THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 282, NO. 40, pp. 29721–29728, October 5, 2007

An Asparaginyl Endopeptidase Mediates *in Vivo* Protein Backbone Cyclization*[S]

Received for publication, June 25, 2007, and in revised form, August 9, 2007 Published, JBC Papers in Press, August 13, 2007, DOI 10.1074/jbc.M705185200

Ivana Saska[‡], Amanda D. Gillon[§], Noriyuki Hatsugai[¶], Ralf G. Dietzgen[‖], Ikuko Hara-Nishimura[¶], Marilyn A. Anderson[§], and David J. Craik[‡1]

From the [‡]Institute for Molecular Bioscience, University of Queensland, Brisbane, Queensland 4072, Australia, the [§]Department of Biochemistry, La Trobe University, Melbourne, Victoria 3086, Australia, the [¶]Graduate School of Science, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan, and the [‖]Queensland Department of Primary Industries and Fisheries, Emerging Technologies, University of Queensland, Brisbane, Queensland 4072, Australia

Proteases can catalyze both peptide bond cleavage and formation, yet the hydrolysis reaction dominates in nature. This presents an interesting challenge for the biosynthesis of backbone cyclized (circular) proteins, which are encoded as part of precursor proteins and require post-translational peptide bond formation to reach their mature form. The largest family of circular proteins are the plant-produced cyclotides; extremely stable proteins with applications as bioengineering scaffolds. Little is known about the mechanism by which they are cyclized *in vivo* but a highly conserved Asn (occasionally Asp) residue at the C terminus of the cyclotide domain suggests that an enzyme with specificity for Asn (asparaginyl endopeptidase; AEP) is involved in the process. *Nicotiana benthamiana* does not endogenously produce circular proteins but when cDNA encoding the precursor of the cyclotide kalata B1 was transiently expressed in the plants they produced the cyclotide, together with linear forms not commonly observed in cyclotide-containing plants. Observation of these species over time showed that *in vivo* asparaginyl bond hydrolysis is necessary for cyclization. When AEP activity was suppressed, either by decreasing AEP gene expression or using a specific inhibitor, the amount of cyclic cyclotide in the plants was reduced compared with controls and was accompanied by the accumulation of extended linear species. These results suggest that an AEP is responsible for catalyzing both peptide bond cleavage and ligation of cyclotides in a single processing event.

Backbone-cyclized (circular) proteins have been identified in bacteria, plants, and mammals (1, 2). The largest family of circular proteins are the cyclotides, which combine a cyclic backbone with a cystine knot arrangement of three disulfide bonds. The cyclic cystine knot (CCK)[2] framework endows cyclotides with a number of advantages over conventional (*i.e.* acyclic) proteins. Cyclotides are resistant to thermal and biochemical extremes, treatment with endoproteases and, because of the absence of termini, exoproteases (3). These features have motivated the development of cyclotides as stable scaffolds in drug design (4). Fig. 1 shows the compact structure of the prototypic cyclotide kalata B1 (5). Although cyclotides display a broad range of bioactivities (6–12) the ability of the cyclotides kalata B1 and kalata B2 to inhibit the development of the insect pests *Helicoverpa punctigera* and *H. armigera* suggests that their endogenous function is in plant defense (13, 14).

Cyclotides are encoded as part of precursor proteins that have a highly conserved organization. Precursor cDNA clones derived from *Oldenlandia affinis* (Rubiaceae) (13) and *Viola odorata* (Violaceae) (15) exhibit an N-terminal endoplasmic reticulum (ER) signal sequence followed by a pro-region, one or more cyclotide domains and a short hydrophobic C-terminal tail sequence. In precursors that contain multiple cyclotide domains each one is preceded by a repeated portion of the pro-region (~20 residues) that has been designated the N-terminal repeat (NTR) (13). The sequence identity of the NTR between species is low but the motif appears to be a structurally conserved helix (15). Fig. 1 shows the general organization of cyclotide precursor proteins and an expanded portion of the specific precursor that encodes kalata B1, designated Oak1 (for O. affinis kalata B1).

Although the residues preceding the cyclotide domain are not highly conserved, an Asn (occasionally Asp) residue is located at the C terminus of the cyclotide domain across all precursor clones. The position of the Asn residue suggests that it is critical for processing, and possibly cyclization, of the cyclotide domain. Intein-based mechanisms have been used to cyclize cyclotides (16) but the residues at the termini of the cyclotide domain are not reminiscent of those required for *in vivo* intein splicing, suggesting that an enzymatic mechanism is probably involved. Asparaginyl endopeptidases specifically cleave peptide bonds C-terminal to Asn and, less efficiently, after Asp and are widespread in plants where they are com-

* This work was funded in part by a grant from the Australian Research Council, an Australian Postgraduate Award, and a University of Queensland Graduate School Research Travel Award. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Fig. S1.

[1] An ARC Professorial Fellow. To whom correspondence should be addressed: Inst. for Molecular Bioscience, The University of Queensland, Brisbane, QLD, 4072, Australia. Tel.: 61-7-3346-2019; Fax: 61-7-3346-2029; E-mail: d.craik@imb.uq.edu.au.

[2] The abbreviations used are: CCK, cyclic cystine knot; NTR, N-terminal repeat; MES, 4-morpholineethanesulfonic acid; AEP, asparaginyl endopeptidase; VPE, vacuolar processing enzyme; dpi, days postinfiltration; ER, endoplasmic reticulum; MALDI-TOF, matrix-assisted laser desorption/ionization-time of flight.

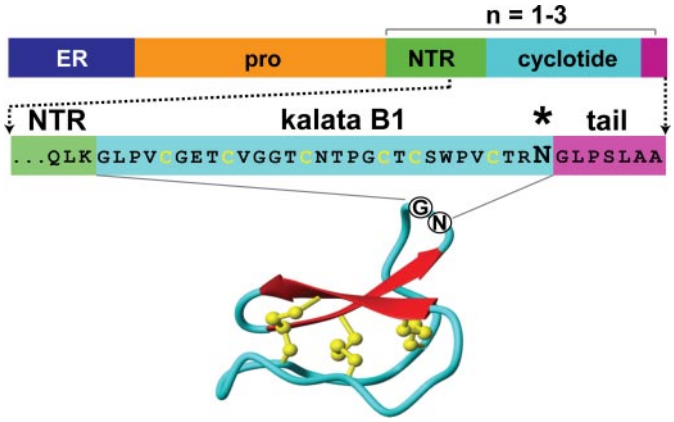

FIGURE 1. **Cyclotide structure and precursor protein organization.** The structure of the prototypic cyclotide kalata B1 (PDB code 1NB1) is shown in a *ribbon representation* with the N and G residues where cyclization takes place *highlighted*. Above this is a schematic representation of the organization of the cyclotide precursors known to date, which encode up to three copies of the same or different cyclotides. The predicted ER signal sequence is followed by a pro-region, a region designated the NTR, the cyclotide domain and a short C-terminal tail region. The NTR always precedes the cyclotide domain in single or multiple copy precursors. An expanded portion of the Oak1 precursor with the residues from kalata B1 and the flanking NTR and tail regions is also shown. The conserved C-terminal Asn is marked with an *asterisk*.

monly called vacuolar processing enzymes (VPEs) and are involved in the activation and degradation of storage proteins (17).

In the current study we present evidence for the role of a plant AEP in the backbone cyclization of cyclotides. We demonstrate that AEP activity is responsible for asparaginyl bond hydrolysis in cyclotide-containing plants and that this activity is linked to the cyclization of cyclotides in *N. benthamiana*. Although *N. benthamiana* does not endogenously produce circular proteins, it is capable of producing correctly folded and backbone-cyclized kalata B1 when transiently expressing the Oak1 precursor, in addition to C-terminally extended linear forms of the cyclotide. Analysis of the different cyclotide forms over time showed that linear kalata B1 was not cyclized *in vivo* despite the presence of a C-terminal Asn, suggesting that hydrolysis of the asparaginyl bond is necessary for cyclization. Knocking down or inhibiting AEP activity significantly reduced the amount of the cyclic protein produced and led to an accumulation of the longer linear cyclotide species. Overall the results appear to implicate an AEP in catalyzing both protein backbone hydrolysis and cyclization.

MATERIALS AND METHODS

Assay for Asparaginyl Bond Hydrolysis—O. affinis and *V. tricolour* leaf extracts in 50 mM sodium acetate buffer (pH 5.5) containing 50 mM NaCl and 1 mM EDTA were preincubated with protease inhibitors (0.2–5 mM in dimethyl sulfoxide) in 20 mM sodium acetate (pH 5.5) containing 0.1 mM EDTA and 0.1 M dithiothreitol at 30 °C for 1 h. Activity was measured against 200 μM benzyloxycarbonyl-Ala-Ala-Asn-methylcoumaryl amide (Z-AAN-mca) at 465 nm for 1 h at 30 °C using a fluorescence spectrophotometer (RF-5000, Shimadzu). Protease inhibitors (Peptide Institute) were phenylmethylsulfonyl fluoride, pepstatin A, acetyl-Tyr-Val-Ala-Asp-aldehyde (Ac-YVAD-CHO), acetyl-Glu-Ser-Glu-Asn-aldehyde (Ac-ESEN-CHO) and L-trans-epoxysuccinyl-leucyl amido(4-guanidino)butane (E-64).

Construction of the Oak1 Expression Plasmid for Transient Expression Assays—Oak1 cDNA, including the ER signal sequence (13), was amplified by PCR using primers EcoRI-MAKF (5′-GGAATTCATGGCTAAGTT-3′) and SLAA*BamHI (5′-GGATCCTTATGCGGCCAAACT-3′) and subcloned into the pART7 vector (18) between the cauliflower mosaic virus (CaMV) ^{35}S constitutive promoter and the 3′ octopine synthase transcriptional termination (ocs) region. To generate pART27-*Oak1*, the *Oak1* expression cassette incorporating the *Oak1* coding sequence and the flanking promoter and termination regions was removed from pART7 as a single NotI fragment and ligated directly into the T-DNA region of the binary vector pART27 (18).

Transient Expression of Oak1 in N. benthamiana—Agrobacterium tumefaciens C58C1 cells transformed with the pART27-*Oak1* plasmid were grown in Luria Broth (LB) media containing 10 mM MES, 20 mM acetosyringone, 100 μg/ml spectinomycin, and 20 μg/ml kanamycin at 28 °C. When the A_{600} reached ~0.6 the cells were pelleted (2000 × *g*, 3 min, 4 °C) and resuspended in an equal volume of MES buffer (10 mM $MgCl_2$, 10 mM MES, pH 5.6) supplemented with 200 μM acetosyringone. The suspension was maintained at room temperature for 2–3 h before use. *N. benthamiana* leaves were infiltrated with the agrobacterium suspensions (agroinfiltration) as described (19).

Detection and Quantification of Cyclotide Proteins using Matrix-assisted Laser Desorption Ionization Time Of Flight Mass Spectrometry (MALDI-TOF-MS)—Leaf extracts (1 mg/ml or 0.1 mg/ml as indicated) were prepared in 50% CH_3CN (acetonitrile), 0.1% trifluoroacetic acid, and combined 1:1 (v/v) with the 4700 Proteomics Analyzer Calibration mixture (1/400 dilution in 50% CH_3CN, 0.1% trifluoroacetic acid; Applied Biosystems) containing a selection of proteins of known mass and concentration. This solution was combined 1:1 (v/v) with CHCA (α-cyano-4-hydroxycinnamic acid) matrix (5 mg/ml in 50% CH_3CN, 0.1% trifluoroacetic acid with 5 mM ammonium phosphate) and spotted (0.6 μl) onto a 192-well sample plate (Applied Biosystems). Mass analysis was carried out on the 4700 Proteomics Analyzer (Applied Biosystems) operated in positive ion reflector mode. Accelerating voltage was set at 20,000 V, the grid voltage was set at 68% of the accelerating voltage, and a 200-ns delay time was used. The low mass gate was set at 500 Da, and data were acquired between 1000 and 5000 Da. 50 spectra at 20 positions were accumulated per spot. Data were analyzed on the accompanying 4000 series Explorer Software.

For relative quantification, the ACTH (18–39 clip) protein from the calibration mixture (average $M+H^+$ of 2,466.72) was used as an internal standard relative to which the amount of cyclotide proteins was measured. The relative quantity of the cyclotide proteins was expressed as the ratio of the summed areas of the first three isotopic peaks of the analytes (cyclotide proteins) to the internal standard. To account for variations in protein expression between plants the ratios were expressed as a percentage of the total relative area of all the cyclotide proteins detected.

Characterization of Cyclotide Proteins Expressed in N. benthamiana—Soluble extracts from *N. benthamiana* leaves 3 days postinfiltration (dpi) with pART27-*Oak1* in agrobacterium were prepared in 50% CH_3CN, 0.1% trifluoroacetic acid, dried, and reduced in 0.1 M NH_4HCO_3, pH 8.5, with 10 mM TCEP (tris(2-carboxyethyl)phosphine) at 55 °C under nitrogen for 30 min. Reactions were quenched with formic acid and desalted using C_{18} ZipTips (Millipore) for analysis by MALDI-TOF MS.

Proteins extracted from *N. benthamiana* leaves 6 dpi in MeOH were purified by reverse phase-high performance liquid chromatography (RP-HPLC) and analyzed by MALDI-TOF MS. Co-elution studies of kalata B1 isolated from *O. affinis* and the 2892 Da protein from *N. benthamiana* were performed on C_{18} at 33% buffer B (90% CH_3CN, 0.05% trifluoroacetic acid) in 0.05% trifluoroacetic acid. Proteins were reduced in 100 mM TCEP for 10 min at 65 °C and 1 h at 37 °C and incubated with trypsin (40 μg/ml in 0.1 M NH_4HCO_3, pH 8.5) for 2 h. The solutions were desalted using ZipTips and analyzed by MALDI-TOF MS. MS/MS fragmentation of the 2914.9 Da precursor ion was performed on the MALDI-TOF 4700 Proteomics Analyzer (Applied Biosystems) operated in positive ion reflector mode. The masses of the y-series (C-terminal) fragment ions were used to determine the partial sequence of the 2892 Da protein.

In planta Inhibitor Assay—Ac-YVAD-CHO (1 mM) or a water (mock) control was infiltrated into *N. benthamiana* leaves 1 h prior to infiltration of the same leaf area with agrobacterium solutions harboring the *Oak1* expression construct, as described above. Eighteen hours after agroinfiltration the same leaves were again infiltrated with the protease inhibitor, and then repeatedly at 24-h intervals for the next 2 days. Infiltrated leaf tissue was collected 3 days after agroinfiltration. Infiltration of Ac-YVAD-CHO and water was conducted on individual leaves of the same plants ($n = 3$). For each treatment the leaf samples from the three different plants were pooled for analysis.

Oak1 Expression in VPE-silenced Plants—Virus-induced gene silencing (VIGS) of *VPE* genes in *N. benthamiana* was carried out with the potato virus X vector construct pPVX:VPE1a-2 described in Ref. 20 and the empty pPVX vector (21) as a control. The viral vectors were transformed into *A. tumefaciens* strain GV3101 and grown on LB-agar supplemented with tetracyclin (20 μg/ml) and kanamycin (50 μg/ml) at 28 °C for 2 days. Inoculation of 2-week-old *N. benthamiana* plants was conducted by piercing the growing leaves with a toothpick covered in the agrobacterium cells. The plants were kept under short day conditions (8 h light/16 h dark) at 22 °C for 3 days and then transferred to long day conditions (16 h light/8 h dark) at 26 °C. Six plants were inoculated with each viral vector. Four weeks after inoculation two upper leaves on each plant were agroinfiltrated with the *Oak1* expression construct as described above and maintained in long day conditions.

RT-PCR for Detection of VPE Expression—Total RNA was prepared using the RNeasy plant RNA extraction kit (Qiagen). Reverse transcriptase (RT)-PCR was carried out using 1 μg of RNA and Ready-to-go RT-PCR beads (Amersham Biosciences) with an oligo-dT primer for cDNA synthesis (Invitrogen). Gene expression was compared using 1 μl of cDNA as a template for PCR with VPE, Oak1, and actin (control) specific primer sets. The VPE (NbVPE common) and actins primer sets have been described previously (20). The EcoRI-MAKF and SLAA*BamHI primers described above were used to amplify Oak1.

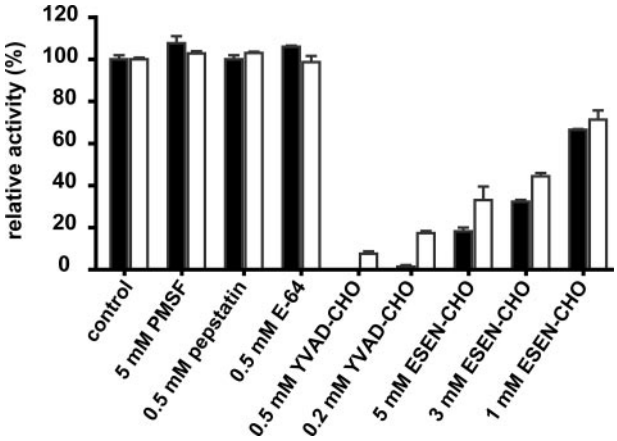

FIGURE 2. **Asparaginyl hydrolysis activity in the leaves of cyclotide-containing plants.** Leaf extracts from *O. affinis* (*black*) and *V. tricolour* (*white*) were preincubated with a range of protease inhibitors before hydrolysis of asparaginyl bonds was assayed using the fluorogenic substrate Z-AAN-mca. Activity detected without the addition of protease inhibitors (control) was standardized to 100%. The standard deviation from two independent experiments is shown.

Immunoblot Analysis of Oak1 Protein Expression—Immunoblot analysis was carried out essentially as described (22). Soluble leaf extracts were prepared in 50% CH_3CN, 0.1% trifluoroacetic acid, and the total protein concentration was standardized using the BCA (bicinchoninic acid) protein assay (Pierce). Aliquots of the samples were fractionated by SDS-PAGE and electrophoretically transferred to nitrocellulose membrane (Amersham Biosciences). The blotted membrane was incubated with polyclonal antibodies raised against recombinant Oak1 (1 mg/ml diluted 5000-fold). A 1/10,000-fold dilution of the Oak1 antibody was used in the immunoblot of VIGS plant extracts. Horseradish peroxidase-conjugated goat antibodies against rabbit IgG (diluted 20,000-fold; Sigma Aldrich) were used as secondary antibodies. Proteins were visualized with an enhanced chemiluminescence kit (Amersham Biosciences).

RESULTS

In Vitro AEP Activity—Protease activity capable of hydrolyzing asparaginyl bonds was assessed in representative cyclotide-containing Rubiaceae and Violaceae plants using the short peptide substrate Z-AAN-mca. Hydrolysis of the asparaginyl bond was detected in both *O. affinis* (Rubiaceae) and *V. tricolour* (Violaceae) leaf extracts as an increase in fluorescence at 465 nm over time, following addition of the substrate. To determine which class of protease(s) was responsible for this activity the cleavage assay was conducted after the plant extracts had been preincubated with a variety of inhibitors, each targeting a different subset of protease activities. As shown in Fig. 2 phenylmethylsulfonyl fluoride, pepstatin, and E-64 had no effect on the ability of either the *O. affinis* or *V. tricolour* leaf extracts to cleave the Asn-containing substrate, but the caspase-1 inhibi-

tor Ac-YVAD-CHO and the specific AEP inhibitor Ac-ESEN-CHO inhibited hydrolysis in a concentration-dependent manner. The inhibition profile observed is characteristic of AEPs, which belong to the CD clan of cysteine proteases (23, 24) and indicated that AEP activity was responsible for Asn-bond hydrolysis in cyclotide-containing plants.

AEP activity was also tested against cyclotide substrates. Cyclic kalata B1 (KB1-c), linear kalata B1 (KB1-l) and a synthetic protein corresponding to linear kalata B1 plus the C-terminal tail region (KB1-tail) were incubated with a commercial preparation of AEP from Jackbean (*Canavalia ensiformis*) seeds (JB AEP) in the presence and absence of dithiothreitol. JB AEP was used because it has been shown to catalyze peptide ligation during the processing of the concanavalin A precursor (25–27). Although JB AEP was active *in vitro* it only cleaved asparaginyl bonds after reduction (and unfolding) of the substrates (see supplemental data Fig. S1 and methods).

Circular Kalata B1 Is Produced by N. benthamiana—Cyclotide-containing plants are not ideal systems for the study of cyclization. The plants are not amenable to genetic manipulation, they endogenously accumulate very high levels of cyclotides (up to 2 mg/g wet plant weight; (28)) that mask small changes in expression, and the processing events appear to occur too rapidly to allow detection of the precursor protein and protein intermediates. Alternatively we expressed Oak1 in *N. benthamiana*, a plant that does not endogenously produce cyclotides and is not known to produce circular proteins, but which contains two characterized AEPs (20). We infiltrated *A. tumefaciens* harboring the pART27 binary vector encoding Oak1 (pART27-*Oak1*) under the control of a ^{35}S CaMV constitutive promoter and *ocs* terminator into *N. benthamiana* leaves. Mature cyclotide production was detected at 24-h intervals using MALDI-TOF MS.

Expression of Oak1 in *N. benthamiana* produced a suite of proteins with masses corresponding to cyclic oxidized kalata B1 and what appeared to be linear forms of the cyclotide not commonly observed in *O. affinis*. None of these masses were observed in uninfiltrated controls. As outlined in Fig. 3*A*, the masses detected correspond to linear kalata B1, linear kalata B1 minus the N-terminal Gly (−G) and linear kalata B1 plus the successive addition of the Gly, Leu, Pro, and Ser residues that constitute the C-terminal tail of Oak1 (+G, +GL, +GLP etc.). A MALDI-TOF mass spectrum of *N. benthamiana* leaves 3 dpi is shown before and after treatment with the reducing agent TCEP in Fig. 3*B*. Reduction caused the mass of each species to increase by 6 Da, indicating that each contained 6 Cys residues and was most likely an oxidized protein variant of kalata B1.

RP-HPLC and mass spectrometry were used to establish that the 2892 Da mass species identified in *N. benthamiana* corresponds to backbone-cyclized kalata B1. As shown in Fig. 4 the 2892 Da protein from *N. benthamiana* co-eluted with kalata B1 extracted from *O. affinis* on RP-HPLC (Fig. 4*A*), yielded the same linear product after digestion with trypsin (Fig. 4*B*) and had the same MS/MS fragmentation pattern as the control protein (Fig. 4*C*). RP-HPLC was also used to characterize the 2910 Da species, which corresponds in mass to both linear oxidized kalata B1 and an N-terminally truncated −G/+G protein. Two peaks of approximately equal area containing the 2910 Da mass

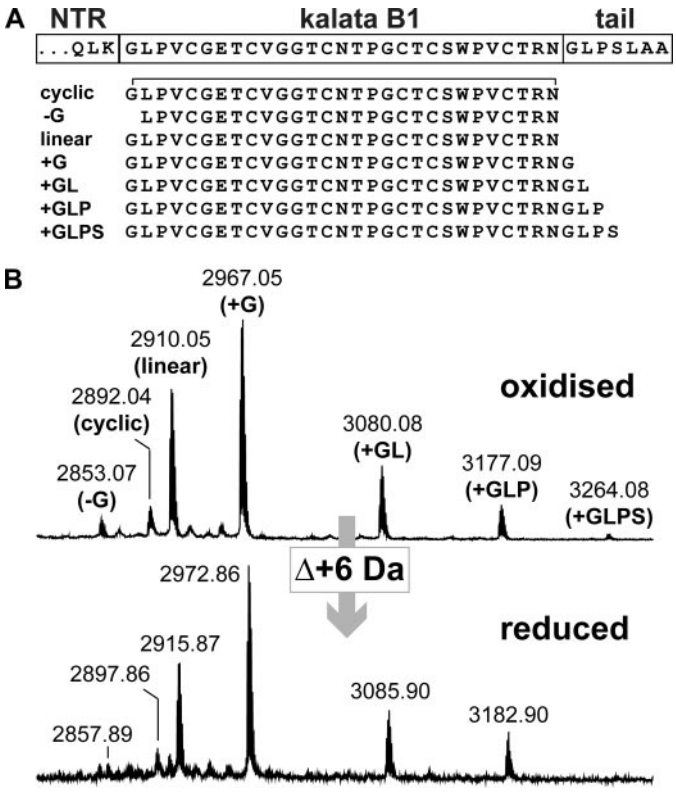

FIGURE 3. **Cyclotide production in *N. benthamiana*.** *A*, amino acid sequences of the putative cyclotide proteins produced in *N. benthamiana* based on their masses. *B*, MALDI-TOF mass spectra of a *N. benthamiana* leaf extract prepared 3 dpi with the Oak1 expression construct in agrobacterium before and after reduction with TCEP. The species are labeled according to *A*.

were isolated from an *N. benthamiana* (6 dpi) extract by RP-HPLC. One of these co-eluted with synthetic oxidized linear kalata B1 while the other contained both −G (2853 Da) and 2910 Da mass species in approximately equal proportions. Consequently we estimated that, at most, the ratio of the linear to −G/+G species in *N. benthamiana* is 2:1. The 2910 Da mass is hereinafter referred to as linear (−G/+G).

A Time Course Analysis of Cyclization—We monitored the progress of cyclotide production in *N. benthamiana* from 1–7 dpi. Fig. 5 shows the relative abundance of the cyclotide protein species at 24-h intervals. At 2 dpi the most abundant species present is +GLP. The linear (−G/+G), +G and +GL species are also in high abundance but cyclic kalata B1 and the N-terminal −G species are only present in low amounts. Over the next several days the sequential tail species disappear as the amount of −G and linear (−G/+G) increases. Although the amount of kalata B1 increases slightly during this period it does not increase in proportion to the decrease observed in the linear form. This suggests that *in vivo* linear kalata B1 (which makes up a majority of the 2910 Da mass peak) is being converted into the −G form rather than cyclic kalata B1 and thus that the Asn residue is not itself sufficient for cyclization.

VPE Inhibition Abolishes Circular Kalata B1 Production—The effect of AEP activity on Oak1 processing *in vivo* was tested using the caspase-1 inhibitor Ac-YVAD-CHO. Leaves of *N. benthamiana* were infiltrated with Ac-YVAD-CHO or a mock water control 1 h prior to agroinfiltration of the same

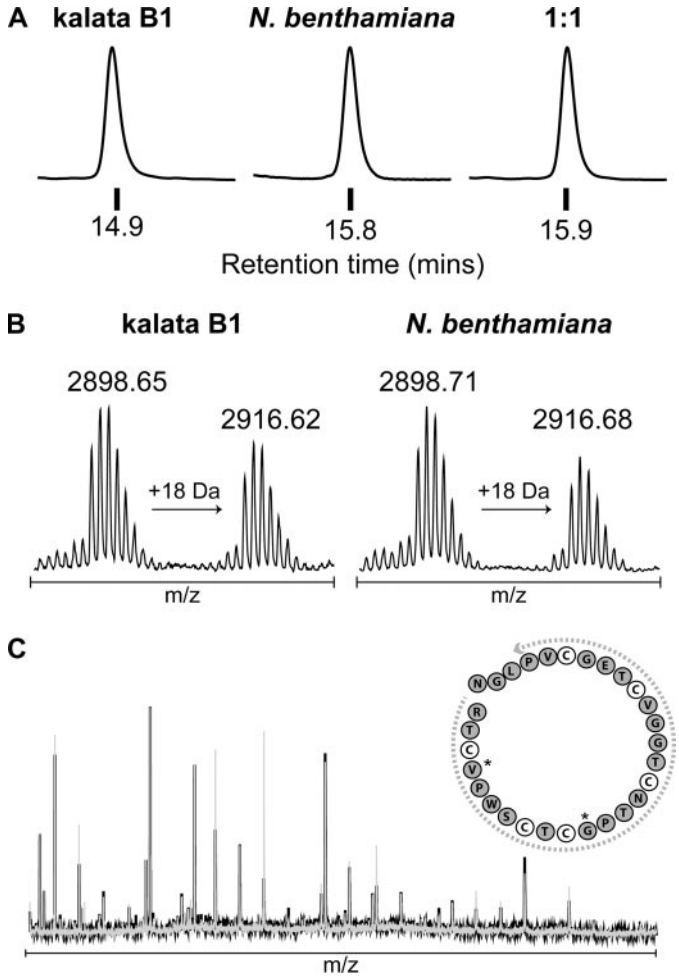

FIGURE 4. **Circular kalata B1 production in *N. benthamiana*.** *A*, co-elution of the 2892 Da protein produced in *N. benthamiana* with native kalata B1 from *O. affinis*. Each protein, as well as a 1:1 mixture of the two was analyzed by RP-HPLC. Each trace is shown at full scale. *B*, trypsin digestion of the 2892 Da protein and native *O. affinis* kalata B1. The proteins were reduced in TCEP and incubated with trypsin. A product 18 Da heavier than the starting material was observed following the incubation. *C*, MS/MS sequencing of the trypsin-digested proteins. The fragmentation fingerprints of both the 2892 Da protein (*black*) and native *O. affinis* kalata B1 (*gray*) are shown overlaid. The primary structure of kalata B1 is represented in the *inset* using the one letter amino acid code. The *dashed arrow* indicates the region of both proteins that was sequenced from the C terminus using y-series fragment ions. *Asterisks* denote fragment ions that were not detected.

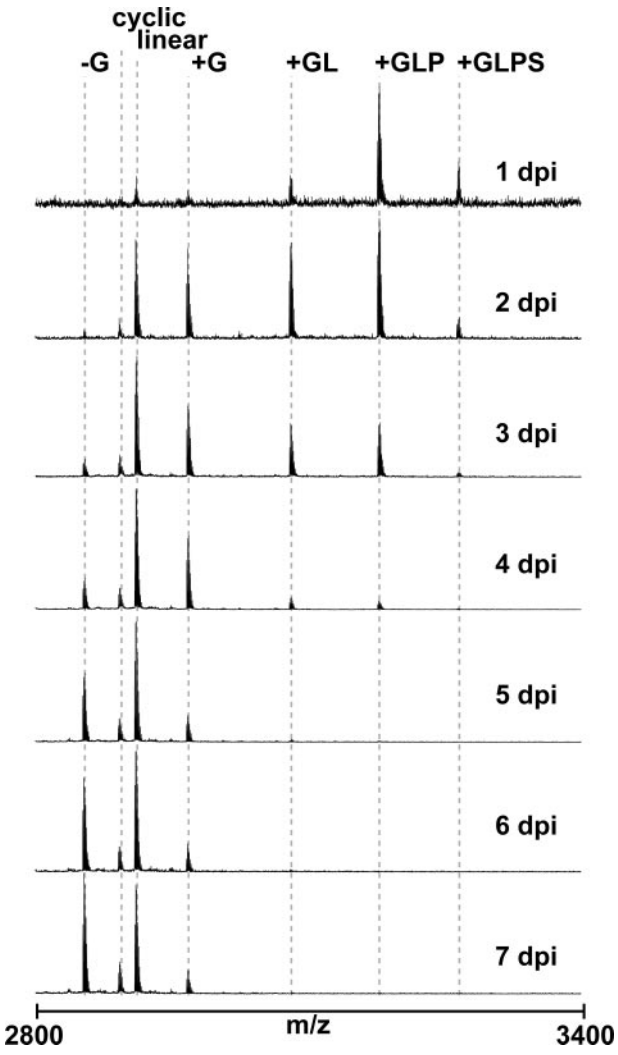

FIGURE 5. **Time course analysis of cyclotide protein production in *N. benthamiana*.** Soluble protein extracts were prepared from *N. benthamiana* leaves 1–7 dpi with pART27-*Oak1* in *A. tumefaciens* and adjusted to 1 mg/ml total protein. MALDI-TOF MS was used to identify kalata B1 proteins. The mass peaks corresponding to the different kalata B1 forms are labeled as in Fig. 3*A*. In each spectrum the most intense peak across the mass range shown is set at 100% relative intensity.

leaves with the pART27-*Oak1* construct. After 18 h the protease inhibitor was re-applied to the leaves and again at 24-h intervals for the next 2 days before the leaf samples were collected. For each treatment infiltrated leaf material from three plants was pooled for analysis. Protein extracts of the treated leaves were prepared in 50% CH_3CN, 0.1% trifluoroacetic acid and subjected to immunoblotting with Oak1 antibodies to examine Oak1 processing and MALDI-TOF MS to assess cyclotide production. Similar levels of the Oak1 precursor and two lower mass intermediates were detected in the Ac-YVAD-CHO treated and control leaves, as shown in Fig. 6*A*, suggesting that initial processing events were not perturbed by the inhibitor. By contrast, the production of cyclotide proteins differed substantially between the samples (Fig. 6*B* and expanded in 6*C*). Cyclic kalata B1 was not detected in leaves infiltrated with Ac-YVAD-CHO and the distribution of the linear cyclotide proteins was changed relative to the control, causing a shift away from the short linear and +G proteins toward accumulation of the longer +GLP species.

Silencing of VPE in N. benthamiana Decreases Production of Circular Kalata B1—Virus-induced gene silencing (VIGS) was used to generate *N. benthamiana* plants in which expression of the major forms of AEPs, termed VPEs in *N. benthamiana*, was suppressed. Silencing was targeted toward a region of cDNA from *N. benthamiana VPE-1a*, which is conserved in the other *N. benthamiana* AEP, *VPE-1b* (20). Plants were inoculated with agrobacterium carrying the region of VPE cDNA encoded in an antisense orientation in the potato virus X vector (pPVX:VPE1a-2). Control plants were inoculated with the empty viral vector pPVX and appeared phenotypically identical to the silenced plants (Fig. 7*A*). Four weeks after inoculation, leaves of both the silenced and control plants were agroinfiltrated with the Oak1 expression vector.

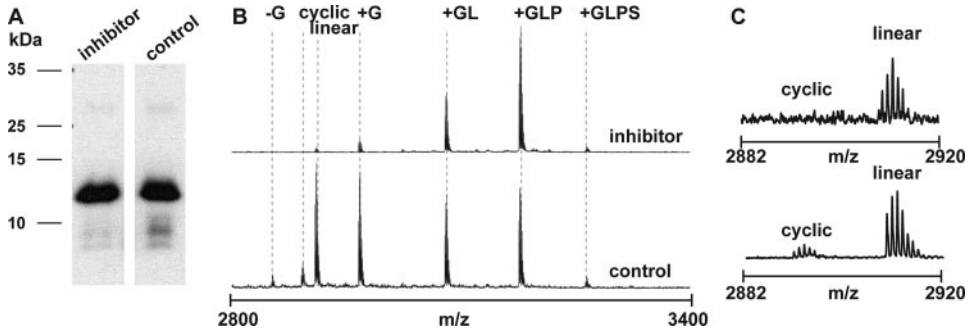

FIGURE 6. **Precursor processing and cyclotide production in *N. benthamiana* leaves infiltrated with an AEP inhibitor.** *N. benthamiana* leaves expressing Oak1 were repeatedly infiltrated with the Ac-YVAD-CHO inhibitor or a water control over the course of 3 days. Soluble leaf extracts (1 mg/ml) from three plants were subjected to immunoblotting (15 μg) with Oak1 antibodies (*A*) and MALDI-TOF MS (*B*). The kalata B1 cyclotide protein masses were detected in each treatment but at different distributions. Kalata B1 species are identified ± flanking residues. An expanded portion of the trace encompassing the mass range of linear and cyclic kalata B1 is shown in *C*. No cyclic kalata B1 was detected in inhibitor infiltrated leaves.

Total RNA was extracted from infiltrated leaves and subject to RT-PCR using a primer set targeting a region common to both *N. benthamiana* AEPs. Weak *VPE* gene expression was detected in the silenced plants but at lower levels than the expression detected in the control plants (Fig. 7*B*). To determine whether the decrease in *VPE* expression was affecting precursor production, *Oak1* infiltrated leaf samples 3 dpi were subject to immunoblot analysis with antibodies against Oak1. Fig. 7*C* shows Oak1 protein production in the six control plants compared with the six silenced plants. A stark increase or decrease in the amount of precursor was not characteristic of either category.

To determine whether decreased VPE expression affects cyclotide production we compared the amount of kalata B1 in leaf extracts from VIGS plants 3 days after agroinfiltration of *Oak1*, using MALDI-TOF MS. For each plant the area of the cyclic kalata B1 mass peaks was calculated relative to an internal standard protein (relative area), the ionization of which we have found to be directly proportional to kalata B1 (data not shown). Total cyclotide protein expression levels (total relative area) did not vary significantly between the silenced and control plants but a significant ($p = \leq 0.0001$) decrease in cyclic kalata B1 production (>40%) was observed in the silenced plants compared with the controls (Fig. 7*D*). As in the inhibitor studies, a decrease in cyclic kalata B1 was accompanied by a relative increase in the abundance of the longer linear cyclotide protein species (Fig. 7*E*).

DISCUSSION

In this study we report that AEP activity is required for the cyclization of cyclotides. Our findings suggest that AEP catalyzes protein backbone cyclization by coupling asparaginyl bond hydrolysis at the C terminus of the cyclotide domain with peptide bond ligation. These results provide an initial insight into the little understood but common phenomenon of protein cyclization in eukaryotes. Naturally occurring circular proteins are known to be particularly stable (29, 30) and there is evidence that cyclization can improve the stability and function of normally acyclic proteins under physiological conditions, thereby improving their applicability as drugs (31, 32). Knowledge of *in vivo* biosynthesis of cyclotides therefore has the potential to impart these promising features to other potential drug leads as part of bioengineering initiatives.

Proteolytic activity capable of hydrolyzing asparaginyl bonds was analyzed in cyclotide-containing plants and found to have the same characteristics as AEP activity. We therefore concluded that specific hydrolysis of asparaginyl bonds in cyclotide-containing plants was catalyzed by AEPs. Despite displaying strict specificity toward asparaginyl and, to a lesser extent, aspartyl bonds AEPs do not target all Asn (Asp) residues in a substrate (33, 34). For example *A. thaliana* γVPE will cleave the fluorogenic substrates Ac-ESEN-mca and Ac-YVAD-mca but not Ac-ESED-mca or Ac-DEVD-mca (35). Given these differences it was not unexpected that JB VPE was unable to cleave cyclotide substrates in their native oxidized state. Instead we focused on *in vivo* approaches to analyzing cyclization.

N. benthamiana does not endogenously produce cyclotides and is not known to naturally produce other circular proteins. Despite this, circular kalata B1 was detected in *N. benthamiana* following the expression of the Oak1 precursor in the leaves. Our findings confirmed that the protein produced in *N. benthamiana* contained a cyclized backbone with the same disulfide connectivity and identical chromatographic properties as the native protein. Previous studies of kalata B1 folding pathways (6, 36) have shown that the structure and disulfide connectivity of cyclotide derivatives affect their retention times on RP-HPLC. The co-elution of native oxidized kalata B1 with kalata B1 produced in *N. benthamiana* under isocratic conditions therefore indicates that the cyclotide produced in *N. benthamiana* has the same three-dimensional fold as the native protein. The ability of *N. benthamiana* to produce correctly folded circular proteins demonstrated that protein backbone cyclization is not a mechanistically unique process limited to cyclotide-containing plants but rather that it has parallels to a process already carried out in the plant.

Unlike the production of kalata B1 in *O. affinis*, where only the circular protein is observed, expression of Oak1 in *N. benthamiana* yields circular kalata B1 as well as a series of proteins comprising the linear cyclotide domain, or the linear cyclotide domain minus an N-terminal Gly, plus C-terminal residues from the tail region (GLPSLAA). All of the species identified were in the oxidized state, indicating that formation of the cystine knot occurs prior to C-terminal processing. From 1–7 dpi a progressive shift toward the shorter proteins was observed, consistent with sequential C-terminal hydrolysis or trimming of the longer linear proteins. Although it is not clear that C-terminal trimming is involved in the endogenous production of cyclotides, the incidence of the processing in *N. benthamiana* provided us with a valuable opportunity to analyze cyclization *in vivo*. By following the evolution of each cyclotide species over several days (Fig. 5) it became evident

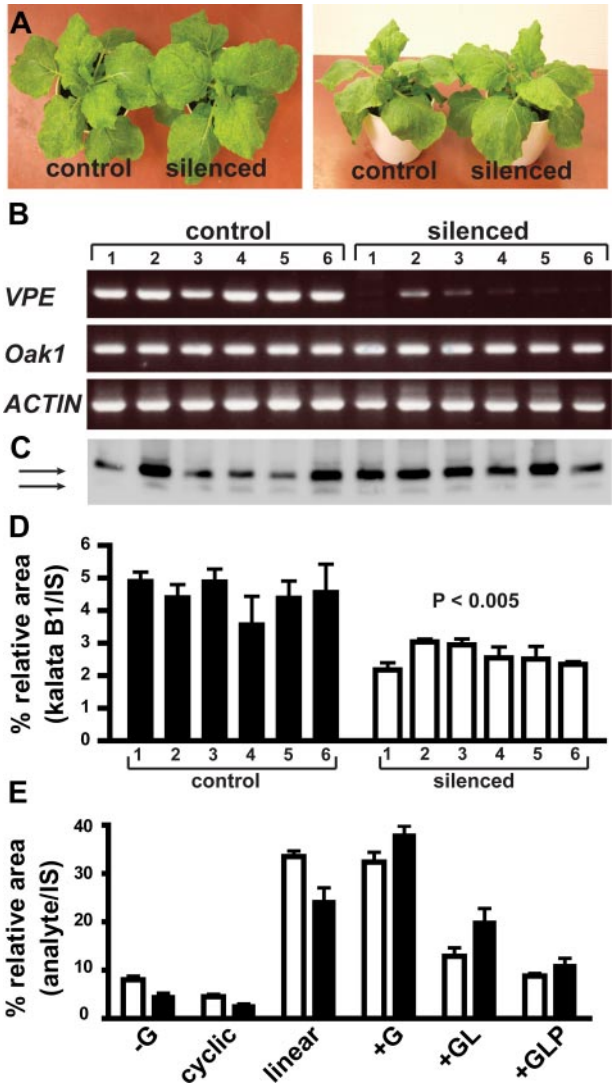

FIGURE 7. **Silencing of AEP activity decreases cyclic kalata B1 production.** VIGS was used to knock down *VPE* gene expression in *N. benthamiana* plants. *A*, AEP-silenced and control plants 4 weeks after inoculation with the PVX vectors. *Top* and *side view*. *B*, RT-PCR analysis of relative mRNA levels of *VPE*, *Oak1*, and *ACTIN* (loading control) in leaves of VIGS plants agroinfiltrated with the *Oak1* construct. *C*, Oak1 immunoblot analysis of leaf extracts (20 μg) at 3 dpi. *Arrows* show the Oak1 precursor protein and a lower molecular weight intermediate. *D* and *E*, quantification of cyclotide protein production in AEP-silenced and control plants. Leaf extracts (0.1 mg/ml) were prepared from each plant ($n = 6$) at 3 dpi and sampled by MALDI-TOF MS in duplicate. *D*, relative area of kalata B1 in each plant ± S.D. Cyclic kalata B1 production was significantly lower in the AEP-silenced plants ($p < 0.0001$). Significance ($p < 0.05$) was calculated using a Student's *t* test. *IS*, internal standard. *E*, average production of each cyclotide species produced in the silenced and control plants, represented as a percentage of total cyclotide protein expression (% relative area).

that the shorter linear (−G/+G) and −G species were being produced from the longer linear forms at a much faster rate than the production of circular kalata B1. This seemed to indicate that the shorter species, including linear kalata B1, were not acting as substrates for cyclization. Indeed at 3 dpi linear kalata B1 was accumulating in the leaves at a level that far exceeded the amount of circular kalata B1 produced and at a much faster rate than it was possible to convert it to the aberrant −G form. This suggested that the asparaginyl bond is integral to the cyclization process and that cyclization is linked to the hydrolysis of this bond.

Alternatively, it is possible that linear kalata B1 is the substrate for cyclization but that in *N. benthamiana* it is consumed by an aberrant process before cyclization can take place. Indeed, the removal of the N-terminal Gly appears to be such an aberrant processing event, producing a protein form (−G) that cannot be cyclized. However, the time course data does not show that the N-terminal trimming process is competing with cyclization for the linear substrate. If this were the case, the linear form would be unlikely to accumulate at high levels in the system and we would expect the −G form to be produced before the circular form. This is because the N-terminal processing event would necessarily have to out compete the cyclization reaction to account for the high accumulation of −G over time. Instead the circular form appears before the −G form in the time course and does not increase in proportion to the increase in the −G form, suggesting that the cyclization event occurs prior to N-terminal trimming.

Both the application of the caspase-1 inhibitor Ac-YVAD-CHO directly into leaves and VIGS of AEPs caused a decrease in the levels of cyclic kalata B1 produced upon expression of Oak1 in *N. benthamiana*. The inhibition of AEP activity achieved by the protease inhibitor abolished cyclic kalata B1 production and caused a shift in the relative abundance of the linear cyclotide proteins toward the longer +GLP species. Taking into account the competing C-terminal trimming process observed in *N. benthamiana* it would appear that the substrate for cyclization is at least as long as the +GLP species. VIGS directed toward the cDNA region of *NbVPE-1a* resulted in a decrease in *VPE* mRNA levels indicative of partial suppression of *VPE* expression in the silenced plants. Accordingly, a moderate decrease in the level of cyclic kalata B1 produced in the silenced plants was observed together with a less pronounced shift toward the longer linear cyclotide species. Upstream processing of the Oak1 precursor in *N. benthamiana* was not perturbed by either treatment, confirming that AEP activity was restricted to the final stages of cyclotide processing.

It cannot be discounted that a decrease in AEP activity elicits a decrease in cyclic kalata B1 production by an indirect pathway. However, several findings point toward the direct involvement of an AEP in the cyclization process. The activity assay conducted in the presence of a range of enzyme inhibitors confirmed that as in *N. benthamiana* (20), asparaginyl bond hydrolysis in cyclotide-containing plants is carried out by an enzyme with characteristics of an AEP. Thus if cleavage of the C-terminal Asn bond of cyclotides occurs, it is most likely carried out by an AEP. This would account for the high conservation of this residue. An alternative possibility is that the conserved Asn serves to terminate processing from the C terminus by a carboxypeptidase and that cyclization occurs in a separate process. However, the observation that linear kalata B1 does not appear to be cyclized *in vivo* strongly implicates hydrolysis of the asparaginyl bond by an endopeptidase in the cyclization process. Finally, the accumulation of longer linear cyclotide species in plants that have decreased AEP activity, together with the persistence of the C-terminal trimming process, identifies the asparaginyl bond in cyclotides as the site of enzyme action.

JB VPE catalyzes protein ligation of the ConA precursor by coupling the reaction to asparaginyl bond hydrolysis (25, 27). In this

transpeptidation event the energy released from peptide bond hydrolysis is utilized for subsequent peptide bond formation. A similar process occurs in the only other reported example of post-translational peptide bond formation in eukaryotes, the re-arrangement of antigenic peptides by the proteasome (37–39). It is conceivable that cyclization of cyclotides occurs via the same mechanism except rather than ligating two peptides together, the ends of a single polypeptide chain are joined. We speculate that in the Ac-YVAD-CHO-treated and *VPE*-silenced plants, this transpeptidation was unable to occur. This allowed more of the extended intermediate form (cyclotide plus tail) into the C-terminal trimming pathway, causing a slight accumulation of the longer truncated forms. The mechanism of processing that occurs at the N terminus of the cyclotide domain to produce the extended linear cyclization substrate remains unclear but must occur prior to processing at the C-terminal Asn and cyclization.

In conclusion, cyclotides are a large and rapidly growing family of plant proteins (40). We have shown that AEPs are implicated in catalyzing cyclization in a post-translational event that appears to couple the cleavage of an asparaginyl bond in a linear intermediate with peptide bond formation. Efforts are currently underway to identify the AEP involved in cyclization from a cyclotide-containing plant. The ability of an enzyme to catalyze the reverse of its normal reaction when presented with a cyclizable substrate would provide a mechanism for the evolution of circular proteins, driven by their higher intrinsic stability compared with linear counterparts. The recent discovery of linear ancestors to cyclotides in economically important cereal crops (41, 42) suggests that simple mutations that introduce appropriately located Asn/Asp residues are sufficient to drive the cyclization process, as opposed to wholesale changes to the biochemical machinery of organisms. Indeed the identification of a pseudogene in humans whose peptidic product prevents HIV infection (43), and is homologous to circular antimicrobial defensins in monkeys (30) suggests that circular proteins are probably more common than previously thought. Potentially, any protein with termini in close proximity is cyclizable. The current findings serve as the basis for new applications of protein cyclization, with its attendant advantages, to the bioengineering of proteins of medical and agricultural importance.

Acknowledgments—We thank Roger Mitchell for assistance with agroinfiltrations and Shane Simonsen for synthesizing KB1-tail.

REFERENCES

1. Craik, D. J. (2006) *Science* **311,** 1563–1564
2. Trabi, M., and Craik, D. J. (2002) *Trends Biochem. Sci.* **27,** 132–138
3. Colgrave, M. L., and Craik, D. J. (2004) *Biochemistry* **43,** 5965–5975
4. Craik, D. J., Cěmažar, M., and Daly, N. L. (2006) *Curr. Opin. Drug Discov. Devel.* **9,** 251–260
5. Saether, O., Craik, D. J., Campbell, I. D., Sletten, K., Juul, J., and Norman, D. G. (1995) *Biochemistry* **34,** 4147–4158
6. Daly, N. L., Love, S., Alewood, P. F., and Craik, D. J. (1999) *Biochemistry* **38,** 10606–10614
7. Gran, L. (1970) *Medd. Nor. Farm. Selsk.* **12,** 173–180
8. Gustafson, K. R., Sowder, R. C. I., Henderson, L. E., Parsons, I. C., Kashman, Y., Cardellina, J. H. I., McMahon, J. B., Buckheit, R. W. J., Pannell, L. K., and Boyd, M. R. (1994) *J. Am. Chem. Soc.* **116,** 9337–9338
9. Lindholm, P., Göransson, U., Johansson, S., Claeson, P., Gulbo, J., Larsson, R., Bohlin, L., and Backlund, A. (2002) *Mol. Cancer Ther.* **1,** 365–369
10. Tam, J. P., Lu, Y. A., Yang, J. L., and Chiu, K. W. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 8913–8918
11. Göransson, U., Sjogren, M., Svangard, E., Claeson, P., and Bohlin, L. (2004) *J. Nat. Prod.* **67,** 1287–1290
12. Witherup, K. M., Bogusky, M. J., Anderson, P. S., Ramjit, H., Ransom, R. W., Wood, T., and Sardana, M. (1994) *J. Nat. Prod.* **57,** 1619–1625
13. Jennings, C., West, J., Waine, C., Craik, D., and Anderson, M. (2001) *Proc. Natl. Acad. Sci. U. S. A.* **98,** 10614–10619
14. Jennings, C. V., Rosengren, K. J., Daly, N. L., Plan, M., Stevens, J., Scanlon, M. J., Waine, C., Norman, D. G., Anderson, M. A., and Craik, D. J. (2005) *Biochemistry* **44,** 851–860
15. Dutton, J. L., Renda, R. F., Waine, C., Clark, R. J., Daly, N. L., Jennings, C. V., Anderson, M. A., and Craik, D. J. (2004) *J. Biol. Chem.* **279,** 46858–46867
16. Kimura, R. H., Tran, A.-T., and Camarero, J. A. (2006) *Angew. Chem. Int. Ed. Engl.* **45,** 973–976
17. Muntz, K., Blattner, F. R., and Shutov, A. D. (2002) *J. Plant Physiol.* **159,** 1281–1293
18. Gleave, A. P. (1992) *Plant Mol. Biol.* **20,** 1203–1207
19. Goodin, M. M., Dietzgen, R. G., Schichnes, D., Ruzin, S., and Jackson, A. O. (2002) *Plant J.* **31,** 375–383
20. Hatsugai, N., Kuroyanagi, M., Yamada, K., Meshi, T., Tsuda, S., Kondo, M., Nishimura, M., and Hara-Nishimura, I. (2004) *Science* **305,** 855–858
21. Chapman, S., Kavanagh, T., and Baulcombe, D. (1992) *Plant J.* **2,** 549–557
22. Harlow, E., and Lane, D. (1999) *Using Antibodies: A Laboratory Manual*, pp. 267–309, Cold Spring Harbor Laboratory Press, New York
23. Barrett, A. J., Kembhavi, A. A., Brown, M. A., Kirschke, H., Knight, C. G., Tamai, M., and Hanada, K. (1982) *Biochem. J.* **201,** 189–198
24. Rawlings, N. D., Morton, F. R., and Barrett, A. J. (2006) *Nucleic Acids Res.* **34,** D270–272
25. Bowles, D. J., Marcus, S. E., Pappin, D. J., Findlay, J. B., Eliopoulos, E., Maycox, P. R., and Burgess, J. (1986) *J. Cell Biol.* **102,** 1284–1297
26. Carrington, D. M., Auffret, A., and Hanke, D. E. (1985) *Nature* **313,** 64–67
27. Min, W., and Jones, D. H. (1994) *Nat. Struct. Mol. Biol.* **1,** 502–504
28. Gran, L. (1973) *Lloydia* **36,** 174–178
29. Cobos, E. S., Filimonov, V. V., Galvez, A., Valdivia, E., Maqueda, M., Martinez, J. C., and Mateo, P. L. (2002) *Biochim. Biophys. Acta-Proteins Proteomics* **1598,** 98–107
30. Tang, Y. Q., Yuan, J., Osapay, G., Osapay, K., Tran, D., Miller, C. J., Ouellette, A. J., and Selsted, M. E. (1999) *Science* **286,** 498–502
31. Clark, R. J., Fischer, H., Dempster, L., Daly, N. L., Rosengren, K. J., Nevin, S. T., Meunier, F. A., Adams, D. J., and Craik, D. J. (2005) *Proc. Natl. Acad. Sci. U. S. A.* **102,** 13767–13772
32. Tam, J. P., Wu, C. W., and Yang, J. L. (2000) *Eur. J. Biochem.* **267,** 3289–3300
33. Dando, P. M., Fortunato, M., Smith, L., Knight, C. G., McKendrick, J. E., and Barrett, A. J. (1999) *Biochem. J.* **339,** 743–749
34. Jung, R., Scott, M. P., Nam, Y.-W., Beaman, T. W., Bassuner, R., Saalbach, I., Muntz, K., and Nielsen, N. C. (1998) *Plant Cell* **10,** 343–358
35. Kuroyanagi, M., Yamada, K., Hatsugai, N., Kondo, M., Nishimura, M., and Hara-Nishimura, I. (2005) *J. Biol. Chem.* **280,** 32914–32920
36. Daly, N. L., Clark, R. J., and Craik, D. J. (2003) *J. Biol. Chem.* **278,** 6314–6322
37. Hanada, K., Yewdell, J. W., and Yang, J. C. (2004) *Nature* **427,** 252–256
38. Vigneron, N., Stroobant, V., Chapiro, J., Ooms, A., Degiovanni, G., Morel, S., van der Bruggen, P., Boon, T., and Van den Eynde, B. J. (2004) *Science* **304,** 587–590
39. Warren, E. H., Vigneron, N. J., Gavin, M. A., Coulie, P. G., Stroobant, V., Dalet, A., Tykodi, S. S., Xuereb, S. M., Mito, J. K., Riddell, S. R., and Van den Eynde, B. J. (2006) *Science* **313,** 1444–1447
40. Simonsen, S. M., Sando, L., Ireland, D. C., Colgrave, M. L., Bharathi, R., Goransson, U., and Craik, D. J. (2005) *Plant Cell* **17,** 3176–3189
41. Basse, C. W. (2005) *Plant Physiol.* **138,** 1774–1784
42. Mulvenna, J. P., Mylne, J. S., Bharathi, R., Burton, R. A., Shirley, N. J., Fincher, G. B., Anderson, M. A., and Craik, D. J. (2006) *Plant Cell* **18,** 2134–2144
43. Cole, A. M., Hong, T., Boo, L. M., Nguyen, T., Zhao, C., Bristol, G., Zack, J. A., Waring, A. J., Yang, O. O., and Lehrer, R. I. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99,** 1813–1818