Acta Biochimica Polonica

Vol. 43 No. 3/1996

431–444

QUARTERLY

Minireview

Squash inhibitor family of serine proteinases

Jacek Otlewski[a] and Daniel Krowarsch

Institute of Biochemistry, University of Wrocław, Tamka 2, 50-137 Wrocław, Poland

Received: 5 June, 1996

Key words: squash seeds, protein proteinase inhibitors, serine proteinases

Squash inhibitors of serine proteinases form an uniform family of small proteins. They are built of 27–33 amino-acid residues and cross-linked with three disulfide bridges. The reactive site peptide bond (P1-P1') is between residue 5 (Lys, Arg or Leu) and 6 (always Ile). High resolution X-ray structures are available for two squash inhibitors complexed with trypsin. NMR solution structures have also been determined for free inhibitors. The major structural motif is a distorted, triple-stranded antiparallel β-sheet. A similar folding motif has been recently found in a number of proteins, including: conotoxins from fish-hunting snails, carboxypeptidase inhibitor from potato, kalata B1 polypeptide, and in some growth factors (e.g. nerve growth factor, transforming growth factor β2, platelet-derived growth factor). Squash inhibitors are highly stable and rigid proteins. They inhibit a number of serine proteinases: trypsin, plasmin, kallikrein, blood clotting factors: X_a and XII_a, cathepsin G. The inhibition spectrum can be much broadened if specific amino-acid substitutions are introduced, especially at residues which contact proteinase. Squash inhibitors inhibit proteinases *via* the standard mechanism. According to the mechanism, inhibitors are substrates which exibit at neutral pH a high k_{cat}/K_m index for hydrolysis and resynthesis of the reactive site, and a low value of the hydrolysis constant.

Protein proteinase inhibitors form a large and well-characterized group of proteins [1–3]. The action of the inhibitor is almost always restricted only to one out of four mechanistic classes of proteinases, hence there are protein inhibitors of serine, aspartic, cysteine and metallo proteinases. Of these four groups, serine proteinase inhibitors are most widely represented in nature and are the most popular subject of investigations. They are common in plant and animal tissues as well as in microorganisms. Within this group of inhibitors one can easily distinguish "small" inhibitors (27–200 amino-acid residues) and serum proteinase inhibitors (glycosylated proteins containing about 400 amino-acid residues). "Small" inhibitors do not form an evolutionarily uniform group. Instead, they can be divided into about 15 different families on the basis of amino-acid and/or nucleotide sequence similarity, topology of disulfide bonds and global spatial structure [1, 3, 4]. Also the position of the reactive site peptide bond, which is primarily responsible for inhibitory action, is fully preserved within the inhibitor family [5]. At present high-resolution crystallographic structures are available for representatives of almost all inhibitor families, either in a free state or in a complex with a cognate proteinase [3].

[a]To whom all the correspondence should be addressed; tel. 71- 402824; fax. 71- 402608; e-mail: otlewski@ii.uni.wroc.pl.

Proteinase inhibitors isolated from seeds of different squash plants are the smallest of "small" inhibitors. They are built of 27–32 amino-acid residues and they are cross-linked with three disulfide bridges. So far they were isolated exclusively from squash seeds (the *Cucurbitaceae* family). They were discovered in late 70s independently by our group at the University of Wrocław and by the group of the late Prof. J.J. Pisano at the National Institutes of Health. "Our" squash seeds inhibitor was discovered as a bovine trypsin inhibitor [6–10]. The protein isolated from pumpkin seeds in Bethesda inhibited bovine trypsin and Hageman factor [11, 12]. Amino-acid sequences of both proteins appeared to be identical [7]. Since that time our group isolated and characterized inhibitors from the seeds of summer squash [9, 10, 13], zucchini [8, 13], watermelon [14, 15], red bryony [14, 15], cucumber [13, 14], figleaf gourd [14] and spaghetti squash [14]. The squash inhibitors were also described in more exotic squash plants from South Africa [16], Japan [17–20), China [21] and France [22]. At present they form a well established inhibitor family: 41 squash inhibitors were sequenced, spatial structures of 3 inhibitors were studied by high resolution X-ray crystallography, 4 inhibitors were investigated by multidimensional NMR, 31 inhibitor analogues were chemically synthesized. Due to their small size, rigidity and stability of the molecule, squash inhibitors provide a particularly encouraging material for studying the serine proteinase-protein inhibitor interactions.

ISOLATION OF SQUASH INHIBITORS

Pure samples of squash inhibitors are relatively easy to obtain as they are very stable and harsh purification techniques (e.g. reverse phase HPLC, organic solvents) can be applied. Typically, the isolation protocol includes defatting of seeds with acetone, extraction, ammonium sulfate and/or organic solvent fractionation. Further purification steps involve different chromatographic techniques: affinity chromatography on immobilized trypsin or anhydrotrypsin, classic low-pressure ion-exchange or molecular sieve chromatography or high pressure reversed phase HPLC. Finally, one can obtain pure inhibitors in quantities varying from 1 mg to approximately 100 mg per 1 kg of seeds, depending on plant species and inhibitor form to be purified. Detailed purification protocols are presented in references [8, 9, 11, 14, 16–18, 20, 22]. So far, the seeds of figleaf gourd (*Cucurbita ficifolia)* are the richest known source of squash inhibitor. Our purification procedure [14] yields about 110 mg of pure inhibitor from 1 kg of seeds. The amino-acid sequence of this protein is identical with that of *Cucurbita maxima* trypsin inhibitor I (CMTI I), which is the best characterized representative of the whole inhibitor family (see below).

PROPOSED NOMENCLATURE

The nomenclature of squash inhibitors [13] uses the latin names of the plant species e.g. *Cucurbita maxima* trypsin inhibitor is CMTI and *Momordica charantia* elastase inhibitor is MCEI. The designation of different inhibitor forms first employs Roman numerals according to the order of trypsin inhibitory activity peaks obtained after initial chromatography. If, upon subsequent chromatography in other systems these are shown to be heterogenous, the subfractions are designated by lower case letters: a, b, c, e.g. CSTI IIb. Other groups use this abbreviation system with minor modifications. Numbering of amino-acid sequences of squash inhibitors is based on the sequence of the CMTI I inhibitor.

AMINO-ACID SEQUENCES AND LOCATION OF DISULFIDE BONDS

All so far published amino-acid sequences of squash inhibitors are presented in Fig. 1. Forty one inhibitors have been sequenced and 39 amino-acid sequences are unique. Within these 41 sequences, there are 7 which differ by an extended N-terminus: His-Glu-Glu, Glu-Glu, Glu or pyroglutamic acid. There is one case of Val-Ser extension at the C-terminus (CSTI IIb). Evidently, squash inhibitors are synthesized as longer polypeptide chains and are proteolytically processed both from the N- and C-terminus. This is in agreement with the recently determined cDNA and gene sequences of TGTI II inhibitor, which revealed that mature inhibitor sequence is flanked by sequences of 21 and

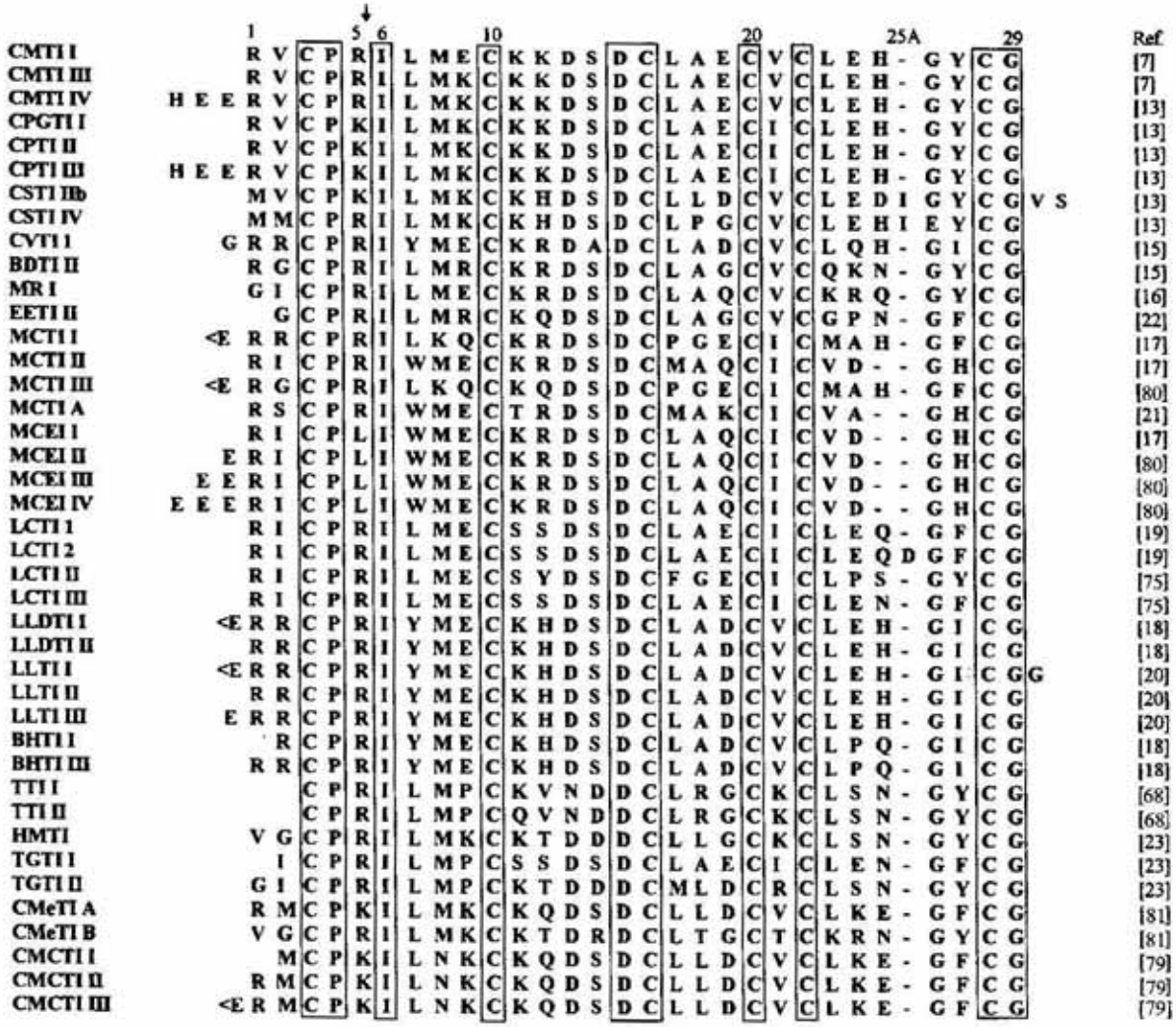

Fig. 1. Amino-acid sequences of squash inhibitors of serine proteinases. An arrow indicates the reactive site peptide bond.

13 residues of pre- and pro-inhibitor, respectively [23]. This type of posttranslational processing exists also in other plant inhibitors [24].

The shortest of squash inhibitors contain 27 amino-acid residues, the longest 33 residues. No carbohydrate moiety was observed in the inhibitors (there are no potential N-glycosylation sites in squash inhibitors). The amino acid sequences of squash inhibitors are homologous. Since the inhibitors contain six totally conserved half-cystines, alignment of their sequences is particularly easy. Squash inhibitors always contain one aromatic residue (Trp, Tyr or Phe) either at position 7 (P2') or 27 (P22'). Surprisingly, simultaneous lack or presence of both chromophores at these positions have never been observed. The reactive site (P1-P1' peptide bond in notation of Schechter & Berger [25]) is located between residues 5 and 6. Residue 5, which mainly ensures proper recognition by proteinase, is almost invariably Lys or Arg. Thus far, the only exception detected is Leu present in the sequences of MCEI I, II, III and IV. The Lys5 or Arg5 containing squash inhibitors always strongly inhibit bovine trypsin, whereas the MCEI inhibitors do not inhibit trypsin but porcine elastase. Taking into account the number of sequenced representatives, squash inhibitors form the third largest family of inhibitors.

Laskowski and coworkers [4, 26] found for a set of 125 amino-acid sequences of orthologous avian ovomucoid third domains that the proteinase contact region exhibits greater variability of amino-acid residues than that of noncontact residues. Later, a similar hypervari-

ability of contact residues has been noticed in the case of 29 proteins paralogous to bovine pancreatic trypsin inhibitor (BPTI) [27].

Variability of amino-acid residues within squash inhibitor family is shown in Fig. 2. It is clear from this presentation that at variance with the results reported for vertebrates the positions of squash inhibitors which contact trypsin do not exhibit hypervariability. Besides six half-cystines, only four other positions are invariant. Three of them (Pro4, Ile6 and Gly29) are in contact with proteinase. Of the seven most variable positions, only two (P4 and P20') are in contact with the enzyme. It might be argued that the possible hypervariability of squash inhibitors was not evidenced, since many investigators have screened plant material only for antitrypsin activity. This is, however, not the case, since our group (unpublished data) and others [6] looked without success also for inhibition of other proteinases. Moreover, since the energy of association with trypsin is dominated by P1 residue (Arg5 is involved in almost half of the contacts with trypsin), we suppose that other contact residues could be relatively unconstrained to evolve.

1. *Arg Met Gly Val
2. *Val Gly Met Arg Ile Ser
3. *Cys
4. *Pro
5. *Arg Lys Leu
6. *Ile
7. *Leu Tyr Trp
8. Met Lys Asn
9. *Glu Lys Arg Gln Pro
10. Cys
11. Lys Gln Thr Ser
12. Lys His Arg Gly Ser Val Tyr Thr
13. Asp Asn
14. Ser Ala Asp
15. Asp
16. Cys
17. Leu Pro Met Phe
18. Ala Leu Pro Gly Arg Thr
19. Glu Asp Gly Gln Lys
20. Cys
21. Val Ile Lys Arg Thr
22. Cys
23. Leu Gln Lys Gly Met Val
24. Glu Gln Lys Arg Pro Ala Asp Ser
25. *His Asp Asn Gln Ser
26. *Gly Glu
27. *Tyr Ile Phe His
28. *Cys
29. *Gly

Fig. 2. Variability of amino-acid residue positions within the squash inhibitor family.
The sequence of *Cucurbita maxima* trypsin inhibitor I (CMTI I) is shown vertically. The reactive site is between Arg5 and Ile6. The residues that are in contact with bovine trypsin, according to the X-ray structure of bovine β-trypsin-CMTI I [16] are indicated by an asterisk.

It is striking that the proteins isolated from a single plant family differ so much in their sequences. Despite sequence differences, overall folding of squash inhibitors is very well preserved. It seems that almost only structural constraints determining conformation of squash inhibitors are disulfide bridges. It should be noted that hypervariable sequences have been recently reported for toxins isolated from venoms of cone snails [28], and moreover molecular size, disulfide pairing and global folding of these toxins are similar to those of squash inhibitors.

The most important structural feature of squash inhibitors is the tight cross-linking of a relatively short polypeptide chain by three disulfide bridges. The topology of disulfide bonds has been established experimentally in some laboratories by completely non-equivalent methods: high resolution X-ray crystallography [29], 2D NMR [30, 31], sequencing of cyanogen bromide-thermolysin generated peptides [17] and thermospray HPLC/mass spectroscopy analysis of enzymatic digest [32].

STRUCTURE OF SQUASH INHIBITORS IN AQUEOUS SOLUTION

Squash inhibitors are ideal candidates for NMR studies for two reasons. First, despite their small size, squash inhibitors possess a stable and rigid spatial fold. In addition, due to compactness and rigidity of the molecule, a very large number of structural constraints (NOEs) can be recorded. Thus far, three representatives of squash inhibitors have been investigated by multidimensional ^{1}H NMR: *Cucurbita maxima* trypsin inhibitor I — CMTI I [30, 33–35], *Cucurbita maxima* trypsin inhibitor III (CMTI III) [36, 37] and *Ecballium elaterium* trypsin inhibitor II — EETI II [31]. Moreover, several other forms of squash inhibitors either enzymatically cleaved or chemically modified or

synthesised have also been analyzed by NMR. High resolution structures have been determined for CMTI I [30, 33–35] and EETI II [31, 38] by a combination of distance geometry and molecular dynamics methods.

The solution structure of CMTI I based on various two-dimensional spectra has been very well defined (324 interproton distance constraints for 29 residues) (Fig. 3). The major structural feature of CMTI I are three buried disulfide bridges, which make the most important contribution to the stability and rigidity of the molecule. CMTI I inhibitor has few regular secondary structure elements. There are two reverse turns in the molecule-type II′ for residues 17 to 20 and type I for residues 23 to 26. Ala18 in the second position of Leu17-Cys20 turn is in a strained conformation since its ϕ angle = 60°. A short fragment (one turn) of 3_{10}-helix is formed from Asp13 to Cys16. The C-terminal part of the inhibitor (Cys20-Gly29) can be considered as a small antiparallel β-sheet. The hydrophobic loop that bears the reactive site peptide bond (segment Val2-Met8) is the least well-defined part of the molecule. Similarly, the binding loops of uncomplexed inhibitors representing other inhibitor families were shown to exhibit the largest B factor values [39, 40]. The structure of CMTI I has been further improved by the relaxation matrix refinement calculations to account for spin diffusion effects [34] and by addition of over 200 new distance constraints extracted from 3D NOE-NOE spectrum [33]. Compared to the initial structure based on 2D spectra, both refined structures of the inhibitor are closer to the X-ray structure of the inhibitor complexed with trypsin. The refined structures show smaller bonds and angles deviations from ideality. They exhibit a uniquely defined conformation of all three disulfide bonds. In the initial structure Cys3-Cys20 and Cys16-Cys28 bonds showed mirror images at the γ-sulphur atoms. The quality of the refined structure corresponds to that of high resolution structure as defined by X-ray crystallography.

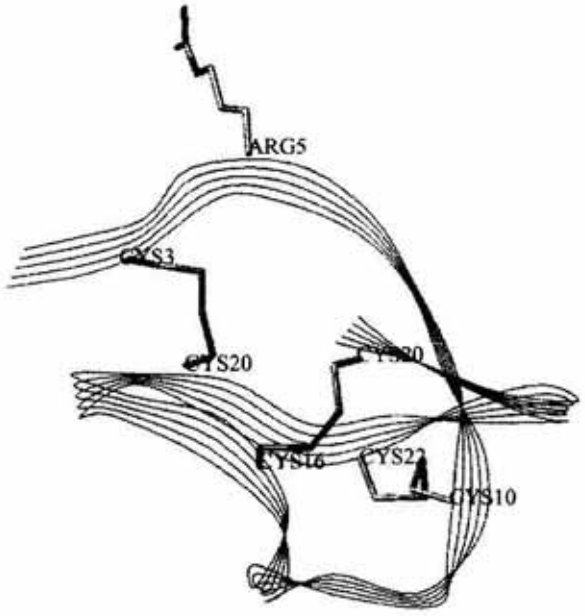

Fig. 3. NMR solution structure of CMTI I [30]. Only three disulfide bonds and the side chain of Arg5 (P1) is shown. For details see text.

The structure of EETI II refined by restrained molecular dynamics in water is essentially identical to that of CMTI I [38], despite 8 differences in its amino-acid sequence. Some deviations which occur in the binding loop conformation are rather difficult to interpret due to relatively imprecise NOE distance constraints.

Another reported NMR structure concerns a hybrid protein which inhibits trypsin and carboxypeptidase A. Squash inhibitors are spatially similar to the carboxypeptidase A inhibitor from potato (PCI) [29]. The major structural difference concerns functional sites: the C-terminal segment which is responsible for carboxypeptidase inhibition, is absent in squash inhibitors, and the proteinase binding loop in PCI is in a completely different conformation. French group synthesised a variant of EETI II intentionally extended at the C-terminus by a four amino-acid residues peptide to inhibit carboxypeptidase A a metallopeptidase. The synthesised protein inhibited both enzymes. NMR structure of the hybrid was closely similar to that of EETI II [41].

Another chemically synthesised variants of EETI II were studied by Nielsen *et al.* [42]. They found a significant decrease of affinity to trypsin upon Ile6⇨Leu substitution at the P1′ position. The possible explanation for this observation is that the reduced binding is due to increased motional disorder of the binding loop upon substitution. Ile at position 6 is fully conserved in squash inhibitors. Another form of EETI II synthesised from all D-amino acids (mirror image of EETI II), in agreement with theoretical expectations, had spectrum identical to that of native form.

A recent NMR study of CMTI III inhibitor (CMTI III differs from CMTI I by a single amino acid substitution Glu9 ⇨ Lys) was performed at neutral pH range [36]. The inhibitor conformation in the pH range 6–11 points clearly to the His25-Tyr27 interaction [36]. According to CMTI I structure, which was determined at pH 4.3, at low pH both aromatic rings are oriented 10 Å away from each other [36]. As the pH rises, the two side chains form a 3 Å hydrogen bond (with a pK around 6.0). This pK is perturbed by Glu9 ⇨ Lys substitution and by hydrolysis of the reactive site peptide bond [43]. As the binding loop and His25 are separated by more than 10 Å, these results suggest that conformational changes can be propagated throughout the whole molecule. In agreement with this, changes of chemical shifts (Δppm up to 0.2) upon hydrolysis of the Arg5-Ile6 reactive site were observed in large part of the protein [43]. The secondary structure of the inhibitor has not been however, affected.

HIGH RESOLUTION CRYSTAL STRUCTURE OF THE TRYPSIN-SQUASH INHIBITOR COMPLEX

Two complexes formed between squash inhibitor and trypsin have been analyzed by high resolution X-ray crystallography: CMTI I-bovine β-trypsin [29] and MCTI-A-porcine β-trypsin [19]. The general scheme of trypsin--inhibitor recognition resembles that observed in virtually all serine proteinase-protein inhibitor complexes [2, 3].

The general view of the CMTI I-trypsin complex is shown in Fig. 4. Thirteen residues of CMTI I inhibitor are in van der Waals contacts (below 4.0 Å) with trypsin: the binding loop (Arg1-Leu7 and Glu9) and the His25-Gly29 segment. The C-terminus of inhibitor makes, however, only 7 contacts of the total number of 138. On the other hand, ultimately important for the proper recognition is Arg5 residue, which participates in almost half of the contacts with trypsin.

The binding loop of the CMTI I forms a short fragment of antiparallel β-sheet with the proteinase. The main chain-main chain hydrogen bonds are formed between Cys3 (P3) and Gly216; Arg5 (P1) and Ser214, Gly193 and Ser195 and between Leu7 (P2') and Phe41. The reactive site peptide bond Arg5-Ile6 is almost planar, its carbonyl carbon is in short van der Waals distance with O^{γ} nucleophile of Ser195, again similarly to the situation observed in other complexes [2, 3]. The architecture of protein-protein interface in MCTI A-trypsin complex is very similar [19].

The structure of inhibitor in complex with trypsin is almost identical with that in solution determined by NMR in terms of global folding and secondary structure. The root-mean-square difference for backbone atoms between the refined NMR structure and the X-ray structure is only 0.61 Å [35]. A characteristic feature of CMTI I is the ion pair between the guanidyl group of Arg1 and the α-carboxyl group of Gly29. This electrostatic interaction was, however, absent in the crystal structure of MCTI A inhibitor. The most unusual structure feature of CMTI I inhibitor elucidated by crystallography is the presence of three well-fixed water molecules, which form an integral part of the inhibitor structure. These solvent molecules link binding loop region to the central part of the inhibitor structure *via* a network of hydrogen bonds. A similar stabilization of the binding loop occurs in other inhibitor families. In those cases, however, it is accomplished by a system of hydrogen bonds and/or electrostatic inter-

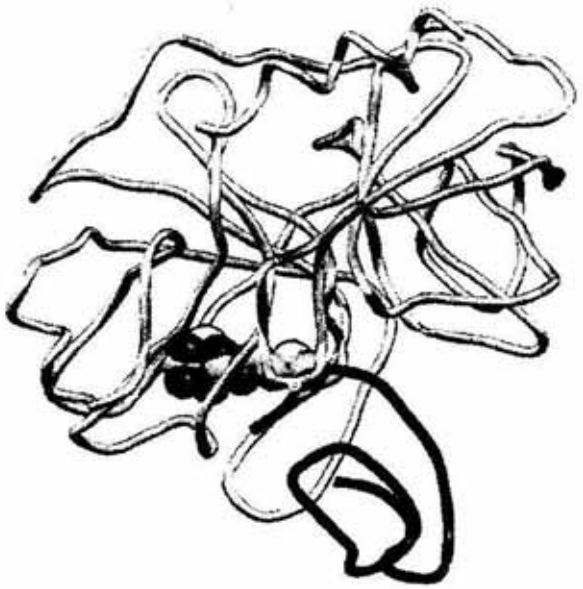

Fig. 4. General view of the CMTI I-bovine β-trypsin complex, according to 2.0 Å crystal structure of the complex [16].
The Asp189-Arg5 interaction is shown as space-filling atoms. The inhibitor's main chain is shown in darker color.

actions formed by evolutionarily conserved amino-acid side chains [40, 44]. The importance of structural water molecules is further stressed by the fact that they occupy similar sites in MCTI A inhibitor, despite large differences in sequences of the two proteins (10 substitutions plus 1 deletion). There is also one water molecule in solution structure of EETI II refined by restrained molecular dynamics, it is however, in a somewhat different position [38].

The dynamic behaviour of the proteinase binding loop (Val2-Met8) is completely different in free and complexed state. In NMR solution structure this region is least constrained and shows the greatest variability [35]. On the other side, the main chain atoms of the binding loop upon complexation with trypsin exhibit the smallest temperature B factors observed in the whole molecule. It should be stressed, however, that despite these dynamic differences, the two structures exhibit an extremely similar conformation of the loop [35]. Similar comparisons between the NMR (free inhibitor) and X-ray (liganded inhibitor) structures were presented for bovine pancreatic trypsin inhibitor [45] and barley inhibitor 2 [46]. In both cases, the local variability of the binding loop was demonstrated.

SPECIFICITY OF SQUASH INHIBITORS

Almost all natural variants of squash inhibitors are potent inhibitors of bovine trypsin. The reported values of association constants should be considered with caution for at least three reasons. Since the constants are very high and approach 10^{12} M^{-1}, they are inherently very difficult to measure. For CMTI I and CMTI III the reported values differ by more than three orders of magnitude [12, 13]. Secondly, not all investigators specified the conditions of measurement. Lastly, a pure form of trypsin (i.e. β or α) should be used. Our measurements of inhibitor association with bovine β-trypsin at pH 8.3 were summarized in references [13] and [47].

Interestingly, removal of a His-Glu-Glu tripeptide from the N-terminus leads to an about 10-fold increase of the association constant [13]. Also one additional pyroglutamic acid residue at the N-terminus of LLDTI I — a *Lagenaria leucantha* inhibitor I decreased the association constant by a factor of 2.5 as compared to LLDTI II [18].

Strong inhibition of porcine trypsin by squash inhibitors has also been reported [21]. Recently, we have showed that some other serine proteinases of trypsin-like specificity, including human trypsin, human plasma kallikrein and human thrombin (very weakly) are also inhibited by squash inhibitors [48]. The latter two enzymes, consistently with their substrate specificity, are much more strongly inhibited by the P1 Arg inhibitor (CMTI I) than by P1 Lys inhibitor (CPTI II). In agreement with early data [11, 12], squash inhibitors bind rather strongly to human factor XII_a [49]. The interaction is 62-fold stronger for CMTI III ($K_a = 3.3 \times 10^8$ M^{-1}, Lys at P4') than for CMTI I (Glu at P4'). The N-terminal tripeptide His-Glu-Glu (CPTI III *versus* CPTI II) again weakens 46-fold energy of association with this proteinase. These results stress the importance of secondary interactions in the case of highly specific proteinases.

Rather surprisingly, a quite strong association (K_a slightly above 10^8 M^{-1}) was demonstrated for human cathepsin G [48]. The proteinase exhibits a preference for bulky hydrophobic side chains at P1 position. Other serine proteinases, including subtilisin BPN', porcine pancreatic and human leukocyte elastases and bovine chymotrypsin, form very weak complexes with the inhibitors.

MECHANISM OF INHIBITION

It is generally accepted that "small" inhibitors of serine proteinases share a common, so called standard, mechanism of inhibition [1]. The mechanism was formulated after careful kinetic and thermodynamic analysis of bovine β-trypsin-soybean trypsin inhibitor (Kunitz) system and later verified for a few more inhibitor families [5, 50–52]. The mechanism can be expressed by the following scheme:

$$E + I \underset{k_{off}}{\overset{k_{on}}{\rightleftarrows}} C \underset{k^*_{on}}{\overset{k^*_{off}}{\rightleftarrows}} E + I^* \qquad (1)$$

where: E, proteinase; I, intact inhibitor; I^*, reactive site cleaved inhibitor; C, stable Michaelis-type complex; k_{on} and k^*_{on}, second order rate constants for the complex formation between proteinase and intact or cleaved inhibitor, re-

spectively; k_{off} and k_{on} and k^*_{off}, first order dissociation rate constants of the complex.

Our group investigated bovine β-trypsin-CMTI I interaction for two reasons. First, we felt that the standard mechanism should be verified on other inhibitor families. Second, availability of large amounts of CMTI I and CPTI II inhibitors which differ by Arg5 ⇨ Lys change allowed us to study the effect of this substitution on their substrate and inhibitory properties. General conclusions for squash inhibitors are similar to those for soybean trypsin inhibitor (Kunitz)-trypsin interaction. Bovine β-trypsin hydrolyses and resynthesises exclusively Arg5-Ile peptide bond, which serves as the reactive site. The equilibrium value of the reactive site conversion, defined as hydrolysis constant, $K_{hyd} = [I^*]/[I]$, was approached both from intact inhibitor side and from cleaved inhibitor side in hydrolysis and resynthesis reactions, respectively. The equilibrium values were within 2% the same. The value of K_{hyd} at pH 6.0 is independent of pH and equals 1.84 for CMTI I (P1 Arg) and 1.61 for CPTI II (P1 Lys). The values are extremely similar, reflecting that K_{hyd} depends mostly on tertiary interactions maintaining proper conformation of the binding loop, and little on chemical nature of P1 and P1′ positions [53]. The time dependence of hydrolysis and resynthesis reactions served to calculate steady-state parameters $k_{cat,f}$ and $K_{m,f}$ and $k_{cat,r}$ and $K_{m,r}$ [47]. We also measured the values of k_{on} and k^*_{on} for lysyl and arginyl inhibitors. The conclusions can be analyzed using the following equations:

$$k_{cat}/K_m)_f/(k_{cat}/K_m)_r = K_{hyd} = (k_{on}/k^*_{on})/(k_{off}/k^*_{off}) \quad (2)$$

The left side of the equation is the Haldane equation, which relates to the substrate properties of the inhibitor. The right side describes kinetics of the association/dissociation ratio of the enzyme-inhibitor complex. Both sides are related to the hydrolysis constant which is a purely inhibitor dependent feature. At pH 8.3 we know all individual rate constants which are present in the equation, since at neutral pH: $k_{cat,f} = k_{off}$ and $k_{cat,r} = k^*_{off}$ [53]. These rate constants were used to calculate K_{hyd} value, according to equation 2. The agreement between the measured and calculated hydrolysis constant was within ±20%. Taking into account huge differences in powers of individual constants and conceptually different methods of their measurement, this agreement provides particularly strong evidence for validity of the standard mechanism.

The height of the energy barrier for hydrolysis of the reactive site peptide bond can be estimated from k_{cat}/K_m specificity indexes or from k_{on}/k_{on} and k^*_{on}or k_{off}/k_{on} and k^*_{off} ratios. For squash inhibitors all these values indicate for a very low hydrolysis energy barrier. The specificity indexes are the same for arginyl and lysyl inhibitors. They are equal to about 2.0×10^6 $M^{-1}s^{-1}$ for the hydrolysis reaction and 2.4×10^5 $M^{-1}s^{-1}$ for the resynthesis. In our opinion, this shows that essentially the only requirement for the excellent stabilization of the transition state in the trypsin-catalyzed reaction is the presence of arginyl or lysyl side chain in position P1. In the case of squash inhibitors this stabilization is identical for both side chains. It should be stressed that stabilization of lysine side chain in the Michaelis type complex C (reflected in lower $K_{m,f}$ value) is 2.7-fold stronger. Lower energy of the Michaelis complex causes however, a slower rate of hydrolysis detected for lysyl inhibitor.

INTERACTION OF SQUASH INHIBITOR WITH TRYPSINOGEN

The interaction between squash inhibitors and bovine trypsinogen is about 2×10^6 times weaker than with bovine β-trypsin [54]. This factor, equivalent to 9.9 kcal/mol at room temperature, reflects the Gibbs energy utilized for rigidification of the activation domain in trypsinogen [55].

A similar effect was reported for the PSTI- and BPTI-trypsinogen interactions [56]. The effect of Lys ⇨ Arg substitution (2.1-fold) was within experimental error identical with that reported for β-trypsin (see the preceding chapter). This suggests that the interactions of both side chains with trypsinogen are highly cooperative, independent of the stereochemical nature of the P1 position and preserved upon the transition to trypsin. The affinity of squash inhibitors to the zymogen increases almost 3-fold upon saturation of the calcium binding site. A much stronger effect (more than 100-fold) on the inhibitor binding to trypsinogen was observed in the presence of 20 mM L-Ile-L-Val dipeptide. It was also possible to resynthesize

the reactive site peptide bond in CMTI I inhibitor using trypsinogen saturated with this dipeptide. Although the reaction proceeded very slowly (10^4 slower than with β-trypsin), it has been proved that the squash inhibitor-trypsinogen interaction obeys the standard mechanism.

HYDROLYSIS OF THE LEU7-MET PEPTIDE BOND WITH PEPSIN

Although native CMTI I is extremely resistant to proteolytic cleavages, its binding loop can be selectively hydrolyzed by a few proteinases. Besides trypsin, thermolysin at 49°C hydrolyzed selectively the reactive site peptide bond (Otlewski, unpublished). The Ile6-Leu bond is the predominant site of proteinase K action (Dryjanski & Otlewski, unpublished). *Penicillium camemberti* aspartic proteinase and porcine pepsin hydrolyze exclusively the Leu7-Met peptide bond [57]. These results show that the binding loop of squash inhibitors resembles somewhat the bait region of α_2-macroglobulin [58] or the reactive site loop of α_1-proteinase inhibitor [59], although cleavages in CMTI I proceed much slower.

Our group probed the binding loop of CMTI I by limited proteolysis with porcine pepsin [57]. The reaction was performed in the pH range 2.0–4.8. Some features of the Leu7-Met peptide bond resembled those of the reactive site Arg5-Ile6. The reaction could be carried out both from the intact inhibitor (hydrolysis) and from the cleaved inhibitor sides. From both sides the same equilibrium was attained. The value of pH-independent hydrolysis constant (K_{hyd}) equals 0.77. This value is similar to that for the reactive site peptide bond (1.84) in the same inhibitor. This reflects the fact that both peptide bonds are situated within the same covalent loop, linked through Cys3-Cys20 and Cys10-Cys22 disulfide bridges. Similar energy of both cleaved forms is prerequisite for the standard mechanism. The second condition, however, is stability of the complex. For the pepsin-CMTI I interaction K_m values for the hydrolysis and resynthesis reactions depend little on pH and are of order of 10^{-4} M. Thus, CMTI I is not a pepsin inhibitor.

CHEMICAL SYNTHESIS OF SQUASH INHIBITORS AND THEIR VARIANTS

In view of their small size, squash inhibitors are ideal candidates for total chemical synthesis. Thus far, more than 30 inhibitor variants were synthesized in several laboratories. Our project was performed in cooperation with two peptide chemistry groups of Prof. G. Kupryszewski, Gdańsk (Poland) and of Prof. U. Ragnarsson, Uppsala (Sweden).

In 1986 CMTI III and CMTI I were synthesized using the solid-phase procedure [60]. The synthetic product was shown to be identical with the inhibitor of natural origin. In the following years, several analogs of CMTI III were obtained by our group [61–64] and by the group from Monsanto Company [65]. Analogs were also obtained for EETI II inhibitor [66, 67] and for TTI I inhibitor [68].

Scientific interest was focused mainly on producing variants with altered P1 positions. Val5-CMTI III turned out to be potent inhibitor of human leukocyte elastase (HLE) [61]. Compared to natural CMTI III (Arg at P1) Val5-CMTI III inhibits this proteinase at least 10^7-fold stronger. A systematic study revealed that also variants with P1 Val, Ile or Gly inhibit HLE, however, no quantitative data have been reported [65]. The screening indicated that Leu, Ala, Phe and Met variants inhibited cathepsin G and chymotrypsin [65]. Our study showed that Phe5-CMTI III is a weak inhibitor of chymotrypsin ($K_a = 10^6$ M^{-1}) [62]. In order to improve binding to chymotrypsin, the analog of CMTI III was prepared in which the binding loop of a strong chymotrypsin inhibitor — turkey ovomucoid third domain — was introduced [62]. Indeed, compared to Phe5-CMTI III, the variant Gly2,Thr4,Phe5,Glu6,Tyr7, Arg8-CMTI III inhibited trypsin 6×10^5-fold stronger. Quantitative data were also reported for EETI II variants. The analogs with Ala, Nle (norleucine), and Val at P1 and Nle at P3′ moderately (K_a about 10^6–10^7 M^{-1}) inhibited porcine pancreatic elastase, whereas variants with Phe at P1 and Nle at P3′ also inhibited chymotrypsin although rather moderately ($K_a = 5 \times 10^5$ M^{-1}) [66].

Besides the binding loop variants, our group prepared a few analogs with altered scaffold of the inhibitor. The variants of CMTI III shortened at the N-terminus by Arg1 and/or Val2 and/or at C-terminus by Gly29 exhibited as strong inhibition of bovine trypsin as did native CMTI III [62]. In order to further simplify the inhibitor molecule, each of three disulfide bridges was separately eliminated [64]. For each of the three variants a huge decrease, by 5–6 orders of magnitude in association constant was observed. The results underline the essential role of disulfide bridges in maintaining the native conformation of squash inhibitors.

Recently, a French group investigated a reductive and oxidative folding pathway of EETI II inhibitor [69]. A stable folding intermediate lacking the Cys3-Cys20 disulfide bond has been characterized using 2D NMR. The intermediate exhibits clearly a disorder in the binding loop segment, which is the probable explanation for the weak inhibiting properties observed in our CMTI III variant lacking the Cys3-Cys20 disulfide bond.

Further NMR studies on a chemically synthesised linear variant of EETI II (three disulfide bonds replaced with six Ser residues) revealed a population of secondary structure elements in regions 10–15, 16–19 and 22–25, which might serve as nucleation sites during early stages of the oxidative folding [70].

PERSPECTIVES

When three-dimensional structure of the CMTI I in complex became available in 1989 there was only one known protein with a similar fold namely a carboxypeptidase A inhibitor from potato (PCI) [29]. Nowadays we know a much greater number of proteins, called knottins, exhibiting a similar fold [71]. The characteristic structural feature of this superfamily is the presence of three disulfide bonds forming a knot in such a way that the first two disulfides (1–4 and 2–5) and the intervening polypeptide backbone form a ring through which the third disulfide (3–6) passes. The proteins are sequentially unrelated and have diverse biological activities: neurotoxic activity is a feature of cone snails conotoxins, kalata B1 polypeptide exhibits a strong uterine contracting activity, anti-endoproteinase or anti-exoproteinase activity characterizes squash inhibitors and PCI, respectively. Moreover, a rather similar structural motif is present in one family of growth factors [72]. Possible engineering of knottin proteins might involve grafting novel activity epitopes. The first example is a successful combination of antitrypsin and anticarboxypeptidase activity in EETI II by the addition of extra C-terminal residues [67]. The stable and rigid fold of knottins seems to be extremely tolerant to sequence substitutions, as found in conotoxins [28] and in squash inhibitors (Fig. 1).

Another possibility includes designing of new potent and specific antiproteolytic activities for squash inhibitors. The energy of the inhibitor-serine proteinase interaction is highly additive, i.e. energy of individual interactions in the contact area sum up [73, 74]. Since wild types of squash inhibitors block many different medically important serine proteinases [11, 12, 48, 75] increasing of the strength of inhibition should be possible with the help of three dimensional structures of respective proteinases and knowledge of substrate preferences of these enzymes. Very recently three groups reported successful expression of squash inhibitors in *E. coli* [76–78], which is a prerequisite for production of inhibitor variants.

REFERENCES

1. Laskowski, M., Jr. & Kato, I. (1980) Protein inhibitors of proteinases. *Annu. Rev. Biochem.* **49**, 593–626.
2. Read, R. & James, M.N.G. (1986) Introduction to the proteinase inhibitors: X-ray crystallography; in *Proteinase Inhibitors* (Barrett, A.J. & Salvesen, G., eds.) pp. 301–336, Elsevier, Amsterdam.
3. Bode, W. & Huber, R. (1992) Natural protein proteinases inhibitors and their interaction with proteinases. *Eur. J. Biochem.* **204**, 433–451.
4. Laskowski, M., Jr. (1986) Protein inhibitors of serine proteinases — mechanism and classification; in *Nutritional and Toxicological Significance of Enzyme Inhibitors* (Friedman, M., ed.) pp. 1–17, Plenum, New York.
5. Ardelt, W. & Laskowski, M., Jr. (1985) Turkey ovomucoid third domain inhibits eight different serine proteinases of varied specificity on the same ... Leu^{18}-Glu^{19} ... reactive site. *Biochemistry* **24**, 5313–5320.
6. Polanowski, A., Wilusz, T., Nienartowicz, B., Cieslar, E., Slonimska, A. & Nowak, K. (1980)

Isolation and partial amino acid sequence of the trypsin inhibitor from the seeds of *Cucurbita maxima*. *Acta Biochim. Polon.* **27**, 371–382.

7. Wilusz, T., Wieczorek, M., Polanowski, A., Denton, A., Cook, J. & Laskowski, M., Jr. (1983) Amino-acid sequence of two trypsin isoinhibitors, ITD I and ITD III from squash seeds (*Cucurbita maxima*). *Hoppe-Seyler's Z. Physiol. Chem.* **364**, 93–95.
8. Leluk, J., Otlewski, J., Wieczorek, M., Polanowski, A. & Wilusz, T. (1983) Preparation and characteristics of trypsin inhibitors from the seeds of squash (*Cucurbita maxima*) and zucchini (*Cucurbita pepo var. Giromontia*). *Acta Biochim. Polon.* **30**, 127–138.
9. Otlewski, J., Polanowski, A., Leluk, J. & Wilusz, T. (1984) Trypsin inhibitors in summer squash (*Cucurbita pepo*) seeds. Isolation, purification and partial characterization of three inhibitors. *Acta Biochim. Polon.* **31**, 267–278.
10. Otlewski, J. & Wilusz, T. (1985) The serine proteinases inhibitors from summer squash (*Cucurbita pepo*): Some structural features, stability and proteolytic degradation. *Acta Biochim. Polon.* **32**, 285–293.
11. Hojima, Y., Pierce, J.V. & Pisano, J.J. (1980) Plant inhibitors of serine proteinases: Hageman factor fragment, kallikreins, plasmin, thrombin, factor X_a trypsin, and chymotrypsin. *Thromb. Res.* **20**, 163–171.
12. Hojima, Y., Pierce, J.V. & Pisano, J.J. (1982) Pumpkin seed inhibitor of human factor XII_a (activated Hageman factor) and bovine trypsin. *Biochemistry* **21**, 3741–3746.
13. Wieczorek, M., Otlewski, J., Cook, J., Parks, K., Leluk, J., Wilimowska-Pelc, A., Polanowski, A., Wilusz, T. & Laskowski, M., Jr. (1985) The squash inhibitor family of serine proteinase inhibitors. Amino acid sequences and association equilibrium constants of inhibitors from squash, summer squash, zucchini, and cucumber seeds. *Biochem. Biophys. Res. Commun.* **126**, 646–652.
14. Polanowski, A., Cieslar, E., Otlewski, J., Nienartowicz, B., Wilimowska-Pelc, A. & Wilusz, T. (1987) Protein inhibitors of trypsin from the seeds of *Cucurbitaceae* plants. *Acta Biochim. Polon.* **34**, 395–406.
15. Otlewski, J., Whatley, H., Polanowski, A. & Wilusz, T. (1987) Amino-acid sequences of trypsin inhibitors from watermelon (*Citrullus vulgaris*) and red bryony (*Bryonia dioica*). *Biol. Chem. Hoppe-Seyler* **368**, 1505–1507.
16. Joubert, F.J. (1984) Trypsin inhibitors from *Momordica repens* seeds. *Phytochemistry* **23**, 1401–1406.
17. Hara, S., Makino, J. & Ikenaka, T. (1989) Amino acid sequences and disulfide bridges of serine proteinase inhibitors from bitter gourd (*Momordica charantia* Linn.) seeds. *J. Biochem. (Tokyo)* **105**, 88–92.
18. Matsuo, M., Hamato, N., Takano, R., Kamei-Hayashi, K., Yasuda-Kamatani, Y., Nomoto, K. & Hara, S. (1992) Trypsin inhibitors from bottle gourd (*Lagenaria leucantha* Rusby var. *Depressa* Makino) seeds. Purification and amino acid sequences. *Biochim. Biophys. Acta* **1120**, 187–192.
19. Hakateyama, T., Hiraoka, M. & Funatsu, G. (1991) Amino acid sequence of the two smallest trypsin inhibitors from sponge gourd seeds. *Agric. Biol. Chem.* **10**, 2641–2642.
20. Hamato, N., Takano, R., Kamei-Hayashi, K. & Hara, S. (1992) Purification and characterization of serine proteinase inhibitors from gourd (*Lagenaria leucantha* Rusby var. *Gourda* Makino) seeds. *Bioscience Biotechnol. Biochem.* **56**, 275–279.
21. Huang, Q., Liu, S. & Tang, Y. (1992) Refined 1.6 resolution crystal structure of the complex formed between porcine β-trypsin and MCTI-A, a trypsin inhibitor of the squash family. *J. Mol. Biol.* **229**, 1022–1036.
22. Favel, A., Mattras, H., Coletti-Previero, M.A., Zwilling, R., Robinson, E.A. & Castro, B. (1989) Protease inhibitors from *Ecballium elaterium* seeds. *Int. J. Peptide Protein Res.* **33**, 202–208.
23. Ling, M.-H., Qi, H. & Chi, C. (1993) Protein, cDNA and genomic DNA sequences of the towel gourd trypsin inhibitor. *J. Biol. Chem.* **268**, 810–814.
24. Ikenaka, T. & Norioka, S. (1986) Bowman-Birk family of serine proteinases; in *Proteinase Inhibitors* (Barrett, A.J. & Salvesen, G., eds) pp. 360–374, Elsevier, Amsterdam.
25. Schechter, I. & Berger, A. (1967) On the size of the active site in proteases. *Biochem. Biophys. Res. Commun.* **27**, 157–162.
26. Laskowski, M., Jr., Kato, I., Ardelt, W., Cook, J., Denton, A., Empie, M.W., Kohr, W.J., Park, S.J., Parks, K., Schatzley, B.L., Schoenberger, O.L., Tashiro, M., Vichot, G., Whatley, H.E., Wieczorek, A. & Wieczorek, M. (1987) Ovomucoid third domains from 100 avian species. Isolation, sequence and hypervariability of enzyme-inhibitor contact residues. *Biochemistry* **26**, 202–221.
27. Creighton, T.E. & Charles, I.G. (1987) Biosynthesis, processing, and evolution of bovine pancreatic trypsin inhibitor. *Cold Spring Harbor Symp. Quant. Biol.* **52**, 511–519.

28. Olivera, B.M., Hillyard, D.R., Marsh, M. & Yoshikami, D. (1995) Combinatorial peptide libraries in drug design: Lessons from venomous cone snails. *Trends Biotechn.* **13**, 422–430.

29. Bode, W., Greyling, H.J., Huber, R., Otlewski, J. & Wilusz, T. (1989) The refined 2.0 Å X-ray crystal structure of the complex formed between bovine β-trypsin and CMTI-I, a trypsin inhibitor from squash seeds (*Cucurbita maxima*). *FEBS Lett.* **242**, 285–292.

30. Holak, T.A., Gondol, D., Otlewski, J. & Wilusz, T. (1989) Determination of the complete three--dimensional structure of the trypsin inhibitor from squash seeds in aqueous solution by nuclear magnetic resonance and a combination of distance geometry and dynamic simulated annealing. *J. Mol. Biol.* **210**, 635–648.

31. Heitz, A., Chiche, L., Le-Nguyen, D. & Castro, B. (1989) 1H 2D NMR and distance geometry study of the folding of *Ecballium elaterium* trypsin inhibitor, a member of the squash inhibitors family. *Biochemistry* **28**, 2392–2398.

32. Stachowiak, K., Otlewski, J., Polanowski, A. & Dyckes, D.F. (1990) Monitoring protein cleavage and concurrent disulfide bond assignment using thermospray LC/MS. *Peptide Res.* **3**, 148–154.

33. Holak, T.A., Habazettl, J., Oschkinat, H. & Otlewski, J. (1991) Structure of proteins in solution derived from homonuclear three--dimensional NOE-NOE nuclear magnetic resonance spectroscopy. High resolution structure of squash trypsin inhibitor. *J. Am. Chem. Soc.* **113**, 3196–3198.

34. Nilges, M., Habazettl, J., Brünger, A.T., Holak, T. A. (1991) Relaxation matrix refinement of the solution structure of squash trypsin inhibitor. *J. Mol. Biol.* **219**, 499–510.

35. Holak, T.A., Bode, W., Huber, R., Otlewski, J. & Wilusz, T. (1989) Nuclear magnetic resonance and X-ray structures of squash trypsin inhibitor exhibit the same conformation of the proteinase binding loop. *J. Mol. Biol.* **210**, 649–654.

36. Krishnamoorthi, R., Lin, C.-L.S., Gong, Y.X., VanderVelde, D. & Hahn, K. (1992) Proton NMR studies of *Cucurbita maxima* trypsin inhibitors: evidence for pH-dependent conformational change and His25-Tyr27 interaction. *Biochemistry* **31**, 905–910.

37. Likos, J.J. (1989) ^{1}H NMR studies of squash seed trypsin inhibitor. *Int. J. Peptide Protein Res.* **34**, 381–386.

38. Chiche, L., Gaboriaud, C., Heitz, A., Mornon, J.-P., Castro, B. & Kollman, P.A. (1989) Use of restrained molecular dynamics in water to determine three dimensional protein structure: Prediction of the three dimensional structure of *Ecballium elaterium* trypsin inhibitor II. *Proteins* **6**, 405–417.

39. Bode, W., Epp, O., Huber, R., Laskowski, M., Jr. & Ardelt, W. (1985) The crystal and molecular structure of the third domain of silver pheasant ovomucid (OMSVP3). *Eur. J. Biochem.* **147**, 387–395.

40. McPhalen, C.A. & James, M.N.G. (1987) Crystal and molecular structure of the serine proteinase inhibitor CI-2 from barley seeds. *Biochemistry* **26**, 261–269.

41. Chiche, L., Heitz, A., Padilla, A., Le-Nguyen, D. & Castro, B. (1993) Solution conformation of synthetic bis-headed inhibitor of trypsin and carboxypeptidase A: New structural alignment between the squash inhibitor and the potato carboxypeptidase inhibitor. *Protein Eng.* **6**, 675–682.

42. Nielsen, K.J., Alewood, D., Andrews, J., Kent, S.B.H. & Craik, D.J. (1994) An ^{1}H NMR determination of the three dimensional structures of mirror image forms of a Leu-5 variant of the trypsin inhibitor from *Ecballium elaterium* (EETI II). *Protein Science* **3**, 291–302.

43. Krishnamoorthi, R., Gong, Y., Lin, C.-L.S. & VanderVelde, D. (1992) Two-dimensional NMR studies of squash family inhibitors. Sequence specific proton assignment and secondary structure of reactive site hydrolyzed *Cucurbita maxima* trypsin inhibitor III. *Biochemistry* **31**, 898–904.

44. Papamokos, E., Weber, E., Bode, W., Huber, R., Empie, M.W., Kato, I. & Laskowski, M., Jr. (1982) Crystallographic refinement of Japanese quail ovomucoid, a Kazal-type inhibitor and model building studies of complexes with serine proteases. *J. Mol. Biol.* **158**, 515–537.

45. Wagner, G., Braun, W., Havel, T.F., Schaumann, T., Go, N. & Wuthrich, K. (1987) Protein structures in solution by nuclear magnetic resonance and distance geometry. The polypeptide of the basic pancreatic trypsin inhibitor determined using two different algorithms, DISGEO and DISMAN. *J. Mol. Biol.* **196**, 611–639.

46. Clore, G.M., Gronenborn, A.M., James, M.N.G., Kjaer, M., McPhalen, C.A. & Poulsen, F.M. (1987) The determination of the three-dimensional structure of barley serine proteinase inhibitor 2 in solution: A study using nuclear magnetic resonance, distance geometry and restrained molecular dynamics. *Protein Eng.* **1**, 313–318.

47. Otlewski, J. & Zbyryt, T. (1994) Single peptide bond hydrolysis/resynthesis in squash inhibitors of serine proteinases. I. Kinetics and

thermodynamics of the interaction between squash inhibitors and bovine β-trypsin. *Biochemistry* **33**, 200–207.

48. Otlewski, J., Zbyryt, T., Krokoszynska, I. & Wilusz, T. (1990). Inhibition of serine proteinases by squash inhibitors. *Biol. Chem. Hoppe-Seyler* **371**, 589–594.
49. Wynn, R. & Laskowski, M., Jr. (1990) Inhibition of human β-factor XII_A by squash family serine proteinase inhibitors. *Biochem. Biophys. Res. Commun.* **166**, 1406–1410.
50. Estell, D.A., Wilson, K.A. & Laskowski, M., Jr. (1980) Thermodynamics and kinetics of the hydrolysis of the reactive site peptide bond in pancreatic trypsin inhibitor (Kunitz) by *Dermasterias imbricata* trypsin. *Biochemistry* **19**, 131–137.
51. Quast, U., Engel, J., Heumann, H., Krauze, G. & Steffen, E. (1978) Kinetics of the interaction of bovine pancreatic trypsin inhibitor (Kunitz) with α-chymotrypsin. *Biochemistry* **13**, 2512–2520.
52. Hiromi, K., Akasaka, K., Mitsui, Y., Tonomura, B. & Murao, S. (1985) *Protein protease inhibitor — The case of Streptomyces subtilisin inhibitor (SSI)*; Elsevier, Amsterdam.
53. Ardelt, W. & Laskowski, M., Jr. (1991) Effect of single amino acid replacements on the thermodynamics of the reactive site peptide bond hydrolysis in ovomucoid third domains. *J. Mol. Biol.* **220**, 1041–1053.
54. Zbyryt, T. & Otlewski, J. (1991) Interaction between squash inhibitors and bovine trypsinogen. *Biol. Chem. Hoppe-Seyler* **372**, 255–262.
55. Huber, R. & Bode, W. (1978) Structural basis of the activation and action of trypsin. *Acc. Chem. Res.* **11**, 114–122.
56. Antonini, E., Ascenzi, P., Bolognesi, M., Gatti, G., Guarneri, M. & Menegatti, E. (1983) Interaction between serine (pro)enzymes, and Kazal and Kunitz inhibitors. *J. Mol. Biol.* **165**, 543–558.
57. Otlewski, J., Zbyryt, T., Dryjanski, M., Bulaj, G. & Wilusz, T. (1994) Single peptide bond hydrolysis/resynthesis in squash inhibitors of serine proteinases. II. Limited proteolysis of *Cucurbita maxima* trypsin inhibitor I (CMTI I) by pepsin. *Biochemistry* **33**, 208–213.
58. Sottrup-Jensen, L., Sand, O., Kristensen, L. & Fey, G.H. (1989) The α-macroglobulin bait region. *J. Biol. Chem.* **264**, 15781–15789.
59. Mast, A.E., Enghild, J.J. & Salvesen, G. (1992) Conformation of the reactive site loop of α_1-proteinase inhibitor probed by limited proteolysis. *Biochemistry* **31**, 2720–2728.
60. Kupryszewski, G., Ragnarsson, U., Rolka, K. & Wilusz, T. (1986) Solid-phase synthesis of trypsin inhibitor CMTI III from squash seeds (*Cucurbita maxima*). *Int. J. Peptide Protein Res.* **27**, 245–250.
61. Rolka, K., Kupryszewski, G., Ragnarsson, U., Otlewski, J., Wilusz, T. & Polanowski, A. (1989) Synthesis of an elastase inhibitor by monosubstitution of Arg5 with Val at the reactive site in a trypsin inhibitor from squash seeds (CMTI III). *Biol. Chem. Hoppe-Seyler* **370**, 499–502.
62. Rolka, K., Kupryszewski, G., Ragnarsson, U., Otlewski, J., Krokoszynska, I. & Wilusz, T. (1991). New synthetic analogues of trypsin inhibitor CMTI III from squash seeds; in *Peptides 1990. Proceedings of the Twenty-First European Peptide Symposium* (Giralt, E. & Andreu, D., eds.) pp. 768–771, ESCOM Science Publ.
63. Rolka, K., Kupryszewski, G., Ragnarsson, U., Otlewski, J., Krokoszynska, I. & Wilusz, T. (1991) Chemical synthesis of new trypsin, chymotrypsin and elastase inhibitors by amino-acid substitutions in a trypsin inhibitor from squash seeds (CMTI I). *Biol. Chem. Hoppe-Seyler* **372**, 63–68.
64. Rolka, K., Kupryszewski, G., Rozycki, J., Ragnarsson, U., Zbyryt, T. & Otlewski, J. (1992) New analogues of *Cucurbita maxima* trypsin inhibitor III (CMTI I III) with simplified structure. *Biol. Chem. Hoppe-Seyler* **373**, 1055–1060.
65. McWherter, C.A., Walkenhorst, W.F., Campbell, E.J. & Glover, G.I. (1989) Novel inhibitors of human leukocyte elastase and cathepsin G. Sequence variants of squash seed protease inhibitor with altered protease selectivity. *Biochemistry* **28**, 5708–5714.
66. Favel, A., Le-Nguyen, D., Coletti-Previero, M.A. & Castro, B. (1989) Active site chemical mutagenesis of *Ecballium elaterium* trypsin inhibitor II: New microproteins inhibiting elastase and chymotrypsin. *Biochem. Biophys. Res. Commun.* **162**, 79–82.
67. Le-Nguyen, D., Mattras, H., Coletti-Previero, M.A. & Castro, B. (1989) Design and chemical synthesis of a 32 residues chimeric microprotein inhibiting both trypsin and carboxypeptidase A. *Biochem. Biophys. Res. Commun.* **162**, 1425–1430.
68. Huang, Z., Wu, M., Qi, Z. & Chi, C. (1990) Total synthesis of *Trichosanthes* trypsin inhibitor and its analogue. *Science in China B*, **33**, 1192–1200.
69. Le-Nguyen, D., Heitz, A., Chiche, L., Hajji, M.E. & Castro, B. (1993) Characterization and 2D NMR study of the stable [9–21, 15–27] 2 disulfide intermediate in the folding of the 3

disulfide trypsin inhibitor EETI II. *Protein Science* **2**, 165–174.

70. Heitz, A., Chiche, L., Le-Nguyen, D. & Castro, B. (1995) Folding of the squash trypsin inhibitor EETI II. *Eur. J. Biochem.* **233**, 837–846.
71. Pallaghy, P.K., Nielsen, K.J., Craik, D.J. & Norton, R.S. (1994) A common structural motif incorporating a cystine knot and a triple--stranded β-sheet in toxic and inhibitory polypeptides. *Protein Science* **3**, 1833–1839.
72. McDonald, N.Q. & Hendrickson, W.A. (1993) A structural superfamily of growth factors containing a cystine knot motif. *Cell* **73**, 421–424.
73. Wells, J.A. (1990) Additivity of mutational effects in proteins. *Biochemistry* **29**, 8509–8517.
74. Laskowski, M., Jr., Park, S.J., Tashiro, M. & Wynn, R. (1989) Design of highly specific inhibitors of serine proteinases. *UCLA Symposia on Molecular and Cellular Biology, New Series, 80* (Hutchens, T.W., ed.) pp. 149–168, Liss, Inc., New York.
75. Hayashi, K., Takehisa, T., Hamato, N., Takano, R., Hara, S., Miyata, T. & Kato, H. (1994) Inhibition of serine proteinases of the blood coagulation system by squash family protease inhibitors. *J. Biochem. (Tokyo)* **116**, 1013–1018.
76. Chen, X., Qian, Y., Chi, C., Gan, K., Zhang, M. & Chang-Qing, C. (1992) Chemical synthesis, molecular cloning, and expression of the gene coding for the *Trichosanthes* trypsin inhibitor — a squash family inhibitor. *J. Biochem. (Tokyo)* **112**, 45–51.
77. Rempola, B., Wilusz, T., Markiewicz, W. & Fikus, M. (1995) Synthesis, cloning, and expression in *Escherichia coli* of the gene coding for the trypsin inhibitor from *Cucurbita pepo*. *Acta Biochim. Polon.* **42**, 109–114.
78. Bolewska, K., Krowarsch, D., Otlewski, J., Lukaszewski, L., & Bierzynski, A. (1995) Synthesis, cloning, and expression in *Escherichia coli* of a gene coding for the Met8 → Leu CMTI I — a representative of the squash inhibitors of serine proteinases. *FEBS Lett.* **377**, 172–174.
79. Nishino, J., Takano, R., Kamei-Hayashi, K., Minakata, H., Nomoto, K. & Hara, S. (1992) Amino-acid sequences of trypsin inhibitors from oriental pickling melon (*Cucumis melo* L. var. *Conomon* Makino) seeds. *Biosci. Biotech. Biochem.* **56**, 1241–1246.
80. Hamato, N., Koshiba, T., Pham, T., Tatsumi, Y., Nakamura, D., Takano, R., Hayashi, K., Hong, Y. & Hara, S. (1992) Trypsin and elastase inhibitors from bitter gourd (*Momordica charantia* Linn.) seeds: Purification, amino acid sequences, and inhibitory activities of four new inhibitors. *J. Biochem. (Tokyo)* **117**, 432–437.
81. Lee, C. & Lin, J. (1995) Amino acid sequences of trypsin inhibitors from the melon *Cucumis melo*. *J. Biochem. (Tokyo)* **118**, 18–22.