Primary Structure of Potato Kunitz-Type Serine Proteinase Inhibitor

**Tatyana A. Valueva**1,*, Tatyana A. Revina2, Vladimir V. Mosolov1 and Reinhard Mentele2

1 Bach Institute of Biochemistry, Russian Academy of Sciences, 171071 Moscow, Russia
2 Department of Clinical Chemistry and Clinical Biochemistry, University of Munich, D-80336 Munich, Germany

*Corresponding author

The serine proteinase inhibitor (PSPI-21) isolated from potato tubers (Solanum tuberosum L.) comprises two protein species with pl 5.2 and 6.3, denoted as PSPI-21-5.2 and PSPI-21-6.3, respectively. They were separated by anion exchange chromatography on a Mono Q FPLC column. Both species tightly inhibit human leukocyte elastase, whereas their interaction with trypsin and chymotrypsin is substantially weaker. The sequences of both PSPI-21-5.2 and PSPI-21-6.3 were determined by analysis of overlapping peptides obtained from the oxidized or reduced and S-pyridylethylated proteins after digestion with trypsin or pepsin. Both species of PSPI-21 are composed of two chains, named chains A and B, which are linked by a disulfide bridge between Cys(146) and Cys(157). The other disulfide bridge is located within the A chains between Cys(48) and Cys(97). The amino acid sequences of the large A chains of the two forms, consisting of 150 amino acids residues each, differ in a single residue at position 52. The small chains B, containing 37 and 36 residues in PSPI-21-6.3 and PSPI-21-5.2, respectively, have nine different residues. The entire amino acid sequences of the two inhibitors show a high degree of homology to the other Kunitz-type proteinase inhibitors from plants.

Key words: Chymotrypsin (EC 3.4.21.1)/ Human leukocyte elastase (EC 3.4.21.36)/ Trypsin (EC 3.4.21.4).

**Introduction**

Protein proteinase inhibitors are widely distributed in living organisms. The modern classification of these proteins is based on their amino acid sequences, active site structure, and mechanisms of action. Plant proteinase inhibitors are classified into 11 different structural groups (families) (Garcia-Olmedo et al., 1987; Valueva and Mosolov, 1999). One of them is the superfamily of the Kunitz soybean trypsin inhibitor. Kunitz-type proteinase inhibitors are mostly 20 – 24 kDa single chain proteins with four cysteine residues forming two disulfide bridges, and with one reactive site (Valueva and Mosolov, 1999). They have been found in plants belonging to diverse taxonomic groups (Garcia-Olmedo et al., 1987; Ishikawa et al., 1994; Valueva and Mosolov, 1999). Potato tubers contain various Kunitz-type proteinase inhibitors with a wide variety of specificities for target enzymes, such as trypsin and/or chymotrypsin inhibitors (Walsh and Twitchell, 1991; Suh et al., 1991), subtilisin inhibitor (Suh et al., 1991), cathepsin D and trypsin inhibitors (Mares et al., 1989; Strukelj et al., 1992), and papain and/or cathepsin L inhibitor (Krizaj et al., 1993). These proteins belong to a subfamily named the potato Kunitz-type proteinase inhibitor (PKPI) family (Ishikawa et al., 1994). The physiological role of PKPIs has not been established conclusively. They are likely to be involved in the protection of the potato plants by inhibiting proteases of invading organisms, various phytopathogens and insects (Walsh and Twitchell, 1991; Suh et al., 1991; Hildmann et al., 1992; Hansen and Hannapel, 1992; Valueva et al., 1998). Recently we reported the isolation from potato tubers and partial characterization of a 21 kDa protein, denoted as PSPI-21, that acts as a potent inhibitor of serine proteinases (Valueva et al., 1998). PSPI-21 deserves special attention since it is predominantly accumulated in potato tubers infected with Phytophthora infestans zoospores (Valueva et al., 1998). The PSPI-21 molecule consists of two polypeptide chains: a large A chain (16.5 kDa) and a small B chain (4.5 kDa), linked by a disulfide bond, and contains two reactive sites acting independently (Valueva et al., 1998, 1999). The present report describes the determination of the amino acid sequences of both chains of PSPI-21, and the localization of two disulfide bridges in its molecule, leading to the elucidation of its complete primary structure. The homology of PSPI-21 with other plant Kunitz-type proteinase inhibitors is discussed.

**Results**

The purified PSPI-21 isolated from potato tubers appeared to be homogeneous according to SDS-PAGE, reversed phase (RP) HPLC, and N-terminal sequence analysis (Valueva et al., 1998). However, two major protein species with pl 5.2 and 6.3, denoted as PSPI-21-5.2 and PSPI-21-6.3, respectively, were separated by isoelectric focussing in a pH gradient from 3 to 9 (Figure 1). Using ion exchange chromatography on a Mono Q FPLC column at pH 8.0, the inhibitory activity was separated into two major protein fractions with pl values of 6.3 and 5.2 (data not shown). Mass spectrometry revealed that the molecular
masses of the PSPI-21-6.3 and PSPI-21-5.2 are 20.262 ± 0.020 and 20.174 ± 0.020 kDa, respectively. Thus, the molecular masses of both proteins are close to the value of 21 ± 1 kDa determined previously for PSPI-21 by SDS-PAGE (Valueva et al., 1998) and gel chromatography (Valueva et al., 1999).

Both species of PSPI-21 were reduced with β-mercaptoethanol and alkylated with 4-vinylpyridine. Each protein was cleaved into two fragments, denoted S-pyridylethyl-(S-PE)-chains A and B, which were separated by RP-HPLC (data not shown). Thus, both PSPI-21-6.3 and PSPI-21-5.2 are composed of two polypeptide chains linked by a disulfide bond. According to mass spectrometry the molecular masses of the separated S-PE-chains A and B of PSPI-21-6.3 were 16.078 ± 0.016 and 4.282 ± 0.004 kDa and those of PSPI-21-5.2 were 16.396 ± 0.016 and 4.182 ± 0.004 kDa, respectively. The sum of these values are very close to the molecular masses of the native proteins (see above).

The strategy used for the determination of the amino acid sequence of the isolated chains of both proteins is shown in Figure 2. The N-terminal amino acid residues of chain A and B (NT) were determined by sequencing native PSPI-21-6.3 and PSPI-21-5.2. The amino acid sequences were completed and confirmed using overlapping peptides generated by proteolytic fragmentation of the S-PE-chains A with trypsin (PET) or pepsin (PEP), of the oxidized A chains with trypsin (OxT), and the S-PE-chains B modified with phenylisothiocyanate (PITC), with trypsin (PTH-PET).

The 32 N-terminal amino acid residues of the A chains of both proteins were identical, while the 30 N-terminal residues of the B chains were differed at positions 8, 25, and 26.

The peptide maps obtained by RP-HPLC separation of tryptic peptides of the S-PE-chains A of both PSPI-21-6.3 and PSPI-21-5.2 are shown in Figure 3. The elution profiles of the peptides of the two protein species are similar and differ only in three fractions, two of which are present in PSPI-21-6.3, and one in PSPI-21-5.2. Sequencing of the peptides marked with arrows in Figure 3 showed that the unique fraction found in a tryptic hydrolyzate of PSPI-21-5.2 (Figure 3, curve 1) contained one peptide with the sequence SPNSDAPCANGIFR. However, both fractions

---

**Fig. 1** Isoelectric Focussing of PSPI-21 Using a pH 3–9 Gradient. Lane 1, PSPI-21 (4 μg); lane 2, pl protein markers (top to bottom): trypsinogen, lectin B, bovine myoglobin, horse myoglobin, human carbonic anhydrase B, conalbumin, β-lactalbumin A (1 μg of each protein); pl values are given on the right.

**Fig. 2** Amino Acid Sequences of PSPI-21-6.3 (1) and PSPI-21-5.2 (2) and Strategy of Their Determination. (A), PSPI-chains A, (B), PSPI-chains B. The overlapping fragments used to deduce the amino acid sequences of both forms are represented by solid lines with arrows. N-terminal sequencing of the inhibitor molecules is indicated as NT. Peptides designated as OxT were obtained by enzymatic cleavage of oxidized chains A with trypsin. PET and PEP peptides resulted from tryptic and peptic fragmentation of S-PE-chains A, respectively. PE-PTHT peptides were obtained by tryptic cleavage of PTH-derivatives of S-PE-chains B. Peptides designated as NP1 and NP2 are from peptic hydrolysis of the native inhibitor molecules. * and **, disulfide bonds. Differing amino acids are boxed.
specific for PSPI-21-6.3 (Figure 3, curve 2) contained a single 14 residue peptide of the sequence SPNSDAP-CANGVFR. These peptides differ from each other by a single substitution of Ile(12) for Val(12) and correspond to positions 41 – 54 in the PSPI-21 protein sequence (Figure 2). Accordingly, the sequences of the A chains of both species differ only in the amino acid residue at position 52.

The peptide maps of the tryptic fragments of S-PE-chains B of the two species differed significantly and were not further analyzed. These data indicate that the major differences in amino acid sequences of the two species are expected in just this region of the molecules.

Amino acid analysis of the B chains showed that both species contain two Arg residues located at positions 19 and 20, as determined by N-terminal sequencing (Figure 2). The ε-amino groups of Lys residues of the S-PE-chains B were modified with PITC. Then the S-PE-PTh-chains B were selectively cleaved with trypsin at Arg residues. Sequencing the peptides obtained (PTh-PET in Figure 2) allowed us to reconstruct the amino acid sequence of both species of the B chains. The sequence of the B chain of PSPI-21-6.3 is different from that of PSPI-21-5.2 in nine positions: Leu(8) → Ser(8), Lys(25)-Asp(26) → Asn(25)-Glu(26), Ile(31)-Ser(32) → Val(31)-Leu(32), Lys(34)-Gln(35)-Val(36)-Gln(37) → Gln(34)-Glu(35)-Val(36).

In order to locate the disulfide bonds, native PSPI-21-6.3 was extensively digested with pepsin. The peptic cleavage mixture was resolved by RP-HPLC (Figure 4). In aliquot of each fragment preparation was subjected to quantitative amino acid analysis and N-terminal sequencing. Only two fragments (marked with arrows in Figure 4) contained two half-cystine residues and two different N-terminal sequences. Both fragments were oxidized by performic acid treatment and separated by RP-HPLC into two peptides each. The resolved peptides (NPEP1 and NPEP2, Figure 2) were then characterized by analysis of their amino acid sequences and assigned to the deduced sequence of the PSPI-21-6.3. In this way, an internal disulfide loop connecting cysteine residues at positions 48 and 97 of chain A was unequivocally established. A second disulfide bond connects the two polypeptide chains via Cys(146) of chain A and Cys(7) of chain B.

Both PSPI-21-6.3 and PSPI-21-5.2 inhibited human leukocyte elastase, trypsin, and chymotrypsin. The Kᵢ values for each target enzyme, calculated using the equation for slow-tight binding inhibition, are represented in Table 1. Obviously, both proteins are tight-binding inhibitors of HLE, but their interaction with chymotrypsin and trypsin is substantially weaker. The estimated Kᵢ values for PSPI-21-6.3 and PSPI-21-5.2 are very similar.

Discussion

The present sequence analysis revealed that the two forms of PSPI-21 described in our previous work (Valueva
et al., 1998) show a high degree of identity (95%) of their primary structures. They display a similar specificity towards target serine proteinases. These results allow us to suggest that PSPI-21-6.3 and PSPI-21-5.2 are iso-inhibitors encoded by two alleles at the same gene locus.

The sequence of PSPI-21-6.3 was used to search the GenBank NCBI database of protein sequences (Altschul et al., 1990). Ten known amino acid sequences of Kunitz-type serine proteinase inhibitors are compared in Figure 5. The sequence of PSPI-21-6.3 reveals a high degree of homology to those of the Kunitz-type proteinase inhibitors from potato tubers (Strukelj et al., 1992; Maganja et al., 1992; Hildmann et al., 1992; Ishikawa et al., 1994), soybeans (Hoffman et al., 1984), winged beans (Odani et al., 1996), latex of melon tree (Habu et al., 1992), cacao seeds (Tai et al., 1991), and sporamin A from sweet potatoes (Hattori et al., 1989). There are 45 invariant amino acid residues (24% identity) among the ten proteinase inhibitors. All of the 4 half-cystine residues of the Kunitz-type inhibitors and their location in the polypeptide chains are conserved. Furthermore, the proteins show conservation of residues that are characteristic of Kunitz-type inhibitors (Figure 5, shadowed residues). When PSPI-21-6.3 was compared individually with the inhibitors from potato tubers, such as two inhibitors of cathepsin D (PIGENI and PCDI), a proteinase inhibitor (pKEN14-28) and a wound-induced aspartic proteinase inhibitor from potato leaves, the number of invariant amino acid residues increased to 178 (95% identity), 172, 144, and 143 residues, respectively. On one hand, this confirms our previous suggestion (Valueva et al., 1998) that PSPI-21 belongs to the subfamily of PKPIs; on the other hand, it suggests that PSPI-21, PCDI, and pKEN14-28 may have evolved from a common ancestral inhibitor at a similar time.

As had been shown previously, PSPI-21 has a two-chain structure, which is different from those of PKPIs (pKEN14-28, PIGENI, PCDI and the wound-induced aspartic proteinase inhibitor), each of which has a single-chain structure. However, the Kunitz-type inhibitors in

![Fig. 5 Comparison of Amino Acid Sequences of PSPI-21 and Other Kunitz-Type Inhibitors.](image-url)

The amino acid sequences of PSPI-21-6.3 (1), a potato tuber proteinase inhibitor, pKEN14-28 (2) (Ishikawa et al., 1994), potato tuber cathepsin D inhibitors PIGENI and PCDI (3 and 4, respectively) (Strukelj et al., 1992; Maganja et al., 1992), a wound-induced potato leaf aspartic proteinase inhibitor (5) (Hildmann et al., 1992), a winged bean (Psophocarpus tetragonolobus L.) chymotrypsin inhibitor, WCl (6) (Odani et al., 1996), a serine proteinase inhibitor from of melon tree Carica papaya L. (7) (Habu et al., 1992), a 21 kDa protein trypsin inhibitor from Theobroma cacao L. seeds (8) (Tai et al., 1991), a sporamin A, a trypsin inhibitor from Ipomoea batatas L. (9) (Hattori et al., 1989), and a classical Kunitz-type soybean trypsin inhibitor (10) (Hoffman et al., 1984), were aligned. Amino acid residues are numbered according to the sequence of PSPI-21-6.3. Identical residues in these inhibitors are boxed; conserved residues are shadowed. The disulfide bonds between Cys(48)-Cys(97) and Cys(146)-Cys(157) are marked with * and **, respectively. The arrows mark the P1 residues of inhibitor reactive sites.
leguminous plants of the Mimosoideae subfamily are composed of two polypeptide chains (16 and 5 kDa) linked by a disulfide bond, which is located in the C-terminal region between residues 125 and 155 (Wu and Lin, 1993; Odani et al., 1996; Souza-Pinto et al., 1996). PSPI-21 and WCI, belonging to this inhibitor group (Odani et al., 1996), have a significant homology with 45 invariant amino acid residues (24% identity).

Kunitz-type serine proteinase inhibitors obey the 'canonical' mechanism of inhibition and bind enzymes in a substrate-like manner (Bode and Huber, 1992). The specific conformation of the reactive site loop, containing a scissile peptide bond P1-P1', is conserved by a hydrogen bonding network (Onesti et al., 1991). Kunitz-type inhibitors generally have one reactive site, although PIGENI, PCDI, and WCI might have two reactive sites (Shibata et al., 1988; Maganja et al., 1992; Strukelj et al., 1992). Our previous report (Valueva et al., 1999) indicated that PSPI-21 is a double-headed inhibitor. The PSPI-21 molecule contains two independent reactive sites, one of which is responsible for binding trypsin and has an Arg residue at the P1 position of the scissile bond. The other site binds HLE (or chymotrypsin) and has a Met residue at the P1 position (Valueva et al., 1999). According to the alignment, the Arg(67)-Phe(68) peptide bond (marked by an arrow in Figure 5) is located in the first reactive site of PSPI-21, which is disposed in the peptide loop, linked by the Cys(48)-Cys(97) disulfide bond. It should be noted that the sequences around the scissile bond located between the P3-P2 residues of the reactive site are identical in the structure of potato inhibitors (Figure 5). Therefore, the occurrence of Thr(64)-Pro(65) residues at the P3-P2 positions seems important because these residues are supposed to play an essential role in the formation of the specific structure of the reactive loop in 'canonical' inhibitors (Apostoluk and Otlewski, 1998).

There is a single Met(115) residue in the PSPI-21 molecule. On this basis we conclude that the P1 and P1' residues forming a second reactive site which binds HLE (or chymotrypsin) are Met(115)-Leu(116). The amino acid sequences between the Thr(114)-Leu(117) residues of the potato inhibitors are identical (Figure 5). We assume that the Cys(146)-Cys(157) disulfide bond connecting the A and B chains of PSPI-21 forms a specific reactive site loop. The Met(115)-Leu(116) peptide bond of the putative second reactive site is located within this loop. The formation of such a structure is supported by the ability of PSPI-21 to form triple complexes with target proteinases, in which one molecule of the inhibitor binds simultaneously one molecule trypsin and one molecule of chymotrypsin (Valueva et al., 1999). However, in order to establish the structure of the complexes of PSPI-21 with enzymes binding to the second reactive sites and the mechanism of their formation further studies are needed.

Materials and Methods

Materials

 Mature potato tubers (Solanum tuberosum L. cv. Isinskii) were stored at 4 °C for 2 to 4 months in the dark prior to use in experiments. PSPI-21 was isolated from potato tubers and purified to homogeneity according to the procedures described previously (Valueva et al., 1998). Enzymes: trypsin and pepsin of sequencing grade for protein digestion were purchased from Boehringer Mannheim GmbH Biochemical (Mannheim, Germany). The inhibitor activities were assayed with the following enzymes: HLE, trypsin, and α-chymotrypsin containing 85, 51, and 81% of active enzymes, respectively (Sigma Chemical, St. Louis, USA). α1-Antitrypsin containing 49.5% of active inhibitor, p-nitrophenyl-p'-guanidino-benzoate, N-trans-cinnamoylimidazole, and Nα-tosyl-phenylalaninechloromethylketone were from Sigma Chemical. Substrates were as follows: N-Suc-L-Ala-L-Ala-L-Val-pNa, Nα-Bz-L-Arg-pNa, and N-Suc-L-L-Ala-L-Val-pNa, (Ba-

Columns: Mono Q HR 5/5 column was from Pharmacia (Uppsala, Sweden), Superspher R P 8 column (125 × 2 mm) was from Merck (Darmstadt, Germany), Aquapore RP 300 column (30 × 2.1 mm) was from Applied Biosystems (Weiterstadt, Germany).

Separation of PSPI-21 Forms

Multiple forms of PSPI-21 were separated by ion exchange chromatography on a Mono Q FPLC column equilibrated with 20 mM Tris-HCl buffer, pH 8.0. Elution was performed with a linear gradient from 0 to 0.3 M NaCl in the starting buffer. After rechromatography under the same conditions the proteins were desalted and lyopholized.

Isoelectric Focussing

Isoelectric focussing in a pH gradient from 3 to 9 was carried out on the PhastSystem (Pharmacia) as described in the manufacturer’s instruction.

Assay of Enzyme Inhibitory Activity

The inhibitory activities were determined by measuring the residual activity of the target enzymes after preincubation with the inhibitor. The tested proteinases were assayed using suitable chromogenic substrates and optimal conditions (pH, temperature, salt concentration) (Erlanger et al., 1961; Wenzel et al., 1980; Powers et al., 1985). The actual concentration of active enzymes was determined by active site titration: trypsin with p-nitrophenyl-p'-guanidino-benzoate (Chase and Shaw, 1970); chymotrypsin with N-trans-cinnamoylimidazole (Shonbaum et al., 1961); HLE, with α1-antitrypsin, the active inhibitor concentration of which was determined by previous titration with trypsin of a known active site concentration. The inhibitor concentration and equilibri um dissociation constant (K) were determined by incubation of the enzymes with increasing inhibitor concentrations. Apparent K values were calculated assuming a slow tight-binding mechanism by non-linear fitting with the Enzfitter program (Knight, 1986).

Enzymatic Digestion

The pure inhibitors were oxidized with performic acid (Moore, 1963) or reduced with β-mercaptoethanol in 6 M guanidine chloride and afterwards S-alkylated with 4-vinylpyridine (Friedman et al., 1970). Oxidized or S-Pε-chains A were digested with trypsin (treated with Nα-tosyl-phenylalanine-chloromethylketone) in 25 mM Tris-HCl buffer, pH 8.0, for 2 h at 37 °C. S-Pε-
chains A and the native protein were digested with pepsin in 5% formic acid for 2.5 h at room temperature. To modify Lys residues, S-PE-chains B were treated with PITC, and phenylthiohydantion (PTH) derivatives were hydrolyzed with trypsin as described.

HPLC
The A and B chains and all peptide mixtures were separated by RP-HPLC on a Superspher RP 8 column using various linear gradients of 60% acetonitrile containing 0.1% trifluoroacetic acid. The flow rate was 0.3 ml/min. Absorbance was monitored at 206, 254, and 280 nm.

Determination of the Molecular Mass
The molecular mass of proteins and peptides was measured by mass spectrometry. Mass spectra were obtained with a tandem quadrupole mass spectrometer API III (Sciex, Thornhill, Canada) equipped with an atmospheric pressure ionization source. Average molecular masses were calculated from spectra containing multiple charged ions (Mann et al., 1989). The instrument was calibrated with the ammonium adduct ions of polypropylene glycol.

Amino Acid Analysis
Amino acid analysis was performed on a Biotronik LC 5000 high performance analyzer system (Puchheim, Germany) after acid hydrolysis by 5.7 M HCl at 110 °C for 20 h.

Sequence Analysis
The amino acid sequences were determined on an Applied Biosystems model 475A protein gas-liquid sequencer. PTH amino acids were identified in a model 120A PTH amino acid analyzer (Applied Biosystems).

Acknowledgements
Thanks are due to Prof. Hans Fritz for hosting T.A. Valueva at his laboratory of the Munich University (Germany) during an extended period of time and fruitful discussions and to Dr. Ennes A. Auerswald (Munich, Germany) for many helpful suggestions and discussions.

References


Received July 18, 2000; accepted October 26, 2000