Seddon, J. M., Gregor, C., Kaye, R. D., & Marsh, D. (1984) *Biochemistry 23*, 2634-2644.
Shyamsunder, E., Gruner, S. M., Tate, M. W., Turner, D. C., So, P. T. C., & Tilcock, C. P. S. (1988) *Biochemistry 27*, 2332-2336.
Siegle, D. P. (1986a) *Biophys. J. 49*, 1155-1170.
Siegle, D. P. (1986b) *Biophys. J. 49*, 1171-1183.
Susi, H. (1981) *Chem. Phys. Lipids 29*, 359-368.
Tardieu, A., & Luzzati, V. (1970) *Biochim. Biophys. Acta 219*, 11-17.
Tate, M. W., & Gruner, S. M. (1989) *Biochemistry 28*, 4245-4253.
Thomas, L. C., & Chittenden, R. A. (1963) *Spectrochim. Acta 20*, 489-502.
Thomas, L. C., & Chittenden, R. A. (1970) *Spectrochim. Acta 26A*, 781-800.
Träuble, H., & Eibl, H. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 214-219.
Van Dijck, P. W. M., de Kruijff, B., Verkleij, A. J., Van Deenen, L. L. M., & De Gier, J. (1978) *Biochim. Biophys. Acta 512*, 84-96.
Van Venetië, R., & Verkleij, A. J. (1981) *Biochim. Biophys. Acta 645*, 262-269.
Verkleij, A. J., Mombers, C., Gerritsen, W. J., Leunissen-Bijvelt, L., & Cullis, P. R. (1979) *Biochim. Biophys. Acta 555*, 358-361.
Verkleij, A. J., De Maagd, R., Leunissen-Bijvelt, L., & de Kruijff, B. (1982) *Biochim. Biophys. Acta 684*, 255-262.
Wong, P. T. T., Capes, S. E., & Mantsch, H. H. (1989) *Biochim. Biophys. Acta 980*, 37-41.
Yellin, N., & Levin, I. W. (1977) *Biochim. Biophys. Acta 489*, 177-190.

Mass Spectrometric Identification of the Amino Donor and Acceptor Sites in a Transglutaminase Protein Substrate Secreted from Rat Seminal Vesicles†

Raffaele Porta,*,‡ Carla Esposito,§ Salvatore Metafora,∥ Antonio Malorni,⊥,# Pietro Pucci,⊥,▽ Rosa Siciliano,⊥,# and Gennaro Marino⊥,◇

Istituto di Chimica Biologica, Facoltà di Medicina e Chirurgia, Università di Palermo, Palermo, Italy, Dipartimento di Biochimica e Biofisica, Università di Napoli, Napoli, Italy, Istituto Internazionale di Genetica e Biofisica, CNR, Napoli, Italy, Servizio di Spettrometria di Massa, CNR, Università di Napoli, Napoli, Italy, Istituto di Chimica delle Molecole di Interesse Biologico, CNR, Arco Felice, Napoli, Italy, Dipartimento di Chimica, Università della Basilicata, Potenza, Italy, and Dipartimento di Chimica Organica e Biologica, Università di Napoli, Napoli, Italy

Received April 16, 1990; Revised Manuscript Received November 19, 1990

ABSTRACT: Four different transglutaminase-modified forms of a protein secreted by the rat seminal vesicles (SV-IV) were synthesized in vitro and characterized. FAB maps of both the native protein and its derivatives, produced by the purified guinea pig liver enzyme in the presence or absence of the polyamine spermidine, were obtained by mass spectrometric analysis after proteolytic digestions. Two differently derivatized SV-IV molecular forms, both possessing only one glutamine residue out of two (Gln-86) cross-linked to endogenous lysine residues, were produced when spermidine was omitted from the reaction mixture: (i) an insoluble homopolymer in which Lys-2, -4, -59, -78, -79, and -80 were involved in the linkage; (ii) a soluble form of the protein with an intramolecular ε-(γ-glutamyl)lysine isopeptide bond between Gln-86 and Lys-59. Two species of SV-IV–spermidine adducts were obtained when the protein was treated with transglutaminase in the presence of high concentrations of the polyamine. The first one was characterized by one spermidine molecule covalently bound to Gln-86 and the second one by two spermidine molecules respectively bound to Gln-9 and Gln-86.

Rat seminal vesicle (SV)[1] epithelium is an androgen-dependent tissue that synthesizes five major secretory proteins designated SV-I through SV-V depending on their migration in an SDS–PAGE system (Higgins et al., 1976; Ostrowski et al., 1979; Wagner & Kistler, 1987). Among these, SV-IV (M_r = 9.758) is the most extensively studied protein; the sequence of its 90 amino acids and the general features of the gene coding for it have been determined (Pan et al., 1980; Pan & Li, 1982; Mansson et al., 1981; Harris et al., 1983; Kandala et al., 1983, 1985). As for the biological properties of the protein, SV-IV was recently found to possess immunosuppressive and anti-inflammatory activities (Metafora et al., 1989a,b; Galdiero et al., 1989). Moreover, it was suggested that SV-IV might be one of the clotting proteins that serve as substrates for transglutaminase (TGase, EC 2.3.2.13) secreted by the rat anterior prostate in the formation of the

† This work was supported by grants (to R.P. and G.M.) from the Consiglio Nazionale delle Ricerche (CNR), Roma, Italy.
* Address correspondence to this author at Dipartimento di Biochimica e Biofisica, Via Constantinopoli, 16 I-80138 Napoli, Italy.
‡ Università di Palermo.
§ Dipartimento di Biochimica e Biofisica, Università di Napoli.
∥ Instituto Internazionale di Genetica e Biofisica, CNR.
⊥ Servizio di Spettrometria di Massa, CNR, Università di Napoli.
CNR, Arco Felize.
▽ Università della Basilicata.
◇ Dipartimento di Chimica Organica e Biologica, Università di Napoli.

[1] Abbreviations: SV, seminal vesicle; TGase, transglutaminase; FABMS, fast atom bombardment mass spectrometry; Spd, spermidine; TFA, trifluoroacetic acid.

0006-2960/91/0430-3114$02.50/0

copulatory plug (Williams-Ashman, 1984; Fawell et al., 1986). It has been also demonstrated that purified SV-IV is transformed by TGase in vitro into a high molecular weight polymer able to bind to rat epididymal spermatozoa and to suppress their immunogenicity (Paonessa et al., 1984; Metafora et al., 1987; Porta et al., 1988a, 1989, 1990).

The aim of the present investigation was to determine the specific amino-acceptor and amino-donor sites occurring in the SV-IV sequence and involved in the TGase-catalyzed cross-link formation. The protein has two glutamine residues (Gln-9 and Gln-86) and nine lysine residues (Lys-2, -4, -6, -34, -39, -59, -78, -79, and -80) that could, in principle, act as amino-acceptor and amino-donor sites, respectively. The identification of the glutamine and lysine residue or residues involved in TGase-mediated reactions is, at present, a laborious and long-lasting task, needing more than a single analytical system to yield unambiguous results. The main procedure so far developed consists of the chromatographic isolation and structural characterization of the isodipeptide ϵ-(γ-glutamyl)lysine following exhaustive proteolytic digestion of the reaction products (Griffin & Wilson, 1984; Loewy, 1984).

Since the introduction of fast atom bombardment (FAB) ionization mode by Barber et al. (1981), mass spectrometric methodologies have been successfully applied to several structural problems in protein and peptide chemistry [for reviews see Biemann and Martin (1987), Biemann and Soble (1987), and Morris and Greer (1988) and references therein].

Nowadays, FABMS analyses allow structural characterization of peptides in the low picomole range (Janes, 1990; Hunt, 1989); moreover, recent developments in the ionization modes such as laser desorption (Karas & Hillenkamp, 1988; Karas, 1989; Beavis & Chait, 1989) and electrospray (Mann, 1989; Loo, 1990; Chowdhury, 1990) allow accurate mass measurement of the molecular mass of proteins as large as 60 kDa.

As FABMS was employed in the analysis of TGase-mediated modifications of substance P and β-endorphin (Porta et al., 1988b; Pucci et al., 1988), this methodology has been used in this work to characterize the structure of the TGase reaction products of SV-IV obtained in vitro under different experimental conditions.

The main results reported in this paper show that Gln-86 is the only acceptor site involved in the SV-IV polymerization, whereas Lys-2, -4, -59, -78, -79, and -80 are the amino donors. As opposed to this, when the polyamine spermidine (Spd) is present in the assay mixture as an external amino donor, both Gln-9 and Gln-86 react, giving rise to a dipolyaminated form of the protein. The TGase-catalyzed formation of an intramolecular ϵ-(γ-glutamyl)-lysine isopeptide bond is finally reported, for the first time as far as it is known to us. In fact, a monomeric modified form of SV-IV, possessing an internal cross-link between Gln-86 and Lys-59, was identified following the incubation of the native protein with active TGase.

Experimental Procedures

Materials. SV-IV was purified to homogeneity from adult rat (Fisher-Wistar) SV secretion as described (Ostrowski et al., 1979). Guinea pig liver TGase, trypsin, *Staphylococcus aureus* V-8 protease, dithiothreitol, glycerol, thioglycerol, and Spd were purchased from Sigma Chemical Co. HPLC-grade solvents and reagents were obtained from Carlo Erba and Sep-pak C-18 cartridges from Waters.

TGase-Mediated Modification of SV-IV. A total of 0.5 mg of SV-IV was incubated at 37 °C overnight with 10 μg of TGase in 1.0 mL of 125 mM Tris-HCl buffer, pH 8.0, containing 2.5 mM $CaCl_2$, 10 mM dithiothreitol, and 0.2 M Spd. Blanks were simultaneously prepared in the absence of either calcium or enzyme. At the end of the incubation, the reaction mixture was centrifuged at 10000 rpm for 10 min, and the resulting modified forms of SV-IV, contained in the supernatant, were separated by ion-exchange chromatography on the FPLC apparatus (Pharmacia) using an analytical Mono S HR 5/5 (10 × 0.3 cm, 10 μm) column. The column was equilibrated with 50 mM sodium phosphate, pH 7.6 (buffer A), and the elution was performed with the same buffer containing 0.5 M NaCl (buffer B). Separation of the native and modified forms of SV-IV was accomplished by a linear gradient of buffer B from 0% to 60% in 10 min (flow rate 1 mL/min). 0.2-mL fractions were collected and the different protein peaks separately pooled and analyzed by SDS–PAGE and FABMS.

Cross-linked forms of SV-IV were prepared by incubation of the native protein under the same experimental conditions described above but in the absence of Spd. Soluble and insoluble reaction products were separated by centrifugation at 10000 rpm for 10 min and finally analyzed by FABMS. The soluble material was also analyzed by FPLC, carried out under the Experimental conditions described above, and by SDS–PAGE; FPLC peaks were submitted to FABMS analysis.

FABMS Analysis. Tryptic hydrolysis of native and modified SV-IV (0.5 mg) was carried out in 0.4% ammonium bicarbonate, pH 8.5, at 37 °C for 4 h (1:50 w/w). V-8 protease digestion (1:50 w/w) was performed on the tryptic peptides in the same buffer (pH 8.0), at 40 °C for 6 h, following inactivation of trypsin by boiling the sample for 5 min. Before mass spectrometric analysis, all of the samples were desalted on Sep-pak C-18 cartridges equilibrated with 0.1% trifluoroacetic acid (TFA). The cartridges were washed with 5 mL of 0.1% TFA; the peptides were then eluted with 3 mL of 70% acetonitrile containing 0.1% TFA. The latter fractions were collected and evaporated to dryness, and the dry samples were finally dissolved in 5% acetic acid; 2 μL of each sample (3–5 nmol) was loaded onto a glycerol–thioglycerol-coated probe tip. FAB mass spectra were recorded on a VG ZAB 2 SE double-focusing mass spectrometer fitted with a VG cesium gun operating at 25 keV (2 μA). The amplification of the electrical signal was decreased during the magnet scan from high mass to low mass according to the intensity of the mass signals as they were observed on the oscilloscope; spectra were recorded on UV paper and manually counted. To confirm the assignments, manual Edman-degradation steps were performed on the peptide mixture by using 5% phenyl isothiocyanate in pyridine as a coupling agent, and the truncated peptides were analyzed by FABMS.

SDS–PAGE. Electrophoretic analyses of the soluble SV-IV molecular forms were performed as previously described (Laemmli, 1970; Porta et al., 1990).

Results

FAB Mapping of Native SV-IV. The peptide map of native SV-IV was obtained by direct FABMS analysis of the peptide mixture generated by enzymatic digestions of the protein with trypsin and V-8 protease (Figure 1). The FAB spectrum of the peptide mixture is shown in Figure 2 together with the assignment of the mass signals to the corresponding peptides along the SV-IV amino acid sequence, made on the basis of molecular weight and confirmed by FABMS analysis after a single Edman-degradation step. The native SV-IV was almost completely mapped in a single experiment, the only undetected fragment being the tripeptide 57–59 (Naylor et al., 1986).

The two glutamine residues possibly involved in TGase-catalyzed reactions as amino-acceptor sites (Gln-9 and Gln-86)

```
                    5                   10                  15                  20
-ARG-LYS-THR-LYS-GLU-LYS-TYR-SER-GLN-SER-GLU-GLU-VAL-VAL-SER-GLU-SER-PHE-ALA-SER

                   25                  30                  35                  40
-GLY-PRO-SER-SER-GLY-SER-SER-ASP-GLU-LEU-VAL-ARG-ASP-LYS-PRO-TYR-GLY-PRO-LYS-VAL

                   45                  50                  55                  60
-SER-GLY-GLY-SER-PHE-GLY-GLU-GLU-ALA-SER-GLU-GLU-ILE-SER-SER-ARG-ARG-SER-LYS-HIS

                   65                  70                  75                  80
-ILE-SER-ARG-SER-SER-GLY-GLY-SER-ASN-MET-GLU-GLY-GLU-SER-SER-TYR-ALA-LYS-LYS-LYS

                   85                  90                  95                 100
-ARG-SER-ARG-PHE-ALA-GLN-ASP-VAL-LEU-ASN
```

FIGURE 1: Amino acid sequence of rat seminal vesicle SV-IV protein. Peptides underlined were detected in the FAB mass spectrum of the native protein following trypsin and V-8 protease digestions. Experimental details are given in the text.

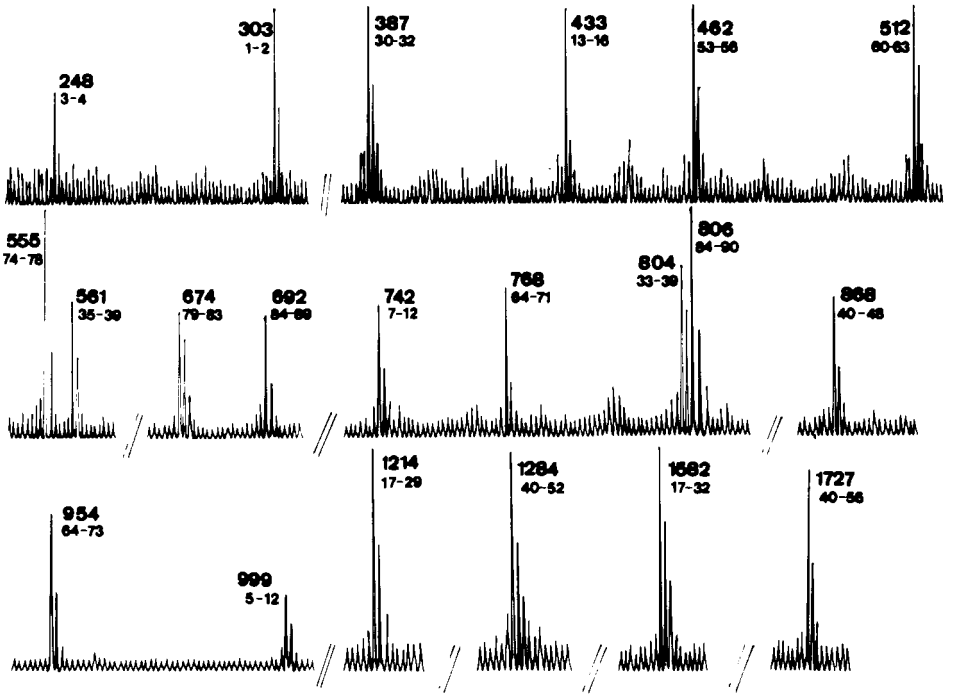

FIGURE 2: FAB map of native SV-IV following hydrolysis with trypsin and V-8 protease. The signals were associated to the corresponding peptides on the basis of their molecular weight. The relative abundances of the mass signals are on an arbitrary scale, since the gain was varied during the scan.

were mapped in four peptides, namely fragments 5–12 as well as 7–12 and 84–90 as well as 84–89, the latter resulting from an aspecific (chymotryptic-like) cleavage at the level of Leu-89. Meanwhile, the putative amino-donor lysine residues were present in several fragments of the mixture (peptides 1–2, 3–4, 5–12, 33–39, 35–39, 57–59, 74–78, and 79–83). Therefore, the map of native SV-IV was used as a reference for analyzing a TGase-modified form or forms of the protein.

It should be noted that several cleavage sites were only partially hydrolyzed by the proteolytic enzymes, giving rise to several peptides mapping the same protein region (see, for example, fragments 17–29 and 17–32, resulting from an incomplete cleavage at the level of Glu-29). No further attempts were carried out to modify the hydrolysis conditions because all the residues of interest had been detected.

TGase-Catalyzed Synthesis and Characterization of (γ-glutamyl)-Spd Derivatives of SV-IV. The modification of SV-IV by polyamines was carried out by incubating the protein in the presence of TGase, Ca^{2+}, and Spd under the conditions reported under Experimental Procedures. The separation of the reaction products was obtained by FPLC on a cation-exchange column (Figure 3). The protein peaks 2 and 3, de-

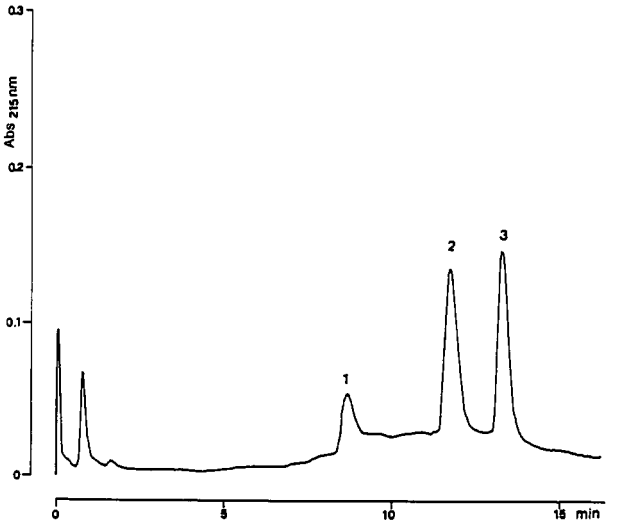

FIGURE 3: FPLC separation of the protein species (50 μg) contained in the reaction mixture following incubation of the native SV-IV with TGase and Ca^{2+} in the presence of Spd. Peak 1, native SV-IV; peaks 2 and 3, Spd derivatives of SV-IV. Further details are given in the text.

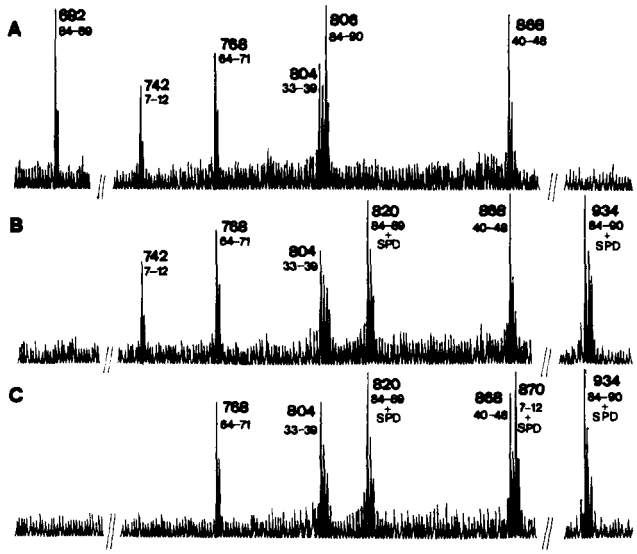

FIGURE 4: Partial FAB mass spectra of the proteolytic digests of native SV-IV (panel A) and of protein peaks 2 (panel B) and 3 (panel C) from FPLC. The relative abundances of the mass signals are on an arbitrary scale, since the gain was varied during the scan.

tectable only when the enzyme was activated by Ca^{2+}, were separately examined by SDS–PAGE and by FABMS following proteolysis with trypsin and V-8 protease. The electrophoretic pattern (data not shown) demonstrated the occurrence in both peaks of protein bands having a mobility identical with that of native SV-IV, thus indicating that both fraction 2 and fraction 3 were modified monomeric forms of SV-IV.

The FABMS characterization of the two FPLC fractions confirmed this hypothesis, demonstrating that fraction 2 corresponded to SV-IV with one Spd molecule covalently bound to Gln-86, whereas the protein eluted in fraction 3 had two Spd molecules linked to Gln-9 and Gln-86, respectively. The partial mass spectrum of FPLC fraction 2 (Figure 4, panel B) shows the disappearance of the signals at m/z 692 and 806, corresponding to the Gln-86-containing fragments 84–89 and 84–90, and the appearance of two new signals at m/z 820 and 934. These mass values refer to the above peptides carrying a single Spd molecule, thus indicating that Gln-86 was recognized and modified by TGase. It should be noted that the mass signal at m/z 742, corresponding to the Gln-9-containing peptide 7–12, was unchanged in this spectrum.

When the FPLC fraction 3 was mass analyzed (Figure 4, panel C), the spectrum showed the presence of three new signals at m/z 820, 870, and 934. Two of them were already described above, while the signal occurring at m/z 870 corresponds to peptide 7–12 in which Gln-9 was covalently bound to Spd. This result, together with the observed disappearance of the signal at m/z 742, demonstrated that this fraction was a modified SV-IV form having glutamine residues at both position 9 and position 86 attached to Spd molecules.

TGase-Catalyzed Synthesis and Characterization of Intermolecularly ε-(γ-Glutamyl)lysine-Cross-Linked SV-IV. The synthesis of the SV-IV polymer was carried out by incubating the protein with TGase and Ca^{2+}, in the absence of Spd, under the conditions reported under Experimental Procedures. At the end of the incubation the insoluble material was recovered from the reaction mixture by centrifugation and submitted to FABMS analysis following digestion with trypsin and V-8 protease. The FAB spectrum of the resulting peptide mixture, reported in Figure 5, shows the occurrence of several signals that were not observed in the spectrum of native SV-IV. The new signals detected at m/z 1063, 1091, 1164, 1219, 1293, 1320, 1401, 1471, and 1515 were assigned to peptides 84–89 or 84–90 cross-linked to various lysine-containing peptides on the basis of their molecular weight; further analysis following a single Edman degradation step confirmed the assignments. Results are summarized in Table I.

This analysis shows that (i) only Gln-86 reacts as an amino acceptor in TGase-mediated modifications of SV-IV when endogenous lysines are the only amino-donor substrates available and (ii) the lysine residues recognized and modified by TGase were Lys-2, Lys-4, Lys-59, Lys-78, Lys-79, and Lys-80 (see Table I). The absence of signals corresponding

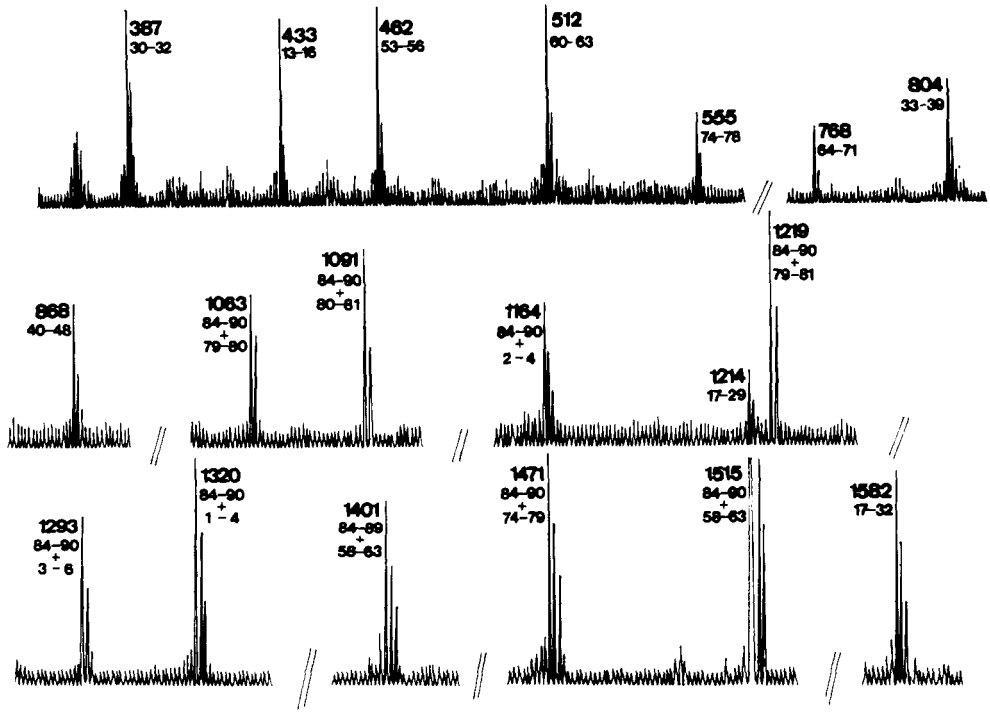

FIGURE 5: FAB mass spectrum of the insoluble protein produced during the incubation of SV-IV with TGase and calcium in the absence of Spd. The relative abundances of the mass signals are on an arbitrary scale, since the gain was varied during the scan.

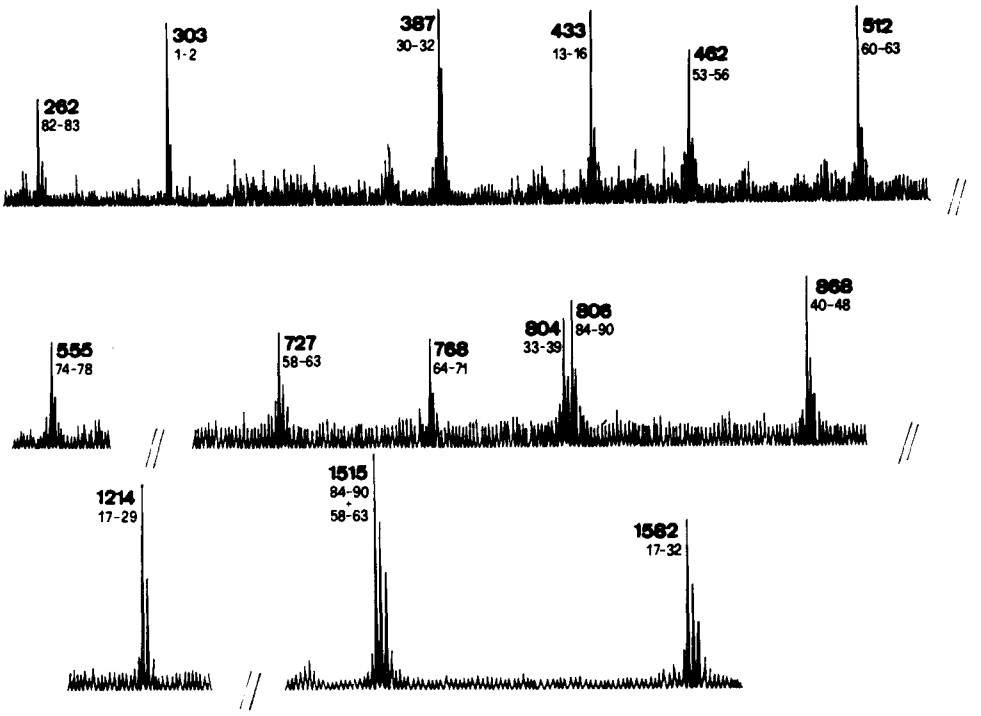

FIGURE 6: FAB mass spectrum of the soluble protein produced during the incubation of SV-IV with TGase and calcium in the absence of Spd. The relative abundances of the mass signals are on an arbitrary scale, since the gain was varied during the scan.

Table I: Cross-Linked Peptides Identified by FABMS Analysis of the Insoluble Material Produced during the Incubation of SV-IV with TGase in the Absence of Spd following Double Digestions with Trypsin and V-8 Protease

MH^+	cross-linked peptides	residues involved in the ε-(γ-glutamyl)-lysine isopeptide bond
1515	(84–90) + (58–63)	Gln-86–Lys-59
1471	(84–90) + (74–79)	Gln-86–Lys-78
1401	(84–89) + (58–63)	Gln-86–Lys-59
1320	(84–90) + (1–4)	Gln-86–Lys-2
1293	(84–90) + (3–6)	Gln-86–Lys-4
1219	(84–90) + (79–81)	Gln-86–Lys-79 or -80
1164	(84–90) + (2–4)	Gln-86–Lys-2
1091	(84–90) + (80–81)	Gln-86–Lys-80
1063	(84–90) + (79–80)	Gln-86–Lys-79

to cross-linked peptides involving Lys-6, -34, and -39 suggests that these residues are not recognized as amino donors by the enzyme under the experimental conditions described.

TGase-Catalyzed Synthesis and Characterization of Intramolecularly ε-(γ-Glutamyl)lysine-Cross-Linked SV-IV. The supernatant recovered from the centrifugation of the reaction mixture containing SV-IV incubated with TGase and Ca^{2+} in the absence of Spd was first analyzed by SDS–PAGE. The electrophoretic pattern (not shown) indicated the presence of a major protein or proteins having a mobility very similar to that of native SV-IV. In order to investigate if this protein fraction was constituted by unreacted SV-IV and/or a TGase-modified form or forms of SV-IV, the soluble material was submitted to FABMS analysis following tryptic and V-8 protease hydrolysis. The resulting mass spectrum (Figure 6) showed the presence of a signal at m/z 1515 that was not present in the spectrum of native SV-IV. This mass value corresponds to peptides 84–90 and 58–63 joined together by an ε-(γ-glutamyl)lysine isopeptide bond involving Gln-86 and Lys-59. Since the SDS-PAGE experiments (see above) indicated that essentially only a monomeric form or forms of SV-IV occurred in the analyzed sample, we concluded that the isopeptide bond detected was an intramolecular cross-link.

Further support for this conclusion came from the FPLC analysis of the soluble material, which showed the occurrence of a major protein peak eluted about one minute before native SV-IV (data not shown). The FABMS analysis of this fraction gave the same results shown in Figure 6, thus confirming the TGase-mediated formation of an intramolecularly cross-linked form of SV-IV.

Discussion

This paper reports the preparation and structural characterization of TGase-modified forms of SV-IV, a protein secreted in large amounts from rat seminal vesicles and thought to be a substrate of prostatic TGase in the formation of the copulatory plug (Williams-Ashman, 1984). Thus, the modified forms of the protein were obtained by treating native SV-IV with purified guinea pig liver TGase, an active molecular form of the enzyme widely distributed intracellularly (Folk, 1980; Williams-Ashman & Canellakis, 1980), and calcium ions both in the presence and absence of Spd. a well-known TGase amino-donor substrate. The choice of Spd was made by considering that large amounts of polyamines occur in rat seminal plasma, where TGase and SV-IV are also secreted (Williams-Ashman & Canellakis, 1980; Williams-Ashman, 1984). The presence of the polyamine in the reaction mixture gives rise to the formation of (γ-glutamyl)-Spd derivatives of the protein, whereas its absence allows SV-IV lysine residues to be recognized by the enzyme as amino-donor sites and to form cross-linked structures of the protein.

The characterization of the TGase-catalyzed reaction products is generally considered a long, laborious task, the procedures commonly used sometimes even giving uncertain results (Griffin & Wilson, 1984; Loewy, 1984). Recently, we suggested to reveal TGase-catalyzed reactions by detecting the products by FABMS. This technique has so far been applied to identify the amino-acceptor and amino-donor sites in two peptides, namely substance P (Porta et al., 1988b) and β-endorphin (Pucci et al., 1988). The present findings confirm

that FABMS allows the identification of the glutamine and lysine residues specifically recognized by TGase, even when the substrate is a polypeptide chain of higher molecular weight.

It is also demonstrated that, when an external amino donor is absent in the reaction mixture, an insoluble intermolecularly cross-linked complex of SV-IV is formed. This complex, containing ε-(γ-glutamyl)lysine isopeptide bonds between one specific glutamine (Gln-86) and six out of nine lysine residues, is most probably an SV-IV branched homopolymer in which two or more lysines of each protein molecule react with the Gln-86 of other molecules. In addition, a second modified form of SV-IV was identified as a product of the TGase-catalyzed reaction carried out in the absence of Spd. It is a soluble monomeric form of the protein, characterized by a single intramolecular ε-(γ-glutamyl)lysine cross-link between Gln-86 and Lys-59. To our knowledge, this is the first time that evidence of the formation of a TGase-catalyzed intramolecular isopeptide bond is reported.

The experiments performed by treating SV-IV with TGase in the presence of Spd led to the synthesis of two different (γ-glutamyl)polyamine derivatives of the protein. The first derivative possessed only one Spd molecule covalently bound to Gln-86, whereas the second one had both Gln-9 and Gln-86 linked to Spd molecules, thus indicating that, under the experimental conditions used, both SV-IV glutamine residues acted as amino acceptors. However, these data and those from the analysis of the polymeric form of SV-IV indicate that TGase exhibits faster kinetics of recognition and modification toward Gln-86. In fact, a single Spd-modified SV-IV form that has Gln-9 involved in the linkage with the polyamine was never observed; furthermore, the analysis of SV-IV modified by TGase in the absence of Spd showed that only Gln-86 was involved in the formation of cross-links with endogenous lysine residues.

Further support for this theory came from the FABMS analysis of the purified CNBr fragment 1–70 of SV-IV, following its incubation with TGase under different experimental conditions (unpublished data). Since this fragment is only 20 amino acids shorter than SV-IV, it should be structurally related to the native protein. Fragment 1–70, where only Gln-9 can act as an amino-acceptor site, was indeed shown to be a TGase substrate in both the presence and the absence of Spd, thus confirming the ability of the enzyme to recognize both glutamine residues in spite of their different kinetics.

Finally, it should be underlined that the structural features of a glutamine residue that is recognized and modified by TGase are still unclear at present. Further work is needed to fully understand all of the aspects of TGase specificity, and FABMS could play a very important role in this field.

References

Barber, M., Bordoli, R. S., Sedgwick, R. D., & Tyler, A. N. (1981) *J. Chem. Soc., Chem. Commun.* 325–327.

Beavis, R. C., & Chait, B. T. (1989) *Rapid Commun. Mass Spectrom. 3*, 432–435.

Biemann, K., & Martin, S. A. (1987) *Mass Spectrom. Rev. 6*, 1–76.

Biemann, K., & Scoble, H. A. (1987) *Science 237*, 992–998.

Chowdhury, S. K., Katta, V., & Chait, B. T. (1990) *Biochem. Biophys. Res. Commun. 167*, 686–692.

Fawell, S. E., Pappin, D. J. C., McDonald, C. F., & Higgins, S. F. (1986) *Mol. Cell. Endocrinol. 45*, 205–213.

Folk, J. E. (1980) *Annu. Rev. Biochem. 49*, 517–531.

Galdiero, F., Tufano, M. A., De Martino, L., Capasso, C., Porta, R., Ravagnan, G., Peluso, G., & Metafora, S. (1989) *J. Reprod. Immunol. 16*, 269–284.

Griffin, M., & Wilson, J. (1984) *Mol. Cell. Biochem. 58*, 37–49.

Harris, S. E., Mansson, P. E., Tully, D. B., & Burkhart, B. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 6460–6464.

Higgins, S. F., Burchell, J. M., & Mainwaring, I. P. (1976) *Biochem. J. 158*, 271–282.

Hunt, D. F., Yates, J. R., III, Shabanowitz, J., Bruns, M. E., & Bruns, D. E. (1989) *J. Biol. Chem. 264*, 6580–6586.

Janes, S. M., Mu, D., Wemmer, D., Smith, A. J., Kaur, S., Maltby, D., Burlingame, A. L., & Klinman, J. P. (1990) *Science 248*, 981–987.

Kandala, J. C., Kistler, M. K., Lawther, R. P., & Kistler, W. S. (1983) *Nucleic Acids Res. 11*, 3169–3186.

Kandala, J. C., Kistler, W. S., & Kistler, M. K. (1985) *J. Biol. Chem. 260*, 15959–15964.

Karas, M., & Hillenkamp, F. (1988) *Anal. Chem. 60*, 2299–2301.

Karas, M., Ingendoh, A., Bahr, U., & Hillenkamp, F. (1989) *Biomed. Environ. Mass Spectrom. 18*, 841–843.

Laemmli, U. K. (1970) *Nature 227*, 680–685.

Loewy, A. G. (1984) *Methods Enzymol. 107*, 241–257.

Loo, J. A., Edmonds, C. G., & Smith, R. D. (1990) *Science 248*, 201–204.

Mann, M., Meng, C. K., & Fenn, J. B. (1989) *Anal. Chem. 61*, 1702–1708.

Mansson, P. E., Sugino, A., & Harris, S. E. (1981) *Nucleic Acids Res. 9*, 935–946.

Metafora, S., Peluso, G., Ravagnan, G., Gentile, V., Fusco, A., & Porta, R. (1987) in *Morphological Basis of Human Reproductive Function* (Spera, G., & Kretser, D. M., Eds.) pp 187–196, Plenum Press, New York.

Metafora, S., Peluso, G., Persico, P., Ravagnan, G., Esposito, C., & Porta, R. (1989a) *Biochem. Pharmacol. 38*, 121–131.

Metafora, S., Porta, R., Ravagnan, G., Peluso, G., Tufano, M. A., De Martino, L., Ianniello, R., & Galdiero, F. (1989b) *J. Leukocyte Biol. 46*, 409–416.

Morris, H. R., & Greer, F. M. (1988) *Trends Biotechnol. 6*, 140–147.

Morris, H. R., Panico, M., & Taylor, G. W. (1983) *Biochem. Biophys. Res. Commun. 117*, 299–305.

Ostrowski, M. C., Kistler, M. K., & Kistler, W. S. (1979) *J. Biol. Chem. 254*, 383–390.

Pan, Y. C. E., & Li, S. S. L. (1982) *Int. J. Pept. Protein Res. 20*, 177–187.

Pan, Y. C. E., Silverberg, A. B., Harris, S. E., & Li, S. S. L. (1980) *Int. J. Pept. Protein Res. 16*, 143–146.

Paonessa, G., Metafora, S., Tajana, G., Abrescia, P., De Santis, A., Gentile, V., Porta, R. (1984) *Science 226*, 852–855.

Porta, R., Esposito, C., Persico, P., Peluso, G., & Metafora, S. (1988a) *Adv. Exp. Med. Biol. 231*, 153–160.

Porta, R., Esposito, C., Metafora, S., Pucci, P., Malorni, A., & Marino, G. (1988b) *Anal. Biochem. 172*, 499–503.

Porta, R., Esposito, C., Siciliano, R., Pucci, P., Ravagnan, G., Peluso, G., & Metafora, S. (1989) in *Biochemical Aspects on the Immunopathology of Reproduction* (Spera, G., Mukherjee, A., Ravagnan, G., & Metafora, S., Eds.) pp 93–103, Acta Medica, Rome.

Porta, R., Esposito, C., Gentile, V., Mariniello, L., Peluso, G., & Metafora, S. (1990) *Int. J. Pept. Protein Res. 35*, 117–122.
Pucci, P., Malorni, A., Marino, G., Metafora, S., Esposito, C., & Porta, R. (1988) *Biochem. Biophys. Res. Commun. 154*, 735–740.
Wagner, C. L., & Kistler, W. S. (1987) *Biol. Reprod. 36*, 501–510.
Williams-Ashman, H. G. (1984) *Mol. Cell. Biochem. 58*, 51–61.
Williams-Ashman, H. G., & Canellakis, Z. N. (1980) *Physiol. Chem. Phys. 12*, 457–472.

Characterization of a Partially Denatured State of a Protein by Two-Dimensional NMR: Reduction of the Hydrophobic Interactions in Ubiquitin†

Margaret M. Harding,‡,§ Dudley H. Williams,*,‡ and Derek N. Woolfson‡,∥
University Chemical Laboratory, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, U.K., and Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, U.K.
Received February 21, 1990; Revised Manuscript Received October 19, 1990

ABSTRACT: A stable, partially structured state of ubiquitin, the A-state, is formed at pH 2.0 in 60% methanol/40% water at 298 K. Detailed characterization of the structure of this state has been carried out by 2D NMR spectroscopy. Assignment of slowly exchanging amide resonances protected from the solvent in the native and A-state shows that gross structural reorganization of the protein has not occurred and that the A-state contains a subset of the interactions present in the native state (N-state). Vicinal coupling constants and NOESY data show the presence of the first two strands of the five-strand β-sheet that is present in the native protein and part of the third β-strand. The hydrophobic face of the β-sheet in the A-state is covered by a partially structured α-helix, tentatively assigned to residues 24–34, that is considerably more flexible than the α-helix in the N-state. There is evidence for some fixed side-chain–side-chain interactions between these two units of structure. The turn-rich area of the protein, which contains seven reverse turns and a short piece of 3_{10} helix, does not appear to be structured in the A-state and is approaching random coil.

Ubiquitin is a cytoplasmic protein of 76 amino acids (8565 Da) found in all eukaryotic cells (Goldstein et al., 1975); it participates in a wide variety of cellular activities (Hershko et al., 1982, 1984; Chin et al., 1982). The X-ray crystal structure of human erythrocytic ubiquitin has been determined (Vijay-Kumar et al., 1985, 1987) and shows the extremely compact nature of the protein with approximately 87% of the polypeptide chain involved in hydrogen-bonded secondary structure (Figure 1) involving a 5-strand mixed β-sheet, 3.5 turns of an α-helix, a short piece of 3_{10} helix, and 7 reverse turns. The β-sheet has a characteristic left-handed twist, and the α-helix fits into the concavity formed by the sheet. There are several unusual features of the structure, namely two β-bulges, one of which is parallel, a highly contorted turn-rich area of 21 residues that contains four reverse turns, and a short piece of 3_{10} helix that forms two interlocked type III reverse turns. ^{1}H NMR[1] assignments of human ubiquitin have been carried out by two groups using sequential (Weber et al., 1987) and main-chain-directed (Di Stefano & Wand, 1987) structure analysis. Both solution structures, which were carried out at pH 4.7 and 5.8, respectively, agree with the solid-state structure. The relatively small size of ubiquitin, the X-ray crystal structure, and the assigned NMR spectrum provide a protein ideal for studying aspects of protein structure and folding by NMR spectroscopy.

Significant advances in protein engineering (Fersht, 1987) and the application of 2D NMR methods to protein structures (Wüthrich, 1989a,b) during the last decade have provided techniques that allow fundamental questions regarding protein structure and folding to be addressed. However, characterization of states other than the native or random-coil state of proteins has met with limited success. A knowledge of partially structured states, intermediate between the two extreme states, should provide insight into the relative strengths and different types of interactions that can stabilize a given section of the polypeptide chain and may provide information regarding the interactions responsible for directing the folding pathways of proteins.

Structural information about states present on the folding pathways of proteins has not generally been directly attainable as folding intermediates are invariably short-lived and not highly populated (Kim & Baldwin, 1982), and it has been necessary to develop methods whereby such kinetic intermediates can be trapped (Creighton, 1978; Biringer & Fink, 1982; Roder et al., 1988; Udgaonkar & Baldwin, 1988). More recently, much attention has been directed toward the so-called

† This work was supported in part by the U.K. Science and Engineering Research Council Molecular Recognition Initiative Programme and Pfizer Central Research, Sandwich. M.M.H. acknowledges receipt of a Ramsay Memorial General British Fellowship.
‡ University Chemical Laboratory.
§ Present address: Department of Organic Chemistry, University of Sydney, Broadway, NSW, 2006 Australia.
∥ Department of Biochemistry.

[1] Abbreviations: NMR, nuclear magnetic resonance; 1D and 2D, one and two dimensional; COSY, two-dimensional spin correlated spectroscopy; DQF, double-quantum filtered; NOE, nuclear Overhauser effect; NOESY, two-dimensional nuclear Overhauser enhancement spectroscopy; HOHAHA, two-dimensional homonuclear Hartmann–Hahn spectroscopy; DANTE, solvent suppression pulse sequence; CD, circular dichroism; MLEV, two-dimensional isotropic mixing experiments employing MLEV-17 or modified MLEV-17 mixing pulses; N-state, native state of the protein; U-state, unfolded state of the protein; A-state, partially folded state of the protein.

0006-2960/91/0430-3120$02.50/0