Journal of Protein Chemistry, Vol. 9, No. 4, 1990

Amino Acid Sequence of the Acidic Kunitz-Type Trypsin Inhibitor from Winged-Bean Seed [*Psophocarpus tetragonolobus* (L.) DC]

J. Bruce Caldwell,[1] Phillip M. Strike,[1] and Alexander A. Kortt[1,2]

Received February 26, 1990

The primary sequence of trypsin inhibitor-2 (WBTI-2) from *Psophocarpus tetragonolobus* (L.) DC seeds was determined. This inhibitor consists of a single polypeptide chain of 182 amino acids, including four half-cystine residues, and an N-terminal residue of pyroglutamic acid. The sequence of WBTI-2 showed 57% identity to the basic trypsin inhibitor (WBTI-3) and 50% identity to the chymotrypsin inhibitor (WBCI) of winged bean, and 54% identity to the trypsin inhibitor DE-3 from *Erythrina latissima* seed. The similarity to the soybean Kunitz trypsin inhibitor (40%) and the other Kunitz-type inhibitors from *Adenanthera pavonina* (30%) and wheat (26%) was much lower. Sequence comparisons indicate that the *Psophocarpus* and *Erythrina* inhibitors are more closely related to each other than to other members of the Kunitz inhibitor family.

KEY WORDS: Trypsin inhibitor; Kunitz-type inhibitor; amino acid sequence; winged bean; *Psophocarpus tetragonolobus* (L.) DC.

1. INTRODUCTION

Winged bean [*Psophocarpus tetragonolobus* (L.) DC] is a tropical legume with considerable potential for development as a high-protein crop, as it contains as much protein and oil in its seed as does soybean (Sri Kantha and Erdman, 1984). Winged-bean seed contains trypsin and chymotrypsin inhibitors which comprise about 5% of the total seed protein (Kortt, 1979, 1980). Two Kunitz-type trypsin inhibitors (WBTI-2 and WBTI-3) were isolated from winged-bean seed by affinity chromatography on trypsin-Sepharose 4B and were separated by ion-exchange chromatography (Kortt, 1979). Amino-terminal sequence analysis of WBTI-3 (Kortt *et al.*, 1983) showed that the inhibitor was related to the soybean Kunitz trypsin inhibitor. Yamamoto *et al.* (1983) isolated and determined the amino acid sequence of a basic Kunitz-type inhibitor from winged bean, which was identical to WBTI-3 (Kortt, 1979; Kortt *et al.*, 1983). Winged bean trypsin inhibitor-2 (WBTI-2), an acidic protein, with an amino acid composition distinct from that of WBTI-3 (Kortt, 1979), has a blocked N-terminus (Kortt *et al.*, 1983). No sequence data were determined in these earlier studies.

The complete sequences of Kunitz-type trypsin inhibitors from soybean (Koide and Ikenaka, 1973; Kim *et al.*, 1986), *Erythrina latissima* seed (Joubert *et al.*, 1985), *Erythrina caffra* seed (Joubert and Dowdle, 1987), and *Adenanthera pavonina* seed (Richardson *et al.*, 1986) have been determined. The sequences of a Kunitz-type α-amylase inhibitor from wheat (Maeda, 1986), and a Kunitz-type chymotrypsin inhibitor (WCI-3) from winged bean (Shibata *et al.*, 1988) have also been reported. The sequence of the acidic trypsin inhibitor (WBTI-2) from winged-bean seed was determined to establish its relationship to the other Kunitz-type seed proteins of winged bean and to the other members of this family.

[1] CSIRO, Division of Biotechnology, Parkville Laboratory, 343 Royal Parade, Parkville, Victoria 3052, Australia.
[2] To whom all correspondence should be addressed.

0277-8033/90/0800-0493$06.00/0

2. EXPERIMENTAL

Winged-bean trypsin inhibitor-2 (WBTI-2) was isolated from mature seeds of *Psophocarpus tetragonolobus* (L.) DC variety UPS-122, purified as described (Kortt, 1979), dialysed against distilled water, and recovered by lyophilization. The inhibitor (10 mg) was dissolved in 0.5 M Tris-HCl, 6 M guanidine/HCl, 2 mM EDTA, *p*H 8.6, reduced with dithiothreitol (0.1 M) at 50°C for 1 hr under nitrogen, and alkylated with iodoacetic acid (0.3 M) for 1 hr at 25°C in the dark. Excess dithiothreitol was added to remove excess alkylating reagent, the reaction mixture was dialysed against distilled water and the protein recovered by lyophilization. The reduced and alkylated inhibitor preparation used for sequencing gave a single band of M_r ~20,000 on SDS-polyacrylamide gel electrophoresis. This preparation therefore did not contain material that was cleaved at the reactive site (Ozawa and Laskowski, 1966; Yamamoto *et al.*, 1983) during purification as found previously for some inhibitor preparations (Kortt, 1979).

Aliquots (25 nmol) of the S-carboxymethylated protein in 0.5 ml 0.05 M ammonium bicarbonate buffer, *p*H 8.0, were digested at 37°C for 4 hr with one of the following proteases at an enzyme/substrate ratio of 1:50 (by mass): trypsin (Worthington), α-chymotrypsin (Worthington), *Staphylococcus aureus* V8 protease (Pierce), lysyl endopeptidase (Wako), and clostripain (Sigma). Clostripain was reduced with dithiothreitol (1 mM) and clostripain, tryptic and chymotryptic digests were carried out in the presence of 1 mM $CaCl_2$. The digests were dried at 50°C in a Savant dryer, redissolved in 0.5 ml of 0.1% trifluoroacetic acid (TFA), and centrifuged. The mixtures of peptides produced by the different enzyme digests were fractionated by reversed-phase high-performance liquid chromatography (RP-HPLC) on a Vydac 218 TP54 column (4.6 × 250 mm) at 45°C using a linear gradient of acetonitrile in 0.1% TFA. A gradient of 0–35% acetonitrile over 30 min was used routinely for the separation of peptides. Peptides which co-eluted or were poorly resolved were purified by rechromatography using a shallower gradient or by isocratic elution.

Sequences of longer peptides were determined with a Model 470A gas-phase sequencer (Applied Biosystems, U.S.A.). Most of the peptides were sequenced manually by a modified Edman degradation method (McKern *et al.*, 1983) using 3–5 nmol of peptide. Phenylthiohydantoins were identified by RP-HPLC. Peptides (~5 nmol) for amino acid analysis were hydrolysed under vacuum in 5.7 M HCl for 22 hr at 110°C. Analyses were performed on a Waters HPLC amino acid analyzer.

The S-carboxymethylated inhibitor protein and its clostripain digest peptide R1, which had a blocked N-terminus, were digested with pyroglutamate aminopeptidase (Boehringer-Mannheim) as described by Crimmins *et al.* (1988). The digest of clostripain peptide R1 was purified by RP-HPLC using the same conditions as for the original digest, to yield a main peak eluting at lower concentration of acetonitrile than the undigested peptide.

3. RESULTS AND DISCUSSION

The sequence of trypsin inhibitor-2 from *P. tetragonolobus* seed is shown in Fig. 1 along with the details of the overlapping peptides isolated and sequenced to establish the covalent structure. Digestion with trypsin and chymotrypsin yielded peptides from which most of the sequence of WBTI-2 could be readily established. Several regions of the sequence proved difficult to align, and lysyl endopeptidase and clostripain peptides provided the overlaps not readily obtained from the other digests. For example, lysyl endopeptidase peptide K2 provided overlaps for peptides T2 and T3a, and R1 and R2a, to confirm the alignment in the N-terminal region of the molecule (Fig. 1). Peptides T4 and T5 were overlapped with peptide R2b, which was produced by cleavage of the Arg^{24}-Pro^{25} bond by clostripain. Cleavage of Arg-Pro by clostripain has been described previously (Mitchell, 1968). Clostripain also produced a major cleavage of the Lys^{127}-Met^{128} bond. Unexpectedly, partial cleavage of the Arg^{24}-Pro^{25} bond was observed with trypsin, yielding peptides T3a and T3b (Fig. 1), and this cleavage was confirmed by the compositions of these two peptides. Both trypsin and chymotrypsin cleaved the Lys^{46}-Ser^{47} bond as noted previously in the case of WBA-1, a related Kunitz-type seed protein (Kortt *et al.*, 1989). The carboxyl-terminal residue was established as Thr^{182} from the sequence and composition data of T22, C23, and S14 (Fig. 1 and Table I).

The N-terminus of the protein was blocked and several peptides (K1, R1, and S[1+2]) with blocked N-termini were isolated. Amino acid analyses suggested that these peptides were deriived from the N-terminal region of the protein (Table I), and the composition of blocked peptide R1 corresponded to the sum of the compositions of blocked peptide K1

and peptide T2, which was aligned in this region. Digestion of peptide R1 with pyroglutamate aminopeptidase generated a peptide with a free-amino terminus, demonstrating that the N-terminal residue of blocked R1 was pyroglutamic acid. The pyroglutamate-aminopeptidase-digested peptide provided the sequence of residues 2–6 and overlapped T2 and S3 (Fig. 1), demonstrating unequivocally that R1 was the N-terminal peptide of the protein. In contrast, pyroglutamate-aminopeptidase digestion of the intact S-carboxymethylated protein did not remove the pyroglutamic acid from the N-terminus.

The sequence determination showed that WBTI-2 contains 182 amino acid residues, including four half-cystine residues, and has a molecular mass of 20,375 D, in agreement with estimates of M_r from SDS-PAGE and ultracentrifugal analysis (Kortt, 1979). The sequence composition agrees closely with the determined amino acid composition of the protein (Table I).

Although the N-terminus of WBTI-2 was blocked, gas-phase sequencer analysis of some preparations of the protein yielded traces of the sequence STFIP. This sequence occurs within the polypeptide

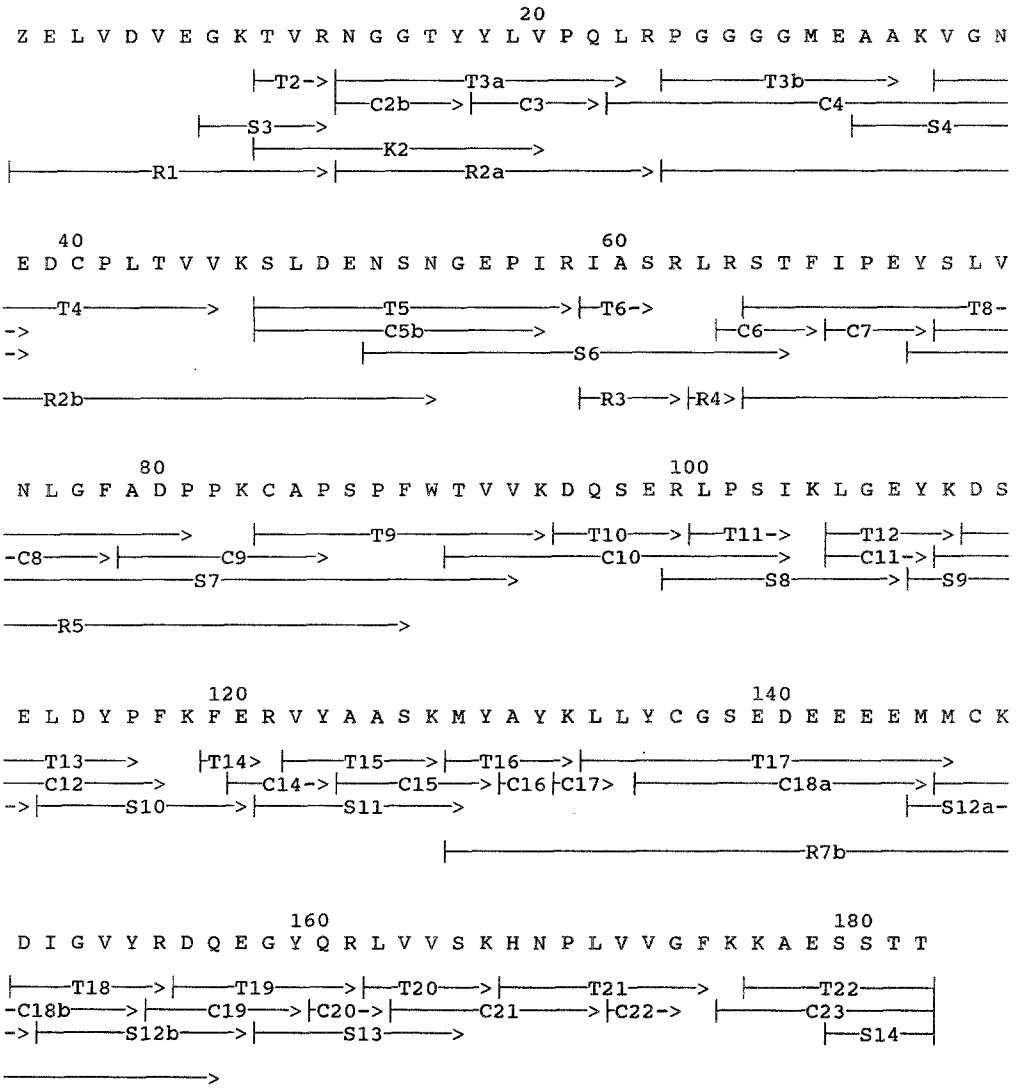

Fig. 1. Amino acid sequence of acidic winged-bean trypsin inhibitor (WBTI-2). Peptides sequenced to establish the structure are shown by solid lines. The notation Z is used for pyroglutamic acid (residue 1). Peptides are numbered from the N-terminus of the protein and T, C, S, K, and R signify tryptic, chymotryptic, staphylococcal protease, lysyl endopeptidase, and clostripain peptides, respectively.

Table I. Amino Acid Compositions of the Winged Bean Trypsin Inhibitor-2 (WBTI-2) and the Amino- and Carboxyl-Terminal Peptides

		N-terminal peptides			C-terminal peptides	
Amino acid	WBT1-2[a]	S(1+2)	K1	R1	T22	C23
Lys	13.6 (14)		1.0 (1)	0.9 (1)	1.1 (1)	1.9 (2)
His	1.1 (1)					
Arg	9.2 (9)			1.0 (1)		
Asp	16.7 (10),	1.0 (1)	1.0 (1)	0.9 (1)		
Asn	(6)					
Thr	7.0 (7)			0.9 (1)	1.9 (2)	1.8 (2)
Ser	13.7 (14)				1.8 (2)	1.8 (2)
Glu	23.1 (18)	3.2 (3)	3.3 (3)	2.8 (3)	1.2 (1)	1.1 (1)
Gln	(5)					
Pro	12.9 (12)					
Gly	16.0 (15)		1.1 (1)	1.1 (1)		
Ala	9.3 (9)				1.0 (1)	1.0 (1)
$\frac{1}{2}$Cys	4.0 (4)					
Val	15.7 (16)	1.9 (2)	1.9 (2)	2.8 (3)		
Met	3.0 (4)					
Ile	5.0 (5)					
Leu	15.1 (15)	1.0 (1)	1.1 (1)	0.9 (1)		
Tyr	11.1 (11)					
Phe	6.0 (6)					
Trp	1.0 (1)					
Total	182	7	9	12	7	8

[a] Amino acid analysis taken from Kortt (1979). Values are given as mol residue/mol peptide after 22 hr hydrolysis in 5.7 M HCl and are not corrected. Numbers in parentheses show values obtained from sequence studies.

chain (residues 65-69), and indicates that these inhibitor samples contain polypeptide chains cleaved at Arg^{64}-Ser^{65}. Cleavage at the reactive site of trypsin inhibitors by catalytic amounts of trypsin at acid *p*H has been reported (Ozawa and Laskowski, 1966; Joubert *et al.*, 1979). As the winged-bean trypsin inhibitors were isolated by elution from trypsin-sepharose at *p*H 2.0 (Kortt, 1979), this result provides direct evidence that Arg^{64} in WBTI-2 is the reactive site residue (Fig. 1). Sequence comparisons with other Kunitz-type trypsin inhibitors (Fig. 2) are consistent with this assignment.

The complete amino acid sequences of a number of Kunitz-type seed trypsin inhibitors have been determined, including three variants of the well-known inhibitor from soybean (Kim *et al.*, 1986), winged bean (Yamamoto *et al.*, 1983), *Erythrina latissima* (Joubert *et al.*, 1985), *Erythrina caffra* (Joubert and Dowdle, 1987), and *Adenanthera pavonina* (Richardson *et al.*, 1986). The sequence of a related Kunitz-type wheat endogenous α-amylase inhibitor (Maeda, 1986) and that of a specific Kunitz-type chymotrypsin inhibitor from winged bean (Shibata *et al.*, 1988) have also been described. The amino acid sequence of WBTI-2 is compared with some of the other inhibitor members of this family in Fig. 2. The four half-cystine residues in all these proteins are conserved; and in the cases of the Kunitz soybean trypsin inhibitor (Koide and Ikenaka, 1973), the winged-bean chymotrypsin inhibitor (Kortt, unpublished data), and the wheat endogenous α-amylase inhibitor (Maeda, 1986), the disulfide bridge-pairing has been established as Cys^{40}-Cys^{84} and Cys^{136}-Cys^{147} (Fig. 2). The same disulfide bridge arrangement is expected for WBTI-2. A comparison of the two trypsin inhibitors from winged bean shows that the acidic inhibitor (WBTI-2) is 10 residues longer than the basic inhibitor (WBTI-3). Five of the extra residues, four of which are glutamic acid, are located in the putative smaller disulfide-bonded loop in the C-terminal third of the WBTI-2 molecule. Two of the other additional residues are found at the C-terminus, and another two are at positions 124-125 in WBTI-2. These two winged-bean trypsin inhibitors

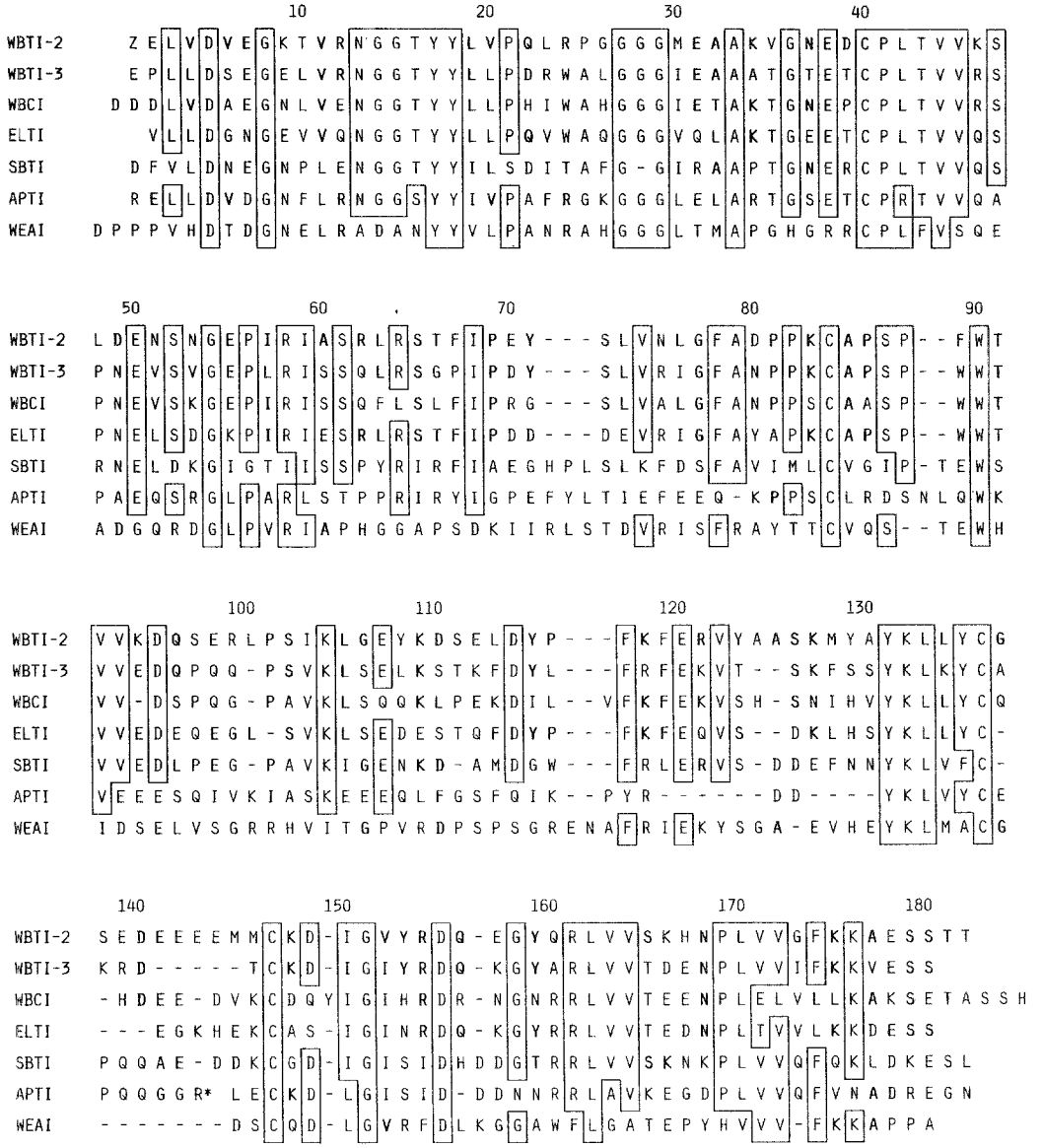

Fig. 2. Comparison of the amino acid sequences of acidic winged-bean trypsin inhibitor-2 (WBTI-2) with other seed Kunitz-type inhibitors. WBTI-3, basic winged-bean trypsin inhibitor-3 (Yamamoto *et al.*, 1983); WBCI, winged-bean chymotrypsin inhibitor (Shibata *et al.*, 1988); ELTI, *E. latissima* (DE-3) trypsin inhibitor (Joubert *et al.*, 1985); SBTI, soybean trypsin inhibitor Ti[a] (Kim *et al.*, 1986); APTI, *A. pavonina* trypsin inhibitor (Richardson *et al.*, 1986); and WEAI, wheat endogenous α-amylase inhibitor (Maeda, 1986). Residues conserved in at least 5 of the inhibitors are boxed and other residues identical in WBTI-2 and one or more of the other inhibitors are in bold type. Gaps (-) are introduced to maximize similarity. The reactive site arginine in the trypsin inhibitors (at position 64) is indicated by a cross; the C-terminus of the large chain of the *A. pavonina* trypsin inhibitor is indicated by an asterisk. The numbering of the residues refers to the WBTI-2 sequence.

show about 57% identity. WBTI-2 is one residue shorter than WBCI, the winged-bean chymotrypsin inhibitor, and has 50% identity with WBCI (Shibata *et al.*, 1988). The sequences of the three winged-bean Kunitz-type inhibitors sequenced to date show considerable diversity with only about 46% identity. This is in contrast to the three Kunitz soybean trypsin-inhibitor variants, which show only 1 and 8 residue differences (Kim *et al.*, 1986), and the two Kunitz-type trypsin inhibitors from *E. latissima* (Joubert *et*

Table II. Sequence Similarities in the Kunitz-Type Family of Seed Inhibitors

	Percent identity[a]					
	WBTI-3	WBCI	ELTI	SBTI	APTI	WEAI
WBTI-2 (182)	57	50	54	40	31	26
WBTI-3 (172)		54	59	37	33	26
WBCI (183)			52	40	30	25
ELTI (172)				38	31	26
SBTI (181)					37	29
APTI (176)						23
WEAI (180)						

[a] The percent identity between each pair of inhibitors was based on the 190 positions compared, including gaps introduced to maximize alignment of the sequences shown. Values in parentheses are the total number of residues in each inhibitor. Abbreviations for the inhibitors are defined in Fig. 2.

al., 1985) and *E. caffra* (Joubert and Dowdle, 1987), which have only 4 residue differences in their sequences.

Comparison of the seven Kunitz-type inhibitors (Fig. 2) shows that out of 190 sites compared, including gaps, 21 sites contain identical amino acid residues. The five inhibitors from the Lotoideae subfamily of the leguminosae share 54 sites with identical amino acids, but this decreases to 36 sites when the *A. pavonina* trypsin inhibitor from the Mimosoideae subfamily is included in the comparison. The sequence similarities between the different members of this family of Kunitz-type inhibitors are summarized in Table II. The winged-bean trypsin inhibitors show 54% (WBTI-2), 59% (WBTI-3), and 52% (WBCI) identity to the *E. latissima* trypsin inhibitor, but only about 40% identity to the soybean Kunitz trypsin inhibitor. This suggests that the *Psophocarpus* and *Erythrina* inhibitors are more closely related to each other than to the *Glycine* (soybean) inhibitor. For example, the sequence (residues 62–70) around the reactive site Arg^{64} in WBTI-2 is identical to that in ELTI (Fig. 2), suggesting that the inhibitory specificity of WBTI-2 will be the same as that described for ELTI (Joubert *et al.*, 1985). The similarity to the two-chain Mimosoideae trypsin inhibitor from *A. pavonina* is only about 30% and to the Kunitz-type inhibitor from wheat is only about 26% (Table II). The Mimosoideae subfamily of the leguminosae is considered to be the most primitive of the legumes (Heywood, 1971); and, therefore, the two-chain inhibitors of this subfamily can be considered to be more ancient members of the Kunitz-type inhibitor family.

The comparison in Fig. 2 shows that the reactive site Arg in SBTI (residue 64 in the alignment in Fig. 2) is followed by Ile–Arg as found for the *A. pavonina* trypsin inhibitor (APTI) and for the silk tree (*Albizzia julibrissin*) trypsin inhibitor, another Mimosoideae legume (Odani *et al.* 1980). This suggests that SBTI has retained the more "primitive" reactive site residues of Arg–Ile as observed in the Mimosoideae inhibitors, and that a change of Ile to Ser has occurred in other Lotoideae inhibitors. It is noteworthy that members of the Bowman–Birk family of double-headed protease inhibitors, which have been isolated from the Lotoideae subfamily of the leguminosae, have the reactive site residue followed in all cases by a serine residue (Joubert *et al.*, 1979). Another link between SBTI and APTI is the conservation of the sequence P^{138}–Q^{140} in the smaller disulfide loop of these inhibitors (Fig. 2). As more sequences of Kunitz-type inhibitors from the subfamilies of the leguminosae and other families, such as cereals, become available, the various relationships between these proteins will become clearer.

ACKNOWLEDGMENTS

The authors thank Mr. N. A. Bartone for carrying out the amino acid and PTH analyses.

REFERENCES

Crimmins, D. L., McCourt, D. W., and Schwartz, B. D. (1988). *Biochem. Biophys. Res. Commun.* **156**, 910–916.

Heywood, V. H. (1971). In *Chemotaxonomy of the Leguminosae* (Harborne, J. B., Boulter, D., and Turner, B. L., eds.), Academic Press, New York, pp. 1–29.

Joubert, F. J., and Dowdle, E. B. D. (1987). *Thromb. Haemostatis* **57**, 356–360.

Joubert, F. J., Kruger, H., Townshend, G. S., and Bates, D. P. (1979). *Eur. J. Biochem.* **97**, 85–91.

Joubert, F., Heussen, C., and Dowdle, E. B. D. (1985). *J. Biol. Chem.* **260**, 12,948–12,953.
Kim, S.-H., Hara, S., Hase, S., Ikenaka, T., Kitamura, K., and Kaizuma, N. (1986). *J. Biochem.* **98**, 435–448.
Koide, T., and Ikenaka, T. (1973). *Eur. J. Biochem.* **32**, 417–431.
Kortt, A. A. (1979). *Biochim. Biophys. Acta* **577**, 371–382.
Kortt, A. A. (1980). *Biochim. Biophys. Acta* **624**, 237–248.
Kortt, A. A., Tao, Z. T., and Rubira, M. (1983). *Phytochemistry* **22**, 767–768.
Kortt, A. A., Strike, P. M., and de Jersey, J. (1989). *Eur. J. Biochem.* **181**, 403–408.
McKern, N. M., O'Donnell, I. J., Inglis, A. S., Stewart, D. J., and Clarke, B. L. (1983). *FEBS Lett.* **164**, 149–153.
Maeda, K. (1986). *Biochim. Biophys. Acta* **871**, 250–256.
Mitchell, W. M. (1968). *Science* **162**, 374–375.
Odani, S., Odani, S., Ono, T., and Ikenaka, T. (1979). *J. Biochem.* **86**, 1795–1805.
Ozawa, K., and Laskowski, M., Jr. (1966). *J. Biol. Chem.* **241**, 3955–3961.
Richardson, M., Campos, F. A. P., Xavier-Filho, J., Macedo, M. L. R., Maia, G. M. C., and Yarwood, A. (1986). *Biochim. Biophys. Acta* **872**, 134–140.
Shibata, H., Hara, S., and Ikenaka, T. (1988). *J. Biochem.* **104**, 537–543.
Sri Kantha, S., and Erdman, J. W., Jr. (1983). *J. Amer. Oil Chem. Soc.* **61**, 515–525.
Yamamoto, M., Hara, S., and Ikenaka, T. (1983). *J. Biochem.* **94**, 849–863.