Assignment of Free and Disulfide-Bonded Cysteine Residues in Testis Angiotensin-Converting Enzyme: Functional Implications†

Edward D. Sturrock,‡,§ X. Christopher Yu,‡ Zhuchun Wu,|| Klaus Biemann,|| and James F. Riordan*,‡

Center for Biochemical and Biophysical Sciences and Medicine, Harvard Medical School, 250 Longwood Avenue, Boston, Massachusetts 02115, and Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received January 31, 1996; Revised Manuscript Received April 25, 1996⊗

ABSTRACT: Human testicular angiotensin-converting enzyme (tACE) is an extracellular protein that contains seven cysteine residues. The cysteines occur in a sequential distribution that is precisely mimicked in the tACE from rabbit and mouse, and in both domains of all known species of somatic ACE. One of the cysteines in human tACE, Cys_{496}, is present in the reduced form as shown by labeling it with 5-[[2-(iodoacetyl)amino]ethylamino]naphthalene-1-sulfonic acid, isolating the fluorescent peptide from enzymatic digests by HPLC, and analyzing its sequence by matrix-assisted laser desorption ionization-mass spectrometry (MALDI-MS) . Thiol reagents have no significant effect on the activity of tACE, indicating that this Cys is not involved in catalysis. The other six cysteines exist as three disulfides. Mass spectral analysis of cyanogen bromide peptides has established that the cystine connectivities follow a nearest-neighbor, *aabbcc*, pattern i.e., Cys_{152}−Cys_{158}, Cys_{352}−Cys_{370}, and Cys_{538}−Cys_{550}, in which the disulfides form three small loops of five, 17, and 11 residues, respectively. Although these disulfide loops constitute less than 5% of the total sequence of the protein, they contribute to the overall structural stabilization of tACE.

Angiotensin-converting enzyme (ACE)[1] is a zinc peptidyl-dipeptidase which plays a critical role in blood pressure regulation. It converts the inactive decapeptide angiotensin I (AI) to the vasopressor octapeptide angiotensin II (AII), which in turn interacts with a variety of receptors that regulate fluid and electrolyte homeostasis. The enzyme also inactivates the vasodilating peptide bradykinin. There are two isozymes of ACE that are encoded by the same gene but which are generated from different transcription initiation sites in a tissue-specific manner (Howard et al., 1990; Hubert et al., 1990). Somatic ACE, which is involved in blood pressure regulation, is an ectoprotein anchored to the membrane of endothelial, epithelial, and neuroepithelial cells. The isozyme is also present as a soluble form in blood and seminal plasma. The other isozyme, testis ACE (tACE), is found only in postmeiotic cells and sperm (Langford et al., 1993). Both forms of ACE are catalytically similar in converting AI to AII *in vitro*. The substrate of the testis ACE *in vivo*, however, is not clear, and the function of this isozyme is unknown.

There is increasing evidence that ACE might play a role in the regulation of fluid and electrolyte balance in testis, as part of the renin angiotensin system (RAS). Several important components of the RAS have been localized to the testis, including renin mRNA (Pandey et al., 1984), AI, AII, and AIII immunoactivities (Pandey et al., 1984), and AII receptors (Millan & Aguilera, 1988; Khanum & Dufau, 1988). It has been suggested that tACE may play a more direct role in sperm function, by affecting sperm motility and capacitation through bradykinin inactivation (Hohlbrugger & Dahlheim, 1983). The presence of captopril, an ACE inhibitor, reduced the percentage of zona-free hamster oocytes penetrated by human sperm *in vitro* (Foresta et al., 1991). In a recent study, when somatic and tACE in mice were both inactivated by an insertional mutation in exon 14, the fertility of homozygous male mutants was drastically reduced (Krege et al., 1995). In addition, although both male and female heterozygotes showed decreased ACE activity in serum, only males had blood pressures that were ∼15% lower than normal. Such findings, along with studies in testicular ACE regulation during spermatogenesis and AII receptors in testis, may have important implications for human fertility and contraception.

Based on the sequence homology between the putative active site of ACE and the active site of carboxypeptidase A whose crystal structure has been determined, a variety of ACE inhibitors were developed in the late 1970s and early 1980s (Cushman et al., 1977; Ondetti et al., 1977; Patchett et al., 1980). They have been widely used in the treatment of hypertension. Since these inhibitors are not tissue specific and inhibit several other proteases nonselectively, most of them have characteristic side effects. Rational design of

† This study was supported in part by funding from NIH grants to J.F.R. (HL34704) and K.B. (RR00317 and GM05472).

* Address correspondence to this author at the Center for Biochemical and Biophysical Sciences and Medicine, Harvard Medical School, 250 Longwood Avenue, Boston, Massachusetts 02115.

‡ Harvard Medical School.

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

|| Massachusetts Institute of Technology.

⊗ Abstract published in *Advance ACS Abstracts,* June 15, 1996.

[1] Abbreviations: ACE, angiotensin-converting enzyme; tACE, testis angiotensin-converting enzyme; AI, angiotensin I; AII, angiotensin II; RAS, renin angiotensin system; IAEDANS, 5-[[2-(iodoacetyl)amino]-ethylamino]naphthalene-1-sulfonic acid; MALDI, matrix-assisted laser desorption ionization; HPLC, high-performance liquid chromatography, DTT, dithiothreitol; NEM, *N*-ethylmaleimide; MALDI-TOF-MS, matrix-assisted laser desorption ionization time-of-flight mass spectrometry; PSD, post-source decay; PIR, Protein Information Resource.

S0006-2960(96)00243-7 CCC: $12.00

highly specific and selective ACE inhibitors requires further understanding in the structure of ACE, including the ultimate determination of its crystal structure.

In the human, tACE is identical to the C-terminal half of the somatic ACE, but differs in the first 36 residues at the N terminal. It is thought to contain one catalytic domain, one transmembrane domain, and it requires a single zinc ion to be enzymatically active. As is found *in vivo*, the extracellular portion of the tACE is released into the medium in tissue culture. The cleavage site is 24 residues N-terminal to the transmembrane domain (Ramchandran et al., 1994); however, the identity and mechanism of the enzyme involved in the cleavage are not clear.

One interesting structural feature in ACE is the conservation of cysteine residues in the extracellular domain, both among species and between the two halves in the somatic isozyme. All seven cysteines in rabbit, mouse, and human tACE are located at the same positions in the consensus sequence of the three proteins. Human somatic ACE has 14 cysteines, seven in each half of the enzyme. We suggest that the presence of at least one unpaired cysteine in tACE and the potential specific pairing of cysteines are critical to the enzyme structure and function. In this report, we demonstrate that one cysteine residue in human tACE can be specifically labeled by the alkylating agent 5-[[2-(iodoacetyl)amino]ethylamino]naphthalene-1-sulfonic acid (IAEDANS). It was also determined by desorption ionization time-of-flight mass spectrometry (MALDI-TOF-MS) that three disulfide bonds link the remaining six cysteines.

MATERIALS AND METHODS

Reagents. Endoproteinase Lys-C was from Boehringer Mannheim Biochemicals. TPCK-treated trypsin, IAEDANS, iodoacetamide, cyanogen bromide, and trifluoroacetic acid were from Sigma Chemical Co.

Enzyme Purification. Recombinant tACE was purified from transfected Chinese hamster ovary cell culture medium as described previously (Ehlers et al., 1991). Protein concentrations were determined by HPLC separation and fluorescence detection of the 6-(aminoquinolyl)-*N*-hydroxysuccinimidyl carbamate derivatized amino acids generated by acid hydrolysis of tACE. Enzyme activity was measured with Fa-Phe-Gly-Gly as substrate as described by Holmquist et al. (1979).

Fluorescence Labeling of the Free Cysteinyl Residue. tACE (200 μg, 2 nmol) was incubated with the alkylating agent IAEDANS (200 nmol) in 500 mM Tris·HCl, pH 8.0, for 16 h at 25 °C. The reaction was carried out in the dark under nitrogen. The excess IAEDANS was removed by repeated centrifugal ultrafiltration using a Centricon 30 cartridge (Amicon). The alkylation was monitored by HPLC using a C_8 Microsorb-MV column, 5 μm, 4.6 × 250 mm (Rainin), equilibrated with 0.1% (v/v) trifluoroacetic acid in water. Elution with a 20%−60% gradient of 0.08% (v/v) trifluoroacetic acid in acetonitrile at a flow rate of 1 mL/min was monitored by UV absorbance (214 nm) and fluorescence (excitation at 340 nm; emission at 500 nm). The stoichiometry of the labeling was determined spectroscopically using the molar absorption coefficient $\epsilon = 6100$ at 337 nm for IAEDANS (Hudson & Weber, 1973).

The alkylated protein was denatured with 6 M guanidine hydrochloride, 500 mM Tris·HCl, pH 8.0, and 1.4 mM dithiothreitol (DTT), and the remaining cysteine residues were carboxymethylated with a 1.1 molar excess of iodoacetamide. The reaction mixture was passed through a small desalting Sephadex G-25 size-exclusion column (Pharmacia) previously equilibrated with 100 mM ammonium bicarbonate.

Trypsin Digest. tACE, derivatized with IAEDANS, was digested with 4% (w/w) TPCK-treated trypsin in 100 mM ammonium bicarbonate, pH 8.5, for 16 h at 37 °C. The peptides were chromatographed by reversed-phase HPLC using a C_{18} Delta-Pak column, 5 μm, 3.9 × 150 mm (Waters), and eluted with a 0.1% TFA in H_2O−0.08% TFA in acetonitrile gradient. Typically, samples containing a peptide mixture (up to 200 μg) in a volume of 50−100 μL were applied to the reverse-phase HPLC column.

Lys-C Digest. The buffer used for the alkylation with IAEDANS was exchanged with 100 mM ammonium bicarbonate, pH 8.5. Endoproteinase Lys-C (0.1 mg/mL in H_2O) was added to tACE (4%, w/w), and the digestion proceeded for 16 h at 37 °C.

Cyanogen Bromide Digest. Purified tACE (4 nmol) was lyophilized and dissolved in 70% trifluoroacetic acid (1 mL). Cyanogen bromide (40 mg) was added, and the reaction mixture incubated at 25 °C for 4 h. The digestion was stopped by the addition of ice cold water (1 mL) and kept on ice for 1 h before lyophilizing. The dried sample was then dissolved in 0.1 M ammonium bicarbonate and chromatographed by reverse-phase HPLC on a C_8 column as described above.

Cysteine Modification with Other Thiol Reagents. tACE (0.66 nmol) in 0.1 M Tris·HCl, pH 8.0 was incubated with *N*-ethylmaleimide (NEM) or iodoacetamide (66 nmol) for 2 h at 25 °C. Aliquots of 10 μL were removed at different times , diluted with 90 μL of cold 0.1 M Tris·HCl, pH 7.5, and assayed immediately.

For *p*-hydroxymercuribenzoate, didansyl-L-cystine, and dithiothreitol, tACE (0.5 nmol) in 0.1 M Tris·HCl, pH 8.0, was incubated with the thiol reagent, at various concentrations, at 37 °C for 1 h. Enzyme activity was measured in a standard assay system with Fa-Phe-Gly-Gly as substrate (Holmquist et al., 1979).

Mass Spectrometry. All mass spectrometry was carried out on a MALDI-TOF instrument (Voyager-Elite Biospectrometry Workstation, PerSeptive Biosystems, Inc., Framingham, MA). A nitrogen laser (337 nm) was used for desorption ionization. Experiments were carried out either in the linear or the reflectron mode with mass accuracy of 0.1% and 0.01%, respectively. For post-source decay (PSD) spectra the voltages of the single-stage reflectron mirror are controlled by the data system and can be stepped down in proportion to the accelerating voltage, so that the fragments can be reflected and focused to the reflector detector. In a typical experiment, a complete fragment ion spectrum is recorded by decreasing the mirror voltage in 10 steps down to 8% of the accelerating voltage.

3,5-Dimethoxy-4-hydroxycinnamic acid (sinapinic acid) or α-cyano-4-hydroxycinnamic acid (Aldrich) were used as matrices. About 1 μL of sample 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 (1−10 pmol peptide or peptide mixture) of the above solution was loaded on the sample plate and left to dry. Calibration standards angiotensin, insulin,

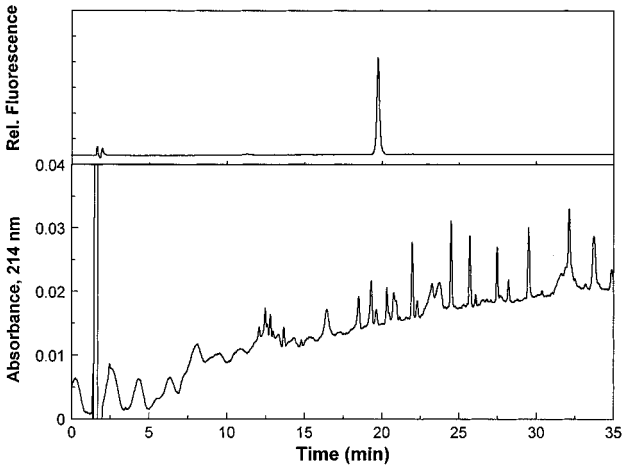

FIGURE 1: HPLC chromatograms of an endoproteinase Lys-C digestion of IAEDANS-labeled tACE. The bottom panel shows the UV absorbance at 214 nm, while the top panel is the fluorescence (excitation at 340 nm, emission at 500 nm). The HPLC column is a C_{18} Delta-Pak, 5 μm, 3.9 × 150 mm (Waters). The gradient is 10%−66% solvent B, linear in 35 min. Solvent A is water containing 0.1% trifluoroacetic acid. Solvent B is acetonitrile/water (9:1, v/v) containing 0.08% trifluoroacetic acid.

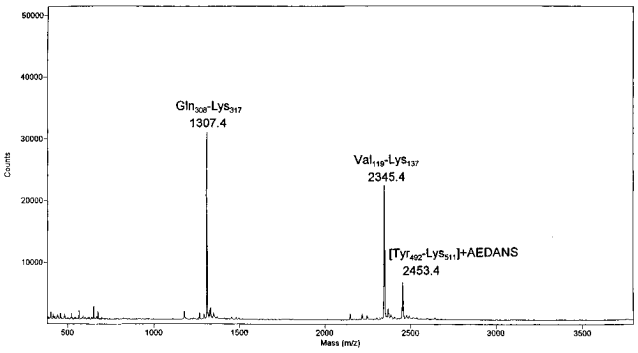

FIGURE 2: MALDI-TOF mass spectrum of the fluorescent HPLC fraction from the Lys-C digestion of the alkylated tACE.

myoglobin, and oxidized insulin B-chain were purchased from Sigma.

RESULTS

Identification of Unpaired Cysteine(s) with IAEDANS Labeling. Purified ACE[2] labeled with IAEDANS was digested with endoproteinase Lys-C, and a single fluorescent peak was observed in an HPLC separation of the entire digest (Figure 1). The MALDI-TOF mass spectrum (Figure 2) of this fraction showed three $[M + H]^+$ ion peaks, one of which (at *m*/*z* 2453.4) agrees well with that expected for the protonated AEDANS-derivatized peptide $Tyr_{492}-Lys_{511}$ (*m*/*z* 2453.7). The other two peptides that coeluted with the fluorescent peptide were identified as Lys-C fragments $Gln_{308}-Lys_{317}$ (expected $[M + H]^+ = m/z$ 1307.6, observed *m*/*z* 1307.4) and $Val_{119}-Lys_{137}$ (expected *m*/*z* 2345.6, observed *m*/*z* 2345.4). The result indicates that Cys_{496} in the fluorescent peptide was alkylated by IAEDANS during the *in vitro* incubation.

A trypsin digest of IAEDANS-labeled ACE was carried out to confirm the above result. All tryptic peptides were

[2] The protein was not deglycosylated in an effort to obviate any disulfide rearrangement that might take place at a higher pH and under the denaturing conditions required for complete removal of the glycans (Ehlers et al., 1992).

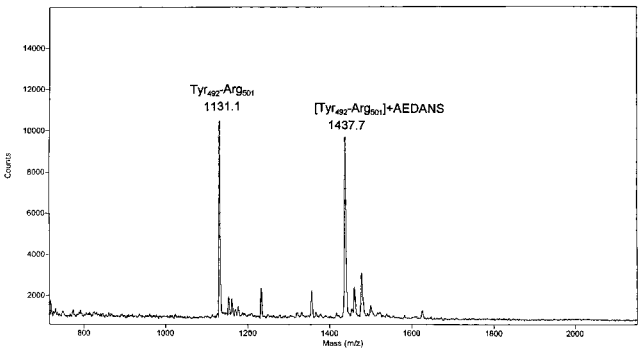

FIGURE 3: MALDI-TOF mass spectrum of the fluorescent HPLC fraction from the trypsin digestion of the alkylated tACE.

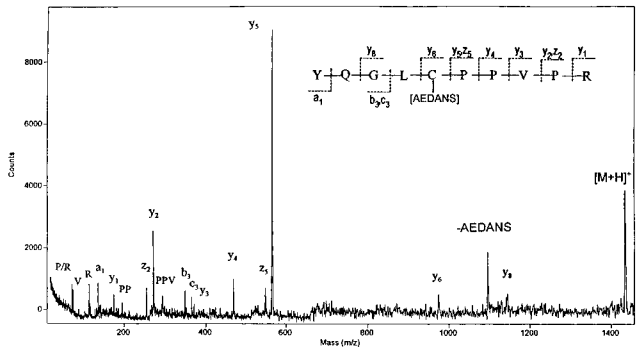

FIGURE 4: Post source decay mass spectrum of the AEDANS-labeled tryptic peptide $Tyr_{492}-Arg_{501}$. The peak labeled "−AEDANS" is the fragment ion that has lost the AEDANS moiety upon ionization.

separated by HPLC and the single fluorescent fraction was rechromatographed and analyzed by MALDI-TOF-MS (Figure 3). The peak at *m*/*z* 1437.7 was found to be due to the AEDANS derivative of peptide $Tyr_{492}-Arg_{501}$ (expected *m*/*z* 1436.6). The other major peak in the spectrum, *m*/*z* 1131.1, was found to correspond to the same peptide without the AEDANS moiety (expected *m*/*z* 1130.3). The magnitude of this peak varied in replicate analyses (data not shown) and is believed to be an indication of the lability of the fluorescent peptide in acidic solvents. However, the possibility that this ion is formed during the ionization process of MALDI analysis cannot be ruled out.

The size of the tryptic peptide $Tyr_{492}-Arg_{501}$ was sufficient to perform an additional experiment to localize the exact site of alkylation by IAEDANS. The PSD mass spectrum of the fluorescent peptide is shown in Figure 4. The series of ions observed due to fragmentation at each amide bond (y series) was consistent with the known sequence of the peptide. Peak y_1 indicates the positively charged C-terminal Arg_{501}, and each successive peak in the series corresponds to a peptide fragment with one additional residue from the C-terminal [for fragment nomenclature, see Biemann (1990)]. The mass difference between ions y_5 and y_6 unequivocally identifies Cys_{496} as the site of AEDANS attachment. We therefore conclude that the single free cysteine in the recombinant tACE is Cys_{496}.

Effects of Thiol Reagents. The treatment of tACE with a 100-fold molar excess of three thiol-alkylating agents (IAEDANS, iodoacetamide, and NEM) had no significant effect on enzyme activity as measured by the specific hydrolysis of Fa-Phe-Gly-Gly (Table 1). The mercaptide-forming reagent, *p*-hydroxymercuribenzoate, and disulfide exchange reagent, didansyl-L-cystine, likewise, had no effect

Table 1: Effect of Thiol Reagents on the Activity of ACE[a]

compound	% inhibition 0.1 mM	1 mM	10 mM
N-ethylmaleimide		5	
iodoacetamide		0	
IAEDANS		4	
p-hydroxymercuribenzoate	8	14	
didansyl-L-cystine	0	0	
dithiothreitol	7	29	78

[a] The effects of various alkylating, mercaptide forming, and other thiol reagents on ACE activity were examined. Each compound was added to the protein at the indicated concentration and as described under Materials and Methods. Incubation was carried out for 60 min with 50−100 μg of tACE prior to enzyme assays. Fa-Phe-Gly-Gly was used as substrate in a standard assay system (Holmquist et al., 1979). For the alkylations the incubations were also left for 16 h and showed no significant change in enzyme activity.

Table 2: Expected and Observed $[M + H]^+$ Ions of Peptides Derived from CNBr Cleavage of tACE

peptide[a]	expected *m/z*[b]	observed *m/z*[b]	difference	%
143−169	2876.2 or 2874.2	2874.0[c]	2.2 or 0.2	0.08 or 0.01
170−223	6320.1	6319.5	0.6	0.01
224−278	6430.3	6430.5	0.2	0.00
306−315	1225.4	1225.8	0.4	0.03
300−315	1874.1	1875.3	1.2	0.06
341−385	5156.8 or 5154.8	5154.9[c]	2.0 or 0.0	0.03 or 0
386−392	848.0	849.0	1.0	0.12
393−448	5994.7	5993.6	1.1	0.02
393−450	6254.0	6254.7	0.7	0.01
451−566	13439.4 or 13437.4	13450	10.6 or 12.6	0.08 or 0.09
567−578	1371.6	1372.5	0.9	0.07
593−626	4105.6	4103.8	1.8	0.04

[a] Residues are numbered beginning with residue 1 of the mature protein and without the 31-residue N-terminal signal peptide. [b] Isotopically averaged *m/z*, calculated for C-terminal homoserine lactone. [c] Internal calibration.

on enzyme activity. Dithiothreitol decreased enzyme activity by 7% at 0.1 mM and 78% at 10 mM concentration (900-fold molar excess).

Identification of Disulfide Linkages in ACE. Recombinant tACE was cleaved with cyanogen bromide, and the resulting peptides were separated by HPLC. Each fraction was analyzed by MALDI-TOF-MS, and the identities of peptides were established by comparing the observed *m/z* values with those expected for all CNBr fragments. The seven cysteines in tACE were located in three separate CNBr fragments, among 12 CNBr fragments identified (Table 2).

As shown in Figure 5, two cysteines from the center of the ACE sequence, Cys_{352} and Cys_{370}, were found to be located within CNBr peptide Leu_{341}−Met_{385}. Accurate mass measurement by internal calibration proved essential for assignment of a disulfide bond between the two cysteines, because the expected $[M + H]^+$ ions of the peptides with and without a disulfide bond differ only by two mass units, *m/z* 5154.8 vs *m/z* 5156.8. By using the singly and doubly charged ions of insulin (bovine) peaks as internal standards (*m/z* 5734.5 and 2867.8), the $[M + H]^+$ ion of the peptide was found at *m/z* 5154.9, indicating the presence of a disulfide.

All CNBr fragments were further treated with Lys-C and the resulting peptides were separated by HPLC and analyzed

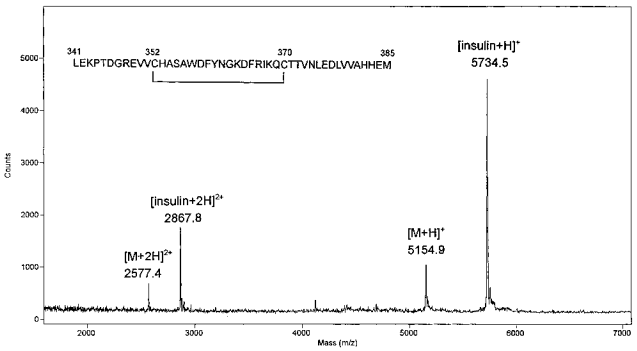

FIGURE 5: MALDI-TOF mass spectrum of the peptide Leu_{341}−Met_{385} containing an internal disulfide bond. Bovine insulin was added to the sample, and the singly and doubly charged ions were used for internal calibration.

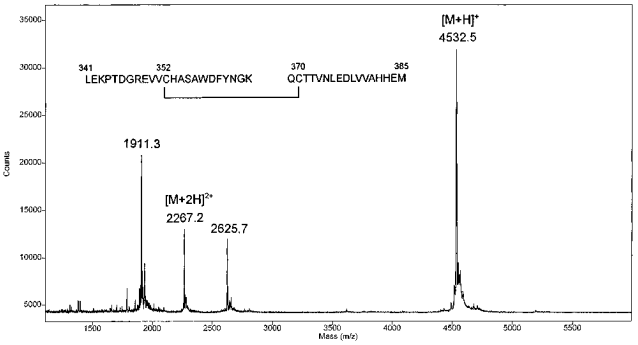

FIGURE 6: MALDI-TOF mass spectrum of an HPLC fraction from a Lys-C digest of a cysteine-containing CNBr peptide, Leu_{341}−Met_{385}. The peak at *m/z* 4532.5 is the disulfide-linked peptide, Leu_{341}−Lys_{363} plus Gln_{369}−Met_{385}. The peaks at *m/z* 2625.7 and 1911.3 correspond to Leu_{341}−Lys_{363} and Gln_{369}−Met_{385}, respectively.

by MALDI-TOF-MS. Peptide Asp_{364}−Lys_{368}, expected to be released from peptide Leu_{341}−Met_{385} by Lys-C, was found at *m/z* 678.8 (expected *m/z* 678.8, data not shown). The remainder of the original peptide, residues 341−363 plus residues 369−385, was found as a single species of *m/z* 4532.5 (expected 2621.9 + 1908.2 + 1 = 4531.1) (Figure 6), indicating that the two peptides generated by Lys-C are linked by a disulfide bond. Furthermore the individual components (*m/z* at 2625.7 and 1911.3) of the linked peptide were also present as a result of either in-source reduction or disulfide bond cleavage. This phenomenon has been reported previously (Crimmins et al., 1995). The peak at *m/z* 2267.2 is the doubly charged ion of *m/z* 4532.5.

Three cysteines at the C-terminal half of tACE, Cys_{496}, Cys_{538}, and Cys_{550}, were found in a large CNBr fragment, Ala_{451}−Met_{566} (*m/z* value 13 450, Table 2, spectrum not shown). Although we have previously shown by fluorescence labeling that Cys_{496} is the free cysteine in tACE, independent confirmation that the two other cysteines were disulfide-linked was still required. The peptide was isolated by HPLC and treated with trypsin. The resultant tryptic peptides were separated by HPLC and analyzed by MALDI-TOF-MS (Figure 7). Peptides Phe_{512}−Lys_{549} and Cys_{550}−Lys_{556} were found to be present as a single species of $[M + H]^+$ = *m/z* 5242.3 (expected *m/z* 5243.1), indicating that the two peptides are linked by a disulfide between Cys_{538} and Cys_{550}. The two components of this linked peptide were also observed, at *m/z* 856.6 and 4388.0, further confirming the identity of this peptide.

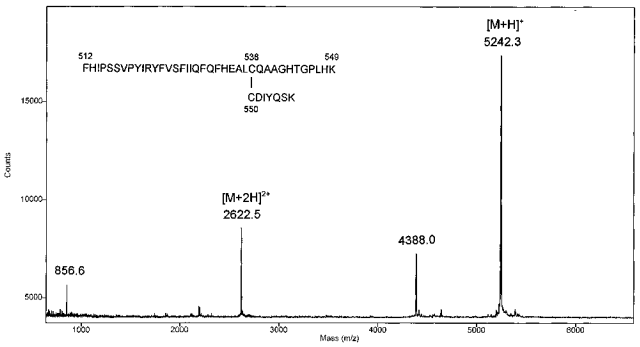

FIGURE 7: MALDI-TOF mass spectrum of an HPLC fraction from a trypsin digestion of a cysteine-containing CNBr peptide, Ala_{451}−Met_{566}. The peak at *m*/*z* 5242.3 is the disulfide-linked peptide, Phe_{512}−Lys_{549} plus Cys_{550}−Lys_{556}. The peaks at *m*/*z* 856.6 and 4388.0 correspond to Cys_{550}−Lys_{556} and Phe_{512}−Lys_{549}, respectively.

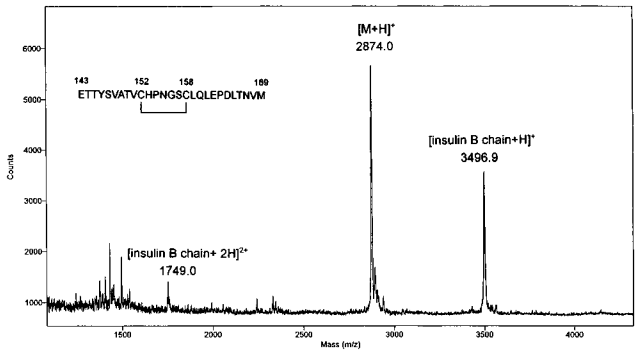

FIGURE 8: MALDI-TOF mass spectrum of the peptide Glu_{143}−Met_{169} containing an internal disulfide bond. Bovine insulin B-chain (oxidized) was added as an internal standard.

The two cysteines from the N-terminal half of tACE are located in one CNBr peptide, Glu_{143}−Met_{169}, shown in Figure 8. Once again by using the singly and doubly charged peaks of oxidized bovine insulin B-chain as internal calibrant, the $[M + H]^+$ ion of this disulfide-linked peptide was found at *m*/*z* 2874.0 (expected *m*/*z* 2874.2). We conclude that Cys_{152} and Cys_{158} are linked by a disulfide bond, as there are no suitable sites for further proteolytic cleavages in this CNBr fragment.

Thus, three disulfide linkages were found to be present in recombinant tACE between Cys_{352}−Cys_{370}, Cys_{538}−Cys_{550}, and Cys_{152}−Cys_{158}. The remaining cysteine, Cys_{496}, was identified as the only unpaired cysteine, using fluorescence labeling as well as mass spectrometry.

DISCUSSION

Molecular cloning and determination of the cDNA sequence of tACE prompted an interest in the seven cysteine residues and their structure−function implications. The possibility of at least one unpaired cysteine in an extracellular environment was tested in this study. The alkylation of free cysteines in ACE was carried out at near-physiological conditions, with an excess of the alkylating reagent. A single free cysteine, Cys_{496}, was identified, and its potential role in enzyme activity was determined. Surprisingly, all thiol-modifying reagents had little or no effect on specific activity, measured by the hydrolysis of Fa-Phe-Gly-Gly. Most likely the cysteine has no clear functional importance in the mature protein and in fact may be buried in the protein as is the case with the single thiol of serum albumin, where Cys_{34} is

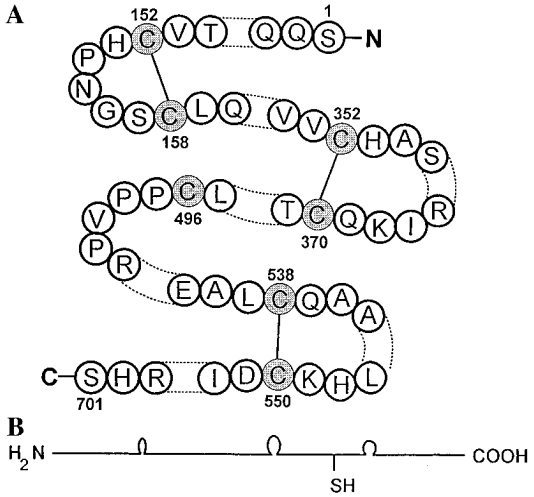

FIGURE 9: Schematic representations of tACE (A) showing the position of the free cysteine and the disulfide connections and (B) drawn on a proportional scale to illustrate the relative size of the disulfide "loops".

located in a crevice of the protein structure (Carter & He, 1992). This is suggested by the fact that stoichiometric alkylation of tACE requires a large excess of the labeling reagent and an extended reaction time. Furthermore the possibility exists that the cysteine is partly blocked by disulfide bonding to cysteine or glutathione. The extent of such coupling will depend on the redox state in which the enzyme is located or the cellular environment from which the soluble enzyme is being produced. If a significant portion of the enzyme is coupled in this fashion, it could limit the rate of cysteine alkylation. Recent NMR solution studies suggest that albumin exists in two structural forms dependent on whether Cys_{34} is a free thiol or a mixed disulfide (Christodoulou et al., 1994). Thus an equilibrium exists between the buried free cysteine and the bound cysteine in an exposed environment.

The diverse roles of free thiols in proteins are not fully understood. These include catalysis (activity), metal binding, and proenzyme activation as exemplified by the cysteine switch mechanism in collagenase (Windsor et al., 1991) and in protein folding. Thus in bovine pancreatic trypsin inhibitor a free cysteine has been implicated as a participant in protein folding promoting thiol-disulfide interchange (Weissman & Kim, 1992). The free thiolate is the active species that attacks an intramolecular disulfide, giving rise to the nucleophilic displacement of a different free cysteine and a new disulfide. This "disulfide shuffling" is required to generate the disulfide pattern of the native protein. Although the propeptide containing the free cysteine must be linked to the mature peptide to accelerate folding, a single cysteine residue tethered to the C-terminal end can also facilitate folding. This suggests a more general implication for free cysteines in protein folding. The potential role that the free cysteine in tACE might play in protein folding and expression remains to be investigated.

Three adjacent disulfide linkages (*aabbcc*) with a free cysteine positioned between the C-terminal and the middle disulfide bonds were identified in tACE using chemical and enzymatic cleavage followed by MALDI-TOF-MS (Figure 9A). This pattern of disulfide bonds bridging neighboring cysteines is very common and has been found to constitute approximately 50% of the disulfide patterns of proteins in

Table 3: Proteinases with Adjacent Disulfide Connections (*aabbcc*) Similar to Those Found in tACE[a]

code	proteinase	disulfides
	Exopeptidase Cleavage	
CPBYY	serine-type carboxypeptidase, *Saccharomyces*	304−318, 335−344
CPSMMU	muramoylpentapeptide carboxypeptidase, *Streptomyces*	45−123, 136−184, 212−253
	Endopeptidase Cleavage	
PRSMAG	proteinase A, *Streptomyces*	130−150, 247−274
PRSMBG	proteinase B, *Streptomyces*	128−148, 249−276
JQ0380	proteinase T, *Lysobacter*	46−137, 192−262
TRSMG	trypsin, *Streptomyces*	58−74, 177−192, 204−233
TRCY1	trypsin, crayfish	30−46, 159−174, 185−213
TRWV3Y	trypsin-like proteinase, mosquito	53−69, 178−194, 205−229[b]
KYVH20	chymotrypsin, hornet	25−40, 146−159, 169−193
KCUF	collagenolytic proteinase U, crab	26−42, 151−164, 174−200[b]
SUTIKA	endopeptidase K, *Tritirachium*	34−124, 179−248
PEMQAJ	pepsin A, monkey	92−97, 253−257, 296−329
CMUMF	mucorpepsin, *Rhizomucor*	120−126, 341−385
A26681	rhizopuspepsin II, *Rhizopus*	116−119, 320−353
CMBO	chymosin, bovine	105−110, 265−269, 308−341
PEMQCG	gastricsin, monkey	88−93, 251−255, 394−327
KHPGD	cathepsin D, pig	46−53, 220−224, 263−300[b]
HYBSPA	pseudolysin, *Pseudomonas*	221−255, 467−494
LYYXB4	beta-lytic metalloendopeptidase, *Lysobacter*	65−111, 155−168[b]
S22387	cuticle-degrading proteinase, *Metarhizium*	143−233, 288−360[b]
JS0260	serine proteinase, *Drosophila*	63−79, 189−201, 211−239[b]

[a] Disulfide connections were obtained from experimental data or [b]were predicted on the basis of sequence homology.

the Brookhaven National Laboratories protein structure data base (Benham & Jafri, 1993).

Human somatic ACE contains 14 cysteines, and the overall internal homology and, particularly, the conservation of cysteines between the two halves suggest that similar disulfide patterns exist in the N and C domains. This would leave two cysteine residues, one from each half, and there is evidence that indicates that they do not form a disulfide. Rabbit pulmonary somatic ACE has been cleaved by exposure to 1 N NH_4OH to give two fragments, one 82 kDa and the other 62 kDa (Iwata et al., 1983). By peptide mapping and terminal sequence analysis these were identified as being derived from the N and C domains, and it is unlikely that these fragments would be generated if the two cysteines were linked. It is possible that the pattern in the N domain differs from that found in tACE since the N and C domains of ACE have been reported to differ somewhat in substrate specificity and chloride activation (Ehlers & Riordan, 1991; Jaspard et al., 1993). The most striking example is the endopeptidase cleavage of luteinizing hormone-releasing hormone. However, it seems unlikely that this difference in endopeptidase cleavage could be ascribed to a difference in disulfide patterns. Structure−function relationships of various families and subfamilies of proteins have been examined in the light of their cystine homology (Baudyš et al., 1991; Bradshaw et al., 1994). The disulfides in cathepsin were used as a basis for a model of steric regulation of accessibility of the binding regions and discrimination between exo- and endopeptidase activities (Baudyš et al., 1991). We examined the cysteine connectivities (experimental and predicted) in all the proteinases in the Protein Information Resource (PIR) data base to see if there were any recurrent patterns, with a view to structure−function prediction. Proteinases with an *aabb* or *aabbcc* pattern similar to that found in tACE constituted 19% of the group (Table 3). The majority of this group displayed endopeptidase activity; however, some families had adjacent as well as "non-adjacent" disulfide patterns, e.g., trypsins. No cysteine homology exists between tACE and the other zincins or metalloproteinases in contrast to the catalytic binding site that is characteristic to almost all of them (Soubrier et al., 1988).

The spacing of the half-cysteines involved in disulfide bridges in proteinases culled from the PIR data base (Table 3) indicates that there is a preference for shorter connections, and they could be divided into three groups in terms of the number of residues separating linked cysteines. Some 21% were separated by less than 10 residues, with the shortest spacing being three. Another 57% were separated by less than 24 residues, and 18% were separated by more than 40 with the largest spacing being 90 residues. In an earlier survey of disulfides in proteins (Thornton, 1981) 49% of the observed disulfides were separated by less than 24 residues, and interestingly the frequency of observed disulfides with short connections was much greater than that expected by chance. This suggests that it is half-cystines close together in sequence that will preferentially form disulfide bridges, as first proposed by Kauzman (1959). In tACE the disulfide "loops" are a small part of the ACE structure (Figure 9B). With only 5, 17, and 11 spacer residues in the three loops, respectively, less than 5% of the 701 residues of tACE are so restricted. Nevertheless dithiothreitol abolishes activity either by reducing the disulfide(s) or even by removing zinc. Earlier findings showed that dithiothreitol had a similar effect on somatic ACE, and activity could not be restored after removal of the reducing agent (Chen & Riordan, 1990).

The formation of disulfide connections described in this study between adjacent cysteines generating intramolecular bridges suggests that these domains form independent folding structures which impact on protein function. The disulfide mapping of tACE should provide a useful basis for the mapping of the somatic ACE and further structural studies of the germinal isoform.

ACKNOWLEDGMENT

We thank D. Strydom for helpful discussions and computational assistance in the disulfide analysis of the proteases

drawn from the PIR data base and J. Brito for amino acid analyses.

REFERENCES

Baudyš, M., Meloun, B., Gan-Erdene, T., Fusek, M., Mareš, M., Kostka, V., Pohl, J., & Blake, C. C. F. (1991) *Biomed. Biochim. Acta 50*, 569−577.

Benham, C. J., & Jafri, M. S. (1993) *Protein Sci. 2*, 41−54.

Biemann, K (1990) *Methods in Enzymology*. (McCloskey, J. A., Ed.) Vol. 193, Appendix 5, pp 886−887.

Bradshaw, R. A., Murrayrust, J., Ibanez, C. F., McDonald N. Q., Lapatto, R., Blundell, T. L. (1994) *Protein Sci. 3*, 1901−1913.

Carter, D. C., & He, X.-M. (1992) *Nature 358*, 209−215.

Chen, Y.-N. P., & Riordan, J. F. (1990) *Biochemistry 29*, 10493−10498.

Christodoulou, J., Sadler, P. J., & Tucker, A. (1994) *Eur. J. Biochem. 225*, 363−368.

Crimmins, D. L., Saylor, M., Rush, J., & Thoma, R. S. (1995) *Anal. Biochem. 226*, 355−361.

Cushman, D. W., Cheung, H. S., Sabo, E. F., & Ondetti, M. A. (1977) *Biochemistry 16*, 5484−5490.

Ehlers, M. R. W., & Riordan, J. F. (1991) *Biochemistry 30*, 7118−7126.

Ehlers, M. R. W., Chen, Y.-N. P., & Riordan, J. F. (1992) *Biochem. Biophys. Res. Commun. 183*, 199−205.

Foresta, C., Mioni, R., Rossato,M., Varotto, A., & Zorzi, M. (1991) *Int. J. Androl. 14*, 333−339.

Gavras, H. (1990) *Circulation 81*, 381−388.

Hohlbrugger, G., & Dahlheim, H. (1983) *Adv. Exp. Med. Biol. 156B*, 845−853.

Holmquist, B., Bünning, P., & Riordan, J. F. (1979) *Anal. Biochem. 95*, 540−548.

Howard, T. E., Shai, S.-Y., Langford, K. G., Martin, B. M., & Bernstein, K. E. (1990) *Mol. Cell Biol. 8*, 4294−4302.

Hubert, C., Houot, A.-M., Corvol, P., & Soubrier, F. (1991) *J. Biol. Chem. 266*, 15377−15283.

Hudson, E. N., & Weber, G. (1973) *Biochemistry 12*, 4154−4161.

Iwata, K., Blacher, R., Soffer, R. L., & Lai, C.-Y. (1983) *Arch. Biochem. Biophys. 227*, 188−201.

Jaspard, E., Wei, L., & Alhenc-Gelas, F. (1993) *J. Biol. Chem. 268*, 9496−9503.

Kauzmann, W. (1959) in *Sulfur in Proteins* (Benesch, R., Benesch, R. E., Boyer, P., Klotz, I., Middlebrook,W. R., Szent-Gyorgyi, A., & Schwarz, D. R., Eds.) pp 93−108, Academic Press, New York.

Khanum, A., & Dufau, M. L. (1988) *J. Biol. Chem. 263*, 5070−5074.

Krege, J. H., John, S. W. M., Langenbach, L. L., Hodgin, J. B., Hagaman, J. R., Bachman, E. S., Jennette, J. C., O'Brien, D. A., & Smithies, O. (1995) *Nature 375*, 146−148.

Langford, K. G., Zhou, Y., Russell, L. D., Wilcox, J. N., & Bernstein, K. E. (1993) *Biol. Reprod. 48*, 1210−1218.

Millan, M. A., & Aguilera, G. (1988) *Endocrinology 122*, 1984−1990.

Ondetti, M. A., Rubin, B., & Cushman, D. W. (1977) *Science 196*, 441−444.

Pandey, K. N., Melner, M. H., Parmentier, M., & Inagami, T. (1984) *Endocrinology 115*, 1753−1759.

Patchett, A. A., Harris, E., Tristram, E. W., Wyvratt, M. J., Wu, M. T., Taub, D., Peterson, E. R., Ikeler, T. J., Broeke, J., Payne, L. G., Ondeyka, D. L., Thorsett, E. D., Greenlee, W. J., Lohr, N. S., Hoffsommer, R. D., Joshua, H., Ruyle, W. V., Rothrock, J. W., Aster, S. D., Maycock, A. L., Robinson, F. M., Hirschmann, R., Sweet, C. S., Ulm, E. H., Gross, D. M., Vassil, T. C., & Stone, C. A. (1980) *Nature 288*, 280−283.

Ramchandran, R., Sen, G. C., Misono, K., & Sen, I. (1994) *J. Biol. Chem. 269*, 2125−2130.

Soubrier, F., Alhenc-Gelas, F., Hubert, C., Allegrini, J., John, M., Tregear, G., & Corvol, P. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 9386−9390.

Thornton, J. M. (1981) *J. Mol. Biol. 151*, 261−287.

Weissman, J. S., & Kim, P. S. (1992) *Cell* 71 , 841−851.

Windsor, L. J., Birkedal-Hansen, H., Birkedal-Hansen, B., & Engler, J. A. (1991) *Biochemistry 30*, 641−647.

BI960243X