The complete amino acid sequence of the bifunctional α-amylase/trypsin inhibitor 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 City, DH1 3LE, England*

Received 3 January 1983

The complete amino acid sequence of the bifunctional α-amylase/trypsin inhibitor from ragi seeds was determined by analysis of fragments and peptides derived from the protein by cleavage with cyanogen bromide and by digestion with trypsin, chymotrypsin, thermolysin, the *S. aureus* V8 protease and a Pro-specific protease. The molecule consists of a single polypeptide chain of 122 amino acids which exhibits sequence homology with trypsin inhibitors from barley and maize and with an α-amylase inhibitor from wheat.

Ragi seed *Bifunctional α-amylase/trypsin inhibitor* *Complete amino acid sequence* *Reactive (trypsin inhibitory) site* *Homology to other enzyme inhibitor*

1. INTRODUCTION

Protein inhibitors of enzymes are known to be of widespread occurrence in the plant kingdom [1,2]. The primary structures of numerous inhibitors of different classes of proteinases have been determined [3], and this information has revealed the existence of a number of families of inhibitors whose members are related by extensive homology and topological relationships between their disulphide bridges and the location of the reactive (inhibitory) sites [4]. However, as yet there is only one complete amino acid sequence of an inhibitor of an enzyme other than a proteinase that is the α-amylase inhibitor from wheat [5].

Recently an unusual bifunctional protein capable of inhibiting both α-amylase and trypsin has been purified to homogeneity from seeds of ragi (Indian finger millet) [6]. This paper presents the complete amino acid sequence of the ragi bifunctional inhibitor which exhibits sequence homology not only with the α-amylase inhibitor from wheat [5] but also with trypsin inhibitors from barley [7,8] and maize [9], and thus clearly establishes the existence in cereals of a new family of inhibitors active against entirely unrelated enzymes.

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 bifunctional inhibitor was purified from the finely milled seeds essentially as described previously [6]. Trypsin-TPCK, α-chymotrypsin and the *S. aureus* V8 protease were obtained from Worthington Biochem. Corp. (NJ); thermolysin from Daiwa Kasei KK, Osaka, Japan; a proline specific endoproteinase (EC 3.4.21.26) from Seikagaku Kogyo Co., Tokyo, and carboxypeptidase A from Sigma Chem. Co.

* To whom correspondence should be addressed

Published by Elsevier Biomedical Press
00145793/83/0000–0000/$3.00

2.2. *Cleavage of the inhibitor with CNBr*

The inhibitor (20 mg) was reduced and *S*-carboxymethylated as in [10]. The product was then treated with CNBr (150-fold molar excess over MET residues) in 2.0 ml 70% formic acid for 24 h at 20°C. The resulting fragments were fractionated on a column (1 × 200 cm) of Biogel P-4 in 0.05 M pyridine/acetate buffer at pH 5.4 and further purified by reverse phase HPLC in a Varian model 5000 HPLC fitted with a μ-Bondapack C-18 column (0.5 × 25 cm) (from HPLC Technology Ltd., Macclesfield, Cheshire), and using a linear gradient of 0 to 70% acetonitrile (HPLC grade S, Rathburn Chem. Co., Peebles, Scotland) in 0.1% trifluoroacetic acid (sequencer grade, Rathburn Chem. Co.) [11].

2.3. *Enzyme digestions and separation of peptides*

The fragments resulting from cleavage with CNBr, and other samples (10–20 mg) of the intact reduced and *S*-carboxymethylated protein were digested separately with trypsin, chymotrypsin, thermolysin and carboxypeptidase A as described in [12], with the protease from *S. aureus* strain V8 as in [13] and with the proline-specific endoproteinase (2% w/w) in 0.1 M phosphate buffer, pH 7.0. Mixtures of peptides obtained by these methods were initially fractionated by chromatography on a column (1 × 200 cm) of Biogel P-4 in 0.05 pyridine/acetate buffer, pH 5.4. Further purification of peak fractions was achieved by reverse phase HPLC as described above. In the case of the Pro-specific enzyme digests the peptides required were purified by HPLC without prior fractionation.

2.4. *Sequence determination*

The intact reduced and *S*-carboxymethylated inhibitor, CNBr fragments and peptides 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 described previously [14]. Amino acid analyses were obtained using a Locarte amino acid analyzer.

2.5. *Identification of the reactive (trypsin inhibitory) site*

A sample (20 mg) of the native inhibitor was subjected to limited hydrolysis with a catalytic amount (0.5 mg) of trypsin in 0.2 M acetic acid containing 0.04 M $CaCl_2$ and adjusted to pH 2.5 with HCl for 20 h at 37°C. After reduction and *S*-carboxymethylation the modified fragments which resulted were purified by column chromatography and HPLC as described above. The inhibitory activity of the native inhibitor 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 [15]. The concentration of active trypsin used was estimated by titration of the active site with *p*-nitrophenyl-*p*′-nitrophenyl-*p*′-guanidinobenzoate HCl as described in [16].

3. RESULTS AND DISCUSSION

The purified ragi inhibitor gave a single band when examined by disc electrophoresis on 10% polyacrylamide gels at pH 8.3 [17] and also after isoelectric focussing [18] which indicated that it was a basic protein with a pI of above 10. The protein was also homogeneous in its amino-terminal sequence as determined by both the DABITC and dansyl–Edman methods.

The complete amino acid sequence of the inhibitor is shown in fig.1 together with the details of the CNBr fragments and the overlapping peptides from which it was deduced. The protein contains 122 amino acids which corresponds to an approximate M_r of 13300. This is in marked contrast to the previous estimates made by other workers of 9000 and 14300 based on gel permeation chromatography and SDS polyacrylamide gel electrophoresis respectively [6]. The discrepancy probably results from the anomalous chromatographic behaviour of proteins with such a basic pI [19].

Treatment of the reduced and *S*-carboxymethylated protein with an excess of CNBr gave rise to five fragments indicating that in addition to cleavage at the expected 3 Met residues an anomalous cleavage had also occurred at the single TRP residue as has been reported previously [5]. This was confirmed by sequence analysis.

Approximately equal amounts of two forms of fragment CNBr3 were separated by reverse phase HPLC. Subsequent sequence analysis revealed that these fragments differed in three amino acid posi-

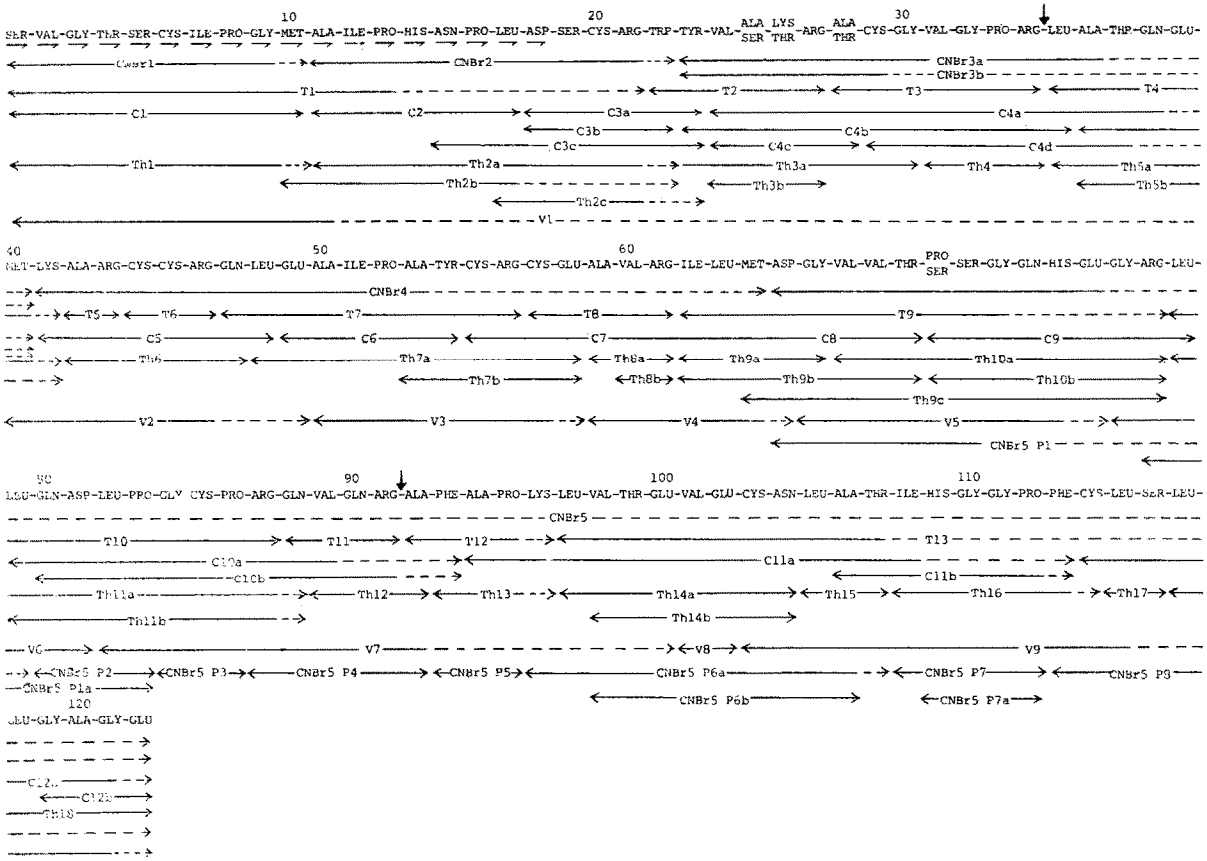

Fig.1. The amino acid sequence of the ragi trypsin/α-amylase inhibitor. T = tryptic peptides, C = chymotryptic peptides, Th = thermolysin peptides, V = peptides from digestion with *S. aureus* V8 protease, CNBr = fragments resulting from cleavage with cyanogen bromide, P = peptides from digestion of CNBr5 with Pro-specific enzyme; (——) results from DABITC method applied to intact protein; (⟷) peptides sequenced by the DABITC method and/or the dansyl–Edman method; (←- -→) regions of peptides which were not sequenced or gave unsatisfactory results. (↓) Trypsin reactive (inhibitory) sites.

tions: CNBr3a having the N-terminal sequence: Tyr-Val-Ser-Thr-Arg-Thr-Cys-, and CNBr3b the sequence: Tyr-Val-Ala-Lys-Arg-Ala-Cys-. Confirmation of the significant level of heterogeneity in this region of the protein was obtained from the analyses of other relevant peptides. The only other example of heterogeneity in this protein occurred in position 70 where both Pro and Ser were found in approximately equal amounts. Where Ser occurred in this position thermolysin hydrolysed the peptide bond Thr^{69}-Ser^{70} as has been reported previously [20].

The specificity of the Pro-enzyme used in the sub-digestion of fragment CNBr5 was not as expected [21] in that in addition to cleavage of Pro-X and Ala-X bonds, anomalous hydrolysis of Gly-X, Thr-X, Leu-X and Ile-X peptide bonds was also detected.

Previous workers [6] have shown that the ragi inhibitor forms a 1:1 molar complex with trypsin and this observation was confirmed during this investigation by titration of the native inhibitor with trypsin. However, attempts to identify the trypsin reactive (inhibitory) peptide bond in the usual way [4] by limited hydrolysis of the native inhibitor with a catalytic amount of trypsin at pH 2.5 for 20 h led to hydrolysis of both the Arg^{34}-Leu^{35} and Arg^{91}-Ala^{92} peptide bonds. The Arg^{34}-Leu^{35} bond lies in a part of the sequence which corresponds exactly to the position of the reactive bond in the homologous barley trypsin inhibitor (fig.2), but both peptide bonds appear to obey the proposed

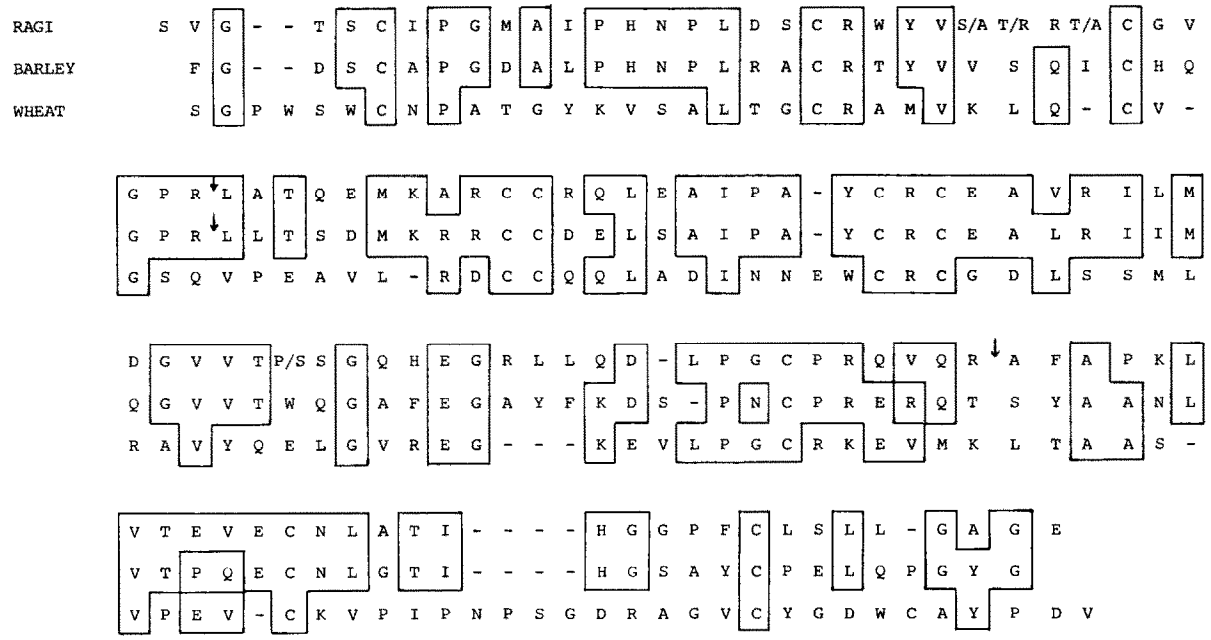

Fig.2. Comparison of the amino acid sequence of the ragi bifunctional trypsin/α-amylase inhibitor with that of the barley trypsin inhibitor [8,9] and the wheat α-amylase inhibitor [5]. Identical residues in the sequences are boxed. (——) A gap inserted in order to obtain maximum homology. (↓) Trypsin reactive (inhibitory) sites.

tentative rules for reactive sites [4]. It may be that the ragi inhibitor possesses two reactive inhibitor sites against trypsin as has been suggested for other similar trypsin inhibitors in barley [22] but in ragi they may be mutually exclusive. Further work is in progress to clarify this finding.

Figure 2 shows the alignment of the sequence of the ragi bifunctional trypsin/α-amylase inhibitor with the sequences of the barley trypsin inhibitor [7,8] and wheat α-amylase inhibitor [5]. There is evident homology between these cereal proteins. A comparison of the ragi and barley proteins reveals approximately 60% homology (69 identical residues out of a possible 110 matches) whilst the ragi/wheat comparison (31 identical in 113) and wheat/barley comparison (33 identical in 113) both show approximately 30% homology. The fragmentary nature of the sequence data for the maize inhibitor [9] does not permit a very meaningful comparison but the limited information available suggests that it also is homologous.

Figure 2 also shows that 26 residues can be regarded as highly conserved (invariant): in particular these include all of the 10 Cys residues possibly involved in the formation of intrachain disulphide bridges. Also noteworthy is the fact that the wheat α-amylase inhibitor, which has no inhibitory activity towards trypsin, does not contain a recognisable reactive site sequence in the region of the sequence which corresponds with the location of the tentatively identified trypsin reactive sites in the barley and ragi proteins.

These findings clearly establish the existence of a new family of enzyme inhibitors to be added to the 10 or so previously classified by Laskowski and Kato [4]. The novel feature of this new family is that it contains members which possess inhibitory activity against entirely unrelated enzymes. It remains to be seen whether these proteins evolved from a common ancestor by a process of gene duplication and mutation, or whether they are the result of gene fusion.

ACKNOWLEDGEMENTS

The authors wish to thank the Science and Engineering Research Council, the CNPq (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 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. (1980) Food Chemistry 6, 235–253.
[3] Ramshaw, J.A.M. (1982) Encyclopedia of Plant Physiol. 14A, 229–290. (New series.)
[4] Laskowski, M.J. and Kato, I. (1980) Ann. Rev. Biochem. 49, 593–626.
[5] Kashlan, N. and Richardson, M. (1981) Phytochemistry 20, 1781–1784.
[6] Shivardj, B. and Pattabiraman, T.N. (1981) Biochem. J. 193, 29–36.
[7] Odani, S., Koide, T. and Ono, T. (1982) FEBS Lett. 141, 279–282.
[8] Odani, S., Koide, T. and Ono, T. (1982) J. Biochem. (Tokyo), in press.
[9] Hochstrasser, K., Illchmann, K. and Werle, E. (1970) Hoppe-Seyler's Z. Physiol. Chem. 351, 721–728.
[10] Crestfield, A.M., Moore, S. and Stein, W.H. (1963) J. Biol. Chem. 238, 622–627.
[11] Mahoney, W.C. and Hermodson, M.A. (1980) J. Biol. Chem. 255, 11199–11203.
[12] Richardson, M. (1974) Biochem. J. 137, 101–112.
[13] Houmard, J. and Drapeau, G.R. (1972) Proc. Natl. Acad. Sci. USA 69, 3506–3509.
[14] Richardson, M. (1979) FEBS Lett. 104, 322–326.
[15] Erlanger, B.F., Kokowski, N. and Kohen, W. (1961) Arch. Biochem. Biophys. 95, 271–278.
[16] Chase, T.J. and Shaw, E. (1967) Biochem. Biophys. Res. Commun. 29, 508–514.
[17] Clarke, J.T. (1964) Ann. N.Y. Acad. Sci. 121, 428–436.
[18] Wrigley, C. (1968) Sci. Tools 15, 17–23.
[19] Starkey, P.M. and Barrett, A.J. (1976) Biochem. J. 155, 255–263.
[20] Emmens, M., Welling, G.W. and Beintema, J.J. (1976) Biochem. J. 157, 317–323.
[21] Yoshimoto, T., Walter, R. and Tsuru, D. (1980) J. Biol. Chem. 255, 4786–4792.
[22] Boisen, S. and Djurtoft, R. (1982) J. Sci. Food Agric. 33, 431–440.