Specificity-determining Regions of a Lepidopteran-specific Insecticidal Protein Produced by Bacillus thuringiensis*

H. Ernest Schnepf, Kathleen Tomczak, Jose Paz Ortega, and H. R. Whiteley‡

From the Department of Microbiology, University of Washington, Seattle, Washington 98195

The lepidopteran-specific, insecticidal crystal proteins of Bacillus thuringiensis vary in toxicity to different species of lepidopteran larvae. We report studies of CryIA(a) and CryIA(c), two related proteins that have different degrees of toxicity to Heliothis virescens yet very similar degrees of toxicity to Manduca sexta. The amino acid differences between these proteins are located primarily between residues 280 and 722. We have constructed a series of chimeric proteins and determined their specificities to both insects. The most significant findings arise from the replacement of three segments of the cryIA(c) gene with homologous portions of the cryIA(a) gene: codons 332–428, 429–447, and 448–722. Each of these segments contributed substantially and largely additively toward toxicity to H. virescens. However, replacement of the 429–447 segment of cryIA(c) gene with the cryIA(a) sequence resulted in a 27–50-fold reduction in toxicity toward M. sexta whereas the reduction in toxicity to H. virescens was only 3–4-fold. Subdivision of the 429–447 segment and replacements involving residues within this segment reduced toxicity to M. sexta by 5–10-fold to more than 2000-fold whereas toxicity to H. virescens was only reduced 3–10-fold. These observations indicate that: 1) different but overlapping regions of the cryIA(c) gene determine specificity to each of the two test insects; 2) some of the examined gene segments interact in determining specificity; and 3) different sequences in the cryIA(a) and cryIA(c) genes are required for maximal toxicity in M. sexta.

Bacillus thuringiensis produces proteins that cause a lethal intoxication of specific types of insects. Each B. thuringiensis strain usually carries several toxin genes that determine the range of insects affected. Of the toxin sequences published to date, all but one show some significant sequence similarity to other toxin genes, indicating that these proteins are encoded by a gene family (1). Although most of the isolated B. thuringiensis strains affect only lepidoptera, some strains are specific only to diptera or to coleoptera, and a few are lethal to both lepidoptera and diptera. Many of the B. thuringiensis strains that affect lepidoptera have quite different efficacies among the CryIA(a), CryIA(b), and CryIA(c) proteins. Specificity among the CryIA(a), CryIA(b), and CryIA(c) proteins is required for the toxicity of CryIA(a) to Bombyx mori (silkworm). Our current studies concern the analysis of the lepidopteran specificities of CryIA(a) and CryIA(c) toward Manduca sexta (tobacco hornworm) and Heliothis virescens (tobacco budworm). Höfte et al. (3) have shown that these two proteins have roughly equivalent activity on M. sexta whereas CryIA(c) is about 50 times more toxic to H. virescens than is CryIA(a). One approach to determining the different activities of related gene products is the construction of hybrid, or chimeric, genes followed by assays of the chimeric gene products. This type of analysis relies upon the observation that related genes have the same overall three-dimensional structure (e.g. Ref. 15) so that linear combinations of related gene products often lead to relatively functional, stable proteins. This paper reports the toxicity to M. sexta and H. virescens of proteins synthesized by Escherichia coli after transformation with plasmids bearing the cryIA(a), cryIA(b), and cryIA(c) genes or chimeric recombinants between them. We have identified a segment within this variable region which is required for the toxicity of CryIA(a) to Bombyx mori (silkworm). Our current studies concern the analysis of the lepidopteran specificities of CryIA(a) and CryIA(c) toward Manduca sexta (tobacco hornworm) and Heliothis virescens (tobacco budworm). Höfte et al. (3) have shown that these two proteins have roughly equivalent activity on M. sexta whereas CryIA(c) is about 50 times more toxic to H. virescens than is CryIA(a).

* This research was supported in part by Public Health Service Grant GM-20784 from the National Institute of General Medical Sciences and by grants from the Washington Technology Center.

† Recipient of Research Career Award K6-GM-442 from the National Institute of General Medical Sciences.

Received for publication, June 15, 1990.
between codons 429 and 447. Overall, the data show that different sequences in a given \textit{B. thuringiensis} insecticidal protein determine toxicity to different insects and also that toxicity to a given insect may involve different regions in closely related proteins.

**MATERIALS AND METHODS**

**Bacterial Strains and Plasmids—**E. coli strains MC1000 (16), RZ1032 (17), and \textit{B. thuringiensis} HD-1-Dipel (18) have been described previously. Plasmids pUC118 and pUC119 (19) were used to manipulate portions of the toxin genes. A derivative of pUC119, pUC119A, was created by digesting with EcoRI and Smal, filling in of the EcoRI site, and re-ligation. M13K07 (19) was used to produce sequencing templates from pUC118, pUC119, and derivatives. Plasmids pMAR4, pWK9, and pWK9' (7) as well as pPS41 (6) have been described.

**Insect Sources—**The tobacco hornworm \textit{M. sexta} was obtained from Drs. L. Riddiford and J. Truman, Dept. of Zoology, University of Washington. The tobacco budworm \textit{H. virescens} was obtained from the USDA-ARS, Stoneville, MS.

**Plasmid Constructions—**Commonly used procedures in molecular biology were performed essentially as described (20). The plasmid pCAa contains the Ndel-Ndel fragment having the cryIA(a) gene from pHH241 in pUC119A (see Fig. 1, A and B). Plasmid pCAb, containing the cryIA(b) gene from pWK29, has the filled-in Ndel fragment containing the toxin gene cloned into the filled-in XbaI site of pUC119A (Fig. 1, A and B). Plasmid pCAc, containing the cryIAc gene from pWK29, has the filled-in XbaI fragment containing the toxin gene cloned into the filled-in XbaI site of pUC119A (Fig. 1, A and B). In all three cases the crystal protein genes are oriented in the same direction as the lac promoter. Plasmids pCS3 and pCS4 contain the indicated EcoRI-SacI fragment of the cryIA(a) and cryIAc (c) genes, respectively, in the same sites of pUC119 (Fig. 1C); pCS1 and pCS2 were formed by excision of the respective ClaI-SacI subfragments from pCS3 and pCS4. pCS1, pCS2, pCS3, and pCS4 were formed by introducing the EcoRI-SacI fragments from pCS1, pCS2, pCS3, respectively, into pCAc that had been partially methylated with EcoRI methylase, digested with EcoRI, digested with SacI, and the appropriate fragment purified. Construction of plasmids pCS5, pCS6, pCS7, and pCS4 proceeded by replacement of the SacI-KpnI fragments of pCC1, pCC2, pCC3, and pAC3, respectively, with the SacI-KpnI fragment of cryIA(a) (Fig. 1D).

Oligonucleotide-directed mutagenesis (17) was used to introduce Poul II sites (see Fig. 6A) into the C4la-SacI fragment of cryIA(c) contained in pCS4 (Fig. 1C). The sites were introduced at codons 439-441 of cryIA(c) corresponding to the same codons in cryIA(a) (pCS5 and pCS6) or at codons 442-445 of cryIA(c) corresponding to gap introduced into the cryIA(c) sequence to optimize the alignment between the two genes (pCS6; see Fig. 6A). The changes were verified by sequencing the entire EcoRI-SacI fragments. The C4la-Poul II fragment of pCS6 (from cryIA(a)), pCS5, and pCS6 were purified and ligated into pCS4 in the following combinations: pCS6 C4la-Poul II plus pCS3 Poul II-SacI to form pCS7; pCS5 C4la-Poul II plus pCS3 Poul II-SacI to form pCS8; pCS3 C4la-Poul II plus pCS6 Poul II-SacI to form pCS9; and pCS3 C4la-Poul II plus pCS6 Poul II-SacI to form pCS10. The EcoRI-SacI fragments from pCS5, pCS6, pCS7, pCS8, and pCS9, and pCS10 were then inserted into pCA3 to form pCC1, pCC10, pCC12, pCC14, pCC13, and pCC15, respectively, by the methods shown in Fig. 1, B-D, to make pCC1, pCC2, and pCC3.

**Purification and Concentration Estimation—**Parasporal crystals of \textit{B. thuringiensis} HD-1-Dipel were purified as described previously (21). Inclusions containing crystal protein were purified from \textit{E. coli} by repeated sonication and washing in 0.01 M Tris, 0.01 M EDTA, pH 7.3, 1 mg/ml lysozyme. A 0.5 mg NaCl wash removed the rest of the adsorbed lysozyme. Samples were stored in H2O at 4 °C, and an additional wash with H2O was necessary.

Soluble protoxin was prepared from inclusions or crystals by extraction with 0.1 M Na2CO3, 0.025 M b-mercaptoethanol, pH 9.5. The resulting supernatant was generally 90% or more of approximately 133-kDa protoxin, with most of the residue consisting of antigenically related fragments of greater than 70 kDa. Protein concentrations were determined by using the method of Bradford (22) and by scanning Coomasie Blue-stained sodium dodecyl sulfate-polyacrylamide gels.

**Bioassay and Data Reduction—**Bioassays were performed in 24-well culture plates with 2-cm2 wells, using the diet of Bell et al. (23). Dilutions of purified protoxin were spread on the surface of the diet and allowed to dry before placing one larva (less than 1 day old) in each well. Eight to twelve larvae were tested per dilution, and after 7 days, immobile, unresponsive larvae (regardless of size) were scored as dead.

Mean lethal concentrations were estimated for each test preparation using the probit model (24), which estimates a parameter equal to the log LC50 and confidence limits that are linearly spaced with the log of the dose. Each group of bioassays included a control dose-response assay using purified solubilized crystals from \textit{B. thuringiensis} strain HD-1-Dipel. All constructions in the set were assayed at the same time, using the same batch of diet and were normalized to the control by subtraction of the control log LC50 from the sample log LC50 according to Finney (24). Each recombinant protoxin was tested three to eight times. Our results showed that the confidence limits determined by probit analysis (24) for the LC50 values of individual experiments clearly underestimated the variation observed between repeated experiments; therefore, the LC50 values reported are the averages of the log LC50 values for several determinations relative to the HD-1-Dipel protoxin standard. On average, the one standard deviation confidence limits determined in this way equalled approximately the two standard deviation confidence limits determined as recommended by Finney (24) and represent more conservative estimates of the accuracy of our results. Since it is the log of the dose, rather than the dose, which is used to estimate the LC50 and the associated confidence limits, the one standard deviation confidence interval for the LC50 (ng/cm2) is presented as 10LC50 ± 0.3 log units. In the case of CryIA(b), which was only assayed once, the one standard deviation confidence limits were derived directly from the LC50 calculation (24).

When the differences between the mean values were tested for statistical significance, several instances were observed in which LC50 values differing by less than 2-fold were significant at the p < 0.05 level. Since we do not believe that observed differences of this magnitude can be considered reliable without numerous replications, testing the significance of differences of observed significance for LC50 values differing by more than 2-fold and having significance at p < 0.05 by t test are referred to as “significant.”

**RESULTS**

**Constructions, Protein Purification, and Assays—**To investigate the contribution of individual portions of the CryIA(a) and cryIA(c) genes to insecticidal specificity, each was subcloned so that it could be expressed efficiently in \textit{E. coli} (Fig. LA). The cryIA(b) gene was included in these studies for specific comparisons as described below. Chimeric genes were then constructed as described in Fig. 1 and under “Materials and Methods.” The positions of some of the amino acid differences in CryIA(a), CryIA(b), and the chimeric proteins relative to CryIA(c) as well as the sources of DNA for the chimeric genes are diagrammed in Fig. 2. The chimeric constructions shown in Fig. 2 test the effects of differences between the cryIA(a) and cryIA(c) genes in the following segments: codons 1-331 plus 723-1156 and in codons 332-498, 498-447, and 448-799. Several constructions, including CC3 and CC8, allowed a test of the effect of sequences in cryIA(b) which differ from cryIA(a) and cryIA(c).

For each chimeric gene, inclusions synthesized in \textit{E. coli} were purified, and the protoxin was solubilized and tested for toxicity to \textit{M. sexta} and \textit{H. virescens}. No significant differences were found in the amounts of inclusions produced from the chimeric genes or in the solubility of the inclusions. In all assays, the toxicity results were normalized to the results obtained with similarly treated crystals from strain HD-1-Dipel. The latter strain is a derivative of strain HD-1, which is used frequently as a standard in assays of \textit{B. thuringiensis} insecticidal activity. Crystals from strain HD-1 contain the CryIA(a), (b), and (c) proteins; crystals from strain HD-1-Dipel contain only the CryIA(a) and CryIA(c) proteins (7). Strains HD-1-Dipel and 11D-1 also express the cryIA(a) gene; however, its product has a much higher LC50 of about 150-
250 µg/cm² on the two test insects¹ (26) and should not contribute appreciably to toxicity. The results are presented as LC₅₀ values in ng/cm² and are plotted on a log scale with the associated one standard deviation confidence interval, since toxicity results are linear with the log of the dose.

**Toxicity of CryIA(a), (b), and (c) to H. virescens and M. sexta**—Table I shows that all three protoxins have similar toxicity to M. sexta whereas the toxicities to H. virescens are quite different, with CryIA(a) being much less toxic than CryIA(b), which is less toxic than CryIA(c). Because of variations in the susceptibilities of different lots of insects, these average values are 2–4-fold higher than values reported previously (3) for the individual proteins tested with M. sexta and 4–8-fold higher than results observed previously (3) for H. virescens. However, the relative potencies of the three proteins to the two insects agree with the prior values. Assays performed with different batches of insects (e.g. Table II in a later section) yielded values that agreed with published values (3).

**Effect of Differences in the cryIA(b) Gene Versus the cryIA(a) and cryIA(c) Genes**—Comparison of the differences in the amino acid sequence diagrammed in Fig. 2 indicates that the cryIA(b) gene can be viewed essentially as a recombinant between the cryIA(a) and cryIA(c) genes at approximately codon 406 with a deletion of 26 codons at codon 793 and a divergent sequence beginning at codon 1065 which includes a 4-codon insertion (25). One of the chimeras between CryIA(a) and CryIA(c), CC3 (Fig. 2), has a sequence very similar to CryIA(b) prior to codon 722. A comparison of the toxicity of CC3 and CryIA(b) on both H. virescens and M. sexta shows no significant differences, demonstrating that the sequence differences distal to codon 722 have no detectable effect on toxicity toward these insects.

The sequences of CryIA(a) and CryIA(b) between codons 448 and 722 differ by 3–4 residues, depending on the source of the cryIA(b) gene. Chimeric protein CC8 is CryIA(a) with the CryIA(b) sequence for residues 448–722. A comparison of toxicities of CryIA(a) and CC8 indicates little difference toward M. sexta and a difference in toxicity to H. virescens.

**Table I**

<table>
<thead>
<tr>
<th>Protein*</th>
<th>H. virescens</th>
<th>M. sexta</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LC₅₀*</td>
<td>1 S.D. C.I.</td>
</tr>
<tr>
<td><strong>ng/cm²</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B.t. HD-1-Dipel</td>
<td>18</td>
<td>11–28</td>
</tr>
<tr>
<td>CryIA(c)</td>
<td>16</td>
<td>12–21</td>
</tr>
<tr>
<td>CryIA(b)</td>
<td>43</td>
<td>27–68</td>
</tr>
<tr>
<td>CryIA(a)</td>
<td>472</td>
<td>226–988</td>
</tr>
<tr>
<td>CC8</td>
<td>240</td>
<td>118–488</td>
</tr>
<tr>
<td>CC9</td>
<td>131</td>
<td>87–196</td>
</tr>
</tbody>
</table>

250 µg/cm² on the two test insects¹ (26) and should not contribute appreciably to toxicity. The results are presented as LC₅₀ values in ng/cm² and are plotted on a log scale with the associated one standard deviation confidence interval, since toxicity results are linear with the log of the dose.

**Toxicity of chimeric proteins constructed by exchanging segments of CryIA(a) and CryIA(c) between codons 332 and 772**

<table>
<thead>
<tr>
<th>Protein*</th>
<th>Segment*</th>
<th>H. virescens</th>
<th>M. sexta</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>X Y Z</td>
<td>LC₅₀*</td>
<td>1 S.D. C.I.</td>
</tr>
<tr>
<td><strong>ng/cm²</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC1 +</td>
<td>49</td>
<td>43–55</td>
<td>56</td>
</tr>
<tr>
<td>CC2 +</td>
<td>59</td>
<td>42–82</td>
<td>537</td>
</tr>
<tr>
<td>CC3 +</td>
<td>34</td>
<td>16–72</td>
<td>23</td>
</tr>
<tr>
<td>CC4 +</td>
<td>147</td>
<td>115–187</td>
<td>23</td>
</tr>
<tr>
<td>CC5 +</td>
<td>842</td>
<td>704–1007</td>
<td>36</td>
</tr>
<tr>
<td>CC6 +</td>
<td>106</td>
<td>56–202</td>
<td>83</td>
</tr>
<tr>
<td>CC7 +</td>
<td>605</td>
<td>573–638</td>
<td>19</td>
</tr>
</tbody>
</table>

*Source of DNA: cryIA(a) = cryIA(b) = cryIA(c) = cc1 = cc2 = cc3 = cc4 = cc5 = cc6 = cc7 = cc8 = cc9 = ccm' = LGO 1 S.D. C.I.

¹ Solubilized crystals from B. thuringiensis (B.t.) subspecies kurstaki HD-1-Dipel or solubilized inclusions from cloned genes expressed in E. coli. Compositions of chimeric proteins CC1–CC9 are shown in Fig. 2.

*ng/cm² of diet.

¹ One standard deviation confidence interval; see “Materials and Methods” for details.

* CryIA(a) segment exchanged into CryIA(c): X = residues 332–426; Y = residues 429–447; Z = residues 448–772.
which is not significant (1.9-fold, p ≤ 0.12; Table I). Three additional constructions (not shown in Fig. 2) were made which were identical to CC5, CC6, and CC7, respectively, except that the codon 448-722 segment was from the cryIA(b) gene rather than cryIA(a). In each case, the potency of the CryIA(b)-containing protein was very similar to that of the CryIA(a)-containing protein for both insects (data not shown). Therefore, the few differences between CryIA(a) and CryIA(b) between residues 448 and 722 do not account for their different toxicities toward H. virescens.

Comparison of CryIA(a) and CryIA(c) Prior to Codon 331 and Distal to Codon 722—There are 10 differences in sequence between CryIA(a) and CryIA(c) prior to residue 331 and 2 differences distal to codon 722 (Fig. 2). Two pairs of proteins, CC4 versus CC9 and CC7 versus CryIA(a), allow comparison of the effects on toxicity of these regions. There were no significant differences in toxicity to either insect for either pair of proteins (Table I), indicating that in the contexts used, these two segments do not contribute to insecticidal specificity.

Combinatorial Analysis of CryIA(a) Substitutions into CryIA(c) from Codons 332-722—The majority of the differences between CryIA(a) and CryIA(c) are located between codons 332 and 722. As stated above, chimeric protein CC1, in which amino acids 332-722 of CryIA(c) are replaced by those in CryIA(a), had very similar toxicity to CryIA(a) for both H. virescens and M. sexta. To analyze which portions of this region determine the greater toxicity to H. virescens, the region was subdivided into three parts: codons 332-428 (segment X), 429-447 (segment Y), and 448-722 (segment Z). Each of the six combinations (CC1-CC6, Fig. 2) of these segments from CryIA(a) was used to replace the corresponding segment of CryIA(c). Figs. 3 and 4 show a graphic analysis of the effect of the substitutions on the LC50 to each insect. Each of the three panels in Figs. 3 and 4 shows the LC50 values of two of the six different pathways by which CryIA(c) can be converted to CC7 by substitution of one of the three segments followed by a second and third segment.

With H. virescens, substitutions of one segment of CryIA(a) into CryIA(c) resulted in a drop of 2.1-3.7-fold in toxicity, regardless of the segment (UC12, UC2, UC3 in Fig. 3, A-C). However, CC3 (substituting segment C; p = 0.06) did not meet our criteria for a significant difference in toxicity from CryIA(a) (Fig. 3C). Double substitutions involving segments X and Y (CC4) or Y and Z (CC6) resulted in a further 1.8-4.4-fold reduction in toxicity (Fig. 3, A-C), with CC6 not differing significantly from CC2 (p = 0.14, Fig. 3B). The substitution involving segments X and Z (CC5) was the least toxic protein to H. virescens, being marginally less toxic than CC7 (Fig. 3, A and C; Table I). With the exception of combination pathways that initiated with segments X and Z (CC5, Fig. 3, A and C), there was a general linear trend to the data indicating that each of the segments tested contributed to specificity toward H. virescens.

Cooperativity between the CryIA(c) 332-428 (X) and 429-447 (Y) Segments—When the same set of chimeric proteins was tested with M. sexta, quite different results were obtained. Whereas segment Z caused an insignificant change in toxicity (Fig. 4C, CC3), segment X caused a 3-fold drop in toxicity (Fig. 4A, CC1), and segment Y caused a 27-fold drop in toxicity (Fig. 4B, CC2) when compared with CryIA(c). For double substitutions, segments X and Y restored toxicity fully relative to the single substitutions (Fig. 4, A and B, CC4), segments X and Z caused a statistically significant (p = 0.002) but marginal (2-fold) reduction from parent protein toxicity (Fig. 4, A and C, CC5), and segments Y and Z retained a 4.3-fold reduction in toxicity (Fig. 4, B and C, CC6) compared with the parent proteins.

Thus, with the exception of CC5, which was 2-fold less toxic than the parent proteins, separation of the CryIA(c) 332-428 and 429-447 segments (segments X and Y) led to a dramatic loss of toxicity on M. sexta (Fig. 4, A-C). Notably, CC2, containing only codons 429-447 from CryIA(a), was 27-fold less toxic than the parent proteins, suggesting a key role for these relatively few residues of the cryIA(c) gene product in the intoxication of M. sexta. Addition of codons 332-428 (segment X) from CryIA(a) to form CC4 resulted in complete restoration of parental toxicity whereas the alternative addition of codons 448-722 to form CC6 (Fig. 4B) yielded a protein that was 3.6-4.3-fold less toxic than the parent proteins. The addition of the CryIA(a) 448-722 segment (segment X) had little effect by itself either before (CC3) or after (CC7) the addition of the whole 332-447 CryIA(a) segment (Fig. 4, A-C). We conclude therefore that the decreased toxicity created by replacement of the CryIA(c) 429-447 segment with that from CryIA(a) was suppressed most effectively by the addition of the 332-428 segment from the CryIA(a).

When tested on M. sexta, the CryIA(a) segments X and Z together had little dependence on the CryIA(a) Y segment since the latter could be replaced with the corresponding CryIA(c) region (to form CC5) with little effect on toxicity. The strong dependence of the CryIA(c) but not CryIA(a) variable region on its own segment Y suggests that different portions of these two toxins are critical for their interaction.
Crystal Protein Specificity-determining Regions 20927

A. +I0

B. %O

C.

FIG. 4. Effect of substitution order on toxicity to M. sexta.
The figure shows the LC50 values of chimeric proteins as a function of the number of CryIA(a) segments introduced into CryIA(c). The first point of each panel is a different segment substitution. Vertical bars show one standard deviation confidence intervals.

FIG. 5. A plot of the ratios of LC50 values of control and chimeric proteins CC1–CC7 to two insects as a function of combined LC50 values. Filled circles denote the parent protein values. Horizontal lines indicate one standard deviation confidence intervals. ms, M. sexta; Hu, H. virescens.

with the M. sexta midgut. Interestingly, CC5 was marginally the least toxic of the chimeric toxins to H. virescens (Fig. 3A), suggesting that the combined action of the CryIA(a) X and Z segments is somewhat impeded by the Y segment of CryIA(c).

Discrimination Versus Activity of the Toxins—An alternative way of viewing the activity of each of the insecticidal proteins to two different insects is shown in Fig. 5. This figure presents a plot of the ability of each toxin to discriminate between the two insects as indicated by the ratio of the LC50 values versus combined LC50, i.e., 10^(-average log LC50) of each protein toward both insects. The differential efficacy of B. thuringiensis strains toward insects has been characterized previously by the ratio of LC50 values to two insects (e.g., Ref. 27). Although this is an adequate measure of the ability of B. thuringiensis insecticidal proteins to discriminate between two insect targets, it does not provide the absolute level of toxicity toward the two insects (i.e., two toxins could have the same ratio of toxicities but be more or less toxic to both insects). The comparison method introduced in Fig. 5 uses the ratio of LC50 as the horizontal axis, preserving this information, and enhances the comparison by including a measure of overall toxicity, the combined LC50. This facilitates comparison of the toxins and shows clearly, for example, that CryIA(c) and CC1 have similar discrimination although CC1 is clearly less toxic. Similarly, CC2 and CC5 have comparable combined LC50 values but a 950-fold difference in discrimination between the two insects.

In terms of the current analysis, the points near the line connecting the value for CryIA(c) with that for CryIA(a) in Fig. 5 indicate those proteins (i.e., CC3, CC4, CC7, CC8, CC9, CryIA(b), and perhaps CC5) in which the increase in combined LC50 and decrease in discrimination are attributable to the loss of toxicity to H. virescens. For CC1 and CC6, the combined LC50 value is increased, but the discrimination value is little changed from CryIA(c), indicating that the changes in these proteins have caused primarily a loss of toxicity to both insects and not an altered specificity. The unique location of CC2 in Fig. 5 demonstrates the large shift in discrimination caused by the much larger loss of toxicity to M. sexta than H. virescens, a result unexpected from the activities of either parent protein.

Fine Structure Analysis of the 429–447 Region (Segment Y)—A further subdivision of this segment of the cryIA(c) gene was made by using site-directed mutagensis to create PvuII restriction sites at two positions of potential homology to the PvuII site in cryIA(a) (Fig. 6A). These sites, introducing substitution mutations, were used to form a set of cryIA(a)/cryIA(c) recombinants between codons 490 and 497 which either incorporate or omit a two-amino acid insertion relative to the CryIA(a)-derived sequence of CC2, as shown in Fig. 6B. The purified, soluble protoxins were bioassayed three to four times on M. sexta and H. virescens in parallel with the control crystal proteins from B. thuringiensis HD 1 Dipel, CryIA(c), and CC2, as shown in Table II. In this series of bioassays, the M. sexta larvae were generally more susceptible to the toxins than those used for the bioassays in Table I, and the LC50 of CC2 was 50-fold higher than that of CryIA(c).

The first construct, CC10, with the PvuII site at codons

![Graph](https://example.com/graph.png)

FIG. 6. Construction of chimeric proteins CC10–CC15. A, nucleotide and deduced amino acid sequences of cryIA(a), cryIA(c), and base changes in two constructs introduced by site-directed mutagensis. PvuII sites are overlined. B, amino acid sequences of CryIA(c) and chimeric proteins CC2 and CC10–CC15. Highlighted sequences are from CryIA(c).
442 and 443, caused two amino acid substitutions (Asn → Ala at residue 442 and Ser → Gly at residue 443; Fig. 6) and resulted in an approximately 4-fold reduction in toxicity to both M. sexta and H. virescens when compared with CryIA(c) (Table II). The second construct, CC11, with the PovII site at codons 439–441, yielded three amino acid substitutions (Gly → Ala at residue 439, Phe → Ala at residue 440, and Ser → Gly at 441; Fig. 6), and this protein had an approximately 110-fold lower toxicity to M. sexta although the toxicity to H. virescens showed a 10-fold decrease from that of the parent protein.

Chimeric proteins CC12 and CC14 were constructed by combining the cryIA(c) gene from codon 429 to each of the two PovII sites with the remainder of the Y segment coming from cryZA(a). Reversing the sources of the genes used to prepare the CC12 and CC14 constructs yielded chimeric proteins CC13 and CC15 (Fig. 6B). As shown in Table II, proteins CC12 and CC14 were 3.5–5-fold less toxic to H. virescens than CryIA(c) whereas the toxicity to M. sexta dropped approximately 120-fold (the near identity of the toxicities of these two proteins seems remarkable since they differ by a two-amino acid deletion and one substitution (Fig. 6B)). An even greater loss in toxicity to M. sexta was observed with chimeric protein CC15 (about 260-fold), and a >1600-fold loss of toxicity to this insect was found with CC13. Large amounts of CC13 inclusions were capable of killing M. sexta, so this protein cannot be regarded as completely nontoxic. The toxicity of chimeric proteins CC15 and CC13 to H. virescens was reduced only about 4-fold and about 10-fold, respectively. The improved toxicity to M. sexta of CC15 compared with the parent proteins and several chimeric constructs indicated that differences in the amino acid sequences prior to codon 331 are shown in Fig. 6B).

Fig. 7 shows a plot similar to that of Fig. 5 displaying the combined LC50 values as a function of the ratio of the LC50 values for the two insects using the constructions in Table II. As noted in Fig. 5, CC2 has a higher combined LC50 than CryI(a) and a dramatically altered discrimination between M. sexta and H. virescens. The relationship between these two proteins remains similar in Fig. 7. The two-amino acid substitution in CryI(a), CC10, shows an increased combined LC50 with little change in discrimination compared with CryI(a). By contrast, CC11, differing by only three amino acid substitutions from CryI(a), has the same discrimination for the two insects as CC2 with a higher combined LC50 than CC2. Three cryI(a)/cryI(c) recombinants, CC12, CC14, and CC15, have values that cluster relatively close to CC2 and CC11. Chimeric protein CC13, with an undeterminably high LC50 on M. sexta, has a higher combined LC50 and higher M. sexta to H. virescens LC50 ratio than any of the other proteins tested.

### DISCUSSION

Although significant advances have been made in characterizing the genes coding for *B. thuringiensis* insecticidal proteins (1), relatively little is known about the mechanisms leading to the death of susceptible insects. It has been demonstrated that specific receptors on the epithelial gut cells bind different toxins (4, 5), and it would be expected that the number of such binding sites and the binding affinities would be important in determining host range and potency. The interaction of the toxin with the epithelial gut cells is followed by changes in permeability, as documented by electron microscopy and the release of small molecules (28). Two studies (14, 29) have identified partially overlapping specificity-determining regions, with different *B. thuringiensis* proteins, but the precise functions of these regions have not been determined, and it remains unknown whether the same region in each is required for toxicity to insects other than the one that was tested.

In the present investigation, we have analyzed two related lepidopteran-specific proteins, CryI(a) and CryI(c), to determine the regions controlling the specificity of these proteins toward *M. sexta* and *H. virescens*. CryI(a) is equally toxic to both insects whereas CryI(a) is significantly less toxic to *H. virescens* than to *M. sexta*. To locate the determinants of toxicity to *H. virescens*, chimeric genes were constructed, and the gene products were assayed for insecticidal activity. Pairwise comparisons of data obtained with the parent proteins and several chimeric constructs indicated that differences in the amino acid sequences prior to codon 331 and following codon 722 had no effect on toxicity to either insect. However, since we observed some unpredicted changes in toxicity depending on the precise combination of sequences in experiments combining segments of these genes between codons 332 and 722, we cannot rule out the possibility that exchanging the 1–331 and 722–1156/78 segments in some of the untested combinations would result in a difference.

The main variability in sequence between CryI(a) and CryI(c) lies between residues 332 and 722. Deletion analysis and protein sequencing data suggest that the carboxyl terminus of the toxicity and possibly also the specificity determinants lies between codons 607 and 623 (9, 13). The codon...
The 21-50-fold increase in LC50 on M. sexta observed with the chimeric protein CC2 relative to the CryIA(c) parent prompted a further analysis of the codon 429-447 sequence. Additional manipulation to substitute and subdivide in the codon 429-447 segment resulted in three proteins with significant differences in toxicity. First, CC10, with substitution mutations at residues 442 and 443, displayed a modest loss of toxicity with no change in discrimination between the two insects from the parent CryIA(c). Second, CC11, having three amino acid changes (codons 439-441), showed toxicity and discrimination closely approximating those of CC2, which has 11 substitutions and two insertion/deletions relative to CryIA(c). Third, CC13, having seven substitutions in the 434-441 segment relative to CryIA(c), was sufficiently nontoxic to M. sexta to prevent determination of an LC50 value with soluble protoxin, yet this protein retained moderate toxicity to H. virescens. The broad range of toxicity and specificity demonstrated by the few amino acid changes in this second set of chimeric proteins suggests that a limited number of single amino acid substitutions will allow the identification of residues controlling specificity of CryIA(c) toward M. sexta. Furthermore, preliminary experiments indicate that CC2 is as stable to proteolysis as CryIA(c) in the presence of either M. sexta or H. virescens brush border membrane vesicles so that it should be possible to determine whether the effect on CC2 toxicity is caused by reduction in binding to a receptor or to some later stage in intoxication (the possibility of enhanced susceptibility to proteolysis has not yet been ruled out for the other mutations in the 429-447 segment).

For H. virescens, replacement of the cryIA(c) segments with counterparts from cryIA(a) resulted in a progressive reduction in toxicity depending on the number of segments replaced. An exception was construct CC5, which was marginally less toxic than CryIA(a) but contained only two of the three segments (332-428 and 448-722). However, these two segments represent the great majority of the variable sequence and may exert an overall dominant effect on toxicity.

Thus, the response of the two insects to the same set of chimeric gene products falls into two classes: \textit{uoext} dependent toxicity (\textit{M. sexta}) or progressive conversion of one parental type to another with increasing amounts of sequence substitution (\textit{H. virescens}). In contrast, Ge et al. (14), who studied the same two genes, showed that the determinants of toxicity of CryIA(a) to \textit{B. mori} were located solely in the codon 332-447 region. Our preliminary results with \textit{Trichoplusia ni}\textsuperscript{2} suggest that specificity of CryIA(c) to the latter insect can also be accounted for by the codon 332-447 segment. In all, these results indicate the existence of three different patterns of toxicity determinants for specific lepidopteran larvae involving different segments of the codon 330-620 region of the toxin-encoding portion of the proteins: 1) incremental and involving the entire variable region (as in the toxicity of CryIA(c) to \textit{H. virescens}); 2) context dependent (as in the toxicity of CryIA(c) to \textit{M. sexta}); and 3) a single, independent specificity domain (as in the toxicity of CryIA(a) to \textit{B. mori}).

A modification of the third pattern of toxicity was found in studies of chimeric proteins containing segments of CryIA and CryIB (29). CryIIA ("P" toxin; Ref. 30), a major component of cuboidal crystals, is toxic to mosquito larvae (\textit{Aedes aegypti}) and lepidopteran larvae; the closely related CryIB protein is toxic only to lepidopteran larvae. Analyses of the insecticidal activities of chimeric CryIIA/CryIB proteins showed that residues 307-382 were required for dipteran toxicity. However, the data obtained with CryIIA/CryIB chimeras indicated that distally located segments also influenced toxicity, possibly by affecting the overall protein conformation. Structural incompatibilities may also explain our current results with CryIA(c)/CryIA(a) chimeric proteins. \textit{M. sexta} toxicity of the codon 332-431 segment of CryIA(a) had moderate dependence on the source of the 432-447 segments; however, \textit{M. sexta} toxicity of the codon 332-431 segment of CryIA(c) was critically dependent on the 432-447 segment of CryIA(c), implying close structural interactions between the latter two segments. By contrast, exchanges of the codon 332-447 and 448-722 segments yielded proteins with relatively predictable activities and discriminations, as illustrated in Fig. 5, suggesting that these segments are relatively independent structural elements. This notion is reinforced by the existence of the naturally occurring CryIA(b) containing an apparent crossover point between CryIA(c) and CryIA(a) at codon 460.

Only one of the chimeric insecticidal proteins we have constructed was essentially nontoxic, and this protein, which had seven amino acid substitutions, displayed the tremendous loss in toxicity only to one of the two test insects. In contrast, Ge et al. (14), working with the same genes but assaying only for toxicity to \textit{B. mori}, noted that exchanges at a point in the 448-722 segment resulted in total inactivation of the chimeric proteins. Koblika et al. (31) observed that a subset of chimeric \textit{adrenergic} receptors had lost activity, and they suggested that specific combinations of transmembrane helices made essential interactions that could not be supplied by the related protein in the chimera. Analogously, we take the data of Ge et al. (14) to mean that the 448-722 segment contains within-segment structural interactions that can be perturbed easily by internal crossovers between these related genes. Interestingly, three of the highly conserved regions found in most \textit{B. thuringiensis} insecticidal proteins (Ref. 1 and Fig. 8) are in this segment of the protein. Taken together with our current results regarding the \textit{M. sexta} specificity of CryIA(c), these data support the proposal that the codons 332-447 and 448-607 segments of the crystal protein gene are independently acting protein structures, if not separate domains of the protein; each variant of each of these two segments can be combined with the other to yield a protein with a combined activity that is predictable from its discrimination between the two insects, and internal exchanges in these segments are disruptive, resulting in less than expected toxicity.

Fig. 8 presents a diagram that shows the location of the specificity determining regions in CryIA(c) to \textit{H. virescens}, \textit{M. sexta}, and \textit{T. ni}, in CryIA(a) to \textit{B. mori} and in CryIA to \textit{A. aegypti} as discussed above (see Ref. 1 for alignments of the proteins). Fig. 8 also includes the region of toxicity of the

\footnote{H. E. Schnepf and K. Tomczak, unpublished data.}
CryIA(a) and CryIA(b) proteins to *Mamestra brassicae* as inferred from results reported by Höfte et al. (3); it has not been determined whether toxicity depends on the entire region or one or more segments thereof. Also shown are results obtained by Haider and Ellar (32) for a variant CryIA(b) protein toxic to both lepidoptera and diptera. Their data indicate that these three amino acid differences in the codon 537–566 segment of CryIA(b) of *B. thuringiensis* subspecies *azazea* IC1 specify novel dipteran and lepidoptera protease cleavage sites that allow the production of the toxin for the respective insect targets. They have also shown that deletion of codons 242–523 of this protein (open box in Fig. 8) allows synthesis of a dipteran but not a lepidoptera toxin. It is interesting that although all of the specificity-determining domains or regions fall within the variable region of the cryIA genes (13) the specificity regions differ in size and position. Also, the amino-terminal half of the variable region (corresponding to residues 280–447 of CryIA(a)) is located between the second and third of the five conserved regions revealed by comparison of the sequences of 13 crystal proteins (Ref. 1 and Fig. 8). The latter half of the variable region, however, contains conserved regions 3–5 (Fig. 8). If, like immunoglobulins, the conserved sequences represent structural scaffolding whereas the variable or hypervariable regions confer differential specificity, the conserved regions of the toxins may be involved more with protein conformation or some common determinant of toxin activity rather than with specificity-determining interactions with target cells. This could explain, in part, the loss of toxicity observed with chimeric proteins having exchange points in the carboxy-terminal half of the variable region (14).

The finding that the same set of chimeric proteins elicits different responses by two different insects suggests that the most active of the toxins to these two insects, CryIA(c), interacts differently with the two insects. It follows then that if alterations are made in the specificity-determining region, not all of the species that are susceptible to CryIA(c) may be affected to the same extent. Conversely, if insect resistance developed to an insecticidal protein, such resistant insects would not necessarily be resistant to a related protein. In fact, it has been demonstrated recently that the Indian meal moth (*Plodia interpunctella*), in which resistance was developed to CryIA(b), is more sensitive to the CryIC protein than the nonresistant parental insects (33).
Specificity-determining regions of a lepidopteran-specific insecticidal protein produced by Bacillus thuringiensis.
H E Schnepf, K Tomczak, J P Ortega and H R Whiteley


Access the most updated version of this article at [http://www.jbc.org/content/265/34/20923](http://www.jbc.org/content/265/34/20923)

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at [http://www.jbc.org/content/265/34/20923.full.html#ref-list-1](http://www.jbc.org/content/265/34/20923.full.html#ref-list-1)