Discrimination of Differentially Inhibited Cysteine Proteases by Activity-Based Profiling Using Cystatin Variants with Tailored Specificities

ARTICLE *in* JOURNAL OF PROTEOME RESEARCH · OCTOBER 2012

Impact Factor: 4.25 · DOI: 10.1021/pr300699n · Source: PubMed

5 AUTHORS, INCLUDING:

Frank Sainsbury
University of Queensland
35 PUBLICATIONS **826** CITATIONS

SEE PROFILE

Marie-Claire Goulet
Laval University
20 PUBLICATIONS **261** CITATIONS

SEE PROFILE

Barend Juan Vorster
University of Pretoria
31 PUBLICATIONS **224** CITATIONS

SEE PROFILE

Journal of proteome research

Article

pubs.acs.org/jpr

Discrimination of Differentially Inhibited Cysteine Proteases by Activity-Based Profiling Using Cystatin Variants with Tailored Specificities

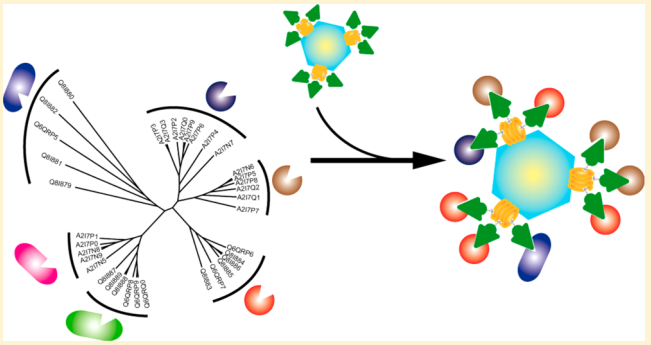
Frank Sainsbury,[†] Ann-Julie Rhéaume,[†] Marie-Claire Goulet,[†] Juan Vorster,[‡] and Dominique Michaud*[,†]

[†]Département de Phytologie, Université Laval, Québec, Canada

[‡]Department of Plant Production and Soil Science, University of Pretoria, Pretoria, South Africa

Supporting Information

ABSTRACT: Recent research has shown the possibility of tailoring the inhibitory specificity of plant cystatins toward cysteine (Cys) proteases by single mutations at positively selected amino acid sites. Here we devised a cystatin activity-based profiling approach to assess the impact of such mutations at the proteome scale using single variants of tomato cystatin SlCYS8 and digestive Cys proteases of the herbivorous insect, Colorado potato beetle, as a model. Biotinylated forms of SlCYS8 and SlCYS8 variants were used to capture susceptible Cys proteases in insect midgut protein extracts by biotin immobilization on avidin-embedded beads. A quantitative LC−MS/MS analysis of the captured proteins was performed to compare the inhibitory profile of different SlCYS8 variants. The approach confirmed the relevance of phylogenetic inferences categorizing the insect digestive Cys proteases into six functionally distinct families. It also revealed significant variation in protease family profiles captured with N-terminal variants of SlCYS8, in line with *in silico* structural models for Cys protease−SlCYS8 interactions suggesting a functional role for the N-terminal region. Our data confirm overall the usefulness of cystatin activity-based protease profiling for the monitoring of Cys protease−inhibitor interactions in complex biological systems. They also illustrate the potential of biotinylated cystatins to identify recombinant cystatin candidates for the inactivation of specific Cys protease targets.



KEYWORDS: *activity-based proteomic profiling, plant cystatins, insect digestive proteases, Cys proteases, Colorado potato beetle intestains*

INTRODUCTION

A large body of literature has been published over the last 25 years discussing the potential of protease inhibitors as pesticidal compounds to protect crop plants from herbivory or pathogenic infection.[1,2] Serine (Ser) protease inhibitors, in particular, have been readily identified as promising candidates for the development of insect-resistant transgenic crops,[3−6] and their usefulness to reduce pesticide loads in the field has been documented.[7,8] Most protease inhibitors in plants are competitive protein inhibitors acting as pseudosubstrates to enter the active site cleft of target proteases.[9] Inhibited enzymes can no longer cleave peptide bonds, thereby causing a sharp decrease of dietary protein hydrolysis in target insects leading to amino acid shortage and detrimental growth delays.[10]

Conclusive cases of insect resistance have been reported for a number of plants expressing recombinant protease inhibitors, but the pesticidal effects of these proteins proved negligible in many other instances despite protease inhibitory effects observed *in vivo*.[2] Insect herbivores have evolved a range of strategies to cope with dietary protease inhibitors, notably involving complex digestive systems with proteases from diverse functional classes and families,[11−14] target protease overexpression to outnumber the inhibitory proteins,[15,16] and the production of protease isoforms weakly sensitive to the inhibitors ingested.[15,17−26] Other adaptations to protease inhibitors include their degradation by nontarget midgut proteases,[21,27−32] the up-regulation of proteases from alternative functional classes,[21,33−35] and a reallocation of metabolic resources toward inhibitor-induced compensatory responses.[24,36] It is now well established that protease−inhibitor interactions in plant−insect systems are the result of a long, coevolutive process triggering the continuous diversification of insect protease and plant protease inhibitor functions.[37−39] In most, if not all, cases the ectopic expression of a protease inhibitor *in planta* not only alters the activity of constitutively expressed target proteases in the midgut of naive herbivores but also induces a significant remodelling of both the midgut transcriptome and digestive protease complement.[22,24,33,36]

Received: July 27, 2012

dx.doi.org/10.1021/pr300699n | *J. Proteome Res.* XXXX, XXX, XXX−XXX

A straightforward way to harness the full potential of protease inhibitors in plant protection may be to use combinations of inhibitors active against different sets of proteases in the target pest, or to use inhibitor variants engineered to exhibit a large spectrum of protease targets, in such a way as to broaden the range of susceptible proteases and overcome the eventual compensatory effects of proteases synthesized *do novo* following inhibitor intake. Fusion proteins integrating complete or partial sequences of Ser and cysteine (Cys) protease inhibitors were shown to exhibit improved potential against target herbivores compared to the inhibitors alone,[40−44] similar to transgene stacking—or "pyramiding"—approaches allowing for the simultaneous expression of two or more inhibitors.[45−47] Protein engineering involving site-directed or random mutations in functionally relevant regions of the native protein has also been used to improve the potency protease inhibitors against herbivorous pests.[39,48−53] Single mutations at positively selected amino acid sites, in particular, were shown to induce functional diversity against different proteases, useful to rapidly generate inhibitor variants with improved or complementary protease inhibitory activities.[39] For instance, single variants of the wound-inducible Cys-type inhibitor tomato cystatin SlCYS8 were engineered using this approach and shown to present a wide range of inhibitory specificities against plant and insect Cys proteases.[53] The challenge now is to develop a rational framework for the effective selection of broad-spectrum or complementary inhibitor variants, taking into account the complex dynamics of protease−inhibitor interactions in plant−insect systems and the striking diversity of midgut protease-encoding sequences in herbivorous insect genomes.[54]

As a first step toward this goal, we devised an activity-based profiling approach to compare the overall potency and inhibitory spectrum of plant cystatins at the proteome scale using single variants of SlCYS8 tailored to inhibit different sets of digestive Cys proteases in the coleopteran insect, Colorado potato beetle.[53] This major pest of potato crops uses proteases from several functional classes to digest dietary proteins, notably including a variety of C1A Cys proteases, the so-called "intestains".[30] These proteases are all sensitive to the small chemical inhibitor *trans*-epoxysuccinyl-L-leucylamido-(4-guanidino)butane (E-64), but differentially sensitive to a variety of plant and animal cystatins considered as possible candidates for plant genetic improvement.[23,30,55−57] Whereas E-64 ingestion is toxic to potato beetle larvae,[58] recombinant cystatins expressed in potato leaves show no negative effects on larval growth or development despite strong protease inhibitory effects *in vivo*.[15,59,60] Here, we produced biotinylated forms of SlCYS8 functional variants to capture, isolate and identify their specific digestive protease targets in crude protein samples of potato beetle larvae. Our data confirm the usefulness of cystatin activity-based protease profiling for the identification of cystatins with broadened or complementary inhibitory activities against herbivorous insect digestive Cys proteases.

EXPERIMENTAL PROCEDURES

Protease Binding Subsites

Intestain amino acid residues involved in protein substrate binding were inferred by homology to the tertiary structure of human cathepsin L in complex with *S*-benzyl-*N*-(biphenyl-4-ylacetyl)-L-cysteinyl-*N*∼5∼-(diamino-methyl)-D-ornithyl-*N*-(2-phenyl-ethyl)-L-tyrosinamide (Protein Data Bank 3BC3) (http://www.wwpdb.org). Cathepsin L binding residues were first identified using the Discovery Studio software, v. 3.0 (Accelrys Software), and the information then transferred to the corresponding positions in intestain primary sequences.

Phylogenetic Analysis

Phylogenetic relationships among intestains were assessed with the Geneious software, v. 5.4.4 (Biomatters), using 36 complete or partial intestain sequences available in the UniProt protein database (Supporting Information Table 1). A multiple sequence alignment was created with the MUSCLE algorithm,[61] from which a maximum likelihood tree was calculated using the PHYML plugin,[62] based on the "JTT" amino acid substitution model.[63]

Protease−Inhibitor Interactions

Protein−protein docking simulations were performed *in silico* for intestains and SlCYS8 variants, after building structure homology models for wild-type SlCYS8 (GenBank Accession No. AF198390) and intestain D4 (IntD4) (GenBank EF154436). Twenty tentative models were built for both SlCYS8 and IntD4 using Modeller, v. 9.7,[64,65] with the NMR structure of oryzacystatin (PDB 1EQK) as a template for SlCYS8, and the crystal structure of human cathepsin L (PDB 1SC8) as a template for IntD4. Stereochemical quality of the models was compared to their template structures with the Procheck program, v.3.5.4 (http://www.ebi.ac.uk/thornton-srv/software/PROCHECK/),[66] and the best models were selected for further analysis. Intestain−cystatin interactions were simulated using the Z-Dock algorithm[67] of Discovery Studio (Accelrys Software) to generate 2000 tentative poses of the resulting complex. The top ranking poses, based on their Z-score, were compared to the solved crystal structure of the Cys protease inhibitor chagasin in complex with human cathepsin L (PDB 2NQD), to confirm the relative binding position and orientation of the interacting proteins in the predicted complexes. Five complexes were chosen and refined through energy minimization using the R-Dock algorithm.[68] The top ranking model was used to model other intestain−cystatin combinations also involving intestain B11 (IntB11) (GenBank AY528229) and P2F, a functional variant of SlCYS8,[39] as described above for the IntD4−SlCYS8 complex. Intestain and SlCYS8 interacting residues at the binding interface were inferred for the different models, as well as total interaction (binding) energies for each enzyme−inhibitor combination.

DNA Cloning and Mutagenesis

Wild-type SlCYS8 and functional variants were expressed in *E. coli* using the glutathione *S*-transferase (GST) gene fusion system (GE Healthcare), with the specifically biotinylated AviTag peptide[69] grafted at the C-terminus for subsequent immobilization on avidin-embedded beads. GoldenGate cloning[70] was adapted to construct seamless protein fusions for purification and labeling. This technique relies on a single type IIs restriction enzyme to cut inside the recognition sites of the donor clones and outside the recognition sites of the acceptor/expression vector. Annotated sequences of oligonucleotide primers used for all cloning steps are given in Supporting Information Table 2. A cloning vector for generating donor clones was created by replacing the ampicillin resistance trait of pUC18 with resistance to streptomycin to create pUC18strep. A PCR fragment for SlCYS8 was generated by the splice-overlap extension method to eliminate an internal *Bsa*I site with terminal extensions containing unique recombi-

nation sites inside of inverted *Bsa*I recognition sites, oriented to cut inside the fragment. The fragment was blunt-end ligated into pUC18strep in the presence of *Sma*I to maintain the specificity of the ligation. Complementary oligonucleotides encoding the AviTag peptide (SGGLNDIFEAQKIEWHE*) similarly flanked by unique recombination sites inside of *Bsa*I recognition sites were annealed and inserted into pUC18strep in the presence of *Sma*I. The pGEX-3X vector for expression of GST fusion proteins (GE Healthcare) was modified to contain the *ccd*B lethal gene within the cloning site flanked by the appropriate recombination sites outside of inverted *Bsa*I recognition sites oriented to cut outside of the *ccd*B gene (pGEX-3X-GG). Recombination reactions between pGEX-3X-GG and donor clones for SlCYS8 and AviTag were driven by the simultaneous application of *Bsa*I and T4-DNA ligase, resulting in GST-SlCYS8-AviTag fusion constructs. Non-recombined pUC18strep and pGEX-3X-GG clones were killed by the presence of ampicillin or the *ccd*B gene, respectively. The SlCYS8 variants P2F, Q47P and T6S were generated by Quickchange mutagenesis (Stratagene) of the donor clone.

Heterologous Expression and Purification

Expression in *E. coli* and affinity purification of the AviTagged *Sl*CYS8 variants were carried out with reduced glutathione-beads and GST as an affinity handle, as described earlier.[53] The GST moiety was removed from the purified cystatins by proteolytic cleavage with human factor X_a (Novagen), following the supplier's instructions.

Protease Inhibitory Activities

Apparent dissociation constants ($K_{d(app)}$) against the model Cys protease papain, and the inhibitory rates of potato beetle cathepsin L-like and cathepsin B-like activities, by the AviTagged *Sl*CYS8 variants were determined with fluorimetric synthetic peptide substrates as described earlier.[53] Papain activity was measured in 50 mM Tris-HCl, pH 7.0, using the peptide *Z*-Phe-Arg-methylcoumarin (MCA) (Sigma-Aldrich). Insect cathepsin B-like and cathepsin L-like activities were measured in 50 mM MES buffer, pH 6.0, containing 10 mM L-cysteine, with the peptides *Z*-Arg-Arg-MCA (Sigma-Aldrich) and *Z*-Phe-Arg-MCA, respectively. Protease activities were monitored using a Fluostar Galaxy fluorimeter (BMG Labtech), with excitation and emission filters of 360 and 450 nm, respectively.

Insect Proteins

Potato beetle proteins for protease profiling were extracted from snap-frozen and powdered second and third instar larvae with three volumes (v/w) of 100 mM citrate phosphate, pH 6.0. The extracts were kept on ice for 10 min, and then clarified by centrifugation at 20 000× *g* for 12 min at 4 °C.

Protease Profiling

The AviTagged cystatins for avidin affinity enrichment were biotinylated during expression in AVB101 *E. coli* cells coexpressing a biotin ligase (Avidity LLC), in the presence of 50 μM D-biotin in the growth medium. Biotinlyated cystatins were applied in excess to an equilibrated TetraLink Avidin resin (Promega), which was then washed twice with 20 volumes of 50 mM Tris, pH 8.0, before storage at 4 °C. The avidin resin was able to support approximately 4 μg of SlCYS8 for each 10 μL of resin. Cys protease capture was performed by incubating 20 μL of cystatin-bound resin with 5.5 mg of insect proteins in 900 μL of 100 mM citrate phosphate, pH 6.0, for 40 min at room temperature. The resin was collected by gentle centrifugation for 2 min at 1000× *g*, and washed three times by resuspension in 900 μL of 100 mM citrate phopshate, pH 6.0, containing 250 mM NaCl and 10 mM L-cysteine (wash buffer), followed by gentle centrifugation for 2 min at 1000× *g*. After the final centrifugation, 10 μL each of wash buffer and SDS-PAGE loading buffer were added to 20 μL of resin. The resulting preparations were submitted to 12% (w/v) SDS-PAGE and Coomassie Blue staining.

Sample Preparation for Mass Spectrometry

Gel slices of Coomassie Blue-stained bands corresponding to the affinity purified proteins were collected for MS analysis. The bands were carefully excised, destained, reduced in 50 μL of 15 mM dithiothreitol, alkylated in 50 μL of 55 mM iodoacetamide, and trypsin-digested for 4.5 h in 25 μL of 100 mM ammonium bicarbonate containing 6 ng/μL of Trypsin Gold (Promega) using a robotic MassPrep Workstation (Waters-Micromass). The resulting peptides were extracted with multiple changes of 50% (v/v) acetonitrile (Acn):1.0% (v/v) formic acid, for a final 60 μL of peptide extract solution at 15% (v/v) Acn:0.5% (v/v) formic acid. Ten microliters of peptide solution was used for LC−MS/MS analysis. Peptide separation was performed using an Easy-nano LC system (Thermo-Scientific).

Mass Spectrometry

Peptide samples for MS analysis were injected onto an Agilent Zorbax 300SC-C18 5 × 0.3 mm column for desalting, and then separated on a NewObjective Biobasic C18 integrafrit column along a 30 min gradient going from 5% to 95% (v/v) Acn:0.1% (v/v) formic acid. Tandem MS/MS spectra were acquired on an LTQ-Velos Orbitrap system (Thermo-Scientific). MS scans in the 350−1600 *m*/*z* range were performed in the Orbitrap at 30 000 resolution. The top 10 multiply charged ions surpassing the 5000 count threshold were selected for collisional induced fragmentation in the Velos-LTQ using a 2 Da isolation window, a normalized collision energy value of 35, an activation window of 0.25 and a 10 ms activation time. Precursor ions within a 10 ppm *m*/*z* tolerance were reselected a second time within a 30 s window from subsequent survey scans before being placed in a dynamic exclusion list for a period of 40 s.

Protein Identification

Tandem MS peaklists were generated with Mascot Distiller, v. 2.3 (Matrix Science), with no deisotoping. The MS/MS samples were analyzed with Mascot 2.3 (Matrix Science)[71] and the Open Source software X! Tandem[72] set up to search a custom database containing all known sequences of Coleoptera (39 285 sequences, as of September 17, 2011), including the fully sequenced genome of *Tribolium castaneum*.[54] X! Tandem searches were carried out on the subset database of Mascot identified proteins within Scaffold and search parameters for protein matching with Mascot were as follows: a fragment ion mass tolerance of 0.8 Da; a parent ion tolerance of 10 ppm; carbamidomethylated Cys residues as fixed modification; and oxidized Met residues as variable modification. Search parameters for X! Tandem were as follows: a fragment ion mass tolerance of 0.8 Da; a parent ion tolerance of 10 ppm; and dehydration of the n-terminus, loss of ammonia at the n-terminus, deamidation of Asn and Gln, methylation of Asp, Glu and the c-terminus, oxidation of Met, and acrylamide adducts on Cys as variable modifications. MS/MS based peptide and protein identifications were validated with Scaffold, v.3.00.07 (Proteome Software, Inc.). Identifications were accepted if they

dx.doi.org/10.1021/pr300699n | J. Proteome Res. XXXX, XXX, XXX−XXX

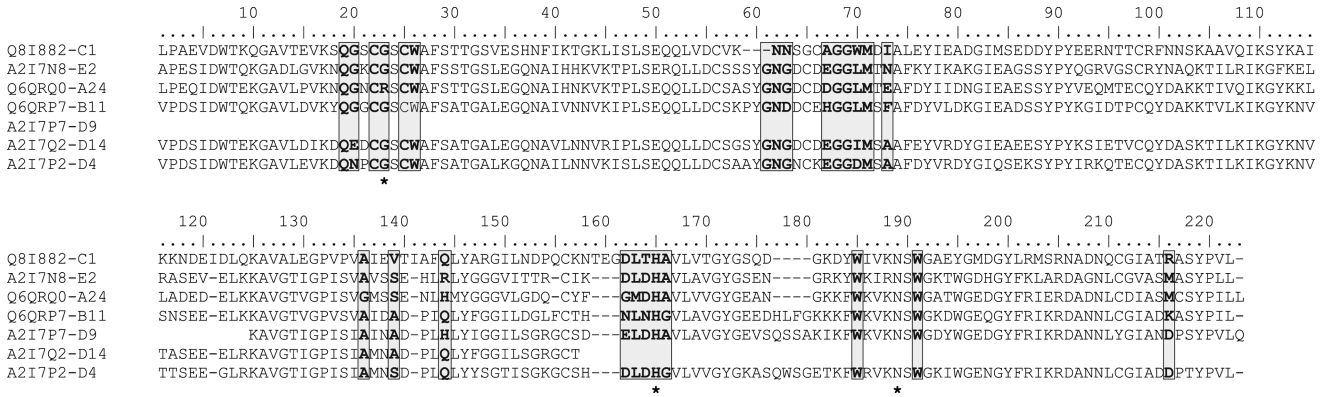


Figure 1. Predicted protein substrate binding subsites in Colorado potato beetle intestains. The binding subsite residues are in bold; asterisks highlight the three residues involved in protein substrate catalysis. Alignment numbering is based on the mature peptide of IntD4. A detailed alignment with all available intestain sequences is provided in Supporting Information Figure 1.

could be established at greater than 95% probability, as specified by the Peptide Prophet algorithm.[73,74]

Spectral Count Analyses

Quantitative analysis of the MS spectra was done using spectral count sampling statistics,[75] on those counts corresponding to peptides that were unique to an intestain subfamily. The spectra assigned to intestains were pooled into evolutive groups according to the intestain phylogenetic tree and predicted functional analysis described above (see Phylogenetic analysis). The spectra obtained from individual bands were combined for each repetition, and only those belonging to peptides unique to a functional group were included in the quantitative analysis. Spectral counts were normalized across repetitions, for which the coefficient of variance was found to be less than 20%. Differential abundances of captured intestains for each SlCYS8 variant were discriminated by pairwise comparison with wild-type SlCYS8, based on a Student's *t*-test with a significance threshold of 5%.

RESULTS

Six Intestain Subfamilies in the Colorado Potato Beetle

Primary structure comparisons were first conducted with currently available intestain sequences to characterize their protein substrate binding residues and phylogenetic relationships for subsequent group assignment purposes. A 3-subfamily grouping was previously proposed for the intestain family, based on an initial analysis of 12 cDNA clones.[23] Thirty-nine distinct cDNA clones are now available in the GenBank database, of which 36 include at least some coding sequence. Most of these sequences are assigned to one of five subfamilies, the intestain A (IntA), IntB, IntC, IntD and IntE subfamilies, each represented by at least one full-length clone encoding proteins of 322−327 amino acids predicted to mature into polypeptides of approximately 24 kDa after removal of the N-terminal proregion. In addition to the five subfamilies, three unclassified homologues, Cpb77, Cpb125 and Cpb160, are found in the database, each representing a unique cDNA clone among the overall complement of intestains. Here we took an *in silico* approach with the 36 intestain-encoding sequences to identify distinguishing primary structure features among the different intestain subfamilies. Intestain substrate binding sites inferred by sequence homology with human cathepsin L revealed a grouping pattern generally matching the current 5-subfamily grouping. A notable exception was found in the IntA and IntD subfamilies, which showed differences among charged residues predicted to be involved in substrate binding (Figure 1, Supporting Information Figure 1).

A phylogenetic analysis of the intestains showed members of the IntA subfamily to group into a single subclade, whereas IntD subfamily members group in two clearly distinct clades suggesting larger sequence heterogeneity (Figure 2). Confirming this, pairwise identity between amino acid sequences was greated than 95% within the IntA, IntB, IntC and IntE subfamilies, compared to 88% only within the IntD family (see 5-group classification in Supporting Information Figure 2). By comparison, identity rates within the two IntD subclades were greater than 95%, compared to 77% between the same two subclades (see 6-group classification in Supporting Information Figure 2). Based on this and the different binding subsite patterns inferred for the two IntD subgroups, we now propose a 6-group (or 6-subfamily) classification for the intestain family, with the introduction of an "IntF" subfamliy including IntD7,

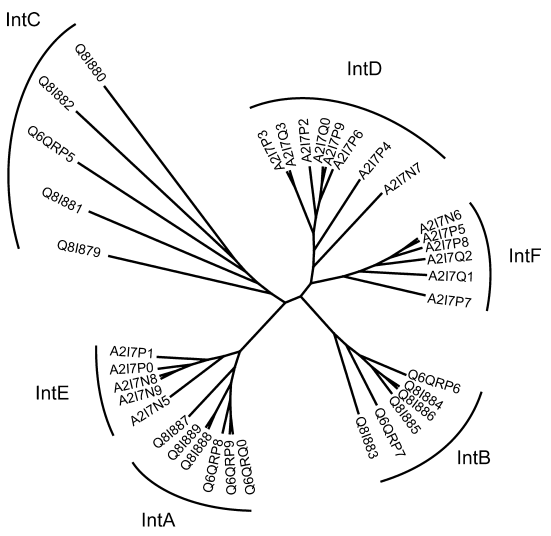


Figure 2. Cladogram representation of Colorado potato beetle intestains phylogeny. An unrooted maximum likelihood tree was built showing the new 6-group classification of the intestain family (for details, see text and Supporting Information Figure 2). Intestains are identified using their UniProt Accession Number (see Supporting Information Table 1). Members of the IntF subfamily were assigned until now to the IntD subfamily.

dx.doi.org/10.1021/pr300699n | J. Proteome Res. XXXX, XXX, XXX−XXX

IntD9, IntD10, IntD13, IntD14 and Cpb125 (Figure 2, Supporting Information Table 1).

Intestain–SlCYS8 Interactions *In Silico*

Protein–protein docking simulations were made to gain some insight on the interactions established between intestains and the model tomato cystatin SlCYS8.[76] An intestain–SlCYS8 complex was first inferred with IntD4 as an example, by homology to the solved structure of human cathepsin L in complex with the parasite inhibitor chagasin.[77] Secondary homology models were then derived from the IntD4–SlCYS8 complex, involving alternative intestains and/or P2F, a Pro-2 to Phe-2 single variant of SlCYS8 known to exhibit stronger activity and a broader inhibitory spectrum against digestive Cys proteases of the potato beetle.[53] A number of amino acid strings on IntD4 and IntB11 were shown to interact with SlCYS8 (Supporting Information Figure 3), closely matching the substrate binding subsites identified above for the insect intestains. In line with the microheterogeneity of intestain substrate binding subsites (Supporting Information Figure 1) and the variable susceptibility of intestain subfamilies to different cystatins,[23,30] calculated interaction [or binding] energies differed for SlCYS8 interacting with different intestains (Table 1). A total energy score of −957 kcal/mol was obtained, for instance, with IntD4, stronger than the scores obtained with IntB11 or with a "consensus" IntF sequence assembled from partial sequences of the new IntF subfamily members. In a similar way, changing the residue Pro-2 for a Phe in SlCYS8 had an intestain-specific impact on total binding energies, generating for instance stabilized complexes with IntD4 and IntF, but having no net effect on the IntB11–SlCYS8 interaction.

Table 1. Total Binding Energies Inferred *In Silico* for IntB11, IntD4 and IntF in Complex with SlCYS8 or its Single Functional Variant P2F

	intestain		
	IntB11	IntD4	IntF*x*[a]
Total interaction energy (kcal/mol)[b]	−726.0	−956.9	−807.8
Weighted mutation energy (SlCYS8 to P2F)[c]	−0.06 (n)	−0.80 (s)	−0.81(s)

[a]IntFx is a "consensus" assembly of partial sequences currently available for the IntF subfamily members (Figure 2). [b]Total binding (or interaction) energies are the sum of energy values inferred for the whole complement of intestatin and cystatin interacting residues (Supporting Information Figure S3). [c]Weighted mutation energies correspond to the total free energy difference between wild-type (SlCYS8) and mutated (P2F) structures, calculated as a weighted sum of the van der Waals, electrostatic, entropy and nonpolar terms. n, neutral effect; s, stabilizing effect.

A closer look at the active site cleft residues of IntB11 and IntD4 interacting with the Pro-2 residue of SlCYS8, or with the alternative Phe-2 residue of P2F, indicated a possible role for hydrogen bond and van der Waals interaction patterns as determinants of intestain–cystatin complex strength (Figure 3). The number of residues interacting with the cystatin, the charge and solvent accessibility of these residues, and their overall spatial orientation in the vicinity of the Pro-2 (or Phe-2) cystatin residue also appeared to vary depending on each intestain–cystatin combination. Cystatin-induced variability of the inferred interactions was expected, given the link

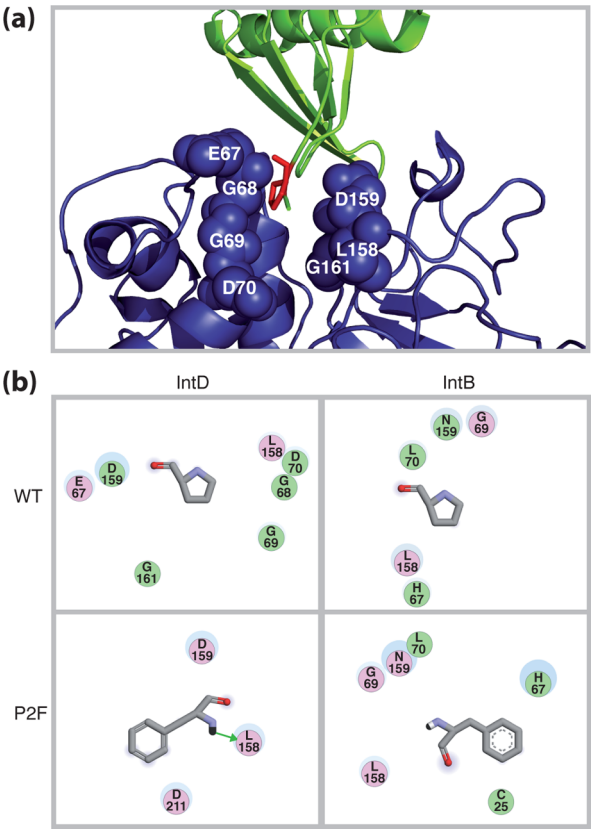


Figure 3. Physicochemical interactions between Colorado potato beetle intestains and the N-terminal region of SlCYS8 variants, as inferred from protein–protein docking with Discovery Studio. (A) General overview of the IntD4–SlCYS8 interaction, showing the N-terminal region and inhibitory loops of SlCYS8 penetrating the active site cleft of IntD4. Amino acids highlighted on IntD4 are those residues identified as interacting with the Pro-2 residue of SlCYS8. (B) Predicted physicochemical interactions between the Pro-2 (or Phe-2) residue of SlCYS8 (or P2F) and the active site cleft residues of intestains D4 and B11 (shown as circles). Pink circles highlight hydrogen bond (specified by the green arrow), charge or polar interactions; green circles highlight van der Waals interactions. Blue halos around the intestain interacting residues represent their solvent accessible surface, proportional to the diameter of the halo.

established earlier between positively selected amino acid sites and the functional diversity of plant cystatins.[39] Intestain-dependent variability was also observed, in line with total binding energy data suggesting a variable susceptibility to cystatin inhibition among intestains (Table 1), such as the IntD and IntB subfamily members, which exhibit sequence heterogeneity at the substrate and inhibitor binding subsites (Supporting Information Figures 1 and 3).

Intestain Capture with Immobilized Biotinylated Cystatins

A cystatin activity-based profiling approach was devised to compare the overall—and presumably divergent—intestain inhibitory spectra of SlCYS8 functional variants at the proteome scale. The technique involved biotinylated cystatin-AviTag peptide fusions[69] as selective affinity ligands for the capture of cystatin-susceptible intestains from potato beetle crude protein extracts. Q47P, a single variant of SlCYS8 lacking a conserved Gln (Q) residue in the first loop consensus motif Q-X-V-X-G,[78] was produced and used as a negative control to optimize the procedure. The AviTag peptide was grafted at the C terminus of the cystatins to prevent interference effects on

dx.doi.org/10.1021/pr300699n | J. Proteome Res. XXXX, XXX, XXX–XXX

functionally active residues of the N-terminal region. Accordingly, an apparent dissociation constant ($K_{d(app)}$) of 4.1 nM was measured for SlCYS8-AviTag inhibiting the model Cys protease papain, similar to the K_d value estimated earlier for untagged SlCYS8.[53] As expected, Q47P showed no activity against papain (not shown), as previously reported for a single mutant of rice cystatin I also lacking the conserved Gln residue.[79]

A major, ~30 kDa protein product was captured with the SlCYS8-AviTag fusion (Figure 4), likely corresponding to the

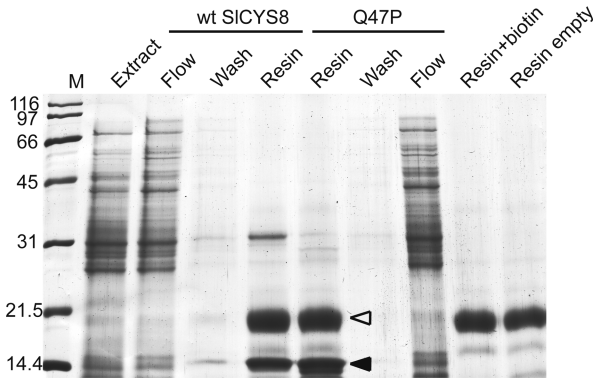


Figure 4. Avidin-affinity chromatography of Colorado potato beetle intestains interacting with a biotinylated SlCYS8-AviTag peptide fusion. Biotinylated AviTagged versions of SlCYS8 and Q47P used as negative control were bound on the avidin resin and then incubated with the insect protein extract for Cys protease capture. Protein fractions from test and control samples were visualized by Coomassie Blue staining, following SDS-PAGE. A major ~30-kDa protein product was captured with SlCYS8, corresponding to the "Ldp30" Cys protease purified earlier with rice cystatin I as an affinity ligand.[80] Open and closed arrowheads point, respectively, to avidin and SlCYS8 recovered from the affinity resin. M refers to commercial molecular weight protein markers (kDa).

proteolytically active "Ldp30" Cys protease of the potato beetle, previously affinity-purified with rice cystatin I as a ligand.[80] A faint, 25-kDa protein signal was also detected, possibly corresponding to a proteolytic fragment of Ldp30 or to an intestain homologue presenting alternative post-translational modifications. As expected, Ldp30 was barely detectable following capture with the Q47P-AviTag fusion, but two bands of ~25 and ~27 kDa were consistently detected across the experiments (Figure 4). These faint bands were the result of specific interactions with the immobilized cystatins, as indicated by the absence of nonspecific binding to the avidin resin in the absence of AviTagged cystatin.

Differential Intestain Capture Using Cystatins with Tailored Specificities

A quantitative analysis of the immobilized proteins was performed to compare the complements of intestains specifically inhibited by SlCYS8, P2F, and Q47P used as an Ldp30-negative control. T6S, a single variant of SlCYS8 presenting a protease inhibitory range comparable to the original inhibitor,[53] was included in the experiments as a neutral control. As expected, similar protein profiles were observed for SlCYS8 and T6S, including Ldp30 as a major band, and the 25-kDa band as a minor protein product (Figure 5A). The same two bands were captured by P2F, along with the 27-kDa protein captured above in lower amounts by Q47P.

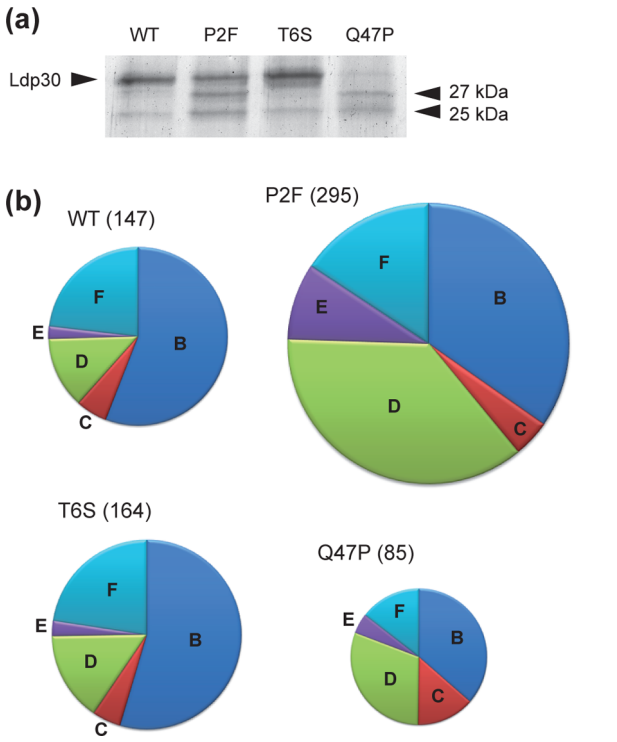


Figure 5. Different target intestain profiles for the cystatin variants. (A) Protein bands revealed by Coomassie Blue staining following protease capture with SlCYS8, P2F, T6S or Q47P and 12% SDS-PAGE. Arrows point to the Ldp30, 27-kDa and 25-kDa protein products observed on gel. (B) Spectral count quantitative analysis of the intestain peptides. The captured proteins were excised from the Coomassie Blue-stained gels, trypsin-digested and submitted to LC–MS/MS analysis. Individual intestain peptides were counted for each band, pooled and assigned to a specific intestain subfamily, based on the 6-group classification presented above (see Figure 2). Total area of the pie charts and areas assigned to each intestain subfamily were scaled to illustrate the average total number of intestain spectra for each repetition (in parentheses). Three repetitions were included in the experimental setup for each cystatin variant.

Tandem MS analyses revealed the presence of 279 unique coleopteran protein-derived peptides in the captured protein samples, of which 120 were confidently assigned to the intestain protease family (Supporting Information Table 3). As expected given the high sequence homology levels among members of each intestain subfamily (Figure 2, Supporting Information Figure 2), only 15 of these peptides could be assigned to an individual intestain sequence. By comparison, 109 peptides were successfully assigned to a single subfamily, which prompted us to treat the intestain subfamilies as individual functional entities. The Ldp30, 27-kDa and 25-kDa protein bands were also treated as a whole, based on analyses indicating the occurrence of intestains from different subfamilies in all three bands (data not shown).

The intestain target profile of each SlCYS8 variant was determined by quantitative analysis of tandem MS spectral counts[75] obtained from three biological replicates (Figure 5B). Specific peptides could be assigned to each intestain subfamily, but comparative analyses included only the IntB, IntC, IntD, IntE and IntF subfamilies given that the only peptides assigned to the IntA subfamily were shared with other intestain subfamilies (Supporting Information Table 3). In line with their almost identical potency and activity spectrum against

dx.doi.org/10.1021/pr300699n | J. Proteome Res. XXXX, XXX, XXX–XXX

potato beetle digestive Cys proteases,[53] SlCYS8 and T6S allowed for the detection of a similar number of intestain peptide spectra, distributed the same way among intestain subfamilies. As expected given the reported efficiency and altered inhibitory spectrum of P2F against potato beetle proteases,[53] P2F-captured samples gave about twice the number of spectra as SlCYS8, distributed differently among intestain subfamilies. Unexpectedly, intestain capture with Q47P gave almost 60% of the number of peptides captured with SlCYS8. While the relative peptides distribution among intestain subfamilies was different for the two cystatins variants, the significant amount of intestain spectra detected with Q47P remained surprising given the negligible activity of this mutant against papain and its inclusion in the experiments as a negative control.

Functional Discrimination of Intestain Classes

An analysis of the Coomassie Blue-stained gels allowed us to estimate that 8 μg of immobilized cystatins capture up to 2 μg of intestains under our experimental conditions. We could also purify up to twice as much intestains with P2F compared to wild-type SlCYS8 (Figure 5), or a similar amount of peptides from half the amount of beetle extract (data not shown). These observations suggested that neither the SlCYS8 variants nor the intestains were limiting under our conditions, and that relative affinity trends could be inferred between the intestain subfamilies and each SlCYS8 variant. An examination of the spectral counts assigned to each intestain subfamily revealed significant differences between the cystatin variants (Figure 6A). For instance, the strong increase in total intestain peptide spectra observed with P2F was due, for the most part, to a significantly increased affinity for members of the IntD and IntE subfamilies. Q47P showed significantly reduced affinity for members the IntB and IntF subfamilies compared to SlCYS8, but comparable affinities to members of other classes. This unexpected observation led us to compare the inhibitory effects of Q47P, SlCYS8 and P2F against canonical Cys protease activities in potato beetle extracts (Figure 6B). As already reported,[53] P2F was by far more effective than SlCYS8 against the insect *Z*-Phe-Arg-MCA-hydrolyzing (cathepsin L-like) and *Z*-Arg-Arg-MCA-hydrolyzing (cathepsin B-like) protease activities. As expected, Q47P was weakly active against the cathepsin L-like activities, but surprisingly as effective as SlCYS8 against cathepsin B-like activities.

DISCUSSION

A key objective of this study was to devise an activity-based protease profiling strategy for the assessment of Cys protease−cystatin interactions in host (plant)−(insect) pest systems. Activity-based protein profiling is a robust and powerful means to capture and identify target enzymes in a variety of tissues and organisms.[81−83] The approach generally involves chemical substrate or pseudosubstrate inhibitor probes, and lends itself to the high-throughput screening and mechanistic assessment of enzyme families.[84] It has proved of particular interest, in recent years, for the structure/function characterization of C1A Cys proteases and the design of potent specific inhibitors.[85−87] The technique described herein, which relies on biotinylated protein inhibitors to capture the target proteases, represents a useful complement to the usual chemical probes approach, notably allowing for the monitoring of protease−inhibitor interactions with a range of natural or engineered protein inhibitors, and for the rational selection of potent inhibitor

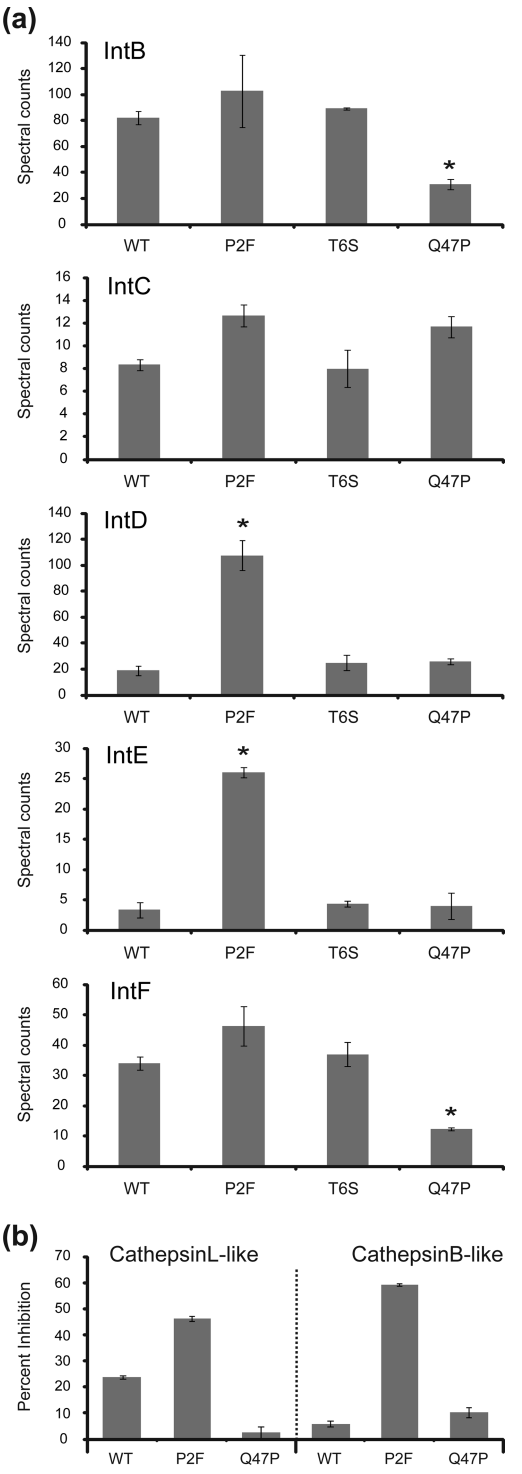


Figure 6. Unique inhibitory spectrum for the strong cystatin variant P2F. (A) Intestain peptide spectral counts for SlCYS8, P2F, T6S and Q47P interacting with different intestain subfamilies. Asterisks (*) indicate significant differences compared to wild-type SlCYS8 (Student's *t*-test; $P < 0.001$). (B) Inhibition of potato beetle midgut Z-Phe-Arg-MCA-hydrolyzing (cathepsin L-like) and Z-Arg-Arg-MCA-hydrolyzing (cathepsin B-like) activities by SlCYS8, P2F and Q47P. Data are presented as relative inhibition rates compared to a non inhibited control, using 625 ng of potato beetle protein with 2.5 pmol of cystatin for cathepsin L activities; or 1.25 μg of potato beetle protein with 0.25 nmol of cystatin for cathepsin B activities. Each bar is the mean of three independent values ± SE.

dx.doi.org/10.1021/pr300699n | J. Proteome Res. XXXX, XXX, XXX−XXX

candidates for host genetic transformation or gene replacement therapeutics.

Immobilised synthetic or protein inhibitors have previously been used as affinity ligands to isolate insect digestive proteases based on their susceptibility to plant-derived inhibitors. For instance, potato proteinase inhibitor I and the first inhibitory-domain of NaPI, a Ser protease inhibitor from the tobacco relative *Nicotiana alata*, were used recently to characterize differentially inhibited chymotrypsins from midgut extracts of the lepidopteran insect *Helicoverpa punctigera*.[88] Similarly, rice cystatin I and ε-aminocaproic acid-Phe-Phe-CH_3, a chemical ligand for cathepsin B-like enzymes, were used to discriminate digestive Cys proteases of the Colorado potato beetle based on their relative susceptibility to the plant cystatin.[80] Extending the use of insect protease affinity chromatography at the proteome scale, biotinylated cystatin-AviTag fusions were used here for protease capture and analysis in combination with a shotgun proteomics assessment of the immobilized proteins. The goal was to compare the target protease complements of SlCYS8 functional variants against digestive proteases of the potato beetle, taking into account the reported impact of positively selected amino acid sites on the inhibitory specificity of plant cystatins.[39]

P2F, a single variant of SlCYS8 bearing an alternative Phe residue in the hypervariable N-terminal region, has been shown to exhibit stronger and broader activity against Colorado potato beetle Cys proteases compared to the native inhibitor, unlike the neutral control T6S showing a conserved inhibitory range.[53] Accordingly, twice the number of peptides was captured, and a different set of intestain targets identified, for P2F compared to SlCYS8, in contrast with T6S capturing a similar set of intestains. These findings confirmed *in silico* inferences suggesting cystatin-specific interactions with potato beetle intestains (this study) and earlier mutagenesis work linking positively selected amino acids in plant cystatins with their functional diversification against herbivorous pest digestive Cys proteases.[39,53] They also support published kinetics data[23,30] and *in silico* inferences on intestain primary structures (this study) suggesting a link between primary sequence micro-heterogeneity and cystatin susceptibility among potato beetle intestains.

From a practical viewpoint, these findings further illustrate the striking diversity of possible protease−inhibitor interactions in plant−insect systems, and underline the relevance of protease inhibitor (e.g., cystatin) activity-based protease profiling to characterize the inhibitory range of recombinant inhibitors considered for pest control. A straightforward way to use cystatins against coleopteran insects may be to combine them with Ser protease inhibitors, in such a way as to significantly broaden the overall spectrum of target protease types and minimize the effectiveness of physiological compensatory processes in the herbivores.[34,43,89] Extending the inhibitory range of recombinant inhibitors to different protease functional classes, however, raises relevant questions on the specificity and compatibility of the resulting control approach in an environmental context.[2] Using Ser-type inhibitors also raises questions on food safety and quality, given with the well described antidigestive effects of these proteins in humans.[90] By comparison, no Cys protease targets are found in the human digestive tract, and the innocuity of plant cystatins expressed in transgenic crops has been documented.[91−93] In this context, an ecologically sustainable alternative for pest control might be to combine complementary cystatin variants acting in concert to reproduce the broad inhibitory range and toxic effect of E-64 against Colorado potato beetle.[58] Work is underway to identify cystatin variants for the broad-spectrum inhibition of potato beetle digestive Cys proteases, using the cystatin activity-based profiling approach described herein. Work is also underway to determine whether the absence of captured peptides specific to the intestain A subfamily in our analyses could be explained by the absence IntA enzymes in the insect extracts, or by a limited susceptibility of these enzymes to cystatin inhibition as suggested earlier.[23]

■ ASSOCIATED CONTENT

Supporting Information

The data associated with this manuscript may be downloaded from the ProteomeCommons.org Tranche repository using the following three hashes: (1) BandA(Ldp30): 0RrjiHuNfJ2iRV-SyFK0Q2i Hk8XddXiNbYjdt3363fGxS6pjISEAb5Hlzt/dXfWLadK/i4stwEqb1i6QhA5SlGD/QyYkAAAAAAAACig==; (2) BandB(27 kDa): GrDuS/TJJaqX8NcJgt2DVtkH-DUiqCN0XG3LSLn1k97 V5M3Txjt37z7kXcdWxDF 5BxKtTC78Hlste+UECce+QAbMJ6wIAAAAAAAACiw==; (3) BandC(25 kDa): iTt5QwYAilX0BSpsA4gC 5i7DGmBfg49rbREOb1LKfQkJhwCZ1I2AzmdWSIgLdUpNc-zi/x9tZwtNJQID42+kBjay3sVUAAAAAAAACfQ==. Supplemental tables and figures. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Telephone: +1-418-656-2131 x5076. Fax: +1-418-656-7856. E-mail: dominique.michaud@fsaa.ulaval.ca.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Line Roy and Marcos Di Falco at Génome Québec (McGill University, Montréal QC, Canada) for valuable help with proteomic analyses and data interpretation, Maxime St-Amand for contributing to the GoldenGate DNA cloning, and Jean-Frédéric Guay for providing the insect material. This work was supported by a Discovery grant from the Natural Science and Engineering Research Council of Canada, and by a Team Project grant from the Fonds Québécois de la Recherche sur la Nature et les Technologies.

■ REFERENCES

(1) Haq, S. K.; Atif, S. M.; Khan, R. H. Protein proteinase inhibitor genes in combat against insects, pests, and pathogens: natural and engineered phytoprotection. *Arch. Biochem. Biophys.* **2004**, *431*, 145−159.

(2) Schlüter, U.; Benchabane, M.; Munger, A.; Kiggundu, A.; Vorster, J.; Goulet, M.-C.; Cloutier, C.; Michaud, D. Recombinant protease inhibitors for herbivore pest control: a multitrophic perspective. *J. Exp. Bot.* **2010**, *61*, 4169−4183.

(3) Hilder, V. A.; Gatehouse, A. M. R.; Sheerman, S. E.; Barker, R. F.; Boulter, D. A novel mechanism of insect resistance engineered into tobacco. *Nature* **1987**, *330*, 160−163.

(4) Johnson, R.; Narvaez, J.; An, G.; Ryan, C. A. Expression of proteinase inhibitors I and II in transgenic tobacco plants: effects on natural defense against *Manduca sexta* larvae. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 9871−9875.

(5) Duan, X.; Li, X.; Xue, Q.; Abo-El-Saad, M.; Xu, D.; Wu, R. Transgenic rice plants harbouring an introduced potato proteinase inhibitor II gene are insect resistant. *Nat. Biotechnol.* **1996**, *14*, 494−498.
(6) Xu, D.; Xue, Q.; McElroy, D.; Mawal, Y.; Hilder, V. A.; Wu, R. Constitutive expression of a cowpea trypsin inhibitor gene, *CpTi*, in transgenic rice plants confers resistance to two major rice insect pests. *Mol. Breed.* **1996**, *2*, 167−173.
(7) Huang, J.; Hu, R.; Rozelle, S.; Pray, C. Insect-resistant GM rice in farmers' fields: assessing productivity and health effects in China. *Science* **2005**, *308*, 688−690.
(8) Qiu, J. Is China ready for GM rice? *Nature* **2008**, *455*, 850−852.
(9) Birk, Y. *Plant protease inhibitors*; Springer: New York, 2003.
(10) Broadway, R. M. The adaptation of insects to protease inhibitors. In *Recombinant protease inhibitors in plants*; Michaud, D., Ed.; Landes Bioscience: Georgetown, TX, 2000; pp 80−88.
(11) Hernandez, C. A.; Pujol, M.; Alfonso-Rubi, J.; Armas, R.; Coll, Y.; Pérez, M.; Gonzalez, A.; Ruiz, M.; Castanera, P.; Ortego, F. Proteolytic gut activities in the rice water weevil, *Lissorhoptrus brevirostris* Suffrian (Coleoptera: Curculionidae). *Arch. Insect Biochem. Physiol.* **2003**, *53*, 19−29.
(12) Prabhakar, S.; Chen, M.-S.; Elpidina, E. N.; Vinokurov, K. S.; Smith, C. S.; Marshall, J.; Oppert, B. Sequence analysis and molecular characterization of larval midgut cDNA transcripts encoding peptidases from the yellow mealworm, *Tenebrio molitor* L. *Insect Mol. Biol.* **2007**, *16*, 455−468.
(13) Vinokurov, K. S.; Elpidina, E. N.; Zhuzhikov, D. P.; Oppert, B.; Kodrik, D.; Sehnal, F. Digestive proteolysis organization in two closely related tenebrionid beetles: red flour beetle (*Tribolium castanuem*) and confused flour beetle (*Tribolium confusum*). *Arch. Insect Biochem. Physiol.* **2009**, *70*, 254−279.
(14) Kiggundu, A.; Muchwezi, J.; Van der Vyver, C.; Viljoen, A.; Vorster, J.; Schlüter, U.; Kunert, K.; Michaud, D. Deleterious effects of plant cystatins against the banana weevil *Cosmopolites sordidus*. *Arch. Insect Biochem. Physiol.* **2010**, *73*, 87−105.
(15) Cloutier, C.; Jean, C.; Fournier, M.; Yelle, S.; Michaud, D. Adult Colorado potato beetles, *Leptinotarsa decemlineata* compensate for nutritional stress on oryzacystatin I transgenic potato plants by hypertrophic behavior and over-production of insensitive proteases. *Arch. Insect Physiol. Biochem.* **2000**, *44*, 69−81.
(16) Ahn, J.-E.; Salzman, R. A.; Braunagel, S. C.; Koiwa, H.; Zhu-Salzman, K. Functional roles of specific bruchid protease isoforms in adaptation to a soybean protease inhibitor. *Insect Mol. Biol.* **2004**, *13*, 649−657.
(17) Jongsma, M. A.; Bakker, P. L.; Peters, J.; Bosch, D.; Stiekema, W. J. Adaptation of *Spodoptera exigua* larvae to plant proteinase inhibitors by induction of gut proteinase activity insensitive to inhibition. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 8041−8045.
(18) Bown, D. P.; Wilkinson, H. S.; Gatehouse, J. A. Differentially regulated inhibitor-sensitive and insensitive protease genes from the phytophagous insect pest, *Helicoverpa armigera*, are members of complex multigene families. *Insect Biochem. Mol. Biol.* **1997**, *27*, 625−638.
(19) Girard, C.; Le Métayer, M.; Zaccomer, B.; Bartlet, E.; Williams, I.; Bonadé-Bottino, M.; Pham-Delegue, M. H.; Jouanin, L. Growth stimulation of beetle larvae reared on a transgenic oilseed rape expressing a cysteine protease inhibitor. *J. Insect Physiol.* **1998**, *44*, 263−270.
(20) Mazumdar-Leighton, S.; Broadway, R. M. Transcriptional induction of diverse midgut trypsins in larval *Agrotis ipsilon* and *Helicoverpa zea* feeding on the soybean trypsin inhibitor. *Insect Biochem. Mol. Biol.* **2001**, *31*, 645−657.
(21) Zhu-Salzman, K.; Koiwa, H.; Salzman, R. A.; Shade, R. E.; Ahn, J.-E. Cowpea bruchid *Callosobruchus maculatus* uses a three-component strategy to overcome a plant defensive cysteine protease inhibitor. *Insect Mol. Biol.* **2003**, *12*, 135−145.
(22) Brunelle, F.; Cloutier, C.; Michaud, D. Colorado potato beetles compensate for tomato cathepsin D inhibitor expressed in transgenic potato. *Arch. Insect Physiol. Biochem.* **2004**, *55*, 103−113.
(23) Gruden, K.; Kuipers, A. G. J.; Guncar, G.; Slapar, N.; Strukelj, B.; Jongsma, M. A. Molecular basis of Colorado potato beetle adaptation to potato plant defence at the level of digestive cysteine proteinases. *Insect Biochem. Mol. Biol.* **2004**, *34*, 365−375.
(24) Liu, Y.; Salzman, R. A.; Pankiw, T.; Zhu-Salzman, K. Transcriptional regulation in southern corn rootworm larvae challenged by soyacystatin N. *Insect Biochem. Mol. Biol.* **2004**, *34*, 1069−1077.
(25) Koo, Y. D.; Ahn, J.-E.; Salzman, R. A.; Moon, J.; Chi, Y. H.; Yun, D.-J.; Lee, S. Y.; Koiwa, H.; Zhu-Salzman, K. Functional expression of an insect cathepsin B-like counter-defence protein. *Insect Mol. Biol.* **2008**, *17*, 235−245.
(26) Ahn, J. E.; Guarino, L. A.; Zhu-Salzman, K. Coordination of hepatocyte nuclear factor 4 and seven-up controls insect counter-defense cathepsin B expression. *J. Biol. Chem.* **2010**, *285*, 6573−6584.
(27) Michaud, D. Avoiding protease-mediated resistance in herbivorous pests. *Trends Biotechnol.* **1997**, *15*, 4−6.
(28) Girard, C.; Le Métayer, M.; Bonadé-Bottino, M.; Pham-Delegue, M. H.; Jouanin, L. High level of resistance to proteinase inhibitors may be conferred by proteolytic cleavage in beetle larvae. *Insect Biochem. Mol. Biol.* **1998**, *28*, 229−237.
(29) Giri, A. P.; Harsulkar, A. M.; Deshpande, V. V.; Sainani, M. N.; Gupta, V. S.; Ranjekar, P. K. Chickpea defensive proteinase inhibitors can be inactivated by podborer gut proteinases. *Plant Physiol.* **1998**, *116*, 393−401.
(30) Gruden, K.; Popovic, T.; Cimerman, N.; Krizaj, I.; Strukelj, B. Diverse enzymatic specificities of digestive proteases, 'intestains', enable Colorado potato beetle larvae to counteract the potato defence mechanism. *Biol. Chem.* **2003**, *384*, 305−310.
(31) Ahn, J. E.; Zhu-Salzman, K. CmCatD, a cathepsin D-like protease, has a potential role in insect defense against a phytocystatin. *J. Insect Physiol.* **2009**, *55*, 678−685.
(32) Yang, L.; Fang, Z.; Dicke, M.; van Loon, J. J. A.; Jongsma, M. A. The diamondback moth, *Plutella xylostella*, specifically inactivates Mustard Trypsin Inhibitor 2 (MTI2) to overcome host plant defence. *Insect Biochem. Mol. Biol.* **2009**, *39*, 55−61.
(33) Rivard, D.; Cloutier, C.; Michaud, D. Colorado potato beetles show differential digestive compensatory responses to host plants expressing distinct sets of defense proteins. *Arch. Insect Biochem. Physiol.* **2004**, *55*, 114−123.
(34) Oppert, B.; Morgan, T. D.; Hartzer, K.; Kramer, K. J. Compensatory proteolytic responses to dietary proteinase inhibitors in the red flour beetle, *Tribolium castaneum* (Coleoptera: Tenebrionidae). *Comp. Biochem. Physiol. C* **2005**, *140*, 53−58.
(35) Vila, L.; Quilis, J.; Meynard, D.; Breitler, J. C.; Marfa, V.; Murillo, I.; Vassal, J. M.; Messeguer, J.; Guiderdoni, E.; San Segundo, B. Expression of the maize proteinase inhibitor (*mpi*) gene in rice plants enhances resistance against the striped stem borer (*Chilo suppressalis*): effects on larval growth and insect gut proteinases. *Plant Biotechnol. J.* **2005**, *3*, 187−202.
(36) Chi, Y. H.; Salzman, R. A.; Balfe, S.; Ahn, J.-E.; Sun, W.; Moon, J.; Yun, D.-J.; Lee, S. Y.; Higgins, T. J. V.; Pittendrigh, B.; Murdock, L. L.; Zhu-Salzman, K. Cowpea bruchid midgut transcriptome response to a soybean cystatin−costs and benefits of counter-defense. *Insect Mol. Biol.* **2009**, *18*, 97−110.
(37) Lopes, A. R.; Juliano, M. A.; Juliano, L.; Terra, W. R. Coevolution of insect trypsins and inhibitors. *Arch. Insect Biochem. Physiol.* **2004**, *55*, 140−152.
(38) Christeller, J. T. Evolutionary mechanisms acting on proteinase inhibitor variability. *FEBS J.* **2005**, *272*, 5710−5722.
(39) Kiggundu, A.; Goulet, M.-C.; Goulet, C.; Dubuc, J.-F.; Rivard, D.; Benchabane, M.; Pépin, G.; van der Vyver, C.; Kunert, K.; Michaud, D. Modulating the proteinase inhibitory profile of a plant cystatin by single mutations at positively selected amino acid sites. *Plant J.* **2006**, *48*, 403−413.
(40) Urwin, P. E.; McPherson, M. J.; Atkinson, H. J. Enhanced transgenic plant resistance to nematodes by dual proteinase inhibitor constructs. *Planta* **1998**, *204*, 472−479.

(41) Inanaga, H.; Kobayasi, D.; Kouzuma, Y.; Aoki-Yasunaga, C.; Iiyama, K.; Kimura, M. Protein engineering of novel proteinase inhibitors and their effects on the growth of *Spodoptera exigua* larvae. *Biosci. Biotechnol. Biochem.* **2001**, *65*, 2259−2264.
(42) Outchkourov, N. S.; de Kogel, W. J.; Wiegers, G. L.; Abrahamson, M.; Jongsma, M. A. Engineered multidomain cysteine protease inhibitors yield resistance against western flower thrips (*Frankliniella occidentalis*) in greenhouse trials. *Plant Biotechnol. J.* **2004**, *2*, 449−458.
(43) Brunelle, F.; Girard, C.; Cloutier, C.; Michaud, D. A hybrid, broad-spectrum inhibitor of Colorado potato beetle aspartate and cysteine proteinases. *Arch. Insect Physiol. Biochem.* **2005**, *60*, 20−31.
(44) Benchabane, M.; Goulet, M.-C.; Dallaire, C.; Côté, P.-L.; Michaud, D. Hybrid protease inhibitors for pest and pathogen control−a functional cost for the fusion partners? *Plant Physiol. Biochem.* **2008**, *46*, 701−708.
(45) Abdeen, A.; Virgos, A.; Olivella, E.; Villanueva, J.; Aviles, X.; Gabarra, R.; Prat., S. Multiple insect resistance in transgenic tomato plants over-expressing two families of plant proteinase inhibitors. *Plant Mol. Biol.* **2005**, *57*, 189−202.
(46) Senthilkumar, R.; Cheng, C.-P.; Yeh, K.-W. Genetically pyramiding protease-inhibitor genes for dual broad-spectrum resistance against insect and phytopathogens in transgenic tobacco. *Plant Biotechnol. J.* **2010**, *8*, 65−75.
(47) Dunse, K. M.; Stevens, J. A.; Lay, F. T.; Gaspar, Y. M.; Heath, R. L.; Anderson, M. A. Coexpression of potato type I and II proteinase inhibitors gives cotton plants protection against insect damage in the field. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 15011−15015.
(48) Urwin, P. E.; Atkinson, H. J.; Waller, D. A.; McPherson, M. J. Engineered oryzacystatin-I expressed in transgenic hairy roots confers resistance to *Globodera pallida*. *Plant J.* **1995**, *8*, 121−131.
(49) Jongsma, M. A.; Bakker, P. L.; Stiekema, W.; Bosch, D. Phage display of a double-headed proteinase inhibitor: Analysis of the binding domains of potato proteinase inhibitor II. *Mol. Breed.* **1995**, *1*, 181−191.
(50) Koiwa, H.; D'Urzo, M. P.; Assfalg-Machleidt, I.; Zhu-Salzman, K.; Shade, R. E.; An, H.; Murdock, L. L.; Machleidt, W.; Bressan, R. A.; Hasegawa, P. M. Phage display selection of hairpin loop soyacystatin variants that mediate high affinity inhibition of a cysteine proteinase. *Plant J.* **2001**, *27*, 383−391.
(51) Ceci, L. R.; Volpicella, M.; Rahbé, Y.; Gallerani, R.; Beekwilder, J.; Jongsma, M. A. Selection by phage display of a variant mustard trypsin inhibitor toxic against aphids. *Plant J.* **2003**, *33*, 557−566.
(52) Melo, F. R.; Mello, M. O.; Franco, O. L.; Rigden, D. J.; Mello, L. V.; Genu, A. M.; Silva-Filho, M. C.; Gleddie, S.; Grossi-de-Sa, M. F. Use of phage display to select novel cystatins specific for *Acanthoscelides obtectus* cysteine proteinases. *Biochim. Biophys. Acta* **2003**, *1651*, 146−152.
(53) Goulet, M.-C.; Dallaire, C.; Vaillancourt, L.-P.; Khalf, M.; Badri, M. A.; Preradov, A.; Duceppe, M.-O.; Goulet, C.; Cloutier, C.; Michaud, D. Tailoring the specificity of a plant cystatin toward herbivorous insect digestive cysteine proteases by single mutations at positively selected amino acid sites. *Plant Physiol.* **2008**, *146*, 1010−1019.
(54) *Tribolium* Genome Sequencing Consortium.. The genome of the model beetle and pest *Tribolium castaneum*. *Nature* **2008**, *452*, 949−955.
(55) Michaud, D.; Nguyen-Quoc, B.; Yelle, S. Selective inhibition of Colorado potato beetle cathepsin H by oryzacystatins I and II. *FEBS Lett.* **1993**, *331*, 173−176.
(56) Michaud, D.; Bernier-Vadnais, N.; Overney, S.; Yelle, S. Constitutive expression of digestive cysteine proteinase forms during development of the Colorado potato beetle, *Leptinotarsa decemlineata* Say (Coleoptera: Chrysomelidae). *Insect Biochem. Mol. Biol.* **1995**, *25*, 1041−1048.
(57) Gruden, K.; Strukelj, B. S.; Popovic, T.; Lenarcic, B.; Bevec, T.; Brzin, J.; Kregar, I.; Herzog-Velikonja, J.; Stiekema, W. J.; Bosch, D.; Jongsma, M. A. The cysteine protease activity of Colorado potato beetle (*Leptinotarsa decemineata* Say) guts, which is insensitive to potato protease inhibitors, is inhibited by thyroglobulin type-I domain inhibitors. *Insect Biochem. Mol. Biol.* **1998**, *28*, 549−560.
(58) Bolter, C. J.; Latoszek-Green, M. Effect of chronic ingestion of the cysteine proteinase inhibitor, E-64, on Colorado potato beetle gut proteinases. *Entomol. Exp. Appl.* **1997**, *83*, 295−303.
(59) Cloutier, C.; Fournier, M.; Jean, C.; Yelle, S.; Michaud, D. Growth compensation and faster development of Colorado potato beetle (Coleoptera: Chrysomelidae) feeding on potato foliage expressing oryzacystatin I. *Arch. Insect Biochem. Physiol.* **1999**, *40*, 69−79.
(60) Bouchard, É.; Cloutier, C.; Michaud, D. Oryzacystatin I expressed in transgenic potato induces digestive compensation in an insect natural predator via its herbivorous prey feeding on the plant. *Mol. Ecol.* **2003**, *12*, 2439−2446.
(61) Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high-throughput. *Nucleic Acids Res.* **2004**, *32*, 1792−1797.
(62) Guindon, S.; Gascuel, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **2003**, *52*, 696−704.
(63) Jones, D. T.; Taylor, W. R.; Thornton, J. M. The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* **1992**, *8*, 275−282.
(64) Sali, A.; Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.* **1993**, *234*, 779−815.
(65) Marti-Renom, M. A.; Stuart, A.; Fiser, A.; Sánchez, R.; Melo, F.; Sali, A. Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 291−325.
(66) Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. *PROCHECK*: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **1993**, *26*, 283−291.
(67) Chen, R.; Li, L.; Weng, Z. ZDOCK: an initial-stage protein-docking algorithm. *Proteins* **2003**, *52*, 80−87.
(68) Li, L.; Chen, R.; Weng, Z. RDOCK: refinement of rigid-body protein docking predictions. *Proteins* **2003**, *53*, 693−707.
(69) Beckett, D.; Kovaleva, E.; Schatz, P. J. A minimal peptide substrate in biotin holoenzyme synthetase-catalyzed biotinylation. *Protein Sci.* **1999**, *8*, 921−929.
(70) Engler, C.; Kandzia, R.; Marillonnet, S. A one pot, one step, precision cloning method with high throughput capability. *PLoS ONE* **2008**, *3*, e3647.
(71) Perkins, D. N.; Pappin, D. J. C.; Creasy, D. M.; Cottrell, J. S. Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* **1999**, *20*, 3551−3567.
(72) Craig, R.; Cortens, J. P.; Beavis, R. C. Open source system for analyzing, validating, and storing protein identification data. *J. Proteome Res.* **2004**, *3*, 1234−1242.
(73) Keller, A.; Nesvizhskii, A. I.; Kolker, E.; Aebersold, R. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* **2002**, *74*, 5383−5392.
(74) Nesvizhskii, A. I.; Keller, A.; Kolker, E.; Aebersold, R. A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* **2003**, *75*, 4646−4658.
(75) Zhang, B.; VerBerkmoes, N. C.; Langston, M. A.; Uberbacher, E.; Hettich, R. L.; Samatova, N. F. Detecting differential and correlated protein expression in label-free shotgun proteomics. *J. Proteome Res.* **2006**, *5*, 2909−2918.
(76) Girard, C.; Rivard, D.; Kiggundu, A.; Kunert, K.; Gleddie, S. C.; Cloutier, C.; Michaud, D. A multicomponent, elicitor-inducible cystatin complex in tomato *Solanum lycopersicum*. *New Phytol.* **2007**, *173*, 841−851.
(77) Ljunggren, A.; Redzynia, I.; Alvarez-Fernandez, M.; Abrahamson, M.; Mort, J. S.; Krupa, J. C.; Jaskolski, M.; Bujacz, G. Crystal structure of the parasite protease inhibitor chagasin in complex with a host target cysteine protease. *J. Mol. Biol.* **2007**, *371*, 137−153.
(78) Turk, V.; Bode, W. The cystatins: protein inhibitors of cysteine proteinases. *FEBS Lett.* **1991**, *285*, 213−219.

(79) Arai, S.; Watanabe, H.; Kondo, H.; Emori, Y.; Abe, K. Papain-inhibitory activity of oryzacystatin, a rice seed cysteine proteinase inhibitor, depends on the central Gln-Val-Val-Ala-Gly region conserved among cystatin superfamily members. *J. Biochem.* **1991**, *109*, 294−298.
(80) Visal-Shah, S.; Vrain, T. C.; Yelle, S.; Nguyen-Quoc, B.; Michaud, D. An electroblotting, two-step procedure for the detection of proteinases and the study of proteinase/inhibitor complexes in gelatin-containing polyacrylmaide gels. *Electrophoresis* **2001**, *22*, 2646−2652.
(81) Cravatt, B. F.; Wright, A. T.; Kozarich, J. W. Activity-based protein profiling: from enzyme chemistry to proteomic chemistry. *Annu. Rev. Biochem.* **2008**, *77*, 383−414.
(82) Uttamchandani, M.; Li, J. Q.; Sun, H.; Yao, S. Q. Activity-based protein profiling: New developments and directions in functional proteomics. *ChemBioChem* **2008**, *9*, 667−675.
(83) Kolodziejek, I.; van der Hoorn, R. A. L. Mining the active proteome in plant science and biotechnology. *Curr. Opin. Biotechnol.* **2010**, *21*, 225−233.
(84) Evans, M. J.; Cravatt, B. F. Mechanism-based profiling of enzyme families. *Chem. Rev.* **2006**, *106*, 3279−3301.
(85) Bogyo, M.; Verhelst, S.; Bellingard-Dubouchaud, V.; Toba, S.; Greenbaum, D. Selective targeting of lysosomal cysteine proteases with radiolabeled electrophilic substrate analogs. *Chem. Biol.* **2000**, *7*, 27−38.
(86) Greenbaum, D.; Medzihradszky, K. F.; Burlingame, A.; Bogyo, M. Epoxide electrophiles as activity-dependent cysteine protease profiling and discovery tools. *Chem. Biol.* **2000**, *7*, 569−581.
(87) Chowdhury, S. F.; Joseph, L.; Kumar, S.; Tulsidas, S. R.; Bhat, S.; Ziomek, E.; Menard, R.; Sivaraman, J.; Purisima, E. O. Exploring inhibitor binding at the S′ subsites of cathepsin L. *J. Med. Chem.* **2008**, *51*, 1361−1368.
(88) Dunse, K. M.; Kaas, Q.; Guarino, R. F.; Barton, P. A.; Craik, D. J.; Anderson, M. A. Molecular basis for the resistance of an insect chymotrypsin to a potato type II proteinase inhibitor. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 15016−15021.
(89) Oppert, B.; Morgan, T. D.; Hartzer, K.; Lenarcic, B.; Galesa, K.; Brzin, J.; Turk, V.; Yoza, K.; Ohtsubo, K.; Kramer, K. J. Effects of proteinase inhibitors on digestive proteinases and growth of the red flour beetle, *Tribolium castaneum* (Herbst) (Coleoptera: Tenebrionidae). *Comp. Biochem. Physiol. C* **2003**, *134*, 481−490.
(90) Gilani, G. S.; Cockell, K. A.; Sepher, E. Effects of antinutritional factors on protein digestibility and amino acid availability in foods. *J. AOAC Int.* **2005**, *88*, 967−987.
(91) Atkinson, H. J.; Johnston, K. A.; Robbins, M. *Prima facie* evidence that a phytocystatin for transgenic plant resistance to nematodes is not a toxic risk in the human diet. *J. Nutr.* **2004**, *134*, 431−434.
(92) Arai, S.; Matsumoto, I.; Emori, Y.; Abe, K. Plant seed cystatins and their target enzymes of endogenous and exogenous origin. *J. Agric. Food Chem.* **2002**, *50*, 6612−6617.
(93) Benchabane, M.; Schlüter, U.; Vorster, J.; Goulet, M.-C.; Michaud, D. Plant cystatins. *Biochimie* **2010**, *92*, 1657−1666.