FEBS Letters 371 (1995) 264–266 FEBS 15971

Complete amino acid sequence of the protease inhibitor from buckwheat seeds

Mikhail A. Belozersky[a,*], Yakov E. Dunaevsky[a], Alexander X. Musolyamov[b], Tsezi A. Egorov[b]

[a]*A.N. Belozersky Institute of Physico-Chemical Biology, Moscow State University, Moscow 119899, Russian Federation*
[b]*V.A. Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Vavilova 32, Moscow 117984, Russian Federation*

Received 24 July 1995

Abstract The complete amino acid sequence of protease inhibitor BWI-1 from buckwheat (*Fagopyrum esculentum* Moench) seeds has been established by automatic Edman degradation and mass spectrometry. The molecule of the inhibitor consists of 69 amino acid residues, with a molecular mass calculated as 7743.8 Da. The active site of the inhibitor containes an Arg^{45}-Asp^{46} bond. Analysis of the amino acid sequence suggests that the buckwheat seed protease inhibitor is a member of the proteinase inhibitor I family.

Key words: Amino acid sequence; Buckwheat; Protease inhibitor

1. Introduction

Protein protease inhibitors are widely distributed in living organisms. The modern classification of these proteins based on their amino acid sequences, active site structures and mechanisms of action involves at least 10 different families [1]. Plant protease inhibitors are classified into three main groups, namely Kunitz, Bowman-Birk and proteinase inhibitor I families. The physiological role of these inhibitors in plants has not been established conclusively and is likely to consist of (1) protection from proteolytic enzymes of various pathogens and insects, (2) regulation of the activity of endogenous proteases, (3) contributing to the pool of storage proteins [2]. We have previously purified and characterized a number of trypsin and chymotrypsin inhibitors from buckwheat seeds [3–6] and demonstrated that one was involved in regulation of endogenous metalloproteinase activity [5] while others suppressed the growth and development of pathogenic fungi [6]. Trypsin inhibitors from buckwheat seeds were also isolated and described in the works of other groups [7–9]. However, they did not study the physiological role of these proteins or their amino acid sequences. In the present paper we describe the complete amino acid sequence of the BWI-1 trypsin inhibitor from buckwheat seeds.

2. Experimental

2.1. Plant material

Dry seeds of buckwheat (*Fagopyrum esculentum* Moench) cv. Shatilovskaya-5 were used

*Corresponding author. Fax: 007-095-939-31-81

Abbreviations HPLC, high-performance liquid chromatography; PE, pyridylethyl; PTH, phenylthiohydantoin; TFA, trifluoroacetic acid.

2.2. Isolation of the inhibitor

The inhibitor was isolated by two independent techniques. In the first case the seed extract was subjected to affinity chromatography on a trypsin-Sepharose column as described previously [6]. The fraction containing the total preparation of buckwheat seed trypsin inhibitors was further chromatographed on a Mono Q HR 5/5 (Pharmacia, Sweden) column by FPLC. The sample (10 mg protein) was applied to the column in 20 mM K, Na-phosphate, pH 6.8. The column was then washed with the same buffer and the protein was eluted with a linear NaCl gradient (0–100 mM, $ml \cdot min^{-1}$, 25 min). The preparation of the BWI-1 inhibitor was finally purified by reversed-phase HPLC on an Aquapore RP-300 (4.6 × 100 mm) (Applied Biosystems, Inc., USA) column using an acetonitryl gradient at a flow rate of 0.5 $ml \cdot min^{-1}$ Solution A: 0.1% (v/v) TFA; solution B: acetonitrile containing 0.08% (v/v) TFA.

In the second procedure the inhibitor preparation was isolated by fractionation of the seed extract on a Sephadex G-75 column [4] followed by chromatography on the Mono Q column and reversed-phase HPLC as described above.

2.3. Assay of the activity of the inhibitors

A preparation of the inhibitors in 100 mM K, Na-phosphate, pH 6.8 (1–10 μl) was added to 5 μl of a trypsin solution (1 $mg \cdot ml^{-1}$) and the mixture was incubated at room temperature for 10 min. The esidual trypsin activity was assayed according to Erlanger et al. [10] with a 0.2 mM solution of *N*-benzoyl-D,L-arginine-*p*-nitroanilide in 500 mM K, Na-phosphate, pH 7.5, as a substrate. One unit of the inhibitor activity produced a decrease in absorption of 0.1 unit of trypsin activity assay.

2.4. Reduction and alkylation of the protein

0.5 mg of the inhibitor preparation were reduced and alkylated by 4-vinilpyridin as described by Thomsen and Bayne [11]. The alkylated protein was dissoluted by reversed-phase HPLC and evaporated in Speedvac (Savant, USA)

2.5. Digestion of the protein by V8 protease

300 μg of reduced and alkylated protein in 100 μl of 100 mM NH_4HCO_3, pH 7.8, were digested by 2 μg of the *Staphylococcus aureus* V8 protease (Sigma, USA) for 16 h at 37°C. The obtained peptides were separated by reversed-phase HPLC and evaporated in Speedvac (Sanvant, USA)

2.6. Amino acid analysis

The amino acid analysis of the intact BWI-1 was carried out according to a standard procedure with a Hitachi 835 amino acid analyzer. Cysteine was determined as cysteic acid after oxidation of the protein with a mixture of H_2O_2 and 88% performic acid (1:9, v/v) followed by hydrolysis with 5.7 M HCl. Tryptophan was determined after hydrolysis of the protein with 4 M methanesulfonic acid containing 0.2% tryptamine.

2.7. Sequence analysis

The amino acid sequence of the protein was determined using a model 816 protein/peptide Sequencer (Knauer GmbH, Germany) equipped with a model 120A PTH-analyzer (Applied Biosystems, Inc., USA)

2.8. Mass spectrometry

The mass spectra were recorded in a ^{252}Cf plasma desorption time-of-flight spectrometer ('Electron', Sumy, Ukraine). Peptides 200–500 pmol

SSDI 0014-5793(95)00899-3

were dissolved in 0.1% TFA and applied to nitrocellulose covered targets. The C-terminal peptide sequences were determined by electrospray mass-spectrometry (ESMS) as described by Tsugita et al. [12].

3. Results and discussion

The inhibitor isolated by the second procedure (gel-filtration, ion-exchange chromatography on Mono Q, reversed-phase HPLC) was intact as demonstrated by the fact that its molecular mass after reduction and alkylation was in good agreement with the calculated value (Fig. 1, II) and from the results of N-terminal sequencing. However, the inhibitor preparation obtained by the first procedure (affinity chromatography, Mono Q chromatography, reversed-phase HPLC) was cleaved on the carboxyl side of Arg^{45} in the active site into two fragments connected by a single disulfide bond (Fig. 1, II). Such cleavage often occurs at acid pH by limited proteolysis with trypsin [13]. In the present case it was demonstrated by the results of N-terminal sequence analysis and mass-spectrometry, which revealed three fragments. The molecular mass of the first fragment (7763.9 Da) coincided with that calculated from the sequence analysis data. The two other fragments, N-terminal (NT) (M_r 4959.6 Da) and C-terminal (CT) (M_r 2804.6 Da), were formed as a result of cleavage of the S–S bond after bombardment of the protein with fission products of ^{252}Cf. N-terminal analysis of the inhibitor preparation showed the presence of two protein sequences corresponding to NT and CT fragments (Fig. 1, II). Reduction and alkylation of the protein preparation resulted in two fragments, which were separated by reversed-phase HPLC (data not shown) and sequenced.

Amino acid analysis of the intact BWI-1 revealed the presence of two half-cysteine residues, that is typical for a certain group of seed protease inhibitors and the absence of histidine and tyrosine (Table 1).

The strategy for determination of the amino acid sequence of the inhibitor is presented in Fig. 2. This sequence was determined both by sequencing from the N-terminus of the reduced and alkylated intact protein and by sequencing of the NT and CT fragments formed during purification of the inhibitor by affinity chromatography on trypsin-Sepharose. The sequence of the NT fragment (45 residues) was determined completely (Fig. 2). In the CT fragment the sequence of 21 out of 24 amino acid residues was established.

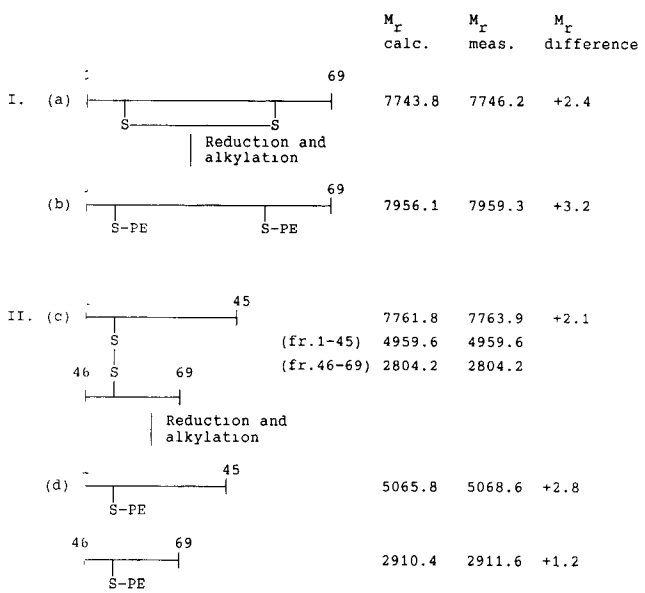

Fig. 1. Structure of the protease inhibitor from buckwheat seeds isolated by gel-filtration and etc. (I) and by affinity chromatography and etc (II). a, c, non-reduced and non-alkylated protein; b, d, reduced and alkylated protein. $M_r$calc., calculated molecular mass, $M_r$meas., measured molecular mass.

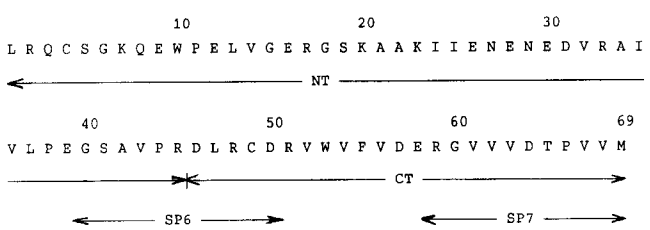

Fig. 2. Sequencing of the protease inhibitor from buckwheat seeds. Peptides NT and CT were obtained after protein isolation by affinity chromatography and etc., followed by its reduction and alkylation and separation by reversed-phase HPLC and peptides SP6 and SP7 were obtained by digestion of the intact protein, isolated by gel-filtration and etc., with V8 protease.

To prove that the peptide bond in the inhibitor molecule was cleaved only at the C-terminus of Arg^{45} during purification by affinity chromatography, the intact protein, obtained without affinity chromatography on trypsin-Sepharose, was digested by the V8 endoprotease. As a result a number of peptides were obtained (data not shown), one of which, SP6, contained an amino acid sequence (Gly^{40}-Asp^{50}) overlapping with the C-terminal part of the NT-fragment and the N-terminal part of CT fragment. These data showed that the peptide bond between Arg^{45} and Asp^{46} is cleaved in the inhibitor obtained by affinity purification on trypsin-Sepharose.

Another peptide, SP7, obtained after V8 protease digestion of the intact inhibitor contained the C-terminal part of the

Table 1
Amino acid composition of intact buckwheat seed protease BWI-1 inhibitor

Amino acid	BWI-1
Asp	6.7 (5)
Asn	(2)
Thr	1.2 (1)
Ser	2.6 (3)
Glu	11.8 (8)
Gln	(2)
Pro	4.1 (4)
Gly	5.2 (5)
Ala	4.1 (4)
1/2 Cys	2.0 (2)
Val	11.3 (12)
Met	0.8 (1)
Ile	1.9 (3)
Leu	3.9 (4)
Tyr	0.0 (0)
Phe	1.1 (1)
Lys	3.2 (3)
His	0.0 (0)
Arg	7.1 (7)
Trp	1.5 (2)
Total	(69)

Numbers in parenthesis are the residues determined after sequence analysis.

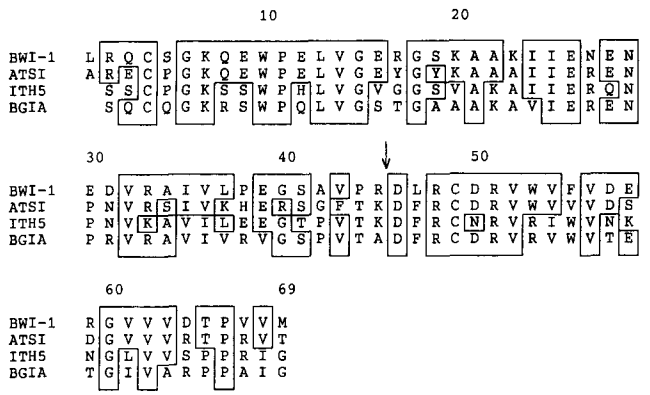

Fig. 3. Comparison of amino acid sequences of some members of the proteinase inhibitor I family. ATSI, inhibitor from *Amaranthus caudatus* [14], ITH5, trypsin inhibitor from pumpkin [15], BGIA, inhibitor from bitter gourd [16]. Identical amino acids in the sequences are boxed. The arrow indicates the P_1 amino acid residue at the reactive sites of the inhibitors.

sequence (Arg^{59}-Met^{69}). In this peptide only the sequence of the last three C-terminal amino acids was not determined. To establish this sequence peptide SP7 was subjected to mass analysis by ESMS. As a result the C-terminal sequence of SP7 was determined: Val^{67}-Val^{68}-Met^{69}. The complete amino acid sequence of the studied inhibitor is shown in Fig. 2. The M_r of the protein (7746.2 Da) determined by mass-spectrometry is in good coincidence with the calculated value (7743.8 Da).

Analysis of the amino acid sequence of the studied buckwheat seed protease inhibitor revealed a 65%, 50% and 49% identity with the trypsin/subtilisin inhibitor from *Amaranthus caudatus* [14], the trypsin inhibitor from pumpkin (*Cucurbita maxima*) [15] and inhibitor of the *Streptomyces griseus* Glu-protease from bitter gourd (*Momordica charantia*) [16], respectively (Fig. 3). These data suggest that BWI-1 belongs to the proteinase inhibitor I family. The reactive site of BWI-1 is proposed to be Arg^{45}-Asp^{46} (Fig. 3) on the basis of its sequence homology to inhibitors of this family. It should be noted that Arg had not been found before at the P_1 position at the active sites of inhibitors which belong to proteinase inhibitor I family.

Acknowledgements: This work was supported by grants from the Russian Foundation for Basic Research, State Program 'Frontiers in Genetics' and from the International Science Foundation.

References

[1] Laskowski, J M. and Kato, J. (1980) Annu. Rev. Biochem. 49, 593–626.
[2] Richardson, M. (1977) Phytochemistry 16, 159–169.
[3] Pokrovsky, S.N., Labasina, N.Y. and Belozersky, M.A. (1978) Bioorg. Chem. (Moscow) 4, 269–275.
[4] Pokrovsky, S.N. and Belozersky, M.A. (1980) Biochemistry (Moscow) 45, 2104–2110.
[5] Voskoboynikova, N.E., Dunaevsky, Y.E. and Belozersky, M.A. (1990) Biochemistry (Moscow) 55, 839–847
[6] Dunaevsky, Y E , Pavliukova, E.B., Beliakova, G.A. and Belozersky, M.A. (1994) Biochemistry (Moscow) 59, 990–996.
[7] Ikeda, K. and Kusano, T. (1983) Agric. Biol. Chem. 47, 481–487.
[8] Kiyohara, T. and Iwasaki, T. (1985) Agric. Biol. Chem. 49, 581–588.
[9] Kiyohara, T. and Iwasaki, T. (1985) Agric. Biol. Chem. 49, 589–594.
[10] Erlanger, B.F., Kokowsky, N. and Cohen, W. (1961) Arch. Biochem. Biophys. 95, 271–278.
[11] Thomsen, J. and Bayne, S.J. (1990) J. Prot. Chem. 7, 295–296.
[12] Tsugita, A., Takamoto, K., Kamo, M. and Iwade, H. (1992) Eur. J. Biochem. 206, 691–696.
[13] Ozawa, K. and Laskowski, M.J. (1966) J. Biol. Chem. 241, 3955–3960.
[14] Hejgaard, J., Dam, J., Pedersen, L C. and Bjoern, S.E. (1994) Biochim. Biophys. Acta 1204, 68–74.
[15] Krishnamoorthi, R., Gong, Y.X. and Richardson, M. (1990) FEBS Lett. 273, 163–167.
[16] Ogata, F , Miyata, T., Fujii, N., Yoshida, N., Makisumi, S. and Ito, A. (1991) J. Biol. Chem. 266, 16715–16721.