The complete amino acid sequence of the α-amylase inhibitor I-2 from seeds of ragi (Indian finger millet, *Eleusine coracana* Gaertn.)

F.A.P. Campos and M. Richardson*

Department of Biochemistry and Molecular Biology, Science Centre, Federal University of Ceará, Fortaleza (CE), Brazil and **Department of Botany, University of Durham, Durham DH1 3LE, England*

Received 30 December 1983

The complete amino acid sequence of the α-amylase inhibitor I-2 from ragi seeds was determined by analysis of peptides derived from the protein by digestion with trypsin, chymotrypsin thermolysin and *S. aureus* V8 protease. The protein consists of a single polypeptide chain of 95 amino acids, with a high content of serine and alanine, and no methionine, phenylalanine, histidine or tryptophan. There is no sequence homology with the bifunctional α-amylase/trypsin inhibitor in the same seeds or with any of the α-amylase inhibitors from other plants. The sequence contains two regions of weakly repetitive internal homology. The predicted secondary structure of the inhibitor is notable for the absence of α-helix and its high content (50%) of β-turn.

Ragi seed *α-Amylase inhibitor* *Amino acid sequence* *Internal homology* *Predicted secondary structure*

1. INTRODUCTION

We recently reported the primary structure of an unusual bifunctional trypsin/α-amylase inhibitor from seeds of ragi (Indian finger millet, *Eleusine coracana* Gaertn.) [1] which exhibited strong sequence homology with an α-amylase inhibitor from wheat [2,3] and with trypsin inhibitors from barley [4,5] and maize [6]. Our suggestion [1] of the existence of a new family of inhibitors active against different classes of enzymes has been confirmed subsequently by other studies which have indicated possible evolutionary relationships between these proteins and the bovine pancreatic secretory trypsin (Kazal) inhibitor [7], the castor bean storage protein [8] and another wheat α-amylase inhibitor [9].

The only other bifunctional inhibitor which has been studied in any detail is the α-amylase/subtilisin inhibitor recently isolated from barley [10]. Preliminary studies of its N-terminal amino acid sequence [11] have revealed that whilst it has little or no similarity with the ragi inhibitor family, it does have strong homology with a rice subtilisin inhibitor [12] and weaker homology with the Kunitz-type protease inhibitors from legume seeds [13].

This evidence for the existence of two separate families of proteins both inactivating α-amylases and other unrelated enzymes prompted us to examine the structure of the second α-amylase inhibitor (I-2) originally isolated and purified from ragi seeds [14,15]. We now report the complete amino acid sequence of this protein together with its predicted secondary structure.

2. MATERIALS AND METHODS

2.1. *Materials*

Seeds of ragi (*Eleusine coracana* Gaertn.) were obtained as a generous gift from the Tropical Products Institute (Culham, Oxford). The α-amylase inhibitor I-2 was purified from finely milled seeds

* To whom correspondence should be addressed

Published by Elsevier Science Publishers B.V.
00145793/84/$3.00

essentially as in [14,15]. Trypsin-TPCK, α-chymotrypsin and the *S. aureus* V8 protease were obtained from Worthington Biochem. Corp. (NJ), thermolysin from the Daiwa Kasei KK (Osaka) and carboxypeptidase A from Sigma.

2.2. *Enzyme digestions and separation of peptides*

Samples (10–20 mg) of the reduced and *S*-carboxymethylated protein were digested separately with trypsin, chymotrypsin, thermolysin, *S. aureus* V8 protease and carboxypeptidase A as in [1]. Mixtures of peptides produced by these methods were initially fractionated by chromatography on a column (1 × 200 cm) of Biogel P-4 in 0.05 M ammonium bicarbonate (pH 8.1). Further purification of peak fractions was achieved by reverse-phase HPLC as in [1].

2.3. *Sequence determination*

The intact reduced and *S*-carboxymethylated inhibitor and the peptides derived from it were subjected to micro-sequence analysis using the 4-*N,N*-dimethylaminoazobenzene-4′-isothiocyanate (DABITC)/phenylisothiocyanate double coupling method and the dansyl-Edman procedure as in [1].

2.4. *Amino acid analysis*

The amino acid composition of the protein and peptides was determined in a Varian 5000 HPLC fitted with a Micropak hydrolysate amino acid (AA) column (4 mm × 15 cm, Varian, Walnut Creek) in the sodium form, a post-column *o*-phthalaldehyde reactor system and a Fluorichrom detector. The buffer system used was the pHIX picobuffer system II (Pierce) and the post-column

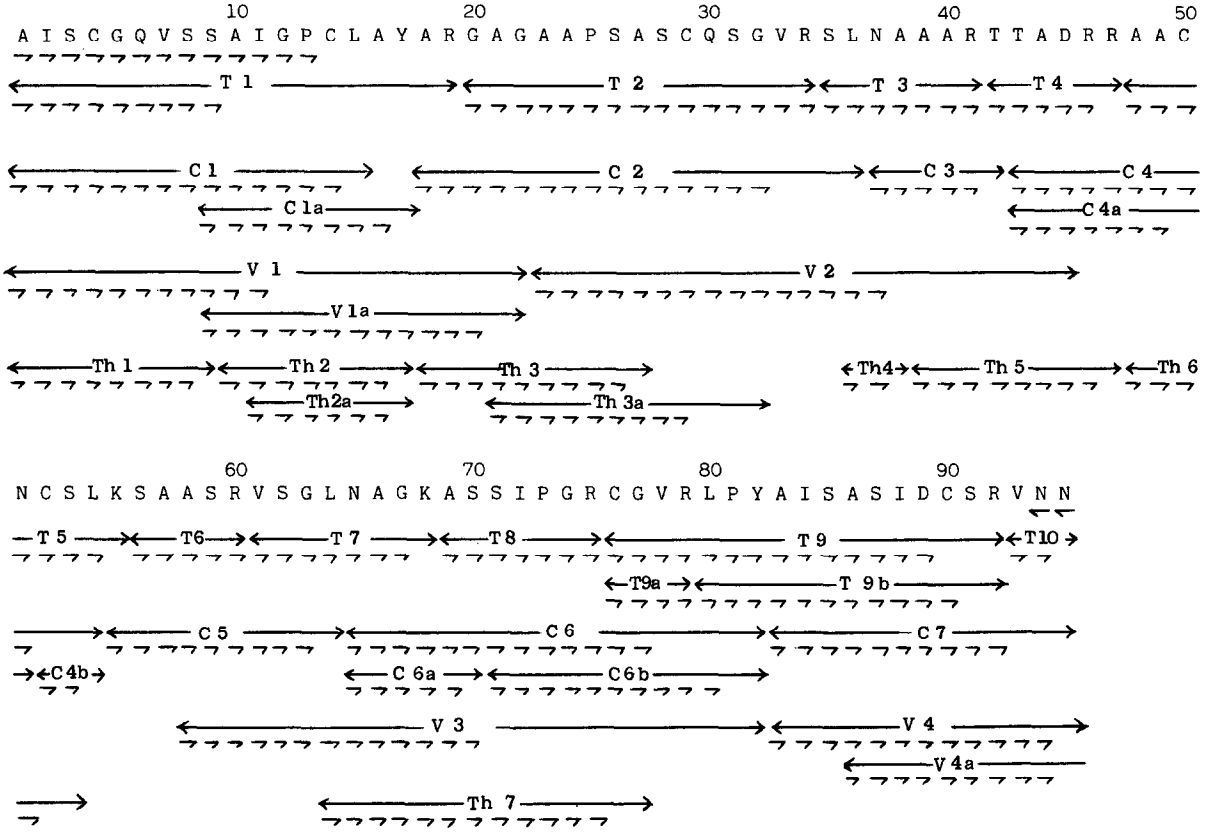

Fig. 1. The amino acid sequence of the α-amylase inhibitor I-2 from seeds of ragi (Indian finger millet, *Eleusine coracana* Gaertn.) T, tryptic peptides; C, chymotryptic peptides; V, peptides from digestion with *S. aureus* V8 protease; Th, thermolytic peptides. (—) Amino acids identified by DABITC micro-sequence method and/or dansyl-Edman method. (—) Residues identified by carboxypeptidase A digestion.

reaction system was as described in the operator's manual (Varian Associates, 1981).

2.5. *Sequence comparisons*

The statistical significance of various alignments of possibly homologous amino acid sequences was tested by using a computer program developed from Gotoh's algorithm [16].

2.6. *Prediction of secondary structure*

The secondary structure of the inhibitor was predicted from the amino acid sequence by using the computerized method in [17] as described in [18].

3. RESULTS AND DISCUSSION

The homogeneity of inhibitor I-2 was confirmed by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) [19], isoelectric focussing [20] and by N-terminal analysis by the dansyl-chloride and by the DABITC/PITC double coupling methods. The final yield of the pure inhibitor was 34 mg/kg ragi flour.

The complete amino acid sequence of inhibitor I-2 is shown in fig.1 together with the details of the peptides from which it was deduced. The protein contains 95 amino acids which corresponds to an M_r of 9333. This is in agreement with our estimate of 10 200 by SDS-PAGE and the previous estimate of 8300 from gel filtration [14]. The sequence is also in agreement with the amino acid composition of the protein (table 1) except for the small discrepancies in the values for Asp and Ala.

The sequence is notable for its unusually high content of Ser and Ala, which is unique amongst the extant α-amylase inhibitors, although relatively high levels of Ser have been reported in the Bowman-Birk family of 'double-headed' proteinase inhibitors from various legume seeds [21]. The inhibitor I-2 sequence is devoid of Met, Trp, His and Phe. The latter two amino acids are also missing in the 0.28 inhibitor of α-amylase from wheat [2,3] and present only in low amounts in the 0.53 form [9]. Isoelectric focussing indicating a high p*I* value (above 10) for inhibitor I-2 which reflects its high content of basic residues (9 Arg, 2 Lys) relative to acidic residues (2 Asp). There was no evidence of heterogeneity in any of the residues in inhibitor I-2, unlike that reported for the bifunctional α-amylase/trypsin inhibitor from the same material [1].

Table 1

The amino acid composition of the α-amylase inhibitor I-2 from seeds of ragi (Indian finger millet, *Eleusine coracana*)

Amino acid	Analysis	Sequence
Cm-Cys	6.59	7
Asp	8.29	7
Thr	1.90[a]	2
Ser	15.22[a]	16
Glu	2.53	2
Pro	ND	4
Gly	8.68	9
Ala	17.21	20
Val	4.03[b]	5
Met	0.10	0
Ile	4.46[b]	5
Leu	5.23	5
Tyr	1.68	2
Phe	0.14	0
Lys	2.03	2
His	0.21	0
Arg	8.66	9
Trp	ND	0

[a] Values obtained by extrapolation to zero hydrolysis time

[b] Values obtained from 48 h hydrolysis

ND, not determined

Most of the peptide bonds in inhibitor I-2 which were hydrolyzed were as expected from the usual specificities of the enzymes employed, however, there were some notable exceptions. Chymotrypsin produced anomalous cleavages of the Ser^8-Ser^9 and Ser^{70}-Ser^{71} bonds. The *S. aureus* V8 protease also hydrolyzed Ser^8-Ser^9, together with the Gly^{22}-Ala^{23}, Ala^{57}-Ala^{58} and Tyr^{82}-Ala^{83} bonds. Such non-specific digestion on the C-terminal side of Gly, Ala and Tyr by the *S. aureus* enzyme has been reported in a number of other proteins [22,23].

There is no homology between the sequence of the ragi α-amylase inhibitor I-2 presented here and that of the bifunctional trypsin/α-amylase inhibitor from the same seeds [1]. We are also unable to detect any similarities with the other α-amylase inhibitors which have been sequenced so far

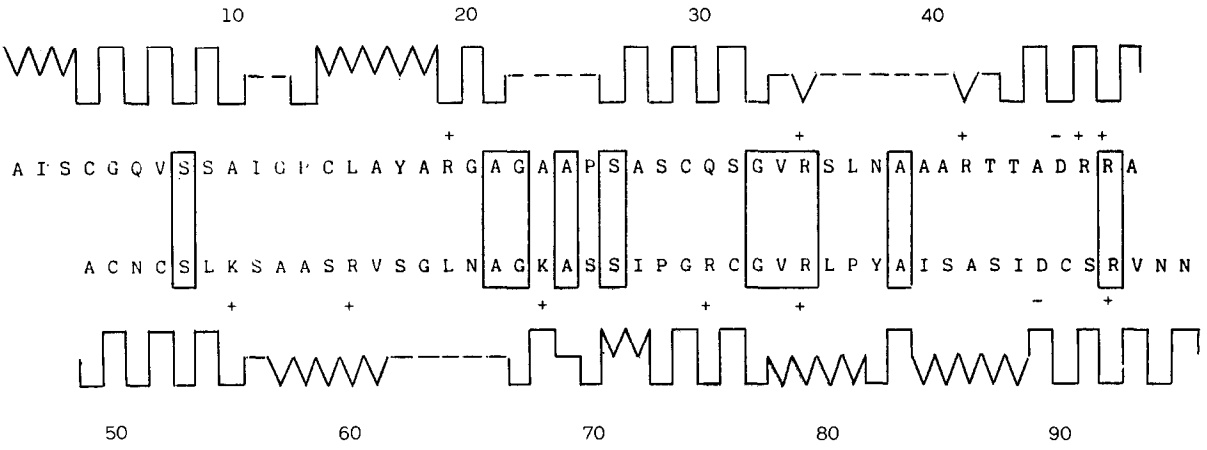

Fig. 2. Alignment of the primary and secondary structures of the α-amylase inhibitor I-2 from seeds of ragi to show the weakly homologous regions. Identical residues in the sequence are linked by vertical boxes. The secondary structure computed as in [18] is shown as W, β sheet; ($\sqcap\!\!\sqcup$) turn structure; (– – –) random coil (there is no α-helix in predicted structure).

[2,3,9,11,24], any of the related inhibitors of proteinases [6,7,11,12], or any other plant protein [25]. The molecule of α-amylase inhibitor I-2 does however contain two regions of weakly repetitive sequence. When the residues between 8 and 47 are compared with those between 53 and 92 (fig.2) there are 10 pairs of matching amino acids out of 40. This observation suggested the possibility that ancestral gene duplication [26] might have occurred during the evolution of inhibitor I-2, as has been suggested for a number of other inhibitors of enzymes [13,21]. However statistical analysis of the alignment of these two regions by the method in [16] to see if the similarity was indicative of homology, i.e., common ancestory, gave an indecisive result as the value of z obtained (-2.893) was just below the value where the hypothesis of homology is accepted [27].

The secondary structure of the inhibitor predicted from the amino acid sequence by using the computerized method in [17] as described in [18] is shown in fig.2. This directional method, presented as an algorithm, was chosen because of its higher rate of predictive success than other methods [28]. The predicted secondary structure is completely lacking in α-helix and has a high content (50%) of β-turn, which at least in part, probably accounts for the relatively high thermal stability of the inhibitor noted previously [14,15].

We were unable to detect any similarities in secondary structure when comparing inhibitor I-2 with the ragi bifunctional trypsin/α-amylase inhibitor or any other α-amylase inhibitors so far sequenced (unpublished). On the other hand comparison of the secondary structures of the two regions of weakly repetitive sequence within inhibitor I-2 does provide further evidence for the concept that they may have arisen by a process of gene duplication. As may be seen in fig.2 there is considerable similarity in the secondary structures of these two regions particularly with respect to the distribution of β-turns. Other workers [29,30] have previously reported a high degree of β-turn conservation in homologous proteins and interpreted their results as an indication that chain reversal regions play an important role in keeping intact the active structural domains of proteins.

ACKNOWLEDGEMENTS

The authors wish to thank the Science and Engineering Research Council, The CNP (Brazilian National Council for Scientific and Technological Development), the CAPES (Brazilian Ministry of Education), the British Council and the Smith Kline Foundation for financial support. We also thank Mr J. Gilroy, Mr D. Bown and Mr P. Miller for invaluable technical

assistance, Mr H. Liesenberg and Dr J. Gatehouse for helpful advice and Professor D. Boulter for the provision of certain facilities.

REFERENCES

[1] Campos, F.A.P. and Richardson, M. (1983) FEBS Lett. 152, 300-304.
[2] Kashlan, N. and Richardson, M. (1981) Phytochemistry 20, 1781-1784.
[3] Richardson, M. and Kashlan, N. (1981) in: Regulators of Intestinal Absorption in Obesity, Diabetes and Nutrition (Berchtold, P. et al. eds) pp. 103-126, Societa Editrice Universo, Rome.
[4] Odani, S., Koide, T. and Ono, T. (1982) FEBS Lett. 141, 279-282.
[5] Odani, S., Koide, T. and Ono, T. (1983) J. Biol. Chem. 285, 7998-8003.
[6] Hochstrasser, K., Illchmann, K. and Werle, E. (1970) Hoppe-Seyler's Z. Physiol. Chem. 351, 721-728.
[7] Odani, S., Koide, T. and Ono, T. (1983) J. Biochem. (Tokyo) 93, 1701-1704.
[8] Odani, S., Koide, T., Ono, T. and Ohnishi, K. (1983) Biochem. J. 213, 543-545.
[9] Maeda, K., Hase, T. and Matsubara, H. (1983) Biochim. Biophys. Acta 743, 52-57.
[10] Mundy, J., Svendsen, I. and Hejgaard, H. (1983) Carlsberg Res. Commun. 48, 81-90.
[11] Hejgaard, J., Svendsen, I. and Mundy, J. (1983) Carlsberg Res. Commun. 48, 91-94.
[12] Kato, I., Tominaga, N. and Kihara, F. (1972) in: Proceedings of the 23rd Conference on Proteins Structure (Iwai, K. ed) pp. 53-56.
[13] Laskowski, M. jr and Kato, I. (1980) Annu. Rev. Biochem. 49, 593-626.
[14] Shivaraj, B. and Pattabiraman, T.N. (1980) Ind. J. Biochem. Biophys. 17, 181-185.
[15] Shivaraj, B. and Pattabiraman, T.N. (1981) Biochem, J. 193, 29-36.
[16] Gotoh, O. (1982) J. Mol. Biol. 162, 705-708.
[17] Robson, B. (1974) Biochem. J. 141, 853-867.
[18] Garnier, J., Osguthorpe, D.J. and Robson, B. (1978) J. Mol. Biol. 120, 97-120.
[19] Laemli, U.K. (1970) Nature 227, 680-685.
[20] Wrigley, C. (1968) Sci. Tools 15, 17-23.
[21] Richardson, M. (1980) Food Chem. 6, 235-253.
[22] Dognin, M.J. and Wittman-Liebold, B. (1977) FEBS Lett. 84, 342-346.
[23] Keil, B. (1982) in: Methods in Protein Sequence Analysis (Elzinga, M. ed) pp. 291-304, Humana Press, NJ.
[24] Aschauer, H., Vertesy, L. and Braunitzer, G. (1981) Hoppe Seyler's Z. Physiol. Chem. 362, 465-467.
[25] Ramshaw, J.A.M. (1982) Encycl. Plant Physiol. (new series) 14A, 229-290.
[26] Doolittle, R.F. (1981) Science 214, 149-159.
[27] Dayoff, O., Barker, W.C. and Hunt, L.T. (1983) Methods Enzymol. 91, 524-545.
[28] Kabsch, W. and Sander, C. (1983) FEBS Lett. 155, 179-182.
[29] Chou, P.Y. and Fasman, G.D. (1979) Biophys. J. 26, 385-400.
[30] Tlomak, P. and Nowak, K. (1981) Acta Biochim. Pol. 28, 241-251.