JOURNAL OF AGRICULTURAL AND FOOD CHEMISTRY

Purification and Characterization of a Chimeric Cry1F δ-Endotoxin Expressed in Transgenic Cotton Plants

YONG GAO,* KRISTIN J. FENCIL, XIAOPING XU, DEBBIE A. SCHWEDLER, JEFFREY R. GILBERT, AND ROD A. HERMAN

Dow AgroSciences LLC, 9330 Zionsville Road, Indianapolis, Indiana 46268

Cotton plants were genetically modified through the introduction of a synthetic gene that encodes a *Bacillus thuringiensis* insecticidal protoxin referred to as Cry1F(synpro). This protoxin is a chimeric, full-length δ-endotoxin of 130 kDa, comprised of the core toxin of Cry1Fa2 protein and parts of the nontoxic portions of Cry1Ca3 and Cry1Ab1 proteins, all of which originated from *Bacillus thuringiensis*. The Cry1F(synpro) expressed in cotton plants confers resistance to lepidopteran pests. The current study was conducted to characterize the Cry1F(synpro) protein expressed in the transgenic cotton event 281-24-236. Results showed that the full-length Cry1F(synpro) produced in the transgenic cotton plants was sensitive to the host cell protease cleavage, resulting in a truncated, biologically active form (core toxin) with an apparent molecular mass of 65 kDa. This truncated toxin was purified by immunoaffinity chromatography from the cotton leaf extract. N-terminal sequencing, peptide mass fingerprinting by MALDI-TOF MS, and internal peptide sequencing by MS/MS confirmed the identity of the truncated core toxin of Cry1F. The mechanism of truncation was explored with Cry1F(synpro) derived from a recombinant *Pseudomonas fluorescens*. The transgenic cotton-produced Cry1F showed equivalent insecticidal activity to that of *Pseudomonas fluorescens*-derived Cry1F.

KEYWORDS: Transgenic cotton; insect resistance; Cry1F(synpro); truncation; *Bacillus thuringiensis*; protein purification and characterization

INTRODUCTION

The tools of genetic engineering allow plant biotechnologists to isolate and select genes that produce desired traits and move the genes from one species to another to produce a new crop variety or cultivar. Genetically modified (GM) crops were first introduced commercially in 1996. In the 9 year period from 1996 to 2004, the global area of GM crops increased 48-fold, from 1.7 million hectares (ha) in 1996 to 81 million ha in 2004 (*1*). On a global basis, 28% of the global 32 million ha of cultivated cotton plants were GM in 2004 (*1*). The dominant traits for GM cotton are insect resistance and herbicide tolerance.

To provide a new transgenic tool for insect resistance in cotton, Dow AgroSciences LLC developed GM cotton lines which express two chimeric, full-length insecticidal δ-endotoxins, referred to as Cry1F(synpro) and Cry1Ac(synpro) (synpro = synthetic protoxin), derived from the common soil bacterium *Bacillus thuringiensis* (Bt). The two cotton lines, Cry1F event 281-24-236 and Cry1Ac event 3006-210-23, were first developed separately and then stacked by conventional breeding processes to produce a combined trait product with the trade name of WideStrike (*2*). Large-scale field trials have demonstrated that WideStrike varieties provide season-long control of a broad spectrum of lepidopteran pests, such as tobacco budworm (*Heliothis virescens*), cotton bollworm (*Helicoverpa zea*), and pink bollworm (*Pectinophora gossyiella*) (*2*). Characterization of transgenes and transgenic proteins is of significant importance to the development and safety assessment of transgenic crops (*3*). The purpose of this study was to purify and characterize the transgenic Cry1F protein expressed in the transgenic cotton plants.

MATERIALS AND METHODS

Cry1F Cotton Event 281-24-236. Many Bt δ-endotoxin protoxin molecules are composed of two segments. The protease-resistant core toxin is the N-terminal segment and corresponds to about the first half of the protein molecule. The second, C-terminal half of the molecule is believed to participate in the protoxin crystal formation (*4*). The core toxin segment and C-terminal segments of various Cry1 toxins can be combined together to create chimeric or fusion toxins for insect control (*5*).

The C-terminal domain of Cry1Ab was reported to be associated with an increased solubility and expression of the protoxin (*6*, *7*). A chimeric endotoxin with a Cry1F core toxin and Cry1Ab C-terminal protoxin segment would take advantage of the insecticidal activity of the Cry1F and the elevated solubility and expression of Cry1Ab. Thus, a plant-optimized *cry*1F(synpro) gene was synthesized encoding a chimeric δ-endotoxin with a sequence of 1148 amino acids (*8*). The corresponding chimeric protoxin is called Cry1F(synpro), of which amino acids 1−604 are from Cry1Fa2 (GenBank, accession number AAA22347), 605−640 from Cry1Ca3 (accession number AAA22343),

* To whom correspondence should be addressed. Phone: (317) 337-5129. Fax: (317) 337-3255. E-mail: ygao2@dow.com.

10.1021/jf0523869 CCC: $33.50
Published on Web 01/14/2006

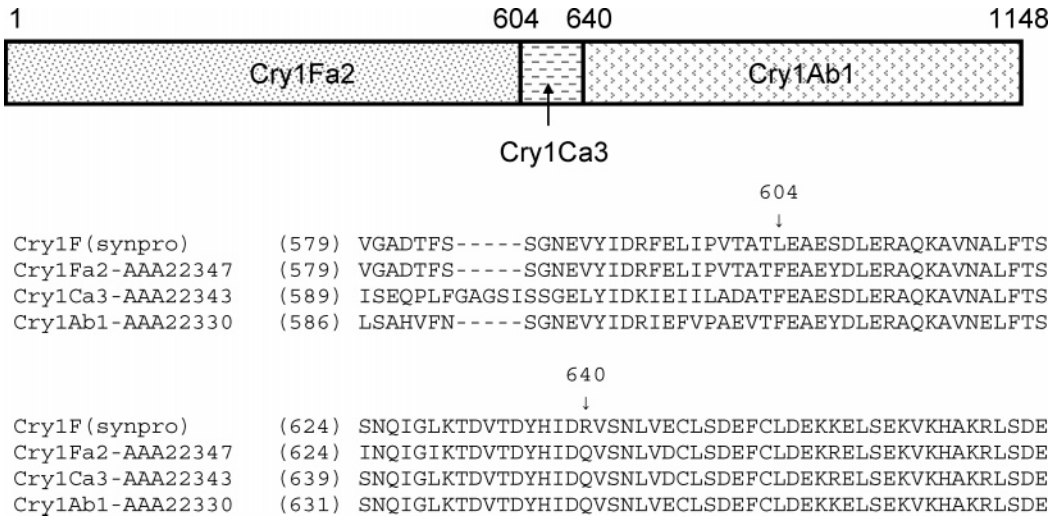

Figure 1. Sequence origin of the chimeric Cry1F(synpro) (top diagram) and partial sequence alignment of Cry1F(synpro), Cry1Fa2, Cry1Ca3, and Cry1Ab1.

and 641−1148 are from Cry1Ab1 (accession number AAA22330) (**Figure 1**). No linker sequence was introduced although two amino acid substitutions were made: no. 604 (F to L) and no. 640 (Q to R) (**Figure 1**). The purpose of making these substitutions was to introduce *Xho* I and *Pvu* I restriction sites in the DNA sequence to facilitate cloning. The no. 604 amino acid substitution (F to L) was based on the fact that a leucine (L) is a homologous site in related Cry1 sequences, such as Cry1Ac1 (accession number AAA22331).

The transgenic cotton (*Gossypium hirsutum*) event 281-24-236 was produced via *Agrobacterium*-mediated transformation of an Acala cotton line (*8*). A binary Ti plasmid vector pAGM281 was constructed (*8*), which is 14 950 bp in total length. The T-DNA insert portion of this vector, which includes the *cry*1F and *pat* genes and their regulatory components, is 8034 bp. Specifically, the T-DNA region of this vector contained two gene expression cassettes organized in a tail-to-tail orientation and sharing one bidirectional termination and polyadenylation sequence. The first cassette contained the coding sequence for Cry1F(synpro), and the transcription of which was controlled by a regulatory element consisting of the mannopine synthase (Δ-mas 2′) promoter from *Agrobacterium tumefaciens* pTi15955 (*9*) and four copies of the octopine synthase enhancer (4OCS) from pTiAch5 (*10*). Polyadenylation sequences were derived from the bidirectional open reading frame-25 (ORF25) terminator from pTi15955 (*9*).

The second cassette of the vector contained a synthetic, plant-codon-optimized version of the *pat* gene which was introduced as a selectable marker for the identification of plant transformants. The *pat* gene, originally isolated from *Streptomyces viridochromogenes* Tu494 (*11*), encodes the phosphinothricin N-acetyltransferase (PAT) which catalyzes the acetylation of herbicide phosphinothricin and confers tolerance to this herbicide (used as a selectable marker). Transcription of the *pat* gene was controlled with a regulatory element consisting of the maize ubiquitin 1 promoter plus exon1 and intron1 (ZmUbi1) (*12*). As with the Cry1F(synpro) encoding sequences, polyadenylation sequences were derived from the bidirectional ORF25 terminator from pTil5955.

Protein Extraction from Cotton Leaves. Leaves of the transgenic cotton were collected from Dow AgroSciences' field station in Woodland, CA, and kept frozen under −20 °C until processed. The extraction process was performed in a cold room maintained at approximately 4 °C. The frozen tissue was homogenized with a stainless steel blender in 3 vol of an extraction buffer containing 50 mM CAPS (pH 10.0), 100 mM NaCl, 2 mM EDTA, 2 mM dithiothreitol, 0.05% (w/v) poly(vinylpyrrolidone), 5 mM ascorbic acid, 0.2 mM Pefabloc SC (Boerhinger Mannheim), and 1.1 μM leupeptin. The extracts were clarified by filtering through two layers of cheesecloth. A 10% (w/v) polyethylenimine solution was then added slowly with stirring to a final concentration of 0.15% (w/v). The solution was further clarified by centrifugation at 10 000*g* for 30 min. Crystalline ammonium sulfate was slowly added to the supernatant to 50% saturation with gentle stirring. The precipitate was collected by centrifugation and suspended in a small volume of the extraction buffer. The resulting solution was dialyzed overnight against 0.5 M ammonium bicarbonate. After dialysis, the material was lyophilized and stored under refrigerated conditions until further analysis or purification. The percentage of the Cry1F toxin to the lyophilized powder weight was 0.053% determined by an enzyme linked immunosorbent assay (ELISA) kit manufactured by Strategic Diagnostics Inc. (Newark, DE).

Bioassay of Insecticidal Activity. Insecticidal activity of the transgenic cotton extract was examined by a diet-overlay bioassay as described elsewhere (*3*). A negative control material was prepared from a nontransgenic, near-isogenic cotton line to assess any background effects. In addition, a recombinant Cry1F(synpro) produced from *Pseudomonas fluorescens* (Pf) was included in the bioassay to examine the biological equivalency between the cotton-produced and Pf-derived Cry1F protein. The Cry1F concentrations in the test materials were determined with ELISA. Three insect pests were selected based on their varied susceptibilities to the Cry1F δ-endotoxin. Tobacco budworm (TBW), *Heliothis virescens*, was chosen as a highly susceptible pest, beet armyworm (BAW), *Spodoptera exigua*, as a susceptible pest, and cotton bollworm (CBW), *Helicoverpa zea*, as a relatively less susceptible pest.

The test and control substances were suspended in 10 mM potassium phosphate buffer (pH 7.5) and then applied to the surface of the artificial diet in 128-well bioassay trays (*13*). Approximately 500 μL of diet was included in each well with an approximate surface area of 1.5 cm^2. Fifty microliters of each formulated test or control material was applied to each of 16 wells on each of the two bioassay dates. The surface of the diet was allowed to dry before infesting each well with a single neonate. The wells were covered with vented lids provided by the bioassay trays. The assay trays were held in a growth chamber of 26 °C after infestation. Mortality and insect-weight data were collected after 6 days.

GI_{80} values (concentration estimated to induce 80% growth inhibition) and 95% confidence limits were calculated by regressing the logarithm of the concentration against the probit of the growth inhibition (*14*). Growth inhibition was calculated based on the average total weight of insects in the control treatments. Where effects on insect growth were seen with the controls, these controls and the corresponding treatments (concentrations) were excluded from the analysis to avoid confounding effects from naturally occurring insecticidal compounds (matrix effects) found in cotton tissue (*15*).

Immunoaffinity Chromatography. Further purification of Cry1F protein from the lyophilized transgenic cotton leaf extract was conducted by immunoaffinity chromatography. The immunoaffinity media was generated by coupling Cry1F-specific rabbit polyclonal antibodies to CNBr-activated Sepharose 4 FF gel (Amersham Biosciences) and was packed into a disposable column. The bed volume of the antibody-

coupled affinity column was approximately 4 mL with a ligand density of 7 mg/mL of the media.

The lyophilized cotton extract (454 mg) was suspended in 45 mL of 20 mM CAPS buffer, pH 10.5. The supernatant was retained after a brief centrifugation. The pH of the supernatant was adjusted to approximately neutral (7−7.5) with 0.2 M citric acid. The supernatant was then diluted by mixing with 1 vol of 20 mM sodium phosphate, pH 7.2, and loaded onto the column which was preequilibrated with a loading buffer (20 mM sodium phosphate, 0.5 M NaCl, and 0.05% Tween 20, pH 7.0). After loading, the column was washed with 17 mL of loading buffer, followed by elution with 12 mL of a McIlvaine buffer of pH 3.2 (prepared by mixing 377 mL of 0.1 M citric acid with 123 mL of 0.2 M dibasic sodium phosphate). Fractions of 1 mL each were collected (40 μL of 1 M Tris base was added to each fraction in advance to neutralize the pH of fractions) and concentrated with Microcon units (MW cutoff 10 kDa, Millipore) to about 150−200 μL each. The purity and immunoreactivity of the Cry1F protein in the fractions were examined by SDS−PAGE and Western blot analysis.

Reference Cry1F Protein. Reference Cry1F protein was generated with a proprietary bacterial expression system of *Pseudomonas fluorescens* (Pf). The full-length Cry1F(synpro) was expressed in the bacterial cells as inclusion bodies. After fermentation of the recombinant strain, the cells were homogenized and inclusions were extracted and washed with 50 mM Tris buffer, pH 7.5, containing 200 mM NaCl, 10% glycerol, 20 mM EDTA, 0.5% Triton X-100, and 1 mM DTT. The inclusions were then dissolved in 25 mM CAPS buffer (pH 9.0) containing 1 mM DTT. The dissolved Cry1F(synpro) was further purified by an anion exchange chromatography (Q-Sepharose). A truncated version of Cry1F was also produced by treating Cry1F(synpro) with trypsin. The truncation reaction was carried out in a microcentrifuge tube containing 100 μL of 10 mM CAPS buffer (pH 10.5), 80 μg Cry1F(synpro), and 20 μg of the trypsin (Sigma, catalog no. T-7168). The reaction cocktail was incubated in a 37 °C water bath for 60 min, and at the end of the incubation, trypsin was deactivated by adding 0.5 M freshly prepared phenylmethylsulfonyl fluoride stock solution to a final concentration of 10 mM. After the trypsinolysis, free amino acids and small peptides (<5 kDa) were removed from the solution using Econo-Pac 10 DG desalting columns (Bio-Rad).

SDS−PAGE and Western Blotting. SDS−PAGE was performed as described previously (*3*). For Western blot analysis, the proteins, after gel electrophoresis, were electroblotted onto nitrocellulose membranes. A specific rabbit polyclonal antibody against Cry1F was used as the primary antibody (Strategic Diagnostics Inc., Newark, DE). A conjugate of goat anti-rabbit IgG (H + L) alkaline phosphatase (Pierce) was used as the secondary antibody. A substrate solution containing 100 mM Tris, 100 mM NaCl, 5 mM $MgCl_2$, 0.025% 5-bromo-4-chloro-3-indolyl phosphate, and 0.05% *p*-nitroblue tetrazolium chloride was used for colorimetric development of the immunoreactive protein bands.

Detection of Glycosylation. The purified Cry1F samples were further separated by SDS−PAGE. After electrophoresis, identical gels were stained either with coomassie brilliant blue (CBB) to visualize all protein bands, or with GelCode glycoprotein staining kit (Pierce) to visualize glycoproteins. The procedure for glycoprotein staining has been described elsewhere (*3*).

N-Terminal Sequencing. The purified preparations of the Cry1F protein were separated from any minor impurities by SDS−PAGE. The proteins were then electroblotted onto a PVDF membrane in 10 mM CAPS (pH 11.0) containing 20% MeOH. Proteins on the membrane were stained with GelCode blue stain (Pierce), and the membrane background was destained briefly with 50% (v/v) MeOH and 10% (v/v) acetic acid. The N-terminal sequences were determined using a Procise protein sequencer (model 494, Applied Biosystems).

Tryptic Peptide Mass Fingerprinting by MALDI-TOF MS. The purified Cry1F protein samples were denatured by heating at 100 °C for 5 min, followed by SDS−PAGE. The respective bands were excised from the gel and digested with sequencing grade trypsin (Roche Diagnostics, Indianapolis, IN) for approximately 16 h at 37 °C. Mass spectral analyses were performed on a PerSeptive Biosystems (Framingham, MA) Voyager-DE STR MALDI-TOF mass spectrometer as described previously (*3*).

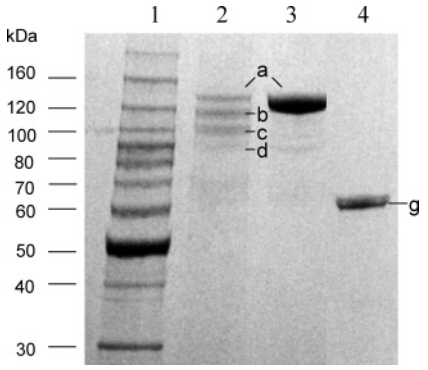

Figure 2. SDS–PAGE (4–15% gradient) of *P. fluorescens*-produced Cry1F protein. **Lane 1**: Invitrogen BenchMark protein ladder. **Lane 2**: recombinant *P. fluorescens*-derived Cry1F(synpro) incubated in CAPS buffer of pH 10.5 for 4 days at 4 °C, 2 μg. **Lane 3**: freshly prepared Cry1F(synpro), 2 μg. **Lane 4**: trypsin-truncated Cry1F(synpro), 1.5 μg.

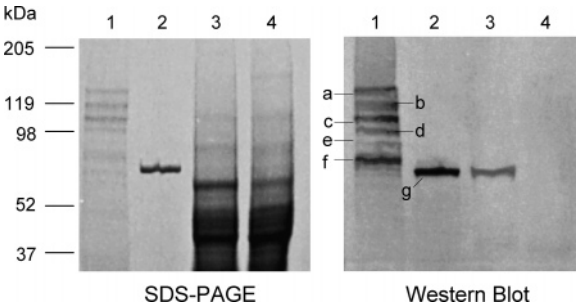

Figure 3. SDS–PAGE (4–15% gradient) and Western blot of *P. fluorescens*-produced and cotton-derived Cry1F proteins. **Lane 1**: recombinant *P. fluorescens*-derived Cry1F(synpro) incubated in CAPS buffer of pH 10.5 for 4 days at 4°C. **Lane 2**: trypsin-truncated Cry1F from *P. fluorescens*. **Lane 3**: transgenic cotton leaf extract. **Lane 4**: nontransgenic cotton leaf extract.

Internal Peptide Sequencing by MS/MS. Selective internal sequencing of peptides generated by trypsin digestion was performed using a Micromass Q-TOF (quadrupole time-of-flight) mass spectrometer as previously described (*3*).

RESULTS AND DISCUSSION

Truncation of Cry1F(synpro). The expected molecular mass of Cry1F(synpro), based on amino acid sequence, is 130 kDa. With recombinant Pf-derived Cry1F, it was demonstrated that the apparent molecular mass of the full-length Cry1F(synpro) was approximately 130 kDa, as expected (**Figure 2**, **lane 3**). The full-length Cry1F(synpro) was very sensitive to protease cleavage once dissolved under high pH conditions. When the Cry1F(synpro) preparation was incubated in 20 mM CAPS buffer (pH 10.5) in a 4 °C refrigerator for 2−4 days, the full-length Cry1F(synpro) band gradually degraded into several partially truncated forms with molecular mass ranging from approximately 70 to 120 kDa on SDS−PAGE (**Figure 2**, **lane 2**) and Western blot (**Figure 3**, **lane 1**). Apparently this is due to the action of trace amounts of host cell proteases that are present in the Cry1F(synpro) preparation. When treated deliberately with trypsin, the full-length and the various forms of partially truncated Cry1F were transformed into a core toxin of approximately 65 kDa in size (**Figure 2**, **lane 4**).

Insecticidal Activity. GI_{80} (80% growth inhibition) values were calculated for the cotton-produced and recombinant Pf-derived Cry1F(synpro). Because only a limited amount of the cotton extract could be incorporated into the insect diet in the bioassay tray, the Cry1F concentration from the cotton materials

Table 1. Potency Comparison of Transgenic Cotton-Produced and Recombinant *Pseudomonas fluorescens*-Derived Cry1F Protein

	Pf-produced Cry1F[a]		cotton-produced Cry1F	
insect	GI_{80} (ng ai/cm^2)[b,c]	(95% CL)[d]	GI_{80} (ng ai/cm^2)[b,c]	(95% CL)[d]
cotton bollworm	237	(99–565)	>353[e]	>104
beet armyworm	43	(26–72)	80	(21–302)
tobacco budworm	5.1	(2.7–9.3)	2.5	(1.9–3.4)

[a] Pf = *Pseudomonas fluorescens*. [b] GI: growth inhibition calculated based on the reduction in total live insect weight in each treatment compared with negative control (nontransgenic cotton extract or buffer) for each test. [c] ai = active ingredient. Cry1F concentration was normalized to the core toxin level. [d] CL = confidence limits. [e] Highest concentration without matrix effects.

was not sufficiently high to allow lethal-concentration estimates (LC_{50}). Therefore, only GI_{80} was derived from this bioassay. The growth-inhibition results clearly demonstrated that the Cry1F protein from both sources (cotton plants and recombinant Pf) appeared to have similar potencies against the three insect species tested (**Table 1**). Both materials are most active against tobacco budworm, followed by beet armyworm and cotton bollworm. The low concentration of the Cry1F protein in the cotton extract and the matrix effects [growth inhibition on insects by cotton intrinsic compounds (*15*)] at high concentrations of cotton extract did not allow a precise estimate of the GI_{80} for cotton bollworm but did allow this insect to be categorized as the least susceptible insect pest tested. It is noted that cotton plants contain a unique group of terpenoids including desoxyhemigossypol, hemigossypol, gossypol, hemigossypolone, and the heliocides that are part of the plant's defense system against pathogenic fungi and insects (*16*, *17*). These compounds might have contributed to the matrix effect observed at high concentrations in the insect bioassay (*15*).

Cry1F Purification from Transgenic Cotton. The Cry1F protein molecular mass profile in fresh cotton tissue was difficult to detect with Western blot due to the low expression level of the protein (3−15 μg/g dry weight in terminal leaves). The need to concentrate and fractionate the extract might alter the Cry1F forms even though several protease inhibitors were included in the extraction buffer. Due to the high sensitivity of Cry1F(synpro) to proteases, one would expect that the expressed Cry1F(synpro) in cotton plants could be truncated or partially truncated by the plant cell proteases. Indeed, the truncated core toxin was detected in the concentrated extract of cotton leaves by Western blot, which had the equivalent molecular size as the truncated Cry1F derived from recombinant Pf (**Figure 3**, **lane 3**).

Immunoaffinity chromatography was conducted on the cotton leaf extract to purify the truncated Cry1F. Elution fractions (1 mL each) were collected from the elution effluent and concentrated to about 150−200 μL. All fractions (nos. 1−12) were examined by SDS−PAGE and Western blot analysis with a Cry1F protein-specific polyclonal antibody. The result showed that fractions nos. 6−12 contained a major protein band at approximately 65 kDa (**Figure 4**), and Western blot analysis confirmed that this major protein band was immunoreactive to the specific polyclonal antibody against Cry1F.

N-Terminal Sequencing. The amino acid residues of various forms of Cry1F (full-length, partially truncated, and fully truncated core toxin) were sequenced. The Edman degradation reaction was performed for the first 5 or 10 cycles depending on protein quantities. The N-terminal sequences obtained from the full-length Cry1F (**band a**, **Figure 2**) and various forms of

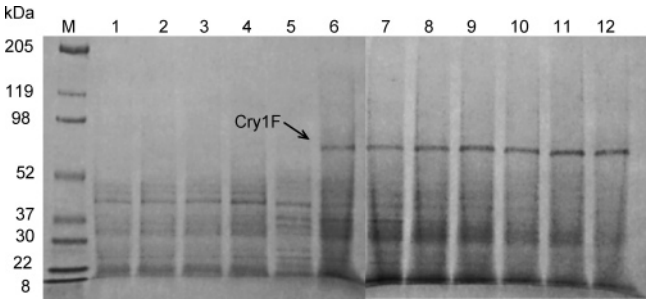

Figure 4. SDS–PAGE (4–15% gel gradient) of the immunoaffinity chromatography fractions of cotton-derived Cry1F(synpro). The numbers above each lane indicate the fraction numbers. The marker lane (M) contained the prestained broad-range molecular weight standards from Bio-Rad.

Table 2. N-Terminal Sequencing

protein samples	N-terminal sequence[a]
Cry1F(synpro), expected sequence	MENNIQNQCVPYNCLNNPEVEILNEER-[28]STGRLPLDIS
Pf Cry1F, full-length, band a[b]	MENNI
Pf Cry1F, partially truncated, band b	MENNI
Pf Cry1F, partially truncated, band c	MENNI
Pf Cry1F, partially truncated, band d	MENNI
Pf Cry1F, partially truncated, band f	XENNI
Pf Cry1F, fully truncated, band g	STGRLPLDIS
cotton Cry1F, naturally truncated	XTGRLPLDIS

[a] X indicates an uncertain amino acid assignment during the sequencing test. [b] Band assignment (a, b, c, d, f, g) refers to **Figures 2** and **3**. Pf = *Pseudomonas fluorescens*.

partially truncated Cry1F (**bands b−d**, and **f**, **Figure 3**) produced by Pf matched the theoretical N-terminal sequence of Cry1F(synpro) (**Table 2**). The quantity of **band e** was too low to be sequenced (**Figure 3**). The 10 amino acid residues determined from the fully truncated Cry1F (**band g**, **Figure 3**) of both Pf (truncated by trypsin) and transgenic cotton (truncated by cotton endogenous proteases) were found to correspond to residues no. 28 through no. 37 of Cry1F(synpro) (**Table 2**).

These results suggest that upon exposure of full-length Cry1F(synpro) to serine proteases such as trypsin, truncation occurs from the C-terminal domain initially, removing peptides in 10−20 kDa segments, resulting in various intermediate or partially truncated forms. Similar stepwise proteolysis of Cry1-type protoxins has been described previously by Choma et al. (*18*). Following complete truncation, almost half of the sequence from the chimeric Cry1F protoxin was cleaved from the C-terminal domain, and the first 27 amino acid residues at the N-terminus were also removed. The remaining 65 kDa protein is or is close to the core toxin form of Cry1F which, in its native state, is resistant to trypsin-like proteases.

It has been well documented that Cry1 toxins, such as Cry1Ab and Cry1Ac, are synthesized as inactive protoxins of around 130 kDa within the Bt cytoplasm (*19*). Upon ingestion by susceptible insects, the crystal protoxin is solubilized and activated through the removal of an N-terminal peptide of 25−30 amino acids and the cleavage of approximately half of the sequences from the C-terminus, resulting in the formation of core toxins of 60−70 kDa in size (*18*−*22*). The role of the C-terminal extension to the active toxin is believed to be in the formation of crystalline inclusion bodies within the Bt bacterium and is dispensable for toxicity (*4*). The 25−30 amino acid residues at the N-terminus play a role in promoting crystallization of the protoxin in the bacterium but do not contribute to

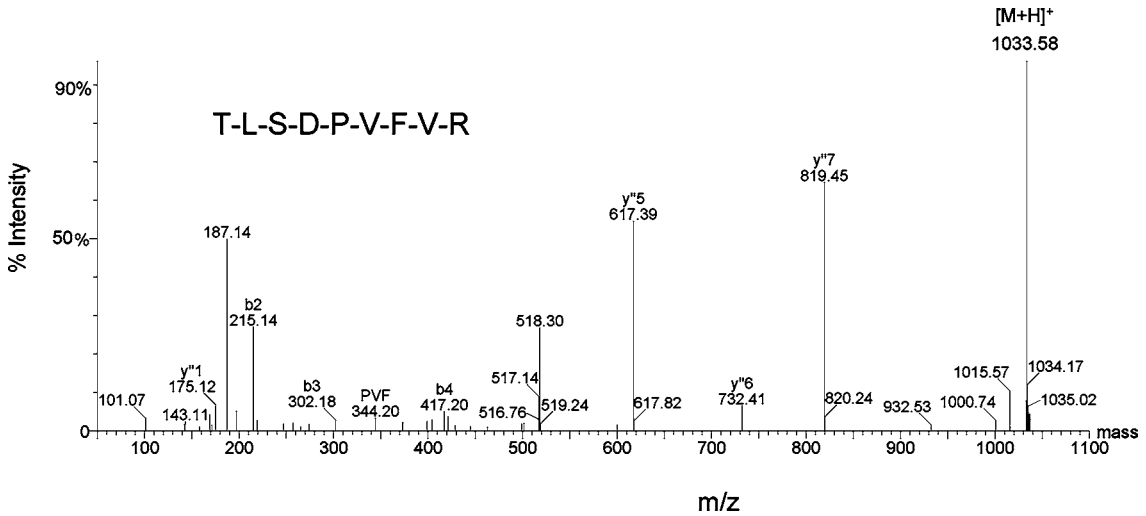

Figure 5. MS/MS spectrum of the proteolytic fragment *m*/*z* 517.3 $[M + 2H]^{2+}$ (*m*/*z* 1033.6 $[M + H]^{+}$) from the transgenic cotton-derived truncated Cry1F.

toxicity (*18*−*21*). Recent studies indicate that proteolytic removal of the N-terminal peptide, including processing of helix α1 of domain I, is an important part of the activation mechanism of the Cry1Ab toxin (*23*). In other words, the N-terminal peptide, even the helix α1 of domain I, is dispensable for Cry1Ab biological activity or toxicity. Bravo et al. recently compared the toxicity, binding, and pore-forming abilities of N-terminus-cleaved, fully activated Cry1Ac with those of an N-terminus-intact Cry1Ac and concluded that proteolytic removal of the N-terminal peptide is essential before the toxin becomes fully active (*21*).

The removal of N-terminal and C-terminal peptides is the result of enzymatic cleavage by proteases in the susceptible insect guts. The major proteases of the lepidopteran insect midgut are trypsin-like or chymotrypsin-like proteases. The activated toxin then binds to specific sites on the brush border membrane of the midgut epithelium of susceptible insects before inserting into the membrane and forming a pore. Apparently, on the basis of the results of the current study, the removal of N-terminal and C-terminal peptides could also happen in transgenic Bt plant cells, which was not unexpected since similar serine proteases are present in plants. The current understanding of the mode of action of Cry toxins involves solubilization of the Cry protoxins, proteolytic activation, receptor binding, membrane insertion, toxin oligomerization, and finally pore formation leading to cell swelling and lysis (*19*, *22*, *23*).

Tryptic Peptide Mass Fingerprinting. Although Cry1F core toxin is resistant to trypsin cleavage in its native state, it can be easily digested by trypsin once it is denatured by heating. Peptide mapping by MALDI-TOF MS following trypsin digestion is a powerful tool in protein identification. The Cry1F proteins derived from both Pf and the transgenic cotton were subjected to heat denaturation and digestion by trypsin. The resulting peptide mixture was analyzed by MALDI-TOF MS. The masses of the detected peptides were compared to those deduced from potential trypsin cleavage sites in the sequence of Cry1F-(synpro). The result showed that in the digest of Pf Cry1F-(synpro), 37 peptides were identified matching the theoretical deduced peptide masses of Cry1F(synpro). These peptides together cover 40% of the Cry1F(synpro) protein sequence. Mass spectra of the peptide mixture resulting from the digestion of a denatured protein by an enzyme (such as trypsin) provide a peptide mass fingerprint (PMF) of a protein with great specificity. In general, a protein identification made by peptide mass fingerprinting is considered to be reliable if the measured coverage of the sequence is 15% or higher with a minimum of five peptide matches (*24*). In the current study, the detected peptide coverage was 40% with 37 peptide matches, suggesting the recombinant Cry1F(synpro) produced by Pf was expressed correctly.

Table 3. Tryptic Peptide Mass Fingerprinting

Cry1F(synpro) residue no.	theoret mass (*m*/*z*)[a]	cotton truncated Cry1F (*m*/*z*)[a]
32–42	1227.72	1227.72
114–125	1441.68	1441.66
172–193	2434.15	2434.04
194–200	878.55	878.55
223–232	1408.73	1408.60
233–251	2177.18	2177.10
252–263	1394.72	1394.71
264–286	2509.21	2509.13
312–324	1413.71	1413.71
358–366	1033.56	1033.56
367–379	1386.72	1386.71
380–392	1416.69	1416.69
431–442	1376.62	1376.61
443–451	1132.53	1132.53
452–463	1301.63	1301.62
464–471	911.59	911.59
472–483	1269.68	1269.68
484–494	1089.56	1089.57
495–515	2289.20	2289.12
516–521	834.50	834.35
530–538	1007.55	1007.54
539–546	924.49	924.49

[a] *m*/*z*: mass-to-charge ratio.

In the trypsin digest of the transgenic cotton-derived truncated Cry1F protein, 22 peptides were identified matching the theoretical deduced peptide masses (**Table 3**). These peptides cover the sequences up to residue no. 546 of Cry1F(synpro). These data further confirm that the remaining sequence in the core toxin was from the N-terminal half of the full-length Cry1F-(synpro).

Internal Peptide Sequencing. The same trypsin-digested peptide mixtures of cotton-derived Cry1F used in the MALDI-TOF MS analysis were analyzed using MS/MS on a Micromass Q-TOF mass spectrometer. Positive electrospray ionization was used to ionize the peptides, and both MS and MS/MS spectra were acquired. Two $[M + 2H]^{2+}$, doubly charged precursor ions, at *m*/*z* of 517.3 (corresponding to the singly charged ion $[M + H]^{+}$ of *m*/*z* 1033.6) and 545.3 (corresponding to $[M + H]^{+}$ of

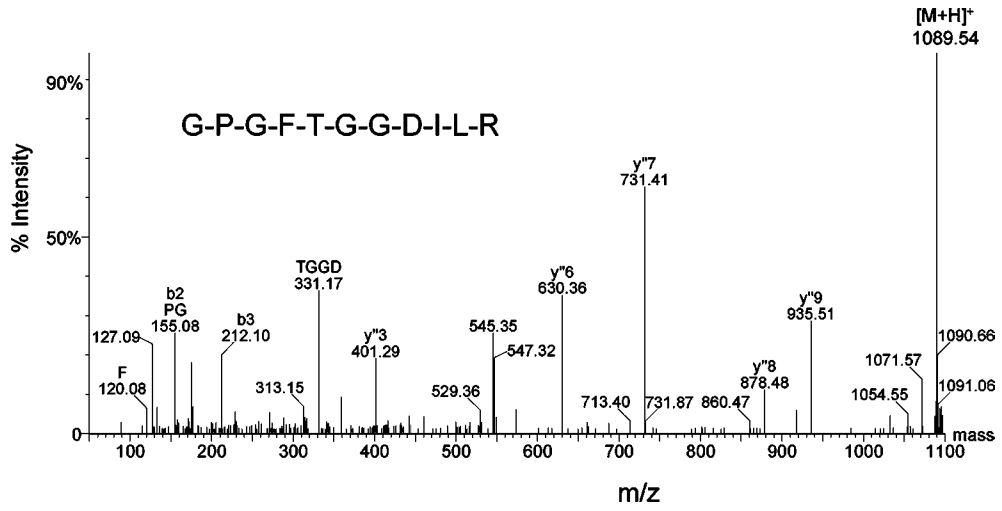

Figure 6. MS/MS spectrum of the proteolytic fragment *m*/*z* 545.3 $[M + 2H]^{2+}$ (*m*/*z* 1089.6 $[M + H]^{+}$) from the transgenic cotton-derived truncated Cry1F.

m/*z* 1089.6), were selected from the digest of the purified transgenic cotton-derived truncated Cry1F preparation. **Figure 5** shows the MS/MS spectra of the ion $[M + 2H]^{2+}$ *m*/*z* 517.3 and the predicted amino acid sequence derived from the MS/MS data. The derived peptide sequence, TLSDPVFVR, matched the theoretical sequence from residue nos. 358−366 of Cry1F-(synpro). **Figure 6** shows the MS/MS spectra of the ion $[M + 2H]^{2+}$ *m*/*z* 545.3 and the predicted amino acid sequence derived from the MS/MS data. The derived peptide sequence, GPGFTGGDILR, matched the theoretical sequence from residue nos. 484−494 of Cry1F(synpro).

Lack of Glycosylation. Detection of potential carbohydrates which might be covalently linked to the Cry1F protein expressed in cotton was assessed with the GelCode glycoprotein staining kit (Pierce). The immunoaffinity-purified transgenic cotton-derived truncated Cry1F protein preparation was separated by SDS−PAGE. A glycoprotein, horseradish peroxidase, was loaded as a positive indicator for glycosylation. A nonglycoprotein, soybean trypsin inhibitor, was employed as a negative control. This staining method is able to detect 0.625 ng of the glycoprotein avidin and 0.16 μg of horseradish peroxidase (Pierce technical information of the test kit). The amount of Cry1F loaded onto the gel was 1 μg (6-fold the detection limit for horseradish peroxidase). The results demonstrated that the cotton-derived Cry1F protein had no detectable carbohydrates.

In conclusion, it was demonstrated that a chimeric full-length Cry1F protoxin expressed in transgenic cotton plants could be truncated by the plant host proteases into a core toxin of approximately 65 kDa, which was equivalent to the truncated Cry1F generated by treating the recombinant *Pseudomonas fluorescens*-derived, chimeric, full-length Cry1F protoxin with trypsin in vitro. The truncation occurred initially from the C-terminal domain, at which 10−20 kDa peptide pieces were successively removed. Upon complete truncation, almost half of the Cry1F protoxin was cleaved from the C-terminal domain. In addition, 27 amino acid residues from the N-terminus were removed. These results indicate that the first two steps (i.e., solubilization of the toxin, proteolytic activation) of the mode of action for the Cry1F protoxin can occur in the transgenic cotton plants prior to ingestion by susceptible insects.

ABBREVIATIONS USED

Bt, *Bacillus thuringiensis*; CAPS, 3-(cyclohexylamino)-1-propanesulfonic acid; CBB, coomassie brilliant blue; DTT, dithiothreitol; EDTA, ethylenediamine tetraacetate; GM, genetically modified; *m*/*z*, mass-to-charge ratio; MALDI-TOF MS, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry; MeOH, methanol; MS/MS: tandem mass spectrometry; MW, molecular weight; Pf, *Pseudomonas fluorescens*; PVDF, polyvinyldifluoride; SDS−PAGE, sodium dodecyl sulfate−polyacrylamide gel electrophoresis; TFA, trifluoroacetic acid.

ACKNOWLEDGMENT

We thank Monty Badger and Kexue Huang for providing the reference Cry1F derived from recombinant *Pseudomonas fluorescens*. We are also grateful to Laura Tagliani, Ken Narva, Steve Evans, and Meibao Zhuang for their support and encouragement during the course of this study.

LITERATURE CITED

(1) James, C. *Global status of commercialized biotech/GM crops: 2004*; ISAAA Briefs 2004, No. 32; International Service for the Acquisition of Agri-biotech Applications (ISAAA): Ithaca, New York, 2004.

(2) Dow AgroSciences LLC. WideStrike insect protection. http://www.widestrike.com (accessed July 2005).

(3) Gao, Y.; Schafer, B. W.; Collins, R. A.; Herman, R. A.; Xu, X.; Gilbert, J. R.; Ni, W.; Langer, V. L.; Tagliani, L. Characterization of Cry34Ab1 and Cry35Ab1 insecticidal crystal proteins expressed in transgenic corn plants and *Pseudomonas fluorescens*. *J. Agric. Food Chem.* **2004**, *52*, 8057−8065.

(4) Park, H. W.; Bideshi, D. K.; Federici, B. A. Molecular genetic manipulation of truncated Cry1C protein synthesis in *Bacillus thuringiensis* to improve stability and yield. *Appl. Environ. Microbiol.* **2000**, *66*, 4449−4455.

(5) Cardineau, G. A.; Stelman, S. J.; Narva, K. E. Plant-optimized genes encoding pesticidal toxins. U.S. Patent No. 6,218,188 B1, 2001.

(6) Aronson, A. I.; Beckman, W.; Dunn, P. *Bacillus thuringiensis* and related insect pathogens. *Microbiol. Rev.* **1986**, *50*, 1−24.

(7) Thompson, M.; Schwab, G. E. δ-endotoxin expression in *Pseudomonas fluorescens*. U.S. Patent No. 5,527,883, 1996.

(8) Narva, K. E.; Palta, A.; Pellow, J. W. *Product characterization data for Bacillus thuringiensis var. aizawai Cry1F(synpro) insect control protein as expressed in cotton. A study report submitted by Dow AgroSciences to United States Environmental Protection Agency (EPA)*; EPA Document No. MRID 45607902, 2001.

(9) Barker, R.; Idler, K.; Thompson, D.; Kemp, J. Nucleotide sequence of the T-DNA region from the *Agrobacterium tumefaciens* octopine Ti plasmid pTil5955. *Plant Mol. Biol.* **1983**, *2*, 335−350.

(10) Ellis, J. G.; Llewellyn, D. J.; Walker, J. C.; Dennis, E. S.; Peacock, W. J. The *ocs*-element: a 16 bp palindrome essential for activity of the octopine synthase enhancer. *EMBO J.* **1987**, *6*, 3203−3208.
(11) Wohlleben, W.; Arnold, W.; Broer, I.; Hillemannn, D.; Strauch, E.; Puhler, A. Nucleotide sequence of the phosphinothricin N-acetyltransferase from *Streptomyces viridochromogenes* Tu494 and its expression in *Nicotiana tabacum*. *Gene* **1988**, *70*, 25−37.
(12) Christensen, A. H.; Sharrock, R. A.; Quail, P. H. Maize polyubiquitin genes: Structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation. *Plant Mol. Biol.* **1992**, *18*, 675−689.
(13) Rose, R. I.; McCabe, J. M. Laboratory rearing techniques for rearing corn rootworm. *J. Econ. Entomol.* **1973**, *66*, 398−400.
(14) SAS Institute. *SAS/STAT User's Guide*, version 6, 4th ed.; SAS Institute: Cary, NC, 1990.
(15) Greenplate, J. T. Quantification of *Bacillus thuringiensis* insect control protein Cry1Ac over time in Bollgard cotton fruit and terminals. *J. Econ. Entomol.* **1999**, *92*, 1377−1383.
(16) Liu, J.; Benedict, C. R.; Stipanovic, R. D.; Bell, A. A. Purification and characterization of *S*-adenosyl-L-methionine: Desoxyhemigossypol-6-O-methyltransferase from cotton plants. An enzyme capable of methylating the defense terpenoids of cotton. *Plant Physiol.* **1999**, *121*, 1017−1024.
(17) Du, L.; Ge, F.; Zhu, S.; Parajulee, M. N. Effect of cotton cultivar on development and reproduction of *Aphis gossypii* (Homoptera: Aphididae) and its predator *Propylaea japonica* (Coleoptera: Coccinellidae). *J. Econ. Entomol.* **2004**, *97*, 1278−1283.
(18) Choma, C. T.; Surewicz, W. K.; Carey, P. R.; Pozsgay, M.; Raynor, T.; Kaplan, H. Unusual proteolysis of the protoxin and toxin from *Bacillus thuringiensis*. Structural implications. *Eur. J. Biochem.* **1990**, *189*, 523−527.
(19) Schnepf, E.; Crickmore, N.; Van Rie, J.; Lereclus, D.; Baum, J.; Feitelson, J.; Zeigler, D. R.; Dean, D. H. *Bacillus thuringiensis* and its pesticidal crystal proteins. *Microbiol. Mol. Biol. Rev.* **1998**, *62*, 775−806.
(20) Gill, S. S.; Cowles, E. A.; Pietrantonio, P. V. The mode of action of *Bacillus thuringiensis* endotoxins. *Annu. Rev. Entomol.* **1992**, *37*, 615−636.
(21) Bravo, A.; Sanchez, J.; Kouskoura, T.; Crickmore, N. N-terminal activation is an essential early step in the mechanism of action of the *Bacillus thuringiensis* Cry1Ac insecticidal toxin. *J. Biol. Chem.* **2002**, *277*, 23985−23987.
(22) Zhuang, M.; Gill, S. S. Mode of action of *Bacillus thuringiensis* toxins. In *Chemistry of Crop Protection, Progress and Prospects in Science and Regulation*; Voss, G., Ramos, G., Eds.; Wiley-VCH: Weinheim, Germany, 2003; pp 213−236.
(23) Gomez, I.; Sanchez, J.; Miranda, R.; Bravo, A.; Soberon, M. Cadherin-like receptor binding facilitates proteolytic cleavage of helix α-1 in domain I and oligomer pre-pore formaion of *Bacillus thuringiensis* Cry1Ab toxin. *FEBS Lett.* **2002**, *513*, 242−246.
(24) Jensen, O. L.; Podtelejnikov, A. V.; Mann, M. Identification of the components of simple protein mixtures by high-accuracy peptide mass mapping and database searching. *Anal. Chem.* **1997**, *69*, 4741−4750.

Received for review September 26, 2005. Revised manuscript received December 12, 2005. Accepted December 12, 2005.

JF0523869