

An Unusual Helix-Turn-Helix Protease Inhibitory Motif in a Novel Trypsin Inhibitor from Seeds of Veronica (*Veronica hederifolia* L.)*

Received for publication, May 10, 2007, and in revised form, June 28, 2007 Published, JBC Papers in Press, July 19, 2007, DOI 10.1074/jbc.M703871200

Rebecca Conners[‡], Alexander V. Konarev^{§¶}, Jane Forsyth^{‡¶}, Alison Lovegrove[¶], Justin Marsh[¶], Timothy Joseph-Horne[‡], Peter Shewry[¶], and R. Leo Brady^{‡¶1}

From the [‡]Department of Biochemistry, University of Bristol, Bristol BS8 1TD, United Kingdom, the [§]All-Russian Institute for Plant Protection, 3 Podbelsky, Pushkin, St. Petersburg 196608, Russia, and the [¶]Rothamsted Research, Harpenden, Hertfordshire AL5 2JQ, United Kingdom

The storage tissues of many plants contain protease inhibitors that are believed to play an important role in defending the plant from invasion by pests and pathogens. These proteinaceous inhibitor molecules belong to a number of structurally distinct families. We describe here the isolation, purification, initial inhibitory properties, and three-dimensional structure of a novel trypsin inhibitor from seeds of *Veronica hederifolia* (VhTI). The VhTI peptide inhibits trypsin with a submicromolar apparent K_i and is expected to be specific for trypsin-like serine proteases. VhTI differs dramatically in structure from all previously described families of trypsin inhibitors, consisting of a helix-turn-helix motif, with the two α helices tightly associated by two disulfide bonds. Unusually, the crystallized complex is in the form of a stabilized acyl-enzyme intermediate with the scissile bond of the VhTI inhibitor cleaved and the resulting N-terminal portion of the inhibitor remaining attached to the trypsin catalytic serine 195 by an ester bond. A synthetic, truncated version of the VhTI peptide has also been produced and co-crystallized with trypsin but, surprisingly, is seen to be uncleaved and consequently forms a noncovalent complex with trypsin. The VhTI peptide shows that effective enzyme inhibitors can be constructed from simple helical motifs and provides a new scaffold on which to base the design of novel serine protease inhibitors.

Plant seeds are rich sources of proteinaceous protease inhibitors. These are believed to form a wide spectrum defense mechanism against fungal pathogens and invertebrate pests but may also play a role in the regulation of metabolism and act as

storage proteins (1). A diverse range of medicinal properties have also been associated with many of these inhibitors, including anti-human immunodeficiency virus activity (2), hemolytic activity (3, 4), and inhibition of neurotensin binding (5). On the basis of their amino acid sequence and target proteinases, plant proteinase inhibitors have been classified into a number of families (6). The two best characterized are the Kunitz and Bowman-Birk families. The Kunitz soybean trypsin inhibitor was the first to be extensively characterized (7) and is an all β -sheet protein of 20 kDa. One exposed surface loop and the N terminus of the protein interact closely with the trypsin molecule, whereas the vast majority of the inhibitor forms no direct contacts with its inhibitory target. The Bowman-Birk family of serine protease inhibitors (reviewed in Ref. 8) are smaller proteins of ~8 kDa that contain seven conserved disulfide bridges. They have two reactive sites that are able to bind to the active sites of a number of serine proteases including trypsin and chymotrypsin from human, animal, and insect sources. The reactive site residues of the inhibitor lie within a β -hairpin region (stabilized by a disulfide bond), which enables them to be presented in the same conformation as the normal peptide substrate (9). Like the Kunitz family of serine protease inhibitors, the Bowman-Birk family are also all- β structures.

In recent years two new types of proteinase inhibitors have been discovered in seeds, both of which are small cyclic peptides. These are the “cyclic knottins” that belong to a large and diverse family of cyclic peptides present in the families Rubiaceae, Violaceae, and Cucurbitaceae (10–13) and the sunflower trypsin inhibitor SFTI-1 (14, 15), which has only been found in seeds of *Helianthus* L. (sunflower) and the related genus *Tithonia* Desf. Ex Juss (16). Both of these families characteristically contain circular (*i.e.* head-to-tail cyclized) peptide structures that are further stabilized through the incorporation of disulfide bonds. Once again, these are essentially composed of β -strands, with the conformation of the residues that insert into the enzyme active site being determined by their location in a loop region. The mechanism(s) by which cyclization is achieved in these peptides within the plant remains unclear (15).

We have carried out extensive surveys of proteinase inhibitors in seeds of the family Compositae (which contains *Helianthus*) (16) and of the clade Asteridae, which comprises some 100 families including the Compositae (18). From these studies

* This work was supported by Biotechnology and Biological Sciences Research Council of the UK Grant C502414 (to R. C., T. J.-H., and R. L. B.) and also by grant-aided support from the Biotechnology and Biological Sciences Research Council to Rothamsted Research. This work was also supported by funds from the Fundamental Researches Programme of the Russian Academy of Agricultural Sciences (to A. V. K.) and by an “Ex-agreement Visit” grant from the Royal Society of London and a Fellowship from Rothamsted International (to P. S. and A. V. K.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 2CMY and 2PLX) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (<http://www.rcsb.org/>).

¹ To whom correspondence should be addressed. Tel.: 44-117-3312150; Fax: 44-117-3312168; E-mail: L.Brady@bris.ac.uk.

we reported the identification of a potentially novel peptide inhibitor of trypsin (VhTI)² from seeds of *Veronica hederifolia* L., a member of the Scrophulariaceae (16). Further studies reported here demonstrate that VhTI represents a completely novel form of peptide inhibitor of trypsin, distinguished by a helix-turn-helix structure not previously described. This characteristic structure appears amenable to ready modification in the design of specific protease inhibitors.

EXPERIMENTAL PROCEDURES

Wild-type Inhibitor Isolation and Purification

Seeds of *V. hederifolia* L. (Scrophulariaceae) were purchased from Herbiseed (Twyford, UK). 100 g of seed were milled and defatted by stirring for 2×12 h at 20 °C with 600 ml of hexane. The meal was then stirred with 1 liter of water for 30 min at 20 °C and centrifuged, and ammonium acetate was added to the supernatant to 0.2 M. After centrifugation 15 ml of trypsin-Sepharose gel (Pharmacia; prepared according to the manufacturer's instructions) was added to the supernatant and placed on an orbital shaker. After 15 min the gel was collected onto a glass filter and washed sequentially with 200 ml of 0.2 M ammonium acetate, water, 0.1 M Na₂CO₃, water, 0.1 M sodium acetate buffer, pH 4.5, and water. The gel was then transferred to a column, and the bound trypsin inhibitors were eluted with 0.015 M HCl. Fractions were lyophilized and analyzed for the presence of trypsin inhibitors using isoelectric focusing combined with the gelatin replicas method (18, 19). Selected fractions were then separated by reverse phase HPLC with a C18 RP Phenomenex column and a gradient of 15–45% (v/v) acetonitrile in 0.1% trifluoroacetic acid. Lyophilized fractions were analyzed for inhibitors of trypsin and subtilisin using gelatin replicas.

Primary Sequence Determination

The N-terminal sequence of purified VhTI was determined at the University of Cambridge Protein and Nucleic Acid Chemistry Facility. Mass Spectrometric analysis was also used to elucidate the peptide amino acid sequence. Following chymotrypsin, subtilisin, and trypsin digestion, the peptides were concentrated and desalted using Zip-Tips (Millipore) and then dried. The peptides were then dissolved in 70% (v/v) methanol containing 1% (v/v) formic acid and sonicated for 3 min. The peptides were loaded into nanoflow tips (Waters) using gel loader tips (Eppendorf). Electrospray ionization-MS was performed on a quadrupole time-of-flight (Q-TOF) I mass spectrometer (Micromass, Manchester, UK) equipped with a z-spray ion source. Instrument operation, data acquisition, and analysis were performed using MassLynx/Biolynx 4.0 software. The sample cone voltage and collision energy were optimized for each sample. The Microchannel plate detector voltage was set at 2800 V. Scanning was performed from *m/z* 100 to 3500.

Synthesis and Refolding of Truncated VhTI

The sequence EQCKVMCYAQRHSSPELLRRCLDNCEK with a free amide group at the C terminus was synthesized by

the University of Bristol Peptide Synthesis Facility and purified by reverse phase HPLC. The peptide was refolded by rapid dilution into 50 mM Tris, pH 8.0, to 0.1 mg/ml with fast stirring for at least 8 h to allow air oxidation of the disulfide bonds.

Determination of Antitrypsin Activity

Inhibitory activity was assayed against bovine trypsin (tosylphenylalanyl chloromethyl ketone treated from Sigma-Aldrich) using the colorimetric trypsin substrate L-BAPNA based on the method described in Ref. 20. Trypsin was dissolved in 1 mM HCl, and final concentrations of 50, 100, and 150 nM were used in the assays. L-BAPNA was dissolved in Me₂SO, and a final concentration of 60 μM was used. Increasing amounts of VhTI peptide (giving final concentrations of 0–250 nM) were incubated for 5 min with the trypsin at 20 °C in 50 mM Tris, pH 8.2, 25 mM CaCl₂ before the reaction was initiated by the addition of L-BAPNA and mixed well, and the change in absorbance at 410 nm followed over 20 min. The VhTI concentration was determined by measuring the amide bond absorbance at 205 nm.

Crystallization

Native VhTI Complexed with Trypsin—Crystals of bovine trypsin were obtained by the vapor diffusion technique as described previously (14). Bovine trypsin (Type III; Sigma) was dissolved to a final concentration of 30 mg/ml in 0.3 M ammonium sulfate, 6 mM calcium chloride, 0.1 M Tris, pH 8.15, 60 mM benzamidine. 0.5 μl of dimethyl formamide was added to the crystallization drop before sealing the wells. The well solution consisted of 1.6–2.1 M ammonium sulfate, 50 mM Tris, pH 8.15, and the crystals grew within 2 weeks.

The crystals were soaked overnight in a backsoak solution (1 ml of 0.1 M sodium phosphate, pH 5.8, 2.5 M ammonium sulfate, 1 mM calcium chloride), and then six further buffer exchanges were performed to completely remove the benzamidine inhibitor (verified by x-ray diffraction analysis of the backsoaked crystals; data not shown). After this procedure, the crystals were transferred to a 20-μl drop, and the pH of the backsoak solution increased to 8.0 by gradual removal of backsoak solution and replacement with binding solution (0.1 M Tris, pH 8.0, 2.5 M ammonium sulfate, 1 mM calcium chloride). The *V. hederifolia* inhibitory peptide was dissolved in this binding solution with the addition of 5% (v/v) dimethyl formamide to improve its solubility. Inhibitor solution was added to the crystal soak to give a final concentration of 10 mM VhTI and incubated for 16 h at 18 °C. A number of attempts at obtaining crystals of the trypsin-inhibitor complex were tried with different fractions of inhibitor from the purification process (see Fig. 1) before a complex was observed.

Synthetic VhTI Complexed with Trypsin—Synthetic VhTI peptide was added to trypsin in a 1:1 molar ratio in 50 mM Tris, pH 8.0. The mixture was incubated overnight at 18 °C and then concentrated to a final concentration of 8.5 mg/ml in a 5-kDa Vivaspinn concentrator (Millipore), which also served to remove any unbound peptide from the mixture. The initial crystals were obtained using the hanging drop vapor diffusion method against a sparse matrix crystallization screen (Crystal Screen HT; Hampton Research), and the conditions were optimized to

² The abbreviations used are: VhTI, *V. hederifolia* trypsin inhibitor; HPLC, high pressure liquid chromatography; MS, mass spectrometry; Q-TOF, quadrupole time-of-flight; L-BAPNA, (*N*^ε-benzoyl-L-arginine *p*-nitroanilide).

TABLE 1

Summary of x-ray diffraction data and refined model statistics for native and synthetic trypsin-VhTI complexes

The values in parentheses are the statistics for the highest resolution shell. The refined models were validated using Molprobit (30).

	Native VhTI-trypsin	Synthetic VhTI-trypsin
Diffraction data		
Resolution range (Å)	50–2.25 (2.33–2.25)	56–1.56 (1.60–1.56)
Completeness (%)	97.5 (84.9)	80.2 (11.5)
Total no. of unique reflections	13342 (1141)	26929 (196)
Redundancy	4.2 (3.6)	3.8 (1.1)
$I/\sigma I$	11.9 (1.4)	30.5 (1.1)
R_{sym}^a	0.123 (0.589)	0.044 (0.611)
$R_{\text{p.i.m.}}^b$	0.062 (0.294)	0.031 (0.037)
$R_{\text{r.i.m.}}^c$	0.128 (0.569)	0.064 (0.076)
Wilson B factor	52.7	14.9
Refinement statistics		
Total no. of reflections	12681	26929
No. of reflections in test set	661	1442
R_{cryst}^d	0.194	0.145
R_{free}^d	0.253	0.188
Root mean square deviation		
Bond lengths (Å)	0.02	0.02
Bond angles (°)	1.89	1.74
Bonded B factors		
Main chain	0.930	0.911
Side chain	2.481	2.160
Protein atoms	1629	1693
Inhibitor atoms	169	219
Solvent molecules	110	318
B-average		
Protein (Å ²)	73.2	15.4
Peptide (Å ²)	82.8	15.6
Residues in most favored regions Ramachandran plot (%)	96.2	97.0

$$^a R_{\text{sym}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl).$$

$$^b R_{\text{p.i.m.}} = \sum_{hkl} [1/(N-1)]^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl).$$

$$^c R_{\text{r.i.m.}} = \sum_{hkl} [N/(N-1)]^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl).$$

$$^d R_{\text{cryst}} = \sum_{hkl} |F_o(hkl) - F_c(hkl)| / \sum_{hkl} F_o(hkl).$$

obtain diffraction quality crystals. The crystals grew from 1.3 to 1.5 M sodium citrate, 0.1 M Na-Hepes, pH 7.5.

Crystal Structure Determination

Native VhTI Complexed with Trypsin—The diffraction data were collected on beamline PX14.1 ($\lambda = 1.488$ Å) at Daresbury SRS (UK). The crystal was cooled to 100 K after a brief soak in crystallization buffer containing 10 mM inhibitor plus 25% (v/v) glycerol for cryoprotection. The data were collected to 2.25 Å resolution and were processed with *HKL2000* (21) and manipulated with the CCP4 suite of crystallographic software (22). The data collection statistics are summarized in Table 1. The crystals belonged to the space group $P2_12_12_1$ with unit cell dimensions $a = 60.7$, $b = 63.9$, $c = 71.7$ Å and one molecule of the complex per asymmetric unit. The structure was solved by molecular replacement using the program *AMoRe* (23) with coordinates from the refined 1.65 Å resolution structure of bovine trypsin complexed with sunflower trypsin inhibitor (14) (Protein Data Bank code 1SFI; inhibitor and water molecules removed) as a search model. The auto-amore setting picked one clear solution after both rotation and translation functions, with a correlation coefficient of 74.4% and an R_{cryst} of 35.5% after subsequent fitting. The structure was refined with iterative cycles of manual model building using *Coot* (24), restrained refinement with *Refmac5* (25), and density improvement with the “atoms update and refinement” mode of *ARP/wARP* (26). A polyaniline model was initially built into the inhibitor electron density, and the side chains gradually mutated as the electron density maps became clearer.

Synthetic VhTI Complexed with Trypsin—The diffraction data were collected on beamline PX14.1 ($\lambda = 1.488$ Å) at Dares-

bury SRS. After cryo-protection in well solution plus 10% glycerol, the crystal was cooled to 100 K, and data were collected to 1.65 Å resolution (based on >50% completion of the highest resolution shell). The data were processed as above, and the data collection statistics are summarized in Table 1. The crystal belonged to the space group $P2_12_12_1$ with unit cell dimensions $a = 113.6$, $b = 41.5$, $c = 51.2$ Å and one molecule of the complex per asymmetric unit. The structure was solved by molecular replacement using the program *Phaser* (27) with the same search model as detailed above. There was one clear solution with Z scores of 23 and 30 after rotation and translation, respectively. Structure refinement was performed as detailed above. The final refinement statistics for both structures are summarized in Table 1.

RESULTS

The trypsin inhibitor fraction prepared from seeds of *Veronica* by trypsin affinity chromatography followed by reverse phase HPLC was separated by isoelectric focusing into over 20 individual components, ranging in pI from ~4.7 to 7.5 (Fig. 1). In this separation the proteins have been transferred from the isoelectric focusing gel to an opaque layer of gelatin (an undeveloped photographic film) that had been incubated on a plate of agarose containing trypsin, allowing zones of activity to be revealed as dark “islands” of undigested gelatin on a transparent background. Preliminary analyses of these components by matrix-assisted laser desorption ionization TOF MS showed that they ranged in mass from ~3,650 to 4,320, with most having masses of ~3,900–4,000. Furthermore, in all cases the components were mixtures of intact peptides and components that had been cleaved (presumably by trypsin during the affinity



FIGURE 1. Isoelectric focusing of the major trypsin inhibitor fraction from seeds of *V. hederifolia*. The inhibitors were separated in the pH range 3–10 and then transferred to a layer of gelatin on a photographic film that was placed onto an agarose gel containing trypsin. The dark “islands” of undigested protein represent zones of inhibitory activity.

chromatography) to give two peptides. For example, band 17 gave masses of 3,934 (intact) and 3,951 (cleaved) when analyzed without reduction with the latter component being replaced by peptides of 2,280 and 1,671 when reduced. Analysis of a number of components by Q-TOF mass spectrometry indicated that all were related, with the differences in mass explained largely by the presence of ragged N and C termini.

The consensus amino acid sequence of the VhTI was determined by Edman and Q-TOF MS analysis. Automated Edman degradation of peptides with masses of 1,455 and 2,280 gave the sequences PEQCKVMCYAQR and HSSPELLRRCLDNCEKEHD, which is consistent with cleavage between Arg and His during preparation. This was confirmed by Q-TOF MS analysis of a peptide generated by subtilisin digestion of intact inhibitor, which gave the sequence AQRHSSPELL. Finally, Q-TOF MS of a tryptic peptide from a component of mass 3,934 gave the sequences NTDPEQCKVMCYAQR and CLDNCEKEHD. This allowed the full sequence to be assembled as NTDPEQCKVMCYAQRHSSPELLRRCLDNCEKEHD and the mass to be calculated as 4,188 Daltons. The four cysteine residues are available to form disulfide bonds, and trypsin cleaves between Arg¹⁵ and His¹⁶, approximately halfway along the peptide of 34 residues (see Fig. 3a).

Synthesis of VhTI and Inhibition of Enzymatic Activity—Attempts to measure the binding constant of the native, full-length VhTI for trypsin were complicated by the very low yields of inhibitor obtained. Because the crystal structure revealed that parts of the inhibitor were disordered and hence did not contribute to its complex with trypsin, a synthetic, truncated form of VhTI was prepared consisting of residues 5–31. The disordered N- and C-terminal regions were omitted in a successful attempt to improve crystal order and to demonstrate that only the core region of the peptide was required for full inhibitory activity. The truncated peptide included all residues in both the $\alpha 1$ and $\alpha 2$ helices as observed in the crystal structure and the (uncleaved) connecting loop that inserts into the active site (see below for details). Under the conditions described this peptide appeared to refold spontaneously, as judged by its ability to associate with trypsin in an equimolar fashion. The inhibition studies (Fig. 2) show an entirely linear inhibition curve as increasing amounts of inhibitor are used. Once the concentration of the inhibitor reaches the same concentration as the trypsin used in the assay, all enzymatic activity is completely abolished, showing infinitely tight binding of VhTI to trypsin. Because of the peptide cannot be detected under the assay conditions, neither a K_D nor K_i value can be determined. Nonetheless, on the basis of the behavior of other peptide inhibitors such as SFTI-1 (14) in this same assay, we estimate the apparent K_D is likely to be subnanomolar (<1 nM). VhTI is an extremely potent inhibitor of trypsin, and the kinetic data we have collected sup-

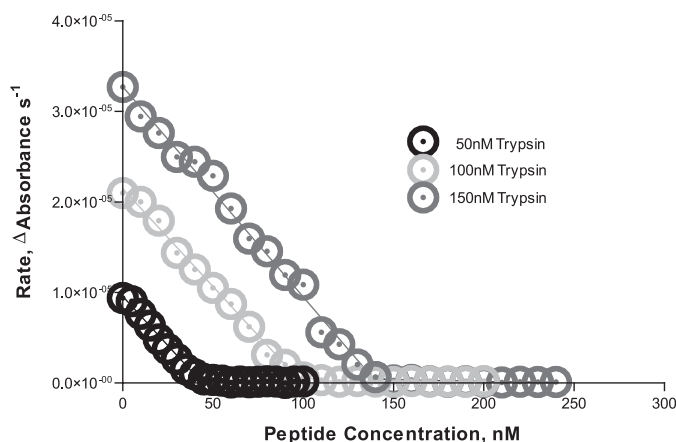


FIGURE 2. Inhibition of bovine trypsin by synthetic VhTI. Trypsin activity was assayed using L-BAPNA as substrate, and the change in OD at 410 nm was monitored, as described under “Experimental Procedures.” Note that for each concentration the rate of reaction reaches 0 when the concentration of inhibitor matches the enzyme concentration, indicating formation of a 1:1 molar complex.

ports either a covalent or a noncovalent interaction between inhibitor and protease.

Structure of VhTI

Native VhTI—The crystal structure shows that VhTI binds to the active site of trypsin and consists essentially of two anti-parallel α -helices that are linked together by two disulfide bonds (Fig. 3). Of the complete 34 residue inhibitor, only residues 7–15 and 18–29 are well ordered in the electron density maps with each stretch of residues forming an α -helix ($\alpha 1$ and $\alpha 2$, respectively). Six residues are disordered at the N terminus of the inhibitor, and five are disordered at the C terminus. Both termini point away from the trypsin molecule, projecting into the solvent region of the crystal where they do not make any specific contacts with the trypsin and do not have defined electron density. In the crystal structure the inhibitor appears to be cleaved at the Arg¹⁵-His¹⁶ P1-P1' scissile bond (referred to using the standard nomenclature (28)), a relatively common occurrence in peptide inhibitor and protease complexes (29). The electron density for residues His¹⁶ and Ser¹⁷, directly after the site of cleavage, is poor, and these residues have not been included in the final model (Fig. 3c). All of the modeled residues of the inhibitor fall within the most favored or additionally allowed regions of the Ramachandran plot as determined by MOLPROBITY (30).

The structure of VhTI is expected to be relatively rigid with two rod-like helices bound together by two disulfide bonds between residues Cys⁷-Cys²⁹, and Cys¹¹-Cys²⁵. These two disulfides are located adjacent to one another, on the same side of each helix with the cysteines each separated by one helical turn. The disulfides form the wall of a small hydrophobic core formed between the helices, also comprising the side chains of Leu²², Leu²⁶, Tyr¹², and the alkyl chain of Lys⁸. No hydrogen bonds (direct or water-mediated) are formed between the two helices.

The P1 residue (Arg¹⁵) is extremely well defined in its binding site within the S1 specificity pocket of trypsin and is situated at the carboxyl end of the $\alpha 1$ helix. The distance from the back-

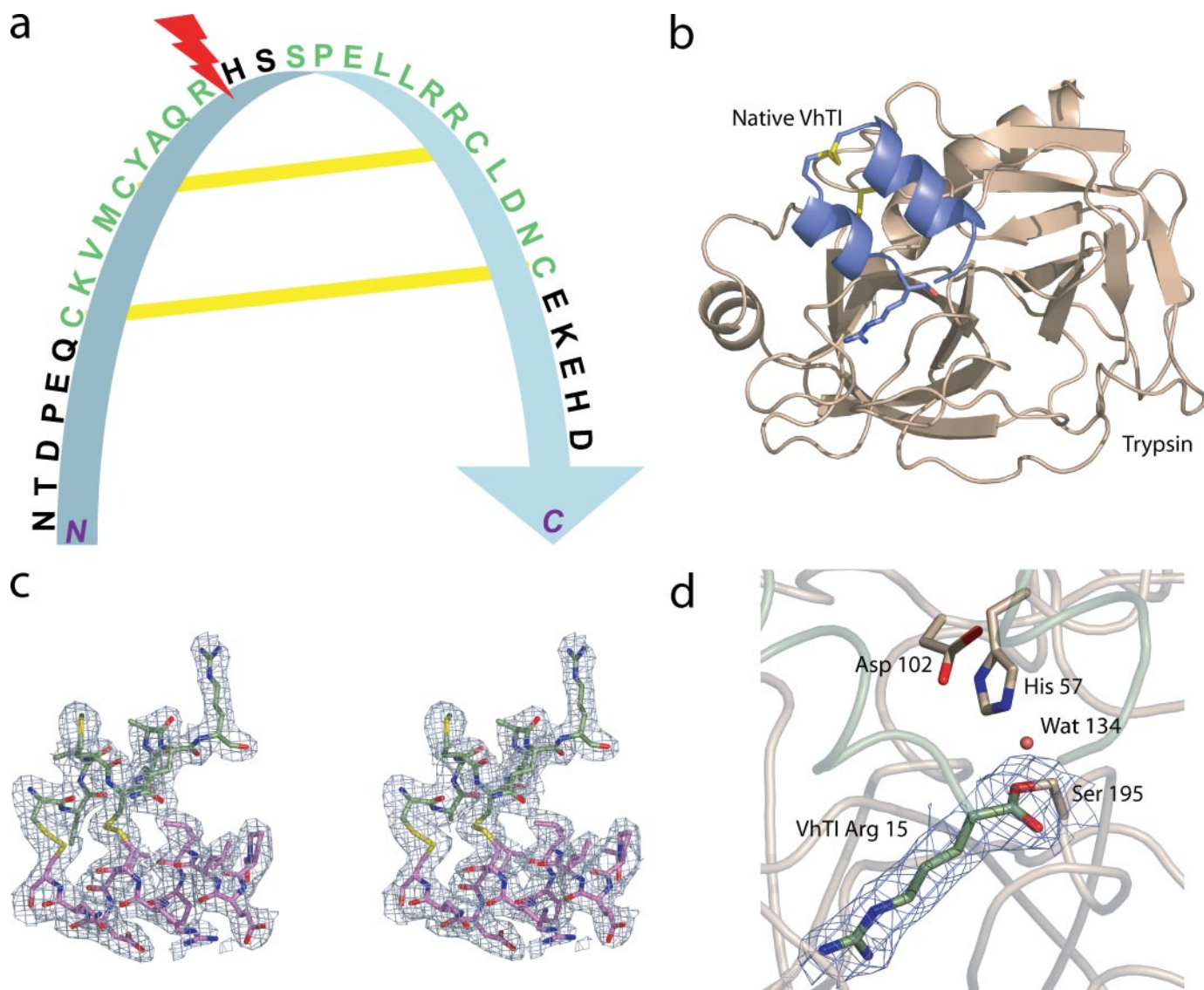


FIGURE 3. Structure of native VhTI. *a*, schematic of the native peptide structure. The two disulfide bonds are shown in yellow, ordered residues visible in electron density are in green, and disordered residues are in black. The scissile bond is depicted by the red lightning bolt. *b*, three-dimensional structure of VhTI reveals two antiparallel α -helices that are linked together by two disulfide bonds to form a rigid structure. The inhibitor is cleaved by trypsin at the P1–P1' scissile bond within the crystal complex. Trypsin is shown as a beige cartoon, VhTI is a blue cartoon, VhTI disulfide bonds are yellow sticks, and the P1 residue (Arg¹⁵) is a blue stick. Only the VhTI residues that were fully ordered in the electron density maps are shown. This and subsequent figures of trypsin and the trypsin active site are represented in the "standard orientation" for serine proteases as defined by Bode (17) with the active site cleft facing the viewer and bound inhibitor running from left to right. *c*, stereo diagram showing $2F_o - F_c$ electron density for VhTI peptide contoured at 1σ . Helix 7–15 (α_1) is shown in green, helix 18–29 (α_2) is in pink, and bridging disulfide bonds are in yellow. *d*, $2F_o - F_c$ electron density contoured at 1σ for ester linkage between Ser¹⁹⁵ of trypsin and Arg¹⁵ of VhTI. The side chains of the catalytic triad (His⁵⁷, Asp¹⁰², and Ser¹⁹⁵) and the P1 residue (Arg¹⁵) are shown as sticks, the backbone trace of trypsin is colored in beige, and the backbone trace of VhTI is in green.

bone carbon atom of Arg¹⁵ to the backbone nitrogen atom of Ser¹⁸ is 13.7 Å and hence too long for the two helices to be directly bridged by the two disordered amino acid residues (15, 16). We therefore conclude that, following cleavage of the scissile bond (Arg¹⁵-His¹⁶), there is likely to be some movement of the N-terminal end of the α_2 helix away from α_1 and the trypsin active site. This is supported by a comparison of the structures of the cleaved native and uncleaved synthetic forms of the peptide (see below). The N-terminal end of this second α -helix points into the disordered solvent region accounting for the lack of observable electron density for residues 16 and 17. We note that the temperature factors for the peptide are lowest for residues most closely associated with the active site (Tyr¹²–

Arg¹⁵) and highest for residues at both the C and N termini of α_2 , the latter comprising the area believed to move away from the active site post cleavage.

Synthetic VhTI—We created a synthetic peptide of VhTI (EQCKVMCYAQRHSSPELLRRCLDNCEK) that lacked the disordered residues 1–4 and 32–34 at the N and C termini, respectively. This peptide was refolded with air oxidation to allow the disulfide bonds to form, and co-crystals were grown of the synthetic VhTI-trypsin complex (Fig. 4*a*). The structure shows that all residues of this synthetic peptide have well ordered electron density with the sole exception of the N-terminal residue, Glu⁵ (numbering as in the original VhTI), which is fully exposed to the solvent. In contrast to

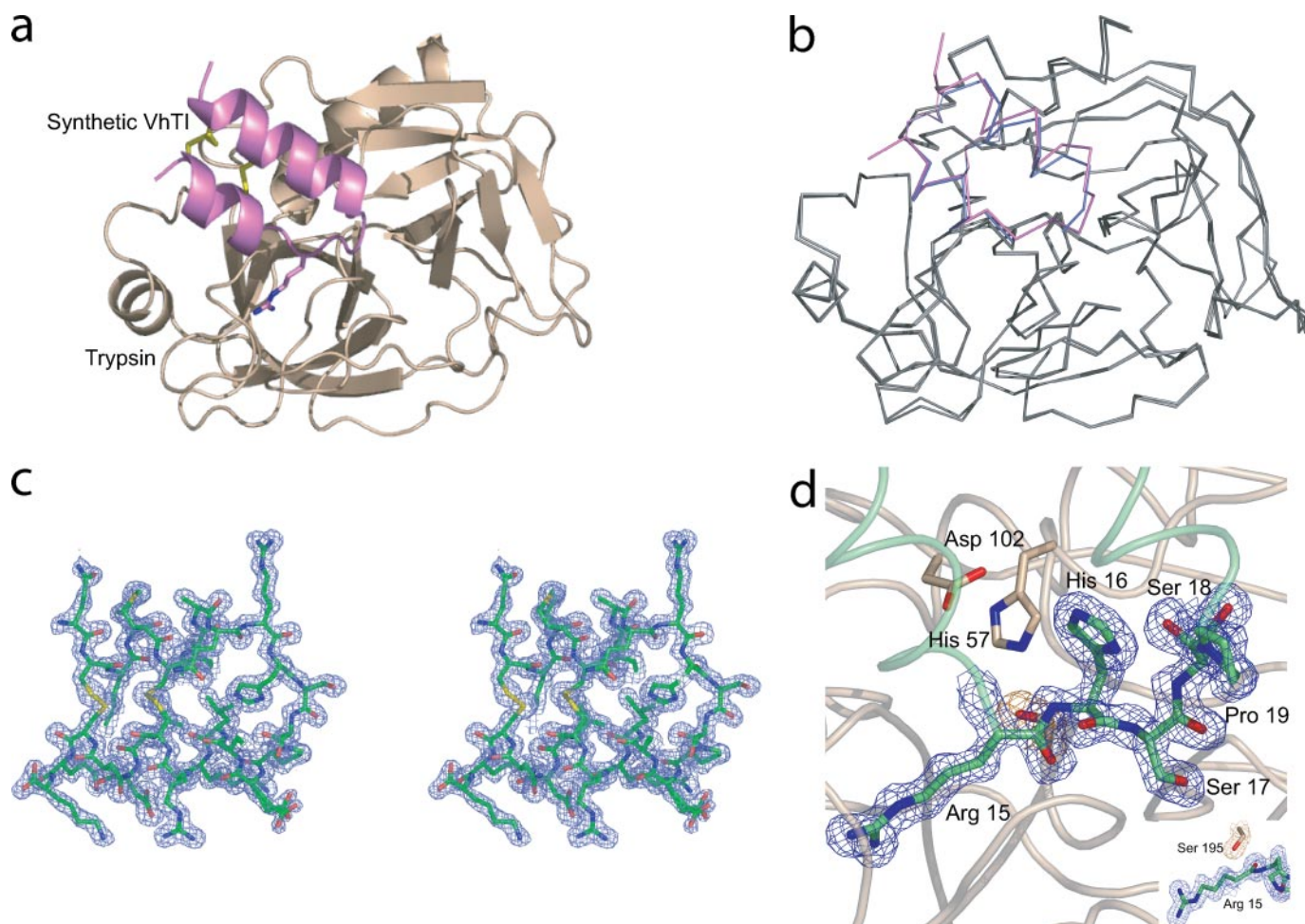


FIGURE 4. Structure of synthetic VhTI. *a*, three-dimensional structure of synthetic VhTI shows the peptide is not cleaved and forms a noncovalent complex with trypsin. Trypsin is shown as a beige cartoon, synthetic VhTI is a purple cartoon, disulfide bonds are yellow sticks, and the P1 residue (Arg¹⁵) is a purple stick. The presence of the uncleaved inhibitor shows that it has a helix-turn-helix motif. *b*, overlay of wild-type VhTI-trypsin complex with the synthetic VhTI-trypsin complex showing the difference between the covalent complex with cleaved peptide (blue trace) and the noncovalent complex with uncleaved peptide (purple trace). *c*, stereo diagram showing $2F_o - F_c$ electron density for synthetic VhTI peptide contoured at 1 σ . Peptide backbone is shown in green, and disulfide bonds are in yellow. *d*, close-up of trypsin active site with synthetic VhTI bound. The side chains of the catalytic triad (His⁵⁷, Asp¹⁰², and Ser¹⁹⁵) are shown as beige sticks. Peptide inhibitor is shown in green with the area spanning the scissile bond highlighted as sticks and showing electron density contoured at 1 σ . The inset shows an absence of connecting $2F_o - F_c$ electron density between trypsin Ser¹⁹⁵ and VhTI Arg¹⁵.

the native complex, the synthetic peptide is not cleaved and residues 16 and 17 (situated immediately after the Arg¹⁵ scissile bond and not visible in the native peptide complex) have clearly defined electron density (Fig. 4*d*). The structure of the synthetic VhTI shows it to have the same helical structure as the native inhibitor and, in addition, reveals that the two helices are joined together by a simple turn. It appears that the peptide can act as an inhibitor of trypsin in either its cleaved or uncleaved state.

Trypsin-Inhibitor Interactions

Native VhTI Complexed with Trypsin—The VhTI inhibitor interacts with trypsin in a substrate-like manner, hence blocking the active site of the protein and competitively preventing substrate from binding. The C-terminal portion of the inhibitor α 1 helix is inserted into the enzyme recognition site at a steep angle, placing the extended Met side chain at position 10 into the enzyme S4 pocket, Ala¹³ in the S3 region, the side chain of Gln¹⁴ in S2, and the side chain of Arg¹⁵ deeply within the S1 pocket (referred to using the standard

nomenclature (28)). Central to this interaction with trypsin is VhTI Arg¹⁵, from which the NH1 and NH₂ atoms of the side chain form a bidentate salt bridge with the OD1 and OD2 atoms of the trypsin Asp¹⁸⁹ side chain, found deep within the substrate specificity pocket, and direct hydrogen bonds with the hydroxyl group of Ser¹⁹⁰ and the carbonyl group of Gly²¹⁹. These interactions have previously been noted in many trypsin-peptide complexes, and in this respect, the VhTI-trypsin interaction mimics that of the natural substrates of the enzyme.

The free carboxyl group of Arg¹⁵ (the result of hydrolysis of the Arg¹⁵-His¹⁶ peptide bond) is seen to lie in very close proximity (1.6 Å) to the O- γ of the catalytic serine 195 side chain. The presence of continuous $2F_{\text{obs}} - F_{\text{calc}}$ electron density between the C-terminal carbon of Arg¹⁵ and the side chain hydroxyl group of Ser¹⁹⁵ is consistent with the structure in the crystals being that of an acyl-enzyme intermediate, with the carboxyl group of the cleaved inhibitor remaining attached to Ser¹⁹⁵ of trypsin via a covalent ester linkage (Fig. 3*d*). The carbonyl oxygen of this ester is stabilized by hydrogen bonds to the

V. hederifolia Trypsin Inhibitor

backbone amides of Ser¹⁹⁵ (2.8 Å) and Gly¹⁹³ (3.0 Å), the oxy-anion hole.

The $\alpha 2$ helix is primarily located away from the enzyme surface and hence makes only limited contacts with trypsin. Because the hairpin arrangement of the helices leads to the $\alpha 2$ helix projecting sharply away from the enzyme surface after the active site, these contacts are limited to extended side chains in the N-terminal end of $\alpha 2$. The Leu²¹ side chain inserts into a depression in the trypsin surface in the vicinity of the S1' pocket. The carboxylate of Glu²⁰ forms a hydrogen bond (2.7 Å) with Tyr³⁹ from the enzyme and is ~ 3 Å from Lys⁶⁰. There is a single symmetry-related contact between the backbone amide of Glu²⁰ of the $\alpha 2$ helix and the backbone carbonyl group of trypsin Gly¹³³ from a neighboring molecule. These limited direct contacts imply that the placement of this helix is likely to be dictated primarily by its interactions with $\alpha 1$ and in particular by the pair of disulfide bonds. Despite its surface location, temperature factors suggest the $\alpha 2$ helix to be only slightly more mobile than $\alpha 1$ (average main chain temperature factors, 78 and 87 Å², respectively). Overall, these observations are consistent with the two helices forming a tightly associated bundle that is expected to provide overall rigidity to the inhibitor.

Synthetic VhTI Complexed with Trypsin—The synthetic VhTI peptide interacts with trypsin in a very similar way to the native peptide (Fig. 4*b*). It makes additional interactions with trypsin via residues 16 and 17 (which are disordered in the native complex) and residues 18 and 19, which lie in different conformations in the two structures. In the synthetic peptide structure (Fig. 4*d*), the side chain of His¹⁶ points away from trypsin and instead forms water-mediated hydrogen bonds to the peptide Gln¹⁴ side chain and Cys¹¹ backbone carbonyl, an interaction expected to increase the rigidity of the helix-turn-helix motif. The Ser¹⁷ side chain lies at the entrance to the S2' pocket and forms a direct hydrogen bond with the carbonyl oxygen of Phe⁴¹ of trypsin in addition to a number of water-mediated contacts. The peptide chain then turns away from the enzyme surface: Ser¹⁸ and Pro¹⁹ form part of the turn that bridges the two helices. The Ser¹⁸ side chain lies in the region of the S1' pocket, where it forms a hydrogen bond with Tyr³⁹ of trypsin. Side chains of residues from the $\alpha 2$ helix form a number of weak hydrogen bonds with a neighboring symmetry-related molecule. These include peptide residues Lys³¹ (3.1 Å), Asn²⁸ (3.1 Å), Asp²⁷ (2.9 Å), and a bifurcated interaction of Arg²⁴ (2.8/3.0 Å).

Overlaying the two complexes (by aligning all C α atoms of trypsin; Fig. 4*b*) reveals that the positioning of helix $\alpha 1$ is well maintained between the two structures. However, a significant displacement of helix $\alpha 2$ is evident. At the C α positions, this displacement varies between 1 and 2 Å and is generally greater at the N-terminal end of the helix (*i.e.* the end closest to the active site). This displacement appears to be most directly associated with cleavage of the connecting loop, implying that the loop creates strain within the peptide that is released on hydrolysis of the position 15–16 peptide bond. Although there are some differences in crystal contacts in this region between the two structures, the positioning of $\alpha 2$ in the native peptide complex prevents inclusion of residues 16 and 17 in this model to form an uncleaved peptide. This implies that the altered crystal

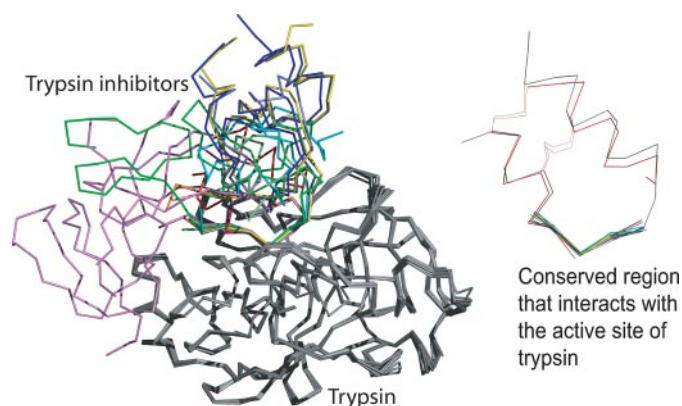


FIGURE 5. Structural comparison of VhTI and other protease inhibitors. Structural alignment of a selection of trypsin-peptide inhibitor complexes highlights the differences observed in inhibitor structure. However, there is also a conserved region at the site of direct interaction with trypsin (*inset*). The native VhTI structure is shown in red, synthetic VhTI is in black, 1AN1 is in cyan, 1BZX is in blue, 1C9P is in green, 1F2S is in light green, 1OX1 is in gray, 1SBW is in salmon, 1TX6 is in violet, 1TAW is in light blue, 1TFX is in yellow, and 1SFI is in orange.

contacts are most likely consequential of movement of the helix rather than causal.

DISCUSSION

Proteases are powerful enzymes, and it is crucial that their activity is tightly controlled. Protease inhibitors are found in many organisms, including plants where they are typically found within the storage tissues such as the seeds, and are thought to protect the plant from pests and pathogens and hence aid in survival of the species. By screening seeds from the clade Asteridae family, we have discovered a novel inhibitor of trypsin in *V. hederifolia* (the ivyleaf speedwell). Characterization of the inhibitor revealed that it comprises a peptide of 34 residues in length, although the truncated, synthetic version demonstrates that the inhibitory motif is entirely located within a 27-amino acid segment.

Analysis of the VhTI peptide sequence by a Blast search (SIB Blast2.0 (31)) failed to identify any significant sequence similarity with other proteins. In addition, comparative analysis of the peptide three-dimensional structure using the Dali server (32) did not reveal any significant structurally similar proteins, although the relatively small size of the peptide may limit the validity of this analysis. Structures of 22 trypsin-peptide inhibitor complexes are present in the Protein Data Bank representing all of the known inhibitor families, and these complexes were examined to ascertain whether any were similar to VhTI. Superimposition of the VhTI complex with a member of each different family (12 structures; Fig. 5) shows that the former displays a distinctive structure, clearly different from any of the other inhibitors. Most other known inhibitors form all- β structures, and the majority are much larger in size than VhTI. A notable feature of this structural alignment is the large variation observed between the overall structures of these individual inhibitor families. Despite these differences, however, the overlay (Fig. 5, *inset*) also highlights the strong conservation of the peptide conformation within the region of the enzyme active site. Although VhTI has an overall structure very different from the other inhibitors, residues 13–15 (found at the end of the

first α helix) lie in a very similar conformation to those seen with other inhibitors. This is surprising because this region is presented at the end of an α helix as a helix-turn-helix motif in VhTI, whereas the equivalent residues from the other inhibitors are located at the ends of β -strands.

A common structural element of protease inhibitors, distinguishing them from peptides that act as substrates, is the incorporation of one or more intramolecular disulfide bonds. These stabilizing bonds lead to retention of the overall peptide structural. Previous attempts to explain the low reactivity of the serine proteases toward these inhibitors have focused on the rigidity of the inhibitor-enzyme complex, suboptimal orientation of reactive groups, or retention of the leaving group $\text{H}_2\text{N-R}_2$ from the acyl-enzyme complex in a position that favors the reverse reaction (reviewed in Ref. 35). A recent study of the chymotrypsin inhibitor CI2 concluded that inhibition arose from retention of the $\text{H}_2\text{N-R}_2$ leaving group in a conformation favoring the religation reaction, proposing that this mechanism is typical for peptidic serine protease inhibitors (35). The surprising observation in the current study that the synthetic form of VhTI forms an uncleaved, noncovalent complex with trypsin, whereas the native peptide is present as an acyl enzyme intermediate in the crystals presents an interesting challenge for these mechanistic explanations. The resolution of both crystal structures is sufficiently high to confirm this distinction. Both crystal structures have been determined at similar pH values (native pH 8.0, synthetic pH 7.5).

A structural comparison of the native and synthetic forms of VhTI (Fig. 4) indicates that, apart from cleavage of the 15–16 peptide bond, the only significant differences are that in the native VhTI-trypsin complex residues 16 and 17 are not visible within the electron density, and the $\alpha 2$ helix moves slightly (1–2 Å) away from the active site. This rearrangement may arise from the release of inherent strain within the peptide concomitant with cleavage of the connecting loop. We assume that these movements are associated with the fundamental difference between these complexes: stabilization of the acyl-enzyme intermediate only in the native peptide complex.

Retention of the covalent intermediate within the native peptide complex crystal is unusual and may be induced by stabilization of this intermediate within the constraints of the crystal lattice. In solution, this form is normally rapidly broken down following nucleophilic attack of the acyl bond by an appropriately placed water molecule, hence allowing the acid component to diffuse out of the active site. In the native peptide complex, a water molecule (water 134) is observed at 50% occupancy approximately equidistant between the carbonyl carbon of the acyl-enzyme linkage (2.9 Å) and the Ne2 nitrogen of the catalytic histidine 57 (3.0 Å). This water appears to be in a suitable position for nucleophilic attack of the ester bond. An overlay of this structure with high resolution structures depicting acyl-enzyme intermediates formed between elastase and a heptapeptide (36, 37) shows that water 134 in the VhTI complex is located in a similar position to related water molecules in these structures (based on an alignment of the active site $\text{C}\alpha$ atoms, the two water molecules are 0.8 Å apart). Nonetheless, hydrolysis of the ester bond does not appear to take place.

Another explanation for stabilization of the native VhTI-trypsin covalent complex is that the geometry of the bound inhibitor may be unfavorable for hydrolysis of the ester bond. A comparison of the active site/inhibitor geometry for all known structures of serine protease inhibitors bound to uncleaved inhibitors has noted tight clustering of the attack angles for the serine O- γ to the carbonyl carbon of the scissile bond, typically $89 \pm 7^\circ$ with a distance of 2.7 ± 0.2 Å (35). This consensus geometry is maintained within the synthetic VhTI-trypsin non-covalent complex in which the distance and angle are 2.6 Å and 93° , respectively. However, in the covalent complex (bond length, now 1.6 Å) formed by the native peptide, the serine O- γ –C=O angle within the ester bond is 116° , and the attack angle of the bound water (2.9 Å from the carbon) is 100° . This implies that hydrolysis of the ester is therefore likely to be a relatively slow reaction, perhaps contributing to the stability of the acyl-enzyme intermediate within the crystals. It is possible that this alteration to the local geometry arises from the $\alpha 2$ conformational change associated with cleavage of the peptide. Stabilization of the ester could therefore arise from geometric restraints imposed by the retained inhibitor that disfavor the forward hydrolysis reaction, combined with spatial placement of residue 16 in a conformation too remote to allow the bond to reform. Because the observed electron density for residues 16 and 17 is very poor, it is also possible that they are not present in the crystal (see below).

A surprising observation is that formation of this acyl-enzyme intermediate is not replicated with the synthetic form of VhTI, which forms a complete and uncleaved peptide within the active site. It is unclear whether the structure of the synthetic peptide complex represents a religated peptide or whether cleavage of this peptide has never occurred. One potentially important difference is that the native peptide has been pre-exposed to trypsin during the purification process and possibly also to a range of endogenous seed proteases, either within the seed itself or during purification. We speculated that the purified native sample may comprise at least a proportion of ready-cleaved peptide (partially evidenced by the sequencing data) that may be selectively bound within the crystals. We therefore incubated two samples of synthetic peptide with trypsin for, respectively, 24 h and 2 months prior to their co-crystallization. However, in both cases the structure of these crystals also confirmed that uncleaved peptide was bound (data not shown). One further possibility is that within the seed further processing may take place, perhaps by other proteases. Because there is little electron density in the native VhTI-trypsin complex for peptide residues 16 and 17, it is unclear whether these amino acids are simply disordered in the crystal or possibly even absent. In the absence of residue 16, religation of the peptide bond is obviously not possible. Even if present, the associated movement of the $\alpha 2$ helix appears to prevent placement of residue 16 sufficiently within the active site for religation to take place. The paucity of crystal contacts with this helix in the native peptide complex suggests this region is free to move within the crystal lattice, whereas the uncleaved synthetic peptide is more restricted by both the presence of the scissile peptide bond and additional crystal contacts formed with $\alpha 2$ helix residues. If the bound form of the native peptide does indeed

comprise a precleaved and trimmed form of the peptide, this implies that the acyl-enzyme intermediate observed in the crystal is formed by the enzymatic reaction running partially backwards after initial binding of product. This is consistent with cyclic peptide inhibitors retaining significant binding to the active site post-cleavage. Intriguingly, if this is the case, VhTI appears to form a very effective inhibitor in both its cleaved and uncleaved forms.

These perplexing results highlight the delicate balance of an enzyme reaction cycle that can be achieved within a crystal lattice, even leading to long-lived stabilization of intermediate states such as the acyl-enzyme form. We suggest that the strained peptide chains presented by cyclic inhibitors such as VhTI, in which the cleaved and uncleaved forms retain conformations with very similar affinities for the active site, might provide valuable tools for examining enzyme mechanistic steps.

Although at this stage we have not explored the selectivity of VhTI for the broader family of serine proteases, the very close association between inhibitor and enzyme in the S1 pocket and its vicinity strongly suggests that effectiveness of the inhibitor, in its present form, will be limited to trypsin and its close homologues. VhTI is significantly smaller than most proteinaceous protease inhibitors and, as demonstrated in this study, is readily amenable to synthetic synthesis. VhTI is larger than SFTI (14 residues) and similar in size to the knottin family of cyclized peptidic inhibitors (26–48 residues), but difficulties in folding and cyclizing synthetic forms of these peptides have been encountered. In contrast, the simple helix-turn-helix motif described here for VhTI has proven straightforward for linear synthesis and folds spontaneously. As such, VhTI may present a novel scaffold for the design of new types of protease inhibitors.

Acknowledgments—We thank the staff of the Daresbury SRS for access to synchrotron data collection facilities, Mike Weldon at University of Cambridge Protein and Nucleic Acid Chemistry Facility for N-terminal sequencing, and John Andralojc and Tony Clarke for advice on enzyme kinetics.

REFERENCES

- Shewry, P. R., and Lucas, J. A. (1997) *Adv. Bot. Res.* **26**, 135–192
- Gustafson, K. R., Sowder, R. C., Henderson, L. E., Parsons, I. C., Kashman, Y., Cardellina, J. H., McMahon, J. B., Buckheit, R. W., Pannell, L. K., and Boyd, M. R. (1994) *J. Am. Chem. Soc.* **116**, 9337–9338
- Tam, J. P., Lu, Y.-A., Yang, J.-L., and Chiu, K.-W. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96**, 8913–8918
- Daly, N. L., Love, S., Alewood, P. F., and Craik, D. J. (1999) *Biochemistry* **38**, 10606–10614
- Wetherup, K. M., Bogusky, M. J., Anderson, P. S., Ramjit, H., Ransom, R. W., Wood, T., and Sardana, M. (1994) *J. Nat. Prod.* **57**, 1619–1625
- Shewry, P. R. (1999) *Seed Proteins* (Shewry, P. R., and Casey, R., eds) Kluwer Academic Publishers, Norwell, MA
- Sweet, R. M., Wright, H. T., Janin, J., Chothia, C. H., and Blow, D. M. (1974) *Biochemistry* **13**, 4212–4228
- Birk, Y. (2003) *Plant Protease Inhibitors: Significance in Nutrition, Plant Protection, Cancer Prevention and Genetic Engineering*, pp. 23–45, Springer-Verlag, Berlin
- McBride, J. D., Watson, E. M., Brauer, A. B. E., Jaulent, A. M., and Leatherbarrow, R. J. (2002) *Biopolymers (Peptide Science)* **66**, 79–92
- Felizmenio-Quimio, M. E., Daly, N. L., and Craik, D. J. (2001) *J. Biol. Chem.* **276**, 22875–22882
- Craik, D. J., Daly, N. L., Bond, T., and Waine, C. (1999) *J. Mol. Biol.* **294**, 1327–1336
- Craik, D. J., Daly, N. L., Mulvenna, J., Plan, M. R., and Trabi, M. (2004) *Curr. Prot. Pep. Sci.* **5**, 297–315
- Simonsen, S. M., Sando, L., Ireland, D. C., Colgrave, M. L., Bharathi, R., Göransson, U., and Craik, D. J. (2005) *Plant Cell* **17**, 3176–3189
- Lockett, S., Santiago Garcia, R., Barker, J. J., Konarev, A. V., Shewry, P. R., Clarke, A. R., and Brady, R. L. (1999) *J. Mol. Biol.* **290**, 525–533
- Mulvenna, J. P., Foley, F. M., and Craik, D. J. (2005) *J. Biol. Chem.* **280**, 32245–32253
- Konarev, A. V., Anisimova, I. N., Gavrilova, V. A., Vachrusheva, T. E., Konechnaya, G. Y., Lewis, M., and Shewry, P. R. (2002) *Phytochemistry* **59**, 279–291
- Bode, W., Turk, D., and Karshikov, A. (1992) *Protein Sci.* **1**, 426–471
- Konarev, A. V., Griffin, J., Konechnaya, G. Y., and Shewry, P. R. (2004) *Phytochemistry* **65**, 3003–3020
- Konarev, A. V. (1986) *Biochemistry (Moscow)* **51**, 195–201
- Erlanger, B. F., Kokowsky, N., and Cohen, W. (1961) *Arch. Biochem. Biophys.* **95**, 271–278
- Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276**, 307–326
- (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50**, 760–763
- Navaza, J. (1994) *Acta Crystallogr. Sect. A* **50**, 157–163
- Emsley, P., and Cowtan, K. (2004) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60**, 2126–2132
- Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **53**, 240–255
- Perrakis, A., Harkiolaki, M., Wilson, K. S., and Lamzin, V. S. (2001) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **57**, 1445–1450
- McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C., and Read, R. J. (2005) *Acta Crystallogr. D Biol. Crystallogr.* **61**, 458–464
- Schechter, I., and Berger, A. (1967) *Biochem. Biophys. Res. Commun.* **27**, 157–162
- Laskowski, M. J. (1986) in *Nutritional and Toxicological Significance of Enzyme Inhibitors in Foods* (Friedman, M., ed) Plenum, New York
- Lovell, S. C., Davis, I. W., Arendall, W. B., III, de Bakker, P. I. W., Word, J. M., Prisant, M. G., Richardson, J. S., and Richardson, D. C. (2003) *Proteins Struct. Funct. Genet.* **50**, 437–450
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. L. (1990) *J. Mol. Biol.* **215**, 403–410
- Holm, L., and Sander, C. (1998) *Nucleic Acids Res.* **26**, 316–319
- Laskowski, M. J., and Kato, I. (1980) *Annu. Rev. Biochem.* **49**, 593–626
- Otlewski, J., Jelen, F., Zakrzewska, M., and Oleksy, A. (2005) *EMBO J.* **24**, 1303–1310
- Radisky, E. S., and Koshland, D. E., Jr. (2002) *Proc. Nat. Acad. Sci. U. S. A.* **99**, 10316–10321
- Wilmouth, R. C., Edman, K., Neutze, R., Wright, P. A., Clifton, I. J., Schneider, T. R., Schofield, C. J., and Hajdu, J. (2001) *Nat. Struct. Biol.* **8**, 689–694
- Katona, G., Wilmouth, R. C., Wright, P. A., Berglund, G. I., Hajdu, J., Neutze, R., and Schofield, C. J. (2002) *J. Biol. Chem.* **277**, 21962–21970