Pergamon

PII: S0031-9422(96)00311-1

Phytochemistry, Vol. 43, No. 2, pp. 327–331, 1996
Copyright © 1996 Elsevier Science Ltd
Printed in Great Britain. All rights reserved
0031-9422/96 $15.00 + 0.00

COMPLETE AMINO ACID SEQUENCES OF TWO TRYPSIN INHIBITORS FROM BUCKWHEAT SEED

MAYA J. PANDYA,* DEBORAH A. SMITH,† ALAN YARWOOD,† JOHN GILROY† and MICHAEL RICHARDSON†

IACR–Long Ashton Research Station, Department of Agricultural Sciences, University of Bristol, Long Ashton, Bristol BS18 9AF, U.K.; †Department of Biological Sciences, Science Laboratories, University of Durham, South Road, Durham DH1 3LE, U.K.

(*Received in revised form* 25 *April* 1996)

Key Word Index—*Fagopyrum esculentum*; Polygonaceae; buckwheat; seed trypsin inhibitors; amino acid sequences.

Abstract—The major trypsin isoinhibitors from seed extracts of buckwheat (*Fagopyrum esculentum* Mönch) were purified by affinity chromatography, anion exchange chromatography, anion exchange HPLC and reversed-phase HPLC, and the complete amino acid sequences of two isoinhibitors, BTI-1 and BTI-2, were established by automated Edman degradation. Each isoinhibitor consists of a single polypeptide chain of 69 amino acids, including two Cys residues. The *N*-terminal sequence of a third isoform, BTI-3, was also determined. The buckwheat trypsin isoinhibitors exhibit clear sequence similarities with the potato chymotrypsin inhibitor I family of serine proteinase inhibitors. Copyright © 1996 Elsevier Science Ltd

INTRODUCTION

Buckwheat (*Fagopyrum esculentum* Mönch) is a crop grown for the floury endosperm of its seeds (achenes), with a widespread cultivation, chiefly in eastern Europe, Asia, North America and Australia. The proteins of these seeds, which are present at relatively high levels of about 15% [1], have high biological value, being rich in Lys, Trp, Arg, Asp and the S-containing amino acids [2]. However, there are strong indications that the digestibility and gastro-intestinal absorption [3–5] of these proteins are poor, probably due to the presence of proteinase inhibitors [6].

Although a number of proteinase inhibitors have been partially characterized from buckwheat seeds [7–12], none have so far had their primary structures determined. We now describe the complete amino acid sequences of two trypsin inhibitors and the partial sequence of a third from seeds of buckwheat which clearly identify all three as being members of the potato chymotrypsin inhibitor I family [13].

RESULTS AND DISCUSSION

The major isoforms of trypsin inhibitor from buckwheat seed were purified to homogeneity using a series of chromatographic steps. A crude seed extract was subjected to heat treatment, ammonium sulphate fractionation, gel filtration and trypsin-affinity chromatography. The purification procedure was monitored by assay for trypsin inhibitory activity, isoelectric focusing (IEF-PAGE) and sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). IEF of the affinity-purified inhibitor preparation (data not shown) indicated that a minimum of five isoforms of buckwheat trypsin inhibitors are present in buckwheat seeds. The affinity-purified fraction was separated into seven protein peaks by ion exchange chromatography on DEAE-Sepharose (Fig. 1). Buckwheat trypsin inhibitors BTI-3, BTI-1 and BTI-2 were purified from peaks

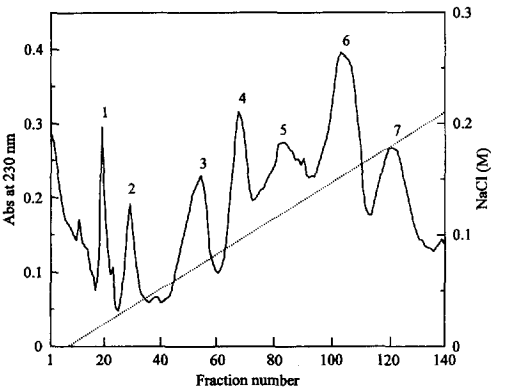


Fig. 1. Anion exchange chromatography of trypsin inhibitors from buckwheat seed on DEAE-Sepharose. The affinity-purified inhibitor preparation was applied to a column (2.5 × 24.5 cm) of DEAE-Sepharose CL-6B in 0.05 M Tris buffer pH 8.4 and eluted by a linear gradient of 0–0.3 M NaCl (over 1 l) in the same buffer. A_{230}, ——; NaCl (M), -----.

*Author to whom correspondence should be addressed.

5–7 (Fig. 1) by anion exchange HPLC. An isoelectric point (pI) of 4.5 was determined for BTI-1 by IEF (data not shown), which is close to the range of pI values (4.6–5.6) determined for the four M_r buckwheat inhibitors isolated by Kiyohara and Iwasaki [11]. The inhibitor samples (BTI-1, 2 and 3) analysed by SDS-PAGE gave single bands under reducing and non-reducing conditions, indicating that the purified trypsin inhibitors exist as monomers. Unique amino acid sequences were obtained during the sequencing of BTI-1, BTI-2 and BTI-3, with no evidence of cleaved forms of inhibitor resulting from trypsin-affinity chromatography. The complete amino acid sequences of isoforms BTI-1 and BTI-2 are shown in Fig. 2, together with details of the overlapping peptides from which the sequences were deduced. In both cases the amino acid sequences were determined by a single run of automated sequencing, in conjunction with automated analyses of selected peptides covering the *C*-terminal regions. The only residue which gave an ambiguous result during the direct automated sequencing on the intact protein was Trp^{10} in BTI-1 which was subsequently identified by automated sequencing of peptide T1. The partial sequence of a third isoform, BTI-3, (also shown in Fig. 2) was the result of a single run of automated sequencing. The isoforms BTI-1 and BTI-2 each consist of a single polypeptide chain of 69 amino acids, with calculated M_r values of 7634 and 7748 respectively. Their sequences show high contents of Val, Arg and Glu, and are in good agreement with the amino acid composition of buckwheat inhibitor I isolated by Kiyohara and Iwasaki [11]. Since the conclusion of this work, other workers [14] have reported the complete amino acid sequence of a protease inhibitor BWI-1 from buckwheat seeds which was established by automated Edman degradation and mass spectrometry and is identical with BTI-2 reported here. Inhibitors BTI-1 and BTI-2 both contain two Cys residues and are identical apart from substitutions at residues 48 and 53. BTI-3 shows considerable sequence identity with BTI-1 and BTI-2 including a Cys (corresponding to residue 5 of BTI-1 and 2), but it lacks two residues at the *N*-terminus.

The buckwheat trypsin inhibitors show strong sequence similarities (up to 65%) with the potato I family of serine proteinase inhibitors [13]. The amino acid sequences of several members of this family, including inhibitors from the monocotyledonous species barley [15, 16] and an invertebrate (the medicinal leech) [17], are aligned with those of the buckwheat trypsin inhibitors in Fig. 3. The presence of a disulphide bridge is a feature of trypsin inhibitors. The existence of inhibitors in barley [15, 16], broad bean [23] and leech [17] which contain no Cys residues, however, indicates that the potato I family are functionally independent of disulphide bridges in their structures. The inhibitory specificity of most serine proteinase inhibitors is determined by a reactive (inhibitory) peptide bond, designated $P_1 - P_1'$ [24], situated within a loop stabilized by a disulphide bridge. Many inhibitors of the potato I family have a hydrophobic residue (Met, Leu, Ala) at the P_1 position and an acidic residue (Asp, Glu) at the P_1' position. Examples, given in Fig. 3, are the inhibitors of chymotrypsin from bitter gourd [20], tomato [21], potato [25], barley [26] and leech [27]. Trypsin inhibitors with the basic amino acid, Lys, at the P_1 position have been isolated from amaranth [18], pumpkin [19] and tomato [28]. The putative reactive site of buckwheat trypsin inhibitors BTI-1 and BTI-2 is Arg^{45}–Asp^{46}, by comparison with other inhibitors of the potato I family. The presence of a basic amino acid at the P_1 position is consistent with their specificity as trypsin inhibitors [9, 10]. Two highly conserved Arg side chains at the P_6' and P_8' positions in the protein core of bound barley chymotrypsin inhibitor (CI-2) and leech eglin c form important electrostatic and hydrogen-bonding interactions with residues at the P_1' (Glu^{60} in CI-2 [29]; Asp^{46} in eglin c [27]) and P_2 (Thr^{58} in CI-2 [29]; Thr^{44} in eglin c [27]) positions of the reactive loop. Protein engineering experiments in which Arg at the P_8' position was replaced with Ala in CI-2 [30], or Lys in eglin [31] resulted in decreased inhib-

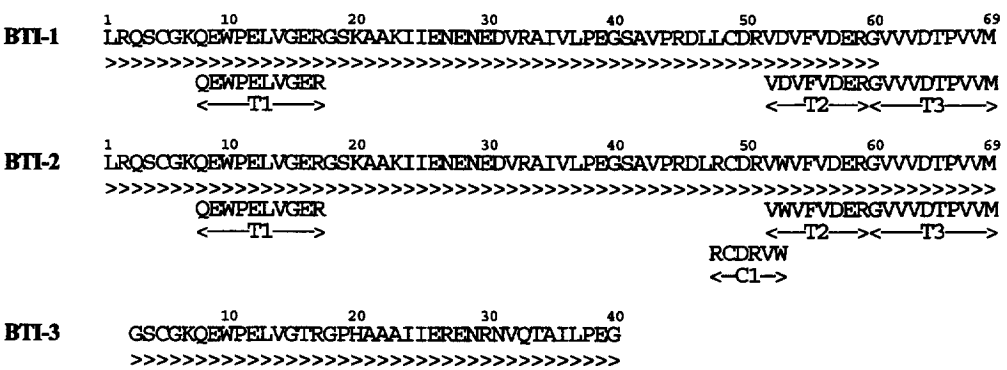
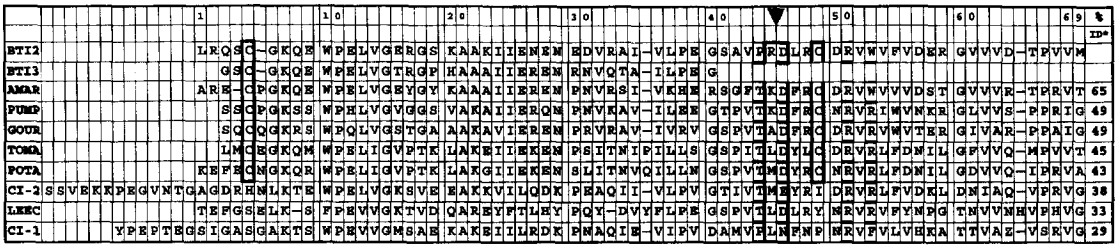


Fig. 2. The complete amino acid sequence of the buckwheat trypsin inhibitors, BTI-1 and BTI-2, and the incomplete sequence of a third isoform, BTI-3. Arrows (>) indicate amino acid residues identified by automated sequencing of the intact *S*-alkylated proteins. T, tryptic peptides; C, chymotryptic peptides.



* % identity compared with BTI-2

Fig. 3. Comparison of the amino acid sequences of buckwheat trypsin inhibitors, BTI-2 and BTI-3, with serine proteinase inhibitors of the potato I family. (BTI2) BTI-2 from buckwheat seed; (BTI3) BTI-3 from buckwheat seed (incomplete sequence); (AMAR) inhibitor of trypsin and subtilisin from amaranth seed [18]; (PUMP) inhibitor of trypsin and activated Hageman factor from pumpkin seed [19]; (GOUR) inhibitor of *Streptomyces griseus* Glu-proteinase from bitter gourd seed [20]; (TOMA) wound inducible leaf proteinase inhibitor from cultivated tomato [21]; (POTA) chymotrypsin inhibitor I subunit D from potato tubers [22]; (CI-2) inhibitor CI-2 of subtilisin and chymotrypsin from barley seed [15]; (LEEC) eglin c inhibitor of elastase, cathepsin G, subtilisin and chymotrypsin from the medicinal leech [17]; (CI-1) inhibitor CI-1 of subtilisin and chymotrypsin from barley seed [16]. The numbering refers to BTI-2; gaps (–) have been introduced into the sequences to improve alignments. Cys residues aligned at positions 5 and 49, and the reactive (inhibitory) site residues at 45 and 46 (determined in potato, barley and leech proteins) are enclosed in boxes. Other residues of interest at positions 51, 53, 58 and 60 are underlined. Values of amino acid identity (%) with BTI-2 are shown at the end of the sequences.

itory activity due to protein destabilization. The buckwheat trypsin inhibitors are unusual in lacking Arg at the P_8' position and containing Pro at position 44 (P_2). BTI-2 has Trp at position 53 (P_8'), a bulky hydrophobic residue which is capable of forming hydrogen bonds. BTI-1 has the acidic residue, Asp, at position 53. The trypsin inhibitor from amaranth [18] also contains Trp at the P_8' position, apparently without adverse effect on inhibitor capacity. The conserved Arg residue is also absent from barley CI-1 [16], which has Phe at the P_8' position and Pro or His at the P_2 position. Barley CI-2 is a stronger chymotrypsin inhibitor than CI-1 using casein as the substrate [32], and is a strong inhibitor of *Bacillus pumilus* proteinase SP226, which barley CI-1 [33] and amaranth trypsin inhibitor [22] only affect weakly. The Asp residue present at position 53 in BTI-1 is unique amongst the known inhibitors of the potato I family. Structural characterization of one of these atypical inhibitors of the potato I family would improve our understanding of the stabilization of the reactive loop both by Cys and Arg residues. It would be of interest to determine how the substitutions present in the putative inhibitory sites of the buckwheat isoinhibitors relate to their activities and specificities.

EXPERIMENTAL

Gel filtration (Bio-Gel P-10, Bio-Gel P-2) and affinity (Affi-Gel 10) column chromatography matrices were purchased from Bio-Rad Laboratories. Sephadex G-50 and DEAE-Sepharose CL-6B matrices were products of Pharmacia LKB Biotechnology. All reagents and supplies used were analytical grade or better.

Purification of inhibitor. Buckwheat seeds (*Fagopyrum esculentum* Mönch) were purchased locally (distributed by Holland & Barrett) and ground to a fine flour in a coffee grinder. 250 g batches of flour were stirred with 2.5 l of 2% (w/v) NaCl for 90 min at room temp. The suspension was filtered through muslin and then centrifuged at 10 000 *g* for 30 min at 4°. The clear supernatant obtained was heated at 80° for 10 min, with stirring. The precipitated proteins were removed by centrifugation. $(NH_4)_2SO_4$ was added to the clear supernatant to give 60% sat. The resulting ppt was collected by centrifugation at 14 000 *g* for 60 min at 4°, then dissolved in a minimal vol. of H_2O, dialysed against H_2O, and lyophilized. This material was dissolved in 6 M urea in 0.1 M HOAc and fractionated by gel filtration chromatography using either Bio-Gel P-10 or Sephadex G-50 equilibrated with 0.1 M HOAc. Proteins were eluted with 0.1 M HOAc and the protein peak (detected by A_{280}) which contained trypsin inhibitory activity was collected and lyophilized. This material was dissolved in 0.1 M Tris buffer pH 8.1 containing 2 mM $CaCl_2$, and applied to a column of Affi-Gel 10, to which trypsin (TPCK treated) was attached, equilibrated with the same buffer. The column was then washed with the starting buffer until the A_{280} reached zero, when a protein peak containing trypsin inhibitory activity was eluted with 0.1 M KCl pH 2.1 and then lyophilized.

The crude inhibitor preparations from affinity chromatography were desalted by gel filtration chromatography using either Sephadex G-50 or Bio-Gel P-2 in 0.1 M HOAc, and then lyophilized. This material was dissolved in 0.05 M Tris buffer pH 8.4 and applied to a column (2.5 × 24.5 cm) of DEAE-Sepharose CL-6B equilibrated with the same buffer. After washing the column with the starting buffer a linear gradient of 0–0.3 M NaCl over 1 l was applied. Seven protein peaks (detected by A_{230}) were collected, lyophilized,

and then desalted either by gel filtration chromatography or by dialysis against H_2O at 4° and lyophilized. These inhibitor peaks were dissolved in 1 ml of 10 mM Na-Pi buffer pH 8.5 and 0.5 ml samples were applied to a weak anion exchange HPLC column (0.41 × 25 cm) of Synchropak AX300 equilibrated in the same buffer. After absorption, the column was eluted with a linear gradient of 0–0.2 M NaCl followed by a linear gradient of 0.2–0.6 M NaCl in the same buffer. The collected protein peaks (detected by A_{216}) were microdialysed, tested for presence of inhibitor, and lyophilized.

Inhibitor assay. Inhibition of the hydrolytic activity of trypsin was measured using α-*N*-benzoyl-D,L-arginine-*p*-nitroanilide as the substrate [34]. The inhibitor sample and trypsin were preincubated together for 30 min prior to the addition of the substrate.

Electrophoresis. Tricine-SDS-PAGE was performed with 16.5% gels under reducing and non-reducing conditions [35]. IEF was performed using PhastGel media, Pharmalyte carrier ampholytes covering the pH ranges 3–9.5 and 4–6.5, and pI calibration proteins (Pharmacia LKB Biotechnology) according to the manufacturer's instructions.

Reduction and S-alkylation. Samples were reduced and *S*-carboxymethylated [36], and then desalted by reversed-phase (RP) HPLC on a Vydac 218TP54 C_{18} column (0.46 × 25 cm) in 0.1% (v/v) aq. TFA. The column was eluted with a linear gradient of 0–60% (v/v) MeCN in 0.1% (v/v) aq. TFA, and the collected protein peaks (detected by A_{216}) were lyophilized.

Sequence determination. The *S*-alkylated inhibitors (2–4 nmol) were sequenced by repeated cycles of Edman degradation using an Applied Biosystems 477A automatic pulsed liquid phase protein sequencer. The phenylthiohydantoin derivatives were identified by RP-HPLC using an Applied Biosystems 120A PTH analyser system. Other samples of the *S*-alkylated inhibitors were digested separately with trypsin and chymotrypsin at 2% (w/w) enzyme/substrate in 0.1 M *N*-ethyl morpholine buffer pH 8.1, for 75 min. Peptides obtained from these digests were purified by RP-HPLC on a Vydac 218TP54 column, lyophilized and sequenced by automated Edman degradation.

Sequence comparison. The amino acid sequences of BTI-1 and BTI-2 were compared with other proteins stored in the Swiss-Prot database using the Genetics Computer Group package.

Acknowledgement—We are grateful to Prof. P. R. Shewry at Long Ashton Research Station for critical reading of the manuscript.

REFERENCES

1. Papotto, G., Virtucio, L. and Mondardini, L. (1990) *Cereal Foods World* **35**, 1157.
2. Pomeranz, Y., Marshall, H. G., Robbins, G. S. and Gilbertson, J. T. (1975) *Cereal Chem.* **52**, 479.
3. Science and Technology Agency – Resources Council (1981) *Data on the Revision of the 4th Standard Tables of Food Composition in Japan (Data no. 92)*, Science and Technology Agency, Resources Council, Tokyo.
4. Javornik, B., Eggum, B. O. and Kreft, I. (1981) *Genetika (Zemun)* **13**, 115.
5. Thacker, P. A., Anderson, D. M. and Bowland, J. P. (1983) *Can. J. Anim. Sci.* **63**, 213.
6. Ikeda, K., Sakaguchi, T., Kusano, T. and Yasumoto, K. (1991) *Cereal Chem.* **68**, 424.
7. Ikeda, K. and Kusano, T. (1978) *Agric. Biol. Chem.* **42**, 309.
8. Pokrovsky, S. N. and Belozersky, M. A. (1980) *Biochemistry (Moscow)* **45**, 2104.
9. Ikeda, K. and Kusano, T. (1983) *Agric. Biol. Chem.* **47**, 1481.
10. Kiyohara, T. and Iwasaki, T. (1985) *Agric. Biol. Chem.* **49**, 581.
11. Kiyohara, T. and Iwasaki, T. (1985) *Agric. Biol. Chem.* **49**, 589.
12. Voskoboynikova, N. E., Dunaevsky, Y. E. and Belozersky, M. A. (1990) *Biochemistry (Moscow)* **55**, 839.
13. Richardson, M. (1991) in *Amino Acids, Proteins and Nucleic Acids*, Methods in Plant Biochemistry (Rogers, L. J., ed), Vol. 5, pp. 259–305. Academic Press, London.
14. Belozersky, M. A., Dunaevsky, Y. E., Musolyamov, A. X. and Egorov, T. A. (1995) *FEBS Letters* **371**, 264.
15. Svendsen, I., Jonassen, I., Hejgaard, J. and Boisen, S. (1980) *Carlsberg Res. Commun.* **45**, 389.
16. Williamson, M. S., Forde, J. and Kreis, M. (1988) *Plant Mol. Biol.* **10**, 521.
17. Seemüller, U., Eulitz, M., Fritz, H. and Strobl, A. (1980) *Hoppe-Seyler's Z. Physiol. Chem.* **361**, 1841.
18. Hejgaard, J., Dam, J., Petersen, L. C. and Bjørn, S. E. (1994) *Biochim. Biophys. Acta* **1204**, 68.
19. Krishnamoorthi, R., Gong, Y. -X. and Richardson, M. (1990) *FEBS Letters* **273**, 163.
20. Ogata, F., Miyata, T., Fujii, N., Yoshida, N., Noda, K., Makisumi, S. and Ito, A. (1991) *J. Biol. Chem.* **266**, 16715.
21. Graham, J. S., Pearce, G., Merryweather, J., Titani, K., Ericsson, L. and Ryan, C. A. (1985) *J. Biol. Chem.* **260**, 6555.
22. Richardson, M. and Cossins, L. (1975) *FEBS Letters* **52**, 161.
23. Svendsen, I., Hejgaard, J. and Chavan, J. K. (1984) *Carlsberg Res. Commun.* **49**, 493.
24. Schechter, I. and Berger, A. (1967) *Biochem. Biophys. Res. Commun.* **27**, 157.
25. Richardson, M., Barker, R. D. J., McMillan, R. T. and Cossins, L. (1977) *Phytochemistry* **16**, 837.
26. Jonassen, I. B. and Svendsen, I. B. (1982) *Carlsberg Res. Commun.* **47**, 199.
27. McPhalen, D. A., Schnebli, H. P. and James, M. N. G. (1985) *FEBS Letters* **188**, 55.

28. Wingate, V. P. M., Broadway, R. M. and Ryan, C. A. (1989) *J. Biol. Chem.* **264**, 17734.
29. McPhalen, C. A. and James, M. N. G. (1987) *Biochemistry* **26**, 261.
30. Jandu, S. K., Ray, S., Brooks, L. and Leatherbarrow, R. J. (1990) *Biochemistry* **29**, 6264.
31. Heinz, D. W., Hyberts, S. G., Peng, J. W., Priestle, J. P., Wagner, G. and Grütter, M. G. (1992) *Biochemistry* **31**, 8755.
32. Boisen, S., Andersen, C. Y. and Hejgaard, J. (1981) *Physiol. Plant.* **52**, 167.
33. Hejgaard, J. (1981) *Analyt. Biochem.* **116**, 444.
34. Campos, F. A. P. and Richardson, M. (1983) *FEBS Letters* **93**, 205.
35. Schägger, H. and von Jagow, G. (1987) *Analytical Biochemistry* **166**, 368.
36. Crestfield, A. M., Moore, S. and Stein, W. H. (1963) *J. Biol. Chem.* **238**, 622.