

**Universidade de Ribeirão Preto – UNAERP
Doutorado em Biotecnologia**

HAGAR CERIANE COSTA CORSINI MACIEL

**ANÁLISE COMPARATIVA DAS SEQUÊNCIAS CODIFICADORAS
DOS GENES CRY DE *Bacillus thuringiensis***

**Ribeirão Preto – SP
2015**

HAGAR CERIANE COSTA CORSINI MACIEL

**ANÁLISE COMPARATIVA DAS SEQUÊNCIAS CODIFICADORAS
DOS GENES CRY DE *Bacillus thuringiensis***

Tese apresentada ao Programa de Pós-graduação
Stricto Sensu da Universidade de Ribeirão Preto
– UNAERP como requisito parcial à obtenção do
título de Doutor em Biotecnologia.

Orientadora: Prof.^a Dr.^a Sonia Marli Zingaretti

**Ribeirão Preto – SP
2015**

Ficha catalográfica preparada pelo Centro de Processamento Técnico
da Biblioteca Central da UNAERP

- Universidade de Ribeirão Preto -

M152a Maciel, Hagar Ceriane Costa Corsini, 1970-
Análise comparativa das sequências codificadoras dos genes
cry de *Bacillus thuringiensis* / Hagar Ceriane Costa Corsini Maciel.
- - Ribeirão Preto, 2016.
86 f.: il. color.

Orientadora: Prof^a. Dr^a. Sonia Marli Zingaretti.

Tese (doutorado) - Universidade de Ribeirão Preto, UNAERP,
Biotecnologia. Ribeirão Preto, 2016.

1. Proteínas CRY. 2. *Bacillus thuringiensis*. 3. Receptores.
I. Título.

CDD 660

Dedico este trabalho em memória ao meu pai e à minha irmã que sempre me apoiaram e acreditaram em mim, que não puderam acompanhar a conclusão do meu doutorado.

AGRADECIMENTOS

Primeiramente agradeço a Deus por ter me dado forças mesmo nos momentos mais difíceis e por ter colocado pessoas especiais ao meu lado que me ajudaram a superar cada obstáculo.

Ao meu marido por estar ao meu lado nos melhores e piores momentos de minha vida.

Aos meus familiares que souberam compreender cada ausência e por confiarem em minha capacidade.

À minha orientadora, professora Dra. Sonia, que soube entender cada percalço e cada dificuldade encontrada no curso me auxiliando e me incentivando a fazer o melhor de mim.

Ao professor Dr. Milton Faria Júnior que esteve sempre disponível para tirar minhas dúvidas.

À Universidade de Ribeirão Preto - UNAERP pela oportunidade de realização do curso e,

Ao Instituto Federal do Sul de Minas, Campus Machado, pela liberação, concessão de bolsa de auxílio, e flexibilização de horário facilitando a realização das atividades no curso enquanto exercia atividade docente na Instituição.

SUMÁRIO

I. INTRODUÇÃO	13
II. REVISÃO BIBLIOGRÁFICA	16
2.1. Proteínas	16
2.2. Aminoácidos	16
2.3. <i>Bt thuringiensis</i>	17
2.4. Características e modo de ação das toxinas CRY.....	18
2.5. União dos receptores	20
2.6. Plantas geneticamente modificadas com genes de <i>Bt</i>	23
2.7. Bioinformática	26
2.7.1 Alinhamento das Sequências	26
2.7.2 Árvores Filogenéticas	26
III. JUSTIFICATIVA	28
IV. HIPÓTESE.....	28
V. OBJETIVOS	28
5.1. Objetivos gerais	28
5.2. Objetivos específicos	28
VI. METODOLOGIA.....	29
6.1. <i>Download</i> e alinhamento de sequências de aminoácidos	29
6.2. Análise filogenética	30
6.3. Análise dos alinhamentos das sequências de aminoácidos	30
6.4. Estrutura secundária das sequências de proteínas	31
VII. RESULTADOS E DISCUSSÃO	32
7.1. Alinhamento e análise das sequências de aminoácidos.....	32
7.2. Construção de árvores filogenéticas	48
7.3. Análise da estrutura secundárias das sequências de proteínas	52
VIII. CONCLUSÕES	76
IX. REFERÊNCIAS	77
X. ANEXOS	84
10.1. Alinhamento do Domínio I das 39 sequências de aminoácidos	84
10.2. Alinhamento do Domínio II das 39 sequências de aminoácidos.....	85
10.3. Alinhamento do Domínio III das 39 sequências de aminoácidos	86

LISTA DE FIGURAS

Figura 1: Representação esquemática da estrutura da toxina CRY3A	19
Figura 2: Representação esquemática do modo de ação da toxina Cry de <i>B. thuringiensis</i>	20
Figura 3: Receptores moleculares das proteínas CRY1A	23
Figura 4: Apresentação geral deste trabalho e suas etapas	29
Figura 5: Posições conservadas dos aminoácidos quanto ao grupo a que pertencem.....	35
Figura 6: Análise comparativa do número de Aminoácidos Idênticos e Aminoácidos semiconservados nos Domínio I, II e III, das 16 seqüências de genes cuja toxina ataca insetos da ordem Lepidóptera/Coleóptera (8) e Coleóptera (8).....	39
Figura 7: Alinhamento aminoácidos idênticos e semi conservados do grupo de genes cuja proteína CRY é toxica para insetos das ordens Lepidóptera e Coleóptera com o grupo das Coleópteras	40
Figura 8: Análise comparativa do número de Aminoácidos Idênticos e Aminoácidos semiconservados nos Domínio I, II e III, das 31 seqüências de genes cuja toxina ataca insetos da ordem Lepidóptera/Coleóptera (8) e Lepidóptera (23).....	42
Figura 9: Alinhamento aminoácidos idênticos e semi conservados do grupo de genes cuja proteína CRY é toxica para insetos das ordens Lepidóptera e Coleóptera com o grupo das Lepidópteras	44
Figura 10: Comparação dos aminoácidos nos 3 domínios para as seqüências (LC e C, LC e L) quanto à similaridade e variabilidade	45
Figura 11: Filogenia obtida pelo método de <i>Neighbor Joining</i> , relativa às 39 seqüências referentes ao domínio I: 8 seqüências das ordens Lepidóptera/Coleóptera, 23 seqüências da ordem Lepidóptera e 8 seqüências da ordem Coleóptera. O valor de bootstrap é 1000. Lc - Lepidóptera/Coleóptera L - Lepidóptera C - Coleóptera.	49
Figura 12: Filogenia obtida pelo método de <i>Neighbor Joining</i> , relativa às 39 seqüências referentes ao domínio II: 8 seqüências das ordens Lepidóptera/Coleóptera, 23 seqüências da ordem Lepidóptera e 8 seqüências da ordem Coleóptera. O valor de bootstrap é 1000. Lc - Lepidóptera/Coleóptera L - Lepidóptera C - Coleóptera.	50
Figura 13: Filogenia obtida pelo método de <i>Neighbor Joining</i> , relativa às 39 seqüências referentes ao domínio III: 8 seqüências das ordens Lepidóptera/Coleóptera, 23 seqüências da ordem Lepidóptera e 8 seqüências da ordem Coleóptera. O valor de bootstrap é 1000. Lc - Lepidóptera/Coleóptera L - Lepidóptera C - Coleóptera.	51
Figura 14: Predição da estrutura secundária da seqüência CRY1Ia1 Representação gráfica da estrutura secundária predita da seqüência CRY1Ia1 utilizando o programa Psipred que considera os níveis de confiança de predição (Conf).	53

Figura 15: Predição da estrutura secundária da sequência CRY1Ib3 Representação gráfica da estrutura secundária predita da sequência CRY1Ib3 utilizando o programa Psipred que considera os níveis de confiança de predição (Conf)	54
Figura 16: Predição da estrutura secundária da sequência CRY1Fa1 Representação gráfica da estrutura secundária predita da sequência CRY1Fa1 utilizando o programa Psipred que considera os níveis de confiança de predição (Conf).	55
Figura 17: Predição da estrutura secundária da sequência CRY9Ba1 Representação gráfica da estrutura secundária predita da sequência CRY9Ba1 utilizando o programa Psipred que considera os níveis de confiança de predição (Conf).	57
Figura 18: Predição da estrutura secundária da sequência CRY7Aa1 Representação gráfica da estrutura secundária predita da sequência CRY7Aa1 utilizando o programa Psipred que considera os níveis de confiança de predição (Conf).	59
Figura 19: Predição da estrutura secundária da sequência CRY3Aa1 Representação gráfica da estrutura secundária predita da sequência CRY3Aa1 utilizando o programa Psipred que considera os níveis de confiança de predição (Conf)..	61
Figura 20: Identificação da estrutura secundária por domínio A - Domínio I.	66
B- Domínio II.	68
C- Domínio III.	72
Figura 21: Análise da Estrutura Secundária da região de β 10 e β 11 das 6 sequências	75

LISTA DE TABELAS

Tabela 1: Plantas transformadas com genes <i>CRY</i> de <i>Bacillus thuringiensis</i> em processo de Comercialização	25
Tabela 2: 39 sequências de aminoácidos para alinhamento	32
Tabela 3: Resumo das 39 sequências analisadas quanto à similaridade e variabilidade dos aminoácidos	33
Tabela 4: Análise comparativa do número de Aminoácidos Idênticos e Aminoácidos semi conservados nos Domínio I, II e III, das 39 sequências de genes avaliadas	34
Tabela 5: Análise comparativa do número de Aminoácidos Idênticos e Aminoácidos semi conservados nos Domínio I, II e III, das 16 sequências de genes cuja toxina ataca insetos da ordem Lepidóptera/Coleóptera (8) e Coleóptera (8)	38
Tabela 6: Análise comparativa do número de Aminoácidos Idênticos e Aminoácidos semi conservados nos Domínio I, II e III, das 16 sequências de genes cuja toína ataca insetos da ordem Lepidóptera e Coleóptera e Lepidóptera	41
Tabela 7: Resumo das análises dos 3 domínios para as sequências (LC e C, LC e L) quanto à similaridade e variabilidade	44
Tabela 8: Múltiplo alinhamento da sequência de aminoácidos de 3 genes <i>CRY</i>	47
Tabela 9: Estruturas secundárias alfa hélices (α) e folhas beta (β) determinadas para os genes, <i>CRY1Ia1</i> , <i>CRY1Fa1</i> e <i>CRY7Aa1</i>	63

LISTA DE SIGLAS OU ABREVIATURAS

(α): Alfa hélices

(β): Folhas Beta

μm : Micrômetro (unidade)

A: Alanina

ALP F: Fosfatase alcalina

APN: Aminopeptidase-N

BBMVs: Brush border membrane vesicles

Bt: *Bacillus Thuringiensis*

C: Cisteína

C: Coleóptera

CADR: Caderina

D: Aspartato

DNA: Deoxyribonucleic Acid (Ácido Desoxirribonucléico)

DOM I: Domínio I

DOM I: Domínio III

DOM II: Domínio II

E: Glutamato

F: Fenilalanina

G: Glicina

GCR: Glicoconjugado

GPI: Glicosilfosfatidil-inositol

H: Histidina

I: Isoleucina

K: Lisina

L: Lepidóptera

L: Leucina

LC: Lepidóptera/Coleóptera

M: Metionina

mRNA: RNA mensageiro

N: Asparagina

OGMs: Organismos Geneticamente Modificados

P: Prolina

PSIPRED: Predição de estrutura Secundária

Q: Glutamina.

R: Arginina

S: Serina

T: Treonina

UPGMA: Método de grupo par com média não ponderada

V: Valina

VIP: Vegetative insecticidal proteins

VMMA: Vesículas de membrana da microvilosidade apical

W: Triptofano

Y: Tirosina

RESUMO

O uso continuado de pesticidas no controle de pragas agrícolas tem preocupado a população em relação aos efeitos nocivos destes compostos sobre os seres humanos e animais, bem como com a possível contaminação dos recursos hídricos e do solo decorrentes desta prática. Uma alternativa de controle tem sido o uso do controle biológico que vem se intensificando nas últimas décadas. Entre os métodos de controle biológicos utilizados uma alternativa viável tem sido o uso da proteína CRY produzida pelo *Bacillus thuringiensis* aplicada diretamente sobre as culturas na forma de um bioinseticida. Mais recentemente as técnicas de biologia molecular possibilitaram a clonagem do gene que codifica essa proteína e sua introdução em plantas gerando as plantas transgênicas conhecidas como plantas *Bt*. As proteínas quando ingeridas pelos insetos atuam nas membranas do intestino levando-os à morte. Atualmente mais de 600 genes que codificam essa proteína, oriundos de diferentes linhagens, foram identificados. Alguns desses genes apresentam especificidade, ou seja, a proteína produzida apresenta-se tóxica apenas para uma determinada ordem enquanto outros, por sua vez, não apresentam especificidade atuando no controle de mais de uma ordem. Neste trabalho foram analisadas e comparadas, utilizando os programas CLUSTALW 2, PSIPRED v3.3 e MEGA 6, as sequências primárias e secundárias das proteínas CRY, especificamente da região dos domínios I, II e III. Foram comparadas as sequências das proteínas codificadas pelos genes CRY1Fa1 e CRY9Ba1 cuja proteína é tóxica para os insetos da ordem Lepidóptera, e CRY7Aa1 e CRY3Aa1 (Coleóptera) e CRY1Ia1 e CRY1Ib3 cuja proteína ataca as duas ordens Lepidóptera e Coleóptera indistintamente. A comparação das sequências primárias e secundárias das proteínas CRY possibilitou a identificação de diferenças na região final do domínio II e começo do domínio III referentes às estruturas β 10 e β 11, assim como, na região entre elas (*loop* 3) que podem ser as responsáveis pela especificidade das proteínas por representarem regiões de ligação ao receptor. Comparando-se as sequências dos genes CRY1Ia1 e CRY1Fa1, especificamente na região do *loop* 2 presentes no domínio II, verificou-se que na proteína que atua sobre os insetos das ordens Lepidóptera e Coleóptera existem 24 aminoácidos DAIGTVHPHPSFTSTTWYNNAPS; enquanto na mesma região do gene CRY1Fa1 específico para Lepidóptera, existem 16 aminoácidos SSVIEDSPVSANIPNG. Essa região é sabidamente responsável pela ligação de receptores e assim as diferenças em número e tipo de aminoácidos encontrados podem influenciar a ligação dos receptores, favorecendo a ligação de diferentes receptores que poderiam ser os responsáveis pela especificidade da ação da proteína.

Palavra-chave: proteínas CRY, *Bt thuringiensis*, receptores

ABSTRACT

The continued use of pesticides in agricultural pest control has concerned the population of the harmful effects of these compounds on humans and animals, as well as the possible contamination of water and soil resulting from this practice. A control alternative has been the use of biological control that has intensified in recent decades. Between the biological control methods used a viable alternative has been the use of Cry protein produced by the *Bacillus thuringiensis* applied directly on crops as a biopesticide. More recently, molecular biology techniques have enabled the cloning of the gene encoding this protein and its introduction into plants generating transgenic plants known as Bt plants. Proteins when ingested by insects act in gut membranes causing them to death. Currently more than 600 genes encoding this protein from different strains were identified. Some of these genes have specificity, ie, the protein produced has become toxic only to a given order while others in turn, do not show specificity acting on the control of more than one order. In this study were analyzed and compared, using the CLUSTALW 2, PSIPRED v3.3 and MEGA 6 programs, the primary and secondary sequences of CRY proteins specifically in the region of domains I, II and III. The sequences of the proteins encoded by the genes CRY1Fa1 and CRY9Ba were compared whose protein is toxic to insects of the Lepidoptera order, and CRY7Aa1 and CRY3Aa1 (Coleoptera) and CRY1Ia1 and CRY1Ib3 protein which attacks both Lepidoptera and Coleoptera orders indiscriminately. Comparison of primary and secondary sequences of CRY proteins allowed the identification of differences in the final region of domain II and domain III beginning related to structures β 10 and β 11, as well as in the area between them (loop 3) that may be responsible for specificity of the proteins by representing the receptor binding. Comparing the sequences of the genes CRY1Fa1 and CRY1Ia1 specifically in this region in the field loop2 II, it was found that the protein acts on insects from the Lepidoptera and Coleoptera orders, there are 24 amino acids DAIGTVHPHPSFTSTTWYNNNAPS; while in the same region of the gene for specific CRY1Fa1 Lepidoptera, there are 16 amino acids SSVIEDSPVSANIPNG. This region is known to be responsible for receptor binding and thus the differences in the number and type of amino acids found can influence the binding of the receptors by promoting the binding of different receptors could be responsible for the specificity of action of the protein.

Keywords: proteins Cry, *Bt thuringiensis*, receptors

1. INTRODUÇÃO

Os insetos são o maior grupo entre os animais, tem grande adaptação aos mais variados hábitos e habitats, sendo grande competidores do homem na terra; pois os insetos herbívoros atacam plantações e também os alimentos que são estocados, ocorrendo perdas econômicas; são vetores de doenças em humanos e animais de criação que são prejudiciais à saúde e ainda como pragas sociais e urbanas como as baratas, moscas e cupins.

Um dos maiores fatores de perda de grãos na agricultura é causada pelos insetos, onde denomina-se Praga Agrícola ou Florestal uma população de organismos capazes de causar danos às plantas, seus produtos e subprodutos. Podem afetar o rendimento do produto ou sua qualidade, através do consumo direto dos tecidos ou órgãos da planta, frutos ou sementes, sucção de seiva, transmissão de doenças, competição por espaço e por nutrientes e ainda deve levar em consideração o custo do controle destas pragas. O menor rendimento das colheitas, o valor depreciado dos produtos pelo efeito do dano causado pelas pragas, aliado ao custo nas medidas de controle, significam para o agricultor uma redução importante em seus lucros.

. Dentro da classe de insetos a grande maioria das pragas agrícolas pertencem às ordens Coleóptera, Lepidóptera, Hemíptera, Himenóptera, Isóptera, Orthoptera, Díptera e Thysanoptera. Jamais um só inseto poderá produzir um dano que compense a sua eliminação da cultura. Apenas quando a densidade populacional atinge população crítica, é que eles irão consumir uma quantidade de alimento que produzirá um prejuízo para a planta explorada pelo homem. Na agricultura, o conceito de inseto-praga está diretamente relacionado com os efeitos econômicos produzidos pela sua alimentação nas plantas.

Para se fazer o controle de pragas ou doenças em qualquer cultura, é preciso que haja perdas de grãos acarretando perda econômica; tudo que faz com que diminua a produtividade da lavoura é motivo de preocupação por parte dos agricultores que chegam, às vezes, ao exagero, tomando medidas drásticas visando à solução do problema, porém com isso, prejudicando o meio ambiente e contaminando alimentos.

Há séculos a qualidade da produção agrícola sempre foi intensamente afetada pelo aparecimento de insetos e ervas daninha. Desta forma, tornou-se necessária a utilização de agrotóxicos ou pesticidas de diversas classes químicas. As suas utilizações na agricultura incluem a elevação da produção com aumento da produtividade, a melhoria da qualidade dos produtos e a redução do trabalho. Sem dúvida esses objetivos foram alcançados nas últimas

décadas. No entanto, esta prática com o uso indiscriminado e pouco criterioso de agrotóxicos trouxe e continua trazendo problemas muito sérios para o ambiente, pois atinge todos os insetos presentes no ecossistema, inclusive os insetos polinizadores que são benéficos as plantações, além de contaminar os lençóis freáticos, podendo ser prejudicial à saúde humana.

Para diminuir estes efeitos ao meio ambiente pesquisas recentes têm mostrado o emprego de controle biológico através do uso de bioinseticidas como uma saída viável no combate às pragas por serem mais específicos que os compostos químicos, menos danosos ao homem e meio ambiente em geral e, quando usados de forma correta, se mostram também eficientes. E também têm sido desenvolvidos outros processos de controle de pragas, como a produção de plantas transgênicas ou Organismos Geneticamente Modificados (OGM) mais resistentes a predação por insetos.

O uso do *Bacillus thuringiensis* (*Bt*) como um bioinseticida é uma alternativa viável para o controle de insetos na agricultura e outras áreas tais como, saúde (controle de vetores de doenças), pois a atividade entomopatogênica dessa bactéria é decorrente da produção da toxina CRY, uma inclusão protéica cristalina que a bactéria gram-positiva produz durante o processo de esporulação.

As proteínas CRY não têm ação entomopatogênica na forma como são sintetizadas, pois na forma original elas tornam-se inacessíveis pela absorção do inseto alvo, mas após a sua dissolução em um ambiente básico, característico do intestino dos insetos, e, processadas por enzimas específicas, proteases, essas proteínas CRY são transformadas em toxinas. Essas toxinas além de serem tóxicas a nematóides, ácaros e protozoários, também se mostram com propriedades entomopatogênicas frente a várias ordens de insetos, tais como Lepidóptera, Díptera, Coleóptera, Hymenoptera, Hemíptera, Orthoptera, Isoptera e Malophaga. A morte da larva do inseto é consequência de uma ligação irreversível que ocorre em receptores específicos da membrana nas células epiteliais do intestino do inseto.

A patogenicidade e a especificidade de uma linhagem de *B. thuringiensis* são determinadas pelos tipos de genes *CRY* funcionais que a mesma apresenta. Uma linhagem de *B. thuringiensis* pode conter uma ou várias cópias de um mesmo gene *CRY* ou de diferentes genes cujos produtos formarão o mesmo cristal.

Até o momento, mais de 600 genes diferentes foram depositados nos bancos de dados no sítio (http://www.lifesci.sussex.ac.uk/home/Neil_Crickmore/Bt/). Dentre esses genes,

verifica-se que alguns codificam proteínas tóxicas para determinadas ordens, enquanto outros não mostram a mesma especificidade podendo ser tóxicas para duas ordens. Essa especificidade pode estar associada a alterações na sequência primária das proteínas, que podem promover alterações na estrutura secundária e terciária. Considerando que os aminoácidos apresentam características diferentes a substituição de um aminoácido com carga diferente pode alterar sua hidrofobicidade, deslocando o posicionamento deste aminoácido dentro da estrutura terciária importante na manutenção da sua função, assim como, alteração a ligação a receptores específicos.

Pela comparação das sequências de aminoácidos de diferentes genes *CRY*, buscamos encontrar indícios que expliquem a especificidade observada em relação à toxina. Essas informações poderão ser importantes na geração de novos cultivares transformados.

Várias culturas, contendo o gene que codifica a proteína cristal oriunda de *B. thuringiensis*, têm sido produzidas e comercializadas. A identificação de novos genes e de sua especificidade possibilitará o desenvolvimento de plantas com resistência específica ou ainda a associação de genes visando à múltipla resistência em plantas, pois plantas modificadas pela adição do gene que codifica a proteína cristal de *B. thuringiensis* oferecem segurança e controle efetivo de insetos.

2. REVISÃO BIBLIOGRÁFICA

2.1 PROTEÍNAS

Segundo LEHNINGER, NELSON e COX (2007), a estrutura primária de uma proteína refere-se à sequência de resíduos de aminoácidos que a compõem. Os aminoácidos por sua vez em se arranjam em estruturas secundárias típicas e através de interações entre seus aminoácidos, ligações covalentes se enovelam sobre si mesmos formando a estrutura terciária, na qual a proteína se apresenta na sua forma ativa.

2.2 AMINOÁCIDOS

Aminoácidos constituem a estrutura básica de uma proteína. Os aminoácidos são compostos por uma cadeia principal de átomos (N, C, O) comum, ligada a uma cadeia lateral denominada cadeia lateral R onde reside a particularidade do aminoácido (RICHARDSON, 1981). Essa cadeia lateral por apresentar diferentes propriedades e características como composição, estruturas e tamanho confere ao aminoácido características próprias de polaridade hidrofobicidade entre outras.

LEHNINGER, NELSON e COX (2007) apresenta uma classificação destes aminoácidos segundo suas características principais de acordo as propriedades do grupo R, principalmente nas características de polaridade e interação com a água em pH biológico próximo de 7. São divididos em 5 grupos: Não Polares e Alifáticos, Aromáticos, Não carregados, mas polares, Positivamente e Negativamente.

- Não Polares e Alifáticos: Possuem apenas carbono e hidrogênio em sua formação, além de não conterem anéis aromáticos, são aminoácidos hidrofóbicos (que possuem menor afinidade pelo solvente polar, água, voltando-se para o interior da proteína) e não polares. Pertencem a este grupo os aminoácidos: Alanina (A), Glicina (G), Prolina (P), Leucina (L), Isoleucina (I), Valina (V) e Metionina (M).
- Aromáticos – que possuem anel aromático em sua composição, com suas cadeias laterais relativamente apolares participam de interações hidrofóbicas. Pertencem a este grupo os aminoácidos: Fenilalanina (F), Tirosina (Y) e Triptofano (W).
- Não carregados, mas polares: seus grupos R são mais solúveis em água (hidrofílicos) pois contém grupos funcionais que tendem a participar de ligações de hidrogênio com

a água. Pertencem a este grupo os aminoácidos: Serina (S), Treonina (T), Cisteína (C), Asparagina (N) e Glutamina (Q).

- Carregados Positivamente: Favoráveis à reações com ânions (íons carregados negativamente). Pertencem a este grupo os aminoácidos: Arginina (R), Histidina (H) e Lisina (K).
- Carregados Negativamente: Favoráveis às reações com cátions (íons carregados positivamente). Pertencem a este grupo os aminoácidos: Aspartato (D) e Glutamato (E).

2.3 *Bacillus thuringiensis*

O *B. thuringiensis* tornou-se um dos microrganismos mais utilizados mundialmente por destacar-se dentre os organismos empregados no controle biológico, em função de sua atividade tóxica contra diversas ordens de insetos dentre elas Lepidóptera, Diptera, Coleóptera, Hemiptera, Hymenoptera, Homoptera e Orthoptera (VAN FRANKENHUYZEN, 2009). A seletividade apresentada pelas toxinas produzidas pelo *Bt* possui pouco efeito sobre insetos não alvos e vertebrados como aves e mamíferos (SIEGEL, 2001) e ainda é uma boa opção para o controle de pragas na agricultura e vetores de doenças (CRICKMORE et al., 2011; PARDOLÓPEZ et al., 2013), atingindo organismos como nematóides (SCHNEPF et al., 1998), contribuindo assim para esta sua aceitação mundial.

No século passado, um bacteriologista japonês, Ishawata, isolou um bacilo da larva morta do bicho da seda *Bombyx mori*, e a morte das larvas foram devido à presença de uma toxina no meio (ISHAWATA, 1901). Um organismo similar foi isolado uma década mais tarde a partir de larvas mortas de *Anagasta kuehniella*. Esse organismo era a bactéria *Bt* que recebeu esse nome pelo fato de ter sido isolado na província de Turíngia, na Alemanha. Essa bactéria apresenta a seguinte classificação: Classe: Firmibacteria; Ordem: Eubacteriales; Família: Bacillaceae; Gênero: Bacillus; Espécie: *B. thuringiensis* (BERLINER, 1911).

B. thuringiensis é uma bactéria de ocorrência cosmopolita, sendo encontrado em todas as partes do mundo, e de ocorrência ubíqua (KRYWUNCZYK; FAST, 1980) encontrado em vários substratos como solo, água, superfície de plantas, insetos mortos, teias de aranha e grãos armazenados (BRAVO et al., 1998). É uma bactéria gram-positiva e aeróbica, podendo facultativamente crescer em anaerobiose, dentro da faixa de 10 a 40°C. Apresenta-se na

forma de bastões isolados, aos pares ou em cadeias com tamanhos que variam entre 0,5 a 2,5 μm de largura por 1,2 a 10,0 μm de comprimento (MIRALLES; PEREZ, 2004) além de possuírem flagelos peritríquios que permitem sua locomoção (GLARE; O'CALLAGHAM, 2000). A atividade entomopatogênica desta bactéria se deve à presença de inclusões proteicas cristalinas, produzidas durante a esporulação. Esses cristais compostos por proteínas denominadas endotoxinas ou proteínas cristal podem ser visualizados por microscopia de contraste de fases (BRAVO et al., 1998; MONNERAT; BRAVO, 2000).

Além das proteínas CRY, *B. thuringiensis* pode produzir uma série de outras toxinas, incluindo proteínas com atividade inseticida como α -exotoxinas, β -exotoxinas, hemolisinas, enterotoxinas, quitinases, fosfolipases (HANSEN; SALAMITOU, 2000) e as VIP (vegetative insecticidal proteins) (ESTRUCH et al., 1996). Os esporos também podem contribuir com a patogenicidade através da ação sinérgica realizada juntamente com as δ -endotoxinas (JOHNSON; McGAUGHEY, 1996).

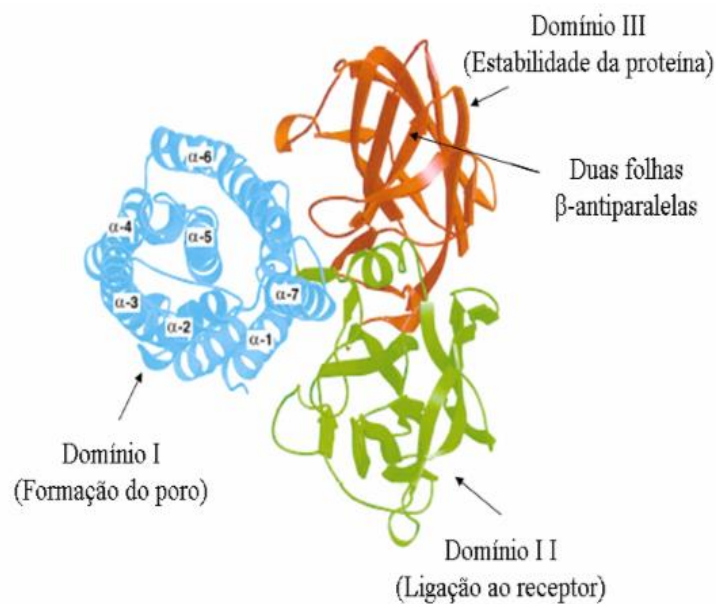
2.4 CARACTERÍSTICAS E MODO DE AÇÃO DAS TOXINAS CRY

A toxicidade das proteínas CRY está associada às suas regiões específicas e estas regiões são utilizadas para a sua classificação. Atualmente, as toxinas de cristal têm sido classificadas com base na homologia de sequência de aminoácidos, no qual cada pró-toxina recebeu um nome que consiste no CRY mnemônico (ou Cyt) e quatro fileiras hierárquicas compostas por números, letras maiúsculas, letras minúsculas e números (por exemplo, CRY25Aa1), de acordo com o seu lugar em uma árvore filogenética. As proteínas CRY com cerca de 50% de homologia em suas sequências são colocadas em primeiro lugar no ranque e quando apresentam 78 e 95% de identidade, constituem o segundo e terceiro lugares no ranque, respectivamente (DE MAAGD et al., 2001).

Quando realizado o alinhamento das toxinas CRY, são encontrados cinco blocos de sequências comuns a uma grande maioria das proteínas. A estrutura tridimensional das formas ativas das toxinas CRY1 (GROCHULSKI et al., 1995), CRY2 e CRY3 (LI et al., 1991), quando analisadas em cristalografia de raio X, mostraram-se muito similares, cada uma apresentando três domínios (Figura 1). O domínio I (N-terminal) é um conjunto de sete α -hélices anfipáticas sendo que a hélice-A5 hidrofóbica é rodeada por outras seis hélices anfipáticas, e este domínio é responsável pela inserção da membrana de poros. O domínio II é composto de onze folhas beta com regiões de loop envolvidos na ligação à receptores

específico do intestino médio e o domínio III é um β -sanduíche (LI et al., 1991; GROCHULSKI et al., 1995; MORSE et al., 2001; GALITSKY et al., 2001; BOONSERM et al., 2005, 2006). As regiões expostas dos domínios III estão envolvidas na ligação ao receptor (BRAVO et al., 2005).

Figura 1 - Representação esquemática da estrutura da toxina CRY3A



Fonte - Li et al., 1991.

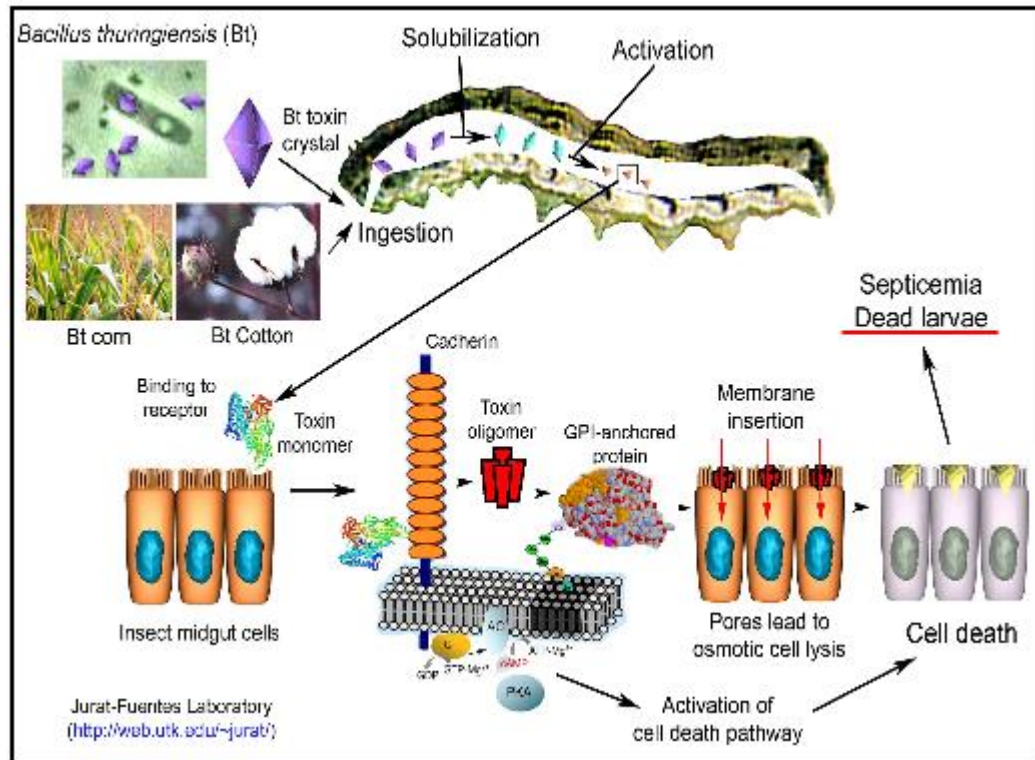
As delta-endotoxinas foram classificadas em vários grupos de proteínas (CRICKMORE et al., 1998) e existem mais de 400 proteínas CRY descritas, que estão classificadas em 55 grupos (http://www.lifesci.sussex.ac.uk/home/Neil_Crickmore/Bt/).

Quando ingerido pela larva o cristal se solubiliza em pH alcalino no intestino médio dos insetos, liberando as pró-toxinas (SCHNEPF et al., 1998). Diferenças na solubilização podem determinar a especificidade do isolado de *B. thuringiensis* à espécie alvo, tanto pela alcalinidade do sistema digestivo quanto pela composição dos cristais de *B. thuringiensis* (ARONSON et al., 1991). Após a solubilização, muitas pró-toxinas são processadas por proteases; essas toxinas ativas ligam-se então a um receptor primário do tipo caderina, presente na membrana das micro vilosidades das células epiteliais do intestino médio.

Após o reconhecimento do receptor, a toxina induz a formação de poros na membrana celular do epitélio intestinal. A formação dos poros na membrana celular provoca o desequilíbrio iônico entre o citoplasma e o meio externo à célula, ocorrendo à intoxicação

dos insetos e a destruição das micro vilosidades, hipertrofia das células epiteliais, vacuolização do citoplasma e lise celular, levando o inseto à paralisa e morte (Figura 2).

Figura 2 - Representação esquemática do modo de ação da toxina Cry de *B. thuringiensis*
Adaptado de <http://web.utk.edu/~jurat/>



2.5 UNIÃO DOS RECEPTORES DE CRY

Os estudos referentes a cinética de união das toxinas CRY às vesículas da membrana da microvilosidade apical dos insetos susceptíveis é bifásica. A primeira etapa é reversível e ocorre pelo reconhecimento do receptor podendo levar à dissociação do complexo toxina-receptor. Esta etapa envolve a interação entre a toxina e seu sítio de união (união reversível), que é um requisito básico para toxicidade, mas não suficiente (SCHNEPF et al., 1998). A segunda etapa ocorre a inserção da toxina na membrana apical da célula tornando a ligação irreversível (VAN RIE et al., 1989). Nesta etapa, a união irreversível a receptores específicos e a inserção na membrana parecem estar mais ligadas com a toxicidade (VAN RIE et al., 1989; MONNERAT; BRAVO, 2000). A interação inicial está relacionada à ligação entre a toxina e o carboidrato do receptor, enquanto a irreversível está associada com uma interação proteína-proteína (LIANG et al., 1995).

As metodologias mais utilizadas são os estudos cinéticos de união utilizando proteínas marcadas com biotina e Vesículas de Membrana da Microvilosidade Apical (VMMA) ou “Brush Border Membrane Vesicles” (BBMVs).

São conhecidos pelo menos quatro tipos de receptores proteicos envolvidos no processo de ligação em diferentes larvas de lepidópteros: a proteína do tipo caderina (CADR), uma fosfatase alcalina (ALP) ancorada a glicosilfosfatidil-inositol (GPI), uma aminopeptidase-N (APN) ancorada a glicosilfosfatidil-inositol (GPI) e um glicoconjugado de 270 kDa (GCR) (Figura 2) (BRAVO et al., 2007).

As proteínas caderina fazem parte de uma grande família de glicoproteínas que são responsáveis dentre outras funções pelo contato inter-celular, organização do citoesqueleto e morfogênese (ANGST et al., 2001). São proteínas transmembranares com um domínio citoplasmático e um domínio extracelular com várias repetições de caderina com comprimento de cerca de 110 aminoácidos (VADLAMUTI et al., 1995).

Em 1993 uma proteína foi isolada do epitélio do intestino médio de *Manduca sexta* a análise de sua sequência de aminoácidos revelou a existência de um peptídeo sinal, 12 repetições de caderina, um domínio extracelular proximal de membrana, um domínio transmembranar e um pequeno domínio citoplasmático (DORSCH et al., 2002).

Posteriormente, novas caderinas de lepidópteros foram identificadas revelando apresentar mesma organização de domínio, dentre elas os receptores BtR175 de *Bombyx mori* e HevCaLP de *Heliothis virescens* (GAHAN et al., 2001; MORIN et al., 2003; WANG et al., 2005; FLANNAGAN et al., 2005).

As proteínas caderinas de lepidópteros foram identificadas na membrana apical das células do epitélio do intestino médio (AIMANOVA et al., 2006), local alvo das toxinas CRY (CHEN et al., 2005), enquanto as caderinas clássicas encontram-se principalmente dentro de junções envolvidas na adesão célula-célula (ANGST et al., 2001),

A interação das proteínas CRY com a CADR é uma interação complexa que envolve três epítomos da CADR correspondentes às regiões extracelulares denominados CR7, CR11 e CR12, em que CR12 é proximal ao domínio de membrana da caderina. Estes epítomos da proteína CADR interagem com as alças expostas 2, 3 e α -8 do domínio II da toxina, promovendo uma clivagem proteolítica adicional da extremidade N-terminal, e também da hélice α -1 do domínio I (GÓMEZ et al., 2002; ATSUMI et al., 2008).

A clivagem da hélice α -1 expõe regiões hidrofóbicas do domínio I, e foi levantada a hipótese de que a clivagem desta hélice é necessária para desencadear a formação da estrutura oligomérica do pré-poro da toxina antes da inserção na membrana (GÓMEZ et al., 2002; PACHECO et al., 2009; ARENAS et al., 2010).

As aminopeptidases N (APN) são exopeptidases e atuam no intestino larval de lepidópteros juntamente com endopeptidases e carboxipeptidases digerindo as proteínas derivadas da dieta do inseto (WANG et al., 2005). As APNs têm sido extensivamente estudadas pelo seu papel na digestão e como receptores das toxinas CRY de *B.thuringiensis*, uma vez que foi mostrado que as proteínas CRY podem se ligar a APN (KNIGHT et al., 1994; SANGADALA., 1994). A interação da proteína CRY com a APN ocorre através da alça-3 exposta do domínio II e a interação com a ALP é estabelecida através da folha- β 16 do domínio III (MASSON et al, 1995; PACHECO et al, 2009; ARENAS et al., 2010).

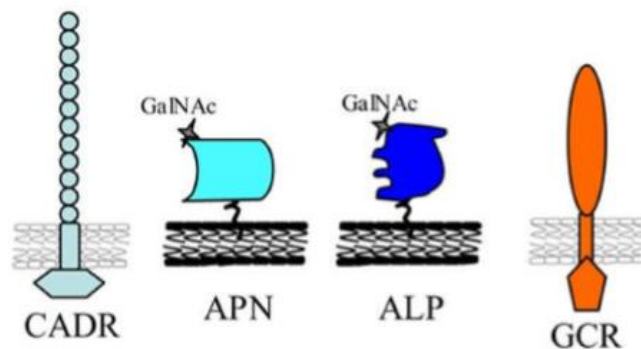
Através do domínio III que reconhece, especificamente, a Nacetilgalactosamina (GalNAc), a toxina CRY1Ac se liga ao receptor APN (MASSON et al., 1995) e após a ligação a toxina CRY se une rápida e irreversivelmente à membrana plasmática das células epiteliais, com subsequente abertura ou formação de poros, ocorrendo, assim, um desequilíbrio osmótico entre o meio intra e extracelular, ocasionando perda da integridade da membrana do intestino das larvas de insetos suscetíveis e consequente morte celular.

As ALPs, APNs e CADR interagem de modo consecutivo com diferentes estruturas da toxina ao longo do mecanismo de ação das proteínas CRY. Em larvas de lepidópteros as toxinas ativas CRY1A primeiramente se ligam aos receptores ALP e APN através de uma interação de baixa afinidade (MASSON et al., 1995; PACHECO et al., 2009; ARENAS et al., 2010). A interação com ALP e APN concentra a toxina ativada na microvilosidades da membrana das células do intestino médio tornando-as aptas a se ligar em através de uma interação de alta afinidade com os receptores CADR (VADLAMUDI et al., 1995; GÓMEZ et al., 2006; PACHECO et al., 2009; ARENAS et al., 2010).

Após a toxina ativa se ligar à CADR ocorre sua primeira mudança conformacional, na qual a região N-terminal da proteína incluindo α -hélice 1 é exposta e clivada por proteases da membrana promovendo a oligomerização das toxinas CRY (na forma de um tetrâmero) que corresponde à estrutura chamada pré-poro (GÓMEZ et al., 2002; ATSUMI et al., 2008).

A estrutura oligomérica CRY apresenta maior afinidade pelos receptores ALP e APN, assim o oligômero se liga a estes receptores e em seguida ocorre a inserção da toxina ativa na membrana de forma irreversível induzindo a abertura ou formação de poros que provocam uma quebra no balanço osmótico da célula e consequente lise celular (PARDO-LÓPEZ et al., 2006; ARENAS et al., 2010; PARDO-LÓPEZ et al., 2013).

Figura 3- Receptores moleculares das proteínas CRY1A.



Fonte - BRAVO et al., 2007.

2.6 PLANTAS GENTICAMENTE MODIFICADAS COM GENES DE *B. thuringiensis*

Atualmente são conhecidas inúmeras espécies de bactérias associadas a doenças em insetos (COSTA et al., 2010), mas nem todas apresentam características importantes e necessárias à aplicação do controle biológico de inseto-praga das plantas cultivadas (FIUZA, 2001). Entre essas bactérias capazes de exercer controle sobre insetos, o gênero *Bacillus* apresenta especial importância no controle biológico de pragas, destacando-se o *B. sphaericuse* o *B. thuringiensis* (PRIEST, 1992; CRICKMORE et al., 2008).

As técnicas de engenharia genética permitiram o desenvolvimento de um sistema que consiste na transferência e expressão de genes exógenos em bactérias e plantas. A utilização destas técnicas permitiu a criação de cultivares geneticamente modificados com características específicas para a resistência a herbicidas, insetos, patógenos e estresse ambiental (LIU et al., 2005); características estas que proporcionam aumento na produção e redução de gastos com controle de doenças e insetos, proporcionando maior rendimento econômico (LIU et al., 2005).

Em 1981, o primeiro gene CRY foi clonado e expresso em *Escherichia coli* (SCHNEPF; WHITLEY, 1981). Devido à complexidade do processo de transformação de

plantas transgênicas, somente em 1987 foram obtidas as primeiras plantas *Bt*, mediante a inserção de genes *CRY* que codificam a síntese de proteínas inseticidas em plantas de tomate (proteína CRY1Ab) e tabaco (proteína CRY1Ac) (FISCHHOFF et al., 1987; POLANCZYK et al., 2003). A partir daí foram introduzidos genes *CRY* em plantas como algodão, arroz, milho, batata, canola e soja (JOUANIN et al., 1998; PARROTT et al., 1994; MACRAE et al., 2005; HOMRICH et al., 2008).

O plantio de cultivares geneticamente modificada está associado às culturas de maior importância econômica como: soja, algodão e canola. Em países aonde esta tecnologia vem sendo desenvolvida, a área utilizada para o cultivo comercial de OGMs cresceu consideravelmente. Atualmente 21 países adotam o plantio de OGMs sendo principais produtores os Estados Unidos, seguidos por Argentina e Brasil, respectivamente (JAMES, 2005). Segundo James (2005), o cultivo de transgênicos nestes três países representa 82% de toda a produção de OGMs no mundo.

Esforços têm sido concentrados no incremento da expressão dos genes de *Bt* em plantas, na seleção de novas variantes de *Bt* mais ativas e/ou na modificação das sequências dos genes *CRY* de maneira a aumentar a produção de toxinas no interior das plantas. As sequências de genes *CRY* de *Bt* apresentam um alto conteúdo de bases adenina/timina (A/T) quando comparados a genes de plantas, que tendem a ter um alto conteúdo de bases guanina/citosina (G/C) (de MAAGD et al., 1999). Como consequência, os códons preferenciais dos genes de *Bt* são ineficientes em plantas, determinando a não tradução ou uma meia-vida do mRNA muito curta, o que leva à expressão reduzida destes genes. A substituição de nucleotídeos na sequência codificadora por meio da síntese química ou da utilização de formas truncadas auxilia a correta leitura para tradução em plantas. Porém, devem-se adequar as modificações a cada espécie vegetal a ser transformada, tendo-se o cuidado de que a síntese (parcial ou total) gere a mesma proteína do gene nativo (ESTRUCH et al., 1997; IANNACONE et al., 1997; de MAAGD et al., 1999). No início da década de 90, foram obtidas plantas com aumento significativo da expressão dos genes de *Bt*, capazes de conferir efetivo controle de pragas do algodão (PERLAK et al., 1990), batata (PERLAK et al., 1993; ADANG et al., 1993) e milho (KOZIEL et al., 1993).

A identificação de novos genes e de sua especificidade irá possibilitar o desenvolvimento de plantas com resistência específica ou ainda a associação de genes visando a múltipla resistência em plantas de valor comercial e segundo Maagd et al. (1999) as plantas

Bt oferecem segurança e controle efetivo de insetos. O gene que codifica a proteína cristal de *Bt* já foi utilizado na transformação de diferentes espécies de plantas cultivadas de interesse comercial. Um exemplo das plantas transformadas utilizando o gene *CRY* está apresentado na Tabela 1.

Tabela 1 - Plantas transformadas com genes *CRY* de *Bacillus thuringiensis* em processo de Comercialização

Nome comum	Nome científico	Inseto alvo
Soja	<i>Glycine canescens</i>	<i>Anticarsia gemmatalis</i>
Laranja	<i>Citrus sinensis</i>	<i>Lymantria dispar</i>
Algodão	<i>Gossypium hirsutum</i>	<i>Heliothis</i> sp
Tabaco	<i>Nicotiana tabacum</i>	<i>Manduca Sexta</i>
Alfafa	<i>Mendicago sativa</i>	<i>Spodoptera</i> spp
Canola	<i>Brassica napus</i>	<i>Agrotis</i> spp
Milho	<i>Zea mays</i>	<i>Ostrinia nubilalis</i>
Milho	<i>Zea mays</i>	<i>Spodoptera frugiperda</i>
Milho	<i>Zea mays</i>	<i>Helicoverpa zea</i>
Arroz	<i>Oryza sativa</i>	<i>Chilo</i> spp
Arroz	<i>Oryza sativa</i>	<i>Spodoptera</i> spp
Batata	<i>Solanum tuberosum</i>	<i>Manduca</i> spp

Fonte - Embrapa- Documentos, número, 54 - ISSN0104-6187

Desde a liberação comercial das plantas-*Bt* nos EUA, os agricultores têm adotado esta tecnologia visando a um efetivo aumento de produção à agricultura sustentável (DUNWELL, 1999). Neste contexto, pode-se mencionar como benefícios do uso desta tecnologia:

- As plantas-*Bt* têm demonstrado vantagens sobre os biopesticidas, pois a eficiência da produção de proteínas *CRY* pelas plantas-*Bt* não é afetada por fatores ambientais como chuva após a aplicação, incidência de radiação solar e altas temperaturas (BETZ et al., 2000).
- As proteínas não são tóxicas para seres humanos e animais domésticos, nem se acumulam nos tecidos adiposos ou persistem no ambiente como alguns inseticidas químicos (BISHOP et al., 1999; SIEGEL, 2001). A exposição de seres humanos e outros animais a estas proteínas é menor do que a exposição aos inseticidas normalmente utilizados. Isto porque os inseticidas químicos e biológicos são normalmente aplicados em grandes quantidades nas folhas, enquanto as proteínas *CRY*, em plantas-*Bt*, são produzidas em pequenas quantidades no interior da planta e aparecem em níveis ainda mais baixos no pólen (BISHOP et al., 1999).

- As proteínas CRY exibem um alto grau de especificidade para os insetos-alvo e espécies relacionadas, devendo ser ingeridas para exercerem seus efeitos, uma vez que não possuem atividade por contato (BETZ et al., 2000).

Nos últimos anos, o impacto negativo dos produtos químicos sobre o meio ambiente tem se agravado com a poluição do solo e mananciais de água doce, acúmulo na cadeia alimentar e com os consequentes problemas de saúde pública. O percentual de redução nas aplicações de inseticidas químicos com o emprego de plantas-*Bt* foi variável, dependendo da cultura.

A proteção inerente das plantas-*Bt* foi traduzida em aumento de produtividade agrícola, um exemplo disso foi verificado em relação às perdas anuais de milho nos EUA devido aos danos causados pelo besouro *L. decemlineata* que variavam em torno de 33 e 300 milhões de toneladas/ano e a redução destas perdas com o emprego do milho-*Bt* levou ao aumento da produção final de milho (BETZ et al., 2000; PEFERÖEN, 1997).

2.7 BIOINFORMÁTICA

A bioinformática tem sido utilizada em diversas áreas como a construção de banco de dados e a mineração de dados; análises de sequências; para identificar gene, predizer suas funções e demonstrar relações entre genes e proteínas; prever a conformação tridimensional das proteínas; construir árvores filogenéticas e modelos evolutivos; construir bibliotecas genômicas; estudar as funções biológicas.

2.7.1 ALINHAMENTO DE SEQUÊNCIAS

A finalidade do alinhamento de sequências (completo ou parcial) de gene ou proteínas é identificar a relação que elas possuem com outros genes ou proteínas. Usando o alinhamento de sequências é possível identificar domínios ou sentidos para DNAs (Ácido Desoxirribonucléico) ou sequências de proteínas que são compartilhados entre um grupo de moléculas. Existem 2 tipos de alinhamentos: em pares e múltiplos. O uso do alinhamento múltiplo de sequências estabelece relações filogenéticas entre diferentes proteínas, encontra domínios conservados e compara proteína de interesse de forma mais detalhada com os membros da mesma família de proteínas.

O programa CLUSTALW2 é um exemplo de programa básico para alinhamento múltiplo de sequências (DNA ou proteínas), produzindo alinhamentos múltiplos biologicamente significativos de sequências divergentes, calculando o melhor score e alinhando as sequências para que identidades, similaridades e diferenças possam ser visualizadas.

O trabalho feito por BOONSERM et al. (2005), utilizando alinhamento múltiplo foi relatada que a proteína CRY4Ba específica da díptera é mais semelhante que a proteína Cry1Aa específica da Lepidóptera do que a proteína Cry3Aa específica da Coleóptera em relação Lepidóptera/Díptera Cry2Aa.

2.7.2 ÁRVORES FILOGENÉTICAS

Devido ao constante aprimoramento e relativa simplicidade na execução as técnicas moleculares têm sido amplamente utilizadas nas análises filogenéticas (BUSO, 2005). E a maior vantagem dos métodos moleculares é a exclusão das influências ambientais e a investigação direta da situação genotípica, permitindo a detecção de variação ao nível de DNA.

Filogenia molecular é o estudo das relações evolutivas entre genes, organismos, ou proteínas combinando as técnicas estatísticas com a Biologia Molecular (POLANSKI; KIMMEL, 2007).

Nos estudos filogenéticos são usados diferentes métodos para se identificar os membros distintos de uma comunidade. Por exemplo, quando se trata de uma família de sequências de genes ou proteínas a análise filogenética determina quanto a família pode ter derivado durante a evolução (MOUNT, 2004).

O padrão de ramificação de uma árvore é a topologia e o objetivo da análise filogenética é identificar as relações entre os galhos de uma árvore e os seus comprimentos (MOUNT, 2004). Para estabelecer estas relações podem ser utilizadas diferentes abordagens na construção de uma árvore filogenética (POLANSKI; KIMMEL, 2007).

No trabalho de BOONSERM et al. (2005), entre as comparações estruturais entre toxinas CRY com atividades específicas contra diferentes ordens de insetos utilizou o recurso de filogenia, criando um dendograma de diversidade estrutural derivada da sobreposição dos três domínios em conjunto e separadamente.

3 JUSTIFICATIVA

Até o momento já foram identificados cerca de 600 genes de *B. thuringiensis* que codificam proteínas tóxicas a diversas ordens de insetos. Essas sequências estão disponíveis no sítio (http://www.lifesci.sussex.ac.uk/home/Neil_Crickmore/Bt/). A ação destas proteínas contra as diferentes ordens é bastante específica, embora haja casos de proteínas cuja ação seja efetiva para duas ordens. A identificação de padrões na sequência de aminoácidos destas proteínas que possam estar associados a sua especificidade, poderá facilitar o uso destes genes na geração de novas plantas transgênicas resistentes as diferentes pragas agrícolas, bem como a construção de plantas piramidizadas, hoje uma realidade.

4 HIPÓTESE

As alterações dos aminoácidos nos domínios da proteína CRY são responsáveis pelas diferentes especificidades nas ordens dos insetos Lepidópteros e Coleópteros.

5 OBJETIVOS

5.1 OBJETIVOS GERAIS

Comparar as sequências de aminoácidos dos diferentes genes *CRY* já disponíveis e baseados na semelhança e diferença encontrada, explicar a especificidade da sequência.

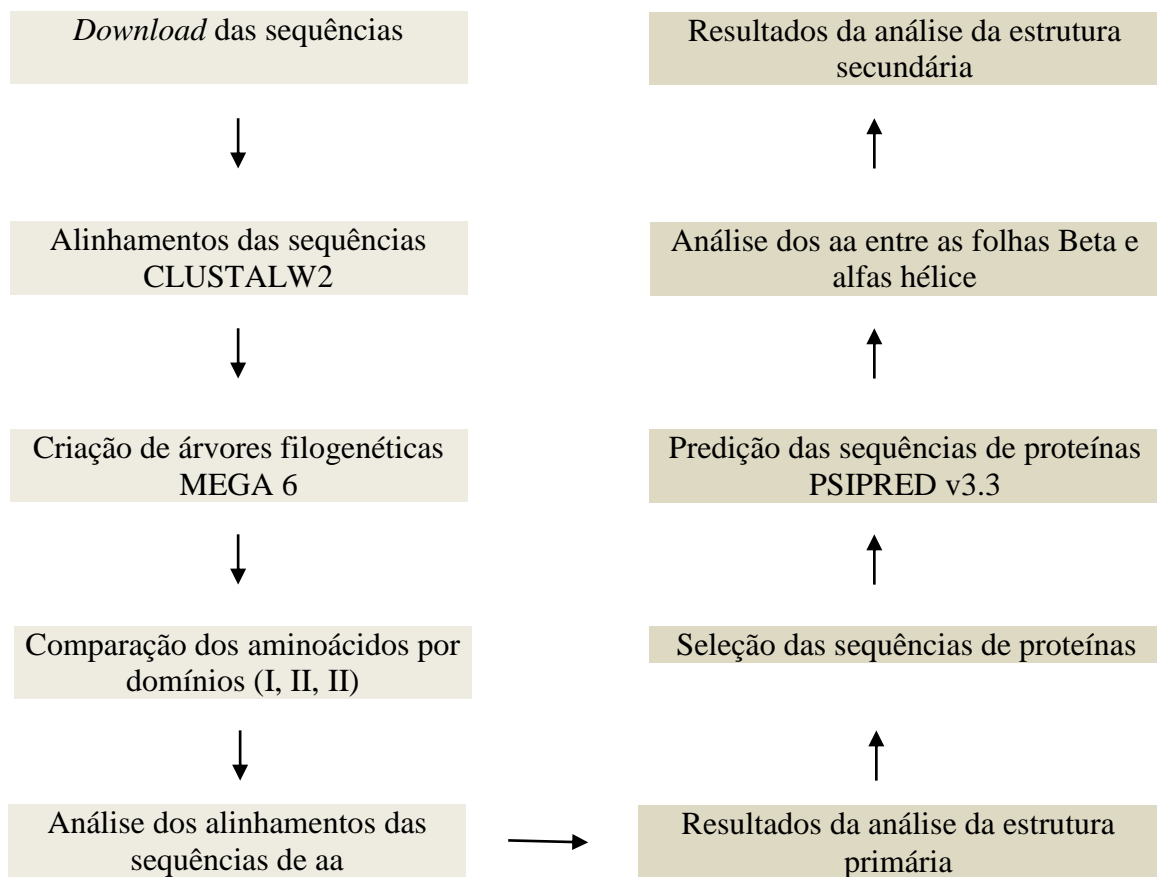
5.2 OBJETIVOS ESPECÍFICOS

- Buscar as sequências de aminoácidos dos genes nos bancos de dados;
- Fazer alinhamento das sequências de aminoácidos;
- Analisar as sequências dos domínios I, II e III entre as diferentes ordens de insetos;
- Verificar as diferenças de aminoácidos entre as sequências e suas características;
- Análise de estrutura secundária;
- Análise filogenética baseada na sequência de aminoácidos.

6 METODOLOGIA

O projeto foi desenvolvido na Unidade de Biotecnologia na Universidade de Ribeirão Preto – UNAERP e no laboratório de Informática do IFSULDEMINAS (Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais) – Campus Machado. O diagrama da (Figura 3) ilustra todas as etapas presentes nesta metodologia, dividida em duas partes: verificação de diferenças e semelhanças na estrutura primária (cinza) e na estrutura secundária (marron-claro).

Figura 4: Apresentação geral deste trabalho e suas etapas



6.1 DOWNLOAD E ALINHAMENTO DAS SEQUÊNCIAS DE AMINOÁCIDOS

Busca e *download* das sequências de proteínas CRY disponíveis no sítio (http://www.lifesci.sussex.ac.uk/home/Neil_Crickmore/Bt/). Os *downloads* foram realizados pelas sequências inteiras de aminoácidos e também separados por domínios (Domínio I, II e III) totalizando 97 sequências de aminoácidos, sendo 4 *downloads* para cada sequência. O

download executado das 97 seqüências em estudo resultou em 17 seqüências de genes, cuja toxina ataca (que apresenta toxicidade) insetos da ordem Lepidóptera e Coleóptera: CRY1Ia1, CRY1Ia2, CRY1Ia3, CRY1Ia4, CRY1Ia5, CRY1Ia7, CRY1Ia8, CRY1Ia11, CRY1Ia12, CRY1Ia13, CRY1Ia14, CRY1Ib1, CRY1Ib2, CRY1Ib3, CRY1Ic1, CRY1Id1, CRY1Ie1; 46 seqüências de genes, cuja toxina ataca insetos da ordem Lepidóptera: CRY1Ba1, CRY1Ba2, CRY1Ba3, CRY1Ba5, CRY1Ba6, CRY1Bb1, CRY1Bc1, CRY1Bd1, CRY1Bd2, CRY1Be1, CRY1Ca1, CRY1Ca2, CRY1Ca3, CRY1Ca5, CRY1Ca6, CRY1Ca8, CRY1Ca9, CRY1Da1, CRY1Da1, CRY1Db1, CRY1Db2, CRY1Ea1, CRY1Ea2, CRY1Ea3, CRY1Ea4, CRY1Ea6, CRY1Ea7, CRY1Ea8, CRY1Eb1, CRY1Fa1, CRY1Fa2, CRY1Fb1, CRY1Fb2, CRY1Fb3, CRY1Fb5, CRY1Fb6, CRY1Fb7, CRY1Ga1, CRY1Ga2, CRY1Gb1, CRY1Gb2, CRY1Ja1, CRY1Jb1, CRY1Ka1, CRY9Aa1, CRY9Aa2, CRY9Ca1 e 34 seqüências de genes, cuja toxina ataca insetos da ordem Coleóptera: CRY3Aa1, CRY3Aa2, CRY3Aa3, CRY3Aa4, CRY3Aa5, CRY3Aa6, CRY3Aa7, CRY3Aa8, CRY3Aa10, CRY3Aa12, CRY3Bb1, CRY3Bb2, CRY3Ca1, CRY7Aa1, CRY7Ab1, CRY7Ab2, CRY7Ab3, CRY7Ab4, CRY7Ab5, CRY7Ab6, CRY8Aa1, CRY8Ab1, CRY8Ba1, CRY8Ca1, CRY8Ca2, CRY8Ca3, CRY8Da1, CRY8Db1, CRY8Ea2, CRY14Aa1, CRY18Aa1, CRY18Ba1, CRY18Ca1, CRY43Ba1.

Após obter as seqüências dessas proteínas o alinhamento foi realizado com a ferramenta Clustaw (THOMPSON et al., 1997) versão para sistema operacional Ms Windows 8.0. A ferramenta Clustaw foi executada utilizando como entrada arquivos Fasta com a seqüência das proteínas. As seqüências foram alinhadas por domínios I, II e III.






6.2 ANÁLISE FILOGENÉTICA

Para a análise filogenética utilizou-se o programa MEGA 6, Molecular Evolutionary Análise Genética (TAMURA et al., 2011) é um programa livre, com a finalidade de realização de análise estatística da evolução molecular e também para a construção de árvores filogenéticas. Baseado no método de *Neighbor joining*, uma árvore filogenética foi gerada para cada domínio I, II, III (POLANSKI; KIMMEL, 2007).

6.3 ANÁLISE DOS ALINHAMENTOS DAS SEQUÊNCIAS DE AMINOÁCIDOS

Os arquivos dos alinhamentos foram exportados para o formato Fasta e abertos no processador de texto Ms Word 10.0 produzido pela Microsoft, utilizando dois comandos: Localizar e Substituir para encontrar o símbolo do aminoácido e substituí-lo por uma cor,

agrupando os aminoácidos por grupos e identificando os diferentes grupos por cores específicas, como na legenda:

 Grupo R não polares e alifáticos	 Grupo R não carregados, mas polares
 Grupo R aromáticos	 Grupo R carregados positivamente
 Grupo R carregados negativamente	

Em seguida com as informações obtidas do alinhamento as posições de aminoácidos idênticos e semiconservados, utilizamos no arquivo do Word o mesmo critério do software Clustalw o símbolo (*) para aminoácidos conservados e o símbolo (:) para aminoácidos semiconservados. Os dados das sequências dos alinhamentos foram comparados de acordo com as semelhanças e diferenças que interferem na característica e função da proteína CRY. A primeira análise comparou-se número de Aminoácidos Idênticos e Aminoácidos semiconservados nos Domínio I, II e III das 39 sequências de genes avaliadas. Na segunda análise comparou-se Aminoácidos Idênticos e Aminoácidos semiconservados nos Domínio I, II e III, das 16 sequências de genes cuja toxina ataca insetos da ordem Lepidóptera/Coleóptera (8) e Coleóptera (8); e por último analisou Aminoácidos Idênticos e Aminoácidos semi conservados nos Domínio I, II e III, das 16 sequências de genes cuja toxina ataca insetos da ordem Lepidóptera/Coleóptera (8) e Lepidóptera (23).

6.4 ESTRUTURA SECUNDÁRIA DAS SEQUÊNCIAS DE PROTEÍNAS

Para a obtenção das estruturas secundárias das sequências de proteínas foi utilizado o programa PSIPRED v3.3 (*Predict Secondary Structure*) no sítio bioinf.cs.ucl.ac.uk/psipred disponibilizado pelo University College London - UCL.

Foram escolhidas aleatoriamente 6 sequências de proteínas, duas sequências cuja toxina ataca insetos das ordens Lepidóptera/Coleóptera, 2 para Lepidóptera e 2 para Coleóptera, a análise foi realizada utilizando a sequência total de aminoácidos. Foram analisadas o número de aminoácidos idênticos e semiconservados nas estruturas secundárias alfa (α) hélices, folhas Beta (β) e *loops* nos 3 domínio (I, II e III); o número de estruturas secundárias alfa (α) hélices, folhas Beta (β) e *loops*; e a comparação dos aminoácidos nestas estruturas por domínio (I, II e III) e por ordem de insetos.

7 RESULTADOS E DISCUSSÃO

7.1 ALINHAMENTO E ANÁLISE DAS SEQUÊNCIAS DE AMINOÁCIDOS

Foram alinhadas inicialmente 97 sequências de aminoácidos por domínio (I 227aa, II 228aa e III 156aa). Após a execução dos alinhamentos, fez-se a comparação de aminoácido por aminoácido entre as proteínas e retirou-se sequências que se apresentaram muito distintas das demais e por aparecerem uma única vez no banco de dados (http://www.lifesci.sussex.ac.uk/home/Neil_Crickmore/Bt/), de forma a fornecer a análise restrita à sequência com mais procedência e desta forma termos mais confiabilidade para as análises. Isto resultou em um grupo de 39 sequências de aminoácidos composto por 8 genes (Tabela 2), cuja toxina ataca insetos da ordem Lepidóptera e Coleóptera: 23 genes, cuja toxina ataca insetos da ordem Lepidóptera e 8 genes, cuja toxina ataca insetos da ordem Coleóptera, os dados do alinhamento estão apresentados na integra no Anexos 1, 2 e 3.

Tabela 2 - 39 Sequências de aminoácidos para alinhamento

Ordens de insetos	Proteínas	TOTAL
L/C	CRY1Ia1, CRY1Ia2, CRY1Ia3, CRY1Ia5, CRY1Ia11, CRY1Ia12, CRY1Ia14, CRY1Ib3	8
L	CRY1Ba1, CRY1Ba5, CRY1Bb1, CRY1Bc1, CRY1Ca3, CRY1Ca5, CRY1Ca6, CRY1Ca8, CRY1Ca9, CRY1Ea1, CRY1Ea2, CRY1Ea3, CRY1Ea6, CRY1Ea7, CRY1Fa1, CRY1Fa2, CRY1Fb1, CRY1Fb2, CRY1Fb3, CRY1Fb5, CRY1Fb7, CRY1Gb1, CRY1Gb2	23
C	CRY3Aa1, CRY3Aa2, CRY3Aa3, CRY3Aa4, CRY3Aa5, CRY3Aa6, CRY3Aa7, CRY3Aa12	8

Analisando o alinhamento das sequências dos aminoácidos dos diferentes domínios, apresentadas nos Anexos 1, 2 e 3, verificou-se que para o domínio I em 33 posições os aminoácidos são conservados (idênticos), 35 posições são semiconservados (pertencem ao mesmo grupo) num total de 68 aminoácidos conservados ou semiconservados e em 159 posições são variáveis o que corresponde a 70,04% respectivamente conforme mostrado na Tabela 3. Os aminoácidos presentes em posições conservadas foram analisados quanto ao grupo a que pertencem e os dados estão apresentados na Tabela 4. Verifica-se que os

aminoácidos pertencem aos mesmos grupos, embora em proporções diferentes. Ressalta-se que o grupo de aminoácidos não polares e alifáticos aparece em maior número (13) que representa 39,39% dos aminoácidos conservados idênticos e 28 (80%) semiconservados. O grupo de aminoácidos aromáticos aparece em segundo lugar, com 11 (33,33%) aminoácidos idênticos e 3 (8,5742%) semiconservados para o domínio I, como representado na Figura 5.

O domínio II apresenta o menor número de aminoácidos conservados e semiconservados (39) quando comparado aos domínios I (68) e III (57) como na Tabela 3. Considerando que o domínio II está diretamente envolvido com a ligação dos receptores, a diversidade de aminoácidos nesta região pode explicar a especificidade das toxinas (WU et al., 2007), uma vez que esse é o primeiro passo descrito no modo de ação das proteínas CRY (LUCENA et al 2014). Em contrapartida o domínio I está envolvido na inserção na membrana e na formação dos poros, apresenta maior similaridade de sequência, muito provavelmente em função do papel que desempenha no modo de ação das toxinas.

Verifica-se que, no domínio III, o grupo de aminoácidos não polares e alifáticos aparece em maior número 19 o que representa 54,29% dos aminoácidos conservados idênticos e 12 (54,55%) semiconservados. O grupo de aminoácidos não carregados, mas polares aparece em segundo lugar, com 5 (14,29%) aminoácidos idênticos e 8 (36,36%) semiconservados para o domínio III. Pelo fato do domínio III apresentar alta porcentagem de aminoácidos não polares e alifáticos e esses aminoácidos apresentarem volumosas cadeias carbônicas, portanto, características hidrofóbicas, considera-se que esse domínio está fortemente envolvido com a estabilidade estrutural da proteína (BRAVO, et al., 2005).

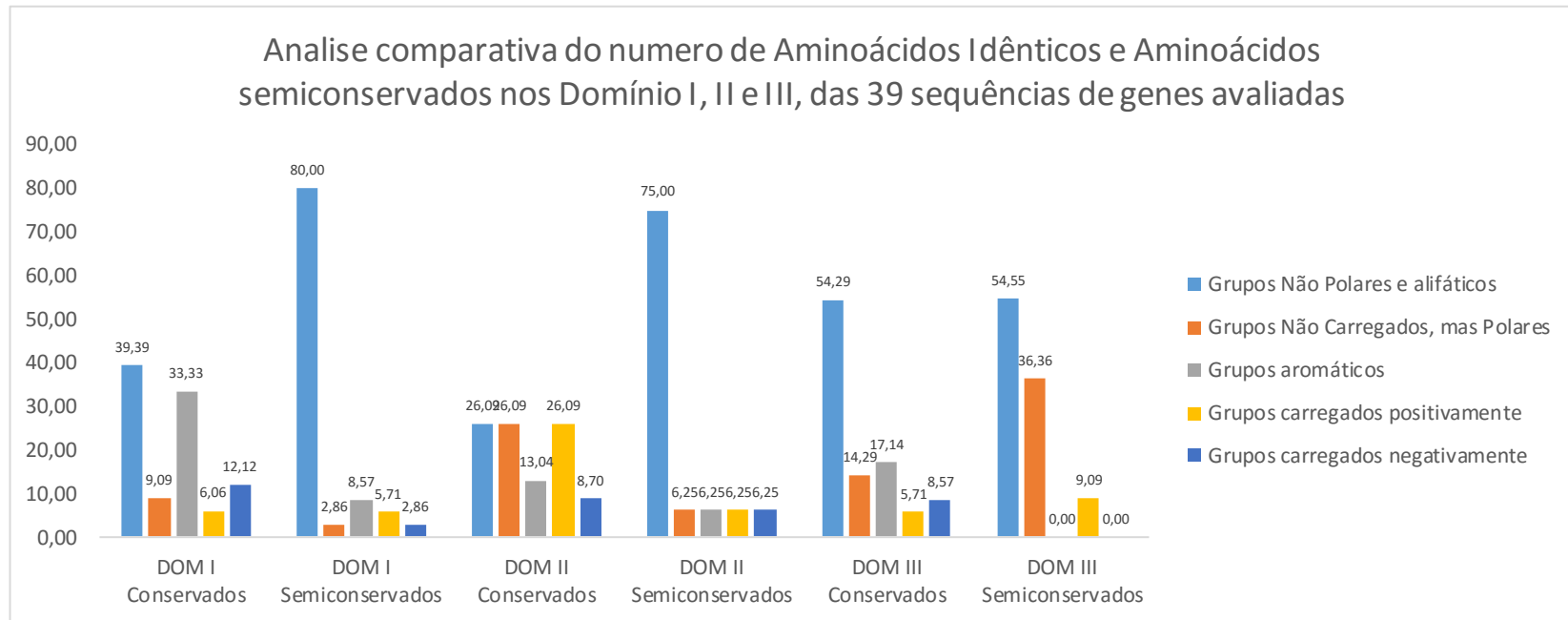
Tabela 3 – Resumo das 39 sequências analisadas quanto à similaridade e variabilidade dos aminoácidos

DOMÍNIO	Nº aa	Conservados	Similaridade %	Variabilidade %
I	227	68	30	70
II	228	39	17	83
III	156	57	36	64

Tabela 4 - Análise comparativa do número de Aminoácidos Idênticos e Aminoácidos semiconservados nos Domínio I, II e III, das 39 sequências de genes avaliadas.

Grupos	Domínio I				Domínio II				Domínio III			
	Posições conservadas		Posições Semiconservadas		Posições conservadas		Posições Semiconservadas		Posições conservadas		Posições Semiconservadas	
	Idênticas				Idênticas				Idênticas			
	Número	%	Número	%	Número	%	Número	%	Número	%	Número	%
Grupos Não Polares e alifáticos	13	39,39	28	80,00	6	26,09	12	75,00	19	54,29	12	54,55
Grupos Não Carregados, mas Polares	3	9,09	1	2,86	6	26,09	1	6,25	5	14,29	8	36,36
Grupos aromáticos	11	33,33	3	8,57	3	13,04	1	6,25	6	17,14	0	0
Grupos carregados positivamente	2	6,06	2	5,71	6	26,09	1	6,25	2	5,71	2	9,09
Grupos carregados negativamente	4	12,12	1	2,86	2	8,70	1	6,25	3	8,57	0	0
TOTAL	33	100,00	35	100,00	23	100,00	16	100,00	35	100,00	22	100,00

Figura 5: Posições conservadas dos aminoácidos quanto ao grupo a que pertencem



Foram comparadas as sequências de aminoácidos das proteínas tóxicas para as ordens Lepidóptera/Coleóptera, com Coleóptera e Lepidóptera/Coleóptera com Lepidóptera com vistas a identificar similaridades entre as mesmas quanto aos grupos de aminoácidos (Figura 6).

Os aminoácidos presentes em posições conservadas foram analisados comparando o grupo da Lepidóptera/Coleóptera com o grupo da Coleóptera e os dados estão apresentados na Tabela 5. Para o domínio I 102 posições que representam 71,33% dos aminoácidos são conservados (idênticos), 41 posições (28,67%) são semiconservados um total de 143 aminoácidos (62,99%), e em 84 posições são variáveis os que correspondem a 37,01%. Ressalta-se que o grupo de aminoácidos não polares e alifáticos aparece em maior número 50 que representa 49,02% dos aminoácidos conservados idênticos e 20 (48,78%) semiconservados. O grupo de aminoácidos aromáticos e não carregados, mas polares aparecem em segundo lugar, com 20 (19,61%) e 15 (14,71%) aminoácidos idênticos e; 4 (9,76%) e 9 (21,94%) semiconservados para o domínio I. Na Figura 7 são mostrados exemplos do alinhamento dos aminoácidos referentes a uma porção da sequência encontrada no Domínio I, II e II. Em A 8 genes cuja toxina ataca insetos das ordens Lepidóptera e Coleóptera foram comparadas com a mesma porção da sequência dos genes cuja toxina ataca somente Coleóptera. Verifica-se que aminoácidos do grupo alifáticos, posições 1, 4, 5 e 9 aminoácidos idênticos, e posições 7 e 8 semiconservadas, onde houve a troca de uma Isoleucina (I) por uma Valina (V) e uma Alanina (A) por Valina (V), respectivamente. Muito embora os aminoácidos tenham grupamentos R de tamanho diferente elas apresentam as mesmas características de hidrofobicidade.

Ainda na Tabela 5, referente ao domínio II verifica-se que em 61 (59,80%) posições os aminoácidos são conservados (idênticos), 41 posições (40,20%) são semiconservados num total de 102 aminoácidos (44,74%), e em 126 posições são variáveis o que corresponde a 55,26%. Novamente o grupo de aminoácidos não polares e alifáticos aparece em maior número 27 que representa 44,26% dos aminoácidos conservados idênticos e 17 (41,46%) semiconservados. O grupo de aminoácidos não carregados, mas polares, aparece em segundo lugar, com 13 (21,31%) aminoácidos idênticos e 16 (39,02%) semiconservados. Observando a Figura 7 (B), onde se apresenta o alinhamento dos aminoácidos referentes a porção para os mesmos genes verifica-se que e aminoácidos do grupo alifáticos, posições 5 e 9 aminoácidos idênticos, e posição 13 semiconservada, onde houve a troca de uma Alanina (A) por uma

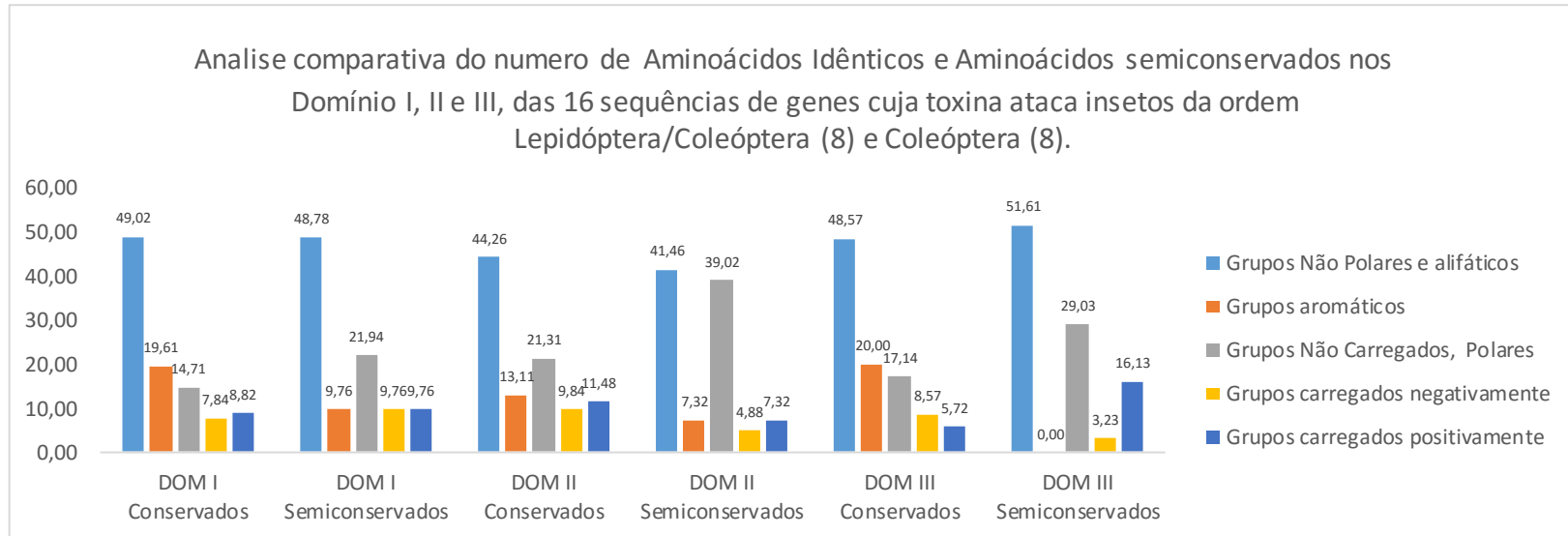
Prolina (P), aumentando as interações dos grupamentos R atribuindo característica mais hidrofóbica à sequência demonstrada.

No domínio III 58,03%, (35 posições), dos aminoácidos são conservados, 31 posições (46,97%) são semiconservados (pertencem ao mesmo grupo) num total de 66 aminoácidos (42,31%), e em 90 posições são variáveis o que corresponde a 57,69%.

Tabela 5 - Análise comparativa do número de Aminoácidos Idênticos e Aminoácidos semiconservados nos Domínio I, II e III, das 16 sequências de genes cuja toxina ataca insetos da ordem Lepidóptera/Coleóptera (8) e Coleóptera (8).

Grupos	Domínio I				Domínio II				Domínio III			
	Posições conservadas idênticas		Posições Semiconservadas		Posições conservadas idênticas		Posições Semiconservadas		Posições conservadas idênticas		Posições Semiconservadas	
	Número	%	Número	%	Número	%	Número	%	Número	%	Número	%
Grupos Não Polares e alifáticos	50	49,02	20	48,78	27	44,26	17	41,46	17	48,57	16	51,61
Grupos aromáticos	20	19,61	4	9,76	8	13,11	3	7,32	7	20,00	0	0,00
Grupos Não Carregados, Polares	15	14,71	9	21,94	13	21,31	16	39,02	6	17,14	9	29,03
Grupos carregados negativamente	8	7,84	4	9,76	6	9,84	2	4,88	3	8,57	1	3,23
Grupos carregados positivamente	9	8,82	4	9,76	7	11,48	3	7,32	2	5,72	5	16,13
TOTAL	102	100,00	41	100,00	61	100,00	41	100,00	35	100,00	31	100,00

Figura 6: Análise comparativa do número de Aminoácidos Idênticos e Aminoácidos semiconservados nos Domínio I, II e III, das 16 sequências de genes cuja toxina ataca insetos da ordem Lepidóptera/Coleóptera (8) e Coleóptera (8).



Como no domínio I e II o grupo de aminoácidos não polares e alifáticos aparecem em maior número 17 que representa 48,57% dos aminoácidos conservados idênticos e 16 (51,66%) semiconservados. O grupo de aminoácidos não carregados, mas polares, aparece em segundo lugar, com 6 (17,14%) aminoácidos idênticos e 9 (29,03%) semiconservados (Tabela 5).

Observou-se que nos domínios I e II (Figura 7 C) os aminoácidos do grupo alifáticos, posições 3 e 8 aminoácidos idênticos, e posição 7 semiconservada, onde houve a troca de uma Arginina (R) por uma Lisina (K), que apresentam estruturas parecidas, ou seja, aminoácidos carregados positivamente, praticamente não interferindo na polaridade do composto. O domínio III apresenta o menor número de aminoácidos conservados e semiconservados (66) quando comparado aos domínios I (143) e II (102).

Figura 7 - Alinhamento aminoácidos idênticos e semiconservados do grupo de genes cuja proteína CRY é tóxica para insetos das ordens Lepidóptera e Coleóptera com o grupo das Coleópteras.

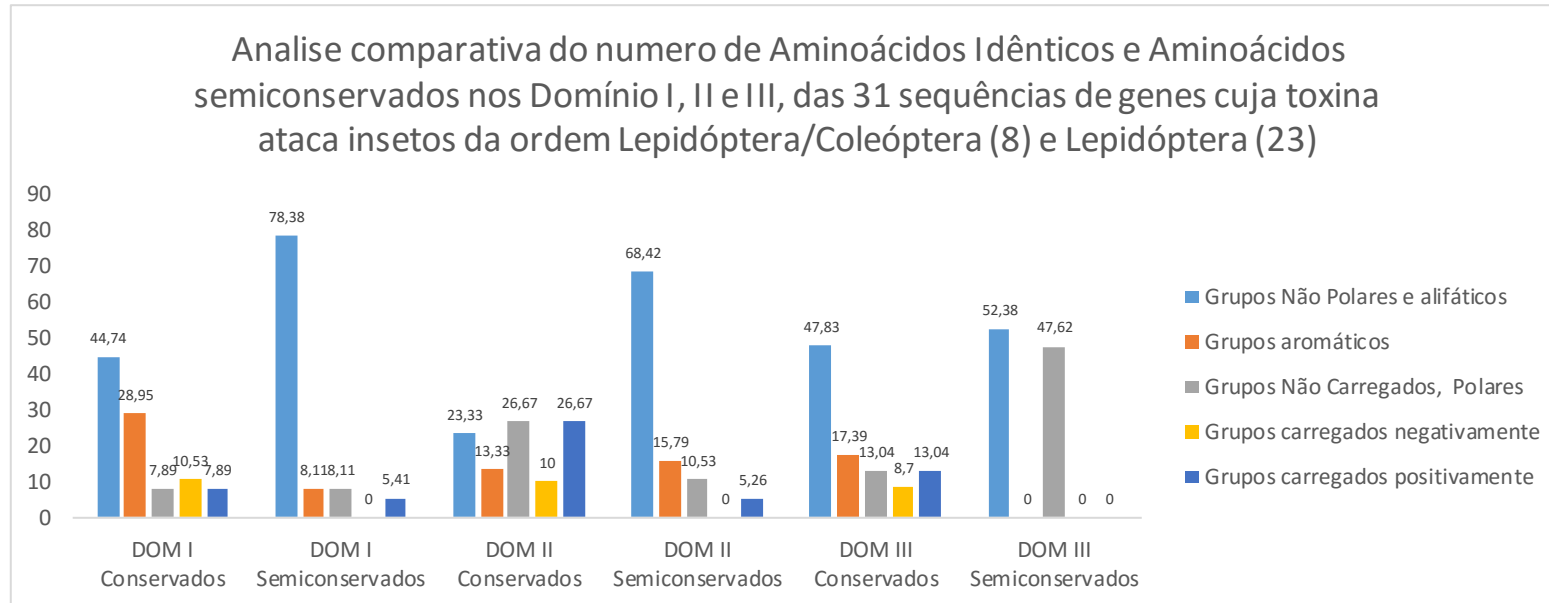
A – DOM I		B – DOM II		C – DOM III	
	1 5 10		5 9 13		3 8 11
1. CRY1Ia1	IQTGIGIAGK	1. CRY1Ia1	TTAQLTREVYTD AJ	1. CRY1Ia1	NTIGPNRITQI
2. CRY1Ia2	IQTGIGIAGK	2. CRY1Ia2	TTAQLTREVYTD AJ	2. CRY1Ia2	NTIGPNRITQI
3. CRY1Ia3	IQTGIGIAGK	3. CRY1Ia3	TTAQLTREVYTD AJ	3. CRY1Ia3	NTIGPNRITQI
4. CRY1Ia5	IQTGIGIAGK	4. CRY1Ia5	TTAQLTREVYTD AJ	4. CRY1Ia5	NTIGPNRITQI
5. CRY1Ia11	IQTGIGIAGK	5. CRY1Ia11	TTAQLTREVYTD AJ	5. CRY1Ia11	NTIDPERINQI
6. CRY1Ia12	IQTGIGIAGK	6. CRY1Ia12	TTAQLTREVYTD AJ	6. CRY1Ia12	NTIDPERINQI
7. CRY1Ia14	IQTGIGIAGK	7. CRY1Ia14	TTAQLTREVYTD AJ	7. CRY1Ia14	NTIDPERINQI
8. CRY1Ib3	IQTGIGIAGK	8. CRY1Ib3	TTSQLTREVYTD AJ	8. CRY1Ib3	NTIDPERINQI
9. CRY3Aa1	IQKGISVVG D:	9. CRY3Aa1	VKTELTRDVLTD PJ	9. CRY3Aa1	NMIDSKKITQL
10. CRY3Aa2	IQKGISVVG D:	10. CRY3Aa2	VKTELTRDVLTD PJ	10. CRY3Aa2	NMIDSKKITQL
11. CRY3Aa3	IQKGISVVG D:	11. CRY3Aa3	VKTELTRDVLTD PJ	11. CRY3Aa3	NMIDSKKITQL
12. CRY3Aa4	IQKGISVVG D:	12. CRY3Aa4	VKTELTRDVLTD PJ	12. CRY3Aa4	NMIDSKKITQL
13. CRY3Aa5	IQKGISVVG D:	13. CRY3Aa5	VKTELTRDVLTD PJ	13. CRY3Aa5	NMIDSKKITQL
14. CRY3Aa6	IQKGISVVG D:	14. CRY3Aa6	VKTELTRDVLTD PJ	14. CRY3Aa6	NMIDSKKITQL
15. CRY3Aa7	IQKGISVVG D:	15. CRY3Aa7	VKTELTRDVLTD PJ	15. CRY3Aa7	NMIDSKKITQL
16. CRY3Aa12	IQKGISVVG D:	16. CRY3Aa12	VKTELTRDVLTD PJ	16. CRY3Aa12	NMIDSKKITQL
	* ** :.*		* . :		* :*.*

A segunda análise foi comparar o grupo dos genes cuja toxina ataca as ordens Lepidóptera/Coleóptera com o grupo das Lepidópteras e os dados estão apresentados na Figura 8. Para o domínio I 38 posições que representam 50,67% dos aminoácidos são conservados (idênticos), 37 posições (49,33%) são semiconservados um total de 75 aminoácidos (33,04%), e em 152 posições são variáveis, o que corresponde a 66,96% (Tabela 6). Observou-se que o grupo de aminoácidos não polares e alifáticos aparece em maior número 17 que representa 44,74% dos aminoácidos conservados idênticos e 29 (78,38%) semiconservados.

Tabela 6 - Análise comparativa do número de Aminoácidos Idênticos e Aminoácidos semiconservados nos Domínio I, II e III, das 31 sequências de genes cuja toxina ataca insetos da ordem Lepidóptera/Coleóptera (8) e Lepidóptera (23).

Grupos	Domínio I				Domínio II				Domínio III			
	Posições conservadas idênticas		Posições Semiconservadas		Posições conservadas idênticas		Posições Semiconservadas		Posições conservadas idênticas		Posições Semiconservadas	
	Número	%	Número	%	Número	%	Número	%	Número	%	Número	%
Grupos Não Polares e alifáticos	17	44,74	29	78,38	7	23,33	13	68,42	22	47,83	11	52,38
Grupos aromáticos	11	28,95	3	8,11	4	13,33	3	15,79	8	17,39	0	0,00
Grupos Não Carregados, mas Polares	3	7,89	3	8,11	8	26,67	2	10,53	6	13,04	10	47,62
Grupos carregados negativamente	4	10,53	0	0,00	3	10,00	0	0,00	4	8,70	0	0,00
Grupos carregados positivamente	3	7,89	2	5,41	8	26,67	1	5,26	6	13,04	0	0,00
TOTAL	38	100,00	37	100,00	30	100,00	19	100,00	46	100,00	21	100,00

Figura 8: Análise comparativa do número de Aminoácidos Idênticos e Aminoácidos semiconservados nos Domínio I, II e III, das 31 sequências de genes cuja toxina ataca insetos da ordem Lepidóptera/Coleóptera (8) e Lepidóptera (23).



O grupo de aminoácidos aromáticos aparece em segundo lugar, com 11 (28,95%) aminoácidos idênticos e 3 (8,11%) semiconservados para o domínio I. Na Figura 9 são mostrados exemplos do alinhamento dos aminoácidos para os Domínios I, II e III, de 8 genes cuja toxina ataca insetos das ordens Lepidóptera e Coleóptera comparadas com 23 sequências do genes cuja toxina ataca somente a ordem Lepidóptera. Verifica-se que aminoácidos do grupo alifáticos, nas posições 5 e 7 na ordem Lepidóptera/Coleóptera todos são Isoleucina(I), enquanto que em algumas sequências nessas mesmas posições na ordem Lepidóptera apareceu também a Leucina (L), como pode ser observada na Figura 9 (A), essa alteração possivelmente não influencia na polaridade do composto, ou seja, na característica da proteína,

No domínio II em 30 posições (61,22%) os aminoácidos são conservados (idênticos), 19 posições (38,78%) são semiconservados, num total de 49 aminoácidos (21,49%), e em 179 posições são variáveis o que corresponde a 78,51%. Novamente o grupo de aminoácidos não polares e alifáticos aparecem em maior número (7) que representa 23,33% dos aminoácidos conservados idênticos e 13 (68,42%) semiconservados. O grupo de aminoácidos não carregados, mas polares, aparece em segundo lugar, com 8 (26,67%) aminoácidos idênticos e 2 (10,53%) semiconservados.

Na comparação do alinhamento no Domínio II, verifica-se que os aminoácidos do grupo alifáticos, na posição 5 são idênticos, em todas as sequências a Leucina (L) ocupa essa posição ainda na posição 9 todos os aminoácidos são Valina (V), enquanto que em algumas sequências nessas mesmas posições na ordem Lepidóptera apareceu também Isoleucina (I) como retrata na Figura 9 (B), essa alteração levou a um aumento do comprimento da cadeia carbônica, aumento das interações hidrofóbicas, e conseqüentemente, a estabilidade da molécula, na sequência analisada.

No domínio III em 46 posições (68,66%) os aminoácidos são conservados, sendo que em 21 posições (31,34%) são semiconservados (pertencem ao mesmo grupo) num total de 67 aminoácidos (42,95%), e em 89 posições são variáveis o que corresponde a 57,05%. Como no domínio I e II o grupo de aminoácidos não polares e alifáticos aparecem em maior número 22 que representa 47,83% dos aminoácidos conservados idênticos e 11 (52,38%) semiconservados. O grupo de aminoácidos não carregados, mas polares, aparece em segundo lugar, com 6 (13,04%) aminoácidos idênticos e 10 (47,62%) semiconservados, como pode ser observado na Figura 9C que evidencia parte da sequência. Nas posições 3, 5 e 8, os aa são

idênticos ocorrendo a Isoleucina (I), Prolina (P) e Isoleucina (I), respectivamente, em todas as sequências avaliadas.

O domínio II apresenta o menor número de aminoácidos conservados e semiconservados (49) quando comparado aos domínios I (75) e III (67) como apresentado da na Tabela 7.

Na Figura 10 são apresentados um resumo das comparações das análises dos 3 domínios para as sequências (LC e C, LC e L) quanto à similaridade e variabilidade.

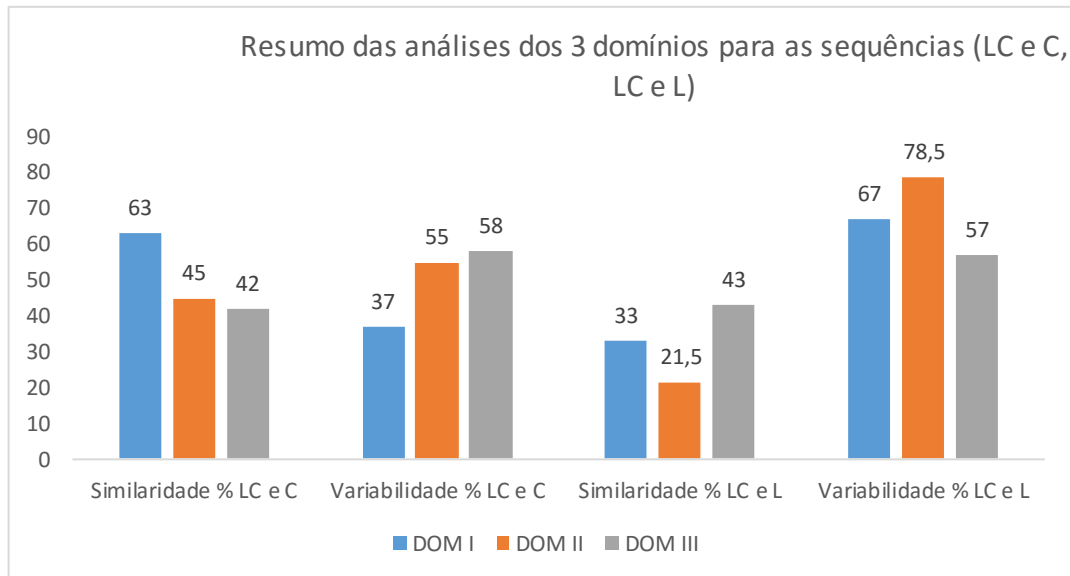
Figura 9 - Alinhamento aminoácidos idênticos e semiconservados do grupo de genes cuja proteína CRY é tóxica para insetos das ordens Lepidóptera e Coleóptera com o grupo das Lepidópteras.

A – DOM I		B – DOM II		C – DOM III	
5 7		5678		3 5 8	
1. CRY1Ia1	IQTGIGIAC	1. CRY1Ia1	TTAQLTREV	1. CRY1Ia1	NTIGPNRIT
2. CRY1Ia2	IQTGIGIAC	2. CRY1Ia2	TTAQLTREV	2. CRY1Ia2	NTIGPNRIT
3. CRY1Ia3	IQTGIGIAC	3. CRY1Ia3	TTAQLTREV	3. CRY1Ia3	NTIGPNRIT
4. CRY1Ia5	IQTGIGIAC	4. CRY1Ia5	TTAQLTREV	4. CRY1Ia5	NTIGPNRIT
5. CRY1Ia11	IQTGIGIAC	5. CRY1Ia11	TTAQLTREV	5. CRY1Ia7	NTIDPERIN
6. CRY1Ia12	IQTGIGIAC	6. CRY1Ia12	TTAQLTREV	6. CRY1Ia11	NTIDPERIN
7. CRY1Ia14	IQTGIGIAC	7. CRY1Ia14	TTAQLTREV	7. CRY1Ia14	NTIDPERIN
8. CRY1Ib3	IQTGIGIAC	8. CRY1Ib3	TTSQLTREV	8. CRY1Ib3	NTIDPERIN
9. CRY1Ba1	VQTGINIAC	9. CRY1Ba1	TSAQLTREV	9. CRY1Ba1	NTIDPERIN
10. CRY1Ba5	VQTGINIAC	10. CRY1Ba5	TSAQLTREV	10. CRY1Ba5	NTINPDII
11. CRY1Bb1	VQTGINIAC	11. CRY1Bb1	TSAQLTREI	11. CRY1Bb1	NTINPDII
12. CRY1Bc1	VQTGINIAC	12. CRY1Bc1	TSAQLTREI	12. CRY1Bc1	NTINPDII
13. CRY1Ca3	IDISLSLV	13. CRY1Ca3	Q---LTREV	13. CRY1Ca3	NTINPDII
14. CRY1Ca5	IDISLSLV	14. CRY1Ca5	Q---LTREV	14. CRY1Ca5	NTINPDII
15. CRY1Ca6	IDISLSLV	15. CRY1Ca6	Q---LTREV	15. CRY1Ca6	NTIDPERIT
16. CRY1Ca8	IDISLSLV	16. CRY1Ca8	Q---LTREV	16. CRY1Ca8	NTIDPERIT
17. CRY1Ca9	IDISLSLV	17. CRY1Ca9	Q---LTREV	17. CRY1Ca9	NIINPNII
18. CRY1Ea1	N-IALEISI	18. CRY1Ea1	TSSQLTREV	18. CRY1Ea1	NIINPNII
19. CRY1Ea2	N-IALEISI	19. CRY1Ea2	TSSQLTREV	19. CRY1Ea2	NIINPNII
20. CRY1Ea3	N-IALEISI	20. CRY1Ea3	TSSQLTREV	20. CRY1Ea3	NIINPNII
21. CRY1Ea6	N-IALEISI	21. CRY1Ea6	TSSQLTREV	21. CRY1Ea6	NIINPNII
22. CRY1Ea7	N-IALEISI	22. CRY1Ea7	TSSQLTREV	22. CRY1Ea7	NTIDPDVI
23. CRY1Fa1	LDISLSLTI	23. CRY1Fa1	TSSQLTREI	23. CRY1Fa1	NTIDPDVI
24. CRY1Fa2	LDISLSLTI	24. CRY1Fa2	TSSQLTREI	24. CRY1Fa2	NTIEPNSIT
25. CRY1Fb1	LDISLSLTI	25. CRY1Fb1	TSSQLTREI	25. CRY1Fb1	NTIEPNSIT
26. CRY1Fb2	LDISLSLTI	26. CRY1Fb2	TSSQLTREI	26. CRY1Fb2	NTIEPNSIT
27. CRY1Fb3	LDISLSLTI	27. CRY1Fb3	TSSQLTREI	27. CRY1Fb3	NTIEPNSIT
28. CRY1Fb5	LDISLSLTI	28. CRY1Fb5	TSSQLTREI	28. CRY1Fb5	NTIEPNSIT
29. CRY1Fb7	LDISLSLTI	29. CRY1Fb7	TSSQLTREI	29. CRY1Fb7	NTIEPNSIT
30. CRY1Gb1	AEIHLKITI	30. CRY1Gb1	TKSQLTREI	30. CRY1Gb1	NTIEPNSIT
31. CRY1Gb2	AEIHLKITI	31. CRY1Gb2	TKSQLTREI	31. CRY1Gb2	NTIEPNSIT
:	:	:	****	:	***

Tabela 7 – Resumo das análises dos 3 domínios para as sequências (LC e C, LC e L) quanto à similaridade e variabilidade

Domínio	Nº aa		Conservados		Similaridade %		Varibilidade %	
	LC e C	LC e L	LC e C	LC e L	LC e C	LC e L	LC e C	LC e L
I	227	227	143	75	63	33	37	67
II	230	228	102	49	45	21,5	55	78,5
III	156	156	66	67	42	43	58	57

Figura 10: Comparação dos aminoácidos nos 3 domínios para as sequências (LC e C, LC e L) quanto à similaridade e variabilidade.



A Tabela 8 apresenta as comparações das sequências por pares após a realização dos alinhamentos múltiplos estruturais. Foram alinhadas 3 sequências de genes cuja toxina ataca insetos das ordens Lepidóptera/Coleóptera (CRY1Ia1), Lepidóptera (CRY1Fa1) e Coleóptera (CRY7Aa1) para posterior análise de sua estrutura secundária. Esta análise indica que quando comparadas as sequências dos domínios individualmente os genes CRY1Fa1 e CRY7Aa1, cuja toxina ataca insetos da ordem Lepidóptera e Coleóptera respectivamente, observou-se que os 3 domínios são menos conservados. Da comparação dos aminoácidos de CRY1Ia1 e CRY1Fa1 com CRY1Ia1 e CRY7Aa1 são iguais no domínio I 63,63% do total de aminoácidos conservados; no domínio II é mais conservado 51,43% para a comparação do gene CRY1Ia1 e CRY7Aa, e no Domínio III a comparação do gene CRY1Ia1 e CRY1Fa1, cuja toxina ataca insetos da ordem Lepidóptera/Coleóptera e Lepidóptera é mais conservado (74,12%). Um trabalho de BOONSERM et al. (2005), que comparou toxinas CRY exibindo atividades específicas contra Lepidóptera (CRY1Aa), Coleóptera (CRY3Aa), Lepidóptera/Díptera (CRY2Aa) e Díptera (CRY4Ba) os resultados mostraram que se compararmos os 3 domínios coletivamente ou separados a CRY1Aa e a CRY3Aa são mais semelhantes; e a CRY4Ba é mais semelhante com a CRY1Aa do que com a CRY3Aa; e a CRY2Aa é a mais distinta das outras. Os resultados da comparação da proteína de Lepidóptera com a proteína de Coleóptera estão de acordo com os resultados apresentados neste trabalho.

Os resultados apresentados indicam que o domínio II é o que apresenta a maior diversidade de sequência, refletido na menor similaridade observada, com valores entre 37 e 51%. Como já descrito o modo de ação das toxinas depende da ligação à receptores e

considerando que diferentes insetos podem ter receptores diferentes espera-se que as proteínas tenham sequências também diferentes no domínio II que é onde ocorre a ligação aos receptores (SCHNEPF, et al, 1998; GOMEZ et al., 2002, 2007).

Nossos resultados corroboram os encontrados por BOONSERM et al. (2005), e podemos inferir que as divergências de sequência no domínio II observadas suportam a hipótese de que esse domínio é determinante a especificidade da toxina. O domínio III foi para as 3 sequências avaliadas o que apresentou maior similaridade de sequência com valores entre 58 e 74% de aminoácidos conservados. Os valores encontrados nesta análise são superiores aos reportados por Grochulsk et al. (1995), que observou uma similaridade de 41% entre as sequências do domínio III referentes aos genes CRY1A(a) e CRY3A cuja toxina é tóxica para insetos das ordens Lepidóptera e Coleóptera respectivamente.

O Domínio I apresentou alta similaridade entre as sequências com valores entre 58 e 63%. A alta similaridade observada neste domínio pode estar relacionada à sua função, pois considerando que esse domínio está envolvido na formação de poros na membrana e que ocorrem apenas após a ligação da proteína aos receptores faz sentido que sejam mais conservadas durante a evolução (CRICKMORE, 2000; WU et al., 2007).

7.2 CONSTRUÇÃO DE ÁRVORES FILOGENÉTICAS

A fim de confirmar se as sequências de proteínas se agruparam de acordo com as suas respectivas ordens foi realizada uma análise filogenética das sequências dos aminoácidos pelos domínios (I, II, III) obtidas através do método *Neighbor-joining*. Árvores filogenéticas foram obtidas através do método de UPGMA e como primeira observação podemos afirmar que as sequências ficaram bem definidas quanto à ordem de insetos.

Na árvore gerada pelas sequências do domínio I observa-se que as sequências foram agrupadas em dois grandes grupos, um grupo contendo 8 sequências de genes cuja proteína é tóxica para insetos das ordens Lepidóptera e Coleóptera; 8 sequências da ordem Coleóptera e 4 sequências da ordem Lepidóptera, evidenciando sua semelhança em nível de sequência primária. Outro grupo formado inclui apenas 19 sequências (82,61%) cuja toxina é específica para insetos da ordem Lepidóptera. Esses dois grandes grupos são compostos por 2 subgrupos que contem sequências da mesma sub classe. Por exemplo grupo 1 apresenta dois subgrupos um que engloba as sequências tóxicas para Lepidópteros/Coleópteros e as específicas para Lepidópteros, enquanto, outro subgrupo engloba apenas as sequências específicas para Coleópteros. Enquanto no grupo 2 encontram-se apenas as sequências específicas para lepidópteros, embora também divididas em dois subgrupos.

As semelhanças e diferenças entre as sequências pode ser a resposta a especificidade das proteínas. Algumas sequências de genes que resultam em proteínas específicas são mais próximas das sequências que produzem proteínas não específicas (Figura 6).

A mesma estrutura de agrupamento se verifica na árvore gerada pelo alinhamento das sequências do domínio II, apresentada na Figura 7. A semelhança na sequência primária do domínio II de quatro genes que codificam proteínas específicas para lepidópteros é maior que as outras 19 que se agruparam num grupo diferente quando comparadas com as sequências que produzem uma proteína não específica para a ordem.

Figura 11 - Filogenia obtida pelo método de *Neighbor Joining*, relativa às 39 sequências referentes ao domínio I: 8 sequências das ordens Lepidóptera/Coleóptera, 23 sequências da ordem Lepidóptera e 8 sequências da ordem Coleóptera. O valor de bootstrap é 1000.

Lc - Lepidóptera/Coleóptera L - Lepidóptera C - Coleóptera

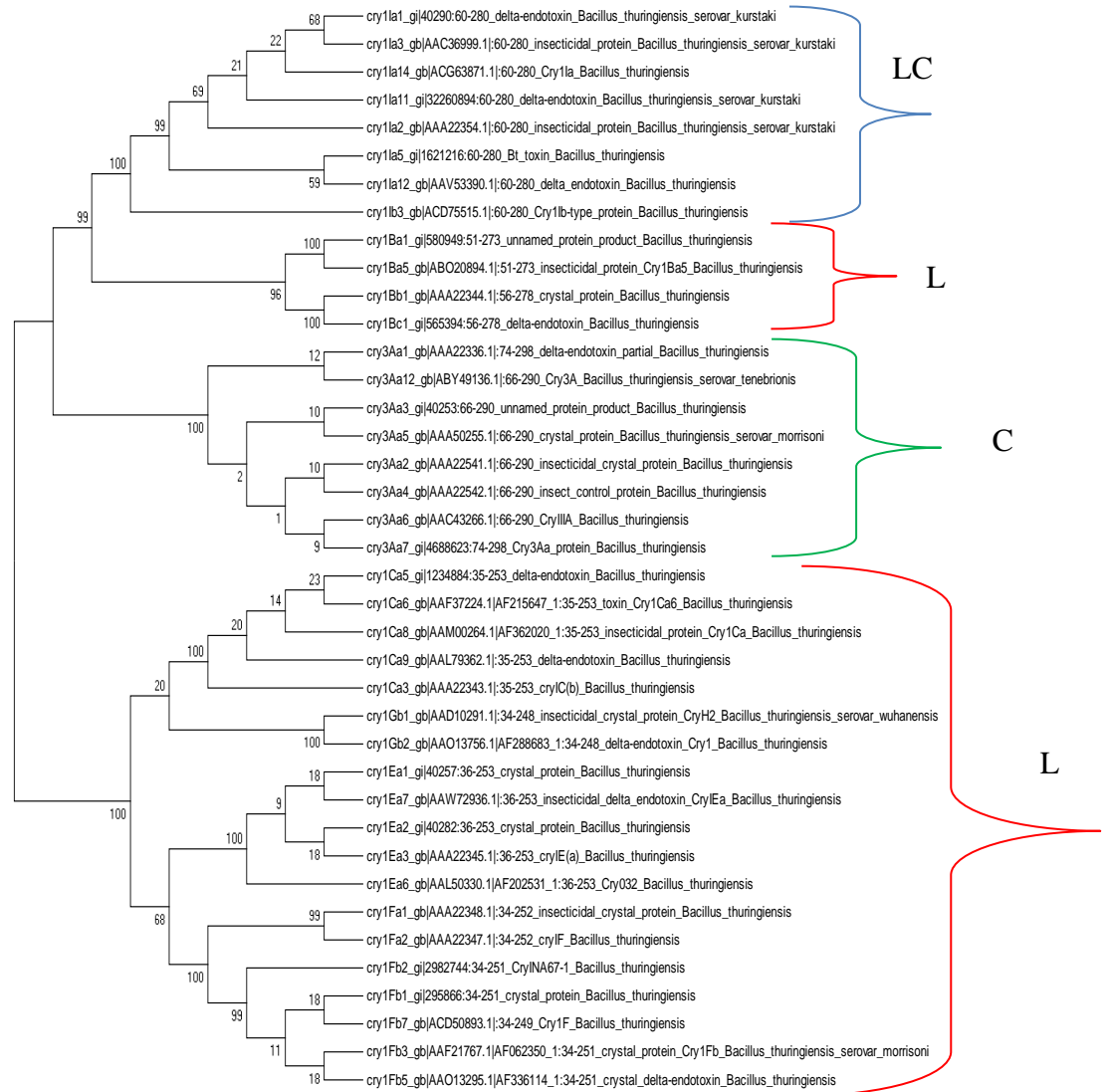
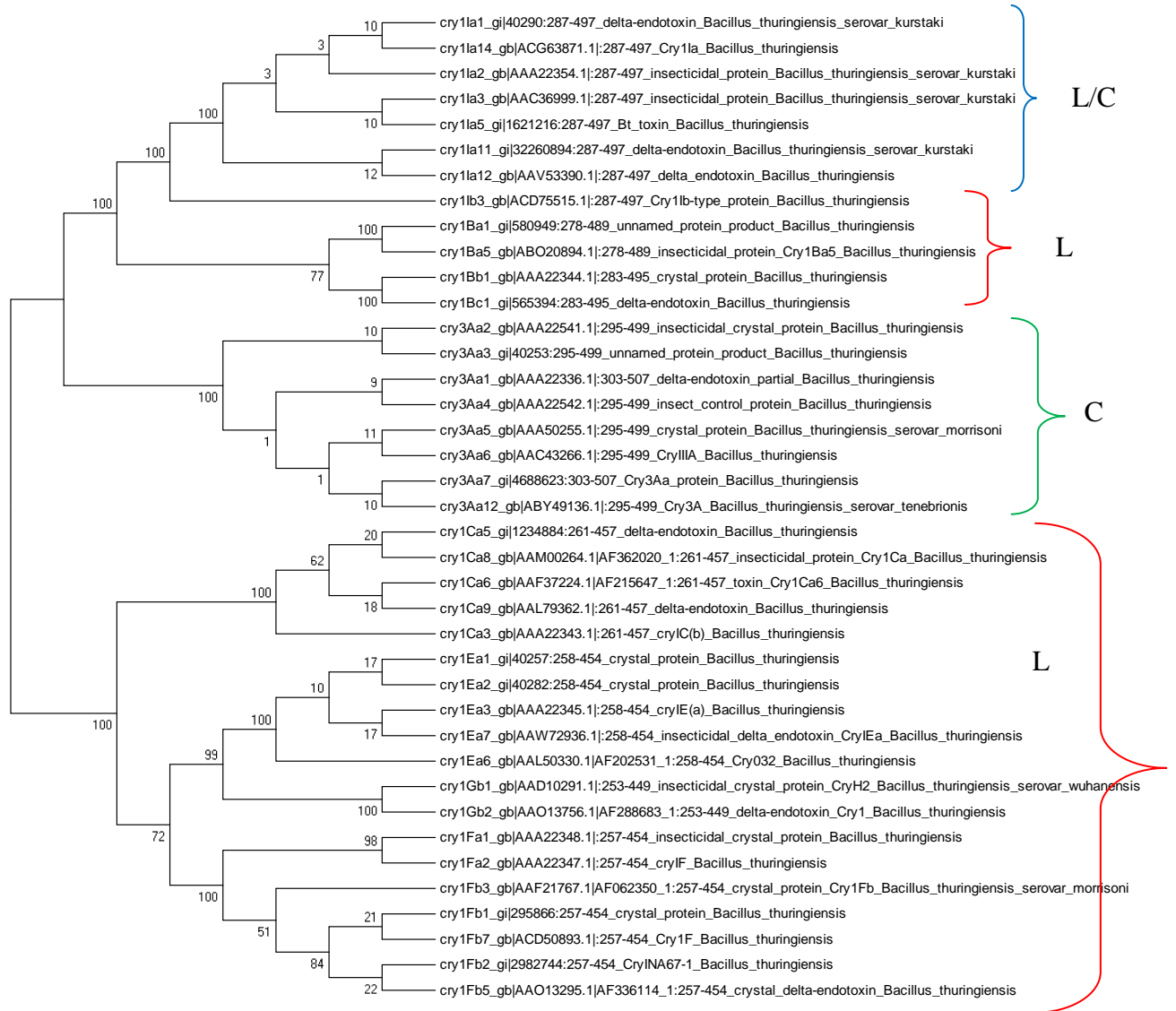


Figura 12- Filogenia obtida pelo método de *Neighbor Joining*, relativa às 39 sequências referentes ao domínio II: 8 sequências das ordens Lepidóptera/Coleóptera, 23 sequências da ordem Lepidóptera e 8 sequências da ordem Coleóptera. O valor de bootstrap é 1000.

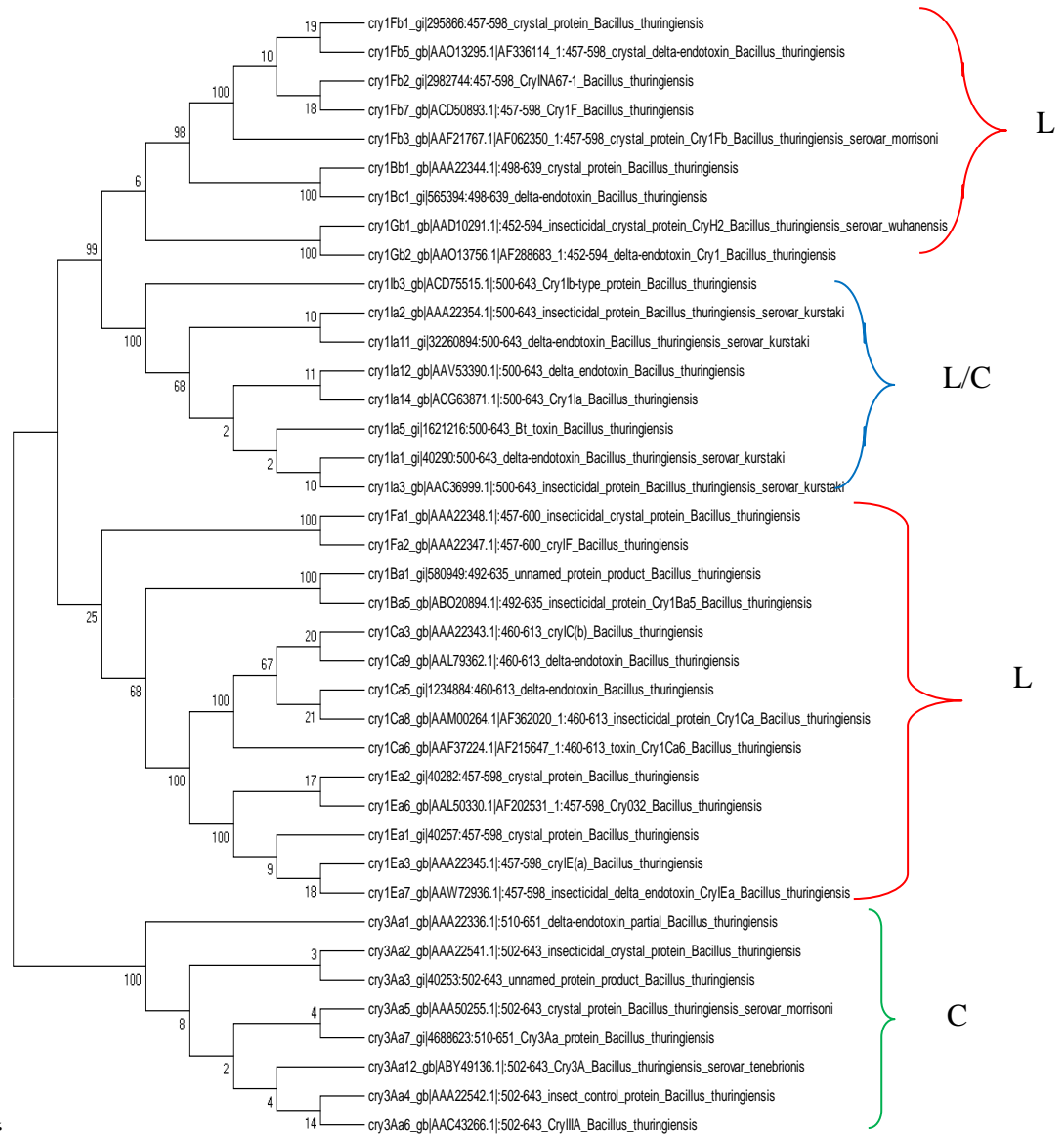
Lc - Lepidóptera/Coleóptera L - Lepidóptera C - Coleóptera



O alinhamento das sequências do domínio III, no entanto, produziram um agrupamento bastante diferente (Figura 8). Dois grupos foram formados um contendo apenas as sequências que codificam toxinas específicas para coleópteros e outro contendo todas as sequências cuja toxina ataca especificamente lepidópteros e lepidópteros/coleópteros, embora as sequências de lepidóptero tenham se subdividido em dois sub grupos. Estranhamente as sequências de coleóptero que se apresentaram semelhantes as sequências de lepidóptero e coleóptero quando do alinhamento I e II e no alinhamento do domínio III agruparam-se separadamente das demais, evidenciando uma grande diferença na sequência primária da proteína.

Figura 13 - Filogenia obtida pelo método de *Neighbor Joining*, relativa às 39 seqüências referentes ao domínio III: 8 seqüências das ordens Lepidóptera/Coleóptera, 23 seqüências da ordem Lepidóptera e 8 seqüências da ordem Coleóptera. O valor de bootstrap é 1000.

Lc - Lepidóptera/Coleóptera L - Lepidóptera C - Coleóptera



7.3 ANÁLISE DAS ESTRUTURAS SECUNDÁRIAS DAS SEQUÊNCIAS DE PROTEÍNAS

A predição foi realizada utilizando o programa PSIPRED, das estruturas secundárias de seis sequências inteiras CRY1Ia1, CRY1Ib3, CRY1Fa1, CRY9Ba1, CRY7Aa1 e CRY3Aa1; duas sequências cujas toxinas são específicas da ordem Lepidóptera e Coleóptera, duas sequências da ordem Lepidóptera e duas sequências apenas para a ordem Coleóptera respectivamente os resultados obtidos estão apresentados nas Figuras 14,15,16,17,18 e 19.

Figura 15 - Predição da estrutura secundária da sequência CRY1Ib3. Representação gráfica da estrutura secundária predita da sequência CRY1Ib3 utilizando o programa PSIPRED que considera os níveis de confiança de predição (Conf).

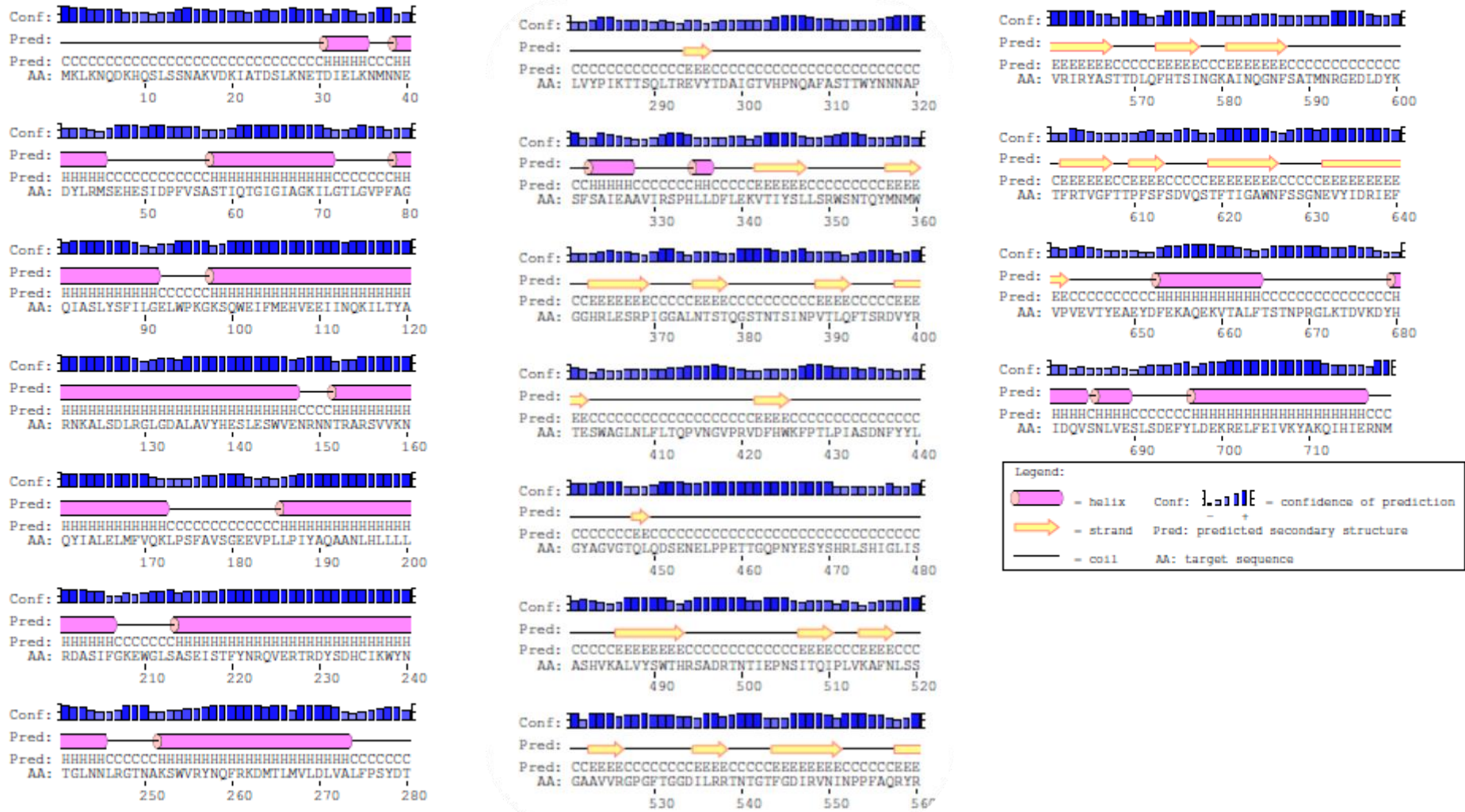


Figura 16 - Predição da estrutura secundária da sequência CRY1Fa1. Representação gráfica da estrutura secundária predita da sequência CRY1Fa1 utilizando o programa PSIPRED que considera os níveis de confiança de predição (Conf).

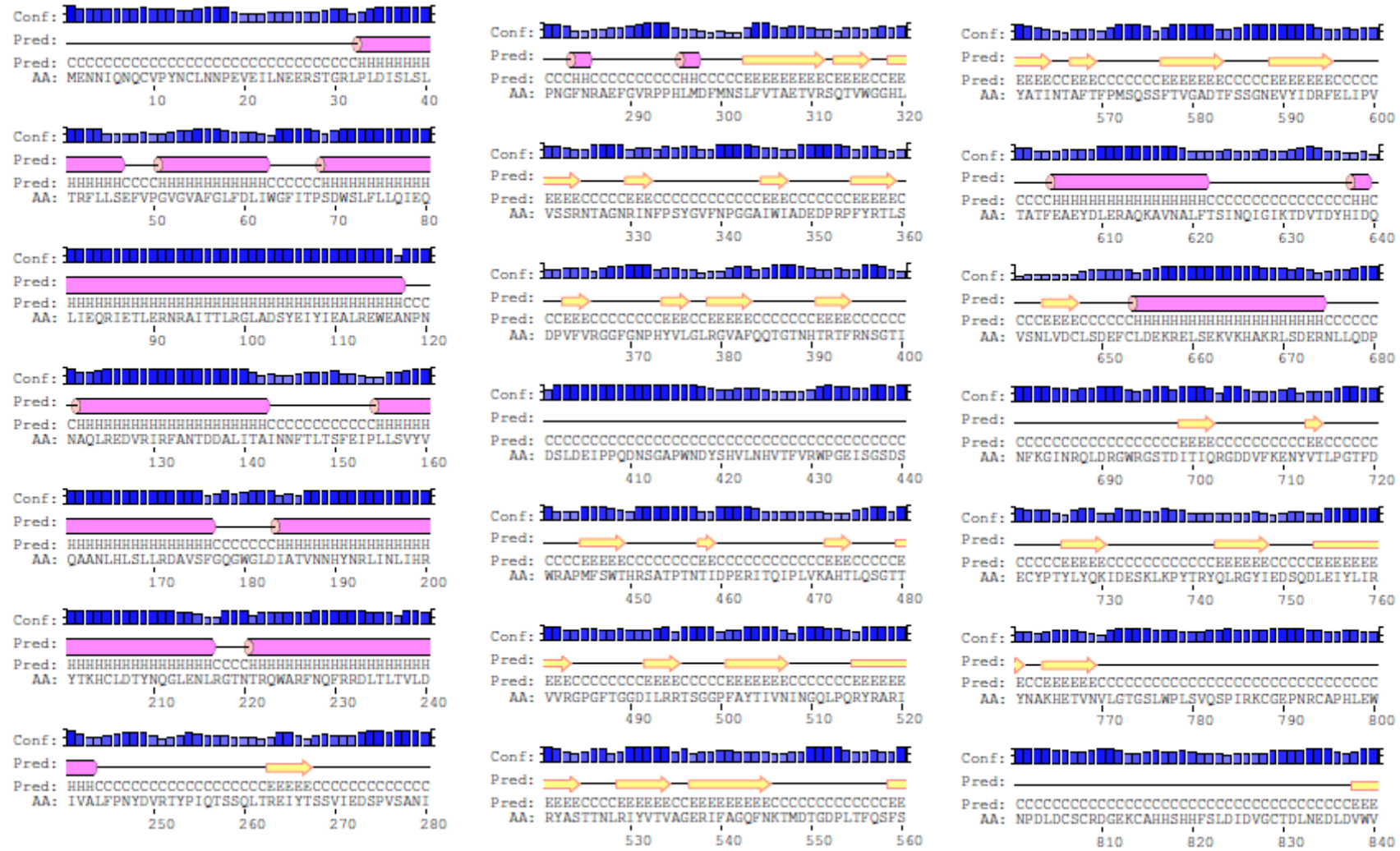


Figura 16 - Continuação

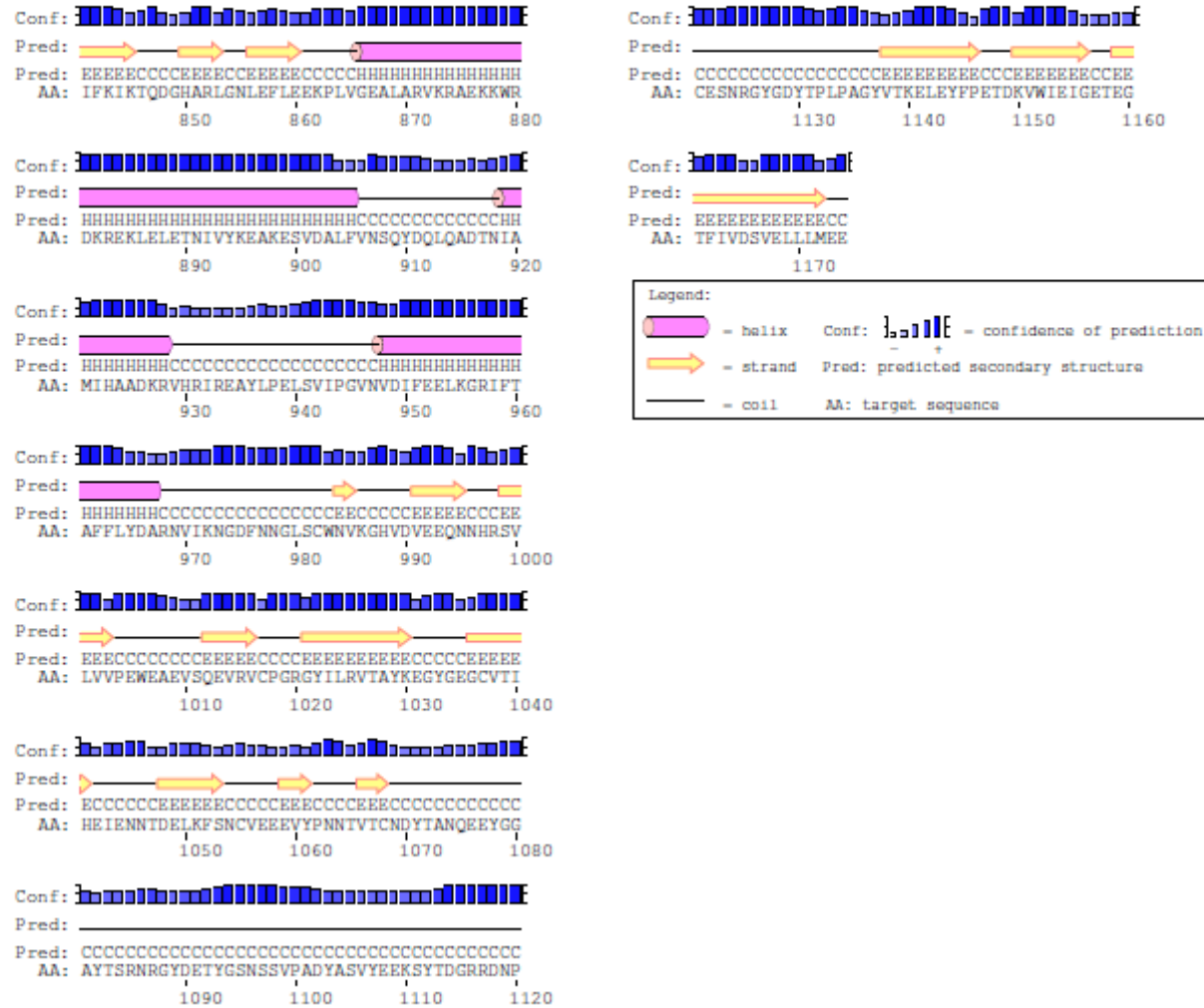


Figura 17 - Predição da estrutura secundária da sequência CRY9Ba1. Representação gráfica da estrutura secundária predita da sequência CRY9Ba1 utilizando o programa PSIPRED que considera os níveis de confiança de predição (Conf).

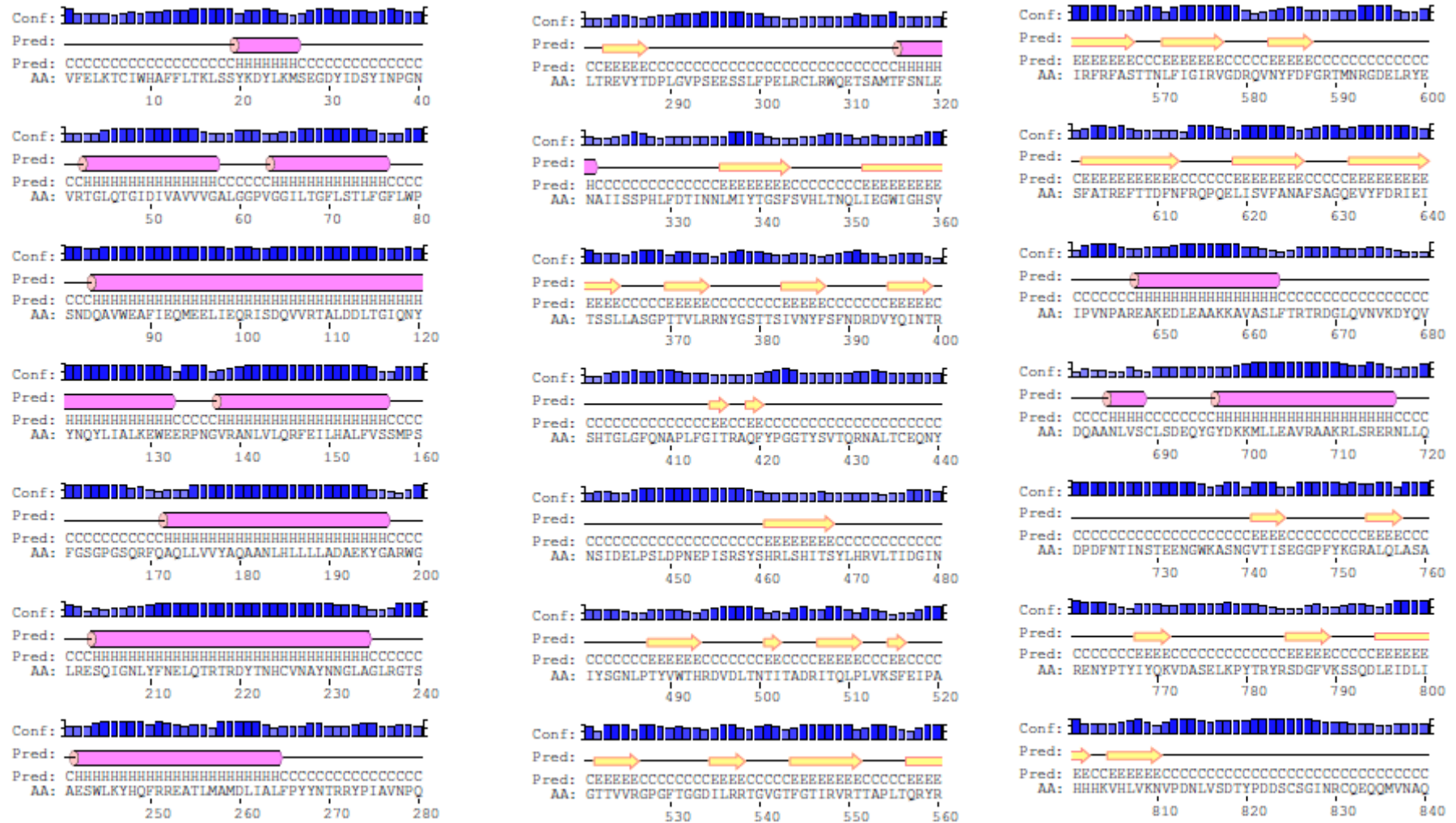


Figura 17 - Continuação



Figura 18 -Predição da estrutura secundária da sequência CRY7Aa1 Representação gráfica da estrutura secundária predita da sequência CRY7Aa1 utilizando o programa PSIPRED que considera os níveis de confiança de predição (Conf).

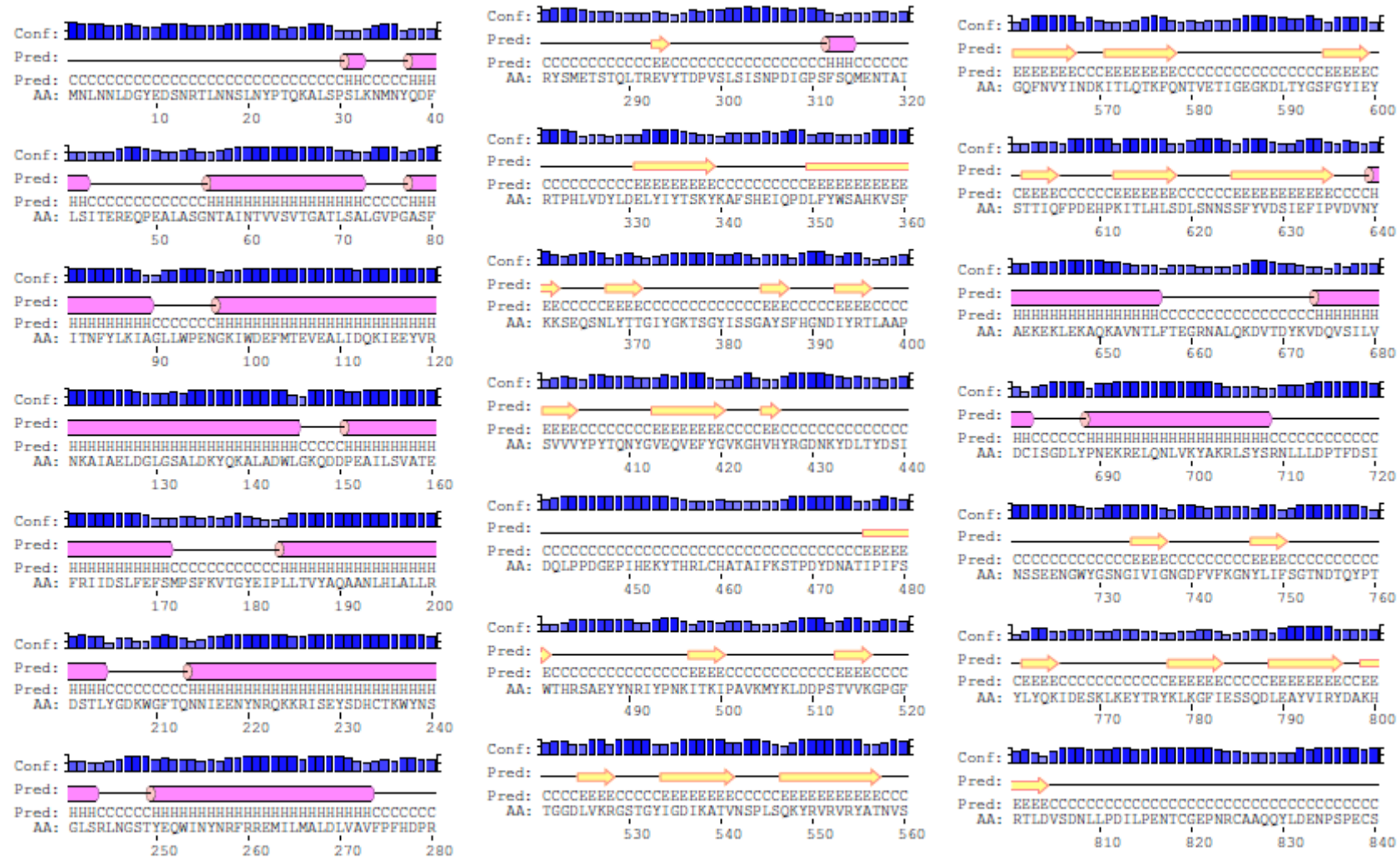


Figura 18 - Continuação

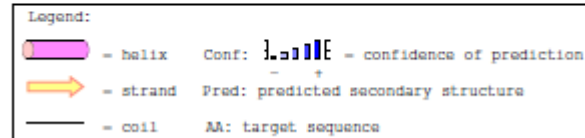
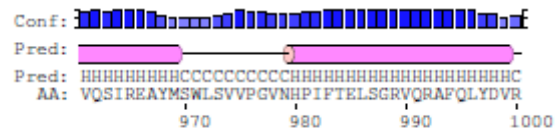
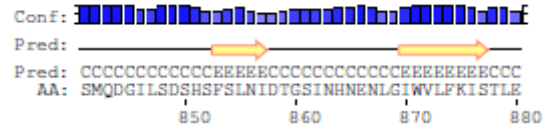
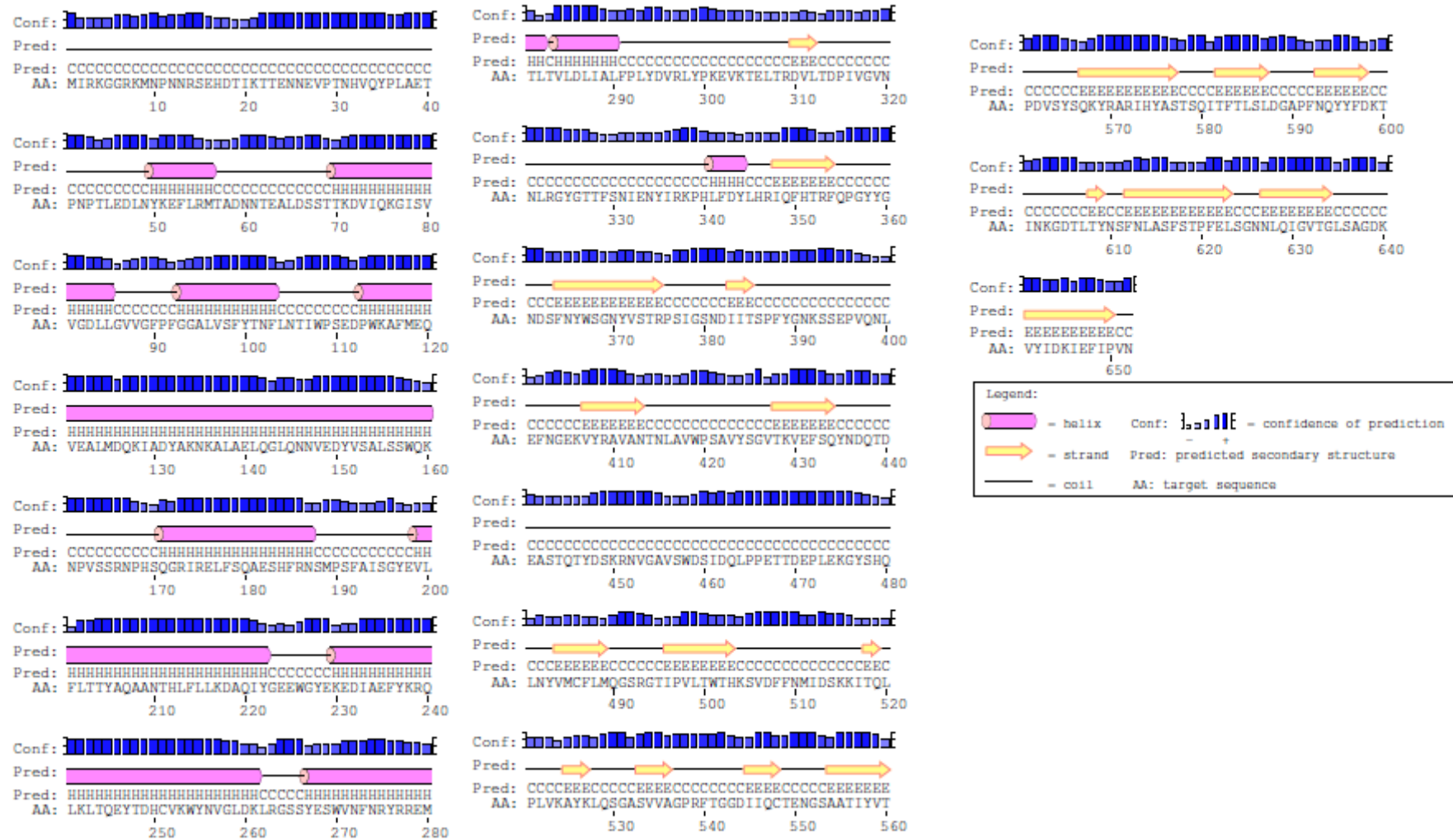


Figura 19 - Predição da estrutura secundária da sequência CRY3Aa1. Representação gráfica da estrutura secundária predita da sequência CRY3Aa1 utilizando o programa PSIPRED que considera os níveis de confiança de predição (Conf).



Um resumo desta predição está apresentado na Tabela 9. Quando se compara o número de alfa hélices encontradas no domínio I verifica-se que a sequência CRY3Aa1 cuja toxina é específica da ordem Coleóptera apresenta 8 alfas hélices enquanto as demais sequências apresentam 7. Percebe-se que na sequência CRY3Aa1 na posição 288 aparece um aminoácido Treonina (T) fazendo com que a alfa hélice se dividisse nas demais sequências nessa posição ocorre uma metionina. Metionina é codificada por ATG e treonina por ACG, é possível que tenha havido uma troca de nucleotídeo C por T na segunda posição, que configuraria uma mutação ou que essa diferença seja apenas um erro de sequenciamento uma vez que a sequência foi retirada do site e não foi realizado novo sequenciamento.

No domínio II, nas sequências cujas toxinas são específicas da ordem Lepidóptera e Coleóptera na CRY1Ia1 e CRY1Ib3 ocorrem 9 e 10 folhas beta (β) e ainda 1 e 2 alfa hélices (α) respectivamente; nas sequências cuja toxinas são específicas da ordem Lepidóptera na CRY1Fa1 e CRY9Ba1 ocorreram 12 e 10 folhas beta (β) e também 2 e 1 alfa hélices (α) respectivamente; e nas sequências cuja toxinas são específicas da ordem Coleóptera na CRY7Aa1 e CRY3Aa1 ocorreram 10 e 8 folhas beta (β) e uma alfa hélice (α) nas duas sequências; no domínio III somente a sequência CRY7Aa1 cuja toxina é específica da ordem coleóptera ocorreu 11 folhas beta e nas outras sequências ocorreram 12 folhas beta (β).

Quando os dados de predição de estrutura da proteína CRY1Fa1 são comparados com a estrutura terciária da proteína CRY1Aa1, obtida por cristalografia (GROCHULSKI et al., 1995), verifica-se que elas apresentam não apresentam o mesmo número de alfa hélices; CRY1Aa1 apresenta 8 alfas hélices (α) no domínio I, e ainda apresentam diferenças no número de estruturas nos domínios II e III.

No domínio II a proteína CRY1Fa1 apresenta 2 alfa hélice e 12 folhas Beta, enquanto que no Domínio III apresenta 12 folhas beta. A proteína CRY1Aa1, por sua vez, apresenta 2 alfas hélice e 11 e 14 folhas beta para os domínios II e III respectivamente. Considerando-se que ambas produzem toxina específica para Lepidóptera, esperava-se que o número de estruturas fosse idêntico em ambas proteínas. No entanto, há que se considerar o método utilizado na obtenção dos dados.

Tabela 9- Estruturas secundárias alfa hélices (α) e folhas beta (β) determinadas para os genes CRY1Ia1, CRY1Ib3, CRY1Fa1, CRY9Ba1, CRY7Aa1 e CRY3Aa1.

Genes	Ordem	Domínio I			Domínio II			Domínio III		
		Nº aa	Alfa	Beta	Nº aa	Alfa	Beta	Nº aa	Alfa	Beta
CRY1Ia1	Lepidóptera/Coleóptera	241	7	-	210	1	9	147	-	12
CRY1Ib3	Lepidóptera/Coleóptera	241	7	-	211	2	10	147	-	12
CRY1Fa1	Lepidóptera	219	7	-	198	2	12	147	-	12
CRY9Ba1	Lepidóptera	228	7	-	221	1	10	146	-	12
CRY7Aa1	Coleóptera	224	7	-	202	1	10	149	-	11
CRY3Aa1	Coleóptera	225	8	-	205	1	8	143	-	12

Após a predição da estrutura secundária foi realizado o alinhamento das sequências de aminoácidos para os seis (6) genes avaliados nos três domínios. Os resultados estão apresentados nas Figuras 20A, B e C. Observou-se nesta comparação que em apenas 3 posições os aminoácidos são conservados e em 16 posições semiconservados; evidenciando a baixa similaridade entre as sequências. Vale considerar que as sequências de aminoácidos das diferentes estruturas também divergem em número.

Duas das 3 posições conservadas e 15 das semiconservadas encontram-se no Domínio I (Figura 20A), sendo este o domínio mais conservado entre os três o mesmo ocorreu quando se avaliou 39 sequências distintas (Tabela 2), considerando a função atribuída a esse domínio, envolvimento na formação de poros na membrana é esperado uma maior similaridade entre as mesmas (WU et al., 2007; LUCENA et al 2014). No entanto, apesar similaridade observada, verifica-se que as sequências apresentam números de aminoácidos diferentes e ainda aminoácidos pertencentes a diferentes grupos. Como exemplo na alfa hélice 2 para as sequências CRY1Fa1 e CRY7Aa1 apresentam 12 aminoácidos; CRY1Ia1, CRY1Ib3 e CRY9Ba1 apresentam um aminoácido a mais uma Glicina (G); e CRY3Aa1 apresenta 11 aminoácidos (Figura 15A).

Também foram observadas trocas de aminoácidos entre as sequências pertencentes a alfa hélice 2. Os genes cuja toxina atua sobre Coleópteros e sobre Lepidópteros/Coleópteros apresentam na oitava posição uma Tirosina (Y), enquanto os genes cuja toxina é específica para lepidópteros apresenta uma Leucina (L) (Figura 20A). A troca de um aminoácido aromático por um alifático não polar manteve a cadeia hidrofóbica, porém mais instável e com maior probabilidade de se ligar a outros aminoácidos (LEHNINGER, NELSON e COX, 2007). Considerando-se as alterações observadas nas alfas hélices 5 e 6, nas posições 4, 10 e 9 e 28, respectivamente, verifica-se quanto ao tipo de aminoácido uma similaridade entre os genes CRY1Ia1 e CRY1Ib3, posto que o mesmo aminoácido está presente nesta posição (Tirosina na posição 4 e histidina na posição 10), enquanto nos genes CRY1Fa1, CRY9Ba1,

CRY7Aa1 e CRY3Aa1 nas mesmas posições apresentam aminoácidos da mesma classe. A presença nesta posição de aminoácidos de classes tão distintas: aromáticos e positivamente carregados podem conferir características diferentes às proteínas que elas codificam, uma vez que os aminoácidos diferem não somente pela polaridade, mas também pelo tamanho de suas moléculas. O mesmo ocorrendo na alfa hélice 6, verifica-se que os aminoácidos pertencem ao mesmo grupo entre os genes CRY7Aa1 e CRY3Aa1 enquanto que para os genes CRY1Ia1, CRY1Ib3, CRY1Fa1 e CRY 9Ba1 apresentam aminoácidos de outro grupo.

Alterações de aminoácidos nesta região também tem sido associada a toxicidade. Tiewsiiri e Angsuhanasombar (2007) promoveram a substituição por Alanina (A) em 4 resíduos aromáticos altamente conservados, W243, F246 e Y249 e F264 presentes na alfa hélice 7 (domínio I) da toxina CRY4B cuja toxina ataca insetos da ordem díptera, o que resultou em uma diminuição na toxicidade contra o mosquito *Stegomyia aegypti*. Nossas análises possibilitaram a identificação das seguintes substituições: na posição 243 os aminoácidos Leucina (L), Alanina (A) e Serina (S), e Serina (S) e Leucina (L); 246 Leucina (L), Prolina (P) e Lisina (K), e Asparagina (N) e Glutamato (E); 249 Treonina (T), Aspartato (D) e Glutamina (Q), e Treonina (T) e Aspartato (D), e 264 Metionina (M), Leucina (L), e Isoleucina (I) e Glutamato (E), respectivamente para às ordens LC, L e C. Sendo que essas posições foram encontradas nas *alfas* 6, 7 e nos *loops* 6 e 7 o que não aconteceu no trabalho de Tiewsiiri e Angsuhanasombar (2007). Observou-se também que a Alanina (A) somente ocorreu na posição 243 da ordem Lepidóptera de apenas uma sequência (CRY1Fa1).

As sequências dos domínios II e III são as mais divergentes como já havíamos determinado no alinhamento das sequências (Tabela 3). No Domínio II apenas um aminoácido é idêntico e nenhum semiconservado entre as diferentes estruturas em folhas beta (Figura 20B). Considerando a função atribuída a esse domínio, ligação a receptores, a divergência de sequência pode refletir a diversidade de receptores nas diferentes presentes nas diferentes ordens de insetos (SCHNEPF, et al, 1998; GOMEZ et al., 2002, 2007).

Nas folhas beta 2, 3 e 4 (Figura 20B) do domínio II foram encontrados na terceira posição somente aminoácidos do grupo Não Polares e Alifáticos para os genes cujas toxinas específicas para lepidópteros/coleópteros e lepidópteros e nesta mesma posição nas folhas beta 2 e 3 houve a troca por aminoácidos aromáticos portanto ocorreu o aumento na cadeia composta, tornando-se mais estável e mais apolar e na folha beta 4 por aminoácidos não carregados, mas polares portanto tornou-se mais hidrofílica, mais solúvel em água, pois apresentam grupos funcionais que formam pontes de hidrogênio com a água.

No Domínio III apenas um aminoácido é semiconservado e nenhum idêntico entre as diferentes estruturas em folhas beta. Na folha beta 1 (Figura 20C) do domínio III foi encontrado na segunda posição somente aminoácidos do grupo Não Carregados, mas polares para os genes cuja toxinas específicas para lepidópteros/coleópteros e coleópteros e nesta mesma posição houve a troca por aminoácidos Alifáticos, não Polares para os genes toxinas específicas para lepidópteros atribuindo características mais hidrofóbicas e o aumento na estabilização das estruturas das proteínas pela promoção de interações hidrofóbicas no seu interior e na folha beta 6 (Figura 20C) encontrado na quinta e sétima posições somente aminoácidos do grupo Carregados Positivamente, para os genes cuja toxinas específicas para lepidópteros/coleópteros e lepidópteros e nestas mesmas posições houve a troca por aminoácidos Alifáticos, não Polares para os genes toxinas específicas para coleópteros, tornando se mais hidrofóbica sendo atraída por moléculas sem carga. Considera-se que esse domínio está fortemente envolvido com a estabilidade estrutural da proteína (BRAVO, et al., 2005).

Figura 20 - Identificação da estrutura secundária por domínio

B - Domínio II

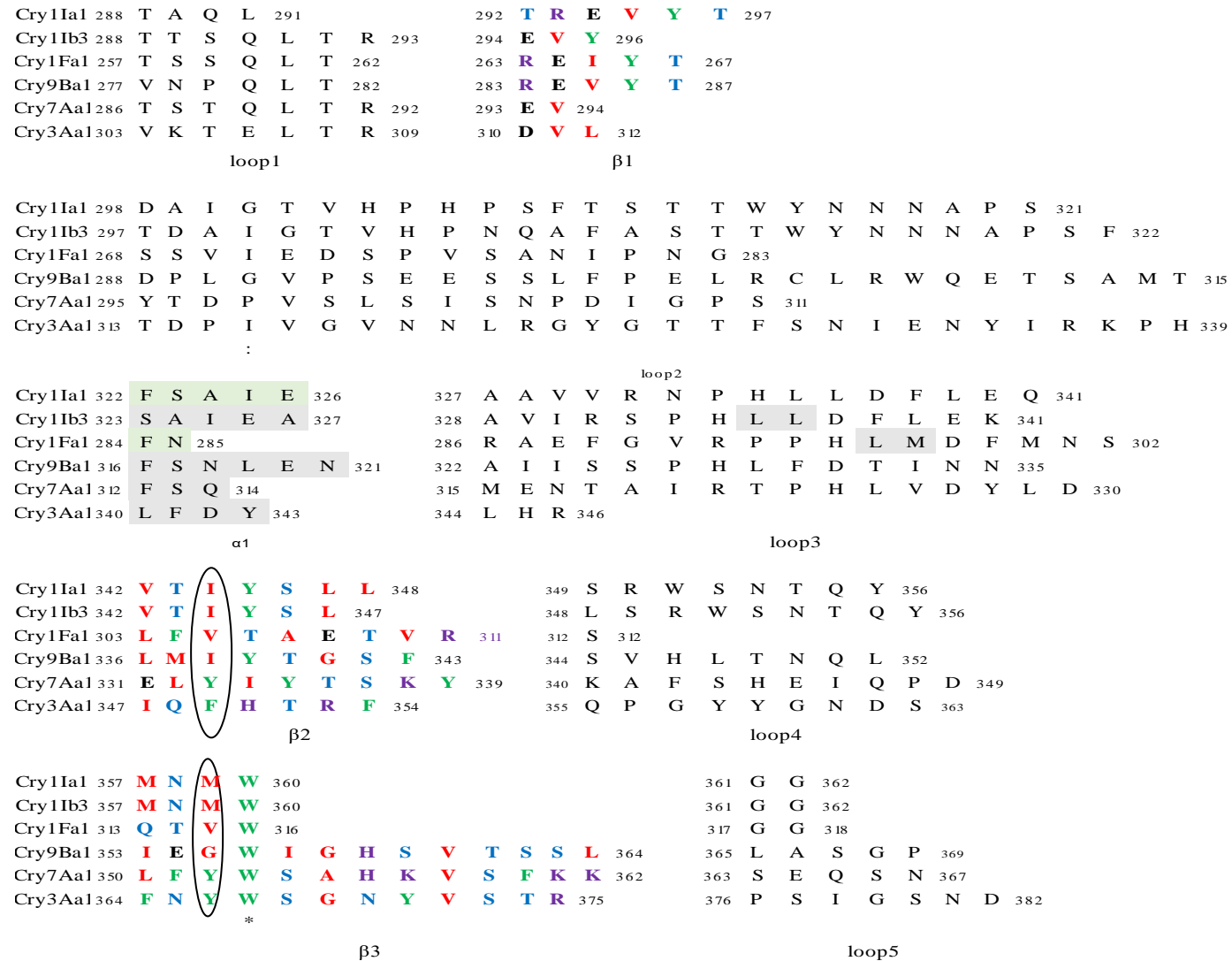


Figura 20B – continuação

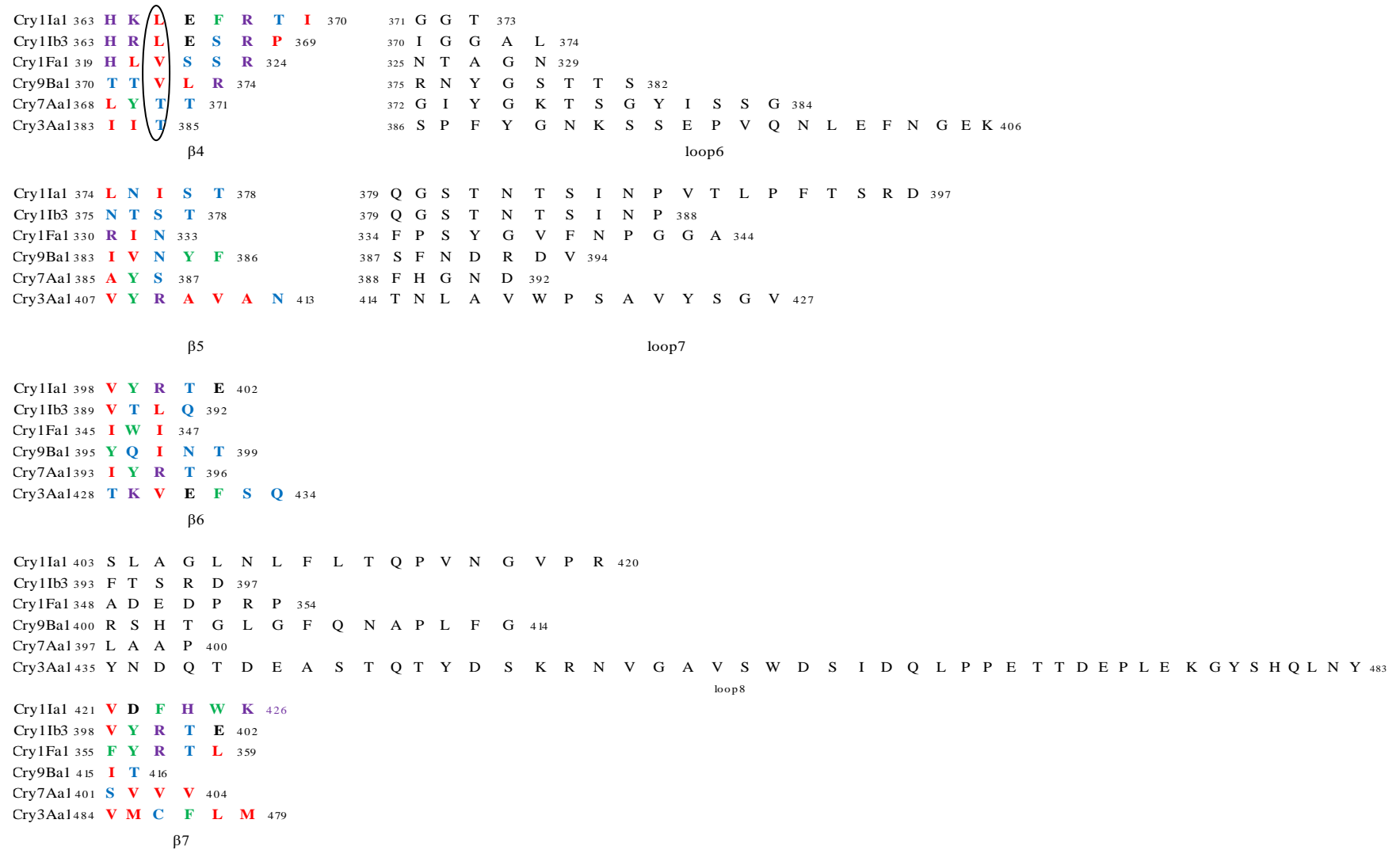


Figura 20B – continuação

Cry1Ia1 427 F V T H P I A S D N F Y Y P G Y A G I G T Q L Q D S E N E L P P E A T G Q P N Y E S Y S H R L S 474
 Cry1Ib3 403 S W A G L N L F L T Q P V N G V P R V 421
 Cry1Fa1 360 S D P 362
 Cry9Ba1 417 R A 418
 Cry7Aa1 405 Y P Y T Q N Y G 412
 Cry3Aa1 490 Q G S R G T 495

loop9

Cry1Ia1 478 **H I G L I** 479
 Cry1Ib3 422 **D F H W** 425
 Cry1Fa1 363 **V F V** 365
 Cry9Ba1 419 **Q F** 420
 Cry7Aa1 413 **V E Q V E F Y G** 420
 Cry3Aa1 496 **I P V L T W T H** 503

 β 8

Cry1Ia1 480 S A S H V 484
 Cry1Ib3 426 K F P T L P I A S D N F Y Y L G Y A G V G T 447
 Cry1Fa1 366 R G G F G N P H 373
 Cry9Ba1 421 Y P G G T Y S V T Q R N A L T C E Q N Y N S I D E L P S L D P N E P I S R S Y S 460
 Cry7Aa1 421 V K G H 424
 Cry3Aa1 504 K S V D F F 507

loop10

Cry1Ia1 485 **K A L V Y S W T** 492
 Cry1Ib3 448 **Q L** 449
 Cry1Fa1 374 **Y V L** 376
 Cry9Ba1 461 **H R L S H I T S** 468
 Cry7Aa1 425 **V H** 426
 Cry3Aa1

 β 9

Cry1Ia1 493 H R S A D 497
 Cry1Ib3 450 Q D S E N E L P P E T T G Q P N Y E S Y S H R L S H I G L I S A S H V K #
 Cry1Fa1 377 **G L** 378
 Cry9Ba1 469 Y L H R V L T I D G I N I Y S G N L P 487
 Cry7Aa1 427 Y R G D N K Y D L T Y D S I D Q L P P D G E P I H E K Y T H R L C H A T A I F K S T P D Y D N A T 475
 Cry3Aa1

loop11

Mohamed et al. (2010), comparando as sequências dos genes CRY1Ab, CRY3Aa e CRY4Ba, cuja toxina ataca as ordens Lepidóptera, Coleóptera e Díptera respectivamente, na região referente ao final domínio II e começo do domínio III indica divergências nas regiões que participam das estruturas $\beta 10$ e $\beta 11$, assim como, na região entre elas (*loop 3*). O tamanho das sequências referentes às $\beta 10$ e $\beta 11$, assim como, do *loop* entre elas diferem bastante entre a sequência do gene cuja toxina ataca lepidópteros CRY1Ab e CRY3A que ataca coleópteros, sendo que em CRY1Ab as folhas beta são menores quando comparadas com a sequência do gene cuja toxina ataca coleópteros. Nossos resultados em relação ao final do domínio II e começo do domínio III entre os 6 genes analisados, corroboram com os resultados reportados por Mohamed et al. (2010) a sequência dos genes CRY1Fa1 e CRY7Aa1 por exemplo referentes a $\beta 10$ e $\beta 11$ não tem o mesmo número de aminoácidos, e diferem no tamanho do *loop* entre elas. Segundo os autores essas divergências podem estar intimamente relacionadas a especificidade da toxina, pois esta região muito provavelmente é um sítio de ligação de receptor. O que foi confirmado por Griko et al, (2007) que comprovou em experimentos de imunoblotting a ligação da porção N terminal desta sequência (“NSSVSIIRAPMFSWIHR”) com o receptor BT-R1. Analisando a presença destes aminoácidos em nossas sequências verificamos que eles estão presentes nos 6 genes e são bem semelhantes; ocorrendo no início das 6 sequências sendo que em 3 posições ocorrem aminoácidos idênticos e em 4 posições semiconservados. Nossa análise também possibilitou a identificação de outra porção que embora altamente conservada (13 em 19 aa são conservados e ou semiconservados) no final apresentando divergências que podem estar relacionadas a especificidade das sequências (Figura 21).

As divergências encontradas nesta análise que podem estar ligadas a especificidade da sequência foram: na décima quinta posição das sequências comparadas ocorre uma Treonina (T) que pertence ao grupo de aminoácidos não carregados, mas polares nas sequências cuja toxina ataca lepidópteros/coleópteros e lepidópteros enquanto que na ordem Coleóptera foi encontrada nesta posição aminoácidos do grupo aromáticos Fenilalanina (F) e Tirosina (Y) tornando essa porção da sequência mais hidrofóbica. Na posição vigésima segunda ocorre uma Serina (S) que pertencem ao grupo de aminoácidos não carregados, mas polares nas sequências cuja toxina ataca lepidópteros/coleópteros enquanto que nas outras sequências cuja toxina ataca lepidópteros e coleópteros foi encontrada nesta posição aminoácidos do grupo Carregados positivamente Arginina (R) e Lisina (K), tornando essa porção da sequência mais hidrofílica, com característica eletropositivas, sendo atraídas por moléculas eletronegativas. Outra diferença encontrada refere-se a trigésima nona posição onde se verifica a presença de

uma (A) Alanina do grupo Alifáticos, não polares nas sequências cuja toxina que ataca lepidópteros/coleópteros e uma Treonina (T) e Serina (S) do grupo Não Carregados, mas polares na mesma posição nas sequência do gene cuja toxina ataca lepidópteros e coleópteros, tornando essa porção da sequência mais hidrofílica, com característica eletropositivas, sendo atraídas por moléculas eletronegativas.

Figura 21 Análise da Estrutura Secundária da região de $\beta 10$ e $\beta 11$ das 6 sequências

```

Cry1Ia1 485 K A L V Y S W T H R S A D R T N T I E P N S I T Q I P L V K A F N L S S G A A V V R G P G F T G G D 534
Cry1Ib3 486 K A L V Y S W T H R S A D R T N T I E P N S I T Q I P L V K A F N L S S G A A V V R G P G F T G G D 535
Cry1Fa1 442 R A P M F S W T H R S A T P T N T I D P E R I T Q I P L V K A H T L Q S G T T V V R G P G F T G G D 491
Cry9Ba1 485 N L P T Y V W T H R D V D L T N T I T A D R I T Q L P L V K S F E I P A G T T V V R G P G F T G G D 534
Cry7Aa1 476 I I P I F S W T H R S A E Y Y N R I Y P N K I T K I P A V K M Y K L D D P S T V V K G P G F T G G D 525
Cry3Aa1 495 T I P V L T W T H K S V D F N M I D S K K I T Q L P L V K A Y K L Q S G A S V V A G P R F T G G D 544
      : : * * * : : * * * : * : * * : : * * * * * * * *

```

Experimentos de metagênese sítio dirigida realizados na toxina codificada por CRY1A demonstraram que o *loop* $\alpha 8$ localizado na junção dos domínios I e II e os *loops* 2 e 3 do domínio II estão envolvidos no reconhecimento e na toxicidade da proteína (LEE, 2006). Assim considerando-se que a ligação dos receptores acontece no *loop* e alterações desses aminoácidos sejam em número ou tipo pode contribuir para a diferença na especificidade.

Na análise feita das 6 sequências, verifica-se que no *loop* 2 do domínio II do gene CRY1Ia1 que atua sobre os insetos das ordens Lepidóptera e Coleóptera existem 24 aminoácidos ²⁰⁸DAIGTVHPHPSFTSTTWYNNAPS³²¹; (Figura 20B) enquanto na mesma região do gene CRY1Fa1 específico para Lepidóptera, existem 16 aminoácidos ²⁶⁸SSVIEDSPVSANIPNG²⁸³. Essa diferença em número de aminoácidos indica que os receptores que provavelmente podem vir a se ligar nesta região devem ser diferentes.

Essa diferença no tamanho dos *loops* do domínio II também foi relatada em relação a proteína CRY4Ba por BOONSERM et al. (2005), os autores relatam que a proteína apresenta *loops* menores no domínio II e que isso afetaria as possíveis interações destes *loops* com os receptores além de afetar a o contato entre os domínios I e II da proteína.

8 CONCLUSÕES

- As análises indicaram maior diversidade de sequência de aa no domínio II.
- As diferenças encontradas na região final domínio II e começo do domínio III referentes às estruturas $\beta 10$ e $\beta 11$, assim como, na região entre elas (*loop 3*) podem ser as responsáveis pela especificidade das proteínas por representarem regiões de ligação à receptor.
- Nessa região, embora apresente 52% de aminoácidos conservados, a substituição da Treonina (não carregados) por Tirosina e Fenilalanina (aromáticos) presente nos genes cuja toxina ataca coleópteros/Lepidópteros e Lepidópteros na posição 490, poderia impedir a ligação da proteína CRY7Aa1 e CRY3Aa1 (Coleóptera) ao receptor.
- Os *loops* participam diretamente da ligação aos receptores e na região referente ao *loop 2* do domínio II foi encontrada uma sequência de aminoácidos divergente em número e tipo de aminoácidos que podem estar relacionadas à especificidade das proteínas uma vez que em CRY1Fa1 específico para Lepidóptera, existem 16 aa, enquanto em CRY1Ia1 que atua sobre os insetos das ordens Lepidóptera e Coleóptera existem 24 aminoácidos.

REFERÊNCIAS BIBLIOGRÁFICAS

- ADANG, M.J.; BRODY, M.S., CARDINEU, G.; EAGAN, N., ROUSH, R.T., SJHEWMAKER, C.K., JONES, A., OAKER, J.V.; MCBRIDE, K.E. The reconstruction and expression of a *Bacillus thuringiensis* CRYIIIa gene in protoplasts and potato plants. **Plant Mol.biol.** 21:2329–2342. 1993.
- AIMANOVA, K.G.; ZHUANG, M.; GILL, S.S. Expression of CRY1Ac cadherin receptors in insect midgut and cell lines. **J. Invertebr. Pathol.** v. 92, p. 178–187, 2006.
- ANGST, B.D.; MARCOZZI, C.; MAGEE, A.I. The cadherin superfamily: diversity in form and function. **J. Cell Sci.** v. 114, p. 629–641, 2001.
- ARENAS, I.; BRAVO, A.; SOBERON, M.; GOMES, I. Role of alkaline phosphatase from *Manduca sexta* in the mechanism of action of *Bacillus thuringiensis* CRY1Ab toxin. **J Biol Chem.** v. 285, p. 12497–503, 2010 [PubMed: 20177063]
- ARONSON, A.I, HAN, E., MCGAUGHEY, W. ; JOHNSON, D. the solubility if the inclusion proteins from *Bacillus thuringiensis* is depend upon protoxin composition and is a factoring toxicity to insects. **Applied and Environmental Microbiology**, 57:981–986. 1991.
- ATSUMI, S. et al. Location of the *Bombyx mori* 175 kDa cadherin-like protein-binding site on *Bacillus thuringiensis* CRY1Aa toxin. **Febs Journal**, v. 275, n. 19, p. 4913-4926, 2008.
- BERLINER, E. 1911. Über die Schlafsucht der Mehlmottenraupe. Z. Gesante Getreidewesen (Berlin). 3: 63- 70,1911.
- BETZ, F. S.; HAMMOND, B. G.; FUCHS, R. L. Safety and advantages of *Bacillus thuringiensis* protected plants to control insect pests. **Regulatory Toxicology and Pharmacology**, v. 32, n. 2, p. 156-173, 2000.
- BISHOP, A.H.; JOHNSON, C.; PERANI, M. the safety of *Bacillus thuringiensis* to mammals investigated by oral and subcutaneous dosage. **World Journal of Microbiology & Biotechnology**, Amsterdam, v.15, p.375-380, 1999.
- BOONSERM, P. et al. J. crystal structure of the mosquito-larvicidal toxin CRY4Ba and its biological implications. **Journal of molecular biology**, London, v. 348, n. 2, p. 363–382, 2005.
- BOONSERM, P.; MO, M; ANGSUTHANASOMBAT, C. H; LESCAR, J.; Structural of the functional form of the mosquito larvicidal CRY4Aa toxin from *Bacillus thuringiensis* at 2.8 resolution. **J. Bacteriol**, v.188, p.33913401, 2006.
- BRAVO, A.; GILI, S.S.; SOBERON, M.; *Bacillus thuringiensis* Mechanisms and Use In: **Comprehensive Molecular Insect Science**. Elsevier BV. v.6, p.175–206, 2005.
- BRAVO, A.; SARABIA, S.; LÓPEZ, L.; ONTIVEROS, H.; ABARCA, C.; ORTIZ, A.; ORTIZ, M.; LINA, L.; VILLALOBOS, F.; PEÑA, NUÑEZ-VALDEZ, G.M.; SOBERON, M.; QUINTERO, R. Characterization of CRY genes in *Bacillus thuringiensis* strain collection. **Appl. Environ.Microbiol.** v. 64, p. 4965-4972, 1998.

BUSO, G.S.C. Marcadores moleculares e análise filogenética. Documentos/Embrapa **Recursos Genéticos e Biotecnologia**, 22p, 2005.

CHEN J, BROWN MR, HUA G, ADANG MJ. Comparison of the localization of *Bacillus thuringiensis* CRY1A delta-endotoxins and their binding proteins in larval midgut of tobacco hornworm, *Manduca sexta*. **Cell Tissue Res.**; 321:123–9. 2005 [PubMed: 15902495]

COSTA, E. L. N.; LUCHO, A. P. R.; FRITZ, L. L.; FIUZA, L. M. Artrópodes e Bactérias Entomopatogênicas. **Biotecnologia, Ciência & Desenvolvimento**. Brasília, v.11, n. 38, p. 04-12, 2010.

CRICKMORE N et al. *Bacillus thuringiensis* toxin nomenclature. http://www.lifesci.sussex.ac.uk/home/Neil_Crickmore/Bt/index.html, 2011.

CRICKMORE, N. et al. Revision of the nomenclature for the *Bacillus thuringiensis* pesticidal Crystal protein. **Microbiology and Molecular Biology Reviews**, v. 62 n (3), p. 807-813, 2008.

CRICKMORE, N.; ZEIGLER, D.R.; FEITELSON, J. SCHNEPF, E.; VAN RIE, J.; LERECLUS, D.; BAUM, J.; DEAN, D.H. **Microbiol. Mol. Biol. Rev.** 62, 807–813 2000.

DE MAAGD, R. A. DE, et al. How *Bacillus thuringiensis* has evolved specific toxins to colonize the insect world. **Trends in Genetics**, v.17, p. 193-199, 2001.

DE MAAGD, R. A.; BOSCH, D.; STIEKEMA, W. *Bacillus thuringiensis* toxin mediated insect resistance in plants. **Trends in Plant Science**, v. 4, n. 1, p. 9-13, Jan 1999.

DORSCH, J.A.; CANDAS, M.; GRIKO, N.B.; MAATY, W.S.; MIDBOE, E.G.; VADLAMUDI, R.K.; BULLA, Jr. L.A. CRY1A toxins of *Bacillus thuringiensis* bind specifically to a region adjacent to the membrane-proximal extracellular domain of BT-R (1) in *Manduca sexta*: involvement of a cadherin in the entomopathogenicity of *Bacillus thuringiensis*. **Insect Biochem. Mol. Biol.** v. 32, p. 1025–1036, 2002.

DUNWELL, J.M. Transgenic crops: the next generation, or an example of 2020 vision. **Annals of Botany**, Kent, v.84, p.269-277, 1999.

ESTRUCH, J. J.; CAROZZI, N.B.; DESAI, N.; DUCK, N.B.; WARREN, G.W.; KOZIEL, M. Transgenic plants: an emerging approach to pest control. **Nature Biotechnology**, n.15, p. 137-141, 1997.

ESTRUCH, J. J.; WARREN, G.W.; MULLINS, M. A.; NYE, G.J.; GRAIG, J. A.; KOZIEL, M.G. Vip 3A, a novel *Bacillus thuringiensis* vegetative insecticidal protein with a wide spectrum of activities against lepidopteran insects. **Proceedings of the National Academy of Sciences of the United States of America**, v.93, p.5389 -5394, 1996

FISCHHOFF, D. A. Insect tolerant transgenic tomato plants. **Bio/Technology**, n.5, p. 807-813, 1987.

FIUZA, L.M. *Bacillus thuringiensis*: Características e o potencial no manejo de insetos. **Acta Biológica Leopoldensia**, v. 23, p. 141-156, 2001.

FLANAGAN, R.D., CAO-GUO, Y., MATHIS, J.P., MEYER, T.E., SHI, X., SIQUEIRA, H.A.A., SIEGFRIED, B.D. Identification, cloning and expression of a CRY1Ab cadherin

receptor from European corn borer, *Ostrinia nubilalis*, (Hübner) (Lepidoptera: Crambidae). *Insect Biochemistry and Molecular Biology*, Oxford, v. 35, p. 3340, 2005.

GAHAN, L.J.; GOULD, F.; HECKEL, D.G. Identification of a gene associated with Bt resistance in *Heliothis virescens*. **Science**, v. 293, p. 857–860, 2001.

GALITSKY, N.; CODY, V.; WOJTCZAK, A.; GHOSH, D.; LUFT, J.R.; PANGBOM, W.; ENGLISH, L. Structure of the insecticidal bacterial endotoxin CRY3Bb1 of *Bacillus thuringiensis*. **Acta CRYstallogr. Sect. D** 57:1101–1109, 2001.

GLARE, T. R.; O'CALLAGHAM, M.; *Bacillus thuringiensis: biology, ecology and safety*. Chichester: John Wiley. p.350, 2000.

GO'MEZ, I.; SA'NCHEZ, J. MIRANDA, R.; BRAVO, A.; SOBERO'N, M. FEBS Lett. v. 513, p. 242–246, 2002.

GO'MEZ, I.; LÓPEZ, L.P. GARAY, C.M.; FERNANDES, L.E.; PÉREZ, C.; SA'NCHEZ, J.; SOBERO'N.; BRAVO, A. Role a receptor interaction in the mode of action of insecticidal Cry and Cyt toxins produced by *Bacillus thuringiensis* 28: 169-173, 2007.

GÓMEZ, I.; ARENAS, I.; BENITEZ, I.; MIRANDA-RIOS, J.; BECEMIL, B.; GRANDE, R.; ALMAGRO, J.C.; BRAVO, A.; SOBERNON, M. Specific epitopes of domains II and III of *Bacillus thuringiensis* CRY1Ab toxin involved in the sequential interaction with cadherin and aminopeptidase-N receptors in *Manduca sexta*. **J. Biol. Chem**, v. 281, p. 34032–34039, 2006.

GROCHULSKI, P.; MASSON, L.; BORISOVA, S.; PUSZTAI-CAREY, M.; SCHWARTZ, J.L.; BROUSSEAU, R.; CYGLER, M. *Bacillus thuringiensis* CRYIA(a) insecticidal toxin: CRYstal structure and channel formation. **J Mol Biol**, v.254, p. 447–464, 1995. [PubMed: 7490762]

HANSEN, B.M. & SALAMITOU, s. Virulence of *Bacillus thuringiensis*. In: Charles, J.F., Delecluse, a. & Nielsen–Le roux, C. (eds). *Entomopathogenic bacteria: from laboratory to field application*. **Netherland. Kluwer academic Publishers**. pp.41–44. 2000.

HOMRICH, M.S., PASSAGLIA, L.M.P., PEREIRA, J.F., BERTAGNOLLI, P.F., SALVADORI, J.R., NICOLAU, M., KALTCHUK-SANTOS, E., ALVES, L.B., BODANESE-ZANETTINI, M.H. 2008. Agronomic performance, chromosomal stability and resistance to velvetbean caterpillar of transgenic soybean expressing cry1Ac gene. **Pesquisa Agropecuária Brasileira**. 43: 801-807.

IANNACONE, R.; GRIECOP, D.; CELLINI, F. Specific sequence modifications of a **CRY3b** endotoxin gene result in high levels of expression and insect resistance. **Plant Molecular Biology**, Dordrecht, v.34, n.3, p.485-496, 1997.

ISHAWATA, S. Investigations on Sotto Bacillus. *Kyoto Sangyo Koshujo Sanji Hokoku*. v.160, p.24-28, 1906.

ISHAWATA, S. On a severe flacherie (sotto disease). *Dainihon Sanshi Kaiho*. v.114, p.15, 1901.

ISHAWATA, S. On Sotto Bacillus. *Dainihon Sanshi Kaiho*. v.160, p.24-28, 1905.

- JAMES, C. Uptake of GM crops in 2005. ISAAA Briefs, ISAAA (**Interactional Service for the Acquisition os Agri-biotech Applications**), Ithaca, no 17, 2005
- JOHNSON, P.E.; JOSHI, M.D.; TOMME, P.; KILBURN, D.G.; McINTOSH, L.P. Structure of the N-terminal cellulose-binding domain of *Cellulomonas fimi* CenC determined by nuclear magnetic resonance spectroscopy. **Biochemistry**, v. 35, p. 14381–14394, 1996.
- JOHNSON, D.E.; MCGAUGHEY, W.H. Contribution of *Bacillus thuringiensis* Spore to Toxicity of Purified Cry Protein Towards Indianmeal Moth Larvae. *Current Microbiology*, v.33, p.54-50, 1996.
- JOUANIN, L.; BONADÉ-BOTTINO, M.; GIRARD, C.; MORROT, G.; GIBAND, M. Transgenic plants for insect resistance. **Plant Science**, n. 131, p.1-11, 1998.
- KNIGHT, P. J.; CRICKMORE, N.; ELLAR, D.J. The receptor for *Bacillus thuringiensis* CRYIA (c) -endotoxin in the brush border membrane of the lepidopteran *Manduca sexta* is aminopeptidase N. **Mol. Microbiol**, v. 11, p.429– 436, 1994.
- KOZIEL, M.G.; BELAND, G.L.; BOWMAN, C.; CAROZZI, N.B. Field performance of elite transgenic maize plants expressing an insecticidal protein derived from *Bacillus thuringiensis*. *Bio/Technology*, n.11, p. 194-200, 1993.
- KRYWUNCZYR, J.; FAST, P.G. Sorological relationships of the CRYstal of *Bacillus thuringiensis* var. israelensis. **Journal of Invertebrate Pathology**, v. 36, p. 139-140, 1980.
- LEHNINGER, A.L.; NELSON, D.L.; COX, M.M. Princípios de Bioquímica. 4ª edição. Editora: Sarvier, 1232p, 2007.
- LI, CARREL, J. J.; ELLAR, D. J. Cristal structure of insecticide delta-endotoxin from *Bacillus thuringiensis* at 2,5 Å resolution. **Nature**, London, v. 353, n. 6.347, p.815-821, 1991.
- LIANG, Y.; PATEL, S.S.; DEAN, D.H. Irreversible binding kinetics of *Bacillus thuringiensis* CRYIA -endotoxins to gypsy moth brush border membrane vesicles is directly correlated to toxicity. **J. Biol. Chem.** v. 270, p. 24719– 24724, 1995.
- LIU, B.; ZENG, Q.; YAN, F.; XU, H.; XU, C. Effects of transgenic plants on soil microorganisms. **Plant and Soil**, v. 271, p. 1-13, 2005.
- LUCENA, W. A.; PELEGRINI, P. B.; MARTINS-DE-SA, D.; FONSECA, F. C. A.; GOMES JÚNIOR, J. E.; MACEDO, L. L. P. de; SILVA, M. C. M. da; OLIVEIRA, R. D.; GROSSI DE SA, M. F. Molecular approaches to improve the insecticidal activity of *Bacillus thuringiensis* cry toxins. **Toxins**, v. 6, p. 2393-2423, 2014.
- MACRAE TC., BAUR ME DJ. FITZPATRICK BJ, GAO AG, GAMUNDI JC, HARRISON LA. KABUYE VT, MCPHERSON RM, MIKLOS JA, PARADISE MS, TOEDEBUSCH AS e VIEGAS A. Laboratory and field evaluations of transgenic soybean exhibiting high-dose expression of a synthetic *Bacillus Thuringiensis* CRYIA gene for control of Lepidoptera. **J Econ Entomol** 98:577-587. 2005.

- MASSON, L.; LU, Y.J.; MAZZA, A.; BROSSEAU, R.; ADANG, M.J. The CRY1A(c) receptor purified from *Manduca sexta* displays multiple specificities. **J Biol Chem**; 270:20309–20315, 1995. [PubMed: 7657602]
- MIRALLES, M.P.; PERES, V.J. Aislamiento y establecimiento de una colección de *Bacillus thuringiensis*. In: BRAVO, A.; CERON, J. (Eds). **Bacillus thuringiensis en el control biológico**. Bogotá, Colombia, p.207-232, 2004.
- MOHAMED Mohamed A. Ibrahima, Natalya Grikoa, Matthew Junkerb & Lee A. Bacillus thuringiensis - A genomics and proteomics perspective. *Bioengineered Bugs*, v. 1, p. 31-50, 2010.
- MONNERAT, R.G. & BRAVO, a. Proteínas bioinseticidas produzidas pela bactéria *Bacillus thuringiensis*: modo de ação e resistência. In: Melo, J. L. (ed). *Controle biológico*. Jaguariúna. **Embrapa Meio ambiente**. 3:163–200. 2000.
- MORIN, S.; BIGGS, R.W.; SISTERTON, M.S.; SHRIVER, L.; ELLERS-KIRK, C.; HIGGINSON, D.; HOLLEY, D.; GAHAN, L.J.; HECKEL, D.G; CARRIERE, Y.; DENNEHY, T.J.; BROWN ; J.K.; TABASHNIK, B.E. Three cadherin alleles associated with resistance to *Bacillus thuringiensis* in pink bollworm. **Proc. Natl. Acad. Sci. USA**, v. 100, p. 5004–5009, 2003.
- MORSE, R.J.; YAMAMOTO, T. STROUD, R.M. Structure of CRY2Aa suggests an unexpected receptor binding epitope. **Structure**, v. 9, p.409–17, 2001. [PubMed: 11377201]
- PACHECO, S. et al. Domain II Loop 3 of *Bacillus thuringiensis* CRY1Ab Toxin Is Involved in a "Ping Pong" Binding Mechanism with *Manduca sexta* Aminopeptidase-N and Cadherin Receptors. **Journal of Biological Chemistry**, v. 284, n. 47, p. 32750-32757, Nov 2009.
- PARDO-LOPEZ, L. et al. Structural changes of the CRY1Ac oligomeric pré-pore from *Bacillus thuringiensis* induced by Nacetylgalactosamine facilitates toxin membrane insertion. **Biochemistry**, v. 45, p. 10329–10336, 2006.
- PARDO-LÓPEZ, L.; SOBERON, M.; BRAVO, A. *Bacillus thuringiensis* insecticidal three-domain CRY toxins: mode of action, insect resistance and consequences for crop protection. **Fems Microbiology Reviews**, v. 37, n. 1, p. 3-22, Jan 2013.
- PARROT, W.A.; ALL, J.N.; ADANG, M.J.; BAILEY, M.A.; BOERMA, H.R.; STEWART, JR.C.N. Recovery and evaluation of soybean plants transgenic of a *Bacillus thuringiensis* var. *kurstaki* insecticidal gene. **In vitro Cellular and Developmental Biology**, v.30, p. 144-149, 1994
- PEFERÖEN, M. Progress and prospects for field use of Bt genes in crops. **Trends in Biotechnology**, Cambridge, v.15, p.173-177. 1997.
- PERLAK, F.J.; DEATON, R.W.; ARMSTRONG, T.A.; FUCHS, R. L.; SIMS,S.R.; GREENPLATE, J.T.; FISCHHOFF, D.A. Insect resistant cotton plants. New York Biotechnology Association, n.8, p. 939- 943, 1990.
- PERLAK, F.J.; STONE, T.B.; MUSKOPF,Y.M; PTERSEN, L.; PARKER, G.B.; McPEHERSON, S.A; WYMAN, J.; LOVE, S.; REED, G.; BIEVER, D.; FISCHHOFF,

- D. Genetically improved potatoes: protection from damage by Colorado potato beetles. **Plant Molecular Biology**, v.22, p.313-321, 1993.
- POLANCZYK, R.; ALVES, S. *Bacillus thuringiensis*: Uma breve revisão. **Agrociência**. vol.VII, n.2, p.1-10, 2003.
- POLANSKI, A.; KIMMEL, M. 2007. Bioinformatics. Springer, Berlim. 1a. Ed.
- PRIEST, F.G. Biological Control of Mosquitoes and other Biting Flies by *Bacillus sphaericus* and *Bacillus thuringiensis*. **Journal of Applied Bacteriology**, v. 72, p. 357-369, 1992.
- RICHARDSON, J.S. The anatomy and taxonomy of protein structure. **Advances in Protein Chemistry**, v. 34, 1981.
- SANGADALA, S.; WALTERS, F.S.; ENGLISH, L.H.; ADANG, M.J. A mixture of *Manduca sexta* aminopeptidase and phosphatase enhances *Bacillus thuringiensis* insecticidal CRYIA(c) toxin binding and 86Rb-K efflux in vitro. **J. Biol. Chem.** v.269, p. 10088–10092, 1994.
- SCHNEPF, E. et al. *Bacillus thuringiensis* and its pesticidal CRYstal proteins. **Microbiology Molecular Biology Review**, v. 62, p. 775–806, 1998.
- SCHNEPF, H.E. & WHITELEY, H.R. Cloning and expression of the *Bacillus thuringiensis* CRYstal protein gene in *Escherichia coli*. **Proc Natl Acad Sci USA**, v. 78, p. 2893–2897, 1981.
- SIEGEL, J. P. The mammalian safety of *Bacillus thuringiensis* Dbased insecticides. **Journal of Invertebrate Pathology**, v.77, p.13-21, 2001.
- TAMURA, K.; PETERSON, D.; PETERSON, N.; STECHER, G.; NEI, M.; KUMAR, S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. **Mol. Biol. Evol.** v. 28, p. 2731-2739, 2011.
- THOMPSON, J.D.; HIGGINS, D.G.; GIBSON, T.J. ClustaW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. **Nucleic Acid Research**, v.22, p.4673-4680, 1997.
- TIEWSIRI, K. ANGSUTHANASOMBAT, C. Structural conserved aromaticity of Tyr249 and Phe264 in helix 7 is important for toxicity of the *Bacillus thuringiensis* Cry4Ba toxin. **J Biochem Mol Biol**, v. 40, p. 163-171, 2007.
- VADLAMUDI, R. K.; WEBER, E.; JI, I.; JI, T.H.; BULLA, Jr, L.A. Cloning and expression of a receptor for an insecticidal toxin of *Bacillus thuringiensis*. **J. Biol. Chem.** v. 270, p. 5490–5494, 1995.
- VAN FRANKENHUYZEN, K. Insecticidal activity of *Bacillus thuringiensis* CRYstal proteins. **Journal of Invertebrate Pathology**, v. 101, p. 1-16, 2009.
- VAN RIE, J.; JANSSENS, S.; HOFTE, H.; DEGHEELE, D.; VAN MELLAERT, H. **Eur. J. Biochem.** 186, 239-247, 1989.

WANG, G.R.; LIANG, G.M.; WU, K.M.; GUO, Y.Y. Gene cloning and sequencing of aminopeptidase N3, a putative receptor for *Bacillus thuringiensis* insecticidal CRY1Ac toxin in *Helicoverpa armigera* (Lepidoptera: Noctuidae). **Eur. J. Entomol.** v.102, p. 13–19, 2005.

WANG, S.; FARFAN-ARRIBAS, D. J.; SHEN, S.; CHOU, T.H.W.; HIRSCH, A.; HE, F.; LU, S. Relative contributions of codon usage, promoter efficiency and leader sequence to the antigen expression and immunogenicity of HIV-1Env DNA vaccine. *Vaccine*, Kidlington, v.24, n.21, p.4531-4540, 2005.

WU, J.Y.; ZHAO, F.Q.; BAI, J.; DENG, G; QIN, S.; BAO, Q.Y. Adaptive evolution of cry genes in *Bacillus thuringiensis*: Implications for their specificity determination. **Genomics Proteomics Bioinform**, v. 5, p. 102-110, 2007.

10ANEXOS

```

1. cry1Ia1 1 5 10 15 20 25 30 35 40 45 55 60 65 70 75 80 85 90 95 105 110 115 120 125 130 135 140 145 155 160 165 170 175 180 185 190 195 205 210 215 220 225
2. cry1Ia2 IQTGIGIAGKILGTLGVVFFAGQVASYLSFILGELWPKGNQWEIFMEHVEEIIINQKISTYARNKALTDLKGGLDALAVYHDSLESWVGNRRNTR---ARSVVKSQYIALELMFVQKLPSPAVSGEEVPLLPYQAANLHLLLRDASIFGKEWGLSSSEISTFYNRQVERAGDYSVHCVKWYSTGLNNLRGTN-AESWVRYNQFRDMTLMVLDLVALFPSYDT--
3. cry1Ia3 IQTGIGIAGKILGTLGVVFFAGQVASYLSFILGELWPKGNQWEIFMEHVEEIIINQKISTYARNKALTDLKGGLDALAVYHDSLESWVGNRRNTR---ARSVVKSQYIALELMFVQKLPSPAVSGEEVPLLPYQAANLHLLLRDASIFGKEWGLSSSEISTFYNRQVERAGDYSVHCVKWYSTGLNNLRGTN-AESWVRYNQFRDMTLMVLDLVALFPSYDT--
4. cry1Ia5 IQTGIGIAGKILGTLGVVFFAGQVASYLSFILGELWPKGNQWEIFMEHVEEIIINQKISTYARNKALTDLKGGLDALAVYHDSLESWVGNRRNTR---ARSVVKSQYIALELMFVQKLPSPAVSGEEVPLLPYQAANLHLLLRDASIFGKEWGLSSSEISTFYNRQVERAGDYSVHCVKWYSTGLNNLRGTN-AESWVRYNQFRDMTLMVLDLVALFPSYDT--
5. cry1Ia11 IQTGIGIAGKILGTLGVVFFAGQVASYLSFILGELWPKGNQWEIFMEHVEEIIINQKISTYARNKALTDLKGGLDALAVYHDSLESWVGNRRNTR---ARSVVKSQYIALELMFVQKLPSPAVSGEEVPLLPYQAANLHLLLRDASIFGKEWGLSSSEISTFYNRQVERAGDYSVHCVKWYSTGLNNLRGTN-AESWVRYNQFRDMTLMVLDLVALFPSYDT--
6. cry1Ia12 IQTGIGIAGKILGTLGVVFFAGQVASYLSFILGELWPKGNQWEIFMEHVEEIIINQKISTYARNKALTDLKGGLDALAVYHDSLESWVGNRRNTR---ARSVVKSQYIALELMFVQKLPSPAVSGEEVPLLPYQAANLHLLLRDASIFGKEWGLSSSEISTFYNRQVERAGDYSVHCVKWYSTGLNNLRGTN-AESWVRYNQFRDMTLMVLDLVALFPSYDT--
7. cry1Ia14 IQTGIGIAGKILGTLGVVFFAGQVASYLSFILGELWPKGNQWEIFMEHVEEIIINQKISTYARNKALTDLKGGLDALAVYHDSLESWVGNRRNTR---ARSVVKSQYIALELMFVQKLPSPAVSGEEVPLLPYQAANLHLLLRDASIFGKEWGLSSSEISTFYNRQVERAGDYSVHCVKWYSTGLNNLRGTN-AESWVRYNQFRDMTLMVLDLVALFPSYDT--
8. cry1Ib3 IQTGIGIAGKILGTLGVVFFAGQIASLYSFLGELWPKGKQWEIFMEHVEEIIINQKILT YARNKALS DLKGLGDALAVYHESLESWVGNRRNTR---ARSVVKQYIALELMFVQKLPSPAVSGEEVPLLPYQAANLHLLLRDASIFGKEWGLSASEISTFYNRQVERTRDYSVHCWKWYSTGLNNLRGTN-AKSWVRYNQFRDMTLMVLDLVALFPSYDT-

9. cry1Ba1 VQTGINIAGRILGVLGVVFFAGQLASFYSLVGLWPRGRDQWEIFLEHVEQLINQQITENARN TALARLQGLGDSFRAYQOSLEDWLENRRDDAR---TRSVLYTQYIALELDFLNAMPLFAIRNQEVPLLMVYAQAANLHLLLRDASLFGSEFGLTSQEIQRYYERQVERTRDYSVYCVWEYNTGLNSLRGTN-AASWVRYNQFRDMTLMVLDLVALFPSYDTRT
10. cry1Ba5 VQTGINIAGRILGVLGVVFFAGQLASFYSLVGLWPRGRDQWEIFLEHVEQLINQQITENARN TALARLQGLGDSFRAYQOSLEDWLENRRDDAR---TRSVLYTQYIALELDFLNAMPLFAIRNQEVPLLMVYAQAANLHLLLRDASLFGSEFGLTSQEIQRYYERQVERTRDYSVYCVWEYNTGLNSLRGTN-AASWVRYNQFRDMTLMVLDLVALFPSYDTRT
11. cry1Bb1 VQTGINIAGRILGVLGVVFFAGQLASFYSLVGLWPSGRDPWEIFLEHVEQLIRQQVTENRNTAIARLEGLGRGYSYQQALETWLDNRNDAR---SRSIILERYVALELDITTAIPLFRIRNEEVPLLMVYAQAANLHLLLRDASLFGSEWGMASSDVNQYQEQIRYTYEYSNHCQWYNTGLNNLRGTN-AESWLRYNQFRDMTLMVLDLVALFPSYDTRT
12. cry1Bc1 VQTGINIAGRILGVLGVVFFAGQLASFYSLVGLWPSGRDPWEIFLEHVEQLIRQQVTENRNTAIARLEGLGRGYSYQQALETWLDNRNDAR---SRSIILERYVALELDITTAIPLFRIRNEEVPLLMVYAQAANLHLLLRDASLFGSEWGMASSDVNQYQEQIRYTYEYSNHCQWYNTGLNNLRGTN-AESWLRYNQFRDMTLMVLDLVALFPSYDTRT
13. cry1Ca3 IDISLSLVQFLVSNF-VPGGGFLVGLIDFVWGI VGPS---QWDAFLVQIEQLINERIAEFARNAAIANLEGLGNFNFIYVEAFKEWEEDPNNPA---TRTRVIDRFRILDGLLERDIPSFRI SGFEVPLLSVYAQAANLHLAALLRDSVIFGERWGLTTINVNENYRRLIRHIDEYADHCANTYRGLNNLP-KSTYQDWITYNRLRRDLTLTVLDIAAFFPNYDNR
14. cry1Ca5 IDISLSLVQFLVSNF-VPGGGFLVGLIDFVWGI VGPS---QWDAFLVQIEQLINERIAEFARNAAIANLEGLGNFNFIYVEAFKEWEEDPNNPA---TRTRVIDRFRILDGLLERDIPSFRI SGFEVPLLSVYAQAANLHLAALLRDSVIFGERWGLTTINVNENYRRLIRHIDEYADHCANTYRGLNNLP-KSTYQDWITYNRLRRDLTLTVLDIAAFFPNYDNR
15. cry1Ca6 IDISLSLVQFLVSNF-VPGGGFLVGLIDFVWGI VGPS---QWDAFLVQIEQLINERIAEFARNAAIANLEGLGNFNFIYVEAFKEWEEDPNNPA---TRTRVIDRFRILDGLLERDIPSFRI SGFEVPLLSVYAQAANLHLAALLRDSVIFGERWGLTTINVNENYRRLIRHIDEYADHCANTYRGLNNLP-KSTYQDWITYNRLRRDLTLTVLDIAAFFPNYDNR
16. cry1Ca8 IDISLSLVQFLVSNF-VPGGGFLVGLIDFVWGI VGPS---QWDAFLVQIEQLINERIAEFARNAAIANLEGLGNFNFIYVEAFKEWEEDPNNPA---TRTRVIDRFRILDGLLERDIPSFRI SGFEVPLLSVYAQAANLHLAALLRDSVIFGERWGLTTINVNENYRRLIRHIDEYADHCANTYRGLNNLP-KSTYQDWITYNRLRRDLTLTVLDIAAFFPNYDNR
17. cry1Ca9 IDISLSLVQFLVSNF-VPGGGFLVGLIDFVWGI VGPS---QWDAFLVQIEQLINERIAEFARNAAIANLEGLGNFNFIYVEAFKEWEEDPNNPA---TRTRVIDRFRILDGLLERDIPSFRI SGFEVPLLSVYAQAANLHLAALLRDSVIFGERWGLTTINVNENYRRLIRHIDEYADHCANTYRGLNNLP-KSTYQDWITYNRLRRDLTLTVLDIAAFFPNYDNR
18. cry1Ea1 N-IALEISRLLASAT-PIG-GILLGLFDAI WSGIGPS---QWDLFLEQIELLIDQKIEEFARNQAI SRLEGISSLYGIYTEAFREWEADPTNPA---LKEEMRTQFNDMNSILVTAIPLFSVQNYQVPLSVYVQAANLHLSVLRDVSVFGQAWGFDIATINSRYNDLTRLPIYTDYAVRWYNTGLDRLPRTGGLRNWARFNQFRRELTISVLDIISFFRNYDSRL
19. cry1Ea2 N-IALEISRLLASAT-PIG-GILLGLFDAI WSGIGPS---QWDLFLEQIELLIDQKIEEFARNQAI SRLEGISSLYGIYTEAFREWEADPTNPA---LKEEMRTQFNDMNSILVTAIPLFSVQNYQVPLSVYVQAANLHLSVLRDVSVFGQAWGFDIATINSRYNDLTRLPIYTDYAVRWYNTGLDRLPRTGGLRNWARFNQFRRELTISVLDIISFFRNYDSRL
20. cry1Ea3 N-IALEISRLLASAT-PIG-GILLGLFDAI WSGIGPS---QWDLFLEQIELLIDQKIEEFARNQAI SRLEGISSLYGIYTEAFREWEADPTNPA---LKEEMRTQFNDMNSILVTAIPLFSVQNYQVPLSVYVQAANLHLSVLRDVSVFGQAWGFDIATINSRYNDLTRLPIYTDYAVRWYNTGLDRLPRTGGLRNWARFNQFRRELTISVLDIISFFRNYDSRL
21. cry1Ea6 N-IALEISRLLASAT-PIG-GILLGLFDAI WSGIGPS---QWDLFLEQIELLIDQKIEEFARNQAI SRLEGISSLYGIYTEAFREWEADPTNPA---LKEEMRTQFNDMNSILVTAIPLFSVQNYQVPLSVYVQAANLHLSVLRDVSVFGQAWGFDIATINSRYNDLTRLPIYTDYAVRWYNTGLDRLPRTGGLRNWARFNQFRRELTISVLDIISFFRNYDSRL
22. cry1Ea7 N-IALEISRLLASAT-PIG-GILLGLFDAI WSGIGPS---QWDLFLEQIELLIDQKIEEFARNQAI SRLEGISSLYGIYTEAFREWEADPTNPA---LKEEMRTQFNDMNSILVTAIPLFSVQNYQVPLSVYVQAANLHLSVLRDVSVFGQAWGFDIATINSRYNDLTRLPIYTDYAVRWYNTGLDRLPRTGGLRNWARFNQFRRELTISVLDIISFFRNYDSRL
23. cry1Fa1 LDISLSLTRFLLSEF-VPGVGVAFGLFDLIWGFITPS---EWSLFLQIEQLIEQRIETLERNRAITTLRGLADSYEVYLEALREWEENPNNAQ---LREDVRIREFANTDDALITAINNFTLTSFEIPLLSVYVQAANLHLSLLRDAVSFGQGWGLDIATVNNHYNRLINLIHRYTEHCLDTYNQGLENLRGTN-TRQWARFNQFRDMTLMVLDLVALFPSYDVRT
24. cry1Fa2 LDISLSLTRFLLSEF-VPGVGVAFGLFDLIWGFITPS---EWSLFLQIEQLIEQRIETLERNRAITTLRGLADSYEVYLEALREWEENPNNAQ---LREDVRIREFANTDDALITAINNFTLTSFEIPLLSVYVQAANLHLSLLRDAVSFGQGWGLDIATVNNHYNRLINLIHRYTEHCLDTYNQGLENLRGTN-TRQWARFNQFRDMTLMVLDLVALFPSYDVRT
25. cry1Fb1 LDISLSLTRFLLSEF-VPGVGVAFGLFDLIWGFITPS---EWSLFLQIEQLIEQRIETLERNRAITTLRGLADSYEVYLEALREWEENPNNAQ---LREDVRIREFANTDDALITAINNFTLTSFEIPLLSVYVQAANLHLSLLRDAVSFGQGWGLDIATVNNHYNRLINLIHRYTEHCLDTYNQGLENLRGTN-TRQWSRFNQFRRELTLMVLDLVALFPSYDAR-
26. cry1Fb2 LDISLSLTRFLLSEF-VPGVGVAFGLFDLIWGFITPS---EWSLFLQIEQLIEQRIETLERNRAITTLRGLADSYEVYLEALREWEENPNNAQ---LREDVRIREFANTDDALITAINNFTLTSFEIPLLSVYVQAANLHLSLLRDAVSFGQGWGLDIATVNNHYNRLINLIHRYTEHCLDTYNQGLENLRGTN-TRQWSRFNQFRRELTLMVLDLVALFPSYDAR-
27. cry1Fb3 LDISLSLTRFLLSEF-VPGVGVAFGLFDLIWGFITPS---EWSLFLQIEQLIEQRIETLERNRAITTLRGLADSYEVYLEALREWEENPNNAQ---LREDVRIREFANTDDALITAINNFTLTSFEIPLLSVYVQAANLHLSLLRDAVSFGQGWGLDIATVNNHYNRLINLIHRYTEHCLDTYNQGLENLRGTN-TRQWSRFNQFRRELTLMVLDLVALFPSYDAR-
28. cry1Fb5 LDISLSLTRFLLSEF-VPGVGVAFGLFDLIWGFITPS---EWSLFLQIEQLIEQRIETLERNRAITTLRGLADSYEVYLEALREWEENPNNAQ---LREDVRIREFANTDDALITAINNFTLTSFEIPLLSVYVQAANLHLSLLRDAVSFGQGWGLDIATVNNHYNRLINLIHRYTEHCLDTYNQGLENLRGTN-TRQWSRFNQFRRELTLMVLDLVALFPSYDAR-
29. cry1Fb7 LDISLSLTRFLLSEF-VPGVGVAFGLFDLIWGFITPS---EWSLFLQIEQLIEQRIETLERNRAITTLRGLADSYEVYLEALREWEENPNNAQ---LREDVRIREFANTDDALITAINNFTLTSFEIPLLSVYVQAANLHLSLLRDAVSFGQGWGLDIATVNNHYNRLINLIHRYTEHCLDTYNQGLENLRGTN-TRQWSRFNQFRRELTLMVLDLVALFPSYD--
30. cry1Gb1 AEIHLKITRLLIENF-LPGGSFAFGLFDLIWGFINED---QWSAFLRQVEELINQRIEFARGQAIQRLVGFGRSYDEYILALKEWENDPDNPA---SKERVTRFRFTDDALLTGVPMAIPGFELATLSVYAQSANLHLALLRDAVFFGERWGLTQTNINDLYSRLKNSIRDYTNHCVRFYNI GLGNLN-VIR---PEYRFQRELTISVLDLVALFPSYDIRT
31. cry1Gb2 AEIHLKITRLLIENF-LPGGSFAFGLFDLIWGFINED---QWSAFLRQVEELINQRIEFARGQAIQRLVGFGRSYDEYILALKEWENDPDNPA---SKERVTRFRFTDDALLTGVPMAIPGFELATLSVYAQSANLHLALLRDAVFFGERWGLTQTNINDLYSRLKNSIRDYTNHCVRFYNI GLGNLN-VIR---PEYRFQRELTISVLDLVALFPSYDIRT

32. cry3Aa1 IQKGI SVVGDLLGVVGFPPGGALVSFYTNFLNTI WPS-EDPWKAFMEQVEALMDQKIADYAKNKALAE LQGLQNNVEDYVSALSSWQKNPVSSRNPHSQGRIRELFSQAESHFRNSMPSFAISGYEVLFLTTYAQAANTHLFLKDAQIYGEWGYEKEDI AEFYKRQKLTQEYTDHCVKWNVGLDKLRGSS-YESWVFNRYRREMTLMVLDLIALFLYDVRL
33. cry3Aa2 IQKGI SVVGDLLGVVGFPPGGALVSFYTNFLNTI WPS-EDPWKAFMEQVEALMDQKIADYAKNKALAE LQGLQNNVEDYVSALSSWQKNPVSSRNPHSQGRIRELFSQAESHFRNSMPSFAISGYEVLFLTTYAQAANTHLFLKDAQIYGEWGYEKEDI AEFYKRQKLTQEYTDHCVKWNVGLDKLRGSS-YESWVFNRYRREMTLMVLDLIALFLYDVRL
34. cry3Aa3 IQKGI SVVGDLLGVVGFPPGGALVSFYTNFLNTI WPS-EDPWKAFMEQVEALMDQKIADYAKNKALAE LQGLQNNVEDYVSALSSWQKNPVSSRNPHSQGRIRELFSQAESHFRNSMPSFAISGYEVLFLTTYAQAANTHLFLKDAQIYGEWGYEKEDI AEFYKRQKLTQEYTDHCVKWNVGLDKLRGSS-YESWVFNRYRREMTLMVLDLIALFLYDVRL
35. cry3Aa4 IQKGI SVVGDLLGVVGFPPGGALVSFYTNFLNTI WPS-EDPWKAFMEQVEALMDQKIADYAKNKALAE LQGLQNNVEDYVSALSSWQKNPVSSRNPHSQGRIRELFSQAESHFRNSMPSFAISGYEVLFLTTYAQAANTHLFLKDAQIYGEWGYEKEDI AEFYKRQKLTQEYTDHCVKWNVGLDKLRGSS-YESWVFNRYRREMTLMVLDLIALFLYDVRL
36. cry3Aa5 IQKGI SVVGDLLGVVGFPPGGALVSFYTNFLNTI WPS-EDPWKAFMEQVEALMDQKIADYAKNKALAE LQGLQNNVEDYVSALSSWQKNPVSSRNPHSQGRIRELFSQAESHFRNSMPSFAISGYEVLFLTTYAQAANTHLFLKDAQIYGEWGYEKEDI AEFYKRQKLTQEYTDHCVKWNVGLDKLRGSS-YESWVFNRYRREMTLMVLDLIALFLYDVRL
37. cry3Aa6 IQKGI SVVGDLLGVVGFPPGGALVSFYTNFLNTI WPS-EDPWKAFMEQVEALMDQKIADYAKNKALAE LQGLQNNVEDYVSALSSWQKNPVSSRNPHSQGRIRELFSQAESHFRNSMPSFAISGYEVLFLTTYAQAANTHLFLKDAQIYGEWGYEKEDI AEFYKRQKLTQEYTDHCVKWNVGLDKLRGSS-YESWVFNRYRREMTLMVLDLIALFLYDVRL
38. cry3Aa7 IQKGI SVVGDLLGVVGFPPGGALVSFYTNFLNTI WPS-EDPWKAFMEQVEALMDQKIADYAKNKALAE LQGLQNNVEDYVSALSSWQKNPVSSRNPHSQGRIRELFSQAESHFRNSMPSFAISGYEVLFLTTYAQAANTHLFLKDAQIYGEWGYEKEDI AEFYKRQKLTQEYTDHCVKWNVGLDKLRGSS-YESWVFNRYRREMTLMVLDLIALFLYDVRL
39. cry3Aa12 IQKGI SVVGDLLGVVGFPPGGALVSFYTNFLNTI WPS-EDPWKAFMEQVEALMDQKIADYAKNKALAE LQGLQNNVEDYVSALSSWQKNPVSSRNPHSQGRIRELFSQAESHFRNSMPSFAISGYEVLFLTTYAQAANTHLFLKDAQIYGEWGYEKEDI AEFYKRQKLTQEYTDHCVKWNVGLDKLRGSS-YESWVFNRYRREMTLMVLDLIALFLYDVRL
: : : : : * : * : : : * : * : : : : : : : : : * * : * * * * : * : * : * *

```

Anexo 1: Alinhamento das 39 seqüências de aminoácidos do domínio I

