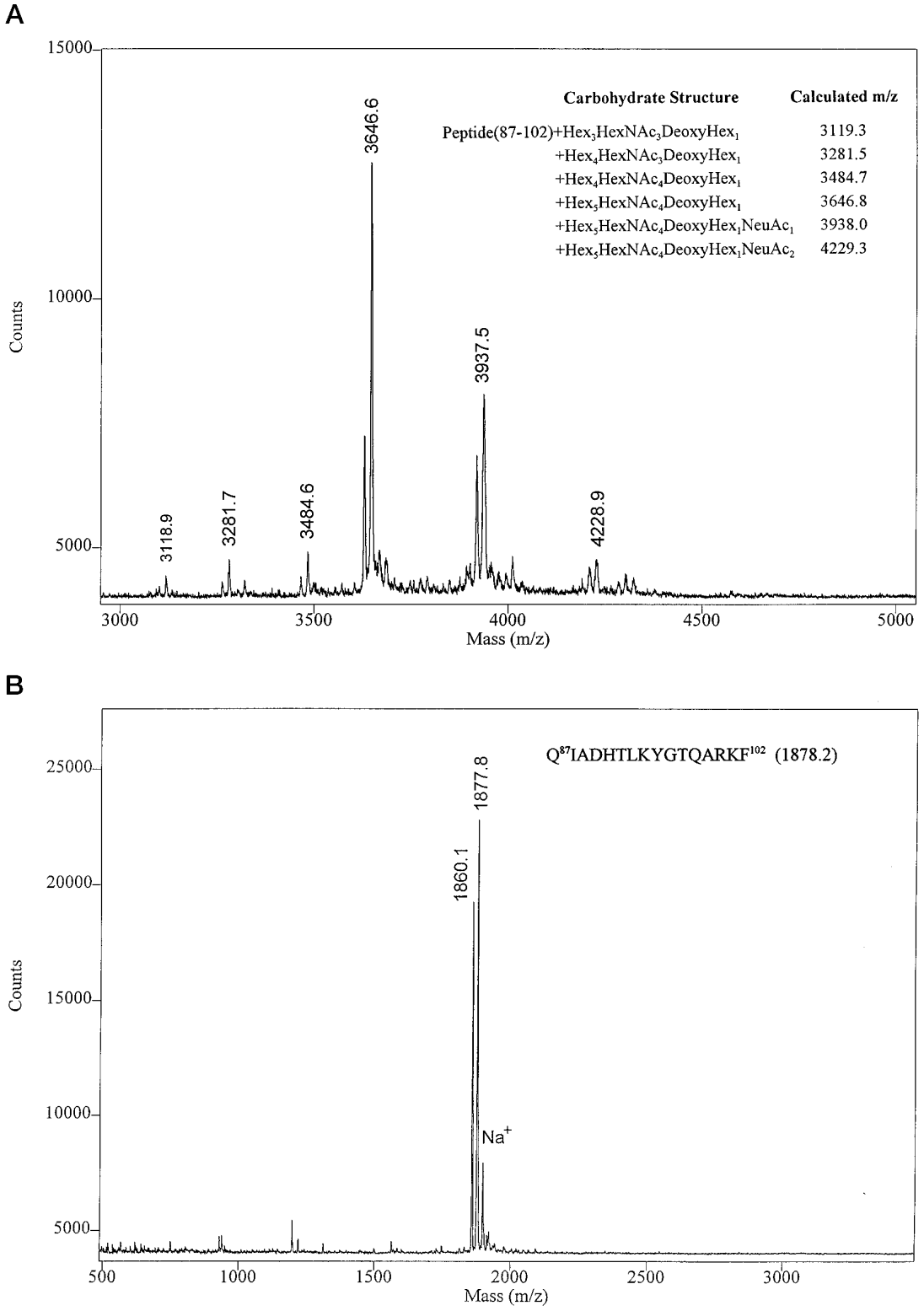

FIG. 5. *Panel A*, MALDI/TOF/MS of a HPLC fraction from an Asp-N digest of CNBr-glycopeptide Gln^{87}-Hse (Met^{142}). The multiple peaks from *m/z* 3000 to 4500 show the glycoform distribution. *Panel B*, MALDI/TOF/MS of the above fraction after treatment with a mixture of glycosidases. The peak at *m/z* 1877.8 is due to protonated peptide Gln^{87}-Phe^{102}. The peak at 1860.1 is the same peptide in which the NH_2-terminal Gln has cyclized to pyroglutamic acid. The 17-Da doublets can also be observed in *panel A*.

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

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—tACEΔ36NJ 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 oligosaccharide chains 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 I 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 interferon-γ 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.

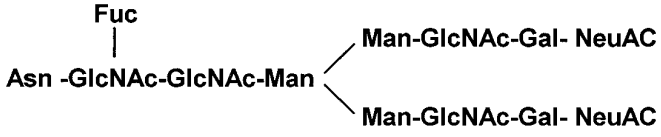

FIG. 6. **Proposed structure of the most extended *N*-linked oligosaccharide.**

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 end 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
Observed $[M+H]^+$ ions of glycosylated peptides in human testis ACE

Sequon	Peptide	Peptide + carbohydrate	Before deglycosylation		After deglycosylation	
			Calc. *m/z*	Obsd. *m/z*	Calc. *m/z*	Obsd. *m/z*
Asn^{72}	47–79[a]	4004.2 + 1769.6-Fuc	5627.7	5626.3		
		+NeuAC	5919.0	2920.4		
		+NeuAC	6210.3	6211.6		
Asn^{90}	87–102[b]	1877.2 + 1769.6	3646.7	3646.6	1878.2	1877.8
		+NeuAC	3938.0	3937.5		
		+NeuAC	4229.3	4228.9		
Asn^{109}	103–120[b]	2140.5 + 1769.6	3910.2	3910.1	2141.5	2142.1
		−(GlcNAc-Gal)	3544.8	3545.0		
		−Gal	3748.0	3748.1		
		−Gal + NeuAC	4039.3	4039.2		
		+NeuAC	4201.4	4201.6		
Asn^{155}	138–188[a]	5902.7 + 1769.6	7673.4	7669.1		
		+NeuAC	7964.7	7960.3		
		+NeuAC	8256.0	8252.2		
Asn^{337}	318–338[a]	2423.7 + 1769.6	4192.3	4216.0[c]	2424.7	2425.1
		+NeuAC	4484.6	4496.8[c]		
		+NeuAC	4775.9	4781.6[c]		
Asn^{586}	568–613[a]	5293.1 + 1769.6	7062.7	7062.2	5294.1	5309.1[d]
		+NeuAC	7354.0	7353.2		
		+NeuAC	7645.3	7644.8		

[a] Peptides were isolated from Lys-C digestion.
[b] Peptides were isolated from CNBr and Asp-N digestions.
[c] Large error due to low signal-to-noise ratio.
[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	LDPGLQPG**NFS**[a]ADEAGAQLFAQSY**NSS**[a]AEQ…
53–67	RQEEAALLSQEFAEA…
74–89	ELYEPIWQ**NFT**[b]DPQLR
91–96	IIGAVR
97–107	TLGSANLPLAK
109–120	QQYNALLS**NMS**[b]R
121–126	IYSTAK
188–199	QDGFTDTGAYWR
296–316	VAEEFFTSLELSPMPPEFWEG…
327–344	EVVCHASAWDFYNRKDFR
447–453	WGVFSGR
480–489	**NET**[b]HFDAGAK
543–557	VLQAGSSRPWQEVLK
573–585	YFQPVTQWLQEQN
691–693	IIK
701–713	AALPAQELEEYNK
776–785	YVELINQAAR
786–797	LNGYVDAGDSWR
798–811	SMYETPSLEQDLER
830–833	ALHR
884–889	QGWTPR
894–914	EADDFFTSLGLLPVPPEF**WNK(S)**[b]
915–924	SMLEKPTDGR
979–1001	EGANPGFHEAIGDVLALSVSTPK
1047–1054	VFDGSITK
1055–1065	ENYNQEWWSLR
1078–1087	TQGDFDPGAK
1190–1203	LGWPQY**NWT**[c]PNSAR

[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.

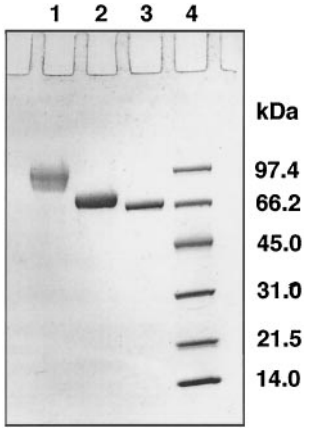

FIG. 7. **SDS-polyacrylamide gel electrophoresis of wild-type and mutant tACE on 4–20% gradient gel stained with Coomassie Brilliant Blue.** *Lane 1*, wild-type tACE; *lane 2*, tACEΔ36NJ; *lane 3*, endoglycosidase H-treated tACEΔ36NJ; *lane 4*, molecular mass markers.

The only unglycosylated site in tACE was identified at Asn^{620}. It has been shown that ACE is released from the plasma membrane of the CHO cells into the medium by a cleavage between Arg^{627} and Ser^{628} (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 Asn^{620} 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, Asn^{1196}, which corresponds to Asn^{620} of tACE, was also not glycosylated. Sequons at the COOH-terminal ends of proteins are often found to be unglycosylated. The present observations

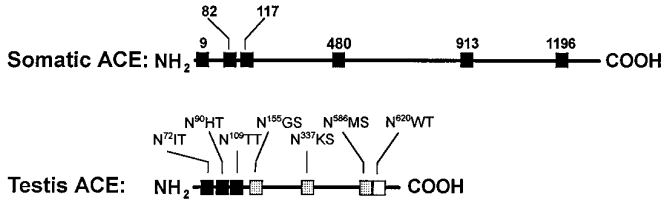

FIG. 8. **Schematic representations of testis and somatic ACE showing the relative positions of the *N*-linked sites identified as being glycosylated only (*closed box*), glycosylated and nonglycosylated (*dotted box*), and nonglycosylated only (*open box*).**

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 NH_2 termini, with the Asn in sequon Asn-Lys-Ser glycosylated in both isozymes (Asn^{337} in tACE and Asn^{913} in somatic ACE). Further to the discussion above, it is likely that carbohydrates attached at the NH_2 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 NH_2 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

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 *E. 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) *Biochemistry* **33,** 6228–6234
5. Naim, H. Y. (1993) *Biochem. J.* **296,** 607–615
6. Sadhukhan, 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., Schwager, 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. (1993) *J. Biol. Chem.* **268,** 570–576
15. Couvineau, A., Fabre, C., Gaudin, P., Maoret, J.-J., and Laburthe, M. (1996) *Biochemistry* **35,** 1745–1752
16. Sareneva, T., Pirhonen, J., Cantell, K., and Julkunen, I. (1995) *Biochem. J.* **308,** 9–14
17. Sturrock, 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 Bulleid, N. J. (1995) *J. Biol. Chem.* **270,** 4797–4804
19. Picard, V., Ersdal-Badju, E., and Bock, S. C. (1995) *Biochemistry* **34,** 8433–8440
20. Ramchandran, R., and Sen, I. (1995) *Biochemistry* **34,** 12645–12652
21. Letourneur, O., Sechi, 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

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

X. Christopher Yu, Edward D. Sturrock, Zhuchun Wu, Klaus Biemann, Mario R. W. Ehlers and James F. Riordan

J. Biol. Chem. 1997, 272:3511-3519.
doi: 10.1074/jbc.272.6.3511