A natural hybrid of a Bacillus thuringiensis Cry2A toxin implicates Domain I in specificity determination


This version is available from Sussex Research Online: http://sro.sussex.ac.uk/id/eprint/69981/

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:
Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.
A natural hybrid of a *Bacillus thuringiensis* Cry2A toxin implicates Domain I in specificity determination.

Changlong Shu\textsuperscript{a}, Fengjiao Zhang\textsuperscript{a}, Guihua Chen\textsuperscript{a}, Lazarus Joseph\textsuperscript{b}, Amina Barqawi\textsuperscript{b}, Jacob Evans\textsuperscript{b}, Fuping Song\textsuperscript{a}, Guoxun Li\textsuperscript{a}, Jie Zhang\textsuperscript{a}, Neil Crickmore\textsuperscript{b}

\textsuperscript{a} State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing, 100193, P. R. China

\textsuperscript{b} School of Life Sciences, University of Sussex, Brighton, BN1 9QG, UK

**Abstract**

A PCR-RFLP method was used to identify *cry2A* toxin genes in a collection of 300 strains of *Bacillus thuringiensis*. From 81 genes identified, the vast majority appeared to be *cry2Aa* or *cry2Ab*, however three showed a different pattern and were subsequently cloned and sequenced. The gene cloned from strain HD395 was named *cry2Ba2*. Since the proteins encoded by the genes cloned from LS5115-3 and DS415 shared >95% sequence identity with existing toxins their genes were named *cry2Aa17* and *cry2Ab29* respectively by the toxin nomenclature committee. Despite this overall similarity these two toxins resembled natural hybrids, with Cry2Ab29 resembling Cry2Ab for the majority of the protein but then showing identity to Cry2Aa for the last 66 amino acids. For Cry2Aa17, Domains II and III most closely resembled Cry2Aa (99% identity) whilst Domain I was identical to that of Cry2Ab. The toxicity of the recombinant toxins was tested against *Aedes aegypti* and *Spodoptera exigua*, and it was found that the toxicity profile of Cry2Aa17 more closely matched the profile of Cry2Ab than that of Cry2Aa, thus implicating Domain I in specificity determination. This association of Domain I with toxicity was confirmed when hybrids were made between Cry2Aa and Cry2Ab.

**Keywords:** Gene evolution; insecticidal toxin; in vivo recombination
1. Introduction

*Bacillus thuringiensis* (Bt) is a gram positive bacterial species that produces insecticidal toxins during sporulation (Raymond et al., 2010). These toxins are encoded by plasmid-borne *cry* genes, which have been incorporated into transgenic crops to provide insect-resistant properties (Pardo-Lopez et al., 2013). Currently over 300 different Cry toxins have been sequenced and named, and the Cry toxins have been classified into 74 primary families (Cry1 to Cry74) based solely on their amino acid identity (Crickmore et al., 1998). Amongst these the Cry2 toxins are an interesting family with 12 subgroups (Cry2Aa-Cry2Ak, Cry2Ba) of which some have dual activity against lepidopteran and dipteran larvae (van Frankenhuyzen, 2013). Since Donovan et al. isolated the first *cry2* gene (Donovan et al., 1988) over 70 *cry2A* toxin genes have now been cloned (Crickmore et al., 2017). Competition binding assays have shown that Cry2A binding sites on the guts of commercially relevant Heliothine species are different from those of Cry1 and Vip3A, supporting the practice of combining Cry2A with Cry1 and Vip3A in transgenic crops to delay the onset of resistance (Gouffon et al., 2011; Hernandez-Rodriguez et al., 2008).

PCR based methods have proved an efficient way to identify Bt toxin genes (Ben-Dov et al., 1997). In particular, PCR-RFLP has been applied widely in *cry* gene identification and is capable of identifying variants of existing genes (Kuo and Chak, 1996; Patel and Ingle, 2012). In the current investigation, we used PCR-RFLP to identify and characterize *cry2* genes in a collection of 300 Bt strains.

2. Materials and methods

2.1 Bacterial strains and plasmids

Ninety six of the *B. thuringiensis* strains used in this study were purchased from the Bacillus Genetic Stock Center (BGSC). The rest were native strains isolated from Shaanxi and stored in our laboratory. *Escherichia coli* DH5α was used for common transformations whilst Rosetta (DE3) was used for the
expressing the cloned genes.

2.2 PCR-RFLP

*B. thuringiensis* strains were incubated overnight at 30°C on Luria-Bertani (LB) agar plates. Genomic DNA was extracted and purified as described previously (Shu et al., 2013a). A 50 μl PCR mixture was composed of 1 μl (10ng) of template DNA, 25 μl Premix Taq Version 2.0 (TaKara), 0.2 μmol l⁻¹ of primers Cry2-FL and Cry2-RL (Table 1), and 22 μl ultrapure water. PCR was carried out for 30 cycles (at 94°C for 1 min, 55°C for 1 min, and 72°C for 2.5 min) with a 2 min pre-denaturation stage at 94°C and a final elongation stage of 7 mins at 72°C. PCR products were digested with both *Hin*f I and *Sau*3A I enzymes in a 30 μl digestion mixture composed of: 10 μl of PCR product, 3 μl 10 x reaction buffer, 15 μl ultrapure water and 1 μl (1U) of each enzyme (TaKara). Digestion was carried out at 37°C for 2 hours before the restriction fragments were separated on a 2% agarose gel.

2.3 Cloning and expression of cry2 genes

Purified PCR products were cloned into the *Ecl*136II site of pEB (Shu et al., 2016) and then introduced into *E. coli* DH5α. A 5’ primer from pEB (TCATAACGTCCCGCGAAAT) and the 3’ primer (Cry2-RL) were used to screen transformants for plasmids containing cry2 genes in the correct orientation for expression. The hybrid toxins were created by amplifying Domain I using primers DIF and DIR (Table 1) and ligating this to the acceptor plasmid in which the entire plasmid, except Domain I, had been amplified using primers DIIF and PlasmidR. To express the Cry2 proteins, *E.coli* Rosetta (DE3) was transformed with the appropriate plasmid and grown in LB medium containing ampicillin (100 mg l⁻¹) overnight. The culture was then diluted (1:100) into fresh LB medium (1000 ml) containing ampicillin and incubated with shaking (37°C, 230 rpm) until an OD₆₀₀ of 0.6–0.8 was reached after which isopropyl-β-d-thiogalactopyranoside (IPTG) was added to a final concentration of 0.5 mmol l⁻¹. Cultures were maintained with shaking (20°C, 12 h, 130 rpm) then centrifuged (10 000 g, 12 min, 4°C). The culture supernatant was discarded, and the pellet suspended in 70 ml 20 mM Tris-HCl (pH 8.0). The suspended cells were sonicated (medium-intensity pulses, 6 min) centrifuged again (10000
g, 15 min, 4°C) and both pellet and soluble fractions collected. The pellet was resuspended in 2 ml of the same buffer (20 mM Tris-HCl, pH 8.0) and stored at −20°C. After sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS–PAGE), an estimation of Cry2 protein concentration was made by densitometry using bovine serum albumin (BSA) as a standard.

2.4 Insecticidal bioassays

For *Aedes aegypti* twenty late 3rd instar larvae and Cry2 inclusions were added to 80 ml of deionized water and incubated at 27°C with a 18 light: 6 hour dark photoperiod. A small amount of larval food (ground TetraMin fish food) was also added to each beaker. The number of dead larvae was counted after a period of 72 hours. Larvae incapable of rising to the surface of the water, or not showing the characteristic diving reaction when the water was disturbed, were counted as dead. If more than 10% of the control larvae pupated, or if more than 20% mortality was recorded in the control, then the test was discarded and repeated.

Insecticidal activity against first-instar larvae of *S. exigua* was measured by incorporating the resuspended Cry2 inclusions into an artificial diet (Jiang et al., 1999). Each concentration was tested with 24 insects and survival was recorded after 7 days. All bioassays were replicated at least three times.

2.5 Nucleotide sequence accession numbers

The nucleotide sequence data for *cry2Ab29*, *cry2Aa17* and *cry2Ba2* were assigned GenBank accession numbers KF860847, KF860848 and KF014123 respectively.

3 Results

3.1 Cloning and characterization of the cry2 genes

Based on conserved regions around the start and stop codons of the *cry2A* genes a pair of universal primers (Cry2-FL, Cry2-RL) was designed (Table 1). Analysis of the gene sequences indicated that digestion of the amplified gene sequence with *Hinf* I and *Sau3A* I could distinguish between the *cry2*
variants (Table 1). 300 strains were screened of which 81 gave a positive PCR signal and RFLP analysis of the PCR products gave five distinct patterns. Two of these patterns corresponded to those predicted for cry2Aa and cry2Ab, whereas the other three appeared unique (Fig 1). Of the 81 amplicons tested 32 gave the cry2Aa pattern and 46 the cry2Ab pattern, the three unique patterns were each identified in only one strain. The three amplicons giving these unique patterns were cloned and sequenced. The deduced amino acid sequence from the gene cloned from HD395 was 99% identical to the Cry2Ba1 toxin and so the gene was named cry2Ba2. The polypeptide encoded by the gene from SD415 was 98% identical with Cry2Ab1 and so its gene was named cry2Ab29. Interestingly all 10 of the amino acid differences between Cry2Ab1 and Cry2Ab29 are found in the last 66 amino acids and Cry2Ab29 only differs from Cry2Aa1 by 1 amino acid in this region. This suggests that Cry2Ab29 is a natural hybrid of Cry2Ab and Cry2Aa. A similar phenomenon was found with the gene cloned from LD5115-3, the first 278 amino acids of the encoded toxin (which corresponds to Domain I and the first two beta-strands of Domain II) matched Cry2Ab1 exactly, whereas the final 355 matched Cry2Aa. The overall sequence identity of the protein was close enough to Cry2Aa for the toxin to be named Cry2Aa17, but it also appears to be a Cry2Aa/Cry2Ab hybrid.

3.2 Domain analysis

Since native Cry2A toxins that appeared to be hybrids of previously identified Cry2A toxins had been identified, an analysis of other publicly available Cry2A sequences was undertaken to look for other putative hybrids. Fig 2 shows the clustering patterns of individual domains of these Cry2A proteins and shows a number of toxins (highlighted by boxes) where a particular domain of that toxin does not cluster with other toxins sharing the same tertiary classification. For example, as noted above, domain I of Cry2Aa17 clusters with the Cry2Ab toxins not the Cry2Aa ones. Of the 10 different tertiary ranks analyzed (Cry2Aa-Cry2Ak) there are eight different Domain I sequences identified along with ten different Domain IIs and seven Domain IIIs. As well as the Cry2Aa17 and Cry2Ab29 hybrids described in this study, eight other natural hybrids were identified (Fig 3).
3.3 Expression of the toxins.

The toxins encoded by the three genes with the variant RFLP patterns were expressed alongside control toxins Cry2Aa (AAO13750), Cry2Ab (AAO13296) and Cry2Ah (EU939453). All six were expressed at a high level and accumulated predominantly in the insoluble fraction (Fig 4), as is common for recombinant crystal proteins expressed in *E. coli* (Agaisse and Lereclus, 1995).

3.4 Insecticidal activity

Since it is well established in the literature that Cry2Aa is toxic to *Aedes aegypti* whereas Cry2Ab is not (van Frankenhuyzen, 2013) we tested these two new hybrids for activity against this mosquito. Using a discriminatory dose of 2µg/ml we confirmed that Cry2Aa was toxic (>50% mortality) while Cry2Ab and Cry2Ah were non-toxic (<10% mortality), however none of three new toxins showed any activity. Since previous work had indicated that a region in Domain II between amino acids 278 and 340 was responsible for determining toxicity towards *A. aegypti* (Liang and Dean, 1994; Widner and Whiteley, 1990) it was unexpected that Cry2Aa17 was non-toxic given that it was identical to Cry2Aa1 from amino acid 279 onwards. To test whether Cry2Aa17 was functional as a toxin a bioassay was performed on *S. exigua* where a dose of 100µg/ml was found to be toxic (>40% mortality) confirming that there were no major functional defects in this toxin. This result implied that residues in the first 278 amino acids could influence specificity against *A. aegypti*. However since there were two amino acids downstream of this point in Cry2Aa17 that matched Cry2Ab1 and not Cry2Aa1 we eliminated the possibility that these could be influencing specificity by creating a mutated form (mCry2Aa17) that contained the mutations E451A and P630S. Using the same concentrations as above mCry2Aa17 was found to retain activity against *S. exigua* but remained inactive against *A. aegypti*. To test whether the differences between Cry2Aa17/Cry2Ab29 and their “type” toxins could affect their activity against *S. exigua* bioassays were performed against this insect. The data (Fig 5) indicate that Cry2Ab and Cry2Ah are more toxic than Cry2Aa and that Cry2Aa17 more closely resembles Cry2Ab. Cry2Ab29 had a similar level of toxicity to Cry2Ab, while Cry2Ba2 had no obvious activity. These results are
consistent with the observation that regions within the first 278 amino acids can influence specificity.

3.5 Creation of Cry2Aa/Cry2Ab hybrids

To test the possibility that Domain I could influence the specificity of Cry2A toxins against *A. aegypti* hybrids were created in which this domain (amino acids 1-263) was exchanged between Cry2Aa and Cry2Ab. The two resulting hybrids (Cry2AaAbAb and Cry2AbAaAa) were expressed and tested against *A. aegypti*. The former was found to be toxic (>50% mortality at 2µg/ml) while the latter had no detectable activity. This result again confirms a potential role of Domain I in determining the specificity of the Cry2A toxins.

4 Discussion

The PCR-RFLP screen used in this study to identify cry2 genes within Bt strains identified two genes that appeared to encode hybrid toxins. It is known that genetic exchange can take place between Cry toxin genes and that this can affect specificity, indeed such recombination in vivo helps explain the large diversity of Bt toxins found in nature (de Maagd et al., 2001). Earlier studies on the Cry2A toxins noted that whilst Cry2Aa had activity against both the Lepidopteran *Manduca sexta* and the Dipteran *Aedes aegypti*, Cry2Ab, despite sharing 87% sequence identity with Cry2Aa, had no activity against *A. aegypti* (Widner and Whiteley, 1989). The same authors later went on to create hybrids between Cry2Aa and Cry2Ab which identified putative blocks, both within Domain II, responsible for activity against each order of insect (Widner and Whiteley, 1990). Liang and Dean later used naturally occurring restriction sites that split Domain II into 3 approximately equal parts and created a number of hybrids (Liang and Dean, 1994). This work indicated that the specificity regions were not co-linear and that whilst part 1 was enough to confer specificity to another lepidopteran insect (*Lymantria dispar*) activity against *A. aegypti* required both parts 1 and 2. When the structure of Cry2Aa was solved (Morse et al., 2001), the authors identified a putative binding epitope in the regions identified in the above hybrid studies. It was later established that
although Cry2Ab does not have activity against *A. aegypti* it is toxic to another species of mosquito, *Anopheles gambiae* (McNeil and Dean, 2011). Amino acids in the putative specificity region that differed between Cry2Aa and Cry2Ab were individually exchanged, in Cry2Ab, and although some mutants lost activity to *A. gambiae* none gained activity to *A. aegypti*.

The data that we obtained on Cry2Ah adds to that previously published (Shu et al., 2013b) and extends its spectrum of activity to *S. exigua*. An interesting observation from our data is that the toxins that have good activity against *S. exigua* (Cry2Ab, Cry2Ah, mCry2Aa17 and Cry2Ab29) have weak or no activity against *A. aegypti* whereas Cry2Aa which is highly active to the mosquito is only weakly active against *S. exigua*. If this association holds for other Cry2A toxins it could help unravel the mechanisms behind specificity determination.

Most work on the mechanism of action of Bt toxins has concentrated on the roles of Domains II and III in binding to the target cells and whilst mutations in Domain I are known to affect toxicity there is no evidence that this domain plays a role in specificity determination (Bravo et al., 2013). Whilst our data indicate an association between Domain I and specificity the mechanism is unclear. It is conceivable that some of the steps involving Domain I, such as toxin stability, oligomerization or pore-formation are executed differently in different insects and thus can affect relative toxicity. Differences in Domain I may result in structural changes in the other two domains and thus indirectly affect toxin binding, alternatively the region of Domain I previously hypothesized to be masking the *A. aegypti* receptor-binding surface (Morse et al., 2001) may actually be contributing directly to that specific binding.

**Acknowledgements**

This study was supported by the National Natural Science Foundation of China (grant numbers. 31301731 and 31201574), the National High Technology Research and Development Program of China (863 Program) (No. 2011AA10A203), and the Opening Foundation of the State Key Laboratory for Biology of Plant Disease and Insect Pests (SKL2012OP05).
References


Table 1 Primers and PCR-RFLP data

<table>
<thead>
<tr>
<th>Gene or primer</th>
<th>Accession no.</th>
<th>Sequence</th>
<th>Product size (bp)</th>
<th>Fragment sizes (bp) with \textit{Hin}I \textit{I} and \textit{Sau}3A I</th>
</tr>
</thead>
<tbody>
<tr>
<td>cry2-FL</td>
<td></td>
<td>ATGAATAGTGATGGAAATAGCGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cry2-RL</td>
<td></td>
<td>TTAATAAAGTGGAAGATTAGTGGG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIF</td>
<td></td>
<td>ATGAATAATGGTAAATAYGGAAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIR</td>
<td></td>
<td>TTTAAAATACGAGACAGATGAKACATA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIIF</td>
<td></td>
<td>TATCAAAGCCTTCTAGTATCTTGYG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PlasmidR</td>
<td></td>
<td>ATAAACAGCTCTTAAATCGAATTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cry2Aa1</td>
<td>AAA22335</td>
<td>1902</td>
<td>597,340,290,140,122,113,</td>
<td></td>
</tr>
<tr>
<td>cry2Ab1</td>
<td>AAA22342</td>
<td>1902</td>
<td>832,290,263,112,105,</td>
<td></td>
</tr>
<tr>
<td>cry2Ac1</td>
<td>CAA40536</td>
<td>1902</td>
<td>499,338,282,244,185,112,</td>
<td></td>
</tr>
<tr>
<td>cry2Ad1</td>
<td>AAFL0583</td>
<td>1902</td>
<td>985,290,263,112,</td>
<td></td>
</tr>
<tr>
<td>cry2Ae1</td>
<td>AAQ52362</td>
<td>1902</td>
<td>654,341,280,254,</td>
<td></td>
</tr>
<tr>
<td>cry2Af1</td>
<td>ABO30519</td>
<td>1902</td>
<td>429,363,290,138,112,107,</td>
<td></td>
</tr>
<tr>
<td>cry2Ag1</td>
<td>ACH91610</td>
<td>1902</td>
<td>522,418,290,263,112,</td>
<td></td>
</tr>
<tr>
<td>cry2Ah1</td>
<td>EU939453</td>
<td>1902</td>
<td>829,338,263,112,105,</td>
<td></td>
</tr>
<tr>
<td>cry2Aii</td>
<td>FJ788388</td>
<td>1902</td>
<td>340,290,244,185,168,140,122,113,</td>
<td></td>
</tr>
<tr>
<td>cry2Ba2</td>
<td>KF014123</td>
<td>1875</td>
<td>532, 341, 308, 180, 140, 130, 112,</td>
<td></td>
</tr>
<tr>
<td>cry2Aa1’</td>
<td>KF860848</td>
<td>1902</td>
<td>832, 290, 140, 122, 113, 105</td>
<td></td>
</tr>
<tr>
<td>cry2Ab29</td>
<td>KF860847</td>
<td>1902</td>
<td>832, 290, 263, 112, 105,</td>
<td></td>
</tr>
</tbody>
</table>
Figure 1. PCR-RFLP patterns of amplified *cry2* genes.

Figure 2 Clustering relationships between individual domains of Cry2 toxins

Relationships were inferred using the Neighbor-Joining method (Saitou and Nei, 1987). The optimal tree with the sum of branch length = 0.80606492 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (500 replicates) is shown next to the branches (Felsenstein, 1985). The tree is drawn to scale, with branch lengths in the same units as those of the distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Poisson correction method (Zuckerkandl and Pauling, 1965) and are in the units of the number of amino acid substitutions per site. The analysis involved 72 amino acid sequences. All positions containing gaps and missing data were eliminated. Evolutionary analyses were conducted in MEGA5 (Tamura et al., 2011).
Figure 3. Domain configurations of the Cry2A toxins.

Each toxin is split into the three domains identified from crystallographic studies of Cry2Aa (Morse et al., 2001). Domains of the same color are those that cluster together at a distance of <0.02 as shown in Fig 2.
Figure 4. SDS-PAGE analysis of recombinant Cry2A toxins

I: insoluble fraction; S: soluble fraction. pEB is a control sample of *E. coli* Rosetta harboring just the expression vector pEB.
Figure 5. Toxicity of the Cry2A toxins towards *Spodoptera exigua*.

First instar larvae were exposed to a dose of 100µg/ml of toxin and mortality assessed after 7 days.

The data represent an average of three replicates and the bars indicate one standard error of the mean.