The distinct properties of natural and GM cry insecticidal proteins

Jonathan R. Latham, Madeleine Love & Angelika Hilbeck

To cite this article: Jonathan R. Latham, Madeleine Love & Angelika Hilbeck (2017) The distinct properties of natural and GM cry insecticidal proteins, Biotechnology and Genetic Engineering Reviews, 33:1, 62-96, DOI: 10.1080/02648725.2017.1357295

To link to this article: http://dx.doi.org/10.1080/02648725.2017.1357295

© 2017 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Published online: 13 Sep 2017.

Article views: 1290

View related articles

View Crossmark data
The distinct properties of natural and GM cry insecticidal proteins

Jonathan R. Latham, Madeleine Love and Angelika Hilbeck

The Bioscience Resource Project, Ithaca, NY, USA; Independent Scholar, Victoria, Australia; Swiss Federal Institute of Technology (ETH), Institute of Integrative Biology, Universitätstrasse, Zurich, Switzerland

ABSTRACT

The Cry toxins are a family of crystal-forming proteins produced by the bacterium Bacillus thuringiensis. Their mode of action is thought to be to create pores that disrupt the gut epithelial membranes of juvenile insects. These pores allow pathogen entry into the hemocoel, thereby killing the insect. Genes encoding a spectrum of Cry toxins, including Cry mutants, Cry chimaeras and other Cry derivatives, are used commercially to enhance insect resistance in genetically modified (GM) crops. In most countries of the world, such GM crops are regulated and must be assessed for human and environmental safety. However, such risk assessments often do not test the GM crop or its tissues directly. Instead, assessments rely primarily on historical information from naturally occurring Cry proteins and on data collected on Cry proteins (called ‘surrogates’) purified from laboratory strains of bacteria engineered to express Cry protein. However, neither surrogates nor naturally occurring Cry proteins are identical to the proteins to which humans or other nontarget organisms are exposed by the production and consumption of GM plants. To-date there has been no systematic survey of these differences. This review fills this knowledge gap with respect to the most commonly grown GM Cry-containing crops approved for international use. Having described the specific differences between natural, surrogate and GM Cry proteins this review assesses these differences for their potential to undermine the reliability of risk assessments. Lastly, we make specific recommendations for improving risk assessments.

Introduction

Bacillus thuringiensis (Berliner) is a bacterium found in diverse ecological niches and may be ubiquitous in distribution (de Maagd, Bravo, & Crickmore, 2001). It produces a varied array of entomopathogenic compounds effective against a
broad range of arthropods. These include crystal proteins (Cry), cytolitic proteins (Cyt), vegetative insecticidal proteins (Vip), secreted insecticidal protein (Sip) and exotoxins – each with different characteristics, specificities and modes of action (de Maagd, Bravo, Berry, Crickmore, & Schnepf, 2003; Økstad & Kolstø, 2012; Schnepf et al., 1998).

Since its discovery over a hundred years ago, research into *B. thuringiensis* has been motivated primarily by its potential for pest control (Hilbeck & Otto, 2015; Økstad & Kolstø, 2012; Sanchis, 2010). Until the 1950s, *B. thuringiensis* was considered taxonomically to be a variety of *Bacillus cereus*, as was *Bacillus anthracis* (Oh, Ham, & Cox, 2012; Økstad & Kolstø, 2012). Based on their genetic analyses, Helgason et al. (2000) still postulated that all three ‘should be considered as belonging to one and the same species’ since the principal difference between *B. cereus*, *B. anthracis* and *B. thuringiensis* is only that the latter produces plasmids encoding crystalline endotoxins. *B. thuringiensis* is considered primarily a gut pathogen of arthropods, nematodes and protozoa (de Maagd et al., 2001; Durmaz, Hu, Aroian, & Klaenhammer, 2016; Wei et al., 2003). However, it can also be a human gut pathogen (McIntyre, Bernard, Beniac, Isaac-Renton, & Naseby, 2008; Oh et al., 2012; Ramarao & Sanchis, 2013; Wilcks et al., 2008). *B. cereus* is a well-known pathogen of mammals, including humans (Ramarao & Sanchis, 2013).

The life cycle of *B. thuringiensis* consists of vegetative and stationary phases (Lambert & Peferoen, 1992). Cells grow in vegetative mode so long as nutrients are available but form endospores within sporangia under unfavourable conditions. Coinciding with sporulation, large inclusion bodies develop that consist of one or more proteins of the crystalline (ie Cry type), or the cytotoxic (Cyt) type (Crickmore et al., 1998; de Maagd et al., 2003). In this review, we focus only on the entomopathogenic Cry (crystal) toxins of *B. thuringiensis*.

The presumptive, but still disputed, biological role of these Cry proteins is to facilitate invasion by *B. thuringiensis* of live host gut tissues (de Maagd et al., 2003; Guillem & Porcar, 2012). A summary of the standard understanding of their mode of action is that the crystal, which is biologically inactive, is progressively disaggregated, solubilised and enzymatically processed, via an inactive but soluble protoxin, into a much-truncated protein capable of binding to insect midgut epithelial receptors. Receptor binding greatly facilitates the creation of pores in the midgut membrane whose result is epithelial lysis and death of the host (Adang, Crickmore, & Jurat-Fuentes, 2014; Vachon, Laprade, & Schwartz, 2012). However, receptor binding is probably not a fundamental requirement for pore formation, at least by Cry1, Cry2, Cry3 or Cry5 proteins since pore formation occurs with synthetic membranes *in vitro* (Kao et al., 2011; Peyronnet et al., 2002; Schwartz et al., 1997; Slatin, Abrams, & English, 1990).

Crickmore has developed a nomenclature for Cry proteins based on amino acid sequence similarities, (Crickmore et al., 1998). So far, 74 Cry classes have been listed in that online database (Crickmore et al., 1998). Within this system, each Cry class is generally considered specific against one (or a few) insect taxonomic
orders. Thus, members of the Cry1 class (such as Cry1Ab) are considered active primarily against larval stages of the order Lepidoptera and toxins of the Cry3 class against larvae of Coleopteran species (Crickmore et al., 1998; de Maagd et al., 2003).

Historically, crystal and spore preparations of distinct strains of *B. thuringiensis* have seen use as biocontrol agents in forestry, in agriculture, and in public health applications against vectors of human diseases, such as mosquitoes (van Frankenhuyzen, 2013). Even in the industrial agriculture systems of North America, *B. thuringiensis*-based insecticides were widely used between the 1980s until the mid 1990s. At that time, heavy reliance on synthetic pesticides had led to pest resistance outpacing the development of new pesticides and so public and private research into Cry proteins experienced an unprecedented surge (Sanchis, 2010).

Subsequently, their use in agriculture was largely supplanted by: (1) the introduction of neonicotinoids (now suspended in the EU) (Kollmeyer et al., 1999) and (2) genetic engineering of crop plants which express Cry proteins within the plant. This GM approach overcomes some of the limitations to the efficacy of natural *B. thuringiensis*-based insecticides, which include their rapid inactivation due to UV light and rain (Behle, McGuire, & Shasha, 1997).

Most commercial Cry toxin-expressing genetically modified (GM) crops, hereafter called Bt crops, are varieties of maize and cotton. Widely used Bt crops are YieldGard maize which is based on insertion event MON810 (Cry1); Syngenta’s Bt11 maize (Cry1), marketed as Agrisure; and Bollgard II cotton whose insertion event (MON15985) contains a cry1 and a cry2 toxin gene. More recently, Bt soybeans have been commercialised in Latin America (Monsanto’s MON87701 and Dow’s DAS-81419–2), while Bt eggplants (aubergines) are undergoing field-testing in Bangladesh.

Most commercial Bt crops utilise proteins of the Cry1, Cry2 or Cry3 classes. In them, *B. thuringiensis*-derived sequences coding for Cry proteins are flanked by promoter and terminator sequences, usually from micro-organisms or viruses. Each transgene insertion (which typically has more than one cry gene) is termed an ‘event’. Each insertion event is normally the subject of an individual regulatory application. This varies, however, depending on the country and nature of the transgene. In the USA, USDA calls these ‘Petitions for deregulation’, while Cry proteins are regulated by EPA as plant-incorporated protectants (PIPs) and successful application results in ‘registration’.

So far, in commercial agriculture, up to six Cry proteins have been combined in a single cultivar, with Cry toxins directed against both lepidopteran and coleopteran pests (Hilbeck & Otto, 2015). In large part, this introduction of multiple Cry proteins in a single cultivar is recent and reflects the need to maintain resistance against pests that are continuously evolving (Carrière, Fabrick, & Tabashnik, 2016). A second approach to circumvent pest resistance has been to create hybrid Cry proteins. An example is Dow’s MXB-13 cotton, which contains elements from three distinct Cry1 proteins (Dow Agrosciences, 2003a; Table 1). Different
<table>
<thead>
<tr>
<th>GM event name/s, crop, company, FSANZ approval date</th>
<th>Transgene (I or II)</th>
<th>Primary reference(s)</th>
<th>Native source protein/s (where applicable) and DNA accession code/s (if applicable)</th>
<th>Source protein amino acid number (where applicable) – putative molecular weight (kDa) (c)</th>
<th>Putative plant protein amino acid number – predicted molecular weight (kDa) (c)</th>
<th>Reported molecular weights (kDa); tissue type (extraction method)</th>
<th>Unremarked proteins observed in the experimental figures (d)</th>
<th>Notes/observations</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bt 176, Maize, Syngenta, 2001</td>
<td>Bt-176</td>
<td>Citia Seeds (1994)</td>
<td>Cry1Ab3 (AAA25651) (1)</td>
<td>1155 aa–130.6 kDa</td>
<td>648 aa–72.6 kDa</td>
<td>Leaf (crude): 6-5 kDa (doublet); 6-5 kDa (2); 36 kDa (4); Pollen (Cry1A): 6-5 kDa (3)</td>
<td>No figures provided (4)</td>
<td>(2) Trypticized fragment; (3) Commerm at aa 25; (4) The 36 kDa band commenced at aa 31</td>
<td>(1) Fischhoff et al. (1987)</td>
</tr>
<tr>
<td>MON 810, Maize, Monsanto (2000)</td>
<td>MON1810</td>
<td>Monsanto (1995)</td>
<td>Cry1Ab10 (A29125.1) (1)</td>
<td>1156 aa–130.7 kDa</td>
<td>615 aa–68.9 kDa</td>
<td>Leaf (crude/fractionated): images provided are indiscernible</td>
<td>(2) Trypticized fragment; (3) Commerm at aa 25; (4) The 36 kDa band commenced at aa 31</td>
<td>(1) Fischhoff et al. (1987)</td>
<td></td>
</tr>
<tr>
<td>T10-40, Cotton, Bayer, 2010</td>
<td>T10-40</td>
<td>Bayer Crop Science (2008)</td>
<td>Cry1Ab5 (CA28405) (1)</td>
<td>1155 aa–130.7 kDa</td>
<td>617 aa–72.6 kDa</td>
<td>Leaf (crude): 63 kDa</td>
<td>No figures provided (4)</td>
<td>(2) Trypticized fragment; (3) Commerm at aa 25; (4) The 36 kDa band commenced at aa 31</td>
<td>(1) Seubert et al. (1996)</td>
</tr>
<tr>
<td>DDT108, Black, Monsanto (2002)</td>
<td>DDT108</td>
<td>Monsanto (2000)</td>
<td>Cry1Ac1 (A22331.1) (1)</td>
<td>1178 aa–133.3 kDa</td>
<td>633 aa–68.7 kDa</td>
<td>Leaf (crude): 66 kDa (2); (3)</td>
<td>No figures provided (4)</td>
<td>(2) Adang et al. (1985)</td>
<td></td>
</tr>
<tr>
<td>COT 246, Cotton, Syngenta, 2009</td>
<td>COT246</td>
<td>Syngenta (2007)</td>
<td>Full length Cry1A (P/Cry1A Delocalised C) (A22331.1) (1) + 26 aa from Cry1Aa (3)</td>
<td>1181 aa–133.6 kDa</td>
<td>611 aa–68.7 kDa</td>
<td>Leaf (crude): 63 kDa (3); Despite the truncation, the full length protein (1.3 kDa) was reported by FDA, FSANZ, and Health Canada (4); Various tissue: 63 kDa core protein; (2) 63 kDa (2); 100 kDa, 70 kDa (2); 69, 43 kDa (2); Images provided are indiscernible</td>
<td>No protein analysis information available onlinePetition to APHIS (5)</td>
<td>(1) Fischhoff et al. (1987)</td>
<td></td>
</tr>
<tr>
<td>MON 531 (Bollgard I)</td>
<td>MON531</td>
<td>Monsanto (1994)</td>
<td>Cry1Ac46 as of Cry1Ab10 (A29125.1) (1) + 712 aa of Cry1Ac1 (A22331.1) (2)(3)</td>
<td>1178 aa–133.1 kDa</td>
<td>617 aa–68.7 kDa</td>
<td>Leaf (crude): 66 kDa (2); (3)</td>
<td>No figures provided (4)</td>
<td>(2) Trypticized fragment; (3) Commerm at aa 25; (4) The 36 kDa band commenced at aa 31</td>
<td>(1) Fischhoff et al. (1987)</td>
</tr>
<tr>
<td>MON 87701, Soy, Monsanto, 2010</td>
<td>MON87701</td>
<td>Monsanto (2000)</td>
<td>Cry1Ac-CTP (98aa) (1) + 496 aa as of Cry1Ab10 (A29125.1) (2) + 712 aa of Cry1Ac1 (A22331.1) (1)(3)(4)</td>
<td>1178 aa–133.1 kDa</td>
<td>617 aa–68.7 kDa</td>
<td>Leaf (crude): 66 kDa (2); (3)</td>
<td>No figures provided (4)</td>
<td>(2) Trypticized fragment; (3) Commerm at aa 25; (4) The 36 kDa band commenced at aa 31</td>
<td>(1) Fischhoff et al. (1987)</td>
</tr>
<tr>
<td>Mon 87701, Soy, Monsanto, 2010</td>
<td>MON87701</td>
<td>Monsanto (2000)</td>
<td>Cry1Ac-CTP (98aa) (1) + 496 aa as of Cry1Ab10 (A29125.1) (2) + 712 aa of Cry1Ac1 (A22331.1) (1)(3)(4)</td>
<td>1178 aa–133.1 kDa</td>
<td>617 aa–68.7 kDa</td>
<td>Leaf (crude): 66 kDa (2); (3)</td>
<td>No figures provided (4)</td>
<td>(2) Trypticized fragment; (3) Commerm at aa 25; (4) The 36 kDa band commenced at aa 31</td>
<td>(1) Fischhoff et al. (1987)</td>
</tr>
<tr>
<td>Mon 89034, Maize, Monsanto, 2008</td>
<td>Monsanto (2006)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------------------</td>
<td>----------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘Cry1A.105’ Can be designated (1) 466 aa of Cry Ab10 (A29125) (2) + 5 aa of Cry CA1 (AAB23130) (3) + 129 aa of Cry F1a (AAB23146) (4) + 573 aa of Cry Ac1 (AAB23133) (5)</td>
<td>α = 133.3 kDa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>na (synthetic fusion protein)</td>
<td>( \approx 133.3 ) kDa</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grain (fractionated); bands &gt;250, 130, 85 kDa (6); 56 kDa (7)</td>
<td>Grain (fractionated); different band patterns dependent on preparation and detection procedures. Including ( \approx 110-140 ) kDa (7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1) Monsanto's description was not specific, e.g. ‘a chimeric protein that consists of domains i and ii from Cry1Ab or Cry1Ac, domain iii from Cry1E and the C-terminal domain from Cry1Ac’ (5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Monsanto noted a substitution (L to S at aa 765) in the Cry1Ac portion of the protein (6)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>referred to as a ‘proteolytic fragment’ (7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>see Figures B-1; B-2; B-3 (7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>‘Cry (Ac) synpro’ 612 aa of Cry1Ac (AAB2331) (1) + 36 aa of Cry1Ca3 (AAB2343) + 508 aa of Cry1Ab1 (AAB2330) (2) (3)</td>
<td>α = 130.7 kDa</td>
</tr>
<tr>
<td>na (synthetic fusion protein)</td>
<td>1156 aa–130.7 kDa</td>
</tr>
<tr>
<td>Grain (crude): 130 kDa (1)</td>
<td>Plant (AP): ‘full-length Cry1Ac could not be detected’ (4) 65 kDa</td>
</tr>
<tr>
<td>Grain (crude): ‘doublet’ at ~68 kDa (1) (2)</td>
<td>Plant (AP): ‘many potential Cry bands’ (6)</td>
</tr>
<tr>
<td>Grain (crude): underexposed gel (3)</td>
<td>Grain (crude): ‘130 kDa and fragments that contain the active core toxin’</td>
</tr>
<tr>
<td>Grain (crude): Many potential Cry bands (4)</td>
<td>Grain (crude): ‘truncated Cry1Ac core protein’</td>
</tr>
<tr>
<td></td>
<td>(1) Conflicting nucleotide advice provided by applicant (h)</td>
</tr>
<tr>
<td></td>
<td>(3) See Figure 50</td>
</tr>
<tr>
<td></td>
<td>(4) N-terminal analysis ‘truncated Cry1Ac sample’ found to begin at aa 29</td>
</tr>
<tr>
<td></td>
<td>(5) Figure 15 (6) Figure 16</td>
</tr>
<tr>
<td></td>
<td>(1) Adang et al. (1985) (2) Wabiko, Raymond, and Bulla (1986)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>‘Cry (Fv3) synpro’ 604 aa of Cry1Fa2 (AAB2347) (1) + 36 aa of Cry1Ca3 (AAB2343) + 508 aa of Cry1Ab1 (AAB2330) (2) (3)</td>
<td>α = 130.2 kDa</td>
</tr>
<tr>
<td>na (synthetic fusion protein)</td>
<td>1148 aa–130.2 kDa</td>
</tr>
<tr>
<td>Leaf, pollen, whole plant, grain (crude): ‘Doublet’ at ~68 kDa (1) (2)</td>
<td>Images provided were indecipherable</td>
</tr>
<tr>
<td>Grain (crude): ‘Doublet’ at 53, 49, 43, 30, 17 kDa (5) (6)</td>
<td>(1) The sequence ‘blocked’ a protein beginning at the 28th amino acid was identified (2) Images were too poor to correlate with the reports. ‘Doublet’ not observable. See Figures 11 and 12</td>
</tr>
<tr>
<td></td>
<td>(3) Conflicting descriptions offered by the applicant (h)</td>
</tr>
<tr>
<td></td>
<td>(4) ‘100% identical in amino acid sequence’ to the synthetic ‘cry1Ac’ protein in MXB-13 lines</td>
</tr>
<tr>
<td></td>
<td>(5) Figure 41</td>
</tr>
<tr>
<td></td>
<td>(6) Gel is unclear due to lack of controls. See Figure 42</td>
</tr>
<tr>
<td></td>
<td>(1) N-terminal sequence ‘blocked’. Protein beginning at the 28th amino acid was identified (2) Images were too poor to correlate with the reports. ‘Doublet’ not observable. See Figures 11 and 12</td>
</tr>
<tr>
<td></td>
<td>(3) Conflicting descriptions offered by the applicant (h)</td>
</tr>
<tr>
<td></td>
<td>(4) ‘100% identical in amino acid sequence’ to the synthetic ‘cry1Ac’ protein in MXB-13 lines</td>
</tr>
<tr>
<td></td>
<td>(5) Figure 41</td>
</tr>
<tr>
<td></td>
<td>(6) Gel is unclear due to lack of controls. See Figure 42</td>
</tr>
<tr>
<td></td>
<td>(1) Wabiko et al. (1986) (2) Conflicting nucleotide advice provided by applicant (h)</td>
</tr>
<tr>
<td></td>
<td>(4) N-terminal analysis ‘truncated Cry1Ac sample’ found to begin at aa 29</td>
</tr>
<tr>
<td></td>
<td>(5) Figure 15 (6) Figure 16</td>
</tr>
<tr>
<td></td>
<td>(1) Adang et al. (1985) (2) Wabiko, Raymond, and Bulla (1986)</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>GM event name/s, crop, company, FSANZ approval date, transgene (I or II)</th>
<th>Primary reference (a)</th>
<th>Native source protein/s (where applicable) and DNA accession code/s (b)</th>
<th>Source protein amino acid number (where applicable) – putative molecular weight (kDa) (c)</th>
<th>Putative plant protein amino acid number predicted molecular weight (kDa) (c)</th>
<th>Reported molecular weights (kDa); tissue type (extraction method)</th>
<th>Unremarked proteins observed in the experimental figures (d)</th>
<th>Notes/observations</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>MON89034, Maize, Monsanto, 2008 (MON89034 produces two novel proteins. See 1st entry above)</td>
<td>Monsanto (2006), FSANZ (2008)</td>
<td>CTP (1) + Cry2Ab2 (CAA87087) (2), or Cry2Ab1 (AA22342) (3) (4)</td>
<td>634aa + 79 aa of CTP – Uncleaved: 79.3 kDa Partially cleaved CTP (3aa) + Cry = 71.2 kDa Cry protein alone = 70.9 kDa</td>
<td>71 kDa, 50 kDa (called band 1 and band 2) (5) (6)</td>
<td>61 kDa (7) (8)</td>
<td>Grain (AMP, fractionated): fainter bands higher and lower than bands 1 and 2 can be seen: 160, 140, 75, 40 kDa, 35, 30, 12.5 kDa (9)</td>
<td>(4) Conflicting nucleotide advice provided by applicant (b) (5) Monsanto never declared an experimentally determined MM for the plant Cry protein (despite reporting MM’s for the transgenic Cry1A.105 protein, also in the crop) (6) N-terminal sequencing: band 1 ‘blocked’ but identified a protein beginning at aa 24, FSANZ describing it as a minor portion of the protein. Analysis of band 2 identified it as beginning at aa 115 (7) FSANZ described the predicted MM as 61 kDa. However, the MW reported by Monsanto and the theoretical MM agree on 71 kDa (8) A sequence starting from the 24th amino acid was observed, which FSANZ described as ‘a minor portion of the protein co-migrating with the full-length protein’ (9) Figure 1-3</td>
<td>(1) Matsuoka, Kano-Murakami, Tanaka, Ozeki, and Yamamoto (1987) (2) Dankocik, Donovan, and Jany (1998) (3) Widmer and Whiteley (1989)</td>
</tr>
<tr>
<td>MON5983, Bollgard II Cotton, Monsanto, 2002 (MON5983 produces two novel proteins. See 1st entry above)</td>
<td>Monsanto (2000), FSANZ (2002)</td>
<td>CTP (1) + Cry2Ab2 (CAA87087) (2), or Cry2Ab1 (AA22342) (3) (4)</td>
<td>634aa + 79 aa of CTP – Uncleaved: 79.3 kDa Fully cleaved: 70.9 kDa</td>
<td>70.9 kDa (5)</td>
<td>No figures in the APHIS documentation. Text reported ‘leaf protein extracts’ of ~63 kDa (5)</td>
<td>(4) Conflicting nucleotide advice provided by applicant (b) (5) FSANZ appeared confused about the engineering of the protein. They mistakenly reported that a 61 kDa protein would be expected after cleavage of the Chloroplast Transitol Peptide</td>
<td>(3) One of the higher MW bands is ‘believed to be a protein dimer and the other is not related to the protein’ (4) Figure 8 (5) Figure 13</td>
<td>(1) Van den Broeck et al. (1985) (2) Dankocik et al. (1990) (3) Widmer and Whiteley (1989)</td>
</tr>
<tr>
<td>GHB19, Cotton, Bayer, 2011</td>
<td>Bayer CropScience (2008)</td>
<td>CTP (1) + Cry2Ae1 (AA212362) (2)</td>
<td>631 aa + 55 aa of CTP – Fully cleaved: 70.9 kDa</td>
<td>Leaf (AMP): 65 kDa and other higher (MW) bands (5) ‘same lower MW bands’ Seeds (AMP): ‘the same as the Cry/A standard’ (6) Seeds (AMP): 65 kDa, also 19 and 17 kDa (5)</td>
<td>Leaf (AMP): also 150, 28, 19 and 17 kDa, also other indirect bands higher and lower (4) Seeds (AMP): 65 kDa, also 19 and 17 kDa (5)</td>
<td>(3) No protein analysis was presented. Very little information available. FSANZ to the plant derived protein sample, a second band was present and was thought to represent a protein that had been degraded in the plant cell” (4) Figure 8 (5) Figure 13</td>
<td>(1) De Almeida et al. (1989) (2) Baum, Chu, Donovan, Gilmer, and Rugar (2003)</td>
<td></td>
</tr>
<tr>
<td>MIR604, Maize, Syngenta, 2006</td>
<td>Syngenta (2006), FSANZ, 2006</td>
<td>mCry3A (1) was re-engineered: The first 47 aa were removed, 3 aa were replaced by 4 aa (2)</td>
<td>598 aa – 67.7 kDa</td>
<td>CBI-Deleted (3)</td>
<td>CBI-Deleted (3)</td>
<td>27</td>
<td>(2) Petition to APHIS provided enough information to deduce the amino acid sequence of the protein, but did not provide the sequence. (3) No protein analysis was presented. (4) No protein analysis was presented. Very little information available. FSANZ to the plant derived protein sample, a second band was present and was thought to represent a protein that had been degraded in the plant cell”</td>
<td>(1) Sekar et al. (1987)</td>
</tr>
<tr>
<td>Bt Event</td>
<td>Genotype</td>
<td>Source</td>
<td>Reference(s)</td>
<td>Description</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>--------</td>
<td>--------------</td>
<td>-------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S307, Maize, Syngenta (2012)</td>
<td>Cry3AAb, 22aa synthetic + 459aa of Cry3A (AAA22356) (1) + 172aa of Cry1Ab3 (AAA22357) (2)</td>
<td>na</td>
<td>Leif (Crude, IAP): A molecular weight consistent with the predicted molecular weight of 73.7 kDa (3) (A)</td>
<td>Leaf (Crude, IAP): 62 kDa (3) Syngenta didn’t name the protein. The only band showing on the Western Blot image was at ~62 kDa (4) Syngenta didn’t state the MW at which its test proteins migrated</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MON863, Maize, Monsanto (2001)</td>
<td>Cry3Bb variant (AAA22334) (1)</td>
<td>652 aa–74.4 kDa 653 aa–74.5 kDa</td>
<td>Grain (AP): Multiple bands ranging from 66.3 to 74.6 kDa at 214000 kDa (5)</td>
<td>Grain (AP): complex band patterns but inadequate controls to verify</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MON88017, Maize, Monsanto (2004)</td>
<td>Cry3Bb variant (AAA22334) (1)</td>
<td>652 aa–74.4 kDa 653 aa–74.5 kDa</td>
<td>Grain (AP): 77.2 kDa (3), 66.2 kDa (3), 55.4 kDa (3), 46 kDa (4)</td>
<td>Grain (AP): bands also visible at 44, 40 kDa (2) (3) (iaP): ? (5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAS-59122, Maize, Dow (2005)</td>
<td>Cry3Ab1 variant (AAA41671) (1)</td>
<td>123 aa–13.6 kDa</td>
<td>Leaf (Crude): 14 kDa (2) (3)</td>
<td>Leaf (Crude): 14 kDa (2) (3) (iaP): 15 (5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DAS-59122, Maize, Dow (2005)</td>
<td>Cry3Ab1 variant (AAA41672) (1)</td>
<td>383 aa–43.8 kDa</td>
<td>Leaf (Crude): 44, 40 kDa (2)</td>
<td>Leaf (Crude): plus bands at 36 and 30 kDa (4) (5)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(a) This is the reference for the entire row unless otherwise specified.
(b) The data in this column required considerable reconciliation. The Neil Crickmore Bt toxin database (Life Sciences, University of Sussex) was a primary and confirmation reference source, augmenting developer and regulator information.
(c) Molecular Weights (MW) estimated by the ExPASy online ProtParam tool based on putative amino acid sequence (Gasteiger et al., 2003).
(d) This column represents the author's interpretations of the same data and figures in the previous column.
(f) These putative protein sizes come from Hernández et al. (2003) and Rosati et al. (2008).
(g) Health Canada (1997); ANZFA (n.d.); FDA (1996).
(h) The discrepancy is 1 nucleotide.
Cry protein domains affect different pests. Therefore, hybrid Cry proteins can combine different domains into a single event, to combat or delay pest resistance and facilitate breeding.

Trends of wider commercialisation and transgene stacking are leading to increasing exposure to Cry toxins which may be produced in green tissues, roots, seeds and pollen. In these tissues, concentrations vary widely, ranging from below the detection limit to 15 μg/g fresh weight (Nguyen & Jehle, 2007; Székács, Lauber, Juracsek, & Darvas, 2010).

The ‘history of safe use’ as incorporated into Cry toxin regulatory safety

The crystalline insecticides purified from *B. thuringiensis* have the common reputation of being fairly safe for the environment due to their limited range of species toxicity (Behle et al., 1997; van Frankenhuyzen, 2013). Following this reputation, regulators and applicants often state, or imply, that standard Cry toxin preparations (hereafter ‘wild-type Cry proteins’) have ‘a history of safe use’ which is presumed to carry over to GM Bt crops. Thus, FDA’s *Biotechnology Consultation Note* on MON810, dated September 1996, reads:

> Monsanto states that the cryIA(b) protein present in MON809 and MON810 is identical to that present in nature and commercial microbial preparations approved by the Environmental Protection Agency (EPA). (FDA, 1996)

And the company itself wrote:

> Using modern biotechnology, Monsanto has developed insect-protected YieldGard corn, event MON 810, that produces the naturally occurring *Bacillus thuringiensis* (Bt) protein, Cry1Ab. (Monsanto, 2002)

The same assumption has been used by regulators and in public communications. In 2011, the Australia/New Zealand GM regulator (FSANZ) equated GM plant and wild-type Cry proteins in a press release:

> [It is] an insecticidal protein Cry1Ab that is produced by the naturally occurring soil bacterium *Bacillus thuringiensis* sub sp. kurstaki (Btk). The gene encoding this protein has been used to genetically modify some crops so that they contain the protein and are thus protected against certain insect pests. The protein is also extensively used in organic and conventional farming as a direct application pesticide. (FSANZ, 2011)

Even more explicit are the opening statements of Dow AgroSciences, 2012 application to USDA DAS-81419-2 soybean (now approved):

> Cry1Ac and Cry1F have a long history of safe use. The proteins originate from the naturally occurring soil bacterium *B. thuringiensis*. The safety of the proteins has been demonstrated in sprayable Bt formulations for pest control in agriculture for over half a century … (Dow AgroSciences, 2012)

Similar statements cover other Cry classes such as Cry3 (eg Monsanto, 2004) and Cry34Ab1 and Cry35Ab1 (Dow Agrosciences, 2004).
Such statements carry the strong implication that data collected on Cry toxins produced in and purified from *B. thuringiensis* are applicable to GM crop risk assessment.

Though much relied upon, this ‘history of safe use’ is not well defined or elaborated by its users in the regulatory system. It will be seen that it consists of both a claim and an assumption. The safe use claim applies to nontarget organisms and human health, and the assumption is that extrapolations can be made from it. The specific claim is not questioned in this review. However, there are reports of human toxicity, allergies and apparent sensitisation to wild-type Cry proteins (Bernstein et al., 1999; Finamore et al., 2008; Mezzomo et al., 2013; Moreno-Fierros, García, Gutiérrez, López-Revilla, & Vázquez-Padrón, 2000; Torres-Martínez et al., 2016). It is the appropriateness of that extrapolation, however, that is the subject of this review.

This assumption has been previously questioned by numerous authors who have noted that extrapolating from wild-type *B. thuringiensis* Cry toxins to toxins produced in a GM plant contradicts the standard theory of Cry toxin activation (eg Goldburg & Tjaden, 1990; Hilbeck, Moar, Pusztai Carey, Filippini, & Bigler, 1998b; Hilbeck & Otto, 2015; Hilbeck & Schmidt, 2006; Székács et al., 2010; Toll, 1988). Prevailing understanding predicts that Cry proteins expressed in Bt crops may have a broader host range and enhanced toxicity than wild-type proteins for two reasons. One reason is that wild-type Cry proteins are tightly bound within crystalline inclusion bodies and are in that form inactive, whereas all GM plant Cry toxins exist in soluble forms. Secondly, wild-type Cry proteins require multiple additional proteolytic steps to convert them into the activated toxin. However, many GM transgenic events (Bt11, Bt-176, TC1507, DBT418 and T304) express potential activated forms. In such plants, no activation steps may be required. Since both solubilisation and proteolysis are activation steps that require highly specific conditions (eg of high pH and specific proteases) that are not met by many potentially affected organisms, GM plant Cry proteins may have broader host ranges or greater toxicity. As Toll (1988) expressed it in reference to solubilisation: ‘[Bt crops] bypass a chemical containment mechanism that limits exposure to a narrow range of species’.

Wild-type Cry protein crystals vary in shape between bipyramidal, cuboidal and rhombooidal forms (Bietlot et al., 1989; de Maagd et al., 2003). Their detailed physical structures have been relatively little studied, but the crystals are known to be complex (Ai, Li, Feng, Li, & Guo, 2013; Clairmont, Milne, Pham, Carrière, & Kaplan, 1998; Schernthaner, Milne, & Kaplan, 2002). As proposed by Clairmont and colleagues, the basic form of naturally occurring crystals is somewhat virus-like in that multiple Cry proteins are attached via their amino termini to a single molecule of DNA that is approximately 20 Kbp in length. The exact nature of this DNA may vary. Xia et al. (2005) reported that the DNA component of the crystal contained ‘the promoter, the coding region, and the terminator of a Cry1Ac gene’, although both Bietlot et al. (1993) and Sun, Wei, Ding, Xia, and
Yuan (2007) reported a heterologous nature for the DNA. Others have reported that the DNA–protein complex contained ‘plasmids harboured by the host strain’ (Chaturvedi, Bhakuni, & Tuli, 2000). The DNA–crystal complex itself, at least for Cry1, is approximately $2 \times 10^6$ Da in size (Bietlot et al., 1993). The full-sized crystal, however, must be an aggregate of these DNA–protein complexes and is held together by disulphide bonds (Clairmont et al., 1998).

This structure is somewhat speculative, but regardless of the details, any stable complex structure implies the existence of a complex activation process that is highly dependent on the physicochemical structure of the crystals and not just on the Cry protein amino acid sequence (Clairmont et al., 1998).

Given these differences between wild-type and GM Cry proteins, a key question is to determine the number of potential containment steps for wild-type Cry proteins of each class.

According to Clairmont, the first activation step is release of the crystal from the bacterial sporophyte as a result of either physical destruction or germination (Clairmont et al., 1998). The second step is for an individual crystal to disaggregate. Disaggregation (in the case of Cry1 proteins) requires a high pH to cleave disulphide bridges between amino acids. The resulting complex is the $2 \times 10^6$ Da unit of 10–20 Cry protein molecules that are organised along the strands of DNA (Bietlot et al., 1993). In the next step, intestinal enzymes (trypsin, chymotrypsin, pepsin and other gut proteases) trim the carboxy terminus of the Cry proteins (Carroll et al., 1997). Also probably required for this step are DNAses to internally cut the DNA scaffold. Thus, the alternating action of DNAses and proteases releases individual Cry molecules that nevertheless still have a short length of DNA attached to them (Bietlot et al., 1993; Clairmont et al., 1998). This structure is further processed at each end to yield the final activated toxin of around 65 kDa (for Cry1) (de Maagd et al., 2003; Vachon et al., 2012).

Focus on the structure and disassembly steps of Cry crystals emphasises that the toxicological differences between solubilised shortened GM plant Cry proteins and wild-type Cry crystals are potentially profound. The complex higher order structure, after all, explains why crystals are inactive, and why they can remain dormant without degrading and why each wild-type Cry toxin is activated only under the highly specific chemical conditions of the gut of susceptible organisms. Even more than Toll (1988) can have known, each disassembly stage is a potential containment step. This is an understanding that reinforces the 1990 recommendation that ‘Activated delta endotoxin as expressed by B.t.k. plants nevertheless should be tested as a new agent’ rather than be assumed to have the toxicity of the wild-type crystal (Goldburg & Tjaden, 1990).

**The role of surrogate proteins in Cry toxin risk assessment**

Surrogate Cry proteins are those purified from GM bacterial strains such as *Escherichia coli* or *Pseudomonas fluorescens*. They are typically proteins intended
to be identical in sequence and length to those expressed in GM plants (Bialy, 1987). (The exceptions are MON 810, Dow’s Cry1F from MXB-13 cotton and the Cry3Bb1 in Monsanto’s MON863—see later). The purpose of surrogates is to obviate the need in risk assessment assays for whole Bt crop tissues or purified Cry proteins isolated from plants, since purifying Cry proteins from plants can be difficult due to their sometimes low abundance (Freese & Schubert, 2004). Surrogate proteins are the test material most commonly used to assess Cry protein biodegradation, Cry protein digestion by mammals and acute toxicity towards mammals and other nontarget organisms. By far the majority of studies contributing to risk assessment thus rely on surrogates (Freese & Schubert, 2004).

The willingness of regulators to accept submissions reliant on surrogate Cry proteins has been repeatedly critiqued for assuming the identity of plant and bacterial Cry proteins. In 2000, the US National Academies of Science wrote: ‘Tests should preferably be conducted with the protein as produced in the plant’. (NAS, 2000). Freese and Schubert (2004) called the use of surrogate proteins a ‘serious mistake’. An external scientific panel convened by EPA (SAP MT, 2000) also criticised the use of surrogate proteins, as has an EU advisory committee and officials from various European environment agencies (Dolezel et al., 2011; EC, 2000).

Despite these criticisms, the acceptance of surrogates and of historical data from wild-type Cry crystals continues at the EPA and in Europe’s EFSA. To give one example, in its 2014 Biopesticides Registration Document for the DAS-81419-2 soybean, EPA accepted a Cry protein purified from P. fluorescens in toxicity studies with the honeybee (Apis mellifera), a parasitic hymenoptera (Nasonia vitripennis), the green lacewing (Chrysoperla rufilabris), other insects, earthworms, a fish and a bird (EPA, 2014). Similarly, when oral toxicity testing of mice with the surrogate had shown no effect, testing DAS-81419-2 soybean for toxicity towards wild mammals was deemed unnecessary.

In the same application, EPA also accepted surrogate Cry proteins (rather than soybean leaves) for a study concluding that Cry1Ac and Cry1F proteins degrade rapidly in soils. EPA also accepted a study simulating mammalian gastric digestion of surrogate Cry protein. The subsequent conclusion of rapid disappearance of surrogate Cry proteins in soils and mammalian guts later became the justification for bypassing various tests on ecosystem toxicity, as well as for EPA’s conclusion that humans would not be exposed to Cry proteins originating from the use of DAS-81419-2 beans.

Thus, except for one field experiment with soybean leaves, EPA, USDA, FSANZ and EFSA were entirely reliant on historical data or on surrogate Cry proteins for their approval of DAS-81419-2 soybeans expressing Cry1Ac and Cry1F.

In summary, the concept of a ‘history of safe use’ and the adoption of surrogate Cry proteins incorporate similar but questionable assumptions; namely that, in comparison to the GM Cry protein present in the Bt crop, surrogate Cry proteins (or the wild-type Cry proteins) are unaltered with respect to diverse parameters
associated with potential harm: toxicity, specificity, allergenicity, bioavailability and persistence.

To examine the use of these assumptions more thoroughly than has so far been attempted, is the purpose of this review. Since, to a significant extent, this is also the purpose of risk assessments for commercial Bt crops, much of the data in this review derives from such risk assessments.

For useful critiques of other aspects of nontarget risk assessment as applied to Bt crop plants, we refer the reader to Andow and Hilbeck (2004), Freese and Schubert (2004), Hilbeck and Schmidt (2006), Lövei, Andow, and Arpaia (2009), Dolezel et al. (2011) and Hilbeck, Meier, and Trtikova (2012).

**Scope, methods, reference materials and selection parameters of this review**

The starting point for this analysis was to catalogue the evidence of intentional or accidental molecular and structural changes to the Cry proteins associated with the engineering process in commercialised Bt maize, cotton and soybean crops. Thus, we were interested in any alterations or changes that potentially could affect risk parameters of Cry proteins and, therefore, invalidate assumptions of identity or equivalence between a GM plant Cry protein and a surrogate Cry protein or between a GM plant Cry protein and wild-type Bt crystals. Changes considered, therefore, included DNA sequence changes, amino acid substitutions, amino acid sequence additions (e.g., chloroplast transit peptides), other protein size changes (due to any cause) and other covalent modifications uncovered during testing (such as glycosylation).

The initial basis for our choice of GM Bt crops were the regulatory summary documents produced by FSANZ (formerly ANZFA), which is the regulatory authority responsible for GM food safety in Australia and New Zealand. A key reason for this choice was that Bt crops approved by FSANZ are the most likely to be mixing in global trade. Where already available to us, the full GM company dossiers submitted to FSANZ were also used (Monsanto’s MON810, MON863 and MON89034, Syngenta’s Cot67B and 5307 and Dow’s TC1507).

On finding that the FSANZ summary documents did not sufficiently nor always accurately reflect the data in GM company applications, additional use was made of documents submitted to the Animal and Plant Health Inspection Service APHIS (part of USDA) and the US EPA, as well as commercial patents and the peer-reviewed scientific literature. The cut-off date for consideration was FSANZ approvals prior to 1 July 2015. This selection process also covers all GM Bt crops submitted to the EFSA. From our extensive past experience in evaluating dossiers submitted to EFSA and the US regulators, we know that the submitted dossiers to all these regulators are essentially scientifically identical, regardless of the differences in regulatory procedures – the US deregulating/exempting or registering – and the EU approving, GMOs (Dolezel et al., 2011).
Table 1 is intended to summarise the regulatory characterisation of each GM plant Cry protein. Table 1 also documents the specific origins of the Cry proteins used in these Bt crops. It additionally contains information on protein detection, size fractionation on polyacrylamide gels and amino acid sequence analysis. It contains the full list of event approvals considered in this study. All were major commodity crops, i.e. maize (corn), soybeans and cotton. Where one event expresses multiple Cry proteins, each protein is accorded a single line of Table 1.

Table 1 presents the data as assessed by the applicant and, less commonly, by the regulator. However, claims by applicants were frequently unsupported, or contradicted by, the data in the application. In the case of MON810, they were contradicted by subsequent peer-reviewed publications as well. Where an alternative interpretation was thought warranted, it is presented in column G, which represents our interpretation of the ‘in planta’ protein data. In spite of this potential for alternative interpretations, the written text (outside of Table 1) always reflects the conclusions drawn by the applicant, unless clearly indicated.

The second aspect of the study was to review this information in the light of a comprehensive understanding of the structures and mechanisms of action of Cry proteins.

This review is not a full critique of Cry toxin risk assessment. Assessments incorporate other data, such as bioinformatic evidence and compositional analysis, as part of their evidence gathering and which is not considered here.

**How Cry proteins in plants differ from wild type**

**DNA base alterations**

DNA from bacterial species typically does not support efficient protein production in plant cells due to codon usage differences (Murray, Lotzer, & Eberle, 1989). Before being inserted into plants, cry gene sequences are therefore extensively mutated (de Maagd et al., 2003; Perlak, Fuchs, Dean, McPherson, & Fischhoff, 1991). These mutations are intended to allow high Cry abundance and thereby to increase toxin efficacy and delay insect resistance. Typical is the cry1Ab transgene introduced into MON810 maize (Monsanto, 1995). It is 2448 DNA base pairs (bp) in length, of which 709 bp differ from the original B. thuringiensis wild-type sequence (Fischhoff et al., 1987; Hernández et al., 2003; Monsanto, 2000). None of the 23 distinct cry transgene containing events approved by FSANZ (see Table 1) thus possesses the DNA sequence of the wild-type source.

Such DNA changes (with the exception of chimaeric and hybrid proteins) typically aim to leave amino acid composition unaltered. An exception is the Q349R substitution in MON863 that was introduced to facilitate DNA manipulation (Monsanto, 2001). Nevertheless, such DNA codon usage substitutions may be of relevance to risk assessment since optimising DNA for eukaryotes affects the potential for horizontal gene transfer. This possibility is not a focus of this review, however.
**Amino acid substitutions in Cry proteins**

Many *cry* transgenes encode amino acid composition substitutions compared to wild type. Some are unintended. The Monsanto Cry1Ac events MON531 (cotton), MON87701 (soy) and MON89034 (maize) all contain an L to S amino acid substitution at position 765 or 766. They are all derived from one mutation described by Monsanto as ‘unintentional’ (see Table 1) (Monsanto, 1994).

In recent years, there has been a trend towards intentional substitutions. An example is MON863. To create MON863 maize the wild-type (*cry3Bb1*) gene was engineered to produce what the Monsanto patent calls a ‘second-generation molecule … with increased activity’ (English et al., 2000). Five amino acids were substituted (D165G, H231R, S311L, N313T and E317K) to create a protein that Monsanto called Cry3Bb1.11098. As a consequence of the Q349R substitution mentioned above and also the insertion of an extra alanine residue at position 2, the intended protein differs by seven amino acids from the wild-type (Monsanto, 2001).

In the same class of enhanced Cry proteins is the Cry3Bb1 of MON88017 which differs from Cry3Bb1.11098 only in lacking the D165G substitution (Monsanto, 2006).

Although not featured in Table 1, the Cry9C protein in Starlink maize (produced by Aventis) was modified to enhance its resistance to degradative enzymes and, therefore, to increase its stability in plant cells (Bucchini & Goldman, 2002). Similarly, Syngenta’s 5307 maize contains a Cry3A variant (called eCry3.1Ab) for which amino acids were substituted to create a cathepsin G protease recognition site. The altered amino acids are V155A, S156A and S157P (Syngenta, 2011).

**Major amino acid changes and chimaeric proteins**

A very common deliberate alteration among GM plant Cry proteins is truncation. Syngenta’s Bt11 maize contains a Cry1Ab10 reduced from a wild-type length of 1156 amino acids (aa) down to 615aa. In Syngenta’s Bt-176 corn the Cry1Ab is cut from 1155aa to 648aa; Dow’s TC1507 corn contains Cry1Fa2 cut from 1174aa to 605aa); Bayer’s T304-40 cotton, containing Cry1Ab5 is reduced from 1155aa to 617aa (Bayer Crop Science, 2008; Ciba-Geigy, 1994; Dow Agrosciences, 2000; Northrup King, 1995). All are Cry1 class proteins and in each case large sections of the carboxy terminal end of the protein were purposefully removed, for a variety of reasons (See also Table 1).

A further significant number of Bt crops contain *cry* transgenes that encode chimaeric proteins (Table 1). Among them is Syngenta’s COT67B cotton. (Syngenta, 2007). The COT67B Cry, designated as ‘FLCry1Ab’, is a hybrid of Cry1Ab3 and 26aas of Cry1Aa. This addition Syngenta considers to be a ‘repair’ of Cry1Ab. Another chimaera is Monsanto’s MON89034 maize (Monsanto, 2006). MON89034 encodes a protein designated as ‘Cry1A.105’. It is a fusion of three
partial proteins, Cry1Ab10, Cry1Ac1 and Cry1Fa1. Another protein chimaera is Dow’s DAS-81419-2. This soybean contains two separate transgenes both of which are chimaeric. One of these, designated Cry1Ac or ‘cry1Ac(synpro)’ encodes a fusion protein synthesised from parts of Cry1Ac1, Cry1Ca3 and Cry1Ab1 (Dow AgroSciences, 2012). The second transgene, though called Cry1F (or sometimes Cry1Fv3), is a chimaera of Cry1Fa2, Cry1Ca3 and Cry1Ab1. Bayer’s GHB119 cotton encodes a Cry2Ae1 protein fused to a chloroplast transit peptide (CTP) from Arabidopsis thaliana (Bayer Crop Science, 2008). Syngenta’s 5307 maize event encodes a fusion of a modified Cry3Aa2, part of a Cry1Ab3 and a 22-amino acid N-terminal synthetic amino acid sequence (Syngenta, 2011). The 22 amino acid sequence is the accidental result of a PCR-induced mutation that created a frame-shift (Syngenta, 2011).

In a somewhat separate class of major alterations is Monsanto’s MON810 maize. The Cry1Ab gene used in the transformation had an open reading frame 3468 bp in length (Table 1). However, Monsanto later informed US regulators and FSANZ that ‘MON810 contains a less than full length CryIA(b) gene’. Independent analysis showed that only 2448 bp had inserted into the plant genome (Hernández et al., 2003). The latter third of the transgene had been lost during the DNA transformation process (Wilson, Latham, & Steinbrecher, 2006). Investigating further, various authors proposed that the truncated Cry1Ab protein extends by 2 or 18 amino acids as a consequence of the open reading frame extending into the maize genome DNA flanking it (Hernández et al., 2003; Rosati, Bogani, Santarlasci, & Buiatti, 2008).

Unintended Cry protein modifications in plants

In the course of typical Bt crop risk assessment, data are collected by applicants on various properties of Cry proteins produced in plants (hereafter ‘GM plant Cry proteins’). The regulatory presumption, sometimes stated explicitly, is that such data will confirm that the GM plant Cry protein is identical either to surrogate protein or to wild-type forms and this identity justifies, for example, not testing the effects of the crop itself on nontarget organisms.

However, as Table 1 makes clear, it is the norm to observe differences that imply plant-specific modification of Cry proteins. This is exemplified by Bayer’s GHB119 cotton (Bayer Crop Science, 2008). GHB119 cotton contains a single cry transgene encoding Cry2Ae fused to a CTP sequence. When extracted from leaf tissues, however, Cry2Ae-specific fragments were observed at five distinct sizes with molecular weights of approximately 150, 65, 28, 19 and 17 kDa on SDS polyacrylamide gels (see Table 1). Except for the fainter 28 and 19 kDa bands, each of the bands is similarly prominent, suggesting cotton leaf tissues contain approximately equal amounts. Other indistinct Cry2Ae-specific bands were also present. Of the five polypeptides, only the 65 kDa band co-migrates with the surrogate Cry2Ae protein (purified from E. coli). Seeds of GHB119 cotton, however,
contain only three Cry2Ae-specific polypeptides (65, 19, 17 kDa) (Bayer Crop Science, 2008). Thus, fragments that are larger or smaller than the surrogate are commonly observed.

It is also common for Bt crops to produce (or at least contain) Cry-specific polypeptides none of which co-migrate with the surrogate on polyacrylamide gels. In extracts from Dow’s MXB-13 cotton, for example, Cry-specific polypeptides were detected for both Cry1Ac1/Cry1Ca3/Cry1Ab1 (sometimes called by Dow Cry1Ac) and Cry1Fa2/Cry1Ca3/Cry1Ab (sometimes called by Dow Cry1F) but no Cry protein prepared from plant material co-migrated with either full-length surrogate (Dow Agrosciences, 2003a, 2003b).

As with Bayer’s GHB119, numerous Cry events provide evidence for the existence in plants of Cry forms with higher than predicted weights. For example, protein extracts from MON89034 maize contained a form of Cry1A.105 at 250 kDa plus other polypeptides at between 56 and 130KD, with some polyacrylamide gel bands being indistinct or otherwise difficult to interpret (Monsanto, 2006). Likewise, DAS-59122 maize Cry34Ab1 exhibits forms at 60, 50, and 42 kDa in addition to the expected 13.6 kDa protein (Dow Agrosciences, 2004). For MON87701 a ‘faint’ band was reported at 250 kDa (Monsanto, 2009). For MON863 maize (Cry3Bb1), Monsanto reported a protein at 220 kDa that was detected by Cry-specific antibodies, although the applicant did provide some evidence this was not a Cry protein (Hileman et al., 2001).

In summary, not one of the events approved by FSANZ exhibited just a single Cry-specific protein band co-migrating with the surrogate. The three potential exceptions are MIR604 maize, DBT418 maize and MON15985 (Bollgard II cotton), where the relevant data are not publicly available. For two additional events, MON810 and TC1507, the information provided was inadequate for any analysis. Thus, five lines lacked the information to form a judgement.

Even when Cry proteins extracted from GM plants migrate as apparent single bands on polyacrylamide gels, more than one protein form may still be present. For example, in the application for MON863, Monsanto extracted from a gel the protein band that migrated at the same weight as the surrogate (produced in E. coli). Using N-Terminal sequence analysis they concluded that, unlike the surrogate, the GM plant Cry protein (Cry3Bb1) had three distinct start sites – at amino acids 19, 25 and 36 (Table 1; Monsanto, 2001).

Applicants typically propose that these unexpected protein forms are artefacts of the extraction process. Smaller forms are often designated as ‘degradation products’ or ‘proteolytic fragments’ or ‘attributable to degradation’, while higher weight forms are often described as ‘protein dimers’ or ‘aggregates’ (eg Bayer Crop Science, 2008; Monsanto, 2001; Syngenta, 2007). Alternatively, applicants may refer to unexplained bands as ‘impurities’, even when detected by their specific antibodies and not present in control lanes (DAS-59122; Dow, 2004). In the single instance where a degradation hypothesis was tested, however, Ciba-Geigy
concluded that degradation during extraction played no part in generating the additional Cry forms observed (Ciba-Geigy, 1994).

Ciba–Geigy’s test implies the need for other explanations for Cry-related polypeptides of unexpected sizes. These might include: (1) unexpected pre-translational processes, such as mRNA splicing; (2) post-translational modification of Cry proteins; and (3) complex or otherwise aberrant transgene insertion events, which are well documented (Wilson et al., 2006). The example of Syngenta’s MIR604 corn suggests how such possibilities might occur. The surrogate used by Syngenta was found to have an alternative upstream translational start site, allowing two protein forms of the mCry3a protein, both equally abundant, to be produced in *E. coli* (EPA, 2010).

Ciba–Geigy’s test of its explanation also highlights a more general issue. The procedures used to determine whether or not GM plant Cry proteins are chemically identical to their surrogates suffer from scientific defects that, at best, prevent independent interpretation and verification. Size markers may be missing (eg Monsanto, 2001 p46), positive and negative controls, especially untransformed controls, are often lacking (eg Dow Agrosciences, 2004; Dow AgroSciences, 2012), procedures and reagents may be incompletely described (all applications), images may be indecipherable or have low resolution (eg Dow Agrosciences, 2000; Dow AgroSciences, 2012; Monsanto, 2001), or key results claimed as confidential (eg Monsanto, 2000; Syngenta, 2006). In consequence, on the basis of the experiments presented, it is often not possible to independently confirm or refute assertions of the applicants.

For example, extracts from gel bands are often further characterised by MALDI-TOF and N-terminal sequencing. The results are typically used to claim that GM plant Cry proteins are ‘similar’ or ‘equivalent’ or ‘substantially similar’. However, as occurred with Bt-11 maize, the plant-derived protein excised from the gel is typically the one that migrated closest to the surrogate Cry protein (Northrup King, 1995). Thus, in almost every case, Cry protein forms that were novel or otherwise unexpected, and thus in explicit need of characterisation, were omitted from further analysis. MON863 represents one of very few cases where N-terminal sequencing was performed on unexpected size fragments. The applicant partially sequenced two bands but still neglected the remainder of the Cry-specific plant polypeptides (Table 1; Monsanto, 2001). Thus, both applicants and regulators appear to have largely lost sight of the fact that establishing the absence of unexpected fragments in plants is the goal. Verification of expected molecules is not.

Further problems of interpretation and verification arise when methods used to extract GM Cry proteins from plants are highly complex or involve the use of antibodies. For example, the initial purification steps of Cry2Ae (in Bayer’s cotton event GHB119) passed plant extracts over an affinity column incorporating a monoclonal antibody raised against the Cry2Ae protein (Bayer Crop Science, 2008). Since monoclonal antibodies recognise just a single epitope, Cry forms produced by the plant but lacking that epitope, eg because they are truncated,
will be absent from all subsequent procedures. Monsanto’s use of a polyclonal antibody raised against the N-terminal 58 amino acids (<9%) of Cry3Bb1 to isolate Cry3Bb1 from plant extracts will similarly miss Cry polypeptides lacking that region (Monsanto, 2001). Monsanto’s use of an antibody raised against the N-terminal 14 amino acids of Cry1A.105 is the most selective of all (Monsanto, 2006).

An equal problem is the use of such selective antibodies at the detection stage, which is usually a western blot procedure. Monoclonal or polyclonal antibodies, especially those raised against a less than full length Cry protein, will again be unable to detect a subset of Cry-derived fragments or forms. Since a different antibody is usually used at the purification stage, the detection of any partial length Cry forms in plants can in principle be made impossible simply by a judicious choice of antibodies, especially if the second antibody is raised against a different terminus. Since many antibody reagents are poorly described it is not possible to know whether this scenario has occurred. In these and the other cases of interpretive ambiguity noted in this review, applicants have the opportunity to make more data available and this would resolve many of these uncertainties.

Therefore, we conclude that, at the very least, all applicants are biasing their analyses of plant tissue extracts to reduce the possibility of detection of partial or otherwise aberrant Cry proteins and so obscuring from regulators the full range of Cry protein forms in plants. A similar critique applies to applicants who treat plant extracts with the protease trypsin. Trypsin treatment was applied to Bt11, Bt-176 and MON810 plant extracts. The rationale for trypsin treatment is that insect guts contain trypsin proteases (eg Northrup King, 1995 Appendix H), but this rationale seems weak, given the disadvantage that trypsinisation risks degrading the unexpected Cry forms whose detection is purportedly the goal of risk assessment.

Lastly, many petitions characterise plant Cry proteins using N-terminal protein sequencing. However, such protein sequencing of GM plant Cry proteins is often found to be ‘blocked’. Such blocking is normally considered indicative of post-translational protein modification, usually acetylation. Blocking was observed for the proteins Cry1Ab5 of T304-40; Cry1Fa2 of TC1507; Cry2Ab2 of MON89034; Cry3Bb1 of MON863; and eCry3.1Ab of 5307 maize. In two of the cases (eCry3.1Ab and MON863), acetylation of the blocked amino acid residue was definitively confirmed (Monsanto, 2001; Syngenta, 2011). One event tested positive for glycosylation. This was Cry2Ae1 from GHB119, though the single datum is open to interpretation (Figure 11; Bayer Crop Science, 2008).

We conclude, therefore, that FSANZ-approved Bt crops usually contain a mixed complement of Cry toxins. No GM plant Cry protein was structurally or chemically identical, either to its wild-type precursor or to its bacterial surrogate. In every case there was at least one verifiable difference, and usually many more.
Are there biological implications of protein differences in GM plants?

As previous authors have pointed out, the logical weaknesses of risk assessment of Bt crops relying on historical evidence and surrogates are twofold. On the one hand, the transfer of a Cry protein into a crop – or its expression within that crop – may result in structural and chemical differences that render the surrogate an unreliable substitute for the plant protein (EC, 2000; Freese & Schubert, 2004; Goldburg & Tjaden, 1990; NAS, 2000; SAP MT, 2000). Secondly, the history of safe use may be invalidated by the differences between wild-type and plant proteins. For example, *B. thuringiensis* Cry proteins are crystalline, whereas GM plant Cry proteins are not. Furthermore, many commercial plant Cry proteins (see Table 1) are truncated in such a way that the carboxy-terminal domain that inhibits toxicity is lost (eg Dolezel et al., 2011; Hilbeck et al., 2012; Hilbeck & Otto, 2015; Toll, 1988). Both solubilisation and truncation are expected to enhance toxicity.

The data described above and in Table 1 amplify and considerably extend such critiques. First, by showing that, beginning from their crystal structure, the activation process of wild-type proteins is more complex and more reliant on specific conditions than risk assessments acknowledge, and second, by showing that applicants are introducing proteins that are more radically changed than commonly supposed.

These differences between GM plant Cry proteins and surrogates or wild-type proteins serve to emphasise why direct comparisons of insecticidal activity of GM plant Cry proteins and surrogates are particularly important, though only in some risk assessments was this test performed. These tests show directly that many surrogates differ toxicologically from the protein purified from the GM plant. The Cry1Ab of Bt11 had identical activity when isolated from plants or bacteria as did the Cry2Ab of MON89034 (Monsanto, 2006; Northrup King, 1995). However, the toxicity of Cry1A.105 purified from MON89034 corn plants was reported to be twice (EC$_{50}$ of 0.0074 ± 0.0017 μg protein/ml diet) that of its surrogate purified from *E. coli* (EC$_{50}$ of 0.012 ± 0.0062 μg) (FSANZ, 2008; p35). Previously, Ciba–Geigy had also reported that Cry1Ab protein isolated from Bt-176 had a much higher toxicity than the surrogate. This result was consistent in tests on European corn borer (*Ostrinia nubilalis*) (5-fold higher), cabbage looper (*Trichoplusia ni*) (10-fold higher) and Corn earworm (*Helicoverpa zea*) (5-fold higher) Ciba-Geigy, 1994). Syngenta reported a similar result with the Cry1Ab of COT67B, with the plant protein having toxicity that was higher by fourfold (an LC$_{50}$ of 1.3 vs. 5.2 ng/cm$^2$) (Syngenta, 2007). In every case where there was a difference, the plant-derived Cry protein had greater activity. Notably, evidence for differential activity of surrogates and GM plant Crys was limited to Cry1 class toxins. Whether differential toxicity of surrogates is limited to the Cry1 class is an important thesis that in future should be tested.

These results concur with other evidence, again primarily for Cry1 class proteins. Wild-type Cry1 proteins are considered by most authors as active only
against Lepidoptera (van Frankenhuyzen, 2013). However, this can depend on protein processing. When Cry1Ba1 crystals were dissolved and truncated in vitro by trypsin digestion, ie to resemble more closely those made in plants, Bradley and colleagues reported a marked increase in toxicity against coleopterans. Most notable was a 50-fold increase in activity against the Colorado Potato Beetle (Leptinotarsa decemlineata) (Bradley, Harkey, Kim, Biever, & Bauer, 1995). These authors also showed that solubilised Cry3A was fourfold more active and Cry1b 10-fold more active against the coleopteran Chrysomela scripta. The authors inferred from these results that host specificity is broadened by the process of activation. Solubilisation does not inevitably enhance toxicity, however. For Cry1B toxicity towards Manduca Sexta the opposite was observed by the same authors (Bradley et al., 1995).

In subsequent work, Hilbeck and colleagues (1998b; Hilbeck, Baumgartner, Fried, & Bigler, 1998a and Hilbeck, Moar, Pusztai-Carey, Filippini, & Bigler, 1999) reported that Cry1Ab from GM maize, also supposedly lepidopteran-specific, was toxic towards a neuropteran, the larvae of green lacewing (Chrysoperla carnea), when administered via nontarget prey (Spodoptera littoralis) that had fed on Cry1Ab GM maize plants. Toxicity was less, however, when the same prey species had fed on artificial diet spiked with a surrogate Cry1Ab toxin or when the surrogate Cry1Ab toxin was administered directly (Hilbeck, 2002). While some have argued that higher lacewing mortality might have resulted from suboptimal prey quality – as a result of its exposure to the Cry toxin (Dutton, Klein, Romeis, & Bigler, 2002), toxic effects were also observed when Cry1Ab toxin was administered directly to the lacewings and no prey species was involved (Hilbeck et al., 1998b). Such an explanation also does not account for the differences in toxicity observed between the different sources of the Cry protein (Hilbeck, 2002; Hilbeck & Schmidt, 2006). These results thus show that toxicity of Cry1Ab is dependent on factors other than its coding sequence and was the highest when produced in a GM plant.

Furthermore, other researchers have also reported a significant toxicity of GM plant-derived Cry1 proteins towards nonlepidopterans. For example, on coccinellid (coleopteran) predators, both when administered directly (Dhillon & Sharma, 2009; Hilbeck et al., 2012; Schmidt, Braun, Whitehouse, & Hilbeck, 2009) or via unaffected aphid prey (Moser, Harwood, & Obrycki, 2008; Zhang, Wan, Lövei, Liu, & Guo, 2006). These results again contradict the expectation that Cry1, when produced in maize, will remain lepidopteran-specific.

Although there has been relatively little regulatory testing of GM Cry1 toxicity outside the order lepidoptera, other research groups have nevertheless expanded the list of insect orders sensitive to GM Cry1 proteins. This includes a significant activity of GM maize debris (MON810, Cry1Ab) against caddisflies (trichopterans) (Chambers et al., 2010; Rosi-Marshall et al., 2007); while others reported activity of MON810 (Cry 1Ab) against water fleas (Daphnia magna) (Bohn, Primicerio, Hessen, & Traavik, 2008; Bohn, Traavik, & Primicerio, 2010; Ferreira-Holderbaum
et al., 2015) and of Bt-176 against the Colembolan Folsomia candida (Ciba-Geigy, 1994) (for an extensive review see Hilbeck & Otto, 2015).

Additionally, there are reports that, when produced in plants, Cry1 proteins can have substantially heightened toxicity towards species already known to be susceptible to them. Thus, Lang and Vojtech (2006) observed that Syngenta’s Bt-176 maize (no longer on the market) had extremely high activity (an LD_{50} of 14 pollen grains) against European swallowtail caterpillars (Papilio machaon L.). Similarly, in a comparison of the toxicity of Cry1Ab from MON810 maize with that of a commercial formulation of Cry1Ab (DiPel) against strains of European corn borers that were normally resistant to DiPel, the GM protein was more than two hundred times more active (Li et al., 2007).

Evidence for altered, enhanced and broadened toxicity compared to wild-type Cry proteins is therefore substantial but largely limited to Cry1 class proteins. Perhaps this is because Cry1 proteins are the most widely used and most studied class in Bt crops.

The implications of these results are threefold: firstly, Cry1 proteins differing only in the organisms from which they were sourced may differ in toxicity, by more than 100-fold. Second, this altered toxicity can manifest either as enhanced toxic activity towards known targets or as novel toxicity towards other insect orders. Thirdly, when compared side-by-side, GM plant Cry proteins have consistently been the more toxic form. Hence, the data broadly agree with the predictions of authors such as Toll (1988) and Dolezel et al. (2011) that progressive activation, whether it occurs by solubilisation or truncation, such as occurs in Bt crops, can broaden specificity and increase the activity of Cry toxins.

The proposition above is that Cry1 toxicity is widened in GM plants. It needs a caveat, however. As documented by van Frankenhuyzen, some wild-type Cry1 toxins contradict the standard expectation of lepidopteran specificity (van Frankenhuyzen, 2013). Other wild-type Cry toxin classes also sometimes show activity towards multiple or unexpected species from diverse classes. Thus, it still needs to be definitively shown that wild-type Cry proteins chosen for commercialisation have the limited specificity claimed for them. Some, but not all, of the apparently broadened specificity of GM plant Cry toxins might therefore be due to inadequate testing of wild-type Cry proteins.

Questions surrounding the appropriateness of surrogate proteins are not limited to direct toxicological implications. Since Bt crops were widely commercialised it has been reported that Cry1Ab embedded in GM plant material can persist in soils for months (eg Zwahlen, Hilbeck, Gugerli, & Nentwig, 2003; Zwahlen, Hilbeck, & Nentwig, 2007). Other authors have concluded that Cry1Ab remains in aquatic systems for long periods (eg Douville, Gagné, Blaise, & André, 2007; Douville, Gagné, Masson, McKay, & Blaise, 2005; Tank et al., 2010). Cry1Ab proteins may also reach the human and foetal bloodstream and are also found in the intestines of pigs (Aris & Leblanc, 2011; Chowdhury et al., 2003; Walsh et al., 2011). By implying lengthy Cry protein survival in real environmental conditions,
these results contradict the invariant conclusions of risk assessments – carried out with surrogate Cry proteins – that concluded persistence of Cry toxins would be measured in days or less, and that Cry protein survival in the gut was brief (EPA, 2010). Other explanations for this discrepancy are possible, for example, the standardised conditions used by applicants may be unrepresentative of the environments they seek to mirror, but the use of surrogate proteins is another potential source of what is an important inconsistency.

A separate reason to doubt toxicological equivalence between wild-type and GM plant Cry proteins is provided by a Monsanto patent (English et al., 2000). In its petition for the transgenic event MON863, Monsanto emphasised that its Cry3Bb1.11098 toxin had 98.9% identity (seven amino acid differences) to the Cry protein of the commercial insecticidal product Raven™. Monsanto stated also that ‘MON863 poses minimal risk to mammals, wildlife, and nontarget insects’ (Monsanto, 2001). EPA thus granted MON863 an ‘exemption from the requirement for a tolerance’ agreeing that the Cry3Bb1.11098 in MON863, though altered, ‘do[es] not differ significantly’ from Cry3 proteins already exempted (EPA, 2001).

Monsanto’s patent on Cry3Bb1.11098, in contrast, claims to have made ‘super’ toxins of which the Cry3Bb1.11098 in MON863 is the most potent (English et al., 2000). This patent measured Cry3Bb1.11098 as being 7.9-fold (690%) more active against corn rootworm species than wild-type Cry3Bb1. This enhancement of toxicity was the basis of the patent’s novelty claim. According to the patent, the proteins in it ‘have the combined advantages of increased insecticidal activity and concomitant broad spectrum activity’ (italics added).

The claim of ‘concomitant broad spectrum activity’ is a plausible one. None of the five amino acid substitutions deliberately made to Cry3Bb1.11098 were introduced with specific pests in mind. One change was intended to reduce the likelihood of nonproductive binding by Cry3Bb1 to random proteins in the insect gut. A second enhanced channel activity and pore formation. In other words, not only did the patent contradict statements the applicant had provided for risk assessment, it also contradicts a ubiquitous proposition of Cry protein risk assessment petitions: that amino acid similarity leads to similarity of toxicity. A second commercial event, MON88017 (maize, see Table 1) also contains an enhanced Cry3 “super” toxin’ covered by the same patent. For this protein too, the company claimed a history of safe use and EPA again granted a regulatory exemption (EPA, 2001; Monsanto, 2004). The patent thus confirms a central proposition of this review, that GM Cry proteins are not equivalent to, and are possibly substantively more toxic than their wild-type ancestors.

Some recommendations

Our findings and conclusions have significance for the regulation of GM crops. The first issue is that current procedures fail to address the use of Cry proteins that have no precise ancestral form, ie those that are mutated or enhanced, are
hybrid proteins or are in other ways synthetic. These are thus novel insecticides. Their purpose is increased toxicity and the thwarting of insect resistance (Brevault et al., 2013; English et al., 2000). But, whereas, they presumably possess either increased toxicological activity or a new host range, petitions make only a rudimentary attempt to acknowledge this novelty. For instance, in no case did we find their changed biological properties precisely described and hardly ever was a rationale for their novelty presented that might have been able to inform risk assessment. Instead, applicants typically assigned names derived from one of the component toxins and appeared to be assuming that toxicity followed from that chosen designation.

Cry proteins in GM plants typically differ in size in comparison to surrogates or wild-type proteins. These size/length differences have diverse origins, from deliberate truncation, accidental truncation of the DNA construct (the case of MON810) or, most commonly of all, they are due to unknown activities associated with Cry gene expression or post-translational modification in Bt crops. What is the toxicological significance of these size differences? The most obvious answer is that all full-length Cry proteins are nontoxic and require shortening to acquire toxicity. Additionally, it is known that size differences can result in differential toxicity. In mosquito (diptera) midguts, Haider, Knowles, and Ellar (1986) observed that Cry1 was processed into a smaller form that is toxic primarily to dipterans, whereas in the gut of lepidopterans the same protein was processed into a larger fragment toxic only to lepidoptera (Haider et al., 1986). Similarly, proteolytic cleavage of the Cry1Ac protein by their respective insect midgut proteases led to differential toxicity against distinct pest species (Lightwood, Ellar, & Jarrett, 2000). Cry3 and Cry9 proteins also demonstrate fragment-specific toxicity (eg Brunet et al., 2010; Guo, Zhang, Song, Zhang, & Huang, 2009; Rausell et al., 2004). Thus, size differences of the kind observed in plants can influence toxicity.

The further reason for focusing on protein fragments and especially unexpected fragment sizes is the need for risk assessment to take into account that transgenic events are selected from among hundreds or thousands for their high activity against pest species. The selection process represents an a priori reason to suppose that any novel or unexpected protein forms are the explanation of this high activity.

For these reasons, regulators should require that surrogate Cry proteins be identical to in size and chemical composition to GM plant Cry proteins before accepting them as substitutes. Since it is known that Cry toxins often vary by size within plants, according to organ and tissue type (see Table 1), diverse GM plant tissues must be examined before it is accepted that GM plant Crys and surrogates are identical in size.

A third significance of our analysis is that surrogate proteins are not the solution to the difficulty of concentrating Cry proteins they appear to be. The first problem is that, in order to be used, surrogates must be shown to be identical. Up to now, identity has typically been left undefined by regulators. In consequence, applicants
use their own terms (such as ‘equivalent’ or ‘substantially similar’) without defining them. In practice, despite clear evidence for differences between surrogates and plant Cry proteins (including primary amino acid differences, additional unexplained protein forms in plants, posttranslational modification and toxin activity in assays), in every case, applicants proposed and regulators accepted, that sufficient ‘identity’ or ‘equivalence’ was established. It seems that no observed difference is sufficient to refute the hypothesis of identity, and lacking a definition, likely never will be. Thus, even when an identical Dow construct (Cry1Fsynpro, having a 130 kDa surrogate) was used to engineer two separate crop species – and the resulting plant proteins were measured as 65 kDa in MXB-13 cotton but as 130 kDa in DAS-81419-2 soybean (see Table 1) – Dow nevertheless concluded that both plant Cry proteins were equivalent in size to the surrogate.

In our view, surrogates should only be used if they are scientifically indistinguishable. Detection of any unexpected or nonidentical Cry-related protein in plants should disqualify their use. Yet establishing identity between surrogates and plant Cry proteins is scientifically problematic since it requires showing the presence of the expected protein form(s), and the absence of unexpected forms, whose identities are, of course, unknown in advance. This is the unacknowledged challenge of using surrogates.

How this difficulty manifests in practice is that any biochemical assay reliant on purified, fractionated or trypsinised samples risks losing unexpected Cry molecule(s) produced in plants in one of the discarded fractions. The use of antibodies, either for purification or for detection, is also ruled out since these too make significant but untestable prior assumptions about the nature of the unknown Cry molecule(s). Without a purification step, however, how will low abundance molecules be detected? The obvious solution might seem to be a total proteomic analysis performed in the most unbiased way feasible (eg Agapito-Tenfen, Guerra, Wikmark, & Nodari, 2013). Such a proteomic approach would have the added advantage of potentially detecting any other unexpected protein molecules arising as a result of plant transformation that would be missed by the use of surrogates. But, while enormously more comprehensive than the risk assessment methodologies currently used, no proteomic analysis guarantees detecting the unexpected. In the final analysis, the use of surrogate proteins is therefore not a logical solution to the difficulty of purifying Cry proteins from plants or using plant materials. Surrogates merely introduce novel difficulties of their own.

The above observations lead us to make the recommendation that, to be scientifically defensible, the use of surrogates should be a last resort. At present, surrogates are sometimes used arbitrarily and without specific justification. An example is Syngenta’s estimate of environmental half-life in soil conducted for MIR604 corn. Syngenta used a purified surrogate toxin when using leaf or other plant tissues would have been both easier and more realistic (Syngenta, 2006).

From this example follows also the recommendation that justifications for the use of surrogates should always be at the level of individual experiments. For
example, it may be impossible to obtain sufficient Cry toxin from plants to dose a mammal but still be possible to dose a small insect species. Every use of a surrogate should specifically justify why whole tissues cannot be used. Whole tissue tests should otherwise be mandatory.

It should thus be noted that the use of surrogates is in strong contradiction to the prevailing science of Cry toxins. Ongoing research into the maturation processes of Cry proteins and their variants implies that truncations and even small changes can greatly alter activity and species specificity (Vachon et al., 2012). Indeed, many of the outcomes of this research are patented as novel and efficacious innovations (e.g., English et al., 2000). We hope that regulators and applicants involved in the risk assessment process for Cry proteins will work to bring their work in line with the prevailing science and revisit past risk assessments to resolve all such inconsistencies.

Our final recommendation, derived from all the above, and as originally suggested by Goldburg and Tjaden (1990), is that a GM plant Cry ‘should be tested as a new agent’ in terms of toxicity towards nontarget species. Historical data serves at present to blur or mask potential hazardous effects. The regulatory system should thus proceed from the assumption that, because GM plant Cry proteins are solubilised compared with wild-type molecules (which are crystalline), and because many are extensively modified by the time they are transferred into plants, and all are altered again inside those plants, they all constitute novel molecules.

We make these recommendations because there are excellent potential remedies for reliance on either surrogate Cry proteins or on data originating from historic use of wild-type proteins. One such example is a standard Caco-2 cell permeability assay. Caco-2 cells are human gut epithelial cells that are used to predict absorption, and inhibition of absorption by small molecules as part of drug discovery (van Breemen & Li, 2005). This assay could easily be adapted to predict Cry protein toxicity and estimate Cry interactions with human intestines. Moreover, such assays would require only small quantities of Cry toxin to be isolated from GM plants. These assays represent inexpensive, easy and direct measures of human toxicity.

In summary, we have demonstrated that GM plant Cry proteins are very different, both to naturally occurring Cry proteins and to surrogate Cry proteins. Secondly, we have outlined a strong *prima facie* case that introduction of a natural Cry protein into a plant can significantly enhance its toxicity towards both target and nontarget species. Indeed, the intention of many of the changes made to natural Cry proteins is to heighten or broaden their toxicity (English et al., 2000). Even quite subtle changes to Cry toxin composition, conformation and size have been shown, by research over the 20 years since Bt crops were first commercialised, to dramatically alter toxicity (Soberon et al., 2007; Vachon et al., 2012). In our view, the regulatory system, by its continuing heavy reliance upon historical evidence and surrogates, has failed to keep up, both with the commercial expansion of Cry
toxin modifications and with scientific developments into the subtleties of Cry toxicity and activity.

When they were first introduced it was widely expected that the introduction of GM Bt crops would lead to reductions in insecticide use. This expectation is still asserted by applicants today as balancing the risks of GM Bt crop introductions. Evidence so far suggests, however, that reductions in the application of conventional insecticides have been uneven and transient (e.g., Benbrook, 2012; Pemsl, Waibel, & Gutierrez, 2005). What our review draws attention to, however, is that measures of insecticide use, and also the claims for reduced insecticide use made by applicants, have disregarded the Cry insecticides in GM crops themselves. These toxins deserve much greater attention and may be of equal or greater concern than conventional pesticides.

Acknowledgements

We thank two anonymous reviewers for constructive comments that greatly improved this paper. We also thank Allison Wilson and Michael Hansen for comments and suggestions. We also appreciate FSANZ for sometimes making documents available to us.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

No external funds were used for this work.

ORCID

Madeleine Love http://orcid.org/0000-0003-1111-3177

References


Carroll, J., Convents, D., Van Damme, J., Boets, A. V., Van Rie, J., & Ellar, D. J. (1997). Intramolecular proteolytic cleavage of *Bacillus thuringiensis* Cry3A δ-endotoxin may facilitate its coleopteran toxicity. *Journal of Invertebrate Pathology, 70*, 41–49. doi:10.1006/jipa.1997.4656


Dhillon, M. K., & Sharma, H. C. (2009). Effects of *Bacillus thuringiensis* δ-endotoxins Cry1Ab and Cry1Ac on the coccinellid beetle, *Cheilomenes sexmaculatus* (Coleoptera, Coccinellidae) under direct and indirect exposure conditions. *Biocontrol Science and Technology, 19*, 407–420. doi:10.1080/09583150902783801


FSANZ. (2011). FSANZ response to study linking Cry1Ab protein in blood to GM foods.


polypeptides of Arabidopsis thaliana. Plant Molecular Biology, 11, 745–759. doi:10.1007/BF00019515


van Frankenhuyzen, K. (2013). Cross-order and cross-phyllum activity of *Bacillus thuringiensis* pesticidal proteins. *Journal of Invertebrate Pathology*, 114, 76–85. doi:10.1016/j.jip.2013.05.010


