Applied and Environmental Microbiology, Sept. 2010, p. 6150–6155 Vol. 76, No. 18
0099-2240/10/$12.00 doi:10.1128/AEM.00797-10

Detection of New *cry* Genes of *Bacillus thuringiensis* by Use of a Novel PCR Primer System▽

Pedro A. Noguera and Jorge E. Ibarra*

Departamento de Biotecnología y Bioquímica, CINVESTAV, Irapuato, Gto., México

Received 31 March 2010/Accepted 11 July 2010

On the basis of the known *cry* gene sequences of *Bacillus thuringiensis*, three sets of primers were designed from four conserved blocks found in the delta-endotoxin-coding region. The primer pairs designed amplify the regions between blocks 1 and 5, 2 and 5, and 1 and 4. *In silico* analyses indicated that 100% of the known three-domain *cry* gene sequences can be amplified by these sets of primers. To test their ability to amplify known and unknown *cry* gene sequences, 27 strains from the CINVESTAV (LBIT series) collection showing atypical crystal morphology were selected. Their DNA was used as the template with the new primer system, and after a systematic amplification and sequencing of the amplicons, each strain showed one or more *cry*-related sequences, totaling 54 different sequences harbored by the 27 strains. Seven sequences were selected on the basis of their low level of identity to the known *cry* sequences, and once cloning and sequencing of the complete open reading frames were done, three new *cry*-type genes (primary ranks) were identified and the toxins that they encode were designated Cry57Aa1, Cry58Aa1, and Cry59Aa1 by the *B. thuringiensis* Toxin Nomenclature Committee. The rest of the seven sequences were classified Cry8Ka2, Cry8-like, Cry20Ba1, and Cry1Ma1 by the committee. The crystal morphology of the selected strains and analysis of the new Cry protein sequences showed interesting peculiarities.

Bacillus thuringiensis has been safely used for the control of insect pests within the orders Lepidoptera, Coleoptera, and Diptera for the last 50 years (19). It is an aerobic, Gram-positive bacterium whose insecticidal activity is based on the presence of parasporal crystalline inclusions formed during the sporulation process. These parasporal bodies or crystals are assembled by the so-called Cry proteins, expressed by the *cry* genes (21).

B. thuringiensis shows great variability, as has been demonstrated by the huge number of strains isolated around the world (19), by the number of serotypes known to date (a total of 84) (20), and by the great number of different *cry* gene sequences accumulated so far (a total of 492) (8), as well as by the number of molecular characterization tools that have been developed, such as sequencing of the flagellin gene and of the *gyrB* and *aroE* genes, the band patterns from repetitive extragenic palindromic-PCR analyses, and the plasmid patterns, among others (17, 18, 25, 27), all indicating the great variability within this species.

In spite of this variability, some similarity has been found in the three-domain Cry proteins, evidenced by the presence of three conserved domains in the tertiary structure of the active toxin (delta-endotoxin), even from Cry toxins showing low levels of identity (i.e., the Cry1-type, Cry3-type, and Cry4-type toxins): one domain with a bundle of α helices and two domains of β sheets (3). This uniformity is, at least in part, a reflection of the five conserved blocks in the gene structure, which are present in almost all the *cry* genes (10).

The great number of sequences known to date is mostly a result of the strong interest in finding novel Cry proteins, with the foci being on three main purposes: (i) the search for a new range of activities, (ii) the search for higher levels of toxicity, and (iii) the search for alternative toxins in case of resistance development. The search for novel Cry toxins has followed different strategies, such as (i) the development of DNA libraries and their expression in acrystalliferous mutants of *B. thuringiensis* (24); (ii) hybridization with degenerate oligonucleotides, partial sequences, or complete *cry* genes used as probes (5, 12); and (iii) PCR amplification of putative novel *cry* genes by using multiple primers (14) or a combination of general primers, followed either by restriction analysis (26) or by a second amplification using a mixture of general and specific primers (13).

These strategies have shown certain levels of efficiency in detecting new *cry* genes, but they also have some limitations. For example, the screening of a large library is time-consuming and requires a high level serendipity, and the use of Southern analysis with either degenerate or specific probes is limited to finding only sequences related to known *cry* genes. The amplification of putatively new *cry* genes by PCR has been widely used in a series of different strategies. However, most of them are limited to finding genes with sequences similar to those used in the primer design. Previous attempts to design universal primers used only nine *cry* types, which significantly limits the search for actual new genes. Besides, the new sequences described were limited to the amplicons and no new *cry* genes were reported (6).

This report describes the design of a system of three sets of primers, based on conserved regions of the *cry* family, potentially able to amplify all the known three-domain *cry* genes. The use of this strategy with *B. thuringiensis* strains showing atypical crystal morphology allowed the identification of three new *cry* types and four more new *cry* holotypes.

* Corresponding author. Mailing address: Departamento de Biotecnologia y Bioquimica, CINVESTAV, Apartado Postal 629, 36500 Irapuato, Gto., Mexico. Phone: 52-462-623-9643. Fax: 52-462-624-5996. E-mail: jibarra@ira.cinvestav.mx.

▽ Published ahead of print on 23 July 2010.

TABLE 1. Universal primers designed from the conserved blocks of the three-domain *cry* genes and results of *in silico* analysis of their efficiency[a]

Set	Block	Sequence	SPE (%)	AE (%)
1	1	5′-TATGCWCAAGCWGCCAATYTWCATYT-3′	90	90
	5	5′-GGRATAAATTCAATTYKRTCWA-3′		
2	2	5′-TTTAGATATTGTTGCAWTATKKYC-3′	68	98
	5	5′-GGRATAAATTCAATTYKRTCWA-3′		
3	1	5′-TATGCWCAAGCWGCCAATYTWCATYT-3′	46	100
	4	5′-CATAACGTAGWYTTAYCTKAWT-3′		

[a] SPE, single-pair efficiency; AE, accumulated efficiency.

MATERIALS AND METHODS

Primer system design. A series of primers were designed with the main purpose of amplifying all the known *cry* gene sequences by PCR. The design was based on the sequences of the five conserved blocks reported as early as 1989 (11), all within the delta-endotoxin-coding region of the so-called three-domain Cry proteins (the five conserved blocks are missing in some *cry* genes, which do not group out in the same dendrogram). The precise delimitation of the five blocks was made by alignment of the nine genes where these blocks were originally detected. Then, each block was located in the rest of the *cry* genes and consensus sequences were obtained from highly related genes, such as *cry1*-type genes. A total of 102 consensus sequences were aligned, and conserved regions within each block were identified. Four degenerate primers, 22 to 26 bases long, were designed after these regions in an attempt to cover as many *cry* sequences as possible, but they were limited to having no more than 5 wobble bases each (Table 1).

In total, three sets of primers were designed on the basis of the use of four primers: two forward primers from blocks 1 and 2 and two reverse primers from blocks 4 and 5. The first pair was designed to amplify the sequences between blocks 1 and 5. The second pair was designed to amplify the region between blocks 2 and 5, and the third pair amplifies the region between blocks 1 and 4 (Table 1). The three sets of primers were designed progressively, as they were tested *in silico*, using the software Amplify (version 3.14; Bill Engels, University of Wisconsin, Madison, WI; http://engels.genetics.wisc.edu/amplify/index.html). Once the theoretical analysis showed that the three-pair system has the potential to amplify 100% of the known *cry* genes coding for the three-domain Cry proteins, the primers were tested with DNA extracted from standard strains (i.e., strains HD-1, IPS-82, and DMS-2803) as well as from uncharacterized strains; and the amplification conditions were optimized in terms of temperature, Mg concentration, and number of cycles (see below).

Strain selection. Once the amplification conditions were optimized, the ability of the new amplification system to identify known and unknown *cry* genes was tested. For the purpose of finding novel *cry* genes, the search was focused on 27 *B. thuringiensis* strains from the CINVESTAV *B. thuringiensis* (LBIT series) collection (16) showing atypical parasporal crystal morphology (Table 2).

Amplification, cloning, and sequencing. Selected strains were subjected to plasmid-enriched DNA extraction by incubating the cells with lysozyme (20 mg/ml) in TE (50 mM Tris base, 10 mM EDTA, 20% sucrose, pH 7.5) at 37°C with shaking at 75 rpm for 2.5 h. Samples were then subjected to alkaline lysis (7).

Amplifications were carried with the following reaction mixture: 2.5 μl 10× reaction buffer, 100 ng each primer, 100 ng DNA, 5 mM $MgCl_2$, 0.4 mM deoxynucleoside triphosphates, 2.5 U *Taq* DNA polymerase (Invitrogen), and water to a final volume of 25 μl. Amplification reactions were carried out in a GeneAmp PCR system 2400 apparatus (Perkin Elmer) under the following conditions: an initial denaturation of 5 min at 94°C, followed by 30 cycles of denaturation at 94°C for 1 min, annealing for 2 min, and polymerization at 72°C for 1.5 min for each cycle. Amplification finished with an extension step at 72°C for 7 min. The annealing temperatures for set 1, set 2, and set 3 of the primers were 50, 56, and 48°C, respectively. Because *B. thuringiensis* strains usually contain more than one *cry* gene and different amplicons showed similar sizes, each amplification product was cloned in pCR2.1 Topo (Invitrogen), following the manufacturer's instructions, followed by restriction analysis of the clones, using EcoRI, EcoRV, HindIII, BamHI, and AluI (Invitrogen), following the manufacturer's instructions, to identify different sequences. The cloned amplicons were sequenced by the Sanger technique, using a 3730 XL DNA analyzer (Applied Biosystems), at the DNA sequence facility of CINVESTAV, Irapuato, Mexico. Sequences were analyzed by using the NCBI BLASTn, NCBI BLASTp, and NCBI BLASTx programs (1; www.blast.ncbi.nlm.nih.gov/Blast.cgi). Sequences showing 65% identity or less with known *cry* genes were selected for full sequencing of the gene.

Complete sequence of *cry* genes. Sequencing of putative new *cry* genes was accomplished by following two techniques: thermal asymmetric interlaced PCR (TAIL-PCR) (15) and a technique that uses the GenomeWalker universal kit (Clontech). Also, one gene required the previously published N-terminal sequence of the protein (16), so specific primers with the N-terminal sequence at the 5′ end of the gene were designed. Once the complete sequence of each *cry* gene was obtained by either technique, the full gene sequence was amplified and cloned into pCR2.1 Topo. The cloned genes were sequenced again, by using the vector's universal and reverse primers, as well as three more primers designed from the internal sequence. Each gene was sequenced at least three times, and a consensus sequence was obtained. The coded amino acid sequence was analyzed by the NCBI BLASTn, NCBI BLASTp, and NCBI BLASTx programs (1; www.blast.ncbi.nlm.nih.gov/Blast.cgi), in order to identify the maximum hit for each sequence. All sequences were registered at GenBank and were submitted to the *B. thuringiensis* Toxin Nomenclature Committee (BTTNC; www.lifesci.sussex.ac.uk/home/Neil_Crickmore/Bt/), which assigned them official names.

MS analysis. Expression of new Cry proteins (those showing less than 45% identity with the known three-domain Cry proteins) was analyzed by matrix-assisted laser desorption ionization–time of flight mass spectrometry (MALDI-TOF MS) and, if required, by Micromass Q-TOF micro-MS. For this purpose, SDS-polyacrylamide gels were stained with Coomassie blue and the bands of interest were excised. Selected gel fragments were serially washed in water, then 50% (vol/vol) acetonitrile in water, then acetonitrile–100 mM ammonium bicarbonate (1:1), and finally, 100% acetonitrile. The proteins in the gel fragments were then reduced, carbamido methylated, and digested with trypsin; and the resulting peptides were extracted from the gel as described by Shevchenko et al. (22). An aliquot of these peptides was analyzed by MALDI-TOF MS in an Ettan MALDI-TOF Pro instrument (General Electric). The mass lists obtained were compared with theoretical lists obtained from databases using the online program Mascot (Matrix Science). When this procedure did not lead to identification of the protein, the peptides were analyzed in a Micromass Q-TOF micromass spectrometer (Waters, Beverley, MA), where some of the peptides were sequenced from their collision-induced dissociation (CID) spectra, using the MassLynx (version 4.0) program provided with the equipment. The sequences obtained were compared with those in databases with the online program MSBlast (http://www.dove.emblheidelberg. de/Blast2/msblast.html) (23).

Nucleotide sequence accession numbers. The nucleotide sequences of the seven complete new *cry* genes were submitted to GenBank, and the accession numbers are given in Table 3.

RESULTS

***In silico* validation of the primer system.** Once the four primers were designed to be used in a three-set system, each set was tested *in silico* using all the known *cry* gene sequences as virtual templates and the software Amplify (version 3.14). The first set of primers showed the potential to amplify 90% of the known three-domain *cry* genes, while the second set of primers was theoretically able to amplify 68% of these *cry* genes. The third set of primers was designed

TABLE 2. Some features from the 27 selected atypical strains of *B. thuringiensis* subjected to amplification by the three-set primer system and the highest hits of their amplicons with the known *cry* genes

LBIT strain no.	Serotype	Origin	Morphology	Gene with highest hit (% identity[a])	
				Set 1	Set 2
LBIT-104	entomocidus	Mexico	Flat squares	*cry1Aa* (99) *cry1Ia* (99)	*cry21Aa* (96)
LBIT-113	entomocidus	Mexico	Flat squares	*cry1Ia* (99)	*cry21Ba* (25)
LBIT-129	thuringiensis	Mexico	Amorphous	*cry1Aa* (99) *cry1Ba* (99) *cry1Ia* (99)	*cry1Ia* (99)
LBIT-133	Unknown	Mexico	Amorphous	No amplification	*cry21Aa* (96)
LBIT-154	thuringiensis	Mexico	Bipyramidal	*cry1Aa* (99)	*cry1Ab* (99) *cry21Ba* (95)
LBIT-154					
LBIT-275	thuringiensis	Mexico	Bipyramidal	*cry1Ab* (99) *cry1Ia* (99)	*cry1Ab* (99) *cry1Ia* (99)
LBIT-279	thuringiensis	Mexico	Bipyramidal	*cry1Ab* (99)	*cry1Ab* (99) *cryIa* (99)
LBIT-442	Unknown	Mexico	Collapsed balloons	*cry1Ia* (99)	*cry5Ba* (95) *cry1Ia* (99)
LBIT-596	darmstadiensis	Argentina	Elongated bypiramidal	No amplification	*cry5Ba* (98)
LBIT-597	kenyae	Argentina	Small spheres	*cry1Ka* (43) *cry1Ia* (99)	*cry1Ia* (99)
LBIT-599	tohokuensis	Argentina	Collapsed balloons	*cry1Ab* (99) *cry1Ba* (99) *cry1Ia* (99)	*cry1Ab* (99) *cry1Ia* (99)
LBIT-772	fukuokaensis	China	Collapsed balloons	*cry1Ia* (99)	*cry5Ba* (80) *cry1Ia* (99)
LBIT-824	canadensis	Kenya	Small spheres	*cry7Ba* (38)	*cry5Ba* (96)
LBIT-935	japonensis	Denmark	Collapsed balloons	*cry1Ca* (63)	*cry5Ba* (93)
LBIT-937	japonensis	Guyana	Collapsed balloons	*cry24Aa* (34)	*cry5Ba* (96)
LBIT-953	mexicanensis	Japan	Small spheres	*cry1Ca* (63)	*cry7Ba* (38) *cry5Ba* (95)
LBIT-955	mexicanensis	Thailand	Small spheres	*cry1Ca* (63)	*cry21Aa* (25)
LBIT-956	monterrey	Spain	Collapsed balloons	No amplification	*cry7Ba* (38)
LBIT-960	amagiensis	South Korea	Small spheres	No amplification	*cry5Ba* (95)
LBIT-976	higo	Canada	Small spheres	*cry27Aa* (99) *cry20Aa* (69)	*cry7Ba* (38)
LBIT-979	kim	Denmark	Small spheres	*cry48Aa* (28)	No amplification
LBIT-980	kim	Spain	Collapsed balloons	*cry24* (34) *cry1Ia* (99)	*cry1Ia* (99)
LBIT-985	iberica	Guyana	Collapsed balloons	No amplification	*cry7Ba* (38) *cry4Ba* (61)
LBIT-1003	Unknown	Mexico	Collapsed balloons	No amplification	*cry1Da* (63) *cry5Ba* (95)
LBIT-1004	Unknown	Mexico	Collapsed balloons	No amplification	*cry4Ba* (61)
LBIT-1045	Unknown	Mexico	Polyhedral	*cry1Ia* (99)	*cry1Ia* (99)
LBIT-1189	Unknown	Mexico	Bipyramidal	*cry1Ca* (63) *cry1Ka* (69) *cry1Ab* (97)	*cry1Ab* (97)

[a] Percent identity with the known *cry* genes.

mostly to amplify the remaining *cry* genes, some of which are found in species other than *B. thuringiensis* (i.e., *Paenibacillus* and *Clostridium* spp.), and has the potential to amplify 46% of the *cry* genes. This means that some genes may be amplified by two or three sets of primers. The overlap between primer sets was higher between set 1 and set 2, and it was estimated that 60% of the known three-domain *cry* genes may be amplified by both sets. The overlaps between sets 1 and 3, sets 2 and 3, and sets 1, 2, and 3 indicated that an estimated 44%, 30%, and 30% of these genes may be amplified by the three sets, respectively. These theoretical results indicated that, after using set 1, the remaining 10% of these genes may be amplified by set 2 (8%) and set 3 (2%). Therefore, the three-pair system has the potential to amplify an accumulated 100% of the known three-domain *cry* genes (Table 1).

Strain selection. Selection of the 27 working strains was based on a combination of the most peculiar crystal morphologies and unusual banding in the SDS-polyacrylamide gels. Some strains showed crystals similar in shape to collapsed footballs, spheres, flat squares, or polyhedrons. Typical bipyramidal crystals associated with atypical inclusions were also detected. Table 2 describes the 27 selected strains, and Fig. 1 and 2 show both the crystal morphology and SDS-polyacrylamide gels for some of these strains, respectively (see below).

Sequence analysis of amplicons. Once the amplification conditions were established and tested on known and unknown strains, DNA from the 27 selected strains was subjected to

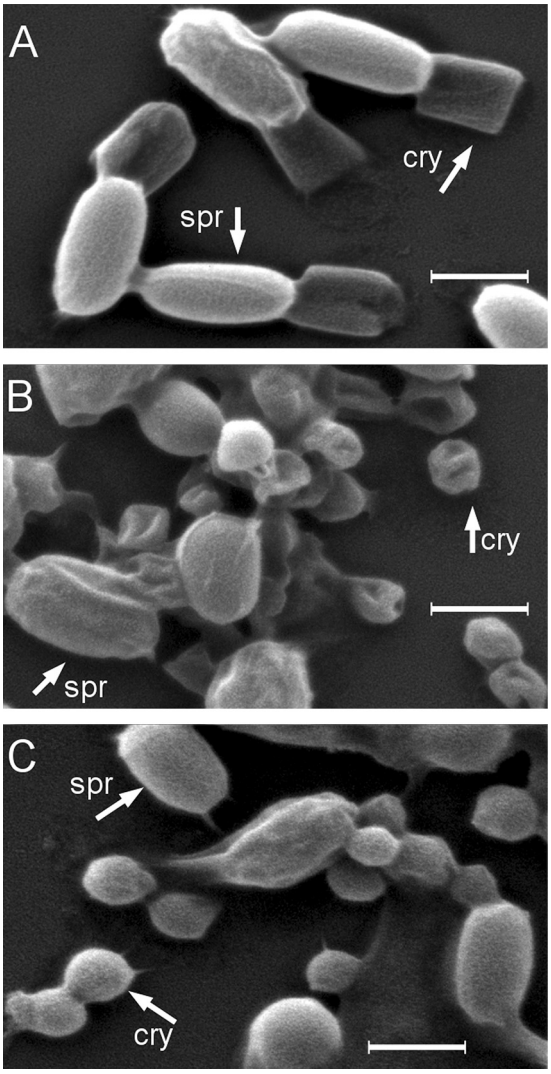

FIG. 1. SEM images of spore-crystal complexes from the three strains where the new *cry* gene types were found. (A) LBIT-113, harboring the *cry58Aa1* gene; (B) LBIT-979, harboring the *cry57Aa1* gene; (C) LBIT-980, harboring the *cry59Aa1* gene. spr, spore; cry, parasporal crystal. Bars, 1 μm.

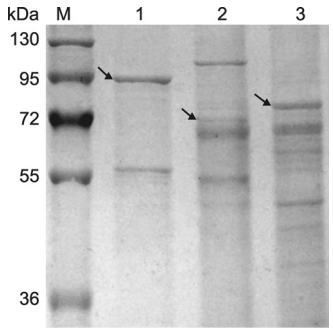

FIG. 2. SDS-polyacrylamide gel showing the new Cry protein types (arrows), identified by their molecular masses and by MALDI-TOF MS and Micromass Q-TOF micro-MS analyses. Lane 1, Cry58Aa1 from the LBIT-113 strain; lane 2, Cry57Aa1 from the LBIT-979 strain; lane 3, Cry59Aa1 from the LBIT-980 strain; lane M, PageRuler Plus prestained molecular size markers (Fermentas).

amplification by using the three-set primer system described above. All the strains selected showed at least one amplification product (Table 2), obtained with either set 1 or set 2. None of the 27 selected strains showed any amplification product when primer set 3 was used. Because different comigrating amplicons in the agarose gel were detected by restriction analysis, the amplification products were cloned and different clones were identified by individual restriction analysis. The number of different amplicons detected in any one strain varied from one to three. The cloned amplicons were subjected to sequencing, and sequences were analyzed with BLAST programs to detect the known *cry* gene to which the gene of the selected strain was the most similar (highest hit) and the level of identity. All sequences were identified as some type of *cry* gene. A total of 66 sequences were analyzed, and the levels of identity with known *cry* genes ranged from 25 to 99%. As expected, *cry1*-type genes were the most frequently found among the 66 sequences analyzed, with a frequency of 59%, followed by the *cry5*-type genes, with a frequency of 14%. Interestingly, five sequences showed highest-hit identities of less than 45%, which indicated potential new *cry* types, and four sequences showed identities of between 45 and 78%, which indicated potential new second-level (new letter)-type sequences. Actual identification and typing of putative new *cry* genes were performed when the complete gene sequences of seven of these genes were obtained. Table 2 shows the gene with the highest hit for each amplicon and its level of identity.

Sequence analysis and characterization of complete *cry* genes. Once techniques such as TAIL-PCR and Genome Walker were used on the seven putatively new *cry* genes, their complete gene sequences were obtained. These are registered in the GenBank database, and basic information for the seven sequenced *cry* genes is shown in Table 3. All seven genes received an official name from the BTTNC (http://www.lifesci.sussex.ac.uk/home/Neil_Crickmore/Bt/), and the set of seven genes included three new *cry* gene types (*cry57Aa1*, *cry58Aa1*, and *cry59Aa1*).

(i) Cry58Aa1 protein. The sequence of Cry58Aa1, a new holotype, should be considered a so-called naturally truncated toxin, as its sequence contains only 837 amino acids, instead of

TABLE 3. Some features of the seven complete new *cry* gene sequences selected from the previously sequenced amplicons by their low levels of identity to the known three-domain Cry proteins[a]

Strain	GenBank accession no.	Size		Highest hit		BTTNC no.
		nt	aa	Gene	% identity	
LBIT-113	FJ770568	2,514	837	*cry21Ba*	28	Cry58Aa1
LBIT-597	FJ770570	3,522	1,173	*cry8Ea*	61	Cry8Ka2
LBIT-824	FJ770571	3,393	E	*cry7Ab*	56	Cry8-like
LBIT-976	FJ884065	2,298	765	*cry20Aa*	63	Cry20Ba1
LBIT-979	FJ770569	1,830	609	*cry48Aa*	27	Cry57Aa1
LBIT-980	FJ770572	1,998	665	*cry24Aa*	35	Cry59Aa1
LBIT-1189	FJ884067	2,097	699	*cry1Bc*	70	Cry1Ma1

[a] nt, number of nucleotides in the gene; aa, number of amino acids in the protein; BTTNC no., code assigned by the *B. thuringiensis* Toxin Nomenclature Committee; E, the nucleotide sequence was not translated due to multiple stop codons in the sequence.

the approximately 1,200 amino acids of the so-called complete toxins, such as Cry1, Cry4, and Cry9, among others. Interestingly, further sequencing of 571 nucleotides, located downstream of the stop codon, showed that the sequence had more similarity with that of the 3′ half of the *cry21Ba* gene. The amino acid sequence of Cry58Aa1, visualized in SDS-polyacrylamide gels (Fig. 2), was corroborated by mass spectrometry. The crystal associated with this protein was a flat square inclusion attached to the spore (Fig. 1), whose characterization has been reported previously (16).

(ii) Cry57Aa1. Cry57Aa1 is another new *cry* gene type whose sequence is also considered to constitute a naturally truncated protein, as it contains only 609 amino acids, as visualized by SDS-PAGE analysis (Fig. 2) and corroborated by mass spectrometry. The crystal associated with Cry57Aa1 looked like a small collapsed balloon by scanning electron microscopy (SEM) (Fig. 1).

(iii) Cry59Aa1 protein. The third new *cry* gene type reported here, Cry59Aa1, shows a sequence with a stop codon in the middle which separates two consecutive open reading frames (ORFs). ORF1 is translated into a 665-amino-acid protein (the Cry59Aa1 protein), as detected by SDS-PAGE analysis (Fig. 2) and corroborated by mass spectrometry, which constitutes a naturally truncated protein. ORF2 generates a 584-amino-acid protein (detected in the SDS-polyacrylamide gels and corroborated by mass spectrometry) whose sequence is 80% similar to the corresponding ORF2 from the *cry39* and *cry40* genes. The crystal associated with Cry59Aa1 looked like a large collapsed balloon (Fig. 1).

(iv) Cry8Ka2. The Cry8Ka2 sequence translates into a 1,173-amino-acid protein, easily detected in the SDS-polyacrylamide gel (data not shown) and corroborated by mass spectrometry. This sequence constitutes a so-called complete toxin. The BTTNC assigned the second 8Ka registration to this sequence, even though the first registered sequence is not available; therefore, this becomes the first report of such a group. Cry8Ka2 is associated with a small polyhedron-like crystal.

(v) Cry8-like sequence. The Cry8-like sequence contains 12 stop codons along its sequence, which may be the reason why no protein was detected in the SDS-polyacrylamide gel. This is also the reason why the BTTNC did not assign a specific name to this toxin. Regardless of whether it is actually expressed or not, the LBIT-824 strain contains small sphere-like crystals which may be formed by other Cry proteins in the strain (Table 2).

(vi) Cry20Ba1. Although the Cry20Ba1 sequence should encode a 765-amino-acid protein, which corresponds to a naturally truncated toxin, detection of such a band in the SDS-polyacrylamide gel was not achieved. Interestingly, the downstream sequence, after the stop codon, was identified as a transposase previously described for the *B. cereus* group (2). Strain LIBT-976 shows a crystal with a collapsed balloon shape, which may be the result of the expression of other *cry* genes present in this strain (Table 2).

(vii) Cry1Ma. The Cry1Ma sequence should be considered to constitute that of a naturally truncated toxin. However, SDS-PAGE and mass spectrometry analyses were unable to identify this protein, mostly because other *cry1* genes and probably one *cry2* gene were present in the LBIT-1189 strain (Table 2). Interestingly, Cry1Ma is related at a 69% identity level to the Cry1Ia toxin, which was, until now, the only naturally truncated Cry1-type protein. Strain LIBT-1189 showed a typical bipyramidal crystal.

DISCUSSION

This report describes the design of a novel primer system putatively able to identify any of the known three-domain *cry* genes of *B. thuringiensis* and proved its ability to recognize three new *cry* gene types. The design of this system was based on the five conserved regions originally located by Höfte and Whiteley (11). Four degenerate primers, used in three sets, were required, due to variation in these conserved regions. It is important to note that, for some reason, some *cry* genes are included in this family (*cry6*, *cry15*, *cry22*, *cry23*, *cry33*, *cry34*, *cry35*, *cry36*, *cry37*, *cry38*, *cry45*, and *cry46*), even though they lack a phylogenetic relationship with the rest of the *cry* genes and show none of the five conserved regions; therefore, their products lack the typical three domains shown by delta-endotoxins (9, 10, 21). Genes encoding these delta-endotoxins are not expected to be amplified by any of the three sets of primers designed here.

The same approach to the design of universal primers to amplify *cry* genes has been reported before (6, 13, 26, 28); however, most of the approaches are limited to amplification of specific groups, or experimental evidence to prove their universality by finding totally new groups within the *cry* family is lacking.

Our approach started by designing a quasiuniversal pair of primers. Theoretically, set 1, which amplifies the sequence from block 1 to block 5, is able to identify 90% of the three-domain *cry* genes. This is the reason why it was so successful at identifying most of the *cry* genes from the 27 selected strains and six out of the seven genes reported here. Set 2 was mostly designed to amplify genes within the groups *cry5*, *cry8*, *cry10*, *cry12*, *cry13*, *cry14*, *cry21*, *cry40*, and *cry44*. Theoretically, this set is able to amplify 8% of the three-domain *cry* genes that set 1 is unable to do, which adds up to a total of 98% of these genes if only these two sets were used. By using set 2, we were able to amplify the new *cry58Aa1* gene, in addition to the *cry5*, *cry21*, and *cry8*-like genes. Set 3 was mostly designed to amplify genes within the *cry16* and *cry18* groups. This set is able to amplify *in silico* the remaining 2% of the three-domain *cry* genes that sets 1 and 2 were unable to amplify, which adds up to amplification of a total of 100% of these genes when all three sets were used. Perhaps due to this specificity, no amplification was detected for any of the 27 selected strains when set 3 was used. Because of these results, set 3 may be excluded in future screening programs. Still, this primer system demonstrated its ability to amplify both known and unknown *cry* genes, which proves its potential to be used in other *cry* gene screening programs.

Previous strategies have been designed to identify new *cry* genes and have been based either on the construction of DNA libraries, on gene hybridization, or on PCR-based methods. The first two strategies have proven to be time-consuming, and previous sequence information is sometimes required. Besides, their use for the screening of great numbers of strains is impractical (12, 24). PCR-based strategies are, in general, simpler, faster, and more efficient; however, their efficiency at recognizing all (or at least most of) the *cry* genes, including unknown *cry* genes, depends on how universal the primers to be used are (13).

Techniques such as PCR-restriction fragment length polymorphism analysis and exclusive PCR (E-PCR) are based on the use of universal primers, although most of these primers are designed to recognize specific groups (i.e., *cry1* genes), or at best, they recognize a few related groups (i.e., *cry1*, *cry8*, and *cry9* genes) (13, 26). Interestingly, both techniques may enhance the ability to identify known *cry* genes, if our proposed primer system is used. Strategies specifically designed to identify unknown *cry* genes have been based either on the hybridization of a mixture of *cry* genes (5) or on the use of nested PCR with universal primers (6). The first strategy is time-consuming and requires the use of radioactivity, and its universality has not been tested. In the second strategy, the universal primers were designed from the sequence of only nine *cry* genes, its universality has not been tested *in silico*, and no new *cry* genes were reported.

The system reported here was able to evaluate 27 strains showing atypical crystals in a fast, easy, and efficient way. A total of 66 *cry* sequences were amplified from the 27 strains; and 7 *cry* genes were completely sequenced, because of their novelty, and 3 of them became new *cry* gene types. These results prove the efficiency of this system to recognize both known and unknown *cry* genes. Indeed, the high proportion of novel sequences (21% of amplicons showed less than 78% identity with the known *cry* sequences) must be related to the previous strain selection, which was focused on those strains harboring parasporal crystals with atypical morphology, that is, morphologies diverging from the typical bipyramidal, cuboidal, flat square, or semispherical crystals, shown by well-known strains active against lepidopteran, coleopteran, or dipteran insects (11). Still, it is important to note that studies are required to associate the atypical crystals observed in the selected strains with the sequences reported here, except for the crystal from strain LBIT-113 (16). On the other hand, as an extension of this approach, the search for new *cry* genes in future work may be focused on species of bacilli other than *B. thuringiensis*, as previous findings of *cry* genes in *Paenibacillus*, *Clostridium*, and *B. sphaericus* support this idea (4, 12, 29). This would be an important challenge for the primer system reported here. Additionally, this primer system may be used as a complementary tool to characterize *B. thuringiensis* strains according to their *cry* gene content.

ACKNOWLEDGMENTS

We thank Regina Basurto, Javier Luévano, José Eleazar Urbina, Aurora Verver, Alicia Chagoya, and Karla Pérez for their excellent technical support. We are also grateful to the Institut Pasteur and Graciela Benintende (INTA, Argentina) for providing some strains used in this work.

This work was partially supported by Group Projects G-36505-N and 182388, CONACYT, Mexico. P.A.N. received a Ph.D. student fellowship from CONACYT.

REFERENCES

1. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
2. **Anderson, I., A. Sorokin, V. Kapatral, G. Reznik, A. Bhattacharya, N. Mikhailova, H. Burd, V. Joukov, D. Kaznadzey, T. Walunas, M. D'Souza, N. Larsen, G. Pusch, K. Liolios, Y. Grechkin, A. Lapidus, E. Goltsman, L. Chu, M. Fonstein, S. D. Ehrlich, R. Overbeek, N. Kyrpides, and N. Ivanova.** 2005. Comparative genome analysis of *Bacillus cereus* group genomes with *Bacillus subtilis*. FEMS Microbiol. Lett. **250:**175–184.
3. **Angsuthanasombat, C., P. Uawithya, S. Leetachewa, W. Pornwiroon, P. Ounjai, T. Kerdcharoen, G. R. Katzenmeier, and S. Panyim.** 2004. *Bacillus thuringiensis* Cry4A and Cry4B mosquito-larvicidal proteins: homology-based 3D model and implications for toxin activity. J. Biochem. Mol. Biol. **37:**304–313.
4. **Barloy, F., A. Delecluse, I. Nicolas, and M. Lecadet.** 1996. Cloning and expression of the first anaerobic toxin gene from *Clostridium bifermentans* subsp. *malaysia*, encoding a new mosquitocidal protein with homologies to *Bacillus thuringiensis* delta-endotoxins. J. Bacteriol. **178:**3099–3105.
5. **Beard, C. E., C. Ranasinghe, and R. T. Akhurst.** 2001. Screening for novel *cry* genes by hybridization. Lett. Appl. Microbiol. **33:**241–245.
6. **Berón, C. M., L. Curatti, and G. L. Salerno.** 2005. New strategy for identification of novel *cry*-type genes from *Bacillus thuringiensis* strains. Appl. Environ. Microbiol. **71:**761–765.
7. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7:**1513–1523.
8. **Crickmore, K., D. R. Zeigler, J. Feitelson, E. Schnepf, J. Van Rie, D. Lereclus, J. Baum, and D. R. Dean.** 1998. Revision of the nomenclature for the *Bacillus thuringiensis* pesticidal crystal protein. Microbiol. Mol. Biol. Rev. **62:**807–813.
9. **de Maagd, R. A., A Bravo, C. Berry, N Crickmore, and H. E. Schnepf.** 2003. Structure, diversity and evolution of protein toxins from spore-forming entomopathogenic bacteria. Annu. Rev. Genet. **37:**409–433.
10. **de Maagd, R. A., A. Bravo, and N. Crickmore.** 2001. How *Bacillus thuringiensis* has evolved specific toxins to colonize the insect world. Trends Genet. **17:**193–199.
11. **Höfte, H., and H. R. Whiteley.** 1989. Insecticidal crystal proteins of *Bacillus thuringiensis*. Microbiol. Rev. **53:**242–255.
12. **Jones, G. W., C. Nielsen-Leroux, Y. Yang, Z. Yuan, V. F. Dumas, R. G. Monnerat, and C. Berry.** 2007. A new Cry toxin with a unique two-component dependency from *Bacillus sphaericus*. FASEB J. **21:**1–9.
13. **Juárez-Pérez, V. M., M. D. Ferrandis, and R. Frutos.** 1997. PCR-based approach for detection of novel *Bacillus thuringiensis cry* genes. Appl. Environ. Microbiol. **63:**2997–3002.
14. **Kalman, S., K. L. Kiehne, J. L. Libs, and T. Yamamoto.** 1993. Cloning of a novel *cryIC*-type gene from a strain of *Bacillus thuringiensis* subsp. *galleriae.* Appl. Environ. Microbiol. **59:**1131–1137.
15. **Liu, Y.-G., and N. Huang.** 1998. Efficient amplification of insert end sequences from bacterial artificial chromosome clones by thermal asymmetric interlaced PCR. Plant Mol. Biol. Rep. **16:**175–181.
16. **Lopez-Meza, J. E., and J. E. Ibarra.** 1996. Characterization of a novel strain of *Bacillus thuringiensis*. Appl. Environ. Microbiol. **62:**1306–1310.
17. **Reyes-Ramírez, A., and J. E. Ibarra.** 2005. Fingerprinting of *Bacillus thuringiensis* type strains and isolates by using *Bacillus cereus* group-specific repetitive extragenic palindromic sequence-based PCR analysis. Appl. Environ. Microbiol. **71:**1346–1355.
18. **Reyes-Ramírez, A., and J. E. Ibarra.** 2008. Plasmid patterns of *Bacillus thuringiensis* type strains. Appl. Environ. Microbiol. **74:**125–129.
19. **Roh, J. Y., J. Y. Choi, M. S. Li, B. R. Jin, and Y. H. Je.** 2007. *Bacillus thuringiensis* as a specific, safe, and effective tool for insect pest control. J. Microbiol. Biotechnol. **17:**547–559.
20. **Roh, J. Y., Q. Liu, D. W. Lee, X. Tao, Y. Wang, H. J. Shim, J. Y. Choi, J. B. Seo, M. Ohba, E. Mizuki, and Y. H. Je.** 2009. *Bacillus thuringiensis* serovar *mogi* (flagellar serotype 3a3b3d), a novel serogroup with a mosquitocidal activity. J. Invertebr. Pathol. **102:**266–268.
21. **Schnepf, E., N. Crickmore, J. Van Rie, D. Lereclus, J. Baum, J. Feitelson, D. R. Zeigler, and D. H. Dean.** 1998. *Bacillus thuringiensis* and its pesticidal crystal proteins. Microbiol. Mol. Biol. Rev. **62:**775–806.
22. **Shevchenko, A., M. Wilm, O. Vorm, and M. Mann.** 1996. Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels. Anal. Chem. **68:**850–858.
23. **Shevchenko, A., S. Sunyaev, A. Loboda, A. Shevchenko, P. Bork, W. Ens, and K. G. Standing.** 2001. Charting the proteomes of organisms with unsequenced genomes by MALDI-quadrupole time-of-flight mass spectrometry and BLAST homology searching. Anal. Chem. **73:**1917–1926.
24. **Shu, C., H. Yu, R. Wang, S. Fen, X. Su, D. Huang, J. Zhang, and F. Song.** 2009. Characterization of two novel *cry8* genes from *Bacillus thuringiensis* strain BT185. Curr. Microbiol. **58:**389–392.
25. **Soufiane, B., and J. Côté.** 2009. Discrimination among *Bacillus thuringiensis* H serotypes, serovars and strains based on 16S rRNA, *gyrB* and *aroE* gene sequence analyses. Antonie Van Leeuwenhoek **95:**33–45.
26. **Tan, F., J. Zhu, J. Tang, X. Tang, S. Wang, A. Zheng, and P. Li.** 2009. Cloning and characterization of two novel crystal protein genes, *cry54Aa1* and *cry30Fa1*, from *Bacillus thuringiensis* strain BtMC28. Curr. Microbiol. **58:**654–659.
27. **Xu, D., and J. Côté.** 2008. Sequence diversity of *Bacillus thuringiensis* flagellin (H antigen) protein at the intra-H serotype level. Appl. Environ. Microbiol. **74:**5524–5532.
28. **Yamaguchi, T., K. Sahara, H. Bando, and S. Asano.** 2008. Discovery of a novel *Bacillus thuringiensis* Cry8D protein and the unique toxicity of the Cry8D-class proteins against scarab beetles. J. Invertebr. Pathol. **99:**257–262.
29. **Zhang, J., T. C. Hodgman, L. Krieger, W. Schnetter, and H. U. Schairer.** 1997. Cloning and analysis of the first *cry* gene from *Bacillus popilliae*. J. Bacteriol. **179:**4336–4341.