THE COMPLETE AMINO ACID SEQUENCE AND THE TRYPSIN REACTIVE (INHIBITORY) SITE OF THE MAJOR PROTEINASE INHIBITOR FROM THE FRUITS OF AUBERGINE (*SOLANUM MELONGENA* L.)

M. RICHARDSON
Department of Botany, University of Durham, Durham City, DH1 3LE, England

Received 4 June 1979

1. Introduction

Plants of the Solanaceae are known to be a rich source of many inhibitors of proteolytic enzymes [1,2]. Although the complete amino acid sequences of the chymotryptic inhibitor I [3–5] and a carboxypeptidase inhibitor [6] from potato tubers have been determined together with the partial sequences of several other inhibitors from this tissue [7–9], there are as yet no details of the structures of any inhibitors from other members of the Solanaceae.

Kanamori et al. [10] first reported the occurrence of a proteinase inhibitor in the exocarps of eggplants (aubergine, *Solanum melongena* L.). The inhibitor was subsequently shown to be a small protein (mol. wt 5300–6200) which inhibited the serine proteinases, and existed in the form of five isoinhibitors which were separable by ion-exchange column chromatography and isoelectrofocusing [11–14]. The present paper reports the elucidation of the complete amino acid sequence and the reactive (inhibitory) site of the major isoinhibitor (pI 4.7).

2. Materials and methods

2.1. *Protein purification*

The proteinase inhibitor was purified from aubergine fruits (obtained commercially) by the methods described previously [15]. The final product gave a single band when examined by disc electrophoresis on 7.5% polyacrylamide gels at pH 8.3 [16] and after isoelectric focusing [17]. The yield of the purified main inhibitor with a pI of 4.7 was 155 mg from 20 kg fresh wt. of fruit.

2.2. *Sequence determination*

A sample (20 mg) of the reduced and *S*-carboxymethylated [18] protein was unblocked at its N-terminus by digestion with 4% (w/w) pyroglutamate aminopeptidase (calf liver, Boehringer Mannheim Biochemicals) for 9 h at 4°C, followed by a further 14 h at 20°C as described in [19].

The unblocked protein was digested with 2% (w/w) trypsin as described previously [3]. Other samples (20 mg) of the reduced and *S*-carboxymethylated protein were digested separately with chymotrypsin, thermolysin, trypsin [3] and with the protease from *Staphylococcus aureus* strain V8 (Miles Laboratories Ltd.) at pH 7.8 as described in [20].

Mixtures of peptides were fractionated by gel filtration on a column (1 × 190 cm) of Biogel P-2 in 0.05 M pyridine acetate buffer at pH 5.4, high-voltage paper electrophoresis and paper chromatography [3].

Peptides and samples of the unblocked protein were subjected to micro-sequence analysis using the 4-*NN*-dimethylaminoazobenzene-4′-isothiocyanate (DABITC)/phenylisothiocyanate double coupling method [21]. The DABITC was prepared as described in [22,23]. The amino acid sequences of peptides were also determined by the dansyl-Edman procedure and by digestion with carboxypeptidase A [3] or with carboxypeptidase Y (Sigma Chem. Co.) as described in [24]. Amino acid analyses were obtained using a Locarte amino acid analyzer.

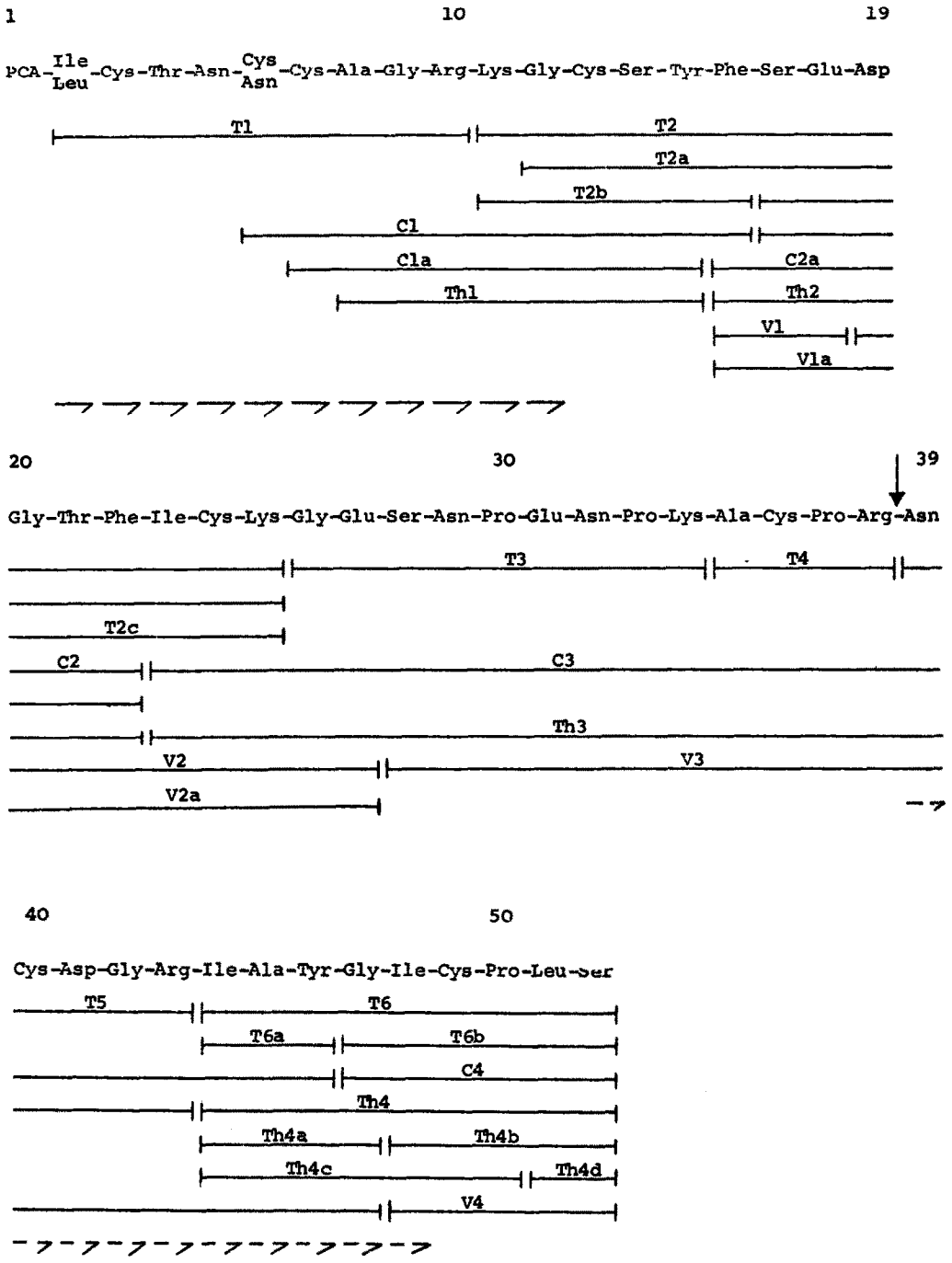

Fig.1. The amino acid sequence of the major proteinase inhibitor from the fruit of aubergine (*Solanum melongena* L.). PCA = pyrrolidone-carboxylic acid, T = tryptic peptides, C = chymotryptic peptides, Th = thermolytic peptides, V = peptides from digestion with *S. aureus* V8 protease; —ⁿ results from DABITC method applied to unblocked (treated with pyroglutamate aminopeptidase) protein; – ⁿ results from DABITC method applied to fragment released by prolonged digestion of native protein with trypsin at pH 2.2; ↓ trypsin reactive (inhibitory) site.

2.3. *Identification of the reactive site*

A sample (12 mg) of the native inhibitor was subjected to limited hydrolysis with a catalytic amount (0.25 mg) of trypsin in 0.2 M acetic acid, containing 0.04 M $CaCl_2$ and adjusted to pH 2.2 with HCl for 24 h. The modified fragments which resulted were purified as described in [25]. The inhibitory activity of native and modified inhibitors against trypsin was determined by measuring the hydrolysis of α-*N*-benzoyl-D,L-arginine-*p*-nitroanilide (Sigma Chem. Co.) at pH 8 as described in [10].

3. Results and discussion

No free N-terminal amino acid could be detected in the native or reduced and *S*-carboxymethylated inhibitor when it was examined by either the dansyl or DABITC method, confirming the earlier report [26] that the terminal amino acid was blocked. Following treatment of the protein with pyroglutamate aminopeptidase for 23 h both isoleucine and leucine were detected as N-terminal amino acids indicating that the native protein had pyrrolidone-carboxylic acid (PCA) as its N-terminal residue and exhibited microheterogeneity in position 2. Further steps of either the dansyl-Edman or the DABITC method yielded the N-terminal sequence of amino acids as shown in fig.1.

Consideration of the amino acid sequences of the peptides derived from the inhibitor by digestion with trypsin, chymotrypsin, thermolysin and the *S. aureus* V8 protease enabled the complete structure of the

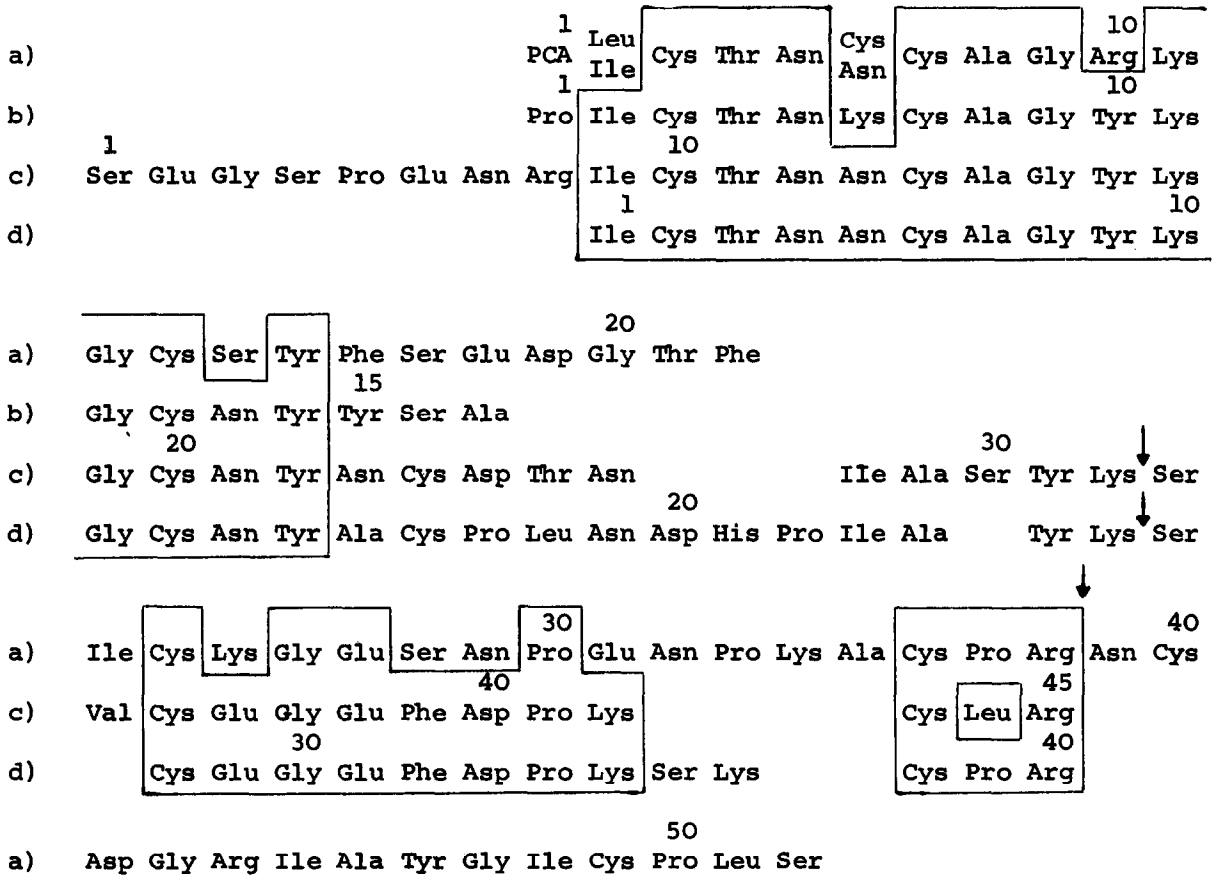

Fig.2. Comparison of the amino acid sequences of (a) the major proteinase inhibitor from aubergine (*Solanum melongena* L.) with (b) the N-terminal region of a chymotryptic inhibitor [7] and (c,d) the active fragments of the proteinase inhibitors IIa [8] and IIb [9] from potatoes (*Solanum tuberosum* L.). Segments of identical sequences are enclosed in boxes; ↓ the trypsin reactive sites.

protein to be deduced as shown in fig.1. In addition to the microheterogeneity observed in position 2 both asparagine and cysteine were found in equal amounts in position 6.

The highly sensitive DABITC micro-sequence method was particularly useful in determining the structure of several of the larger peptides (C3, Th3, and V3) whose size was beyond the limitations of the conventional dansyl-Edman procedure. In addition the method permitted the direct identification of the various acid/amide residues in these peptides which could not be unambiguously assigned by the dansyl-Edman method and consideration of the electrophoretic mobilities of the peptides.

Minor examples of anomalous cleavages by trypsin were observed at the peptide bonds between Phe^{6}–Ser^{7} and Tyr^{46}–Gly^{47}. These were assumed to be due to traces of chymotrypsin contaminating the preparation of trypsin used. The specificity of the digestion by the *S. aureus* protease was as expected, except for the cleavage at Tyr^{15}–Phe^{16}, and the failure to hydrolyse the peptide bond between Glu^{31}–Asn^{32}.

Comparison of the sequence of the aubergine inhibitor with the N-terminal sequence of the chymotrypsin inhibitor [7] and the corresponding regions of the active fragments of the proteinase inhibitors IIa [8] and IIb [9] from potatoes (fig.2) reveals a certain degree of homology, but this similarity does not extend to the regions containing the reactive (inhibitory) sites of the proteins. In the case of the aubergine inhibitor the reactive site was identified as the Arg^{38}–Asn^{39} peptide bond as this was the only bond cleaved during prolonged digestion of the native protein with catalytic amounts of trypsin at pH 2.2, and subsequent treatment of the modified inhibitor with carboxypeptidase B abolished the inhibitory activity towards trypsin.

It is perhaps surprising that the amino acid sequence of the aubergine inhibitor reported here does not contain an Arg–Ser peptide bond as this has been previously reported to be the trypsin reactive site of this inhibitor [26]. On the other hand, it is interesting to note that the sequence of amino acid residues (Ala^{35}–Cys–Pro–Arg–Asn^{39}) surrounding the trypsin reactive site in the aubergine inhibitor is identical with the corresponding regions of the ovoinhibitors from Japanese quail and chicken and the proteinase inhibitor from chicken serum [27].

Acknowledgements

I thank Mr J. Gilroy for invaluable technical assistance and Professor D. Boulter for the provision of certain facilities.

References

[1] Ryan, C. A. (1973) Ann. Rev. Plant Physiol. 24, 173–196.
[2] Richardson, M. (1977) Phytochemistry 16, 159–169.
[3] Richardson, M. (1974) Biochem. J. 137, 101–112.
[4] Richardson, M. and Cossins, L. (1974) FEBS Lett. 45, 11–13.
[5] Richardson, M. and Cossins, L. (1975) FEBS Lett. 55, 161.
[6] Hass, G. M., Nau, H., Biemann, K., Grahn, D. T., Ericsson, L. H. and Neurath, H. (1975) Biochemistry 14, 1334–1342.
[7] Hass, G. M., Venkatakrishnan, R. and Ryan, C. A. (1976) Proc. Natl. Acad. Sci. USA 73, 1941–1944.
[8] Iwasaki, T., Kiyohara, T. and Yoshikawa, M. (1976) J. Biochem. 79, 381–391.
[9] Iwasaki, T., Wada, J., Kiyohara, T. and Yoshikawa, M. (1977) J. Biochem. 82, 991–1004.
[10] Kanamori, M., Ibuki, F., Tashiro, M., Yamada, M. and Miyoshi, M. (1975) J. Nutr. Sci. Vitaminol. 21, 421–428.
[11] Kanamori, M., Ibuki, F., Yamada, M., Tashiro, M. and Miyoshi, M. (1975) J. Nutr. Sci. Vitaminol. 21, 429–436.
[12] Kanamori, M., Ibuki, F., Yamada, M., Tashiro, M. and Miyoshi, M. (1976) Agric. Biol. Chem. 40, 839–844.
[13] Kanamori, M., Ibuki, F., Tashiro, M., Yamada, M. and Miyoshi, M. (1976) Biochim. Biophys. Acta 439, 398–405.
[14] Yamada, M., Santo, S., Yamaguchi, H., Tashiro, M., Ibuki, F. and Kanamori, M. (1977) Agric. Biol. Chem. 41, 2343–2347.
[15] Ibuki, F., Yamada, M., Tashiro, M. and Kanamori, M. (1977) J. Nutr. Sci. Vitaminol. 23, 133–143.
[16] Clarke, J. T. (1964) Ann. NY Acad. Sci. 121, 428–436.
[17] Wrigley, C. (1968) Sci. Tools 15, 17–23.
[18] Crestfield, A. M., Moore, S. and Stein, W. H. (1963) J. Biol. Chem. 238, 622–627.
[19] Podell, D. N. and Abraham, G. N. (1978) Biochem. Biophys. Res. Commun. 81, 176–185.
[20] Houmard, J. and Drapeau, G. R. (1972) Proc. Natl. Acad. Sci. USA 69, 3506–3509.
[21] Chang, J. Y., Brauer, D. and Wittmann-Liebold, B. (1978) FEBS Lett. 93, 205–214.

[22] Chang, J. Y., Creaser, E. H. and Bentley, K. W. (1976) Biochem. J. 153, 607–611.
[23] Chang, J. Y. (1977) Biochem. J. 163, 517–520.
[24] Hayashi, R., Moore, S. and Stein, W. H. (1973) J. Biol. Chem. 248, 2296–2302.
[25] Richardson, M., Barker, R. D. J., McMillan, R. T. and Cossins, L. (1977) Phytochemistry 16, 837–839.
[26] Yamada, M., Tashiro, M., Yamaguchi, H., Yamada, H., Ibuki, F. and Kanamori, M. (1976) J. Biochem. 80, 1293–1297.
[27] Laskowski, M., jr, Kato, I. and Kohr, W. J. (1978) in: Versatility of Proteins (Li, C. H. ed) pp. 307–318, Academic Press, London.