the FEBS Journal

Investigation of peptide splicing using two-peptide-chain analogs of trypsin inhibitor SFTI-1

Natalia Karna, Dawid Dębowski, Agata Gitlin, Anna Łęgowska and Krzysztof Rolka

Department of Bioorganic Chemistry, Faculty of Chemistry, University of Gdańsk, Poland

Keywords
inhibitors; intermolecular disulfide bridge; peptide splicing; serine proteinases; SFTI-1

Correspondence
D. Dębowski, Faculty of Chemistry, University of Gdańsk, Wita Stwosza 63, 80-952 Gdańsk, Poland
Fax: +48 58 523 54 72
Tel: +48 58 523 50 91
E-mail: ddebowski@chem.univ.gda.pl

(Received 10 July 2013, revised 28 August 2013, accepted 17 September 2013)

doi:10.1111/febs.12542

This study examines peptide splicing catalyzed by serine proteinases. A series of two-peptide-chain analogs of trypsin inhibitor SFTI-1 were designed and synthesized via the solid-phase method. All consisted of two peptide chains (also called N- and C-terminal fragments) joined together by one disulfide bridge. The analogs were incubated with bovine β-trypsin or bovine α-chymotrypsin. Analysis of MS data analysis showed that, after enzyme-catalyzed degradation of the single peptide bond between the Lys and Ser residues located at the C-terminus of the C-terminal peptide chain, a new peptide bond was formed. This bond brought together the separated peptide chains, and, as a result, monocyclic SFTI-1 was recovered. This proteolytic route of peptide rearrangement appears to be similar to peptide splicing catalyzed by proteasomes. However, the proteasome is much more complex than 'classical' serine proteinases.

Introduction

SFTI-1, the trypsin inhibitor isolated from the seeds of sunflower (*Helianthus annuus*), is currently the smallest circular peptide consisting of proteinogenic amino acids derived from natural sources [1]. It has attracted the attention of many research groups since its discovery in 1999. Despite its small size (the circular backbone consists of 14 amino acids, and the structure is constrained by the disulfide bridge, Fig. 1), SFTI-1 is a more potent trypsin inhibitor than larger members (60–90 amino acids) of the Bowman–Birk family of inhibitors.

The inhibitor displays a well-defined, rigid three-dimensional structure that is responsible for its high proteolytic stability. Its monocyclic derivative, which lacks the peptide bond between Gly1 and Asp14 (disulfide bridge only), retains strong inhibitory activity [3,4]. A P_1–P_1' reactive site is located between the Lys5 and Ser6 residues. In the case of canonical inhibitors, the P_1 residue interacts within the S_1 cavity of the enzyme, accounting for up to 50% of the inhibitor–enzyme contacts. Therefore, this residue is largely responsible for SFTI-1 specificity. Structure–activity relationship studies have shown that a substitution in this position may dramatically affect specificity [5]. This has opened the way to design inhibitors of physiologically relevant proteinases: chymotrypsin, human neutrophil elastase, cathepsin G, matriptase, β-tryptase, proteinase K and human kallikrein-related peptidase 4. Recently, Cascales *et al.* [6] reported that fluorescent Alexa 488-labeled SFTI-1 internalized into endosomal compartments of human MCF-7 breast cancer cells. The cellular uptake of SFTI-1 is comparable with that of Alexa-labeled *Momordica cochinchinensis* trypsin inhibitor II; however, the mechanism of penetration appears to be different. It has been proposed that SFTI-1 and two other cyclic peptides (*Momordica cochinchinensis* trypsin inhibitor I and kalata B1) constitute a new family of cyclic cell-penetrating peptides. This result is consistent with our preliminary findings [5]. We have observed that analogs of monocyclic SFTI-1 (with the disulfide bridge) with fluorescent tags are able to enter cancer cells [5], and more detailed research is in progress.

Abbreviations
Anb-NH_2, 5-amino-2-nitrobenzoic acid; OMTKY-3, turkey ovomucoid third domain inhibitor; SFTI-1, sunflower trypsin inhibitor 1.

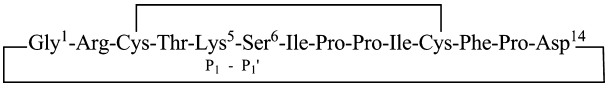

Fig. 1. The primary structure of the native, bi-cyclic trypsin inhibitor SFTI-1. The reactive site is located between Lys5 (the P_1 residue according to Schechter & Berger [2] notation, mostly responsible for inhibitory specificity) and Ser6 (P_1').

Recently, Colgrave *et al.* [7] studied the proteolytic degradation and trypsin-assisted cyclization of a series of disulfide-deficient or acyclic SFTI-1 analogs. Interestingly, for the disulfide-deficient mutant in which both Cys residues were replaced synthetically by 2-Aminobutyric acid (Abu), cleavage after the Arg2 and Lys5 residues yielded a cyclic undecapeptide (Fig. 2A).

Similarly, we have described the proteolytic processing of double-sequence SFTI-1 analogs, cross-linked by a single disulfide bridge between two flanking Cys3 and Cys25 residues [8]. After incubation with an appropriate enzyme (trypsin or chymotrypsin), the structure of monocyclic SFTI-1, with a disulfide bridge only (or its analog [Phe5]SFTI-1) was restored [8] (Fig. 2B). This process resembles peptide splicing in proteasomes [9] (Fig. 3), a large multi-enzymatic complex involved in the degradation of cellular proteins. Splicing contributes to the diversity of peptidic antigens presented by major histocompatibility complex class I [9–11]. Hanada *et al.* [9] analyzed the antigenicity of peptides derived from fibroblast growth factor 5. They found that one particular nonapeptide (NTYASPRFK) was recognized by C2 cytotoxic T lymphocytes. Interestingly, this peptide consisted of two fragments (a pentapeptide and a tetrapeptide) that were separated in the primary sequence of fibroblast growth factor 5 by 40 amino acid residues. Vigneron *et al.* [10] described a second example of peptide splicing. CD8 T lymphocytes were found to recognize a nonapeptide on melanoma cells, presented by human lymphocyte antigen A32. Again, this peptide comprises two non-contiguous fragments of melanocytic glycoprotein gp100^{PMEL17}. The results of MS analysis of products obtained after incubation of proteasomes with a set of appropriately designed peptides led to the conclusion that the energy required for new peptide bond formation is recovered from the ester bond formed between the hydroxyl group of enzyme Thr and the carboxyl group of the tripeptide RTK. Peptide splicing occurs in the catalytic chamber of a proteasome, which prevents rapid diffusion of peptides released from the protein or peptide substrate. Therefore, the concentration of coupled peptides is high. In the case of the peptide splicing described by us [8], two fragments from spliced peptides were covalently linked by the disulfide bridge, whereas, in the study by Colgrave *et al.* [7], splicing occurred on a homodetic peptide. However, the biological relevance (if any) of peptide splicing by serine proteinases is still unknown. It may be speculated to occur *in vivo,* producing cyclic peptides via enzymatic pro-

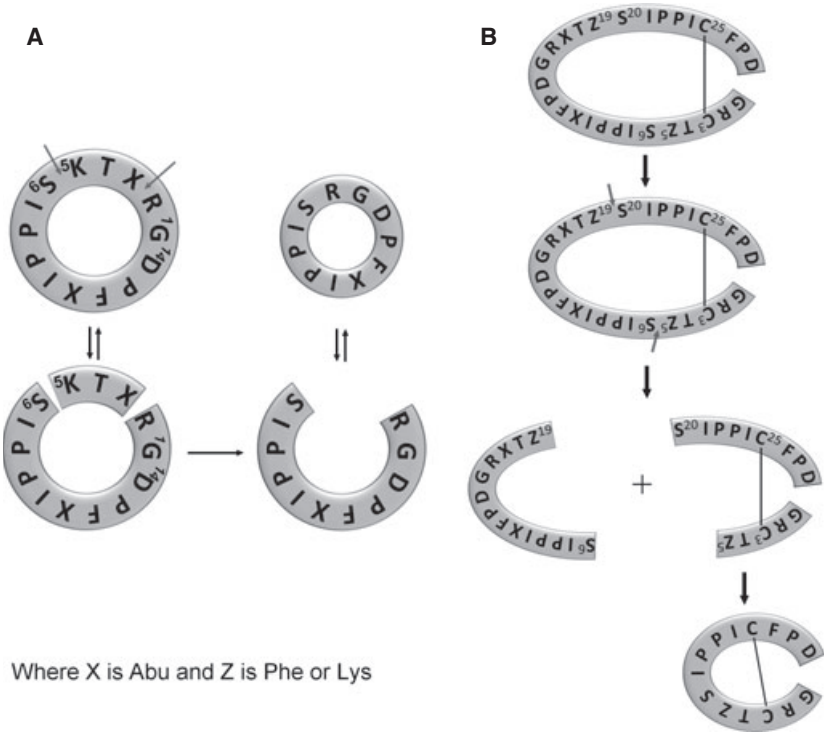

Fig. 2. Trypsin- or chymotrypsin-assisted proteolysis and head-to-tail cyclization of the homodetic SFTI-1 analog [7] (A) and the double-sequence SFTI-1 analogs [8] (B). In the latter case, the two reactive sites P_1–P_1' (Lys/Phe5–Ser6 and Lys/Phe19–Ser20) are cleaved, the middle fragment (the inserted peptide) is removed, and the new reactive P_1–P_1' bond (X5–Ser20) is formed.

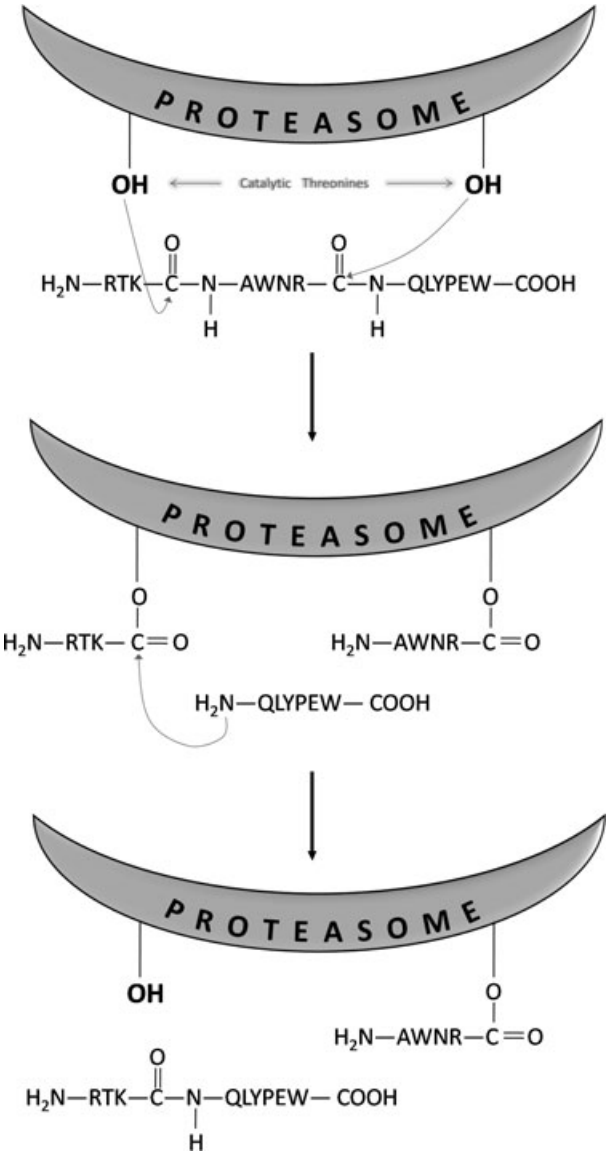

Fig. 3. Proteasome-catalyzed mechanism of peptide splicing [9].

cessing of their precursors generated by standard translation. It has been suggested that protease-catalyzed protein splicing is a type of post-translational modification [12]. On the other hand, such a system may be utilized for introduction of site-selective synthetic moieties into proteins or cancer cell compounds as insert peptides (intervening sequence). Such peptides may serve as therapeutic or diagnostic agents.

As the mechanism of peptide splicing driven by serine proteinases has not been investigated thoroughly, the present work aims to study the mechanism of enzymatic proteolysis and re-synthesis of SFTI-1 analogs. All the designed analogs comprise two separate peptide chains connected by a disulfide bridge (Fig. 4). To the best of our knowledge, these types of compounds, which may be considered heterodimers, have not so far been the subject of enzymatic studies. The starting structure for compounds **1**–**7** was monocyclic SFTI-1 with a disulfide bridge only. Peptide **1** is the analog of monocyclic SFTI-1 with the hydrolyzed P_1–P_1' reactive site. Peptides **2**–**4** contain a 5-amino-2-nitrobenzoic acid (*Anb*-NH_2) moiety attached to either the Lys or Phe residues located in the P_1 position, which determines the substrate specificity. In our previous work, we have shown that *Anb*-NH_2 is an excellent chromophore attached to substrates of serine proteinases [13]. In the present study, we wished to obtain substrates of trypsin (**2** and **3**) and chymotrypsin (**4**), which, after incubation with the enzyme, would be converted to inhibitors. In the remaining three peptides (**5**, **6** and **7**), one or two additional residues (Ser and/or Lys) are attached. These modifications increase the number of peptide bonds that may be cleaved by trypsin.

Results

All the studied SFTI-1 analogs consisted of two peptide chains (the C- and N-terminal peptide fragments) linked by a disulfide bridge. Synthesis of such compounds is not a simple task. The inter-molecular disulfide bond was introduced by a modification of the method described by Hoffmann and Otvos [14]. An active derivative, comprising a C-terminal fragment of SFTI-1 and 2,2′-dithiodipyridine, was combined with an N-terminal fragment of the inhibitor. Finally, an appropriate heterodimer was obtained (Fig. 5). The active derivative was stable, but the total yield of the reaction was moderate (approximately 20%). The

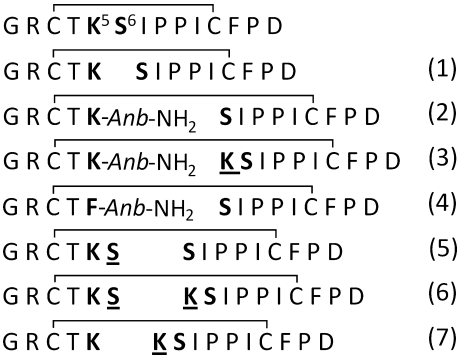

Fig. 4. The primary structures of two-peptide-chain SFTI-1 analogs. The Lys and Ser residues forming the inhibitor's P_1–P_1' reactive site are shown in bold. The inserted amino acid residues and the chromophore moiety *Anb*-NH_2 (amide of 5-amino-2-nitrobenzoic acid) are underlined.

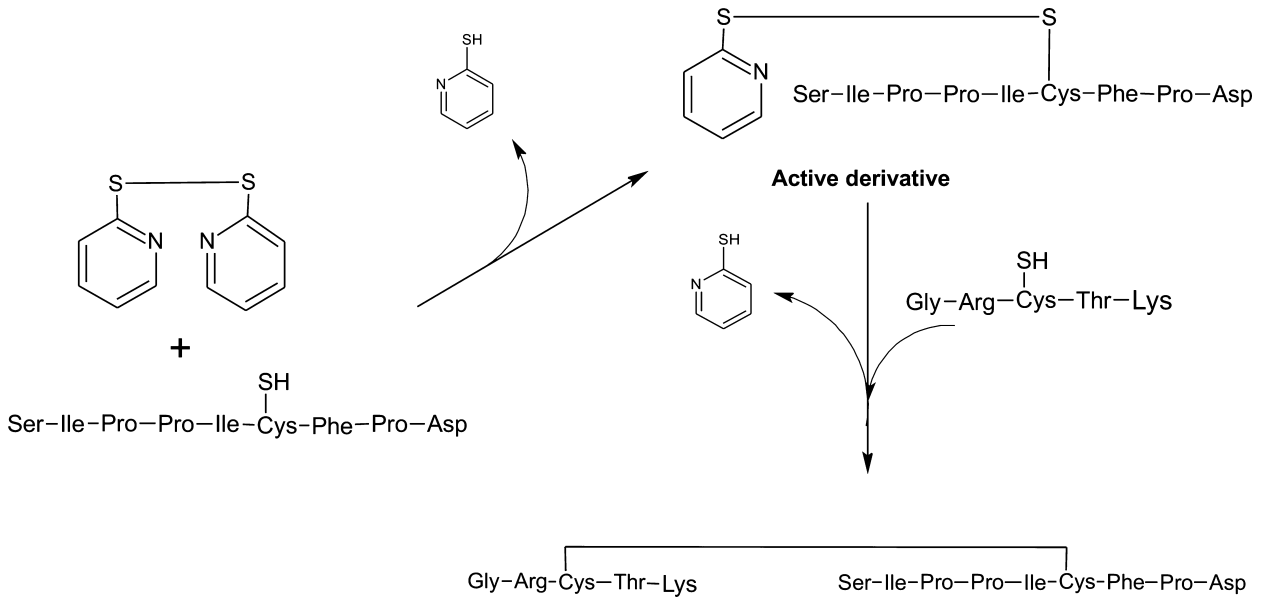

Fig. 5. Scheme of inter-molecular disulfide-bridge formation in peptide **1** using 2,2′-dithiodipyridine.

physico-chemical properties of the synthesized two-peptide-chain SFTI-1 analogs, together with their inhibitory activity, measured as K_a values, are summarized in Table 1.

According to the initial analysis, the analogs displayed varied levels of activity against trypsin (Fig. 6). The most potent were **1** and **5**, and the least active was **3**. Moreover, the inhibitory activities were time-dependent for all the peptides tested and increased clearly over time. The association constants (K_a) were determined for inhibitors **1**, **2** and **5** in assays with trypsin, and for compound **4** in an assay with chymotrypsin (Table 1).

The values of K_a were in the range of 10^5–10^7 M^{-1}, at least two orders of magnitude lower than that of the monocyclic SFTI-1 used as a reference inhibitor (Fig. S1). It is worth noting that most of the analogs studied were proteolytically degraded when incubated with experimental enzymes; therefore, the determined K_a values do not reflect the interaction between the enzyme and the individual inhibitor.

Despite our expectations, analogs **2** and **4** with the *Anb*-NH_2 moiety at the C-termini of their N-terminal fragments were proteolytically resistant. No signal corresponding to *Anb*-NH_2 was visible in the MS spectra after incubation with trypsin and chymotrypsin,

Table 1. Physicochemical properties of SFTI-1 analogs.

Analog	Calculated molecular weight	Molecular weight determined in this experiment[a]	RT (min)[b]	K_a (M^{-1})	
				Trypsin	Chymotrypsin
SFTI-1 (monocyclic)	1531.8	1531.8	18.4	$9.9 \pm 1.1 \times 10^9$	$4.9 \pm 1.4 \times 10^6$
1	1548.8	1549.7	17.8	$9.0 \pm 1.0 \times 10^5$	
2	1713.9	1713.8	17.1	$6.0 \pm 1.1 \times 10^5$	
3	1842.1	1842.8	15.1	ND	
4	1732.9	1733.4	20.0		$1.1 \pm 0.6 \times 10^7$
5	1636.9	1637.3	15.8	$7.2 \pm 1.6 \times 10^6$	
6	1763.8	1763.0	16.4	ND	
7	1678.0	1678.6	14.9	ND	

RT, retention time; ND, not determined.

[a] The molecular weights of the peptides were determined using a Bruker Biflex III MALDI-TOF spectrometer. The mean isotopic values are given.

[b] HPLC was performed as described in Experimental procedures. Linear gradient 10–90% B in 40 min was applied.

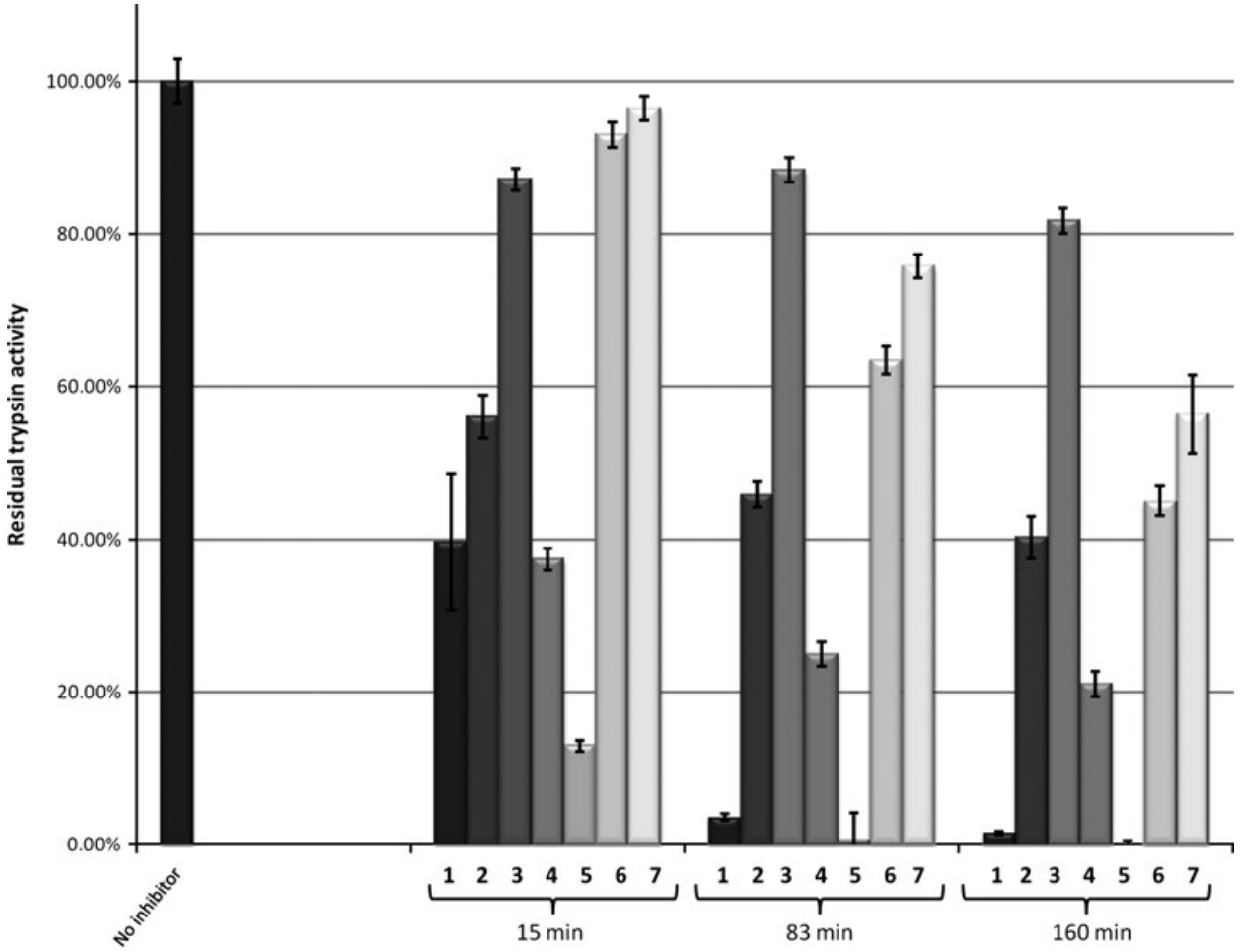

Fig. 6. Residual activity of bovine β-trypsin and chymotrypsin (compound **4**) incubated with SFTI analogs. Three incubation periods were used: 15, 83 and 160 min. The concentrations of all inhibitors were 7.35 μM. The concentration of trypsin was 0.14 μM, i.e. an approximately 52-fold molar excess of inhibitors over the experimental enzyme. The concentration of chymotrypsin was 0.11 μM.

respectively. Our extensive studies on proteinase substrates clearly show that *Anb*-NH_2 is a very useful leaving group when placed in the P_1' position [13]. Moreover, considering monomeric pentapeptides alone (Gly-Arg-Cys-Thr-Lys(Phe)-*Anb*-NH_2), the bonds between Lys or Phe and *Anb*-NH_2 are cleaved rapidly (after 15 min incubation) by trypsin or chymotrypsin, respectively (Fig. S2). Also, the truncated peptide without the N-terminal Gly-Arg dipeptide appears to be a trypsin substrate. Unlike peptides **2** and **4**, compound **3** was cleaved by the enzyme. It is noteworthy that, compared to analog **2**, analog **3** has only one additional Lys residue at the N-terminus of its C-terminal fragment. According to the MS data (Fig. 7A), three additional, distinct signals were identified after 1.5 h incubation with trypsin. The first signal (*m*/*z* 1629.5) corresponded to the truncated product without the N-terminal Gly-Arg dipeptide, and the second one (*m*/*z* 1485.3) corresponded to the product without the N-terminal dipeptide and Lys residue. The third signal (*m*/*z* 1336.8) corresponded to the compound formed after further enzymatic processing (cleavage of the *Anb*-NH_2 leaving group). The MS signal intensity for the second product (*m*/*z* 1485.3) increased with time. It is likely that digestion of the peptide bond between Lys and Ser (at the C-terminal fragment) stimulated enzyme-catalyzed cleavage of the *Anb*-NH_2 group.

Surprisingly, peptide **1**, i.e. the modified version of monocyclic SFTI-1 (with cleaved P_1–P_1' reactive site), remained unchanged during 24 h incubation with bovine β-trypsin under applied experimental conditions (according to the MS results). No enzymatic synthesis of the Lys-Ser peptide bond was observed. Marx *et al.* [15] reported that, in the case of native SFTI-1, the ratio between intact and modified inhibitors is 9:1, regardless of whether or not the inhibitor, intact or

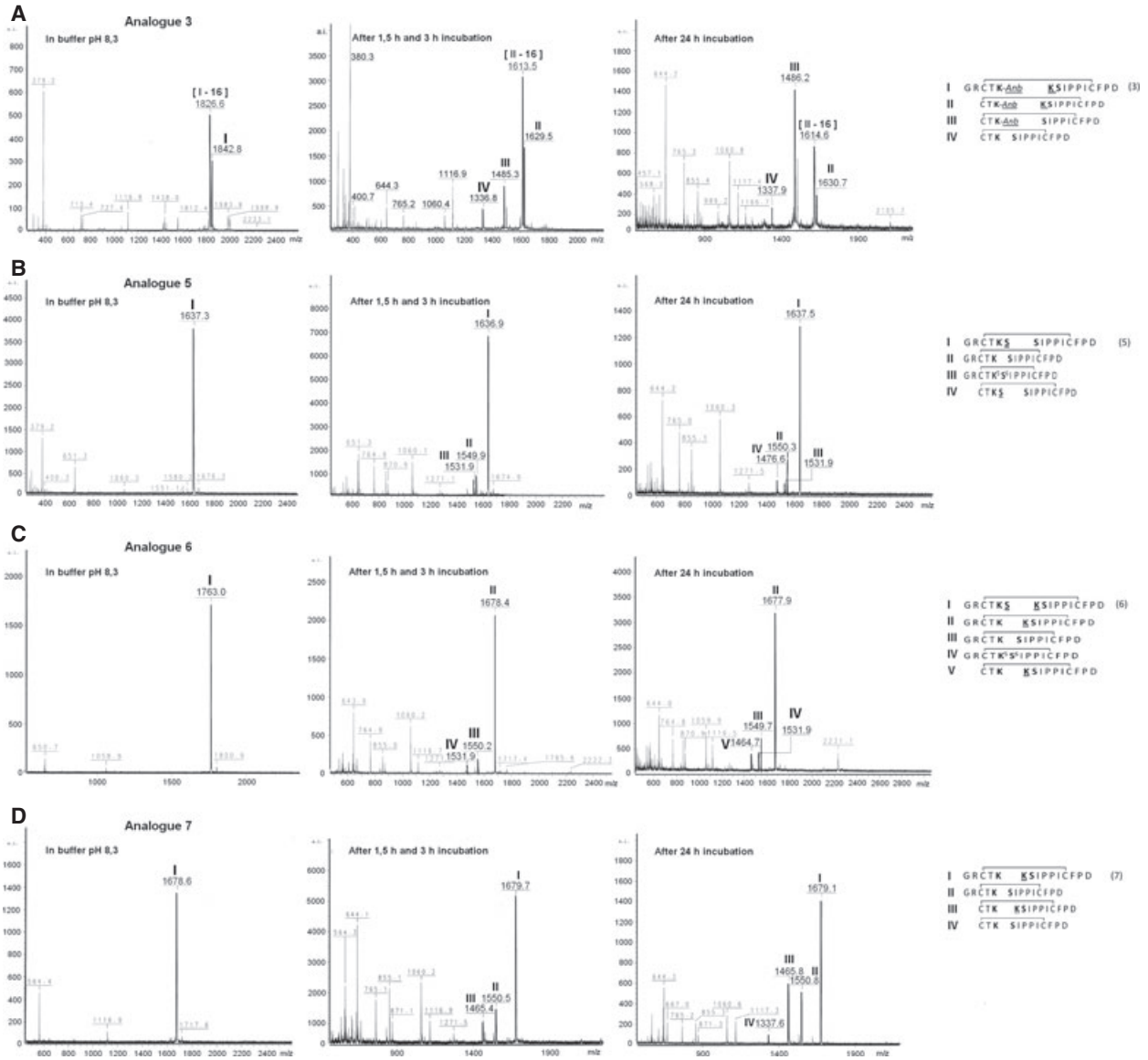

Fig. 7. MS analysis of the products after incubation of SFTI-1 analogs with bovine β-trypsin. The signals that correspond to the starting peptides are indicated as '**I**'. For all compounds with the *Anb* moiety (e.g. Analogue 3), despite their HPLC homogeneity, an additional peak I-16 (*m*/*z* 1826.5) appears in the MS spectra.

modified (without the Lys5–Ser6 bond), was incubated with trypsin. The conformational study indicated that the three-dimensional structure of the modified SFTI-1 analog, although less defined, resembles the structure of the wild SFTI-1. Although the peptide bond between Lys5 and Ser6 was not recovered, peptide **1** reduced the trypsin activity (Fig. 6), with a K_a of 9.0×10^5 M^{-1}. An extension at the C-terminus of the N-terminal pentapeptide by Ser (compound **5**) yielded a trypsin inhibitor with a K_a one order of magnitude higher (7.2×10^6 M^{-1}). MS analysis of **5** revealed that, after incubation with trypsin, two products with lower molecular masses were present (Fig. 7B). The first product (*m*/*z* 1549.9) was formed after removal of Ser residues; the second one (*m*/*z* 1531.9) corresponded to re-synthesized monocyclic SFTI-1.

Enzyme-catalyzed hydrolysis and re-synthesis was also observed in the case of analog **6**, which contains Lys–Ser peptide bonds on both peptide chains (Fig. 7C). After 1.5 h incubation with the enzyme, truncated products were found, either without one Ser residue at the C-terminus of the pentapeptide (*m*/*z* 1678.4) or without both Ser residues (*m*/*z* 1550.2). Moreover, the appearance of a peak at *m*/*z* 1531.8

clearly indicated that monocyclic SFTI-1 was re-synthesized successfully. If incubation time was prolonged up to 24 h, a new peak appeared (*m*/*z* 1464.7), corresponding to the analog of monocyclic SFTI-1 lacking the N-terminal Gly-Arg dipeptide. The proteolytic pattern of analog **7** with only one Lys–Ser peptide bond (at the N-terminus of the C-terminal fragment) was different than that of analogue 6. (Fig. 7D). Comparing MS data recorded for peptides **6** and **7**, the peptide bond between Arg2 and Cys3 was more susceptible to hydrolysis in the latter compound. The corresponding product (*m*/*z* 1465.4), together with the compound without the Ser residue (*m*/*z* 1550.5), was observed after 1.5 h incubation. It is interesting that, in contrast to peptides **5** and **6**, peptide splicing was not observed for peptide **7**.

Discussion

Among the seven two-peptide-chain analogs of SFTI-1 studied, trypsin-catalyzed peptide splicing was observed only for analogs **5** and **6**. In both cases, monocyclic SFTI-1 (with the disulfide bridge) was re-synthesized. The formation of a new Lys–Ser peptide bond (the P_1–P_1' reactive site) was preceded by hydrolysis of the Lys–Ser peptide bond located at the C-terminus of the N-terminal fragment (hexapeptide). According to the catalytic mechanism of serine proteinases, nucleophilic attack by the hydroxyl group of Ser195 of the enzyme on the carbonyl carbon of the peptide bond initially yields an acyl–enzyme intermediate. This is followed by hydrolytic cleavage of an ester bond. For analogs **5** and **6**, the ester bond is formed between trypsin and the carbonyl carbon of the Lys residue on the N-terminal fragment (hexapeptide). The N-terminal amino group of the Ser on the C-terminal fragment in analog **6** after removal of the N-terminal Lys residue is presumably in sufficiently close contact to compete with water molecules during nucleophilic attack on the acyl–enzyme intermediate. Finally, a new Lys–Ser peptide bond is synthesized.

As mentioned above, enzymatic re-synthesis of the monocyclic SFTI-1 is not observed after incubation of analog **7** with trypsin. In contrast to analogs **5** and **6**, the ester bond is formed between trypsin and the carbonyl carbon of the Lys residue on the C-terminal fragment. The bond between Lys and Ser is cleaved, and the Lys residue is lost. The acyl–enzyme intermediate is not covalently connected to the N-terminal peptide; therefore, nucleophilic attack by a free peptide amino group is less likely.

It may be assumed that the mechanism of peptide splicing catalyzed by trypsin is similar to that described for proteasome (see Fig. 3) [9]. In this case, splicing takes place in the enzyme catalytic chamber, which maintains spliced peptides in close contact and high concentrations. Moreover, in contrast to the double-sequence SFTI-1 analogs [8], the splicing observed in analogs **5** and **6** is not a dominant proteolytic route, and peaks corresponding to monocyclic SFTI-1 have low intensity. In the case of double-sequence SFTI-1 analogs, spliced peptides interact with the enzyme via both prime and non-prime segments, whereas such interactions occur through one of the segments only for analogs **5** and **6**. The suggestion that the energy required to establish a new peptide bond is recovered from proteolysis is supported by the fact that splicing was not observed for analog **1**. However, in light of the discussed splicing mechanism, the results obtained for peptides **2**–**4** are rather intriguing. The highly scissile bonds, Lys–*Anb* and Phe–*Anb*, found in proteinase substrates [13] appear to be proteolytically resistant when introduced into the inhibitor binding loop. It may be speculated that the interaction between the proteinase and inhibitor prime positions prevents access of the discussed amide bond (Lys–*Anb* and Phe–*Anb*) to the catalytic triad of the enzyme. As a result, the synthesized SFTI-1 analogs inhibit trypsin (compound **2**) with a K_a of 6.0×10^5 M^{-1} or chymotrypsin (compound **4**) with a K_a of 1.1×10^7 M^{-1}.

It is worth noting that analogs **5** and **6**, which are extended by Ser at the C-terminus of the N-terminal fragment, differ in terms of the hydrolytic rate of the Lys–Ser bond. With regard to analog **5**, the intact peptide is still present after 24 h of incubation, whereas rapid and complete hydrolysis of the Lys–Ser bond (even after a 1.5 h incubation with trypsin) is observed for analog **6**. This suggests that the presence of Lys at the N-terminus of the C-terminal peptide fragment increases the proteolytic susceptibility of the Lys–Ser bond of the N-terminal fragment. We speculate that this may be a result of electrostatic interactions of the Lys residue with Asp102 of the trypsin substrate pocket, which, in turn, brings the Lys–Ser scissile bond closer to the catalytic triad of the enzyme.

The results indicate that the proteasome-catalyzed mechanism of peptide splicing may also be applied to the smaller enzymes, e.g. trypsin. In addition, a new product (a peptide chain) may be formed not from one but from two chains linked by a disulfide bridge. The energy required for new peptide bond formation is recovered from the ester bond of the acyl–enzyme intermediate. The peptide models used for these studies were based on an SFTI-1 inhibitor tightly bound to trypsin. This allows specific interactions between enzymes and the studied peptides.

Experimental procedures

Peptide synthesis

All peptides were synthesized manually via the solid-phase approach using Fmoc (fluorenyl-9-methoxycarbonyl) chemistry. The following amino acid derivatives (GL Biochem Shanghai Ltd, China) were used: Fmoc-Gly, Fmoc-Arg (Pbf), Fmoc-Cys(Trt), Fmoc-Thr(*t*Bu), Fmoc-Ser(*t*Bu), Fmoc-Ile, Fmoc-Pro, Fmoc-Phe and Fmoc-Asp(O*t*Bu). The C-terminal Fmoc-Asp(O*t*Bu) was attached to 2-chlorotrityl chloride resin loading of Cl 1.46 mmol·g^{-1} Calbiochem/Novabiochem AG, Läufelfingen, Switzerland) in the presence of an equimolar amount of N,N-diisopropylethylamine (DIPEA) in anhydrous dichloromethane. Unless stated otherwise, the protected amino acid derivatives were coupled using equimolar mixtures of *N*,*N*′-diisopropylcarbodiimide and 1-hydroxybenzotriazole. Removal of the Fmoc group was performed using 20% piperidine in dimethylformamide (DMF) N-methyl-2-pyrrolidone solution (1:1 v/v) with addition of Triton X-100. In the case of peptides with the amide of 5-amino-2-nitrobenzoic acid (*Anb*-NH_2) attached to their C-termini, TentaGel^{TM} S RAM (loading $\sim$0.23 mmol·g^{-1}) Rapp Polymere GmbH, Tuebingen, Germany.) was used as the solid support. The *Anb* moiety was attached to the resin using the procedure described by Hojo *et al*. (O-(Benzotriazol-1-yl)-N,N,N′,N′-tetramethyluronium tetrafluoroborate (TBTU)/4-dimethylaminopyridine method) [16]. Briefly, 1.4 mmol *Anb* was dissolved in 5 mL DMF, and 1.4 mmol TBTU was added, followed by 0.9 mmol 4-dimethylaminopyridine. The solution was added to the resin (2 g), and 2.8 mmol DIPEA was added after 30 s. The whole mixture was stirred for 3 h at ambient temperature. The procedure was repeated three times until all amino groups on the resin were acylated; the chloranil test was used to monitor free amino groups. Then, the C-terminal amino acid residue (Lys or Phe) was incorporated using $POCl_3$ as the coupling reagent [13]. After completing the syntheses, the fully side-deprotected peptides were cleaved from the resin using a mixture of trifluoroacetic acid/triisopropylsilane/phenol/water (88:5:2:5 v/v/v/v). The crude peptides were purified by HPLC on a Beckman Gold System (Beckman Coulter Inc., CA, USA) using an RP Supelco wide-pore C8 column (10μM particles size, 25 cm length × 10 mm I.D.) (Sigma-Aldrich, St Louis, MO, USA). The solvent system was 0.1% trifluoroacetic acid (A) and 80% acetonitrile in 0.1% trifluoroacetic acid (B). Either isocratic conditions or linear gradients were applied (flow rate 5.5 mL/min, monitored at 226 nm). The purity of the synthesized peptides was checked on an HPLC Pro Star system (Varian, Mulgrave, Australia), using an RP Supelco wide-pore C8 column (5μM particles size, 25 cm length × 4.6 mm I.D.) (Sigma-Aldrich). The solvent system was 0.1% trifluoroacetic acid (A) and 80% acetonitrile in 0.1% trifluoroacetic acid (B), with a linear gradient from 10–90% B for 40 min at a flow rate of 1 mL/min, Absorbance was monitored at 226 nm. The MS analysis was performed using a Biflex III MALDI-TOF spectrometer (Bruker Daltonics, Bremen, Germany) using α-cyano-4-hydroxycinnamic acid matrix (Fig. S3).

Intermolecular disulfide bridge formation

First, 0.1 mmol of a C-terminal fragment of SFTI-1 (for example SIPPICFPD) was dissolved in 4 mL of 0.1 M aqueous acetic acid. Next, 4 mL of 0.02 mM 2,2′-dithiodipyridine in isopropanol was added, and the reaction was stirred for 1 h at ambient temperature. The active (in the thiol function) peptide derivative was purified by reverse-phase HPLC and lyophilized. In the next step, dithiopyridy -peptide and 2 mmol of an N-terminal fragment of SFTI-1 (for example GRCTK) were dissolved in 6 mL of 0.1 M sodium acetate buffer, pH 5. The reaction was performed for 1 h. The resulting two-peptide-chain analogs were purified by reverse-phase HPLC and lyophilized.

Determination of inhibitory activity

The inhibitory activities of the synthesized compounds were determined against two serine proteinases derived from bovine pancreas: β-trypsin and α-chymotrypsin (Sigma-Aldrich). The chromogenic trypsin burst substrate nitrophenyl-4′-guanidinobenzonate and the turkey ovomucoid third domain inhibitor (OMTKY-3) were purchased from Sigma-Aldrich. OMTKY-3 is the mutual inhibitor of trypsin and chymotrypsin. The chromogenic turnover substrates N^{α}-benzoyl-D,L-Arg-*p*NA (for trypsin), Suc-Ala-Ala-Pro-Phe-*p*NA and Suc-Ala-Ala-Pro-Leu-*p*NA (both for chymotrypsin) were purchased from Bachem (Bubendorf, Switzerland). All measurements were performed on a Cary 3E spectrophotometer (Varian). The stock solutions of each enzyme (bovine β-trypsin, 8 mg/mL; bovine α-chymotrypsin, 1.0 mg/mL) were prepared using 1 mM HCl solvent with 20 mM $CaCl_2$ and 0.005% Triton X-100. The stock solutions of OMTKY-3, SFTI-1 and its analogs were prepared with 1 mM HCl. The concentration of bovine β-trypsin (active form) was determined with nitrophenyl-4′-guanidinobenzonate solution using the procedure described by Chase and Shaw [17]. Standardized trypsin solution with a concentration range of 1.5–3.0 × 10^{-5} M was used to titrate SFTI-1 and OMTKY-3 solutions with N^{α}-benzoyl-D, L-Arg-*p*NA. OMTKY-3 (at a concentration of approximately 1.3 × 10^{-4} M) was used to determine the concentration of α-chymotrypsin, using Suc-Ala-Ala-Pro-Leu-*p*NA as the substrate. The concentrations of stock solutions of all the inhibitors were determined by HPLC. The area of the inhibitor peak was integrated and compared to that of SFTI-1.

Determination of association equilibrium constants

Analogs **1**, **2** and **5**, as the most potent trypsin inhibitors, were selected to measure the association constant (K_a) values for their complexes with the enzyme. Similarly, the K_a value for the complex of analog **4** and chymotrypsin was determined. The K_a values were measured using a modified version of the Green and Work method [18], as described by Empie and Laskowski [19] and Otlewski and Zbyryt [20]. All measurements were performed in 50 mM Tris/HCl buffer solution, pH 8.3, 20 mM $CaCl_2$ and 0.005% Triton X-100. Increasing amounts of the inhibitor were added to a constant amount of the enzyme. The total enzyme concentration $[E_0]$ was chosen so as to fulfil the condition $2 < [E_0] \times K_a < 50$. After an appropriate incubation time (120 min), the residual enzyme activity $[E]$ was measured using the chromogenic substrate. The value of K_a was calculated by a two-parameter algorithm using the GraFit program (Erithacus Software, Horley, UK), according to the equation:

$$E = \frac{1}{2}\left([E_0] - [I_0] - K_a^{-1} + \sqrt{\left([E_0] + [I_0] + K_a^{-1}\right)^2 - 4[E_0][I_0]}\right)$$

where $[E_0]$ and $[I_0]$ are the total enzyme and inhibitor concentrations, respectively, and $[E]$ is the residual enzyme concentration.

Proteolytic susceptibility

The SFTI-1 analogs were incubated in 100 mM Tris/HCl buffer (pH 8.3) containing 20 mM $CaCl_2$ and 0.005% Triton X-100 using catalytic amounts of the enzymes (1.5%). The incubation was performed at ambient temperature, and aliquots of the mixture were extracted periodically and subjected to reverse-phase HPLC and MALDI MS analysis.

Acknowledgements

This work was supported by the Polish National Science Centre under grant number UMO-2011/01/B/ST5/03772.

References

1 Luckett S, Garcia RS, Barker JJ, Konarev AV, Shewry PR, Clarke AR & Brady RL (1999) High-resolution structure of a potent, cyclic proteinase inhibitor from sunflower seeds. *J Mol Biol* **290**, 525–533.
2 Schechter I & Berger A (1967) On the size of the active site in proteases. I. Papain. *Biochem Biophys Res Commun* **27**, 157–162.
3 Korsinczky ML, Schirra HJ, Rosengren KJ, West J, Condie BA, Otvos L, Anderson MA & Craik DJ (2001) Solution structures by ^{1}H NMR of the novel cyclic trypsin inhibitor SFTI-1 from sunflower seeds and an acyclic permutant. *J Mol Biol* **311**, 579–591.
4 Zabotna E, Kaźmierczak K, Jaśkiewicz A, Stawikowski M, Kupryszewski G & Rolka K (2002) Chemical synthesis and kinetic study of the smallest naturally occurring trypsin inhibitor SFTI-1 isolated from sunflower seeds and its analogues. *Biochem Biophys Res Commun* **292**, 855–859.
5 Lesner A, Lęgowska A, Wysocka M & Rolka K (2011) Sunflower trypsin inhibitor 1 as a molecular scaffold for drug discovery. *Curr Pharm Design* **38**, 4308–4317.
6 Cascales L, Henriques ST, Kerr MC, Huang YH, Sweet MJ, Daly NL & Craik DJ (2011) Identification and characterization of a new family of cell-penetrating peptides. *J Biol Chem* **286**, 36932–36943.
7 Colgrave ML, Korsinczky MJL, Clark RJ, Foley F & Craik DJ (2010) Sunflower trypsin inhibitor-1 proteolytic studies on a trypsin inhibitor peptide and its analogs. *Biopolymers* **94**, 665–672.
8 Lęgowska A. Lesner A, Bulak E, Jaśkiewicz A, Sieradzan A, Cydzik M, Stefanowicz P, Szewczuk Z & Rolka K (2010) Inhibitory activity of double-sequence analogues of trypsin inhibitor SFTI-1 from sunflower seeds: an example of peptide splicing. *FEBS J* **277**, 2351–2359.
9 Hanada K, Yewdell JW & Yang JC (2004) Immune recognition of a human renal cancer antigen through post-translational protein splicing. *Nature* **427**, 252–256.
10 Vigneron N, Stroobant V, Chapiro J, Ooms A, Degiovanni G, Morel S, Van der Bruggen P, Boom T & Van den Eynde B (2004) An antigenic peptide produced by peptide splicing in the proteasome. *Science* **304**, 587–590.
11 Dalet A, Vigneron N, Stroobant V, Hanada K & Van den Eynde B (2010) Splicing of distant peptide fragments occurs in the proteasome by transpeptidation and produces the spliced antigenic peptide derived from fibroblast growth factor-5. *J Immunol* **184**, 3016–3024.
12 Saska I & Craik DJ (2008) Protease-catalysed protein splicing: a new post-translational modification? *Trends Biochem Sci* **33**, 363–368.
13 Wysocka M, Lęgowska A. Bulak E, Jaśkiewicz A, Miecznikowska H, Lesner A & Rolka K (2007) New chromogenic substrates of human neutrophil cathepsin G containing non-natural aromatic amino acid residues in position P_1 selected by combinatorial chemistry methods. *Mol Divers* **11**, 93–99.
14 Hoffmann R & Otvos L (1996) Selective dimerization of cysteines in glycopeptides and phosphopeptides. *Lett Pept Sci* **3**, 371–377.
15 Marx UC, Korsinczky ML, Schirra HJ, Jones A, Condie B, Otvos L & Craik DJ (2003) Enzymatic cyclization of potent Bowman–Birk protease inhibitor,

sunflower trypsin inhibitor-1, and solution structure of an acyclic precursor peptide. *J Biol Chem* **278**, 21782–21789.
16 Hojo K, Maeda M, Iguchi S, Smith T, Okamoto H & Kawasaki K (2000) Amino acids and peptides. XXXV. Facile preparation of *p*-nitroanilide analogs by the solid-phase method. *Chem Pharm Bull* **48**, 1740–1745.
17 Chase T & Shaw E (1970) Titration of trypsin, plasmin, and thrombin with *p*-nitrophenyl *p*′-guanidinobenzoate HCl. *Methods Enzymol* **19**, 20–27.
18 Green NM & Work E (1953) Pancreatic trypsin inhibitor. *Biochemistry* **54**, 347–352.
19 Empie MW & Laskowski M Jr (1982) Thermodynamics and kinetics of single residue replacements in avian ovomucoid third domains: effect on inhibitor interactions with serine proteinases. *Biochemistry* **21**, 2274–2284.
20 Otlewski J & Zbyryt T (1994) Single peptide bond hydrolysis/resynthesis in squash inhibitors of serine proteinases. 1. Kinetics and thermodynamics of the interaction between squash inhibitors and bovine β-trypsin. *Biochemistry* **33**, 200–207.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web site:

Fig. S1. Inhibition curves of the active inhibitors.

Fig. S2. Proteolytic susceptibility assay of GRCTF-*Anb* with bovine α-chymotrypsin and GRCTK-*Anb* (B) with bovine β-trypsin.

Fig. S3. MS spectrum and HPLC of synthesized analogs of SFTI-1.