Journal of Biotechnology 145 (2010) 215–221

Contents lists available at ScienceDirect

Journal of Biotechnology

journal homepage: www.elsevier.com/locate/jbiotec

Variant Cry1Ia toxins generated by DNA shuffling are active against sugarcane giant borer

Kilvia I.C. Craveiro[a,b,c], José Edilson Gomes Júnior[a,b], Maria Cristina M. Silva[a], Leonardo L.P. Macedo[a,d], Wagner A. Lucena[a,e,f], Marilia S. Silva[g], José Dijair Antonino de Souza Júnior[a,b], Gustavo R. Oliveira[a,e,i], Mariana T. Quezado de Magalhães[a,e], Antônio D. Santiago[h], Maria Fatima Grossi-de-Sa[a,d,*]

[a] *Embrapa Recursos Genéticos e Biotecnologia, PqEB-Final W5 Norte, CEP 70770-900, Brasília, DF, Brazil*
[b] *Departamento de Biologia Celular, Universidade de Brasília, CEP 70910-900, Brasília, DF, Brazil*
[c] *Embrapa Transferência de Tecnologia, PqEB-Final Av. W5 Norte, CEP 70770-901, Brasília, DF, Brazil*
[d] *Pós-Graduação em Ciências Genômicas e Biotecnologia, Universidade Católica de Brasília, SGAN 916-Av. W5 Norte, CEP 70790-160, Brasília, DF, Brazil*
[e] *Universidade Federal do Rio Grande do Sul, Campus do Vale, Av. Bento Gonçalves 9500, Prédio 43421, CP 15005, CEP 91501-970, Porto Alegre, RS, Brazil*
[f] *Embrapa Algodão, Rua Oswaldo Cruz 1143, Centenário, CP 174, Campina Grande, PB, Brazil*
[g] *Embrapa Cerrados, BR 020 Km 18, CP 08223, CEP 73310-970, Planaltina, DF, Brazil*
[h] *Embrapa Tabuleiros Costeiros, Av. Beira Mar 3250, Jardins, CP 44, CEP 49025-040, Aracaju, SE, Brazil*
[i] *Pós-Graduação em Processos Biotecnológicos, Centro Politécnico, Universidade Federal do Paraná, CEP 81531-990, Curitiba, PR, Brazil*

ARTICLE INFO

Article history:
Received 12 August 2009
Received in revised form 3 November 2009
Accepted 12 November 2009

Keywords:
Telchin licus licus
Bacillus thuringiensis
Saccharum officinarum
Brush Border Midgut Vesicles
Phage-display
Molecular modelling

ABSTRACT

Sugarcane giant borer (*Telchin licus licus*) is a serious sugarcane pest in Americas whose endophytic lifestyle hampers effective chemical and biological controls. Therefore, development of alternative control methods is extremely important. Envisaging development of transgenic plants resistant to this pest, we investigated the effect of the *Bacillus thuringiensis* Cry protein Cry1Ia12synth (truncated protein lacking C-terminus with plant codon usage) and variants against *T. l. licus*. *cry1Ia12synth* gene was used to generate mutated variants, which were screened for toxicity toward *T. l. licus*. For that purpose, an innovative technique combining *cry* gene shuffling with phage-display was used to build a combinatorial library comprising 1.97×10^5 Cry1Ia12synth variants. Screening of this library for variants binding to *T. l. licus* Brush Border Midgut Vesicles led to the identification of hundreds of clones, out of which 30 were randomly chosen for toxicity testing. Bioassays revealed four variants exhibiting activity against *T. l. licus* as compared to the non-toxic Cry1Ia12synth. Eight single substitutions sites were found in these active variants. Based on theoretical molecular modelling, the probable implications of these mutations are discussed. Therefore, we have four genes encoding Cry1Ia12synth variants active against *T. l. licus* promising for future development of resistant transgenic sugarcane lines.

1. Introduction

Despite technological advances, sugarcane (*Saccharum officinarum* L.) culture still faces several phytosanitary problems and plant predators. Sugarcane giant borer, *Telchin licus licus* (Drury, 1773) (Lepidoptera: Castniidae), is one of the most important insect pests of the sugarcane crop, occurring in several countries of the Americas (Mendonça et al., 1996). The *T. l. licus* caterpillar develops inside the sugarcane stem, survives from one season to the next and causes significant annual economical losses (Mendonça, 1982). This endophytic lifestyle hampers the effectiveness of chemical, mechanical and biological control methods. The damage caused by this insect pest involves destruction of the basal internodes, reduction of sucrose, and formation of galleries that may compromise the entire diameter of the stem. This facilitates secondary infections by other insects and even more frequently by microorganisms (Mendonça et al., 1996). Therefore, the search for new alternatives for the control of this pest is of great importance for sugarcane producers.

In this context, a *Bacillus thuringiensis* (*Bt*) encoded entomotoxic protein (Cry) has been investigated for potential control of sugarcane giant borer neonate larvae. *Bt* is an aerobic, gram-positive bacterium that synthesizes crystalline inclusions during its sporulation that are composed of one or more Cry toxins and/or Cyt toxins (Höfte and Whiteley, 1989; Silva-Werneck and Ellar, 2008). Cry toxins are presently classified into 58 types (Cry1–Cry58) and many sub-types (e.g. Cry1Aa and Cry1Ba) based on their amino acid sequence similarity. They are active against a limited number of

* Corresponding author at: Embrapa Recursos Genéticos e Biotecnologia, PqEB-Final W5 Norte s/n, CP 02372, CEP 70770-900, Brasília, DF, Brazil.
Tel.: +55 61 3448 4902; fax: +55 61 3340 3624.
E-mail address: fatimasa@cenargen.embrapa.br (M.F. Grossi-de-Sa).

0168-1656/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.jbiotec.2009.11.011

susceptible insect species (including lepidopterans, coleopterans and dipterans) and also against nematodes (Bravo and Soberón, 2008). A major group of Cry toxins is the three-domain (3D)-Cry family, members of which share similarities in sequence and structure. At least two different hypotheses have been proposed to explain the mode of action of these toxins, one relating to formation of pores in the target insect midgut and the other involving signal transduction. For both models, the first step is similar, i.e. the crystals are ingested by the larvae and solubilised in the midgut into protoxins. These are cleaved by midgut proteases to give rise to an active 60 kDa 3D-Cry toxin. The activated toxin binds to a cadherin receptor that is located in the midgut microvilli (Bravo and Soberón, 2008). The pore formation model suggests that this interaction with a cadherin receptor facilitates the proteolytic removal of the Cry α1-helix, triggering toxin oligomerisation that results in pore formation, causing larval death. In contrast, the signal transduction model proposes that binding of monomeric toxin to a cadherin receptor activates an adenylyl cyclase/protein kinase A magnesium-dependent signalling pathway, resulting in cell lysis (Bravo and Soberón, 2008; Zhang et al., 2006).

The spectrum of insects controlled by *Bt* can be broadened by directed molecular evolution techniques, such as DNA shuffling (Lassner and Bedbrook, 2001). DNA shuffling coupled with the phage-display technique has been valuable for the generation of genetic diversity and for selection of variants showing binding affinity to specific protein targets. DNA shuffling is an *in vitro* recombination method that uses small homologous DNA sequence fragments as substrates for PCR reactions, aiming to produce populations of gene variants (Stemmer, 1994; Zhao and Arnold, 1997). This technique has been used for several applications in different fields, including the generation of new molecules conferring resistance to insect pests (Patten et al., 1997). The phage-display approach involves the presentation of peptide and protein libraries on the surface of phage particles for facilitated selection of proteins with high affinity and specificity for a determined target (Willats, 2002).

In a previous study, it was reported that Cry1Ia12 protein exhibits considerable toxicity against the lepidopteran fall armyworm (*Spodoptera frugiperda*), indicating a potential for activity against other lepidopteran species (Grossi-de-Sa et al., 2007). Since the *cry1Ia12* gene was isolated by our research group and our bank of *cry* genes is still under construction, the *cry1Ia12* gene was therefore used to initiate the prospection and development of Cry toxins against *T. l. licus*. For this purpose, the nucleotide sequence of the original *cry1Ia12* gene was first changed to accommodate plant codon usage, resulting in *cry1Ia12synth*. Considering that Cry1Ia12synth is not toxic to *T. l. licus*, the aim of this work was to use the *cry1Ia12*synth gene to generate and select protein variants with toxicity towards *T. l. licus*. Then, DNA shuffling coupled with phage-display was used to generate a *cry1Ia12synth* combinatorial library. Screening of this combinatorial library for *cry1Ia12synth* variants that bind to Brush Border Midgut Vesicles (BBMVs) of *T. l. licus* resulted in the identification of new *cry1Ia12synth* molecules with entomotoxicity towards sugarcane giant borer larvae.

2. Materials and methods

2.1. Midgut dissection

T. l. licus second and third instar larvae, available in the field solely during part of sugarcane season, were collected in a sugarcane plantation in Maceió, AL, Brazil. The larval midguts were extracted and transferred to a microcentrifuge tube with MET buffer (300 mM mannitol, 5 mM EGTA, 17 mM Tris, pH 7.5) containing 1 mM PMSF. Then, midguts were centrifuged at 2500 × *g* for 5 min at 4 °C, the pellet was washed twice with MET buffer. The resulting pellet, enriched with midguts, was stored at −80 °C until use.

2.2. Preparation of brush border membrane vesicles (BBMVs)

Preparation of BBMVs from the dissected midguts of *T. l. licus* larvae was performed by using the precipitation method described by Wolfersberger et al. (1987). The concentration of BBMVs proteins was determined by the Bradford assay (Bradford, 1976) using bovine serum albumin (BSA) as the standard for the calibration curve.

2.3. cry1Ia12synth gene

Previously, the *B. thuringiensis* S811 Brazilian strain, obtained from the collection of EMBRAPA Genetic Resources and Biotechnology, was used to isolate the *cry1Ia12* gene (Grossi-de-Sa et al., 2007) (GenBank accession no. AY788868). For the present work, the original *cry1Ia12* nucleotide sequence coding solely for the truncated toxin (lacking the C-terminus, i.e. consisting of the N-terminus and the domains I–III) was changed in order to accommodate plant genetic codon usage and named *cry1Ia12synth*. The *cry1Ia12synth* gene (1944 bp) was synthesized by Epoch Biolabs, Texas, US and cloned into the pBluescript II vector (Stratagene). The resulting sequence of *cry1Ia12synth* gene was deposited in the NCBI gene databank under accession number FJ938022.

2.4. DNA shuffling

First, the *cry1Ia12synth* gene was excised from the pBluescript II harbouring vector by digestion at 37 °C for 16 h under the following conditions: 5 μg plasmid DNA, 10 U *Not* I (Promega), 20 U *Sal* I (New England Biolabs), 1× Buffer D (Promega), 10 μg/mL BSA, in a 20 μL final volume. The digestion products were analysed by 1% agarose gel electrophoresis and the fragment corresponding to the *cry1Ia12synth* insert (1944 bp) was excised and purified from the agarose gel using the Geneclean II Kit (Bio 101). Then, 10 μg of the purified *cry1Ia12synth* gene was digested with 12.5 U of DNAse I (Invitrogen) at 15 °C for 15 min and then interrupted by addition of 5 μL of 0.5 M EDTA. Resulting fragments of 30–50 bp were jointly purified from a 2.5% agarose gel by using the High Pure PCR Product Purification Kit (Roche). Ten microlitres of the pool of purified fragments was used as template in a PCR without primers in a 25 μL final volume containing 0.4 mM dNTPs, 1 mM $MgSO_4$ and 2.5 U Platinum *Taq* DNA Polymerase High Fidelity (Invitrogen), in 5× Platinum *Taq* DNA Polymerase High Fidelity buffer. The conditions of the primerless PCR were: 2 min 95 °C; 44 cycles: 1 min 95 °C, 1 min 42 °C and 1 min 72 °C (with a 5 s increase in extension time per cycle); with a final step of extension for 7 min at 72 °C. The products of the primerless PCR (1.5 μL) were used as template for a second PCR, containing the forward primer Cry1Ia12synthFOR (5′-CCCGGCCCAGGCGGCCATGAAACTCAAGAAC-3′) and the reverse primer Cry1Ia12synthREV (5′-CCGGCCGGCCTGGCCTTCGTAAGTAACTTC-3′). Both primers encode an *Sfi* I site, which is adequate for later cloning into the pCOMB3X phagemid (Andris-Widhopf et al., 2000). The second PCR, performed in a 100 μL final volume, contained 0.2 mM dNTPs, 2 mM $MgSO_4$, 0.8 μM each primer, 5 U *Taq* DNA Polymerase (Invitrogen) and 5 U Platinum *Taq* DNA Polymerase High Fidelity (Invitrogen), in 1× Platinum *Taq* Buffer. The conditions for the second PCR were: 2 min 95 °C; 10 cycles: 30 s 95 °C, 1 min 42 °C and 1 min 72 °C; followed by 14 cycles: 1 min 95 °C, 1 min 42 °C and 1 min 72 °C (with 20 s increase in extension time per cycle); with a final step of extension for 10 min at 72 °C. The second PCR product corresponding to the DNA shuffling product was analysed by 1% agarose gel electrophoresis

and the only band of approximately 1944 bp, corresponding to a population of *cry1Ia12synth* variants, was excised and purified with the Geneclean II Kit (Bio 101).

2.5. Phage display combinatorial library construction

The DNA shuffling product and the pCOMB3X phagemid were both digested with the restriction enzyme *Sfi* I. The resulting digestion products were separated by agarose gel electrophoresis and purified with the Geneclean II Kit (Bio 101). The purified DNA shuffling product (1 μg) was ligated into the pCOMB3X phagemid (3.5 μg) in a reaction containing 18 U T4 DNA Ligase (Invitrogen) in 200 μL final volume. The ligation product was dialysed, lyophilised, dissolved in 15 μL H_2O and then fractionated into five aliquots. Each aliquot containing 3 μL was used to transform 60 μL *E. coli* XL1-Blue eletrocompetent cells (Stratagene). The procedure was repeated five times and the product of all the bacterial transformations, harbouring the phagemids encoding *cry1Ia12synth* gene variants, were cultivated for 1 h at 37 °C in SB medium (1% MOPS, 2% yeast extract, 3% tryptone, pH 7.0) containing carbenicillin at 20 μg/mL and tetracycline at 10 μg/mL. Afterwards, carbenicillin was added to 30 μg/mL and the cultivation proceeded for 1 h at 37 °C. M13 fusion phages were then produced by the sequential addition of pre-warmed SB medium at 37 °C, carbenicillin to 38 μg/mL, tetracycline to 7.75 μg/mL and 10^{12} plaque-forming units (pfu)/mL of helper phage (VCSM13 Stratagene). The culture was incubated for 2 h at 37 °C under agitation at 250 rpm. After adding 70 μg/mL kanamycin, the culture was incubated for 16 h at 37 °C under agitation at 250 rpm. The resulting M13 phages consisted of fusions between protein III of the phage capsid coding region and the *cry1Ia12synth* gene. In order to determine the library titre, an aliquot of the original culture was plated on selective LB-agar (1% tryptone, 0.5% yeast extract, 0.5% sodium chloride, pH 7.0, containing 1.4% agar and 100 μg/mL carbenicillin). After incubation at 37 °C for 16 h, the number of plaques was counted and used to calculate the library titre.

2.6. Selection of the cry1Ia12synth gene variants (biopanning)

The biopanning procedure was performed as described by Rader et al. (2001), with some modifications. Briefly, at each round, wells in a microtitre plate were coated with BBMV preparation (100 μg) and incubated 16 h at 4 °C. After coating, fusion phages in solution were added and incubated for 2 h at 37 °C. Several washing cycles were performed with PBS (137 mM NaCl, 2.7 mM KCl, 12 mM Na_2HPO_4 and 1.2 mM KH_2PO_4) containing 0.1% Tween 20. The retained M13 fusion phages were eluted (in 0.2 M glycine, pH 2.2) and used for transfection into *E. coli* cells, to be amplified for further selection rounds. In order to determine the titre of the phage specifically bound to BBMVs that were recovered in each selection cycle, an aliquot of the eluted phage was diluted 100 times and plated onto SB-agar containing 100 μg/mL carbenicillin. The biopanning cycle exhibiting the highest number of colony-forming units (cfus) was taken as the "cycle of enrichment of specifically bound phages". In this study, the biopanning procedure consisted of five rounds. The clones isolated from the enriched cycles were analysed to verify the integrity of the inserted *cry1Ia12synth* variant genes via colony PCR, using primers Cry1Ia12synthFOR and Cry1Ia12synthREV. The clones that presented amplicons around 1944 bp in length were further used for expression of Cry1Ia12synth variant proteins.

2.7. Expression of Cry1Ia12synth variants in fusion phage

The *cry1Ia12synth* variants exhibiting size integrity (1944 bp) and the DNA shuffling parental gene *cry1Ia12synth* were expressed on the surface of M13 phage. Bacteria harbouring the fusion phage

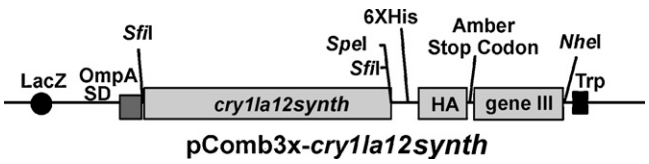

Fig. 1. Schematic representation of the expression cassette of the pCOMB3x-*cry1Ia12synth* phagemid construct, for expression of the protein Cry1Ia12synth displayed on the surface of filamentous M13 phage. Similar constructs containing shuffled *cry1Ia12synth* inserts were built to display the Cry1Ia12synth variants on the surface of M13 phage particles. LacZ: LacZ inducible operon. SD: One ribosome binding sites (Shine Dalgarno) give rise to separate polypeptides that are directed by the pelβ signal peptide to the periplasm. OmpA: Signal peptide. *Sfi*I, *Spe*I and *Nhe*I: Restriction sites. *cry1Ia12synth*: *cry* gene expressed as a fusion with M13 phage protein III (encoded by the virus). 6XHis: Six histidine-tag coding sequence. HA: Epitope for immunodetection by hemaglutinin antibodies. Amber Stop Codon: TAG stop codon that allows production of free and fusion heterologous proteins. gene III: Coding region of M13 phage Protein III. Trp: Transcription termination. Schematic figure adapted from Scott and Barbas (2001).

were cultivated in SB medium containing 100 μg/mL carbenicillin, 5 μg/mL tetracycline and 100 μg/mL kanamycin for 16 h at 37 °C under stirring (250 rpm). The culture supernatant was collected, precipitated with PEG-NaCl (20% Polyethylene–Glycol 8000, 15% NaCl) for 30 min at 4 °C and then centrifuged at 4300 × *g*. The precipitate was dissolved in TBS (5 mM Tris–HCl, 15 mM NaCl, pH 7.5), centrifuged, and the supernatant (containing the fusion phages) was collected and stored at 4 °C until use. Dot blot analysis by immunodetection of the hemaglutinin (HA) epitope (present in the pCOMB3X phagemid – Fig. 1) fused to the variant proteins was used to evaluate the presence or absence of the *cry1Ia12synth* variant fusion genes expressed on the phage. The concentration of the recombinant Cry1Ia12synth variant proteins was quantified by ELISA, using a polyclonal rabbit anti-Cry1I previously obtained in our lab (data not shown).

2.8. Bioassays of the anti-T. l. licus activity of Cry1Ia12synth and its variants

To assess the toxicity of the fusion phage-expressed Cry1Ia12synth and variants to sugarcane giant borer larvae, bioassays were carried out in 96-well microplates. Each well contained a 50 μL mixture of sterilised liquid artificial diet (2% casein, 1% yeast extract, 6% white sugar, 1% ascorbic acid, 1.1% vitamins, 0.75% Wesson salt mixture, 0.03% cholesterol, 0.03% sodium benzoate, 100 μg/mL ampicillin, 0.2% cholin chloride 50% and water) and 10 μg/mL of phage-expressed Cry1Ia12synth or variants deposited on a solid support (1 cm^2 of a commercial porous sponge of 0.5 mm thickness). One *T. l. licus* neonate larva was then placed into each well, and the microplate was incubated at 28 ± 1 °C, 80 ± 10% relative humidity and 12 h photoperiod. For the negative control treatment, the artificial diet without addition of Cry1Ia12synth or variants was used. Each treatment was repeated four times and each replicate contained 10 larvae. After 4 days, percent mortality was recorded for each treatment. Statistical analyses were performed to compare the average percent mortalities by using ANOVA and the Tukey's test at the probability level of 5% (Cruz, 2001).

2.9. DNA sequencing

The *cry1Ia12synth* variant genes that exhibited proteins significant toxicity against *T. l. licus* in the bioassays had their nucleotide sequences determined using a 3130xL Genetic Analyzer automatic sequencer (Applied Biosystems). The forward and reverse primers used for sequencing N and C-terminal coding ends were: MMB4 (5′-GCTTCCGGCTCGTATGTTGTGT-3′)

and MMB5 (5′-CGTCCATTGCATTCTTTTAAT-3′), respectively. The primers TLL3IF (5′-TGTGTCCAGCCAATATATCG-3′) and TLL3IR (5′-GTTGCCGTGTTGGTTCTCCT-3′) were designed to determine the variants' intermediate sequences. The nucleotide sequences were analysed using the BLASTn and BLASTx algorithms (Altschul et al., 1997), available on the web (http://www.ncbi.nlm.nih.gov/blast/Blast.cgi). Sequence alignments were performed by the ClustalW program (Thompson et al., 1994), available on the web (http://www.expasy.org/). In order to assemble the sequenced fragments and identify the positions of the mutated residues, the variant sequences were compared with the *cry1Ia12synth* parental sequence using the Staden Package program (Staden, 1996).

2.10. Molecular modelling

Sequence homology to Cry1Ia12synth was sought in the Protein Database Bank—PDB at the NCBI web site by using BLASTp with default parameters (Altschul et al., 1997). The resulting sequence set was aligned by using MUSCLE (Edgar, 2004) and further analyzed with JALVIEW (Clamp et al., 2004). After manual refinement of the sequence alignment, a homology model of Cry1Ia12synth and four variants was generated using MODELLER, Version 9.6 (Sali and Blundell, 1993). Default regimes of model refinement by energy minimisation and simulated annealing were employed. A rigorous interactive modelling protocol was adopted in which 20 models were constructed and analysed for each variant alignment. These models were analysed for stereochemical properties by using PROCHECK (Laskowski et al., 1993). When no further improvements could be achieved, the model with the best PROCHECK score was taken as the final model. Diagrammatic representations of the structures were generated by using PyMOL (DeLano, 2002).

3. Results

3.1. Generation of a cry1Ia12synth combinatorial library

The *cry1Ia12* gene (Grossi-de-Sa et al., 2007) had its nucleotide sequence modified (hereafter referred to as *cry1Ia12synth*) to conform to plant codon usage, facilitating its future use in plants for transgenic protection against insect pests. Besides changes in codon usage, the *cry1Ia12synth* gene coding sequence was also reduced to comprise only the active part of the original *cry1Ia12* molecule. Thus, the C-terminal domain, known to be dispensable for the toxicity of Cry proteins, was deleted. Consequently, the *cry1Ia12synth* gene is comprised of a 1944 bp open reading frame encoding a predicted protein of 648 residues, whereas the *cry1Ia12* gene is 2160 bp long and encodes a protein of 719 amino acids. For DNA shuffling purposes, the purified *cry1Ia12synth* parental gene was fragmented with DNaseI. The resulting fragments in the 30–50 bp range were recombined by primerless PCR. After a subsequent PCR reaction containing suitable primers to recover mutated *cry1Ia12synth* sequences, a population of *cry1Ia12synth* variants was obtained and visualised as a 1944 bp band (data not shown). This band was excised from the gel, purified and cloned into the pCOMB3X phagemid between two *Sfi* I restriction sites (Fig. 1). XL-1 Blue *E. coli* cells transformed with the resulting pCOMB3X-*cry1Ia12synth* constructs generated a phage-display combinatorial library of *cry1Ia12synth* variants containing 1.97×10^5 colony-forming units (cfu)/mL (Fig. 2).

3.2. Selection of Cry1Ia12synth variants binding to T. l. licus BBMVs (biopanning)

In order to select the clones expressing Cry1Ia12synth variants that bind specifically to the *T. l. licus* midgut, the phage-display combinatorial library was used for biopanning. Five selection rounds were performed. The second and third rounds were chosen since they yielded phage showing the highest binding specificity to *T. l. licus* BBMVs (Fig. 2). Four hundred randomly chosen clones from the second and third rounds exhibited amplicons corresponding to the original gene size of 1944 bp (data not show). The expression of the variant proteins was confirmed by dot blot detection of the hemaglutinin (HA) epitope fused to the variant proteins. A reaction signal was observed in the dot blot for most of the analysed clones (data not shown). Thirty clones that showed an expression signal in the dot blot analyses were randomly selected for further analysis of activity against *T. l. licus*.

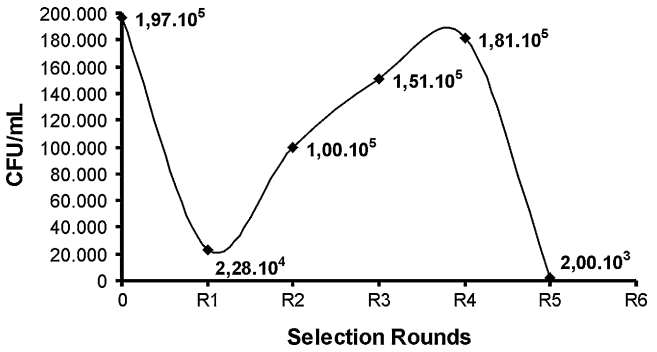

Fig. 2. Biopanning screening of the *cry1Ia12synth* Phage display combinatorial library for interactions of Cry1Ia12synth variants with Brush Border Midgut Vesicles (BBMVs) from sugarcane giant borer larvae (*T. l. licus*). The graph shows the number of phage-infected bacterial colonies obtained from each round of biopanning. In the second and third rounds of biopanning occurred the enrichment of recombinant phages displaying Cry1Ia12synth variants specifically bound to *T. l. licus* BBMVs. The points of the curve indicate the library titres in colony-forming units (cfu)/mL.

3.3. Bioassays of the anti-T. l. licus activity of Cry1Ia12synth and its variants

Bioassays to test the activity of Cry1Ia12synth and the 30 selected Cry1Ia12synth variants against *T. l. licus* were conducted by feeding neonate larvae an artificial diet containing 10 μg/mL of the proteins expressed in fusion with M13 phage. Incubation of larvae either with a Cry1Ia12synth protein-supplemented diet or a diet with no supplementation (negative control) resulted in a mortality rate of about 25%, indicating that Cry1Ia12synth protein is not active against *T. l. licus* at the concentration tested (Fig. 3). On the other hand, four out of the thirty Cry1Ia12synth variants tested had a statistically significant effect on *T. l. licus* as compared to the non-supplemented diet (Fig. 3). The analysis of variance was significant for the mortality rates of these four variants ($F = 20.10$; $p < 0.0001$). Statistical analysis to compare the average mortality rate revealed that the variants had similar toxicity to each other and higher toxicity than Cry1Ia12synth (Fig. 3).

3.4. Analysis of the mutations in the active Cry1Ia12synth variants

The four Cry1Ia12synth variants previously demonstrated to be active against *T. l. licus* were sequenced, and after assembly and alignment analysis, these new variant sequences were deposited in the NCBI gene databank under assigned accession numbers FJ938023 to FJ938026, respectively corresponding to *cry1Ia12synth* variants 1–4. The *cry1Ia12synth* gene and its four shuffled variant genes encode proteins of 648 amino acid residues. Sequence alignments comparing parental *cry1Ia12synth* and its variants identified numerous nucleotide substitutions. Most of these nucleotide substitutions resulted in silent mutations, but a few resulted in amino acid residue substitutions (mutations). No deletions were obtained

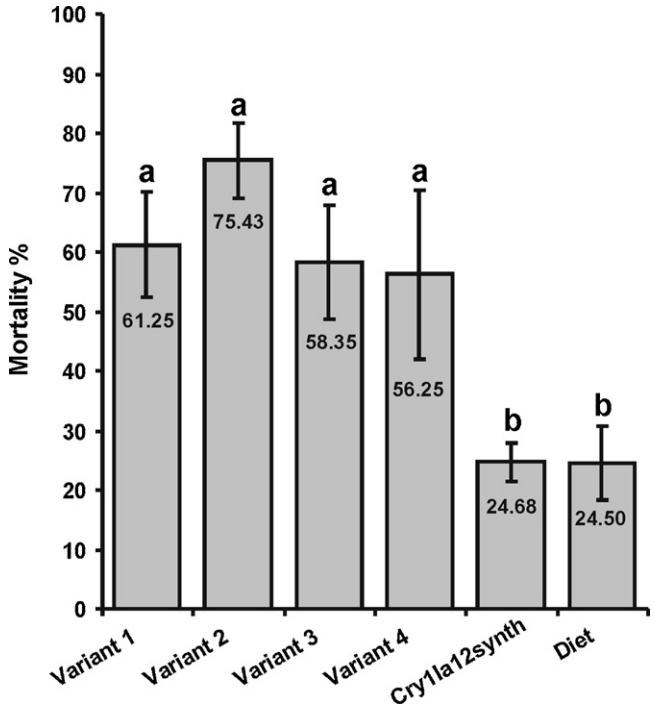

Fig. 3. Bioassays of Cry1a12synth and variants against *T. l. licus*. Treatments consisted of artificial diet supplemented or without Cry1Ia12synth protein or its variants. Average mortality rates (%) are plotted (vertical bars) and corresponding values are indicated. The standard deviation is indicated for each treatment (vertical lines). Average mortality rate values with different letters differ from each other by the Tukey's test at a probability of 5%.

(Table 1). Most of the changes in amino acid sequence were found in domain I, with a few changes found in domain III and none in domain II (Table 1).

To try to predict the effects of the mutations on structure and function, theoretical models were built for Cry1Ia12synth and the four variants by using Cry8Ea1 and Cry3Bb1 (PDB entries 3EB7 and 1JI6, respectively) (Galitsky et al., 2001; Guo et al., 2009) as templates. Models were chosen for Cry1Ia12synth and its variants according to PROCHECK analysis. In the template crystal structures chosen (i.e. E3B7 and 1JI6) there are no defined atomic coordinates to serve as templates to model N- and C-termini of the Cry1Ia12synth variant structures. Therefore, the models of the Cry1Ia12synth variants correspond to residues 62–643 (residue numbers of the Cry primary sequences), lacking the N- and C-termini but containing the three conserved domains of the Cry toxin structure (Fig. 4A).

The theoretical tertiary structure of Cry1Ia12synth depicts the location of the eight substitutions present in variants 1–4 (Fig. 4A). Three substitutions are located in domain I (Fig. 4A): (i) I116T, in variant 3, located in the loop between alpha2b and alpha3; (ii) D233N, in all variants except for variant 3, located in helix α6; and (iii) L266F, in variant 3, located in helix α7. Two mutations were found in domain III: (i) K580R, in variant 3, located in strand ß16; and (ii) E639G, in variant 1, located in strand ß19.

Table 1
Mutations present in Cry1Ia12synth variants.

Variant	Substitution mutation	Domain
1	D233N	I
	E639G	III
2	D233N	I
3	I116T	I
	L266F	I
	K580R	III
4	M45V	N-terminus
	D233N	I

4. Discussion

The gene *cry1Ia12*, previously isolated by our group from *B. thuringiensis* S811 strain, encodes a protein proven to be highly effective against the lepidopteran species *S. frugiperda*. In the present work, we generated a combinatorial phage-display library of variants of *cry1Ia12synth* via DNA shuffling and successfully identified four variants with optimised activity against *T. l. licus*.

In the last decade, directed molecular evolution techniques, such as DNA shuffling, have been employed in the agricultural research to create genes, such as the *cry* genes, encoding insecticidal molecules for the development of transgenic insect-resistant plants (Lassner and Bedbrook, 2001). The DNA shuffling strategy used here was highly efficient for our purposes, as it generated 1.97×10^5 variants of *cry1Ia12synth*, out of which over 400 bound specifically to *T. l. licus* BBMVs. Furthermore, four out of 30 tested variants presented activity against sugarcane giant borer neonate larvae when compared to the non-toxic Cry1Ia12synth.

In comparison to conventional bacterial expression systems, the phage-display system, in addition to providing soluble, ready-to-use protein, has advantages for the selection of variants that include speed, ease of use and low cost. However, previous studies (Kasman et al., 1998; Marzari et al., 1997) have encountered problems with the functionality of proteins displayed on M13 phage. It was found that the Cry1Aa toxin was not properly displayed, resulting in deletions of the fused protein. Moreover, despite the fact that Cry1Ac toxin displayed by M13 phage retained its toxicity against *Manduca sexta* larvae, there was no binding to functional insect midgut receptors *in vitro*, suggesting structural restrictions of the displayed toxin (Pacheco et al., 2006). In spite of these limitations previously related in the literature, our group apparently established optimal conditions that overcame such constraints to successfully express Cry1Ia12synth protein variants with the M13 phage-display system.

When considering the four Cry1Ia12synth variants active against sugarcane giant borer larvae, eight mutation sites of amino acid substitutions were found. Out of these eight, half resulted in substitutions of nonpolar residues, whereas the other half resulted in substitution of charged residues. Moreover, it was observed that most of these mutations were found in domain I of the Cry1Ia12synth active variants. One of the roles of the Cry domain I is the insertion of the protein into the membrane. Dean et al. (1996) using Y153D, Y153A and Y153R substitutions in CryIAb domain I, demonstrated that the negatively charged residue D was less favourable for toxin insertion into membranes than the positively charged R or neutral A residues. Another biological role attributed to Cry domain I is the formation of pores. Girade et al. (2009) demonstrated with Cry1Aa mutations in the α4 and α5 helices and the α3-α4 and α5-α6 adjacent loops that these regions have an important role in the pore formation mechanism. Therefore, mutations located in domain I of the Cry1Ia12synth variants may have influenced their insertion into the membrane or pore formation in the midgut of *T. l. licus*.

In our work, two mutations in active Cry1Ia12synth variants were found in domain III, a domain associated with protection of Cry from proteolysis, with receptor binding and with regulation of pore formation (Flores et al., 1997). In this respect, there is evidence that domain III of Cry proteins is involved in binding to membrane proteins of the *Spodoptera exigua* midgut, demonstrating that domain III of CryIC can play an important role in toxicity to insect pests (de Maagd et al., 1996).

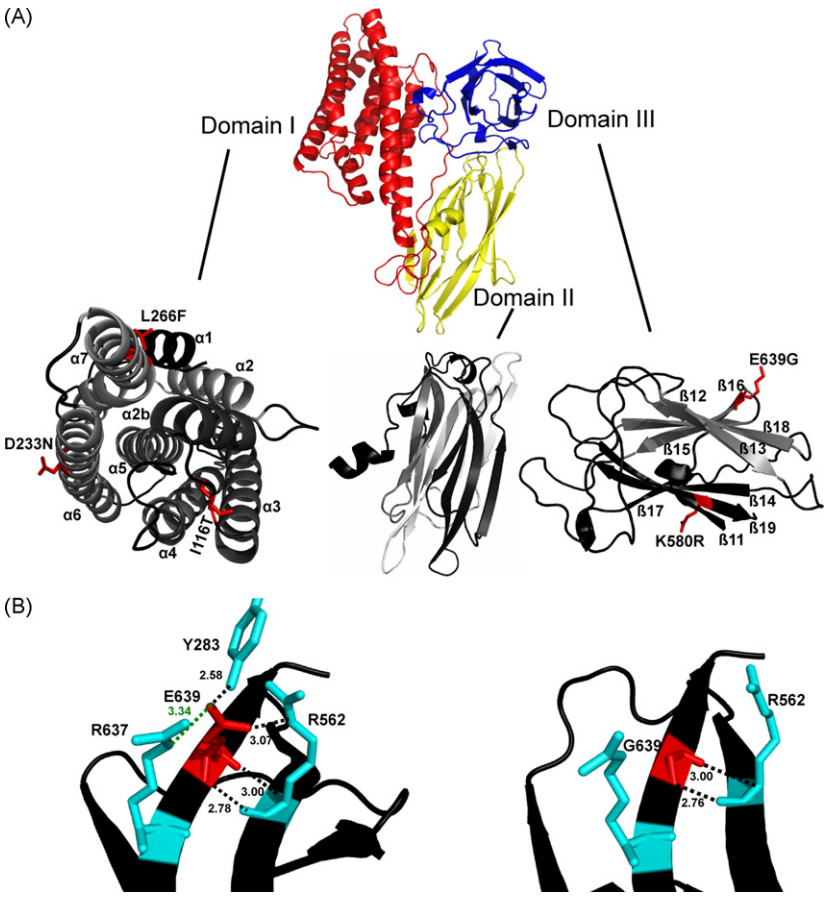

Fig. 4. Theoretical models of Cry1Ia12synth and variants. (A) A Cry1Ia12synth ribbon diagram, as viewed using PyMOL, is depicted at the top of panel A. Domain I is shown in red, domain II in yellow and domain III in blue. Separate models of each Cry1Ia12synth domain is shown in detail at the bottom of panel A. α-helices and β-strands are identified. Amino acid single substitutions present in the variants are coloured in red. The black dashed lines represent hydrogen bonds and the green dashed lines represent salt bridges. (B) Details of an amino acid substitution present in Cry1Ia12synth variant 1 (right), as compared to the parental Cry1Ia12synth (left). The variant 1 presents the mutation E639G, located in the β19 strand in domain III. The model on the left depicts the interactions between residue E639 (in red) and residues Y283 (from block 2), R562 (from block 4) and R637 (from block 5) (in cyan). These interactions are disrupted in Cry1Ia12synth variant 1 (right).

Taken alongside previous data in the literature, our molecular modelling analysis provides potential explanations for the activity of the four variants:

(i) In all Cry proteins, there are three conserved blocks, denoted blocks 2, 4 and 5, that are associated with the functionality of the toxins (Schnepf et al., 1998). In reference to the mutation E639G found in variant 1, G639 interacts solely with residue R562 from block 4 through two hydrogen bonds of around 3 Å, whereas in the Cry1Ia12synth protein, E639 interacts with Y283 from block 2 through a hydrogen bond of 2.58 Å, with R562 from block 4 through three hydrogen bonds of around 3 Å, and with R637 from block 5 through a salt bridge of 3.34 Å (Fig. 4B). Moreover, the mutation E639G of variant 1 is located in the β19 strand of domain III, a domain associated with protection of the toxin from proteolysis, binding to midgut receptors, regulation of pore formation (Flores et al., 1997) and binding to insect midgut membrane proteins (de Maagd et al., 1996).

(ii) Variant 2 contains the single mutation D233N. The residue at this position, be it D233 or N233, invariably presents its side chain to the solvent environment. These residues neither interact with nor belong to any region known to be relevant to the entomotoxic activity of Cry proteins. Nevertheless, the substitution of an oxygen from D233 for a nitrogen in N233 makes this spot on the toxin less negative, which might affect the interactions of the toxin with the solvent environment. To elucidate the role of this mutation in the functionality of variant 2, it would be interesting to perform modelling studies of molecular dynamics in order to better understand the movement of the side chain of residue 233 in relation to simulated changes in environmental conditions. Modelling studies of Cry1Ia12synth or variant 2 docking to the Cry receptor might also be revealing. Moreover, bench experiments involving site-directed mutagenesis of residue 233 and its consequences on interactions with the receptor and protein toxicity could shed light on the role of this residue in Cry functionality.

(iii) Three mutations occurring in variant 3 seem to present an interactive effect on the variant's structure, and consequently probably also on the variant's biological activity. Concerning the mutation I116T, the oxygen of T116 interacts with the beta-carbon of residue A120. This residue corresponds to residue A92 from Cry1Ab, which is involved in the insertion of the Cry toxin into the insect midgut membrane (Chen et al., 1995). Regarding the mutation L266F, L266 has its side chain turned towards the residues E455 (domain II), D270 and R262 (domain I). Thus the aromatic ring of F266, which is a less reactive residue than L266, interferes with interactions between domains I and II. With respect to the mutation K580R, both K580 and R580 are turned towards the neighbouring N578, which corresponds to the Cry1Ac N546 residue demonstrated to be related to Cry1Ac toxicity (Xiang et al., 2009).

Since the first 61 residues corresponding to the N-terminal portion of Cry1Ia12synth were not modelled, the effect of the mutation M45V present in variant 4 could neither be modelled nor analysed.

This work presents a pioneering approach towards the generation and selection of new genes to be used in the development of transgenic sugarcane resistant to the giant borer *T. l. licus*, a pest that is extremely resistant to conventional control methods and for which there is no known natural source of resistance. It should be noted that additional experiments, such as site-directed mutagenesis and modelling studies of molecular dynamics and interactions with the receptor are needed to better understand the effects of the mutations present in the Cry1Ia12synth variants on the biological activity of these proteins against *T. l. licus*. It will be also interesting to test these toxins against to others insect pest and non-target beneficial insects. In conclusion, the Cry1Ia12synth variants 1 to 4 are active against sugarcane giant borer larvae and are therefore promising proteins for the future development of transgenic sugarcane resistant to insect pests.

Acknowledgments

The present work was performed with financial support from CNPq, CAPES and EMBRAPA. The authors thank: Dr. Andréa Maranhão (University of Brasília, Brazil) for kindly providing the pCOMB3X phagemid and support about phage-display approaches; Dr. Dan Rigden (University of Liverpool, UK) for critical reading; Mr. Luiz Avelar Brandão Góis (Phytosanity Supervisor, Usina Triunfo, AL, Brazil) for sending us the giant borer larvae and Dr. Eduardo Alano Vieira (EMBRAPA Cerrados, Planaltina, DF, Brazil) for the aid in statistical analyses.

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25, 3389–3402.

Andris-Widhopf, J., Rader, C., Steinberger, P., Fuller, R., Barbas III, C.F., 2000. Methods for the generation of chicken monoclonal antibody fragments by Phage display. J. Immunol. Methods 242, 159–181.

Bradford, M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72, 248–254.

Bravo, A., Soberón, M., 2008. How to cope with insect resistance to Bt toxins? Trends Biotechnol. 26, 573–579.

Chen, X.J., Curtiss, A., Alcantara, E., Dean, D.H., 1995. Mutations in Domain I of *Bacillus thuringiensis* δ-Endotoxin CryIAb reduce the irreversible binding of toxin to *Manduca sexta* brush border membrane vesicles. J. Biol. Chem. 270, 6412–6419.

Clamp, M., Cuff, J., Searle, S.M., Barton, G.F., 2004. The Jalview Java alignment editor. Bioinformatics 20, 426–427.

Cruz, C.D., 2001. Programa Genes (versão Windows): aplicativo computacional em genética e estatística. Editora UFV, Viçosa.

de Maagd, R.A., Kwa, M.S.G., Klei, V.D., Yamamoto, T., Schipper, B., Vlak, J.M., Stiekema, W.J., Bosch, D., 1996. Domain III substitution in *Bacillus thuringiensis* delta-endotoxin CryIA(b) results in superior toxicity for *Spodoptera exigua* and altered membrane protein recognition. Appl. Environ. Microbiol. 62, 1537–1543.

Dean, D.H., Rajamohan, F., Lee, M.K., Wu, S.-J., Chen, X.J., Alcantara, E., Hussain, S.R., 1996. Probing the mechanism of action of *Bacillus thuringiensis* insecticidal proteins by site-directed mutagenesis—a minireview. Gene 179, 111–117.

DeLano, W.L., 2002. The PyMOL Molecular Graphics System on the World Wide Web.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797.

Flores, H., Soberón, X., Sánches, J., Bravo, A., 1997. Isolated domain II and III from the *Bacillus thuringiensis* Cry1Ab delta-endotoxin binds to lepidopteran midgut membranes. FEBS Lett. 414, 313–318.

Galitsky, N., Cody, V., Wojtczak, A., Ghosh, D., Luft, J.R., Pangborn, W., English, L., 2001. Structure of the insecticidal bacterial delta-endotoxin Cry3Bb1 of *Bacillus thuringiensis*. Acta Crystallogr., Sect. D. 57, 1101–1109.

Girade, F., Vachon, V., Lebel, G., Préfontaine, G., Scgwartz, J.-L., Masson, L., Laprade, R., 2009. Chemical modification of Bacillus thuringiensis Cry1Aa toxin single-cysteine mutants reveals the importance of domain I structural elements in the mechanism or pore formation. Biochim. Biophys. Acta. 1788, 575–580.

Grossi-de-Sa, M.F., Magalhães, M.Q., Silva, M.S., Silva, S.M.B., Dias, S.C., Nakasu, E.Y.T., Brunetta, P.S.F., Oliveira, G.R., Oliveira-Neto, O.B., Oliveira, R.S., Soares, L.H.B., Ayub, M.A.Z., Siqueira, H.A.A., Figueira, E.L.Z., 2007. Susceptibility of *Anthonomus grandis* (Cotton Boll Weevil) and *Spodoptera frugiperda* (Fall Armyworm) to a Cry1Ia-type toxin from a Brazilian *Bacillus thuringiensis* strain. J. Biochem. Mol. Biol. 40, 773–782.

Guo, S., Ye, S., Liu, Y., Wei, L., Xue, J., Wu, H., Song, F., Zhang, J., Wu, X., Huang, D., Rao, Z., 2009. Crystal structure of *Bacillus thuringiensis* Cry8Ea1: An insecticidal toxin toxic to underground pests, the larvae of *Holotrichia parallela*. J. Struct. Biol. 168, 259–266.

Höfte, H., Whiteley, H.R., 1989. Insecticidal crystal proteins of *Bacillus thuringiensis*. Microbiol. Rev. 53, 242–255.

Kasman, L.M., Lukowiak, A.A., Garczynski, S.F., McNall, R.J., Youngman, P., Adang, M.J., 1998. Phage display of a biologically active *Bacillus thuringiensis* toxin. Appl. Environ. Microbiol. 64, 2995–3003.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. 26, 283–291.

Lassner, M., Bedbrook, J., 2001. Directed molecular evolution in plant improvement. Curr. Opin. Plant Biol. 4, 152–156.

Marzari, R., Edomi, P., Bhatnagar, R.K., Ahmad, S., Selvapandiyan, A., Bradbury, A., 1997. Phage display of a *Bacillus thuringiensis* CryIA(a) insecticidal toxin. FEBS Lett. 411, 27–31.

Mendonça, A.F., 1982. A broca gigante Castnia licus Drury, 1770 (Lepidóptera: Castniidae) no Brasil. Saccharum APC, São Paulo.

Mendonça, A.F., Viveiros, A.J.A., Sampaio, F.F., 1996. A broca gigante da cana-de-açúcar, Castnia licus Drury, 1770 (Lep.: Castniidae). Pragas da cana-de-açúcar. Insetos, Cia., Maceió., pp. 133–167.

Pacheco, S., Gómez, I., Sato, R., Bravo, A., Soberón, M., 2006. Functional display of *Bacillus thuringiensis* Cry1Ac toxin on T7 phage. J. Invertebr. Pathol. 92, 45–49.

Patten, P.A., Howard, R.J., Stemmer, W.P.C., 1997. Applications of DNA shuffling to pharmaceuticals and vaccines. Curr. Opin. Biotechnol. 8, 724–733.

Rader, C., Steinberger, P., Barbas III, C.F., 2001. Selection from Antibody Libraries. In: Barbas III, C.F., Burton, D.R., Scott, J.K., Silverman, G.J. (Eds.), Phage Display: A laboratory manual. Cold Spring Harbor Laboratory Press, New York, pp. 10.1–10.20.

Sali, A., Blundell, T.L., 1993. Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. 234, 779–815.

Schnepf, E., Crickmore, N., Van Rie, J., Lereclus, D., Baum, J., Feitelson, J., Zeigler, D.R., Dean, D.H., 1998. *Bacillus thuringiensis* and its pesticidal crystal proteins. Microbiol. Mol. Biol. Rev. 62, 775–806.

Scott, J.K., Barbas III, C.F., 2001. Phage-display Vectors. In: Barbas III, C.F., Burton, D.R., Scott, J.K., Silverman, G.J. (Eds.), Phage Display: A laboratory manual. Cold Spring Harbor Laboratory Press, New York, pp. 2.1–2.19.

Silva-Werneck, J.O., Ellar, D.J., 2008. Characterization of a novel Cry9Bb d-endotoxin from *Bacillus thuringiensis*. J. Invertebr. Pathol. 98, 320–328.

Staden, R., 1996. The staden sequence analysis package. Mol. Biotechnol. 5, 233–241.

Stemmer, W.P.C., 1994. Rapid evolution of a protein in vitro by DNA shuffling. Lett. Nat. 370, 389–391.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Willats, W.G.T., 2002. Phage display: practicalities and prospects. Plant Mol. Biol. 50, 837–854.

Wolfersberger, M., Luethy, P., Maurer, A., Parenti, P., Sacchi, F.V., Giordana, B., Hanozet, G.M., 1987. Preparation and partial characterization of amino acid transporting brush border membrane vesicles from the larval midgut of the cabbage butterfly (*Pieris brassicae*). Comp. Biochem. Phisiol. 86, 301–308.

Xiang, W.F., Qiu, X.L., Zhi, D.X., Min, Z.X., 2009. N546 in β18-β19 loop is important for binding and toxicity of the *Bacillus thuringiensis* Cry1Ac toxin. J. Invertebr. Pathol 101, 119–123.

Zhang, X., Candas, M., Griko, N.B., Taussig, R., Bulla Jr., L.A., 2006. A mechanism of cell death involving an adenylyl cyclase/PKA signaling pathway is induced by the Cry1Ab toxin of *Bacillus thuringiensis*. Proc. Natl. Acad. Sci. U.S.A. 103, 9897–9902.

Zhao, H., Arnold, F.H., 1997. Opmitization for DNA shuffling for high fidelity recombination. Nucleic Acids Res. 25, 1307–1308.