Structural basis of the resistance of an insect carboxypeptidase to plant protease inhibitors

Alex Bayés*, Mireia Comellas-Bigler†, Monica Rodríguez de la Vega*, Klaus Maskos†, Wolfram Bode†, Francesc X. Aviles*, Maarten A. Jongsma‡, Jules Beekwilder‡§, and Josep Vendrell*§

*Departament de Bioquímica i Biologia Molecular, Facultat de Ciències, and Institut de Biotecnologia i de Biomedicina, Universitat Autònoma de Barcelona, E-08193 Bellaterra, Spain; †Proteinase Research Group, Max Planck Institute of Biochemistry, D-82152 Martinsried, Germany; and ‡Plant Research International, Postbus 16, 6700, Wageningen, The Netherlands

Edited by Clarence A. Ryan, Jr., Washington State University, Pullman, WA, and approved September 26, 2005 (received for review June 29, 2005)

Corn earworm (*Helicoverpa zea*), also called tomato fruitworm, is a common pest of many Solanaceous plants. This insect is known to adapt to the ingestion of plant serine protease inhibitors by using digestive proteases that are insensitive to inhibition. We have now identified a B-type carboxypeptidase of *H. zea* (CPBHz) insensitive to potato carboxypeptidase inhibitor (PCI) in corn earworm. To elucidate the structural features leading to the adaptation of the insect enzyme, the crystal structure of the recombinant CPBHz protein was determined by x-ray diffraction. CPBHz is a member of the A/B subfamily of metallocarboxypeptidases, which displays the characteristic metallocarboxypeptidase α/β-hydrolase fold, and does not differ essentially from the previously described *Helicoverpa armigera* CPA, which is very sensitive to PCI. The data provide structural insight into several functional properties of CPBHz. The high selectivity shown by CPBHz for C-terminal lysine residues is due to residue changes in the S1′ substrate specificity pocket that render it unable to accommodate the side chain of an arginine. The insensitivity of CPBHz to plant inhibitors is explained by the exceptional positioning of two of the main regions that stabilize other carboxypeptidase–PCI complexes, the β8-α9 loop, and α7 together with the α7-α8 loop. The rearrangement of these two regions leads to a displacement of the active-site entrance that impairs the proper interaction with PCI. This report explains a crystal structure of an insect protease and its adaptation to defensive plant protease inhibitors.

crystal structure | inhibitor resistance | *Helicoverpa zea* | metallocarboxypeptidase

The larvae of plant-eating insects have a high need for free amino acids and nitrogen during their development. Normally, they meet these requirements by degrading dietary proteins. The digestive enzymes present in the larvae of the lepidopteran insects such as *Helicoverpa zea* are closely related to those secreted from the mammalian pancreas, comprising serine endopeptidases like trypsins and chymotrypsins and exopeptidases such as metallocarboxypeptidases and aminopeptidases (1, 2).

Plants have evolved a defense that interferes with the insects' digestive system by expressing a number of different protease inhibitors (PIs), which are either constitutively present or strongly up-regulated in response to wound damage (3). Most PIs are small proteins that have been found in all plant species investigated thus far and occur in both reproductive and vegetative tissues. In herbivorous insects, they act by inhibiting protein digestive enzymes in the guts of insect larvae or adults, resulting in amino acid deficiencies that lead to serious developmental delay, mortality, or reduced fecundity.

A number of polyphagous insects, among them *H. zea*, a common pest of many Solanaceous plants such as the potato (4), have adapted to the protease inhibitors of their various host plants. Their survival and larval development is not affected by the presence of such molecules in the diet, because of their ability to express specific inhibitor-insensitive digestive proteolytic enzymes depending on the specific inhibitor repertoire encountered in the host plant which they colonize (5, 6). Analyses of trypsin-like enzymes have shown that PI-sensitive and PI-adapted trypsins differ only marginally in amino acid sequence and substrate specificity but exhibit huge differences in sensitivity against PIs (7). However, in the absence of crystal structures, there has been no conclusive mechanistic, structure-based explanation for the apparent insensitivity of PI-adapted insect proteases.

Many studies have focused on the adaptation of insect endoproteases to protease inhibitors, but despite the recent identification of metallocarboxypeptidases (MCPs) in the gut of several lepidopteran species (8–11), very little is known about the effect of the presence of dietary inhibitors on insect carboxypeptidase activity. Similarly to their mammalian counterparts, insect carboxypeptidases (CPs) are members of the A/B subfamily of MCPs, formerly known as "pancreatic-like." MCPs contain an active center Zn^{2+} atom essential for catalysis and are synthesized as proenzymes, which are activated by trypsin-like enzymes through the release of a 90- to 95-residue N-terminal polypeptide. In mammals, two main forms of pancreatic-like CPs, CPA and CPB, acting respectively on peptide substrates with C-terminal aromatic/aliphatic residues and with C-terminal basic residues, are involved in the degradation of dietary proteins. The cDNA encoding for the CPA enzyme activity in the gut lumen of *H. armigera* has been cloned and expressed (8), and its 3D structure has been determined (12).

Carboxypeptidases are also targets of plant-produced antifeedant molecules, such as the potato carboxypeptidase inhibitor (PCI), which is constitutively expressed in the tubers and induced in leaves upon wounding (13). So far, all mammalian members of the A/B subfamily of CPs appear to be highly susceptible to inhibition by PCI, as is CPA from *H. armigera* (14).

In the present work, we describe the 3D structure of a PCI-insensitive carboxypeptidase B purified from larvae of corn earworm (*H. zea*). Cloning, heterologous expression, and purification of the proenzyme form of B-type metallocarboxypeptidase of *H. zea* (PCPBHz) allowed the production of crystals of the active form amenable to x-ray diffraction studies. The structural information explains the substrate

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: CP, carboxypeptidase; CPBHz, B-type CP of *H. zea*; CPBh, B-type carboxypeptidase of human; CPAHa, A-type CP of *H. armigera*; i, inhibitor; LCI, leech carboxypeptidase inhibitor; MCP, metallocarboxypeptidase; PCI, potato carboxypeptidase inhibitor; PCPBHz, B-type metalloprocarboxypeptidase of *H. zea*; PI, protein inhibitor.

Data deposition: The atomic coordinates of CPBHz have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 2C1C). The sequence reported in this paper has been deposited in the GenBank database (accession no. AM085495).

§To whom correspondence may be addressed. E-mail: jules.beekwilder@wur.nl or josep.vendrell@uab.es.

selectivity of the enzyme and unveils the determinants of its resistance to inhibition by plant proteinaceous inhibitors.

Materials and Methods

Isolation, Recombinant Expression, and Purification of PCPBHz. A preliminary purification of B-type carboxypeptidase of *H. zea* (CPBHz) from the gut content of *H. zea* larvae allowed its cloning and recombinant expression in the *Pichia pastoris* system. The cDNA coding for the complete zymogen (PCPBHz) was subcloned into the expression vector pPIC9 and used to transform the KM 71 strain of *P. pastoris* by using the spheroplasts method as described for other procarboxypeptidases (14, 15). The protein was subsequently purified from the supernatant by a first step of hydrophobic interaction chromatography in a butyl-Toyopearl 650M column and a second step of anion-exchange on a Sep-Pack Q column, applying a gradient from 100% buffer A (20 mM Tris, pH 8.0) to 100% buffer B (20 mM Tris/NaCl 1M, pH 8.0) in 30 min. A final gel filtration on a Sephacryl S200 column (20 mM Tris/0.2 M NaCl, pH 8.0) was performed to achieve maximal purity.

Enzymatic Measurements. All substrates used were dissolved in 20 mM Tris/0.1 M NaCl, pH 7.5, and the activity assays were carried out at 25°C. To measure the kinetic constants for CPBHz and B-type carboxypeptidase of human (CPBh), substrates benzoyl-glycyl-arginine and benzoyl-glycyl-lysine were dissolved at a final concentration of 1 mM, and the activity was measured by the increase in absorbance at 254 nm; substrate methoxyphenylazoformyl-Arg was dissolved at a 0.1 mM concentration after the decrease in absorbance at 350 nm; the *N*-(4-furylacryloyl)-Ala-Lys (FAAK) substrate was dissolved at a concentration of 0.2 mM, and the decrease in absorbance at 330 nm was followed. To calculate the equilibrium dissociation constants (K_i) for protein inhibitors, the substrates *N*-[3-(2-furyl)acryloyl]-Phe-Phe and FAAK were used at a final concentration of 0.2 mM to measure CPA and CPB activities, respectively, after the decrease in absorbance at 330 nm. The method described for reversible tight-binding inhibitors that uses presteady-state kinetic analysis (16) was used to calculate K_i values. The enzyme concentration was left constant at 0.8 nM, and increasing amounts of inhibitor were added. At each point, the activity (v_i) was measured against the substrate FAAK. The activity of CPBHz in the absence of inhibitor is defined as v_o, and the parameter a is defined as v_i/v_o. By plotting $[I]/(1 - a)$ against $1/a$, a line is obtained that follows the equation: $[I] = [E](1 - a) + K_{iapp}(1 - a)/a$. To correct for the effect of the substrate on the formation of the complex EI, the following equation is applied: $K_i = K_{iapp}/(1 + [S]/K_m)$, resulting in the final K_i value.

Crystallization and Structure Determination. Purified protein at 5 mg/ml dissolved in 0.2 M NaCl/20 mM Tris·HCl, pH 8.0 was mixed at a 1:1 ratio with the crystallization reservoir solutions. Crystallization occurred upon concentration against the reservoir solution at 293K by using a sitting drop vapor diffusion setup. For data collection, crystals were soaked in a cryobuffer containing 25% (vol/vol) glycerol before freezing in a nitrogen stream at 100K (Oxford Cryosystems Cryostream). Data were collected on a MAR345 imaging-plate detector (MAR Research, Norderstedt, Germany) at the DESY BW6 beamline (Deutsches Elektronen Synchroton, Hamburg, Germany) at a wavelength of 1.050 Å. Indexing and integration of the diffraction data were carried out with MOSFLM 6.11 (17). The data were merged, scaled, and truncated by using programs supported by the CCP4 suite (18). The structure was solved by molecular replacement with the program suite AMORE (19) and the structure of procarboxypeptidase A of *H. amigera* (12) as a template. The program MAIN (20) was used for model building performed on an Silicon Graphics workstation. Positional and individual temperature factor refinement was carried out with the program CNS (21). In later stages, water molecules and ligands were introduced at stereochemically reasonable positions. X-ray data and refinement statistics are summarized in Table 1.

Table 1. X-ray data and refinement statistics

Measurement	Value
Unit cell dimensions	a = 42.012 Å, b = 58.3395 Å, c = 146.5643 Å, β = 89.8739°
Space group	$P2_1$
Limiting resolution, Å	2.3
Total no. of reflections measured	260,322
Unique reflections (20–2.3 Å)	29,050
R_{merge},* % (20.0–2.3 Å)	7.5
Completeness, % (20.0–2.3 Å)	91.4
Completeness, % (2.4–2.3 Å)	90.2
Solvent molecules	348
Zinc ions	2
Yttrium ions	2
Reflections used for refinement	27,624
Sigma cutoff for refinement	2.0
Resolution range used	20–2.3
R value overall,† %	21.46
R_{free},‡ %	29.2
rms standard deviations	
Bond length, Å	0.0093
Bond angles, °	1.6

*$R_{merge} = [\Sigma_h\Sigma_i |I(h, i) - \langle i(h)\rangle|/\Sigma_h\Sigma_i I(h, i)] \times 100$, where $I(h, i)$ is the intensity value of the ith measurement of h and $\langle I(h)\rangle$ is the corresponding mean value of h for all i measurements of h. The summation is over all measurements.
†R value $= (\Sigma|F_o - F_o|/\Sigma F_o) \times 100$.
‡R_{free} was calculated randomly omitting 5% of the observed reflections from refinement and R factor calculation.

Results and Discussion

CPBHz Is a Carboxypeptidase That Resists Inhibition by Protein Inhibitors. The development and survival of *H. zea* larvae are not affected by the dietary presence of the CP inhibitor from potato, PCI. In these conditions, analysis of the CP activities present in the midgut fraction showed that the CPB activity does not vary in response to PCI, whereas the activities of CPA and CPO, a recently described enzyme that specifically cleaves C-terminal glutamate residues (22), are clearly reduced (A.B., J.V., F.X.A., M.A.J., and J.B., unpublished data). The recombinantly expressed CPB enzyme, referred to as CPBHz, was tested for inhibition by four plant- and animal-derived CP inhibitors and compared with two other CPs. The K_i values obtained are summarized in Table 2. PCI is a well characterized inhibitor from potato (25) with K_i values ≈1 nM for the A/B subfamily of carboxypeptidases (26). However, it shows a remarkably high K_i

Table 2. K_i values of carboxypeptidase protein inhibitors against human CPB, *H. armigera* CPAHa, and *H. zea* CPBHz

Carboxypeptidase	K_i, nM			
	PCI	LCI	TCI	Latexin
CPAHa	0.065 ± 0.73*	0.26 ± 0.032*	ND	NI†
CPBh	1.8 ± 0.32	1.7 ± 0.085‡	1.3 ± 0.2	1.1 ± 0.1†
CPBHz	10,200 ± 250	62.3 ± 3.32	9.7 ± 2.7	NI

ND, not determined; NI, no inhibition observed; TCI, tick carboxypeptidase inhibitor.
*Data taken from ref. 14.
†Data taken from ref. 24.
‡Data taken from ref. 23.

BIOCHEMISTRY

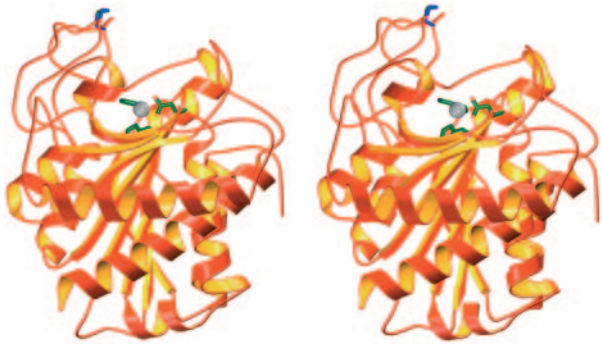

Fig. 1. Stereo ribbon plot representation of CPBHz. The two histidines and the glutamic acid residues coordinating the zinc atom are shown as stick models in green, the zinc atom is shown as a silver sphere, and the single disulfide bridge present in the structure is shown in blue. The polypeptide fold consists of a core of eight twisted β-sheets surrounded by eight α-helices. Figure generated with MOLSCRIPT (31) and RASTER 3D (32).

of 10.2 μM for CPBHz, four orders of magnitude higher than that for the A-type CP of *H. armigera* (CPAHa) (65 pM) (14), an enzyme that belongs to the same subfamily. This high K_i value classifies CPBHz as resistant to PCI under physiological enzyme concentrations (27). The other three inhibitors are of animal origin and are not part of the larval diet. Two of them, the leech carboxypeptidase inhibitor (LCI) (28) and the tick carboxypeptidase inhibitor (TCI) (23, 29), although being structurally different from PCI (30), share with it the same inhibition mechanism consisting of a substrate-like binding of their C-terminal tail into the active site groove of CPs. The K_i values of LCI and TCI for CPBHz are, respectively, 62 nM and 9.7 nM, both of them higher than those for the reference enzyme CPBh (see Table 2) but still indicating that they are probably moderately effective against CPBHz. Hence, CPBHz seems to have evolved a strong and specific resistance to dietary plant inhibitors like PCI, although it is rather susceptible to mechanistically similar inhibitors that are not normally found in the larval diet. The fourth inhibitor, human latexin, does not inhibit CPBHz or CPAHa. Apparently, its molecular architecture and inhibition mechanism, which are completely different from those of the previously discussed inhibitors (24), allow no interaction with the insect enzymes.

Overall Structure of the Carboxypeptidase B from *H. zea*. The carboxypeptidase B from *H. zea* is a 314-residue polypeptide chain that displays a general globular shape. The main fold of mammalian carboxypeptidases, described by a core of eight twisted β-sheets surrounded by eight α-helices (Fig. 1), is also conserved in the structure of the insect enzyme. As shown in the structural alignment (Fig. 2), CPBHz contains the only disulfide bond (Cys-138-Cys-161) fully conserved in CPs during evolution (12, 34, 35) and displays two (Ser-197-Phe-198 and Pro-205-Trp-206) of the three cis-peptide bonds characteristic for mammalian carboxypeptidases. A difference with the conserved structure is the presence of an additional short α-helix (α7).

The structural alignment of CPBHz with its mammalian counterparts required the introduction of several insertions, located at the end of helix α2, β2-β3, α4-α5, α6-β7, α7-α8, and β8-α9 loops connecting secondary structure elements, and a single deletion at the α2-α3 loop (Fig. 2). CPBHz and CPAHa from *H. armigera* are highly homologous (58% identity) in such a way that most loops connecting the secondary structure elements are perfectly superimposable. The most significant differences between both carboxypeptidases are a two-residue deletion in the loop between β8 and α9 and the different conformations adopted by the α7-α8 loop and residues Asn-133-Leu-137 in the α4-α5 loop.

Insensitivity of CPBHz to Natural Protein Inhibitors. The most remarkable functional property of CPBHz is its insensitivity to PCI, a unique feature among members of the A/B subfamily of MCPs. The stability of typical CP–PCI complexes depends on

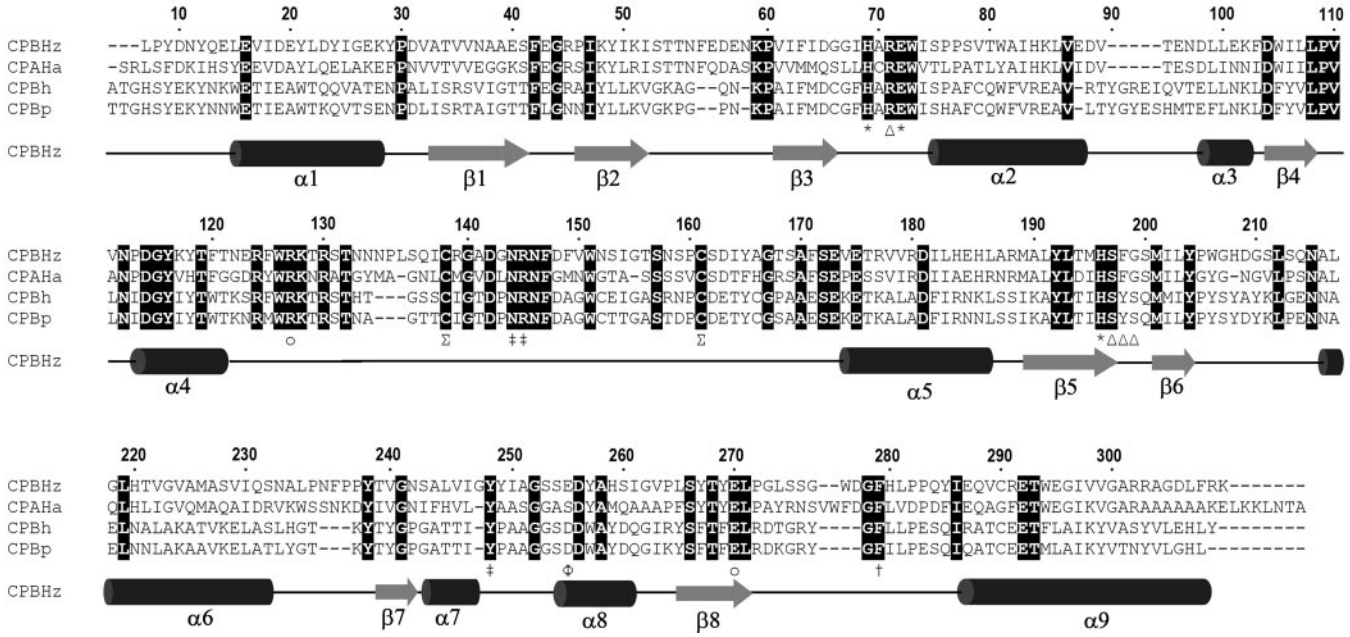

Fig. 2. Structure-based sequence alignment of the carboxypeptidase B from *H. zea* (CPBHz), and with human (CPBh) and porcine (CPBp) B carboxypeptidases and the A form from *H. armigera* (CPAHa). Zinc- binding residues, subsite and substrate specificity-determining residues, and the disulfide bridge are identified. Conserved residues are shown in white with a black background. CPBHz numbering is according to ref. 33. Secondary structures are represented by arrows and cylinders. Symbols correspond with: ∗, zinc-binding residues; ‡, S1′ residues; ○, S1 residues; △, S2 residues; †, S3 residue; Σ, disulfide bridge cysteines; ϕ, substrate specificity residue.

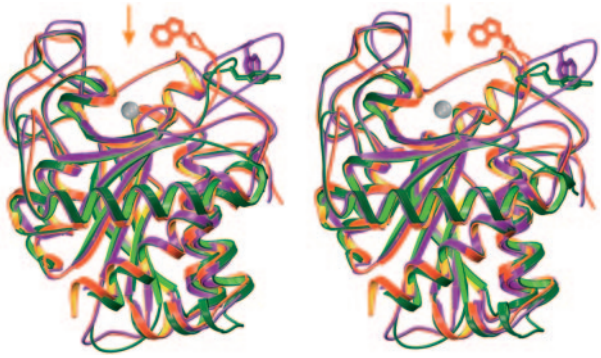

Fig. 3. Stereoview of a structural superimposition of *H. zea* carboxypeptidase B in orange, human carboxypeptidase B in green, and *H. armigera* carboxypeptidase A, in purple. Trp-277A of CPBHz and residues Trp-277B and Tyr-277 in human CPAHa and CPBh, respectively, are represented as stick models; the arrow indicates the entrance to the active-site cleft. The zinc ion is shown in silver. The orientation of the molecule is the same as in Fig. 1.

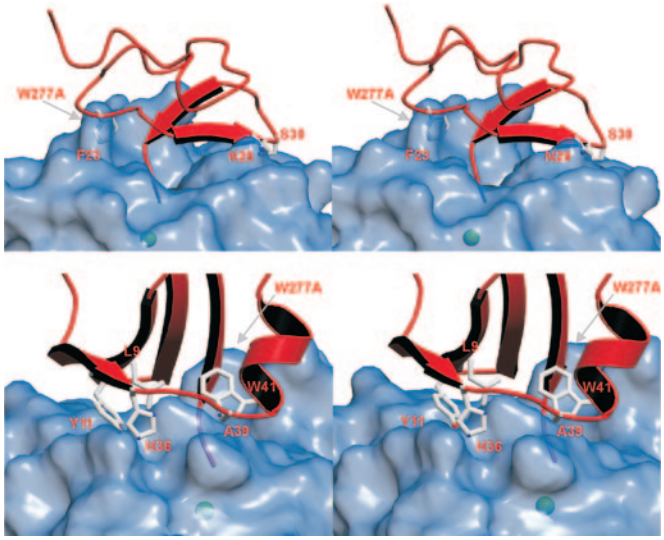

Fig. 4. Structural superimposition of CPBHz with protein inhibitors. (*Upper*) Model complex between the CPBHz and PCI (in red). The CPBHz surface is shown in half-transparent blue. PCI residues, which make contact with the carboxypeptidase, are represented as silver stick models and labeled in red. The zinc ion is represented as a green sphere. (*Lower*) Model complex between CPBHz and LCI. The different elements are organized as in *Upper*. The position of Trp-277A in CPBHz is given as a reference in both parts of the figure to indicate their different orientations relative to Fig. 1.

the interactions at the primary and secondary binding sites. The primary binding site is formed by the C-terminal tail of the inhibitor, which penetrates into the active site of the enzyme where it interacts with subsites S1′, S1, S2, and S3. It has been established that the PCI C-terminal residue Gly39i (i, inhibitor) is cleaved off by the enzyme and the new C-terminal Val38i becomes essential for inhibition (26, 30). The secondary binding site consists of two CP loops (after helix α7 and between β8 and α9) that provide a funnel-shaped access to the primary binding site and interact with complementary surfaces on PCI.

The structural features explaining the insensitive behavior of CPBHz do not directly relate to changes in residues involved in the primary binding site (30). The active site residues differ only at two positions, which do not seem to be relevant to inhibition but rather affect substrate specificity (see below). More pronounced differences occur in the loops forming the secondary binding site. Most significant is the shape adopted by the β8-α9 loop around Trp-277A. The orientation of this loop differs completely from that in other CPs, particularly CPAHa. In CPBHz, this loop is placed toward the active-site cleft and provides a quite exposed Trp-277A side chain directed toward the inner part of the molecule (Fig. 3). The position of this residue, which is not affected by intermolecular crystal contacts, is likely to play a major role in the adaptation of CPBHz to plant carboxypeptidase inhibitors. The second region involved in PCI binding, around the inserted Gly247A (Fig. 2), protrudes toward the active site, compared with other carboxypeptidase structures. Both these changes lead to a change in the shape of the access to the active site of CPBHz, hindering the docking of the C-terminal tail of PCI.

To determine whether these structural features do impair inhibitor binding, the CPBHz structure was superimposed with that of bovine CPA in complex with PCI (30) by using the software MAIN (20) to generate a hypothetical CPBHz–PCI complex. The resulting model (Fig. 4 *Upper*) clearly shows that complex formation is not in accord with the indole moiety of Trp-277A, which occupies the space needed by Phe23i of PCI, a residue that stands for 21% of the interaction area at the secondary binding site of the inhibitor (36). PCI residues Asn29i and Ser30i would also clash with the carbonyl groups of Gly-247A and Val-245, respectively. The structure of the funnel-like access to the active site has a similar general shape and entrance amplitude in both type-A bovine CP (CPAb) and CPBHz. However, the outer rim of the groove is displaced 5–6 Å away from the specificity pocket in CPBHz as compared with CPAb (see supporting information, which is published on the PNAS web site). Therefore, occupancy of the S1′ and S1 subsites in CPBHz by PCI (through the C-terminal residues Gly-39i and Val-38i, respectively) is impossible because of severe steric hindrances. Because of the changes in the secondary binding site, the PCI tail must remain ≈5 Å away from its inhibitory complex position, with, for instance, Val-38i positioned near residues at the S3 subsite, instead of S1 (see supporting information). The lack of contacts at the primary binding site thus results in a destabilization of the complex, which is reflected in a high K_i (10.2 μM). This value corresponds well to the 40 μM K_i shown by a PCI mutant lacking both Val-38i and Gly-39i C-terminal residues when tested with the otherwise PCI-sensitive CPAb (26).

A similar approach was used to understand the moderate K_i (62.3 nM) of LCI against CPBHz (Fig. 4 *Lower*). The superimposition of CPBHz on the complex structure of LCI with human carboxypeptidase A2 (28) showed that the residues between Gly-247A and Tyr-249 would disturb proper complex formation: the imidazole group of His-36i would be quite close to the phenolic side chain of Tyr-249; the phenol group of Tyr11i would clash with the Cβ of Tyr-248; the indole nitrogen of Trp-41i, which establishes a hydrogen bond with the carbonyl group of Ile-247 in CPA2, is too close to that group in our superposition; the carbonyl group of Ala-39i also contacts Gly-247A; and, finally, Leu-9i is quite close to Gly-247A and Tyr-248. Also, of the three hydrogen bonds formed at the secondary binding site of the hCPA2–LCI complex, only the one between the phenol oxygen of Tyr-11i and the amide nitrogen of Tyr-249 could be formed. Interestingly, and in contrast to PCI, the region of Trp-277A in CPBHz (loop β8-α9) would not affect the binding of LCI because there is no interaction with that region of the carboxypeptidase.

The structural basis of the adaptation of insect proteases to plant protease inhibitors has long been a matter of discussion. It has been noted that differences between sensitive and insensitive trypsins are mainly located in the loops around the reactive site (7). Human mast-cell tryptase is one of the rare examples of an inhibitor-resistant protease for which the structural determinants have been solved. Its crystal structure revealed that monomeric units of the protease, which act as a tetramer, recruit

Table 3. Kinetic constants for substrate hydrolysis by CPBHz and CPBh

	CPBHz			CPBh		
Substrate	k_{cat}, S^{-1}	K_M, mM	k_{cat}/K_M ($\times 10^{-3}$), $M^{-1}\,s^{-1}$	k_{cat}, s^{-1}	K_M, mM	k_{cat}/K_M ($\times 10^{3}$), $M^{-1}\,s^{-1}$
Bz-Gly-Arg	NM	NM	NM	230*	0.24*	958*
Bz-Gly-Lys	45.7	0.26	176	170*	90*	1.9*
Methoxyphenylazoformyl-Arg	NM	NM	NM	126	0.06	2,100
Furylacryolyl-Ala-Lys	15.3	0.061	251	14.5	0.1	145

NM, nonmeasurable.
*Data taken from ref. 38.

loops from the neighboring monomers to narrow its substrate binding site, thereby becoming inaccessible to most inhibitors (37). Although insect proteases have been reported to act as dimers (2), there has been no support for a role of dimerization in insensitivity to inhibitors. In the case of CPBHz, the adaptations that are interfering with inhibitor binding also affect loop regions, and they appear to involve redirection of a bulky loop residue (Trp277A) and insertion of a single residue to expand a second loop. These variations involve small changes at the DNA level and have probably been selected during the evolution of polyphagous pest insects that have adapted to the presence of many different inhibitors in their diet.

Active Site and Structural Basis for Substrate Specificity. The overall active-site architecture and the critical residues involved in enzyme catalysis are very similar to other MCPs (Fig. 2). Residue 255 responsible for substrate specificity is, as expected for a CP of the B type, an acidic residue. Interestingly, although all CPBs from the A/B subfamily present in the databases display an aspartate, position 255 is occupied by a glutamate in CPBHz. Also, a serine is found at position 243, where all mammalian CPBs display a glycine; this replacement, also found in the sequence of most other insect CPBs, would reduce the space of the substrate specificity pocket.

Four different substrates ending with a basic residue were used to test and compare the enzymatic activities of CPBHz and CPBh. The kinetic data shown in Table 3 indicate that CPBHz is completely unable to degrade substrates with a C-terminal arginine, although displaying a good catalytic efficiency against substrates ending with a lysine residue. Human CPB is able to hydrolyze all substrates but shows a better efficiency against those ending with arginine. Based on the observed restricted specificity of CPBHz and to predict the interaction with natural substrates, two tripeptides (Trp-Gly-Lys and Trp-Gly-Arg) were modeled into the active site, based on the structure of the complex between PCI and bovine CPA (30). After energy minimization, the model contains all of the interactions described for proper substrate binding (28, 30) (Fig. 5): the negative charge of the carboxyl terminus of the peptide is stabilized by Asn-144, Arg-145, and Tyr-248 (data not shown in the figure); Arg-127 is close enough to the carbonyl group of the scissile peptide bond to polarize it; the guanidinium group of Arg-71 can make a hydrogen bond with the P2 carbonyl oxygen atom; also, the Lys side chain of the substrate can interact with the Glu-255 located at the bottom of the substrate specificity pocket. All of these contacts would produce a rather stable enzyme–substrate complex, which would allow for a proper catalysis, in agreement with the promoted water mechanism for catalysis of carboxypetidases (40). In contrast, the modeling of a peptide containing a C-terminal arginine residue reveals that its bulky guanidyl group would not fit into the space available in the substrate specificity pocket. The presence of Glu-255 makes this pocket slightly smaller, and thereby inadequate, to accommodate an arginine. Fig. 5 shows that the arginine side chain (in dark magenta) would clash with the CPBHz surface. The presence of a serine in position 243 (occupied by a glycine in mammalian CPBs) further narrows this specificity pocket. With a glycine in this position, the lateral chain of the C-terminal Arg residue of a bound substrate might have some space to move away from Glu-255 so that an arginine might be accepted. Thus, the identity of both residues affects CPBHz substrate specificity but seems of no influence on its sensitivity to inhibitors.

The selectivity of CPBHz resembles that of regulatory carboxypeptidases, which, in general, display a high specificity for either arginine or lysine residues. Interestingly, the CPO recently described in *H. armigera* (22) displays a much better cleavage efficiency toward glutamic acid than aspartic acid-containing substrates, which is again atypical for digestive enzymes. In contrast, CPAHa exhibits a very clear activity against basic and acidic residues (14), in addition to typical CPA substrates. This observation suggests that, in *Helicoverpa* species, CPA is a general digestive carboxypeptidase, whereas CPB and CPO have a narrow specificity.

Conclusions

In this work we describe a structure of an insect carboxypeptidase that is resistant to plant protease inhibitors. The structure of CPBHz reveals which structural determinants are important to render it specifically insensitive to the inhibitors that occur in some of the host plants of corn earworm. The fact that inhibitors that originate from leeches or ticks and share the inhibitory mechanism with PCI are much better inhibitors than that of plant origin further suggests that the CP gene has evolved to avoid specifically the effect of the plant molecules. As has been extensively described in refs. 3 and 41, plant protease inhibitors work in the plant defensive

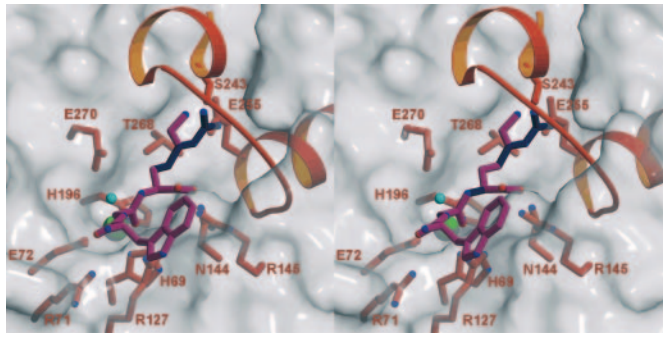

Fig. 5. Stereoview of the active site residues (orange), superimposed on a half-transparent surface (white) calculated for the enzyme lacking the α7 helix and the α7-α8 loop, to give an unrestricted view toward the active site. Two peptides have been modeled and superimposed into the active site representation, with the Trp-Gly-Lys peptide shown in purple and the Arg side chain of Trp-Gly-Arg given in dark blue. The catalytical water molecule is represented as a small blue sphere and the zinc ion as a green sphere. CPBHz residues are labeled in orange. Surfaces were generated with GRASP (39).

machinery against insect attack. Many insect proteases insensitive to plant inhibitors have been described, but a 3D structure has not yet been available for any of them. Therefore, the structure presented here provides a previously uncharacterized model of the way proteases evolve to become insensitive to plant inhibitors and should help us to understand the adaptation strategies of insects to protein inhibitors and provide a basis for the design of new and durable biological methods for pest control.

We thank Francesc Canals, Oscar Conchillo, Joan Lopez Arolas, and Salvador Bartolomé. Human CPB was a generous gift from Roman Bonet (Universitat Autònoma de Barcelona). This work was supported by Spanish Ministry for Science and Technology Grants BIO2001-2046 and BIO2004-05879-C02 and EU projects QLK3-CT-2002-02136 (to W.B. and K.M.) and FAIR6-CT98-4239 (to M.A.J. and J.B.). A.B. acknowledges a fellowship from the Generalitat de Catalunya, a Federation of Biochemical Societies Short-Term Fellowship, and a Boehringer Ingleheim Fonds Travel Allowance.

1. Scriber, J. M. & Feeny, P. (1979) *Ecology* **60,** 829–850.
2. Terra, W. R. & Ferreira, C. (1994) *Comp. Biochem. Physiol., B* **109,** 1–62.
3. Green, T. R. & Ryan, C. A. (1972) *Science* **175,** 776–777.
4. Chakrabarti, S. K., Mandaokar, A. D., Shukla, A., Pattanayak, D., Naik, P. S., Sharma, R. P. & Kumar, P. A. (2000) *Potato Res.* **43,** 143–152.
5. Broadway, R. M. (1995) *J. Insect Physiol.* **41,** 107–116.
6. Jongsma, M. A., Bakker, P. L., Peters, J., Bosch, D. & Stiekema, W. J. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 8041–8045.
7. Volpicella, M., Ceci, L. R., Cordewener, J., America, T., Gallerani, R., Bode, W., Jongsma, M. A. & Beekwilder, J. (2003) *Eur. J. Biochem.* **270,** 10–19.
8. Bown, D. P., Wilkinson, H. S. & Gatehouse, J. A. (1998) *Insect Biochem. Mol. Biol.* **28,** 739–749.
9. Hegedus, D., Baldwin, D., O'Grady, M., Braun, L., Gleddie, S., Sharpe, A., Lydiate, D. & Erlandson, M. (2003) *Arch. Insect Biochem. Physiol.* **53,** 30–47.
10. Johnson, K. S. & Felton, G. W. (2000) *Arch. Insect Biochem. Physiol.* **44,** 151–161.
11. Valaitis, A. P., Augustin, S. & Clancy, K. M. (1999) *Insect Biochem. Mol. Biol.* **29,** 405–415.
12. Estebanez-Perpina, E., Bayes, A., Vendrell, J., Jongsma, M. A., Bown, D. P., Gatehouse, J. A., Huber, R., Bode, W., Aviles, F. X. & Reverter, D. (2001) *J. Mol. Biol.* **313,** 629–638.
13. Graham, J. S. & Ryan, C. A. (1981) *Biochem. Biophys. Res. Commun.* **101,** 1164–1170.
14. Bayes, A., Sonnenschein, A., Daura, X., Vendrell, J. & Aviles, F. X. (2003) *Eur. J. Biochem.* **270,** 3026–3035.
15. Ventura, S., Villegas, V., Sterner, J., Larson, J., Vendrell, J., Hershberger, C. L. & Aviles, F. X. (1999) *J. Biol. Chem.* **274,** 19925–19933.
16. Bieth, J. G. (1995) *Methods Enzymol.* **248,** 59–84.
17. Powell, H. R. (1999) *Acta Crystallogr. D* **55,** 1690–1695.
18. Collaborative Computational Project, Number 4 (1994) *Acta Crystallogr. D* **50,** 760–766.
19. Navaza, J. (1994) *Acta. Crystallogr. A* **50,** 157–163.
20. Turk, D. (1992) Ph.D. thesis (Technische Universität München, Munich).
21. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., *et al.* (1998) *Acta Crystallogr. D.* **54,** 905–921.
22. Bown, D. P. & Gatehouse, J. A. (2004) *Eur. J. Biochem.* **271,** 2000–2011.
23. Arolas, J. L., Lorenzo, J., Rovira, A., Castella, J., Aviles, F. X. & Sommerhoff, C. P. (2005) *J. Biol. Chem.* **280,** 3441–3448.
24. Pallares, I., Bonet, R., Garcia-Castellanos, R., Ventura, S., Aviles, F. X., Vendrell, J. & Gomis-Ruth, F. X. (2005) *Proc. Natl. Acad. Sci. USA* **102,** 3978–3983.
25. Ryan, C. A., Hass, G. M. & Kuhn, R. W. (1974) *J. Biol. Chem.* **249,** 5495–5499.
26. Molina, M. A., Marino, C., Oliva, B., Aviles, F. X. & Querol, E. (1994) *J. Biol. Chem.* **269,** 21467–21472.
27. Bolter, C. & Jongsma, M. A. (1997) *J. Insect Physiol.* **43,** 885–895.
28. Reverter, D., Fernandez-Catalan, C., Baumgartner, R., Pfander, R., Huber, R., Bode, W., Vendrell, J., Holak, T. A. & Aviles, F. X. (2000) *Nat. Struct. Biol.* **7,** 322–328.
29. Arolas, J. L., Popowicz, G. M., Lorenzo, J., Sommerhoff, C. P., Huber, R., Aviles, F. X. & Holak, T. A. (2005) *J. Mol. Biol.* **350,** 489–498.
30. Rees, D. C. & Lipscomb, W. N. (1982) *J. Mol. Biol.* **160,** 475–498.
31. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950.
32. Merritt, E. A. & Bacon, D. J. (1997) *Macromol. Crystallogr. B* **277,** 505–524.
33. Coll, M., Guasch, A., Aviles, F. X. & Huber, R. (1991) *EMBO J.* **10,** 1–9.
34. Gomis-Ruth, F. X., Gomez, M., Bode, W., Huber, R. & Aviles, F. X. (1995) *EMBO J.* **14,** 4387–4394.
35. Guasch, A., Coll, M., Aviles, F. X. & Huber, R. (1992) *J. Mol. Biol.* **224,** 141–157.
36. Arolas, J. L., Lorenzo, J., Rovira, A., Vendrell, J., Aviles, F. X. & Ventura, S. (2004) *Biochemistry* **43,** 7973–7982.
37. Pereira, P. J., Bergner, A., Macedo-Ribeiro, S., Huber, R., Matschiner, G., Fritz, H., Sommerhoff, C. P. & Bode, W. (1998) *Nature* **392,** 306–311.
38. McKay, T. J., Phelan, A. W. & Plummer, T. H., Jr. (1979) *Arch. Biochem. Biophys.* **197,** 487–492.
39. Nicholls, A., Bharadwaj, R. & Honig, B. (1993) *Biophys. J.* **64,** A166.
40. Kim, H. & Lipscomb, W. N. (1990) *Biochemistry* **29,** 5546–5555.
41. Lopes, A. R., Juliano, M. A., Juliano, L. & Terra, W. R. (2004) *Arch. Insect Biochem. Physiol.* **55,** 140–152.