



(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ

(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN)  
CỤC SỞ HỮU TRÍ TUỆ

(11)



1-0029074

(51)<sup>8</sup> C12N 15/82; A01H 5/00; C07K 14/325 (13) B

(21) 1-2017-01396

(22) 15/10/2015

(86) PCT/US2015/055779 15/10/2015

(87) WO2016/061377 21/04/2016

(30) 62/064,994 16/10/2014 US; 62/065,017 17/10/2014 US

(45) 25/08/2021 401

(43) 25/07/2017 352A

(73) MONSANTO TECHNOLOGY LLC (US)

800 N. Lindbergh Boulevard, Mail Zone E1NA, St. Louis, Missouri 63167, United States of America

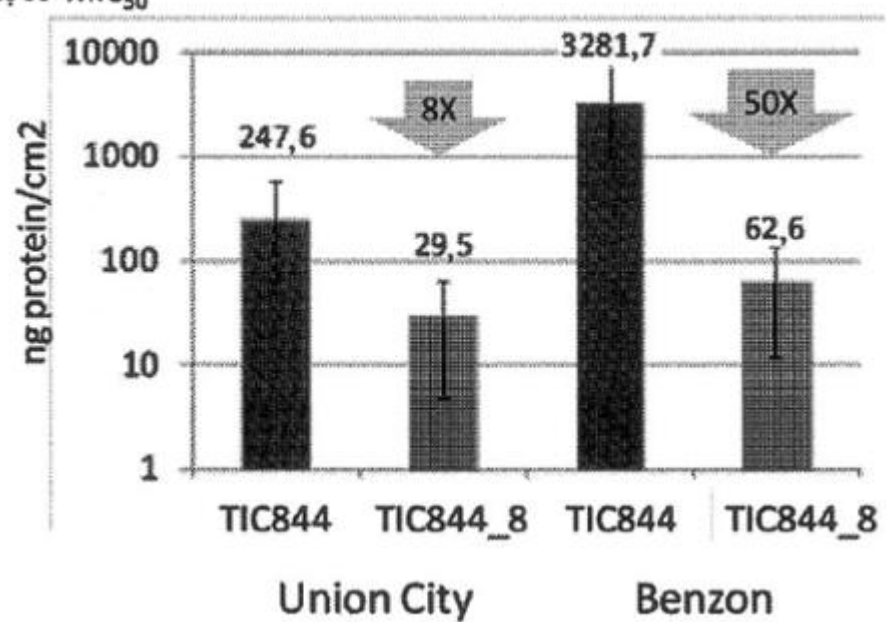
(72) BAUM, James A. (US); FLASINSKI, Stanislaw (US); FU, Xiaoran (US); HOWE, Arlene R. (US); SALVADOR, Sara Ann (US); CERRUTI, Thomas (US).

(74) Công ty Luật TNHH T&G (TGVN)

(54) PROTEIN TRỪ CÔN TRÙNG ĐÃ ĐƯỢC XỬ LÝ, CHẾ PHẨM ỨC CHẾ CÔN TRÙNG CHỨA PROTEIN NÀY VÀ PHƯƠNG PHÁP PHÒNG TRỪ SINH VẬT GÂY HẠI CÁNH VÂY

(57) Sáng chế đề xuất protein Cry1Da đã được xử lý, thể hiện hoạt tính trừ sinh vật gây hại tốt hơn đối với bọ cánh vẩy và phở bọ cánh vẩy gia tăng so với độc tố protein Cry1Da có trong tự nhiên. Sáng chế cũng đề xuất các trình tự polynucleotit được dự định để biểu hiện các protein cải tiến ở thực vật và tế bào chủ chứa các polynucleotit này. Sáng chế cũng đề xuất chế phẩm chứa các protein đã được xử lý với lượng có tác dụng phòng trừ côn trùng, cũng như tế bào thực vật chuyển gen, thực vật chuyển gen, bộ phận của thực vật, và hạt chứa protein đã xử lý này. Sáng chế cũng đề xuất phương pháp phòng trừ hoặc ức chế sinh vật gây hại cánh vẩy, phương pháp sản xuất hạt chứa protein trừ sinh vật gây hại và phương pháp tăng hoạt tính trên bọ cánh vẩy và tăng cường phổ ức chế bọ cánh vẩy của protein.

Trị số MIC<sub>50</sub>



### Lĩnh vực kỹ thuật được đề cập

Nói chung, sáng chế liên quan đến lĩnh vực protein ức chế côn trùng. Nhóm protein này đã được xử lý thể hiện hoạt tính ức chế côn trùng đối với các vật gây hại cho cây trồng và hạt trong nông nghiệp được bộc lộ. Cụ thể, nhóm được bộc lộ này bao gồm các protein có tính ức chế đã được xử lý có hoạt tính trừ côn trùng đối với côn trùng gây hại thuộc bộ cánh vảy. Thực vật, bộ phận của thực vật, và hạt chứa cấu trúc polynucleotit mã hóa một hoặc nhiều protein có hoạt tính ức chế đã xử lý đã bộc lộ được đề xuất.

### Tình trạng kỹ thuật của sáng chế

*Helicoverpa zea* là sinh vật gây hại cánh vảy đáng kể đối với các cây nông nghiệp chính, bao gồm cây ngô, cây bông, và cây đậu tương. Được gọi là sâu đục trái hại cây ngô (corn earworm - CEW), sâu hại cây bông (cotton bollworm - CBW), và sâu hại cây đậu tương (soy podworm - SPW), loài côn trùng ăn tạp này đặc biệt khó phòng trừ bằng các protein trừ côn trùng từ *Bacillus thuringiensis* (Bt) hoặc các loài vi khuẩn khác. *H. zea* được xem là có nguy cơ hình thành tính kháng đối với các đặc tính phòng trừ côn trùng hiện nay, dựa vào khả năng của nó ăn nhiều cây trồng khác nhau và hiện chưa có chiến lược phòng trừ bằng liệu cao. Do đó, cần có các cách thức tác động (MoA) mới để đảm bảo khả năng chịu đựng của thực vật chuyển gen được bảo vệ khỏi sự phá hoại do *H. zea* ăn.

Protein Cry1Da1 là protein có hoạt tính đối với bộ cánh vảy mà được bộc lộ lần đầu tiên trong tài liệu: Hofte, et al. "Nucleotide sequence and deduced amino acid sequence of a new lepidopteran-specific crystal protein gene from *Bacillus thuringiensis*." Nucleic Acides Res. 18(18) (1990): 5545. Protein này thể hiện hoạt tính trừ côn trùng tuyệt vời đối với các loài *Spodoptera* kể cả *Spodoptera frugiperda* (sâu xanh mùa thu, FAW), là vật gây hại đối với vài cây trồng theo luống, bao gồm cây ngô, cây bông và cây đậu tương. Tuy nhiên, Cry1Da1 thể hiện hoạt tính thấp đến trung bình đối với nhiều loài sinh vật gây hại cánh vảy chính khác, bao gồm sâu đục quả (ví dụ, *Helicoverpa armigera* và *H. zea*), sâu đục thân (ví dụ, *Ostrinia nubilalis* và

*Diatraea grandiosella*) và sâu đo đậu tương (*Pseudoplusia includens*). Do phổ trừ côn trùng của nó hẹp và nó không có khả năng bảo vệ ở mức độ thương mại đối với nhiều sinh vật gây hại cánh vảy nông nghiệp quan trọng như CEW, protein trừ côn trùng Cry1Da1 có giá trị hạn chế trong việc dùng làm đặc điểm phòng trừ côn trùng ở thực vật chuyển gen. Kết quả là, hiện không có giống cây thương mại được bảo vệ khỏi côn trùng nào sử dụng Cry1Da1 làm tác nhân bảo vệ đưa vào thực vật.

Mặc dù phổ trừ côn trùng của nó hẹp, Cry1Da1 là protein trừ côn trùng được quan tâm vì dường như protein Cry1Da1 áp dụng MoA khác để phòng trừ một số sinh vật gây hại cánh vảy nhất định. Bằng chứng này có được từ các nghiên cứu với nhiều cụm côn trùng kháng. Ví dụ, các cụm *Plutella xylostella* (sâu tơ) và *Pectinophora gossypiella* (sâu hồng đục quả) lấy từ thực địa mà kháng nhiễm độc Cry1Ac duy trì toàn bộ tính miễn cảm với protein Cry1Da1 (Tabashnik, et al. "Cross-Resistance of Pink Bollworm (*Pectinophora gossypiella*) to *Bacillus thuringiensis* toxins." Appl. Environ. Microbiol. 66 (2000): 4582-4584; Tabashnik, et al. "Cross-Resistance to *Bacillus thuringiensis* Toxin Cry1Ja in strain of diamondback moth adapted to Artificial Diet". J. Invert. Pathol. 76:(2000): 81-83.). Các bằng chứng này cho thấy rằng Cry1Da1 nhận diện các thụ thể ở ruột giữa của bộ cánh vảy khác biệt với các loại nhận diện được bởi các protein hoạt tính đối với bộ cánh vảy hiện đang được triển khai ở các thực vật chuyển gen bao gồm Cry1Ac, Cry1Ab, Cry1A.105, Cry1Fa, Cry2Ae, và Cry2Ab2. Nhờ MoA mới dễ thấy này, việc tối ưu hóa các protein kiểu Cry1Da1 nhằm có hoạt tính tốt hơn đối với phổ các loài *Helicoverpa* rộng hơn trong khi vẫn duy trì hoặc làm tăng hoạt tính trừ côn trùng của chúng đối với *Spodoptera* hẳn sẽ tạo ra tác nhân bảo vệ có giá trị cao kết hợp vào thực vật nhằm quản lý tính kháng của côn trùng.

#### **Bản chất kỹ thuật của sáng chế**

Theo sáng chế, một vài biến thể của trình tự axit amin của các protein khung TIC844 và Cry1Da đã được xác định mà thể hiện hoạt tính tốt hơn đáng kể (so với độc tố Cry1Da1 tự nhiên) đối với *H. zea* trong khi vẫn giữ được hoạt tính tốt đối với *S. frugiperda*. Các biến thể cải tiến của TIC844 và Cry1Da đã được xử lý để được biểu hiện ở cây trồng (ví dụ, cây ngô, cây đậu tương, cây bông, cây mía), và là các lựa chọn mới cho việc quản lý tính kháng *in-planta* và phòng trừ côn trùng gây hại là bộ cánh vảy nhờ phương thức tác động rất khác thường của Cry1Da được kết hợp với việc cải thiện hoạt tính đối với *H. zea* nhờ xử lý.

Các protein đã được xử lý có độc tính đối với bọ cánh vảy được bộc lộ theo sáng chế (được gọi là “các protein của độc tố đã được xử lý”, “các protein đã được xử lý có độc tính”, hoặc “các protein trừ côn trùng đã được xử lý”) là các chất dẫn xuất của độc tố Cry1Da1 có hoạt tính trừ côn trùng có trong tự nhiên của *Bacillus thuringiensis* (SEQ ID NO:2) hoặc thể tương đồng Cry1Da1 khảm, là TIC844 (SEQ ID NO:14), mà chứa độc tố lõi của Cry1Da1 nhưng thay thế tiền độc tố Cry1Ab3 cho miền của tiền độc tố Cry1Da1 có trong tự nhiên. Mỗi protein trừ côn trùng đã được xử lý theo sáng chế có ít nhất một thay thế axit amin, một bổ sung axit amin, hoặc một loại bỏ axit amin so với các protein khung bất kỳ trong số các trình tự nêu trong SEQ ID NO:2 hoặc SEQ ID NO:14. Các protein trừ côn trùng đã được xử lý theo sáng chế là đặc biệt độc đối với các loài côn trùng *Helicoverpa zea* (sâu hại ngô, sâu hại đậu tương, sâu hại bông) và *Spodoptera frugiperda* (sâu xanh mùa thu). Trong khi các protein khung TIC844 (SEQ ID NO:14) và Cry1Da1 (SEQ ID NO:2) thể hiện độc tính thấp đối với *H. zea*, thì các protein trừ côn trùng đã được xử lý theo sáng chế thể hiện hoạt tính trừ côn trùng tốt hơn một cách đáng ngạc nhiên và ngoài mong đợi và phổ trừ côn trùng rộng hơn đối với các côn trùng gây hại là bọ cánh vảy kể cả *H. zea*.

Theo các phương án nhất định, protein trừ côn trùng đã được xử lý chứa trình tự axit amin bất kỳ trong số các trình tự nêu trong SEQ ID NO:44, SEQ ID NO: 40, SEQ ID NO: 12, SEQ ID NO:26, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38 hoặc SEQ ID NO:42, hoặc đoạn có hoạt tính ức chế côn trùng của chúng được bộc lộ. Theo các phương án nhất định, protein trừ côn trùng đã được xử lý này thể hiện hoạt tính ức chế đối với loài côn trùng thuộc bộ bọ cánh vảy. Loài sinh vật gây hại cánh vảy đích bị ức chế bởi các protein có độc tính đối với bọ cánh vảy theo sáng chế bao gồm ít nhất sâu xanh mùa thu (*Spodoptera frugiperda*), sâu ăn lá cây củ cải đường (*Spodoptera exigua*), sâu ăn lá bertha (*Mamestra configurata*), sâu xám (*Agrotis ipsilon*), sâu đo ký sinh trên cải bắp (*Trichoplusia ni*), sâu đo đậu tương (*Chrysodeixis includens*), sâu bướm hại đậu mèo (*Anticarsia gemmatilis*), sâu xanh (*Hypena scabra*), sâu chồi thuốc lá (*Heliothis virescens*), sâu ngài đêm dạng hạt (*Agrotis subterranea*), sâu ăn lá (*Pseudaletia unipuncta*), sâu ngài đêm tây (*Agrotis orthogonia*), sâu đục thân ngô châu Âu (*Ostrinia nubilalis*), bướm hại cam navel (*Amyelois transitella*), sâu kéo màng hại

rễ ngô (*Crambus caliginosellus*), sâu kéo màng ở cỏ (*Herpetogramma licarsisalis*), sâu bướm hướng dương (*Homoeosoma electellum*), sâu đục thân cây ngô nhỏ (*Elasmopalpus lignosellus*), bướm cá tuyết (*Cydia pomonella*), bướm hại cây việt quất (*Endopiza viteana*), bướm hại quả phượng đông (*Grapholita molesta*), bướm hại chồi hướng dương (*Suleima helianthana*), sâu tơ (*Plutella xylostella*), sâu hồng đục quả (*Pectinophora gossypiella*), sâu đục thân hồng (*Sesamia inferens*), sâu bướm gypsy (*Lymantria dispar*), sâu hại lá cây bông (*Alabama argillacea*), sâu cuốn lá cây ăn quả (*Archips argyrospila*), sâu cuốn lá châu Âu (*Archips rosana*), sâu đục thân lúa châu Á, hoặc sâu đục thân lúa (*Chilo suppressalis*), sâu cuốn lá lúa (*Cnaphalocrocis medinalis*), sâu kéo màng hại rễ ngô (*Crambus caliginosellus*), sâu kéo màng ở cỏ xanh (*Crambus teterrellus*), sâu đục thân ngô tây nam (*Diatraea grandiosella*), sâu đục thân mía (*Diatraea saccharalis*), sâu đục quả có gai (*Earias insulana*), sâu đục quả đốm (*Earias vittella*), sâu hại bông cổ (*Helicoverpa armigera*), sâu đục trái hại ngô, sâu hại đậu tương hoặc sâu hại bông (*Helicoverpa zea*), sâu kéo màng ở cỏ (*Herpetogramma licarsisalis*), bướm hại nho châu Âu (*Lobesia botrana*), sâu ăn lá quả có múi (*Phyllocnistis citrella*), bướm trắng to (*Pieris brassicae*), sâu hại cải bắp nhập khẩu, hoặc bướm trắng nhỏ (*Pieris rapae*), sâu ngài đêm hại thuốc lá, hoặc sâu bướm cụm (*Spodoptera litura*), và sâu ăn lá cà chua (*Tuta absoluta*).

Sáng chế còn đề xuất polynucleotit mã hóa protein trừ côn trùng đã được xử lý hoặc đoạn có hoạt tính trừ vật gây hại của nó, trong đó polynucleotit được liên kết theo kiểu hoạt động được với đoạn khởi đầu khác loại và protein trừ côn trùng đã được xử lý chứa trình tự axit amin bất kỳ trong số các trình tự nêu trong SEQ ID NO:44, SEQ ID NO: 40, SEQ ID NO: 12, SEQ ID NO:26, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38 hoặc SEQ ID NO:42.

Theo phương án khác, sáng chế đề xuất polynucleotit mã hóa protein trừ côn trùng đã được xử lý, trong đó polynucleotit chứa trình tự nucleotit mà tùy ý lai trong các điều kiện nghiêm ngặt với trình tự bổ sung ngược của trình tự polynucleotit bất kỳ trong số các trình tự nêu trong SEQ ID NO: 43, SEQ ID NO: 39, SEQ ID NO: 11, SEQ ID NO: 11, SEQ ID NO: 25, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21,

SEQ ID NO: 23, SEQ ID NO:31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37 hoặc SEQ ID NO: 41; hoặc mã hóa protein trừ côn trùng đã được xử lý chứa trình tự axit amin bất kỳ trong số các trình tự nêu trong SEQ ID NO:44, SEQ ID NO: 40, SEQ ID NO: 12, SEQ ID NO:26, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38 hoặc SEQ ID NO:42.

Sáng chế còn đề xuất tế bào chủ chứa polynucleotit bất kỳ trong số các trình tự nêu trong SEQ ID NO: 43, SEQ ID NO: 39, SEQ ID NO: 11, SEQ ID NO: 11, SEQ ID NO: 25, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO:31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37 hoặc SEQ ID NO: 41, trong đó tế bào chủ được chọn từ nhóm bao gồm tế bào chủ vi khuẩn hoặc tế bào chủ thực vật. Các tế bào chủ vi khuẩn được dự định bao gồm các tế bào chủ vi khuẩn được chọn từ nhóm bao gồm *Agrobacterium*, *Rhizobium*, *Bacillus*, *Brevibacillus*, *Escherichia*, *Pseudomonas*, *Klebsiella*, và *Erwinia*, trong đó loài *Bacillus* là *Bacillus cereus* hoặc *Bacillus thuringiensis*, *Brevibacillus* là *Brevibacillus laterosperous*, và *Escherichia* là *Escherichia coli*. Ngoài ra, các tế bào chủ thực vật được dự định bao gồm loại một lá mầm hoặc loại hai lá mầm.

Theo phương án khác nữa, sáng chế đề xuất chế phẩm ức chế côn trùng chứa protein trừ côn trùng đã được xử lý chứa trình tự axit amin bất kỳ trong số các trình tự nêu trong SEQ ID NO:44, SEQ ID NO: 40, SEQ ID NO: 12, SEQ ID NO:26, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38 hoặc SEQ ID NO:42 hoặc đoạn có hoạt tính ức chế côn trùng của nó. Dự định rằng chế phẩm ức chế côn trùng này còn có thể chứa ít nhất một tác nhân ức chế côn trùng khác với protein trừ côn trùng đã được xử lý. Các tác nhân ức chế côn trùng đã được dự định bao gồm protein ức chế côn trùng, phân tử ARN sợi đôi ức chế côn trùng, và hóa chất ức chế côn trùng. Dự định rằng ít nhất một tác nhân trừ vật gây hại khác có thể thể hiện hoạt tính đối với một hoặc nhiều loài vật gây hại thuộc các bộ bọ cánh vảy, bọ cánh cứng, bọ cánh nửa, bọ cánh đều, hoặc bọ cánh tơ.

Sáng chế còn đề xuất hạt chứa lượng hữu hiệu có tác dụng ức chế côn trùng của protein trừ côn trùng đã được xử lý chứa trình tự axit amin bất kỳ trong số các trình tự nêu trong SEQ ID NO:44, SEQ ID NO: 40, SEQ ID NO: 12, SEQ ID NO:26, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38 hoặc SEQ ID NO:42; hoặc polynucleotit bất kỳ trong số các trình tự nêu trong SEQ ID NO: 43, SEQ ID NO: 39, SEQ ID NO: 11, SEQ ID NO: 11, SEQ ID NO: 25, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO:31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37 hoặc SEQ ID NO: 41.

Theo phương án khác, sáng chế còn đề xuất phương pháp phòng trừ sinh vật gây hại cánh vảy, phương pháp này bao gồm bước cho sinh vật gây hại cánh vảy tiếp xúc với protein trừ côn trùng đã được xử lý với lượng có tác dụng ức chế.

Theo phương án khác nữa, sáng chế đề xuất tế bào thực vật, thực vật hoặc bộ phận của thực vật chuyển gen chứa protein trừ côn trùng đã được xử lý. Sáng chế đề xuất các phương pháp phòng trừ sinh vật gây hại cánh vảy, bao gồm việc cho vật gây hại tiếp xúc với tế bào thực vật, thực vật hoặc bộ phận của thực vật chuyển gen chứa protein trừ côn trùng đã được xử lý. Các sản phẩm thương mại thu được từ tế bào thực vật, thực vật hoặc bộ phận của thực vật chứa protein trừ côn trùng đã được xử lý, trong đó sản phẩm chứa protein trừ côn trùng đã được xử lý, với lượng phát hiện được, cũng được dự định. Các sản phẩm thương mại được dự định bao gồm sinh khối thực vật, dầu, bột thô, thức ăn chăn nuôi, bột, cốm, cám, thóc, vỏ, và hạt đã được xử lý.

Phương pháp khác được bộc lộ theo sáng chế là phương pháp sản xuất hạt chứa protein trừ côn trùng đã được xử lý, phương pháp này bao gồm các bước: gieo ít nhất một hạt chứa protein trừ côn trùng đã được xử lý; trồng cây từ hạt đó; và thu hoạch hạt từ cây này, trong đó hạt thu được chứa protein trừ côn trùng đã được xử lý.

Theo phương án khác nữa, sáng chế đề xuất phương pháp ức chế sinh vật gây hại cánh vảy hại cây trồng bao gồm việc cải biến một hoặc nhiều gốc axit amin của trình tự SEQ ID NO: 2 hoặc SEQ ID NO:14 thông qua việc thế một hoặc nhiều gốc axit amin để tạo ra trình tự SEQ ID NO:2 hoặc SEQ ID NO:14 cải biến; và tạo ra SEQ ID

NO: 2 hoặc SEQ ID NO:14 cải biến với lượng ức chế bọ cánh vảy sẵn bên trong, trên bề mặt, hoặc trong vùng lân cận của mô của cây trồng đó; trong đó gốc axit amin đã được cải biến của SEQ ID NO:2 hoặc SEQ ID NO:14 được chọn từ nhóm bao gồm serin ở vị trí 282 được thế bằng lysin hoặc valin, tyrosin ở vị trí 316 được thế bằng serin, isoleuxin ở vị trí 368 được thế bằng prolin hoặc arginin, serin ở 374 được thế bằng arginin, asparagin ở vị trí 375 được thế bằng histidin, và isoleuxin ở vị trí 432 được thế bằng leuxin.

Các phân tử polynucleotit tái tổ hợp mà mã hóa các protein trừ côn trùng đã được xử lý theo sáng chế cũng được đề xuất. Các phân tử polynucleotit tái tổ hợp được dự định bao gồm trình tự polynucleotit được chọn từ nhóm bao gồm các trình tự được nêu trong trình tự SEQ ID NO: 43, SEQ ID NO: 39, SEQ ID NO: 11, SEQ ID NO: 11, SEQ ID NO: 25, SEQ ID NO: 3, SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 9, SEQ ID NO: 15, SEQ ID NO: 17, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 23, SEQ ID NO:31, SEQ ID NO: 33, SEQ ID NO: 35, SEQ ID NO: 37 hoặc SEQ ID NO: 41; và tùy ý trình tự polynucleotit mã hóa tác nhân ức chế côn trùng khác với protein trừ côn trùng đã được xử lý.

Phương pháp khác theo sáng chế là phương pháp làm gia tăng hoạt tính đối với bọ cánh vảy và làm gia tăng phổ ức chế bọ cánh vảy của protein khung, phương pháp này bao gồm việc cải biến một hoặc nhiều gốc axit amin của trình tự SEQ ID NO: 2 hoặc SEQ ID NO: 14 thông qua việc thế (các) gốc axit amin để tạo ra protein trừ côn trùng đã được xử lý, trong đó gốc axit amin đã được cải biến của trình tự SEQ ID NO:2 hoặc SEQ ID NO:14 được chọn từ nhóm bao gồm serin ở vị trí 282 được thế bằng lysin hoặc valin, tyrosin ở vị trí 316 được thế bằng serin, isoleuxin ở vị trí 368 được thế bằng prolin hoặc arginin, serin ở 374 được thế bằng arginin, asparagin ở vị trí 375 được thế bằng histidin, và isoleuxin ở vị trí 432 được thế bằng leuxin. Theo các phương án nhất định của phương pháp này, protein trừ côn trùng đã được xử lý này có khả năng gây chết *Helicoverpa zea* tăng ít nhất tám lần so với protein khung.

Các phương án, dấu hiệu, và các ưu điểm khác theo sáng chế trở nên rõ ràng khi tham khảo phần mô tả chi tiết, các ví dụ thực hiện sáng chế, và các điểm yêu cầu bảo hộ dưới đây.

**Mô tả vắn tắt các hình vẽ**

Hình 1 minh họa các trị số MIC<sub>50</sub> của protein khung TIC844 (SEQ ID NO: 14) so với protein trừ côn trùng đã được xử lý TIC844\_8 (SEQ ID NO: 26) ở hai cụm *Helicoverpa zea* (CEW) khác nhau, Union City và Benzon.

**Mô tả vắn tắt các trình tự**

SEQ ID NO:1 là trình tự nucleotit mã hóa protein Cry1Da1.

SEQ ID NO:2 là trình tự axit amin của độc tố protein Cry1Da1.

SEQ ID NO:3 là trình tự nucleotit mã hóa protein Cry1Da1\_3.

SEQ ID NO:4 là trình tự axit amin của độc tố protein Cry1Da1\_3.

SEQ ID NO:5 là trình tự nucleotit mã hóa protein Cry1Da1\_4.

SEQ ID NO:6 là trình tự axit amin của độc tố protein Cry1Da1\_4.

SEQ ID NO:7 là trình tự nucleotit mã hóa protein Cry1Da1\_5.

SEQ ID NO:8 là trình tự axit amin của độc tố protein Cry1Da1\_5.

SEQ ID NO:9 là trình tự nucleotit mã hóa protein Cry1Da1\_6.

SEQ ID NO:10 là trình tự axit amin của độc tố protein Cry1Da1\_6.

SEQ ID NO:11 là trình tự nucleotit mã hóa protein Cry1Da1\_7.

SEQ ID NO:12 là trình tự axit amin của độc tố protein Cry1Da1\_7.

SEQ ID NO:13 là trình tự nucleotit mã hóa protein TIC844.

SEQ ID NO:14 là trình tự axit amin của độc tố protein TIC844.

SEQ ID NO:15 là trình tự nucleotit mã hóa protein TIC844\_2.

SEQ ID NO:16 là trình tự axit amin của độc tố protein TIC844\_2.

SEQ ID NO:17 là trình tự nucleotit mã hóa protein TIC844\_4.

SEQ ID NO:18 là trình tự axit amin của độc tố protein TIC844\_4.

SEQ ID NO:19 là trình tự nucleotit mã hóa protein TIC844\_5.

SEQ ID NO:20 là trình tự axit amin của độc tố protein TIC844\_5.

SEQ ID NO:21 là trình tự nucleotit mã hóa protein TIC844\_6.

SEQ ID NO:22 là trình tự axit amin của độc tố protein TIC844\_6.

SEQ ID NO:23 là trình tự nucleotit mã hóa protein TIC844\_7.

SEQ ID NO:24 là trình tự axit amin của độc tố protein TIC844\_7.

SEQ ID NO:25 là trình tự nucleotit mã hóa protein TIC844\_8.

SEQ ID NO:26 là trình tự axit amin của độc tố protein TIC844\_8.

SEQ ID NO:27 là trình tự polynucleotit được thiết kế để dùng trong việc biểu hiện protein Cry1Da1 ở thực vật.

SEQ ID NO:28 là trình tự axit amin của độc tố protein Cry1Da1.

SEQ ID NO:29 là trình tự polynucleotit được thiết kế để dùng trong việc biểu hiện protein Cry1Da1\_2.nno ở thực vật.

SEQ ID NO:30 là trình tự axit amin của độc tố protein Cry1Da1\_2.nno.

SEQ ID NO:31 là trình tự polynucleotit được thiết kế để dùng trong việc biểu hiện protein Cry1Da1\_3.nno ở thực vật.

SEQ ID NO:32 là trình tự axit amin của độc tố protein Cry1Da1\_3.nno.

SEQ ID NO:33 là trình tự polynucleotit được thiết kế để dùng trong việc biểu hiện protein Cry1Da1\_4.nno ở thực vật.

SEQ ID NO:34 là trình tự axit amin của độc tố protein Cry1Da1\_4.nno.

SEQ ID NO:35 là trình tự polynucleotit được thiết kế để dùng trong việc biểu hiện protein Cry1Da1\_5.nno ở thực vật.

SEQ ID NO:36 là trình tự axit amin của độc tố protein Cry1Da1\_5.nno.

SEQ ID NO:37 là trình tự polynucleotit được thiết kế để dùng trong việc biểu hiện protein Cry1Da1\_6.nno ở thực vật.

SEQ ID NO:38 là trình tự axit amin của độc tố protein Cry1Da1\_6.nno.

SEQ ID NO:39 là trình tự polynucleotit được thiết kế để dùng trong việc biểu hiện protein Cry1Da1\_7.nno ở thực vật.

SEQ ID NO:40 là trình tự axit amin của độc tố protein Cry1Da1\_7.nno.

SEQ ID NO:41 là trình tự polynucleotit được thiết kế để dùng trong việc biểu hiện protein TIC844\_9.nno ở thực vật.

SEQ ID NO:42 là trình tự axit amin của độc tố protein TIC844\_9.nno.

SEQ ID NO:43 là trình tự polynucleotit được thiết kế để dùng trong việc biểu hiện protein TIC844\_11.nno ở thực vật.

SEQ ID NO:44 là trình tự axit amin của độc tố protein TIC844\_11.nno.

### **Mô tả chi tiết sáng chế**

Các protein trừ côn trùng đã được xử lý mà thể hiện hoạt tính độc ở mức độ cao đáng ngạc nhiên đối với loài bọ cánh vảy và phổ trừ côn trùng rộng hơn so với các protein trừ côn trùng đã biết khác đối với bọ cánh vảy được bộc lộ theo sáng chế. Các protein trừ côn trùng đã được xử lý này có nguồn gốc từ các protein khung có hoạt tính trừ côn trùng, mà có tác dụng làm khuôn cho các cải biến axit amin khác nhau. Các ví dụ về protein khung có hoạt tính trừ côn trùng này bao gồm, nhưng không bị giới hạn ở, Cry1Da1 và TIC844 (thể tương đồng Cry1Da1). TIC844 chứa độc tố lõi của Cry1Da1 (tức là các miền I, II và III) nhưng sử dụng miền của tiền độc tố Cry1Ab3 để đảm bảo biểu hiện tốt ở *Bacillus thuringiensis* (Bt). Mức độ biểu hiện của Cry1Da1 ở Bt là kém khi sử dụng miền của tiền độc tố Cry1Da1 tự nhiên. Tuy nhiên, như được thể hiện trong bản mô tả này, mức độ biểu hiện của độc tố lõi của Cry1Da1 được cải thiện đáng kể ở chủng không kết tinh Bt khi miền tự nhiên của tiền độc tố được loại bỏ và đoạn mã hóa độc tố lõi Cry1Da1 được dung hợp trong khung với đoạn mã hóa miền tiền độc tố Cry1Ab3. Đặc biệt là, các protein khung TIC844 (SEQ ID NO:14) và Cry1Da1 (SEQ ID NO:2) không thể hiện phổ ức chế bọ cánh vảy hữu dụng về mặt thương mại và hoạt tính ức chế bọ cánh vảy tốt hơn quan sát được ở các protein trừ côn trùng đã được xử lý.

Các protein trừ côn trùng đã xử lý bộc lộ theo sáng chế được đề xuất có cải biến axit amin mà nhờ đó các protein đã được cải biến này thể hiện phổ ức chế bọ cánh vảy tốt hơn và/hoặc hoạt tính ức chế bọ cánh vảy tốt hơn so với protein khung gốc, TIC844 hoặc Cry1Da1. Các thuật ngữ “có hoạt tính hơn”, “hoạt tính tốt hơn”, “tính đặc hiệu tốt hơn”, “độc lực cao hơn”, “độc tính cao hơn”, “hoạt tính ức chế bọ cánh vảy tốt hơn”, “hoạt tính ức chế bọ cánh vảy cao hơn”, và “phổ ức chế bọ cánh vảy tốt hơn” được dùng để chỉ sự so sánh hoạt tính của protein trừ côn trùng đã được xử lý với hoạt tính của protein khung (TIC844 hoặc Cry1Da1) đối với côn trùng bọ cánh vảy, trong đó hoạt tính được quy cho protein trừ côn trùng đã được xử lý mạnh hơn hoạt tính được quy cho protein khung. Theo các phương án nhất định, các protein trừ côn trùng đã được xử lý theo sáng chế thể hiện phổ ức chế bọ cánh vảy tốt hơn và/hoặc hoạt tính ức chế bọ cánh vảy tốt hơn hoặc cao hơn so với hoạt tính của protein khung TIC844

hoặc Cry1Da1 khi các loài sinh vật gây hại cánh vảy bao gồm, nhưng không bị giới hạn ở, *Helicoverpa zea* và *Spodoptera frugiperda*.

Các thuật ngữ và cụm từ “có hoạt tính” hoặc “hoạt tính”; “hoạt tính trừ vật gây hại” hoặc “trừ vật gây hại”; hoặc “hoạt tính trừ côn trùng”, “ức chế côn trùng”, “trừ côn trùng”, hoặc “lượng có tác dụng ức chế côn trùng”, được dùng theo sáng chế chỉ hiệu quả của tác nhân có độc tính, như protein trừ côn trùng, trong việc ức chế (ức chế sự sinh trưởng, ăn, sinh sản, hoặc sức sống), ngăn chặn (ngăn chặn sự sinh trưởng, ăn, sinh sản, hoặc sức sống), phòng trừ (phòng trừ sự phá hoại của vật gây hại, phòng trừ vật gây hại ăn cây trồng cụ thể chứa protein trừ côn trùng đã xử lý đã được bộc lộ với lượng hữu hiệu) hoặc diệt (gây bệnh, chết, hoặc giảm mức độ sinh sản) ở vật gây hại. Tương tự, thuật ngữ “lượng ức chế bộ cánh vảy” được dùng để chỉ lượng tác nhân có độc tính, như protein trừ côn trùng, mà ức chế sức sống của bộ cánh vảy, sự sinh trưởng của bộ cánh vảy, sự phát triển của bộ cánh vảy, sự sinh sản của bộ cánh vảy, tập tính ăn của bộ cánh vảy, tập tính giao phối của bộ cánh vảy bất kỳ ở mức độ đo được và/hoặc làm giảm tác dụng phụ bất kỳ ở mức độ đo được do việc bộ cánh vảy hại cây. Các thuật ngữ này được dự định bao gồm kết quả do việc cung cấp tác nhân có độc tính với lượng có tác dụng trừ vật hại hữu hiệu cho vật hại khi việc tiếp cận của vật hại với tác nhân có độc tính này dẫn đến bệnh tật, chết, giảm sinh sản, hoặc còi cọc. Các thuật ngữ này còn bao gồm sự sợ hãi của vật gây hại đối với thực vật, mô thực vật, bộ phận của thực vật, hạt, các tế bào thực vật, hoặc đối với vị trí địa lý cụ thể nơi cây đó có thể sinh trưởng, nhờ kết quả của việc cung cấp tác nhân có độc tính với lượng có tác dụng trừ vật hại hữu hiệu ở hoặc trên cây đó. Nói chung, thuật ngữ hoạt tính trừ vật gây hại được dùng để chỉ khả năng của tác nhân có độc tính là hữu hiệu trong việc ức chế sự sinh trưởng, phát triển, sức sống, tập tính ăn, tập tính giao phối, sức sinh sản, hoặc làm giảm tác dụng phụ bất kỳ ở mức độ đo được gây ra bởi việc côn trùng ăn protein này, đoạn protein, phân đoạn protein hoặc polynucleotit của vật gây hại đích cụ thể, bao gồm, nhưng không bị giới hạn ở, côn trùng thuộc bộ bộ cánh vảy. Tác nhân có độc tính có thể được sản sinh ra bởi thực vật hoặc có thể được dùng cho thực vật hoặc môi trường trong khu vực của thực vật đó.

Lượng có tác dụng trừ vật hại của tác nhân có độc tính, khi được cung cấp trong chế độ dinh dưỡng cho vật gây hại đích, thể hiện hoạt tính trừ vật gây hại khi vật gây hại ăn chất độc này vào. Tác nhân có độc tính có thể là protein trừ vật gây hại hoặc một hoặc nhiều tác nhân hóa học đã biết trong lĩnh vực kỹ thuật này. Các tác nhân hóa học trừ vật gây hại hoặc trừ côn trùng và các tác nhân protein trừ vật gây hại hoặc trừ

côn trùng có thể được sử dụng riêng rẽ hoặc kết hợp với nhau. Các tác nhân hóa học bao gồm, nhưng không bị giới hạn ở, các phân tử ARN sợi đôi hướng đích gen cụ thể để ngăn chặn ở vật gây hại đích, clorua hữu cơ, phosphat hữu cơ, carbamat, pyrethroid, neonicotinoit, và ryanoit. Các tác nhân protein trừ vật gây hại hoặc trừ côn trùng bao gồm các protein trừ côn trùng đã được xử lý nêu trong bản mô tả này, cũng như các tác nhân có bản chất protein có độc tính khác bao gồm các tác nhân mà hướng đích đến loài sinh vật gây hại cánh vảy, cũng như các độc tố protein mà được dùng để phòng trừ vật gây hại thực vật khác như các protein Cry sẵn có trong lĩnh vực kỹ thuật này để dùng trong việc phòng trừ các loài bọ cánh cứng, bọ cánh nửa và bọ cánh đều.

Các thuật ngữ “phân đoạn” hoặc “đoạn” được dùng theo sáng chế chỉ trình tự axit amin hoặc axit nucleic liên tục mà là ngắn hơn trình tự axit amin hoặc axit nucleic hoàn chỉnh mô tả các protein trừ côn trùng đã được xử lý.

Dự định rằng việc đề cập đến vật gây hại, đặc biệt là vật gây hại cho cây trồng, có nghĩa là các côn trùng gây hại cây trồng, đặc biệt là các côn trùng gây hại là bọ cánh vảy mà được phòng trừ bởi các protein trừ côn trùng đã xử lý được bộc lộ. Tuy nhiên, đề cập đến vật gây hại còn có thể bao gồm các côn trùng gây hại thực vật là bọ cánh cứng, bọ cánh nửa và bọ cánh đều, cũng như giun tròn và nấm mốc, khi các tác nhân có độc tính hướng đích đến các vật gây hại này ở cùng khu vực hoặc có mặt cùng với các protein trừ côn trùng đã được xử lý đã bộc lộ.

Các protein trừ côn trùng đã xử lý được bộc lộ thể hiện hoạt tính trừ côn trùng đối với các côn trùng gây hại thuộc loài côn trùng bọ cánh vảy, kể cả con trưởng thành, con nhộng, ấu trùng, và con mới sinh. Côn trùng thuộc bộ bọ cánh vảy bao gồm, nhưng không bị giới hạn ở, sâu ăn lá, sâu ngài đêm, sâu đo, và bướm đêm thuộc họ Noctuidae, ví dụ, sâu xanh mùa thu (*Spodoptera frugiperda*), sâu ăn lá cây củ cải đường (*Spodoptera exigua*), sâu ăn lá bertha (*Mamestra con figurata*), sâu xám (*Agrotis ipsilon*), sâu đo ký sinh trên cải bắp (*Trichoplusia ni*), sâu đo đậu tương (*Pseudoplusia includens*), sâu bướm hại đậu mè (*Anticarsia gemmatilis*), sâu xanh (*Hyponomeuta scabra*), sâu chồi thuốc lá (*Heliothis virescens*), sâu ngài đêm dạng hạt (*Agrotis subterranea*), sâu ăn lá (*Pseudaletia unipuncta*), sâu ngài đêm tây (*Agrotis orthogonia*); sâu đục thân, sâu túi, sâu kéo màng, sâu hình nón, sâu hại cải bắp và sâu gặm thịt lá thuộc họ Pyralidae, ví dụ, sâu đục thân ngô châu Âu (*Ostrinia nubilalis*), bướm hại cam navel (*Amyelois transitella*), sâu kéo màng hại rễ ngô (*Crambus caliginosellus*), sâu kéo màng ở cỏ (*Herpetogramma licarsisalis*), sâu bướm hương dương (*Homoeosoma electellum*), sâu đục thân cây ngô nhỏ (*Elasmopalpus*

*lignosellus*); sâu cuốn lá, sâu đục chồi, sâu hại hạt, và sâu hại quả thuộc họ Tortricidae, ví dụ, bướm cá tuyết (*Cydia pomonella*), bướm hại cây việt quất (*Endopiza viteana*), bướm hại quả phượng đông (*Grapholitamolesta*), bướm hại chồi hướng dương (*Suleima helianthana*); và nhiều bộ cánh vảy khác quan trọng về mặt kinh tế, ví dụ, sâu tơ (*Plutella xylostella*), sâu hồng đục quả (*Pectinophora gossypiella*) và sâu bướm gypsy (*Lymantria dispar*). Các côn trùng gây hại khác thuộc bộ bộ cánh vảy bao gồm, ví dụ, *Alabama argillacea* (sâu hại lá cây bông), *Archips argyrospila* (sâu cuốn lá cây ăn quả), *Archips rosana* (sâu cuốn lá châu Âu) và các loài *Archips* khác, *Chilo suppressalis* (sâu đục thân lúa châu Á, hoặc sâu đục thân lúa), *Cnaphalocrocis medinalis* (sâu cuốn lá lúa), *Crambus caliginosellus* (sâu kéo màng hại rế ngô), *Crambus teterrellus* (sâu kéo màng cỏ blue), *Diatraea grandiosella* (sâu đục thân ngô tây nam), *Diatraea saccharalis* (sâu đục thân mía), *Eariasinsulana* (sâu đục quả có gai), *Earias vittella* (sâu đục quả đốm), *Helicoverpa armigera* (sâu đục quả Mỹ), *Helicoverpa zea* (sâu hại ngô hoặc sâu hại bông), *Heliothis virescens* (sâu chồi thuốc lá), *Herpetogramma licarsisalis* (sâu kéo màng ở cỏ), *Lobesia botrana* (bướm hại nho châu Âu), *Phyllocnistis citrella* (sâu ăn lá quả có múi), *Pieris brassicae* (bướm trắng to), *Pieris rapae* (sâu hại cải bắp nhập khẩu, hoặc bướm trắng nhỏ), *Plutella xylostella* (sâu tơ), *Spodoptera exigua* (sâu hại củ cải đường), *Spodoptera litura* (sâu hại thuốc lá, sâu bướm cụm), và *Tuta absoluta* (sâu ăn lá cà chua).

Thuật ngữ “phân tử ADN phân lập”, hoặc thuật ngữ hoặc cụm từ tương đương, được dùng trong bản mô tả này để chỉ phân tử ADN là loại mà có mặt riêng rẽ hoặc kết hợp với các thành phần khác, nhưng không trong môi trường tự nhiên của nó. Ví dụ, các thành phần axit nucleic như trình tự mã hóa, trình tự intron, trình tự dẫn đầu chưa được dịch mã, trình tự của đoạn khởi đầu, trình tự kết thúc phiên mã, và các thành phần tương tự, mà vốn tìm thấy ở ADN của hệ gen của sinh vật không được xem là “phân lập” chừng nào thành phần này ở trong hệ gen của sinh vật này và tại vị trí trong hệ gen nơi nó được tìm thấy trong tự nhiên. Tuy nhiên, mỗi thành phần trong số các thành phần này, và các phần phụ của các thành phần này, sẽ là “phân lập” trong phạm vi của bản mô tả này chừng nào thành phần đó không nằm trong hệ gen của sinh vật này và tại vị trí trong hệ gen nơi nó được tìm thấy trong tự nhiên. Tương tự, trình tự nucleotit mã hóa protein trừ côn trùng hoặc biến thể bất kỳ có hoạt tính trừ côn trùng trong tự nhiên của protein đó sẽ là trình tự nucleotit phân lập chừng nào trình tự nucleotit đó không nằm trong ADN của vi khuẩn nơi mà trình tự mã hóa protein đó được tìm thấy trong tự nhiên. Trình tự nucleotit tổng hợp mã hóa trình tự axit amin của

protein trừ côn trùng có trong tự nhiên được xem là đã được phân lập nhằm mục đích của sáng chế. Nhằm mục đích của sáng chế, trình tự nucleotit chuyển gen bất kỳ, tức là trình tự nucleotit của ADN được cài xen vào hệ gen của các tế bào của thực vật hoặc vi khuẩn, hoặc có mặt ở vật truyền ngoài nhiễm sắc thể, xem là trình tự nucleotit phân lập phụ thuộc vào việc liệu nó có mặt hay không trong plasmid hoặc cấu trúc tương tự được dùng để biến nạp tế bào, trong hệ gen của thực vật hoặc vi khuẩn, hoặc có mặt với lượng phát hiện được ở mô, cây con, mẫu sinh học hoặc sản phẩm thương mại thu được từ thực vật hoặc vi khuẩn đó.

Như được mô tả rõ hơn trong các ví dụ thực hiện sáng chế, các vòng xử lý, thử nghiệm và lựa chọn lặp lại hơn hai nghìn (2000) biến thể trình tự axit amin của TIC844 và Cry1Da1 dẫn đến việc xác định được một số gốc axit amin nhất định mà có thể được thế, được cài xen vào hoặc được loại bỏ ra khỏi protein khung đã định để tạo ra các protein trừ côn trùng đã được xử lý mà thể hiện phổ phòng trừ bộ cánh vảy rộng hơn và/hoặc hoạt tính ức chế bộ cánh vảy tốt hơn (tức là độc hơn; cần ít protein trừ côn trùng hơn để đạt được cùng một mức độ tiêu diệt) nếu so với phổ và hoạt tính của các protein khung gốc, TIC844 hoặc Cry1Da1. Các vòng xử lý, thử nghiệm và lựa chọn lặp lại này còn dẫn đến việc xác định thay thế, cài xen hoặc loại bỏ gốc axit amin trung tính ở các protein khung TIC844 và Cry1Da1 mà không làm thay đổi phổ hoặc hoạt tính ức chế côn trùng của protein. Các gốc axit amin cụ thể ở khung TIC844 và Cry1Da1 mà có thể được cải biến để tạo ra phổ ức chế bộ cánh vảy tốt hơn và/hoặc hoạt tính ức chế bộ cánh vảy tốt hơn so với TIC844 hoặc Cry1Da1 được xác định trong bản mô tả này. Theo các phương án nhất định, protein trừ côn trùng đã được xử lý theo sáng chế có thể thể hiện hoạt tính ức chế bộ cánh vảy cao hơn khoảng tám lần hoặc nhiều hơn đối với loài sinh vật gây hại cánh vảy so với protein khung của trình tự SEQ ID NO:14 (TIC844) hoặc SEQ ID NO:2 (Cry1Da1).

Việc “xử lý” ở các vòng lặp lại này bao gồm việc xác định các gốc có liên quan ở protein khung để thay đổi nhằm tạo ra protein thử nghiệm cải biến, và việc tách dòng và biểu hiện các protein thử nghiệm cải biến thu được. Cấu trúc nguyên tử của các protein khung được sử dụng để dẫn và hỗ trợ cho phương thức lựa chọn bán ngẫu nhiên các gốc axit amin để cải biến cho việc xử lý. Việc “thử nghiệm” ở các vòng lặp lại này bao gồm việc so sánh hoạt tính đối với loài bộ cánh vảy của protein thử nghiệm đã được cải biến với protein khung gốc của nó. Việc “lựa chọn” ở các vòng lặp lại này bao gồm việc xác định các protein thử nghiệm đã được cải biến có hoạt tính tốt hơn (các biến thể cải tiến) và các gốc có liên quan mà được xử lý để tạo ra các biến thể cải

tiền này (trong bản mô tả này, các biến thể cải tiến này được gọi là “các protein trừ côn trùng đã được xử lý”).

Các ví dụ về phương pháp thử nghiệm và chọn lọc các protein trừ côn trùng đã được xử lý bao gồm dùng cùng một lượng protein thử nghiệm cải biến và protein khung (TIC844 hoặc Cry1Da1) cho côn trùng gây hại trong các điều kiện thử nghiệm được kiểm soát và đo và so sánh hiệu lực của protein thử nghiệm cải biến và protein khung. Phương pháp khác để thử nghiệm và chọn lọc các protein trừ côn trùng đã xử lý bao gồm việc xác định liều lượng protein (ví dụ, nồng độ protein trong chế độ dinh dưỡng) của protein thử nghiệm cải biến và của protein khung (TIC844 hoặc Cry1Da1) mà gây ra các đáp ứng quần thể côn trùng tương đương trong các điều kiện thử nghiệm được kiểm soát (tức là thu được đường cong đáp ứng liều lượng). Trị số đáp ứng liều lượng thô về mặt thống kê được dùng để so sánh sẽ là nồng độ gây chết trung bình ( $LC_{50}$ ) cần thiết để diệt 50% quần thể thử nghiệm hoặc nồng độ ức chế lột xác (“molting inhibition concentration -  $MIC_{50}$ ”), nồng độ trung bình cần thiết để ức chế lột xác 50%).

Theo các phương án nhất định, các protein trừ côn trùng đã được xử lý được bọc lộ theo sáng chế có ít nhất một cải biến axit amin ở các vị trí tương đối nêu dưới đây của TIC844 (SEQ ID NO:14) hoặc Cry1Da1 (SEQ ID NO:2): serin ở vị trí 282 được thế bằng lysin hoặc valin, tyrosin ở vị trí 316 được thế bằng serin, isoleuxin ở vị trí 368 được thế bằng prolin hoặc arginin, serin ở 374 được thế bằng arginin, asparagin ở vị trí 375 được thế bằng histidin, và isoleuxin ở vị trí 432 được thế bằng leuxin. Các protein trừ côn trùng đã được xử lý còn có thể có ít nhất hai, ba, bốn, hoặc nhiều thay thế hoặc loại bỏ axit amin trong cùng một trình tự protein trừ côn trùng đã được xử lý.

Các protein trừ côn trùng đã được xử lý mà có các cải biến axit amin này bao gồm các protein nêu trong SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:10, SEQ ID NO:12, SEQ ID NO:16, SEQ ID NO:18, SEQ ID NO:20, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:32, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:38, SEQ ID NO:40, và SEQ ID NO:44, và các đoạn có hoạt tính ức chế côn trùng của chúng. Mỗi protein trong số các protein trừ côn trùng đã được xử lý này có khối lượng đo được khoảng 132k-Dalton. Các đặc tính riêng lẻ của các protein khung có hoạt tính trừ côn trùng TIC844 và Cry1Da1 và các protein trừ côn trùng đã được xử lý thu được từ đó được thể hiện trong Bảng 1.

Bảng 1. Các đặc tính của TIC844, Cry1Da1 và các protein trừ côn trùng đã được xử lý

| Protein<br>(Tên/SEQ ID. NO.) | Phân tử<br>lượng<br>(Dalton) | Chiều dài<br>axit amin | Điểm<br>đẳng<br>điện | Điện tích ở<br>độ pH 7,0 | Số axit amin có<br>tính kiềm mạnh<br>(-) | Số axit amin có<br>tính axit mạnh | Số axit<br>amin kỵ<br>nước | Số axit<br>amin phân<br>cực |
|------------------------------|------------------------------|------------------------|----------------------|--------------------------|------------------------------------------|-----------------------------------|----------------------------|-----------------------------|
| Cry1Da1/NO:2                 | 132481,87                    | 1165                   | 5,087                | -39,319                  | 113                                      | 156                               | 388                        | 347                         |
| Cry1Da1_3/NO:4               | 132405,77                    | 1165                   | 5,087                | -39,318                  | 113                                      | 156                               | 388                        | 347                         |
| Cry1Da1_4/NO:6               | 132550,98                    | 1165                   | 5,112                | -38,319                  | 114                                      | 156                               | 388                        | 346                         |
| Cry1Da1_5/NO:8               | 132448,80                    | 1165                   | 5,112                | -38,318                  | 114                                      | 156                               | 387                        | 347                         |
| Cry1Da1_6/NO:10              | 132430,82                    | 1165                   | 5,112                | -38,319                  | 114                                      | 156                               | 387                        | 346                         |
| Cry1Da1_7/NO:12              | 132401,78                    | 1165                   | 5,087                | -39,318                  | 113                                      | 156                               | 388                        | 346                         |
| TIC844/NO:14                 | 129182,91                    | 1139                   | 5,026                | -39,540                  | 110                                      | 153                               | 382                        | 340                         |
| TIC844_2/NO:16               | 129129,85                    | 1139                   | 5,048                | -39,373                  | 110                                      | 153                               | 382                        | 339                         |
| TIC844_4/NO:18               | 129106,81                    | 1139                   | 5,026                | -39,539                  | 110                                      | 153                               | 382                        | 340                         |
| TIC844_5/NO:20               | 129118,08                    | 1069                   | 5,325                | -27,535                  | 105                                      | 136                               | 363                        | 321                         |
| TIC844_6/NO:22               | 129252,02                    | 1139                   | 5,050                | -38,540                  | 111                                      | 153                               | 382                        | 339                         |
| TIC844_7/NO:24               | 129149,84                    | 1139                   | 5,050                | -38,539                  | 111                                      | 153                               | 381                        | 340                         |
| TIC844_8/NO:26               | 129102,82                    | 1139                   | 5,026                | -39,539                  | 110                                      | 153                               | 382                        | 339                         |
| Cry1Da1/NO:28                | 132481,87                    | 1165                   | 5,087                | -39,319                  | 113                                      | 156                               | 388                        | 347                         |
| Cry1Da1_2.nno/NO:30          | 132552,95                    | 1166                   | 5,087                | -39,319                  | 113                                      | 156                               | 389                        | 347                         |
| Cry1Da1_3.nno/NO:32          | 132476,85                    | 1166                   | 5,087                | -39,318                  | 113                                      | 156                               | 389                        | 347                         |
| Cry1Da1_4.nno/NO:34          | 132622,06                    | 1166                   | 5,112                | -38,319                  | 114                                      | 156                               | 389                        | 346                         |
| Cry1Da1_5.nno/NO:36          | 132519,88                    | 1166                   | 5,112                | -38,318                  | 114                                      | 156                               | 388                        | 347                         |
| Cry1Da1_6.nno/NO:38          | 132501,90                    | 1166                   | 5,112                | -39,319                  | 114                                      | 156                               | 388                        | 346                         |
| Cry1Da1_7.nno/NO:40          | 132472,86                    | 1166                   | 5,087                | -39,318                  | 113                                      | 156                               | 389                        | 346                         |
| TIC844_9.nno/NO:42           | 129253,99                    | 1140                   | 5,026                | -39,540                  | 110                                      | 153                               | 383                        | 340                         |
| TIC844_11.nno/NO:44          | 129173,90                    | 1140                   | 5,026                | -39,539                  | 110                                      | 153                               | 383                        | 339                         |

Các đoạn của các protein trừ côn trùng đã được xử lý được bộc lộ theo sáng chế có thể ở dạng cắt cụt, trong đó một hoặc nhiều axit amin được loại ra khỏi đầu tận cùng N, đầu tận cùng C, phần giữa của protein này, hoặc kết hợp chúng mà không mất hoạt tính ức chế côn trùng. Các đoạn này sẽ vẫn giữ được hoạt tính ức chế côn trùng của protein trừ côn trùng gốc đã được xử lý.

Các protein mà giống với các protein trừ côn trùng đã được xử lý có thể được xác định bằng cách so sánh với nhau bằng cách áp dụng các thuật toán khác nhau trên máy tính đã biết trong lĩnh vực kỹ thuật này. Ví dụ, nhận diện trình tự axit amin của các protein liên quan đến các protein trừ côn trùng đã được xử lý có thể được phân tích bằng cách áp dụng cách duỗi thẳng Clustal W bằng cách sử dụng các thông số mặc định này: ma trận trọng lượng: blosum, điểm trừ cho việc mở khoảng cách: 10,0, điểm trừ cho việc kéo dài khoảng cách: 0,05, các khoảng cách ưa nước: bật, các gốc ưa nước: GPSNDQERK, điểm trừ khoảng cách đặc hiệu đối với gốc: bật (Thompson, *et al* (1994) *Nucleic Acids Research*, 22:4673-4680). Phần trăm nhận diện axit amin còn được tính theo kết quả của 100% nhân với (số nhận diện axit amin/chiều dài của protein đối tượng). Các thuật toán duỗi thẳng khác cũng sẵn có trong lĩnh vực kỹ thuật này và tạo ra kết quả tương tự như kết quả đạt được bằng cách áp dụng cách duỗi thẳng Clustal W.

Như được mô tả rõ hơn trong phần ví dụ thực hiện sáng chế, các trình tự tổng hợp hoặc nhân tạo mã hóa các protein khung và các protein trừ côn trùng đã xử lý được thiết kế để dùng ở thực vật. Các trình tự nucleotit tổng hợp làm ví dụ mà được thiết kế để dùng ở thực vật được thể hiện ở trình tự SEQ ID NO:27 (Cry1Da1.nno), SEQ ID NO:29 (Cry1Da1\_2.nno), SEQ ID NO:31 (Cry1Da1\_3.nno), SEQ ID NO:33 (Cry1Da1\_4.nno), SEQ ID NO:35 (Cry1Da1\_5.nno), SEQ ID NO:37 (Cry1Da1\_6.nno), SEQ ID NO:39 (Cry1Da1\_7.nno), SEQ ID NO:41 (TIC844\_9.nno) và SEQ ID NO:43 (TIC844\_11.nno).

Các catxet biểu hiện và các vật truyền chứa các trình tự nucleotit tổng hợp hoặc nhân tạo này được xây dựng và đưa vào các tế bào của cây ngô, cây bông và cây đậu tương theo các phương pháp biến nạp và các kỹ thuật đã biết trong lĩnh vực kỹ thuật này. Tế bào đã biến nạp được tái sinh ở các cây đã biến nạp mà được quan sát thấy biểu hiện protein trừ côn trùng đã được xử lý hoặc protein khung. Để thử nghiệm hoạt tính trừ vật gây hại, các thử nghiệm sinh học được thực hiện với sự có mặt của ấu

trùng sinh vật gây hại cánh vảy bằng cách sử dụng phiên lá thực vật thu được từ cây đã biến nạp.

Các chế phẩm phân tử axit nucleic tái tổ hợp mà mã hóa các protein trừ côn trùng đã được xử lý được dự định theo sáng chế. Ví dụ, protein trừ côn trùng đã được xử lý có thể được biểu hiện với cấu trúc ADN tái tổ hợp, trong đó phân tử polynucleotit với ORF mã hóa protein trừ côn trùng đã được xử lý được liên kết theo kiểu hoạt động được với các yếu tố biểu hiện di truyền như gen khởi đầu và yếu tố điều hòa bất kỳ khác cần thiết đối với sự biểu hiện ở hệ mà cấu trúc này được dự định. Các ví dụ không giới hạn phạm vi của sáng chế bao gồm đoạn khởi đầu có chức năng ở thực vật liên kết theo kiểu hoạt động được với protein trừ côn trùng tổng hợp đã được xử lý mã hóa các trình tự dùng để biểu hiện protein trừ côn trùng đã được xử lý ở thực vật hoặc đoạn khởi đầu có chức năng ở *Bt* liên kết theo kiểu hoạt động được với trình tự đã xử lý có hoạt tính trừ côn trùng mã hóa protein dùng để biểu hiện protein ở vi khuẩn *Bt* hoặc loài *Bacillus* khác. Các yếu tố khác có thể liên kết theo kiểu hoạt động được với protein trừ côn trùng đã được xử lý mã hóa các trình tự bao gồm, nhưng không bị giới hạn ở, gen tăng cường, intron, trình tự dẫn đầu chưa được dịch mã, đuôi đánh dấu cố định protein đã được mã hóa (đuôi đánh dấu HIS), các peptit dịch chuyển (tức là các peptit chuyển tiếp thể hạt, các peptit tín hiệu), các trình tự polypeptit dùng cho enzym cải biến sau dịch mã, vị trí liên kết ở ribosom, và vị trí đích ở ARN can thiệp. Các phân tử polynucleotit tái tổ hợp làm ví dụ theo sáng chế bao gồm, nhưng không bị giới hạn ở, đoạn khởi đầu khác loại liên kết theo kiểu hoạt động được với polynucleotit như SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:9, SEQ ID NO:11, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, SEQ ID NO:21, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:31, SEQ ID NO:33, SEQ ID NO:35, SEQ ID NO:37, SEQ ID NO:39 và SEQ ID NO:43, mà mã hóa polypeptit hoặc protein có trình tự axit amin đã nêu trong SEQ ID NO:4 (Cry1Da1\_3), SEQ ID NO:6 (Cry1Da1\_4), SEQ ID NO:8 (Cry1Da1\_5), SEQ ID NO:10 (Cry1Da1\_6), SEQ ID NO:12 (Cry1Da1\_7), SEQ ID NO:16 (TIC844\_2), SEQ ID NO:18 (TIC844\_4), SEQ ID NO:20 (TIC844\_5), SEQ ID NO:22 (TIC844\_6), SEQ ID NO:24 (TIC844\_7), SEQ ID NO:26 (TIC844\_8), SEQ ID NO:32 (Cry1Da1\_3.nno), SEQ ID NO:34 (Cry1Da1\_4.nno), SEQ ID NO:36 (Cry1Da1\_5.nno), SEQ ID NO:38 (Cry1Da1\_6.nno), SEQ ID NO:40 (Cry1Da1\_7.nno) và SEQ ID NO:44 (TIC844\_11.nno). Đoạn khởi đầu khác loại cũng

có thể liên kết theo kiểu hoạt động được với các trình tự mã hóa ADN tổng hợp mã hóa protein trừ côn trùng đã được xử lý thể hạt được hướng đến và protein trừ côn trùng đã được xử lý không được hướng đến. Dự định rằng các codon của phân tử axit nucleic tái tổ hợp mã hóa protein trừ côn trùng đã xử lý được bộc lộ theo sáng chế có thể được thể bằng các codon cùng nghĩa (đã biết trong lĩnh vực kỹ thuật này là thay thế câm).

Phân tử hoặc cấu trúc ADN tái tổ hợp chứa trình tự mã hóa protein có hoạt tính trừ côn trùng đã xử lý còn có thể bao gồm vùng của ADN mà mã hóa một hoặc nhiều tác nhân có độc tính mà có thể được tạo cấu hình để biểu hiện đồng thời hoặc cùng biểu hiện với trình tự ADN mã hóa protein trừ côn trùng đã được xử lý, protein khác với protein trừ côn trùng đã được xử lý, phân tử ARN sợi đôi ức chế côn trùng, hoặc protein phụ. Các protein phụ bao gồm, nhưng không bị giới hạn ở, các đồng tác nhân, các enzym, các đối tác liên kết, hoặc các tác nhân khác mà thực hiện chức năng trợ giúp hiệu quả của tác nhân ức chế côn trùng, ví dụ, bằng cách trợ giúp biểu hiện nó, ảnh hưởng đến tính ổn định của nó ở thực vật, tối ưu hóa năng lượng tự do cho việc tạo ra oligome, làm gia tăng độc tính của nó, và làm tăng phổ hoạt tính của nó. Protein phụ có thể tạo điều kiện thuận lợi cho sự hấp thụ một hoặc nhiều tác nhân ức chế côn trùng, ví dụ, hoặc tạo tiềm năng cho tác dụng độc của tác nhân có độc tính.

Phân tử hoặc cấu trúc ADN tái tổ hợp có thể được lắp ráp sao cho tất cả các protein hoặc các phân tử ARN sợi đôi được biểu hiện từ một đoạn khởi đầu hoặc từng protein hoặc phân tử ARN sợi đôi chịu sự kiểm soát của đoạn khởi đầu riêng hoặc một vài cách kết hợp của chúng. Các protein theo sáng chế có thể được biểu hiện từ một hệ biểu hiện nhiều gen, trong đó protein trừ côn trùng đã xử lý được biểu hiện từ phân đoạn nucleotit chung mà nó còn chứa các khung đọc mở và các gen khởi đầu khác, tùy theo loại hệ biểu hiện đã được chọn. Ví dụ, hệ biểu hiện nhiều gen của vi khuẩn có thể sử dụng một đoạn khởi đầu để điều khiển sự biểu hiện của các khung đọc mở đa liên kết/nối tiếp từ bên trong một operon (tức là biểu hiện đa cistron). Ví dụ khác, hệ biểu hiện nhiều gen của thực vật có thể sử dụng các catxet biểu hiện đa bội không liên kết, mỗi catxet biểu hiện protein khác hoặc tác nhân có độc tính khác như một hoặc nhiều phân tử ARN sợi đôi.

Các phân tử axit nucleic tái tổ hợp hoặc các cấu trúc ADN tái tổ hợp chứa trình tự mã hóa protein có hoạt tính trừ côn trùng đã xử lý có thể được phân phối đến các tế

bào chủ nhờ vật truyền, ví dụ, plasmit, baculovirut, nhiễm sắc thể tổng hợp, virion, cosmit, phagomit, thể thực khuẩn, hoặc vật truyền virut. Các vật truyền này có thể được dùng để đạt được sự biểu hiện ổn định hoặc tạm thời của trình tự mã hóa protein có hoạt tính trừ côn trùng đã xử lý ở tế bào chủ, hoặc sự biểu hiện tiếp theo của polypeptit được mã hóa. Polynucleotit tái tổ hợp hoặc cấu trúc ADN tái tổ hợp ngoại sinh mà chứa trình tự mã hóa trình tự protein có hoạt tính trừ côn trùng đã xử lý và được đưa vào tế bào chủ trong bản mô tả này được gọi là “gen chuyển”.

Vi khuẩn chuyển gen, các tế bào thực vật chuyển gen, các thực vật chuyển gen, và bộ phận của thực vật chuyển gen mà chứa polynucleotit mà mã hóa một hoặc nhiều protein trong số các protein trừ côn trùng bất kỳ đã được xử lý được bộc lộ theo sáng chế. Thuật ngữ “tế bào vi khuẩn” hoặc “vi khuẩn” có thể bao gồm, nhưng không bị giới hạn ở, tế bào của *Agrobacterium*, *Bacillus*, *Escherichia*, *Salmonella*, *Pseudomonas*, hoặc *Rhizobium*. Thuật ngữ “tế bào thực vật” hoặc “thực vật” có thể bao gồm, nhưng không bị giới hạn ở, tế bào hai lá mầm hoặc tế bào một lá mầm. Các cây và các tế bào thực vật được dự định bao gồm, nhưng không bị giới hạn ở, tế bào hoặc cây cỏ linh lăng, chuối, lúa mạch, đỗ, bông cải xanh, cải bắp, cải bẹ xanh, cà rốt, sắn, thầu dầu, sup lơ, cần tây, đậu gà, cải thảo, cây có múi, dừa, cà phê, ngô, cỏ ba lá, cây bông, cây thuộc họ bầu bí, cây dưa chuột, cây thông Douglas, cà tím dài, khuy nh diệp, cây lanh, tỏi, nho, hublong, tỏi tây, rau diếp, cây thông, hạt kê, quả dưa, quả hạch, yến mạch, ôliu, hành, cây cảnh, cò, cỏ cảnh, đậu Hà Lan, lạc, hạt tiêu, đậu triều, cây thông, khoai tây, cây dương, bí ngô, cây thông Radiata, củ cải, hạt cải dầu, lúa, cây thân rễ, lúa mạch đen, cây rum, cây ngải đắng, lúa miến, cây thông nam, đậu tương, rau bina, cây bí, dâu, củ cải đường, mía đường, hướng dương, ngô ngọt, gôm ngọt, khoai lang, củ mềm, chè, thuốc lá, cà chua, tiểu hắc mạch, củ mặt đất, dưa hấu, và lúa mỳ. Theo các phương án nhất định, sáng chế đề xuất các thực vật chuyển gen và bộ phận của thực vật chuyển gen tái sinh từ tế bào của thực vật chuyển gen. Theo các phương án nhất định, các thực vật chuyển gen có thể thu được từ hạt chuyển gen, bằng cách cắt, kẹp vỡ, nghiền hoặc tách theo cách khác phần này ra khỏi cây. Theo các phương án nhất định, bộ phận của thực vật có thể là hạt, quả nang, lá, hoa, thân, rễ, hoặc phần bất kỳ của chúng, hoặc phần không thể tái sinh của bộ phận của thực vật chuyển gen. Thuật ngữ phần “không thể tái sinh” của bộ phận của thực vật chuyển gen được dùng trong bản mô tả này là một phần mà không thể kích thích được để tạo ra cả

cây hoặc không thể kích thích được để tạo ra cả cây mà có khả năng sinh sản hữu tính và/hoặc sinh sản vô tính. Theo các phương án nhất định, phần không thể tái sinh của bộ phận của cây là phần của hạt, quả nang, lá, hoa, thân hoặc rễ chuyển gen.

Sáng chế đề xuất các phương pháp tạo ra thực vật chuyển gen mà chứa protein trừ côn trùng đã được xử lý với lượng ức chế bọ cánh vảy. Các cây này có thể được tạo ra bằng cách đưa polynucleotit mà nó mã hóa các protein trừ côn trùng đã được xử lý theo sáng chế vào tế bào thực vật, và chọn cây thu được từ tế bào thực vật đó mà biểu hiện lượng có tác dụng ức chế côn trùng hoặc bọ cánh vảy của protein trừ côn trùng đã được xử lý. Các cây có thể thu được từ các tế bào thực vật đó bằng cách tái sinh, hạt, phân hoa, hoặc các kỹ thuật biến nạp mô phân sinh. Các phương pháp biến nạp thực vật là đã biết trong lĩnh vực kỹ thuật này.

Cây biểu hiện protein trừ côn trùng đã được xử lý có thể được lai bằng cách nhân giống với thành phần chuyển gen biểu hiện các protein trừ côn trùng khác và/hoặc biểu hiện các tính trạng chuyển gen khác như tính trạng phòng trừ côn trùng khác, gen dung nạp thuốc diệt cỏ, gen tạo năng suất hoặc tính trạng chịu đựng stress, và các yếu tố tương tự, hoặc các tính trạng này có thể được đưa vào một vật truyền sao cho tất cả các tính trạng này được liên kết.

Các sản phẩm thực vật đã xử lý, trong đó sản phẩm đã xử lý chứa lượng phát hiện được của protein trừ côn trùng đã được xử lý, phân đoạn hoặc đoạn của nó có tác dụng ức chế côn trùng, hoặc phần khác biệt bất kỳ của chúng, cũng được bộc lộ trong bản mô tả này. Theo các phương án nhất định, sản phẩm đã xử lý được chọn từ nhóm bao gồm các bộ phận của thực vật, sinh khối thực vật, dầu, bột thô, đường, thức ăn chăn nuôi, bột, cốm, cám, thóc, vỏ, hạt đã xử lý, và hạt. Theo các phương án nhất định, sản phẩm đã xử lý không có khả năng tái sinh. Sản phẩm từ thực vật có thể là hàng hóa hoặc sản phẩm thương mại khác thu được từ thực vật chuyển gen hoặc bộ phận của thực vật chuyển gen, khi hàng hóa hoặc các sản phẩm khác có thể theo dõi được trong thương mại bằng cách phát hiện các đoạn nucleotit hoặc ARN đã được biểu hiện hoặc các protein mà mã hóa hoặc chứa phần khác biệt của protein trừ côn trùng đã được xử lý.

Các phương pháp phòng trừ côn trùng, cụ thể là sinh vật gây hại cánh vảy hại cây trồng, bằng các protein trừ côn trùng đã được xử lý cũng được bộc lộ trong bản mô tả

này. Các phương pháp này có thể bao gồm việc trồng cây chứa protein trừ côn trùng đã được xử lý với lượng ức chế côn trùng hoặc bộ cánh vảy. Theo các phương án nhất định, các phương pháp này còn có thể bao gồm một hoặc nhiều bước: (i) dùng chế phẩm bất kỳ chứa hoặc mã hóa protein trừ côn trùng đã được xử lý cho cây hoặc hạt để tạo ra cây; và (ii) biến nạp cây hoặc tế bào thực vật mà tạo ra cây bằng polynucleotit mã hóa protein trừ côn trùng đã được xử lý. Nói chung, dự định rằng protein trừ côn trùng đã được xử lý có thể được cung cấp ở dạng chế phẩm, cung cấp ở dạng vi sinh vật, hoặc cung cấp ở dạng thực vật chuyển gen để truyền hoạt tính ức chế côn trùng đối với côn trùng là bộ cánh vảy.

Theo các phương án nhất định, protein trừ côn trùng đã được xử lý là thành phần có hoạt tính trừ côn trùng của chế phẩm ức chế côn trùng tạo ra được bằng cách nuôi cấy *Bacillus* tái tổ hợp hoặc tế bào vi khuẩn tái tổ hợp bất kỳ khác đã được biến nạp để biểu hiện protein trừ côn trùng đã được xử lý trong các điều kiện thích hợp để biểu hiện. Chế phẩm này có thể được tạo ra bằng cách làm khô, làm đông khô nhanh, đông nhất hóa, chiết, lọc, ly tâm, kết lắng, hoặc cô dịch nuôi cấy các tế bào tái tổ hợp biểu hiện/tạo ra protein trừ côn trùng đã được xử lý. Quy trình này có thể tạo ra dịch chiết tế bào, huyền phù tế bào, dịch đồng nhất tế bào, dịch phân giải tế bào, dịch nổi tế bào, dịch lọc tế bào, hoặc viên tế bào *Bacillus* hoặc vi khuẩn khác gây bệnh cho côn trùng. Bằng cách thu protein trừ côn trùng đã xử lý tạo ra được theo cách đó, chế phẩm mà chứa protein trừ côn trùng đã được xử lý có thể chứa tế bào vi khuẩn, bào tử vi khuẩn, và các thể vùi của bào tử vỏ và có thể được tạo ra cho các mục đích sử dụng khác nhau, bao gồm làm các sản phẩm xịt ức chế côn trùng trong nông nghiệp hoặc làm các chế phẩm ức chế côn trùng trong thử nghiệm sinh học về chế độ ăn.

Theo một phương án, để làm giảm khả năng hình thành tính kháng, chế phẩm ức chế côn trùng hoặc thực vật chuyển gen chứa protein trừ côn trùng đã được xử lý còn có thể chứa ít nhất một tác nhân bổ sung có độc tính thể hiện hoạt tính ức chế côn trùng đối với cùng một loài côn trùng bộ cánh vảy, nhưng là khác với protein trừ côn trùng đã được xử lý. Các tác nhân bổ sung có độc tính có thể đối với chế phẩm này bao gồm protein ức chế côn trùng và phân tử ARN sợi đôi ức chế côn trùng. Một ví dụ về việc sử dụng các trình tự ribonucleotit này để phòng trừ các côn trùng gây hại được bộc lộ trong tài liệu của Baum và các đồng tác giả (công bố đơn yêu cầu cấp US 2006/0021087 A1). (Các) polypeptit bổ sung như vậy để phòng trừ sinh vật gây hại

cánh vảy có thể được chọn từ nhóm bao gồm protein ức chế côn trùng như, nhưng không bị giới hạn ở, Cry1A (US 5,880,275), Cry1Ab, Cry1Ac, Cry1A.105, Cry1Ae, Cry1B (US 10/525,318), Cry1C (US 6,033,874), Cry1D, Cry1E, Cry1F, và thể khảm Cry1A/F (các US 7,070,982; 6,962,705; và 6,713,063), Cry1G, Cry1H, Cry1I, Cry1J, Cry1K, Cry1L, Cry2A, Cry2Ab (US 7,064,249), Cry2Ae, Cry4B, Cry6, Cry7, Cry8, Cry9, Cry15, Cry43A, Cry43B, Cry51Aa1, ET66, TIC400, TIC800, TIC834, TIC1415, Vip3A, VIP3Ab, VIP3B, AXMI-001, AXMI-002, AXMI-030, AXMI-035, và AXMI-045 (US 2013-0117884 A1), AXMI-52, AXMI-58, AXMI-88, AXMI-97, AXMI-102, AXMI-112, AXMI-117, AXMI-100 (US 2013-0310543 A1), AXMI-115, AXMI-113, AXMI-005 (US 2013-0104259 A1), AXMI-134 (US 2013-0167264 A1), AXMI-150 (US 2010-0160231 A1), AXMI-184 (US 2010-0004176 A1), AXMI-196, AXMI-204, AXMI-207, AXMI-209 (US 2011-0030096 A1), AXMI-218, AXMI-220 (US 2014-0245491 A1), AXMI-221z, AXMI-222z, AXMI-223z, AXMI-224z, AXMI-225z (US 2014-0196175 A1), AXMI-238 (US 2014-0033363 A1), AXMI-270 (US 2014-0223598 A1), AXMI-345 (US 2014-0373195 A1), DIG-3 (US 2013-0219570 A1), DIG-5 (US 2010-0317569 A1), DIG-11 (US 2010-0319093 A1), AfIP-1A và các chất dẫn xuất của nó (US 2014-0033361 A1), AfIP-1B và các chất dẫn xuất của nó (US 2014-0033361 A1), PIP-1APIP-1B (US 2014-0007292 A1), PSEEN3174 (US 2014-0007292 A1), AECFG-592740 (US 2014-0007292 A1), Pput\_1063 (US 2014-0007292 A1), Pput\_1064 (US 2014-0007292 A1), GS-135 và các chất dẫn xuất của nó (US 2012-0233726 A1), GS153 và các chất dẫn xuất của nó (US 2012-0192310 A1), GS154 và các chất dẫn xuất của nó (US 2012-0192310 A1), GS155 và các chất dẫn xuất của nó (US 2012-0192310 A1), SEQ ID NO:2 và các chất dẫn xuất của nó như đã được bộc lộ trong US 2012-0167259 A1, SEQ ID NO:2 và các chất dẫn xuất của nó như đã được bộc lộ trong US 2012-0047606 A1, SEQ ID NO:2 và các chất dẫn xuất của nó như đã được bộc lộ trong US 2011-0154536 A1, SEQ ID NO:2 và các chất dẫn xuất của nó như đã được bộc lộ trong US 2011-0112013 A1, SEQ ID NO:2 và 4 và các chất dẫn xuất của nó như đã được bộc lộ trong US 2010-0192256 A1, SEQ ID NO:2 và các chất dẫn xuất của nó như đã được bộc lộ trong US 2010-0077507 A1, SEQ ID NO:2 và các chất dẫn xuất của nó như đã được bộc lộ trong US 2010-0077508 A1, SEQ ID NO:2 và các chất dẫn xuất của nó như đã được bộc lộ trong US 2009-0313721 A1, SEQ ID NO:2 hoặc 4 và các chất dẫn xuất của nó như đã được bộc lộ

trong US 2010-0269221 A1, SEQ ID NO:2 và các chất dẫn xuất của nó như đã được bộc lộ trong US 7,772,465 (B2), CF161\_0085 và các chất dẫn xuất của nó như đã được bộc lộ trong công bố đơn quốc tế số WO2014/008054 A2, các protein có độc tính đối với bọ cánh vảy và các chất dẫn xuất của chúng như đã được bộc lộ trong các công bố đơn yêu cầu cấp US US2008-0172762 A1, US2011-0055968 A1, và US2012-0117690 A1; SEQ ID NO:2 và các chất dẫn xuất của nó như đã được bộc lộ trong US 7510878(B2), SEQ ID NO:2 và các chất dẫn xuất của nó như đã được bộc lộ trong US 7812129(B1); và các protein tương tự.

Theo các phương án khác, chế phẩm ức chế côn trùng hoặc thực vật chuyển gen có thể còn chứa ít nhất một tác nhân bổ sung có độc tính thể hiện hoạt tính ức chế côn trùng đối với côn trùng gây hại mà không bị ức chế bởi các protein trừ côn trùng đã được xử lý theo sáng chế (như vật gây hại là bọ cánh cứng, bọ cánh nửa và bọ cánh đều), để mở rộng phổ ức chế côn trùng thu được.

Tác nhân bổ sung có độc tính này để phòng trừ vật gây hại là bọ cánh cứng có thể được chọn từ nhóm bao gồm protein ức chế côn trùng như, nhưng không bị giới hạn ở, Cry3Bb (US 6,501,009), các biến thể của Cry1C, các biến thể của Cry3A, Cry3, Cry3B, Cry34/35, 5307, AXMI134 (US 2013-0167264 A1) AXMI-184 (US 2010-0004176 A1), AXMI-205 (US 2014-0298538 A1), axmi207 (US 2013-0303440 A1), AXMI-218, AXMI-220 (US 20140245491A1), AXMI-221z, AXMI-223z (US 2014-0196175 A1), AXMI-279 (US 2014-0223599 A1), AXMI-R1 và các biến thể của chúng (US 2010-0197592 A1, TIC407, TIC417, TIC431, TIC807, TIC853, TIC901, TIC1201, TIC3131, DIG-10 (US 2010-0319092 A1), eHIP (US 2010/0017914), IP3 và các biến thể của chúng (US 2012-0210462 A1), và  $\alpha$ -Hexatoxin-Hv1a (US US2014-0366227 A1).

Tác nhân bổ sung có độc tính để phòng trừ vật gây hại là bọ cánh nửa có thể được chọn từ nhóm bao gồm các protein có hoạt tính đối với bọ cánh nửa như, nhưng không bị giới hạn ở, TIC1415 (US 2013-0097735 A1), TIC807 (US 8609936), TIC834 (US 2013-0269060 A1), AXMI-036 (US 2010-0137216 A1), và AXMI-171 (US 2013-0055469 A1). Các polypeptit bổ sung dùng để phòng trừ các côn trùng gây hại là bọ cánh cứng, bọ cánh vảy, và bọ cánh nửa có thể được tìm thấy trên trang web về danh pháp độc tố của *Bacillus thuringiensis* được duy trì bởi Neil Crickmore (trên trang web toàn cầu tại địa chỉ [btnomenclature.info](http://btnomenclature.info)).

Các trình tự mã hóa protein trừ côn trùng đã được xử lý và các trình tự có tỷ lệ đồng nhất đáng kể tính theo phần trăm với các protein trừ côn trùng đã được xử lý có thể được xác định bằng cách áp dụng các phương pháp mà người có hiểu biết trung bình về lĩnh vực kỹ thuật này đã biết như phản ứng chuỗi polymeraza (PCR), khuếch đại bằng nhiệt và lai. Ví dụ, protein trừ côn trùng đã được xử lý có thể được dùng để tạo ra kháng thể mà liên kết đặc hiệu với protein có liên quan, và có thể được dùng để sàng lọc và tìm ra các protein khác mà liên quan gần.

Hơn nữa, các trình tự nucleotit mã hóa các protein trừ côn trùng đã được xử lý có thể được dùng làm mẫu dò và các đoạn mồi để sàng lọc nhằm xác định các thành viên khác của nhóm này bằng cách áp dụng việc khuếch đại theo chu kỳ nhiệt hoặc đẳng nhiệt và các phương pháp lai. Ví dụ, các oligonucleotit thu được từ các trình tự đã nêu trong SEQ ID NO: 3 có thể được dùng để xác định sự có mặt hoặc sự không có mặt của gen chuyển có hoạt tính trừ côn trùng đã được xử lý ở mẫu axit deoxyribonucleic thu được từ sản phẩm thương mại. Nhờ độ nhạy của một số phương pháp phát hiện axit nucleic nhất định mà sử dụng các oligonucleotit, dự định rằng các oligonucleotit thu được từ các trình tự bất kỳ trong số các trình tự nêu trong SEQ ID NO: 3 có thể được dùng để phát hiện protein trừ côn trùng đã được xử lý tương ứng ở các sản phẩm thương mại thu được từ các nguồn gom được khi chỉ một phân đoạn của sản phẩm thương mại thu được từ thực vật chuyển gen chứa trình tự bất kỳ trong số SEQ ID NO: 3.

Các dấu hiệu khác và các ưu điểm khác của sáng chế sẽ trở nên rõ ràng nếu dựa vào các ví dụ nêu dưới đây và các điểm yêu cầu bảo hộ.

### **Ví dụ thực hiện sáng chế**

Trên cơ sở nội dung nêu trên, người có hiểu biết trung bình về lĩnh vực kỹ thuật này cần phải hiểu rằng có thể được tạo ra thay đổi theo các khía cạnh cụ thể mà đã được bộc lộ và vẫn đạt được cùng một kết quả hoặc kết quả tương tự mà không vượt quá ý tưởng và phạm vi của sáng chế. Do đó, cần hiểu rằng các chi tiết cụ thể về cấu trúc và chức năng đã được bộc lộ theo sáng chế không nhằm giới hạn phạm vi yêu cầu bảo hộ của sáng chế. Cần phải hiểu rằng toàn bộ nội dung của từng tài liệu trích dẫn được đưa vào đây bằng cách viện dẫn.

Ví dụ 1. Thiết kế protein thử nghiệm cải biến và chuẩn bị mẫu để thử nghiệm sinh học đối với côn trùng

Ví dụ này minh họa các phương pháp được thực hiện để xác định các gốc axit amin liên quan ở các protein khung để thay đổi nhằm tạo ra các protein thử nghiệm đã được cải biến, và tách dòng và biểu hiện các protein thử nghiệm đã được cải biến thu được.

Một vài kỹ thuật xử lý ở mức độ phân tử được áp dụng theo phương thức theo bậc để tạo cấu trúc biến thể cải tiến của Cry1Da1 có phổ ức chế bọ cánh vảy tốt hơn và/hoặc hoạt tính ức chế bọ cánh vảy tốt hơn so với các protein khung của Cry1Da1 và TIC844, là thể tương đồng của Cry1Da1. Bậc thứ nhất, hay vòng thiết kế đầu tiên, được thực hiện chủ yếu theo giả thuyết. Bậc thứ hai và bậc thứ ba là các vòng thiết kế được thực hiện theo thống kê. Ví dụ, ở bậc thiết kế thứ hai, các đột biến không hại theo thống kê được kết hợp với các đột biến được cho là có lợi để tạo ra các đột biến kép mà thỏa mãn một số tiêu chí thống kê xác định. Ở bậc thiết kế thứ ba, toàn bộ dữ liệu từ các thử nghiệm trước được phân tích bằng cách áp dụng nhiều phương pháp thống kê. Chỉ các đột biến thể hiện sự cải thiện đáng kể về mặt thống kê ở nhiều hơn một phương pháp thống kê được chọn vào vòng đột biến cuối cùng. Các biến thể thiết kế được ở bậc này chứa nhiều hơn một hoặc hai đột biến tích cực so với các biến thể đã từng được xác nhận là tích cực. Do đó, thiết kế bậc ba nâng cao đáng kể chất lượng của các biến thể có hoạt tính so với bậc thứ nhất và bậc thứ hai. Như được thể hiện trong các Ví dụ tiếp theo, việc áp dụng chiến lược thiết kế ba bậc này xác định cả các đột biến đơn lẻ và các đột biến hiệp đồng mà nâng cao đáng kể hoạt tính đối với CEW đối với các biến thể cải tiến nhất định so với khung TIC844 và Cry1Da1.

Phương pháp mà được áp dụng để tạo ra các protein thử nghiệm đã được cải biến bao gồm, nhưng không chỉ giới hạn ở, cải biến bán ngẫu nhiên, các cải biến có định hướng về dao động trong việc duỗi thẳng TIC844/Cry1Da1 với các protein tự nhiên khác của *Bacillus thuringiensis* (Bt), và thiết kế được trợ giúp bởi cấu trúc/chức năng. Các ví dụ về các kỹ thuật xử lý phân tử được áp dụng bao gồm các kỹ thuật dưới đây.

Liên kết thụ thể

Tính miễn cảm của sinh vật gây hại cánh vảy, đặc biệt là sâu đục trái hại ngô (CEW, *Helicoverpa zea*) đối với các biến thể Cry1Da1/TIC844 cải tiến có thể là do

các thụ thể khác nhau ở ruột được hướng đích đến. Các thiết kế mà được sử dụng để cải thiện khả năng liên kết với các thụ thể ở ruột, nhờ đó gia tăng độc tính, bao gồm: (1) gây đột biến mọi vị trí trong vòng đỉnh của miền II đều thành kiểu axit amin; và (2) đổi mọi cách kết hợp có thể của các vòng đỉnh của miền II với các cách kết hợp từ các thể tương đồng Cry1Da1 khác (ví dụ, Cry1Db1, Cry1Dc1) và các độc tố ba miền có hoạt tính đối với CEW (ví dụ, Cry1Bb1, Cry1Ja1 và Cry2Ab2).

Cách thức trên cơ sở duỗi thẳng

Việc duỗi thẳng Cry1Da1 với các thể tương đồng khác (ví dụ, Cry1Db1 và Cry1Dc1) được dùng để xác định các vùng biến dị. Kết quả của việc duỗi thẳng là xác định được một trăm năm mươi (150) vị trí và hai trăm chín mươi lăm (295) đột biến đơn nhất. Các vị trí này nằm ở cả ba miền. Các vị trí cách nhau trong vòng bốn (4) axit amin được nhóm lại với nhau. Chỉ các đột biến từ cùng các trình tự gốc được đề cử cho từng nhóm vị trí, tạo ra một trăm ba mươi hai (132) biến thể đơn nhất.

Cách gây đột biến bề mặt

Các polynucleotit mã hóa vị trí bề mặt ở miền II và miền III của các protein khung được gây đột biến bằng cách quét. Các gốc axit amin được cải biến thành alanin khi alanin không có sẵn ở protein khung. Ở các vị trí bề mặt nơi các gốc tự nhiên là lysin, các đột biến arginin được đưa vào bổ sung cho các đột biến alanin. Lý do cho đột biến lysin thành arginin là dựa trên quan sát thấy rằng các độc tố có hoạt tính đối với bộ cánh vảy thường có khuynh hướng có rất ít lysin và nhiều arginin và, do đó, đã có giả thiết rằng việc thay đổi các vị trí lysin bề mặt ở miền II và miền III thành arginin sẽ nâng cao hoạt tính đối với bộ cánh vảy của protein thử nghiệm đã được cải biến.

Thay đổi khả năng phân giải protein

Có giả thuyết rằng quá trình phân giải protein là khía cạnh quan trọng về hoạt tính của các độc tố ba miền trong ruột côn trùng bộ cánh vảy. Để thử nghiệm điều này, vài nhóm đột biến được tạo ra để có tiềm năng làm thay đổi sự phân cắt bất kỳ phân giải protein. Các vị trí phân cắt tiềm năng được định vị ở đầu tận cùng N và giữa miền III và tiền độc tố. Các vị trí đột biến bao gồm vùng vòng đã được dự đoán từ đầu tận cùng N đến phần bắt đầu của vòng xoắn 4 và từ đầu tận cùng C của miền III đến ~40 axit amin ở tiền độc tố. Nói chung, có giả thuyết rằng các gốc glyxin thúc đẩy sự phân

giải protein thông qua việc nhận diện vị trí phân giải protein hoặc bằng cách tăng mức độ linh hoạt của protein, nhờ đó khiến nó dễ bị phân giải protein hơn. Ngoài ra, trypsin và chymotrypsin là hai proteaza được chấp nhận rộng rãi làm các proteaza dễ tồn tại ở ruột giữa của bộ cánh vảy. Các gốc lysin là vị trí nhận diện cho trypsin và gốc tyrosin là vị trí nhận diện cho chymotrypsin. Do đó, vị trí đột biến đã chọn trong số các vị trí phân cắt tiềm năng được gây đột biến thành glyxin, lysin hoặc tyrosin.

Các đột biến điểm nóng tiềm năng từ các độc tố khác có hoạt tính đối với CEW

Dữ liệu về hoạt tính và không có hoạt tính đối với CEW cho một nhóm lớn protein (kể cả các thể khảm, các đoạn và các trình tự tự nhiên) được phân tích. Thông tin thu được từ việc phân tích thống kê dữ liệu này được sử dụng để xác định các đột biến cụ thể tiềm năng hoặc các vị trí cho đột biến mà có thể làm tăng hoạt tính CEW ở các protein thử nghiệm đã cải biến thu được.

Các protein thử nghiệm đã được cải biến mà được tạo ra theo các phương pháp luận xử lý phân tử nêu trên được tách dòng bằng cách áp dụng các phương pháp đã biết trong lĩnh vực kỹ thuật này vào vật truyền biểu hiện plasmid *Bt* tái tổ hợp sau đoạn khởi đầu biểu hiện đặc hiệu sự hình thành bào tử và được biến nạp vào tế bào chủ *Bt* phi kết tinh.

Ví dụ 2. Thử nghiệm các protein thử nghiệm đã cải biến trong thử nghiệm sinh học về chế độ dinh dưỡng đối với sinh vật gây hại cánh vảy

Ví dụ này minh họa việc thử nghiệm các protein thử nghiệm đã cải biến tạo ra được từ các cố gắng xử lý đã được bộc lộ trong Ví dụ 1.

Từ các cố gắng xử lý được bộc lộ trong Ví dụ 1, khoảng hai nghìn năm trăm (2500) chủng *Bt* tái tổ hợp được tạo ra mà biểu hiện nhiều hơn hai nghìn ba trăm (2300) protein thử nghiệm khác nhau đã được cải biến. Các protein thử nghiệm đã cải biến này được biểu hiện ở *Bt* và được thử nghiệm về độc tính đối với các loài bộ cánh vảy khác nhau. Các thử nghiệm cho ăn đều được thực hiện với ấu trùng mới nở (<24 giờ sau khi nở) của các loài bộ cánh vảy khác nhau, bao gồm sâu đục trái hại ngô (CEW, *Helicoverpa zea*) và sâu xanh mùa thu (FAW, *Spodoptera frugiperda*). Trùng côn trùng để thử nghiệm CEW thu được từ hai cụm phòng thí nghiệm khác nhau: Benson Research, Carlisle, PA và Monsanto Company, Union City, TN. Tất cả các protein thử nghiệm cải biến đã biểu hiện được thử trên CEW và một vài trong số các

protein thử nghiệm đã cải biến thể hiện hoạt tính tốt hơn đối với CEW so với các protein khung gốc của chúng được thử nghiệm trên FAW, ngoài việc thực hiện các thử nghiệm sinh học bổ sung để xác nhận hoạt tính đối với CEW.

Các quy trình khác nhau để thực hiện các thử nghiệm sinh học và ghi nhận tỷ lệ chết và còi cọc ở côn trùng là đã biết trong lĩnh vực kỹ thuật này. Biến thể của các phương pháp, như đã được bộc lộ trong WO 2012/139004 và trong US 7,927,598, được áp dụng.

Ví dụ 3. Các protein thử nghiệm cải biến thể hiện hoạt tính tốt hơn đối với CEW

Ví dụ này minh họa phát hiện về phổ ức chế bọ cánh vảy tốt hơn và/hoặc hoạt tính ức chế bọ cánh vảy tốt hơn hoặc cao hơn đối với một vài trong số các protein thử nghiệm đã được cải biến so với hoạt tính của các protein TIC844 hoặc Cry1Da1 khung trong nhiều vòng thử nghiệm.

Các protein thử nghiệm đã cải biến tạo ra được từ các cố gắng xử lý nêu trong Ví dụ 1 và được thử nghiệm trong các thử nghiệm sinh học đối với côn trùng như nêu trong Ví dụ 2 được thử nghiệm trong các vòng lặp lại, trong đó hoạt tính của các protein thử nghiệm đã cải biến đối với các loài bọ cánh vảy được so sánh với các protein khung gốc tương ứng của chúng (tức là TIC844 hoặc Cry1Da1). Trong vòng đầu tiên, ba trăm bảy mươi (370) protein thử nghiệm khác nhau đã cải biến thể hiện độc tính đối với CEW cao hơn so với TIC844 hoặc Cry1Da1 trong thử nghiệm sinh học về chế độ ăn. Ở từng thử nghiệm trong số các thử nghiệm sinh học về chế độ ăn này, cùng một lượng protein (protein thử nghiệm đã được cải biến hoặc protein khung) được cung cấp cho CEW trong các điều kiện thử nghiệm một liều được kiểm soát. Hiệu lực của các protein thử nghiệm đã được cải biến và các protein khung được xác định bằng cách đo và so sánh tỷ lệ chết và tỷ lệ còi cọc quan sát được ở từng thử nghiệm sinh học đối với protein thử nghiệm đã cải biến với tỷ lệ chết và tỷ lệ còi cọc quan sát được ở thử nghiệm sinh học đối với protein khung gốc.

Trong số ba trăm bảy mươi (370) protein thử nghiệm đã cải biến mà thể hiện độc tính cao hơn đối với CEW so với các protein khung trong sàng lọc thử nghiệm một liều, khoảng một trăm tám mươi (180) trong số đó được thử nghiệm tiếp trong thử nghiệm sinh học đối với FAW để xác định xem các protein thử nghiệm đã cải biến này có duy trì hoặc thể hiện hoạt tính cao hơn đối với FAW so với các gốc protein khung của chúng hay không. Khoảng bốn mươi (40) đến năm mươi (50) trong số các protein thử nghiệm đã cải biến này thể hiện hoạt tính tương tự hoặc tốt hơn đối với FAW so

với các protein khung gốc của chúng. Các protein thử nghiệm đã cải biến được sàng lọc tiếp này còn được thử nghiệm trong các thử nghiệm bổ sung đối với CEW để xác nhận hoạt tính đối với CEW. Các vòng chọn lọc và thử nghiệm các protein thử nghiệm đã được cải biến này mà thể hiện hoạt tính tốt hơn đối với CEW trong khi vẫn duy trì hoặc cải thiện hoạt tính đối với FAW tạo ra danh mục cuối cùng bao gồm các biến thể cải tiến (trong bản mô tả này còn được gọi là “các protein trừ côn trùng đã được xử lý”). Bảng 2 xác định các protein trừ côn trùng đã được xử lý này và các đột biến axit amin ở từng protein trừ côn trùng đã được xử lý. Bảng 2 còn thể hiện hoạt tính của các protein trừ côn trùng khung và đã được xử lý đối với CEW và FAW (hoạt tính trừ côn trùng được thể hiện theo trị số  $LC_{50}$  (nồng độ của độc tố cần để diệt 50% quần thể côn trùng trong khoảng thời gian tiếp cận đã được ấn định. Trị số  $LC_{50}$  càng thấp, thì độc lực càng mạnh) và trị số  $MIC_{50}$  (nồng độ cần thiết để ức chế sự lột xác đối với lứa tuổi cụ thể ở 50% ấu trùng trong khoảng thời gian tiếp cận đã được ấn định). Bảng này cho thấy rằng các protein trừ côn trùng đã được xử lý có hoạt tính tốt hơn đối với CEW, trong khi vẫn duy trì hoặc cải thiện hoạt tính đối với FAW.

Bảng 2. Các đột biến axit amin và dữ liệu về hoạt tính đối với các protein khung và các protein trừ côn trùng đã được xử lý

| <u>Protein</u><br>(Tên/SEQ<br>ID NO.) | <u>Đột biến axit amin *</u> | $LC_{50}$<br>( $\mu g/cm^2$ ) đối với cụm<br>CEW Benzon với chế<br>phẩm thử nghiệm sinh<br>học bào tử-tinh thể | $MIC_{50}$<br>( $\mu g/cm^2$ ) đối với cụm<br>CEW Benzon với chế<br>phẩm thử nghiệm sinh<br>học bào tử-tinh thể |
|---------------------------------------|-----------------------------|----------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------|
| Cry1Da1/N<br>O:2, 28                  | Không (protein khung)       | không áp dụng**                                                                                                | ~3,0                                                                                                            |
| Cry1Da1_3/<br>NO:4                    | Cry1Da1 + Y316S             | không áp dụng**                                                                                                | không áp dụng**                                                                                                 |
| Cry1Da1_4/<br>NO:6                    | Cry1Da1 + S374R             | không áp dụng**                                                                                                | không áp dụng**                                                                                                 |
| Cry1Da1_5/<br>NO:8                    | Cry1Da1 + Y316S_I368R       | không áp dụng**                                                                                                | không áp dụng**                                                                                                 |
| Cry1Da1_6/<br>NO:10                   | Cry1Da1 +S282K_Y316S_I368P  | không áp dụng**                                                                                                | không áp dụng**                                                                                                 |
| Cry1Da1_7/<br>NO:12                   | Cry1Da1+S282V_Y316S_I368P   | không áp dụng**                                                                                                | không áp dụng**                                                                                                 |
| TIC844/NO:<br>14                      | Không (protein khung)       | 41,90                                                                                                          | 3,73                                                                                                            |
| TIC844_2/N<br>O:16                    | TIC844+Y316S_N375H_I432L    | 0,81                                                                                                           | 0,65                                                                                                            |
| TIC844_4/N                            | TIC844+Y316S                | 0,98                                                                                                           | 0,57                                                                                                            |

|                    |                           |       |      |
|--------------------|---------------------------|-------|------|
| O:18               |                           |       |      |
| TIC844_5/N<br>O:20 | TIC844+S282K_Y316S_I368P  | 0,32  | 0,33 |
| TIC844_6/N<br>O:22 | TIC844+S374R              | 4,09  | 1,39 |
| TIC844_7/N<br>O:24 | TIC844+Y316S_I368R        | 0,93  | 0,61 |
| TIC844_8/N<br>O:26 | TIC844+ S282V_Y316S_I368P | 0,221 | ,064 |

\*Các đột biến axit amin được xác định bằng cách sử dụng mã axit amin chuẩn theo IUPAC. Xem tài liệu: *IUPAC-IUB Joint Commission on Biochemical Nomenclature. Nomenclature and Symbolism for Amino Acids and Peptides. Eur. J. Biochem. 138:9-37(1984)*. Chữ viết tắt trình tự axit amin đầu tiên biểu thị axit amin gốc trong protein khung đã định, số thể hiện vị trí của axit amin, và chữ viết tắt trình tự axit amin thứ hai biểu thị axit amin được đặt ở vị trí đó trong biến thể protein cải tiến.

\*\*Độc tố lõi của Cry1Da1 giống hệt độc tố lõi của TIC844.

Để thể hiện thêm phổ ức chế bọ cánh vảy tốt hơn và hoạt tính ức chế bọ cánh vảy tốt hơn của các protein trừ côn trùng đã được xử lý, khả năng làm chết của protein trừ côn trùng đã được xử lý TIC844\_8 so với protein khung gốc của nó được thể hiện trên Hình 1. Biểu đồ trên Hình 1 thể hiện các trị số MIC<sub>50</sub> của TIC844\_8 so với protein khung TIC844 đối với hai cụm CEW khác nhau, Union City và Benzon. Các kết quả thử nghiệm sinh học được thể hiện trên Hình 1 được tính từ chế phẩm thử nghiệm sinh học đã được tinh chế theo gradien sucroza. Lý do các thử nghiệm sinh học thứ cấp này được thực hiện với chế phẩm chứa protein đã được tinh chế theo gradien sucroza chứ không phải chế phẩm chứa các protein bào tử-tinh thể là để đảm bảo rằng hoạt tính đã cải thiện của TIC844\_8 vẫn được duy trì trong khoảng tinh chế rộng hơn. Ngoài ra, cụm Union City được thử nghiệm để xác nhận hoạt tính tốt hơn quan sát được ở cụm Benzon. Như được thể hiện trên Hình 1, các đột biến ở ba gốc đối với TIC844\_8 (S282V\_Y316S\_I368P), đem lại sự cải thiện lớn gấp 8 lần về khả năng gây chết CEW, so với TIC844, đối với cụm Union City và cải thiện lớn gấp 50 lần về khả năng gây chết CEW, so với TIC844, đối với cụm Benzon.

Để thể hiện tiếp nữa phổ ức chế bọ cánh vảy tốt hơn và hoạt tính ức chế bọ cánh vảy tốt hơn của các protein trừ côn trùng đã được xử lý, biên dạng hoạt tính đối với côn trùng của TIC844 và TIC844\_8 theo các nghiên cứu thử nghiệm sinh học về chế độ dinh dưỡng, được thực hiện đối với phổ rộng bao gồm các loài côn trùng bọ cánh

vảy, được thể hiện trong Bảng 3. Các côn trùng được thử nghiệm theo các nghiên cứu thử nghiệm sinh học trong Bảng 3 bao gồm sâu xám (BCW, *Agrotis ipsilon*), sâu đục trái hại ngô (CEW, *Helicoverpa zea*), sâu xanh mùa thu (FAW, *Spodoptera frugiperda*), sâu nam ăn lá (SAW, *Spodoptera eridania*), sâu đo ký sinh trên cải bắp (CLW, *Trichoplusia ni*), sâu đục thân ngô châu Âu (ECB, *Ostrinia nubilalis*), sâu đục thân ngô tây nam (SWC, *Diatraea grandiosella*), sâu chồi thuốc lá (TBW, *Heliothis virescens*), sâu bướm hại đậu mèo (VBC, *Anticarsia gemmatilis*), sâu đo đậu tương (SBL, *Chrysodeixis includes*), và sâu đục thân cây mía (SCB, *Diatraea saccharalis*). Bảng 3 này thể hiện phổ ức chế bộ cánh vảy tốt hơn của TIC844\_8 so với protein khung gốc TIC844, đặc biệt là với hoạt tính tốt hơn đối với CEW và VBC.

Bảng 3. Phổ hoạt tính đối với côn trùng của TIC844 và TIC844\_8

| SEQ ID NO. | Protein  | BCW | CEW | FAW | SAW | CLW | ECB | SWC | TBW | VBC | SBL | SCB |
|------------|----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 26         | TIC844_8 |     | *   | *   | *   | *   |     |     |     | *   | *   | *   |
| 14         | TIC844   |     |     | *   | *   | *   |     |     |     |     | *   | *   |

\*Có hoạt tính đối với loài côn trùng chỉ định.

Phổ ức chế bộ cánh vảy tốt hơn của các protein trừ côn trùng đã được xử lý còn được thể hiện trong Bảng 4 mà thể hiện biên dạng hoạt tính đối với côn trùng của một số protein trừ côn trùng nhất định đã được xử lý theo các nghiên cứu thử nghiệm sinh học về chế độ dinh dưỡng. Côn trùng đã được thử nghiệm trong các nghiên cứu thử nghiệm sinh học nêu trong Bảng 4 bao gồm sâu hại bông ở châu Âu, châu Á và châu Phi (CBW, *Helicoverpa armigera*), sâu ngài đêm hại thuốc lá (TCW, *Spodoptera litura*), sâu ăn lá cây củ cải đường (BAW, *Spodoptera exigua*), sâu hồng đục quả (PBW, *Pectinophora gossypiella*), sâu đục thân hồng (PSB, *Sesamia inferens*) và sâu đục quả đốm (SBW, *Earias vitella*). Các kết quả được thể hiện trong Bảng 4 thể hiện phổ ức chế bộ cánh vảy tốt hơn của các protein trừ côn trùng đã xử lý được liệt kê so với protein khung Cry1Da1, đặc biệt là với hoạt tính tốt hơn đối với CBW, PBW (kháng Cry1Ac), PBW (gom được từ thực địa) và SBW.

Bảng 4. So sánh biên dạng hoạt tính đối với côn trùng của Cry1Da1 và các protein trừ côn trùng đã được xử lý

| SEQ ID NO. | Protein   | CB W | TC W | BA W | PBW (nuôi trong phòng thí nghiệm) | PBW (kháng Cry1Ac) | PBW (gom trên thực địa) | PSB | SBW |
|------------|-----------|------|------|------|-----------------------------------|--------------------|-------------------------|-----|-----|
| 2          | Cry1Da1   |      | +    | +    | +                                 | +                  | +                       | +   |     |
| 12         | Cry1Da1_7 | +    | +    | +    | +                                 | +                  | +                       |     | +   |
| 18         | TIC844_4  | +    | +    | +    | +                                 | +                  |                         |     | +   |
| 20         | TIC844_5  | +    | +    | +    | +                                 | +                  |                         |     | +   |
| 24         | TIC844_7  | +    | +    | +    | +                                 | +                  |                         |     |     |

+ Có hoạt tính đối với loài côn trùng được chỉ định.

Ví dụ 4. Tổng hợp gen mã hóa protein trừ côn trùng đã được xử lý và các protein khung để biểu hiện ở thực vật

Ví dụ này minh họa quy trình tổng hợp polynucleotit mã hóa các protein trừ côn trùng đã được xử lý và các protein khung để biểu hiện ở thực vật.

Các trình tự nucleotit mã hóa các protein khung và các protein trừ côn trùng đã được xử lý để biểu hiện ở thực vật được thiết kế và tổng hợp theo các phương pháp đã được bộc lộ chung trong US 5,500,365, tránh một số trình tự nhất định có vấn đề về tính độc hại như ATTTA và các trình tự polyadenyl hóa ở thực vật giàu A/T trong khi bảo toàn trình tự axit amin của protein khung gốc hoặc protein trừ côn trùng đã được xử lý. Các trình tự nucleotit cho các gen này mã hóa các protein trừ côn trùng đã được xử lý và các protein khung để biểu hiện ở thực vật được nêu dưới đây trong Bảng 5.

Bảng 5. Các trình tự polynucleotit thiết kế được để dùng ở thực vật mã hóa protein khung và protein trừ côn trùng đã được xử lý.

| Nucleotit SEQ ID NO. | Protein       | Biến thể              |
|----------------------|---------------|-----------------------|
| 27                   | Cry1Da1.nno   | Không (protein khung) |
| 29                   | Cry1Da1_2.nno | Cry1Da1 + A2**        |
| 31                   | Cry1Da1_3.nno | Cry1Da1 + Y316S + A2  |
| 33                   | Cry1Da1_4.nno | Cry1Da1 + S374R + A2  |

|    |               |                                  |
|----|---------------|----------------------------------|
| 35 | Cry1Da1_5.nno | Cry1Da1 + S374R + A2             |
| 37 | Cry1Da1_6.nno | Cry1Da1 + S282K_Y316S_I368P + A2 |
| 39 | Cry1Da1_7.nno | Cry1Da1 + S282V_Y316S_I368P + A2 |
| 41 | TIC844_9.nno  | TIC844 + A2                      |
| 43 | TIC844_11.nno | TIC844 + S282V_Y316S_I368P + A2  |

\*\*Tên gọi của biến thể “A2” biểu thị việc cài xen gốc alanin ở vị trí axit amin 2 so với trình tự tự nhiên nhằm mục đích tách dòng vào các vật truyền biểu hiện ở thực vật.

Ví dụ 5. Các catxet biểu hiện dùng để biểu hiện các protein trừ côn trùng đã được xử lý ở thực vật

Ví dụ này minh họa cấu trúc của các catxet biểu hiện chứa các trình tự polynucleotit thiết kế được để dùng ở thực vật mà mã hóa protein khung và protein trừ côn trùng đã được xử lý.

Nhiều loại catxet biểu hiện ở thực vật được xây dựng với các trình tự polynucleotit mã hóa protein khung và protein trừ côn trùng đã được xử lý thiết kế được để biểu hiện ở thực vật được nêu trong Bảng 5. Các catxet biểu hiện này là hữu ích để biểu hiện tạm thời ở thể nguyên sinh thực vật hoặc biến nạp tế bào thực vật. Các catxet biểu hiện điển hình được thiết kế liên quan đến việc cuối cùng đặt protein vào trong tế bào. Một bộ catxet biểu hiện được thiết kế theo cách để cho phép protein được dịch mã và ở lại trong dung dịch bào tương. Một bộ catxet biểu hiện khác được thiết kế để có peptit chuyển tiếp liên kết với protein của độc tố này để cho phép hướng đích đến hạt cơ quan của tế bào như thể lập lục hoặc thể hạt. Tất cả các catxet biểu hiện được thiết kế để bắt đầu ở đầu tận cùng 5' bằng gen khởi đầu, mà có thể bao gồm nhiều thành phần đoạn khởi đầu, các thành phần gen tăng cường, hoặc các thành phần biểu hiện khác mà người có hiểu biết trung bình về lĩnh vực kỹ thuật này đã biết được liên kết theo kiểu hoạt động được để thúc đẩy sự biểu hiện của gen chuyển. Ngay tiếp sau trình tự của đoạn khởi đầu thường là một hoặc nhiều trình tự dẫn đầu 3' so với đoạn khởi đầu. Trình tự intron thường ở phía 3' so với trình tự dẫn đầu để cải thiện sự biểu hiện của gen chuyển. Trình tự mã hóa cho độc tố hoặc peptit chuyển tiếp và trình tự mã hóa cho độc tố thường nằm ở phía 3' so với đoạn khởi đầu được liên kết theo kiểu hoạt động được, gen dẫn đầu và cấu trúc intron. Trình tự 3'UTR thường nằm ở phía 3' của trình tự mã hóa để tạo điều kiện thuận lợi cho việc kết thúc sự phiên mã và

để tạo ra các trình tự quan trọng đối với việc polyadenyl hóa sản phẩm phiên mã thu được. Tất cả các yếu tố nêu trên được liên kết theo kiểu hoạt động được và bố trí liên tục, thường cùng với các trình tự bổ sung được cung cấp để xây dựng catxet biểu hiện.

Ví dụ 6. Biến nạp vật truyền chứa catxet biểu hiện protein khung hoặc protein trừ côn trùng đã được xử lý

Ví dụ này minh họa việc đưa protein khung hoặc protein trừ côn trùng đã xử lý vào mô thực vật.

Các phương pháp tạo ra thực vật chuyển gen mà biểu hiện phân đoạn axit nucleic mã hóa protein khung hoặc protein trừ côn trùng đã được xử lý có thể được thực hiện bằng cách áp dụng các biến thể của các phương pháp đã biết trong lĩnh vực kỹ thuật này. Nói chung, phương pháp này bao gồm việc biến nạp tế bào chủ thích hợp bằng phân đoạn ADN mà chứa gen khởi đầu được liên kết theo kiểu hoạt động được với vùng mã hóa mà mã hóa một hoặc nhiều protein trừ côn trùng đã được xử lý hoặc protein khung. Vùng mã hóa này thường được liên kết theo kiểu hoạt động được với vùng kết thúc phiên mã, nhờ đó gen khởi đầu này có khả năng điều khiển quá trình phiên mã của vùng mã hóa trong tế bào, và do đó làm cho tế bào có khả năng tạo ra polypeptit *in vivo*. Vật truyền, plasmid, cosmit, và phân đoạn ADN để dùng nhằm biến nạp các tế bào này thường chứa operon, gen, hoặc trình tự có nguồn gốc từ gen, dù tự nhiên hay có nguồn gốc tổng hợp, và đặc biệt là các thành phần mã hóa các protein trừ côn trùng đã xử lý được bộc lộ. Các cấu trúc ADN này có thể còn chứa các cấu trúc như gen khởi đầu, gen tăng cường, nhóm đa liên kết, hoặc các trình tự gen khác mà có thể có hoạt tính điều hòa đối với gen cụ thể đáng quan tâm. Thực vật chuyển gen, bộ phận của thực vật và các tế bào thực vật thu được được thử nghiệm về mức độ biểu hiện và hoạt tính sinh học của protein đã được mã hóa.

Các ví dụ về phương pháp mà có thể được cải biến để thu được các thực vật chuyển gen mà biểu hiện các protein có hoạt tính đối với bộ cánh vảy bao gồm các protein bộc lộ, ví dụ, các protein Cry1A (US 5,880,275), Cry1B (US 10/525318), Cry1C (US 6,033,874), thể khảm Cry1A/F (US 7,070,982; US 6,962,705, và US 6,713,063), và protein Cry2Ab (US 7,064,249).

Ví dụ 7. Hoạt tính đối với bộ cánh vảy của các protein trừ côn trùng đã được xử lý ở cây ngô đã được chuyển nhiễm một cách ổn định

Ví dụ này minh họa hoạt tính ức chế được thể hiện bởi các protein trừ côn trùng đã được xử lý đối với sinh vật gây hại cánh vảy khi được biểu hiện ở cây ngô và được cung cấp làm chế độ dinh dưỡng cho côn trùng gây hại tương ứng.

Các cây ngô chuyển gen R0 biểu hiện các protein Cry1Da1 và Cry1Da1\_7.nno được tạo ra bằng cách sử dụng các vật truyền chứa catxet biểu hiện đã được bộc lộ trong Ví dụ 6. Các cây ngô chuyển gen F1 được trồng từ hạt đã được tạo ra bằng cách thụ phấn bông của cây chứa chất mầm thương mại tự nhiên không biến nạp với phân hoa từ các thể biến nạp R0.

Các tế bào đã được biến nạp được kích thích để tạo ra cây theo các phương pháp đã biết trong lĩnh vực kỹ thuật này. Các thử nghiệm sinh học bằng cách sử dụng phiến lá thực vật được thực hiện tương tự với các thử nghiệm đã được bộc lộ trong US 8,344,207. Cây không biến nạp được dùng để thu mô làm đối chứng âm tính. Nhiều hiện tượng biến nạp từ từng vật truyền lưỡng phân được đánh giá, và các kết quả được lập bảng.

Hoạt tính trừ côn trùng của cây ngô chuyển gen biểu hiện các protein Cry1Da1 và Cry1Da1\_7.nno ở F1 và R0 được thể hiện trong Bảng 6, ngoài hoạt tính đối với cây ngô chuyển gen biểu hiện các protein Cry1Da1 và Cry1Da1\_7.nno ở F1 trên thực địa. Cụ thể, Bảng 6 thể hiện biên dạng hoạt tính đối với bộ cánh vảy của Cry1Da1\_7.nno so với protein khung gốc Cry1Da1 khi được thử nghiệm đối với CEW, FAW, và SWC. Như có thể thấy được từ Bảng 6, không giống Cry1Da1, Cry1Da1\_7.nno thể hiện hoạt tính đối với cả CEW và FAW ở các thử nghiệm sinh học R0 và F1 và các thử nghiệm F1 trên thực địa.

Bảng 6. Biên dạng hoạt tính đối với côn trùng của Cry1Da1 và Cry1Da1\_7.nno được biểu hiện ở cây ngô

| Protein<br>(SEQ ID NO.) | CEW |    |               | FAW |    |               | SWC |    |               |
|-------------------------|-----|----|---------------|-----|----|---------------|-----|----|---------------|
|                         | R0  | F1 | Trên thực địa | R0  | F1 | Trên thực địa | R0  | F1 | Trên thực địa |
| Cry1Da1<br>(28)         | -   | NT | NT            | +   | NT | NT            | -   | NT | NT            |
| Cry1Da1_7.nno<br>(40)   | +   | +  | +             | +   | +  | +             | -   | -  | -             |

+ Có hoạt tính đối với loài côn trùng; - Không có hoạt tính đối với loài côn trùng; NT Không được thử nghiệm

Ví dụ 8. Hoạt tính đối với bộ cánh vảy của protein trừ côn trùng đã được xử lý ở cây bông được chuyển nhiễm một cách ổn định

Ví dụ này minh họa hoạt tính ức chế được thể hiện bởi các protein trừ côn trùng đã được xử lý đối với sinh vật gây hại cánh vảy khi được biểu hiện ở cây bông và được cung cấp làm chế độ dinh dưỡng cho côn trùng gây hại tương ứng.

Cây bông biểu hiện các protein Cry1Da1\_7.nno và TIC844\_11.nno đã được tạo ra bằng cách sử dụng các vật truyền chứa catxet biểu hiện được bộc lộ trong Ví dụ 6. Tế bào đã biến nạp được kích thích để tạo ra cây theo các phương pháp đã biết trong lĩnh vực kỹ thuật này. Mô của lá cây bông được sử dụng trong thử nghiệm sinh học như được mô tả trong Ví dụ 7 và được thử nghiệm đối với CBW, FAW, sâu chồi thuốc lá (TBW, *Heliothis virescens*), và SBL. Bảng 7 thể hiện hoạt tính quan sát được đối với các loài bộ cánh vảy này ở cây bông thể hệ R<sub>0</sub> đã được chuyển nhiễm một cách ổn định. Như có thể thấy được từ Bảng 7, Cry1Da1\_7.nno và TIC844\_11.nno thể hiện hoạt tính đối với hai hoặc nhiều loài sinh vật gây hại cánh vảy ở cây bông thể hệ R<sub>0</sub> đã được chuyển nhiễm một cách ổn định.

Bảng 7. Biên dạng hoạt tính theo thử nghiệm sinh học của Cry1Da1\_7.nno, và TIC844\_11.nno được biểu hiện ở cây bông thể hệ R<sub>0</sub>

| Độc tố                          | CBW | FAW | TBW | SBL |
|---------------------------------|-----|-----|-----|-----|
| Cry1Da1_7.nno<br>(SEQ ID NO:40) | +   | +   | +   | +   |
| TIC844_11.nno<br>(SEQ ID NO:44) | +   | +   | -   | +   |

+ Có hoạt tính đối với loài côn trùng; - Không có hoạt tính đối với loài côn trùng.

Các hiện tượng biến nạp chọn lọc được áp dụng để tạo ra các cây R<sub>1</sub>. Các cây R<sub>1</sub> biểu hiện Cry1Da1\_7.nno được thử nghiệm về mức kháng CBW, FAW và SBL. Mô lá, mô vuông và mô quả nang được dùng trong thử nghiệm sinh học, ngoài các thử nghiệm trên thực địa được thực hiện trong nhà kính. Bảng 8 thể hiện hoạt tính quan sát được ở các thử nghiệm này. Như được thể hiện trong Bảng 8, Cry1Da1\_7.nno thể hiện hoạt tính đối với CBW, FAW và SBL trong thử nghiệm sinh học và các thử nghiệm thực địa.

Bảng 8. Biên dạng hoạt tính đối với côn trùng của Cry1Da1\_7.nno được biểu hiện ở cây bông thể hệ R<sub>1</sub>

| Độc tố                          | CBW |       |          | FAW |       |          | SBL | Nhà kính |     |
|---------------------------------|-----|-------|----------|-----|-------|----------|-----|----------|-----|
|                                 | Lá  | Vuông | Quả nang | Lá  | Vuông | Quả nang | Lá  | CBW      | FAW |
| Cry1Da1_7.nno<br>(SEQ ID NO:40) | +   | +     | +        | +   | +     | +        | +   | +        | +   |

+Có hoạt tính đối với loài côn trùng; - Không có hoạt tính đối với loài côn trùng.

Ví dụ 9. Hoạt tính đối với bộ cánh vảy của các protein trừ côn trùng đã được xử lý ở đậu tương đã được chuyển nhiễm một cách ổn định

Ví dụ này minh họa hoạt tính ức chế thể hiện bởi các protein trừ côn trùng đã được xử lý đối với sinh vật gây hại cánh vảy khi được biểu hiện ở cây đậu tương và được cung cấp làm chế độ dinh dưỡng cho côn trùng gây hại tương ứng.

Cây đậu tương biểu hiện các protein Cry1Da1\_7.nno, TIC844\_9.nno và TIC844\_11.nno được tạo ra bằng cách sử dụng các vật truyền chứa catxet biểu hiện

được bọc lộ trong Ví dụ 6. Mô lá được thu gom và sử dụng trong thử nghiệm sinh học như được mô tả trong Ví dụ 7 hoặc, theo cách khác, mô đông khô nhanh được dùng trong chế độ dinh dưỡng cho côn trùng để thử nghiệm sinh học. Thử nghiệm sinh học được thực hiện đối với các loài bọ cánh vảy khác nhau, bao gồm SAW, SBL và sâu vỏ đậu tương (soybean pod worm - SPW, *Helicoverpa zea*). Bảng 9 thể hiện hoạt tính quan sát được đối với sinh vật gây hại cánh vảy này ở đậu tương thế hệ R<sub>0</sub> đã được chuyển nhiễm một cách ổn định. Như có thể thấy được từ Bảng 9, Cry1Da1\_7.nno và TIC844\_11.nno thể hiện hoạt tính đối với SPW, SAW và SBL. TIC844\_9.nno (TIC844 cộng thêm alanin để tách dòng) không thể hiện hoạt tính đối với SPW.

Bảng 9. Biên dạng hoạt tính theo thử nghiệm sinh học của Cry1Da1\_7.nno, TIC844\_9.nno và TIC844\_11.nno được thể hiện ở đậu tương thế hệ R<sub>0</sub>

| Độc tố                       | SPW | SAW | SBL |
|------------------------------|-----|-----|-----|
| Cry1Da1_7.nno (SEQ ID NO:40) | +   | +   | +   |
| TIC844_11.nno (SEQ ID NO:44) | +   | +   | +   |
| TIC844_9.nno (SEQ ID NO:42)  | -   | +   | +   |

+ Có hoạt tính đối với loài côn trùng; - Không có hoạt tính đối với loài côn trùng.

Các hiện tượng biến nạp chọn lọc được dùng để tạo ra các cây R<sub>1</sub>. Các cây R<sub>1</sub> biểu hiện Cry1Da1\_7.nno được thử nghiệm về tính kháng đối với SAW, SBL, SPW và sâu bướm hại đậu mèo (VBC, *Anticarsia gemmatilis*). Mô lá được thu gom từ các cây thế hệ R<sub>1</sub> và được sử dụng trong thử nghiệm sinh học cho ăn. Bảng 10 thể hiện hoạt tính quan sát được ở các thử nghiệm này. Như được thể hiện trong Bảng 10, Cry1Da1\_7.nno thể hiện hoạt tính đối với SPW, SAW và SBL.

Bảng 10. Biên dạng hoạt tính theo thử nghiệm sinh học của Cry1Da1\_7.nno được biểu hiện ở đậu tương thế hệ R<sub>1</sub>

| Độc tố                       | SPW | SAW | SBL | VBC |
|------------------------------|-----|-----|-----|-----|
| Cry1Da1_7.nno (SEQ ID NO:40) | +   | +   | +   | -   |

+ Có hoạt tính đối với loài côn trùng; - Không có hoạt tính đối với loài côn trùng.

Bảng 11 thể hiện kết quả của thử nghiệm thực địa được thực hiện trong nhà kính với cây đậu tương thế hệ R<sub>1</sub> đã được chuyển nhiễm một cách ổn định biểu hiện

Cry1Da1\_7.nno. Các loài được sử dụng để phá hại thực vật trong nhà kính bao gồm sâu màu đen ăn lá (black armyworm - BLAW, *Spodoptera cosmioides*), bướm đêm gây hại đỗ (BSM, *Crociosema aporema*), sâu gây hại chồi Nam Mỹ (SAPW, *Helicoverpa gelotopoeon*), sâu đo hướng dương (SFL, *Rachiplusia nu*) và VBC. Bảng 11 thể hiện hoạt tính quan sát được ở các thử nghiệm này. Như được thể hiện trong Bảng 11, Cry1Da1\_7.nno thể hiện hoạt tính đối với BLAW, SAPW và SFL.

Bảng 11. Biên dạng hoạt tính của Cry1Da1\_7.nno được biểu hiện ở đậu tương thế hệ R<sub>1</sub> được thử nghiệm ở thử nghiệm thực địa trong nhà kính

| Độc tố                       | BLAW | BSM | SAPW | SFL | VBC |
|------------------------------|------|-----|------|-----|-----|
| Cry1Da1_7.nno (SEQ ID NO:40) | +    | -   | +    | +   | -   |

+ Có hoạt tính đối với loài côn trùng; - Không có hoạt tính đối với loài côn trùng.

Tất cả các chế phẩm và các phương pháp đã được bộc lộ và được yêu cầu bảo hộ theo sáng chế có thể được tạo ra và thực hiện mà không cần thử nghiệm quá mức khi dựa vào bản mô tả này. Trong khi các chế phẩm và các phương pháp theo sáng chế đã được bộc lộ ở các phương án minh họa nêu trên, người có hiểu biết trung bình về lĩnh vực kỹ thuật này biết rõ rằng các biến đổi, thay đổi, cải biến và sửa đổi có thể được áp dụng đối với chế phẩm, phương pháp và ở các bước hoặc trong trình tự của các bước theo các phương pháp được bộc lộ theo sáng chế, mà không vượt quá quan niệm, ý tưởng, và phạm vi thực của sáng chế. Cụ thể, một số tác nhân nhất định mà liên quan cả về mặt hóa học và về mặt sinh lý học rõ ràng có thể được thay thế cho các tác nhân được bộc lộ theo sáng chế trong khi đạt được cùng một kết quả hoặc các kết quả tương tự. Tất cả các thay thế tương tự và các cải biến đó dễ nhận thấy đối với người có hiểu biết trung bình về lĩnh vực kỹ thuật này được coi là nằm trong phạm vi, quan niệm và ý tưởng của sáng chế như được xác định bởi các điểm yêu cầu bảo hộ kèm theo.

## YÊU CẦU BẢO HỘ

1. Protein trừ côn trùng đã được xử lý chứa trình tự axit amin bất kỳ trong số các trình tự nêu trong SEQ ID NO:44, SEQ ID NO:40, SEQ ID NO:12, SEQ ID NO:26.
2. Protein trừ côn trùng đã được xử lý theo điểm 1, trong đó protein trừ côn trùng đã được xử lý này thể hiện hoạt tính ức chế chống lại các loài côn trùng thuộc bộ cánh vảy.
3. Polynucleotit mã hóa protein trừ côn trùng đã được xử lý theo điểm 1, trong đó polynucleotit này được liên kết theo kiểu hoạt động được với trình tự khởi đầu khác loại.
4. Polynucleotit mã hóa protein trừ côn trùng đã được xử lý, trong đó polynucleotit này chứa trình tự nucleotit:
  - a. bất kỳ trong số các trình tự nêu trong SEQ ID NO:43, SEQ ID NO:39, SEQ ID NO:11, hoặc SEQ ID NO:25; hoặc
  - b. mã hóa protein trừ côn trùng đã được xử lý chứa trình tự axit amin bất kỳ trong số các trình tự nêu trong SEQ ID NO:44, SEQ ID NO:40, SEQ ID NO:12, hoặc SEQ ID NO:26.
5. Tế bào chủ chứa ít nhất một polynucleotit theo điểm 3 hoặc điểm 4, trong đó tế bào chủ này được chọn từ nhóm bao gồm tế bào chủ vi khuẩn và tế bào chủ thực vật.
6. Tế bào chủ theo điểm 5, trong đó tế bào chủ vi khuẩn được chọn từ nhóm bao gồm *Agrobacterium*, *Rhizobium*, *Bacillus*, *Brevibacillus*, *Escherichia*, *Pseudomonas*, *Klebsiella* và *Erwinia*.
7. Chế phẩm ức chế côn trùng chứa protein trừ côn trùng đã được xử lý theo điểm 1.
8. Chế phẩm ức chế côn trùng theo điểm 7, trong đó chế phẩm này còn chứa thêm ít nhất một tác nhân ức chế côn trùng khác với protein trừ côn trùng đã được xử lý.
9. Chế phẩm ức chế côn trùng theo điểm 8, trong đó ít nhất một tác nhân ức chế côn trùng được chọn từ nhóm gồm protein ức chế côn trùng, phân tử dsARN ức chế côn trùng, và hóa chất ức chế côn trùng.

10. Hạt chứa:

- a. protein trừ côn trùng đã được xử lý theo điểm 1; hoặc
- b. polynucleotit mã hóa protein trừ côn trùng đã được xử lý theo điểm 1.

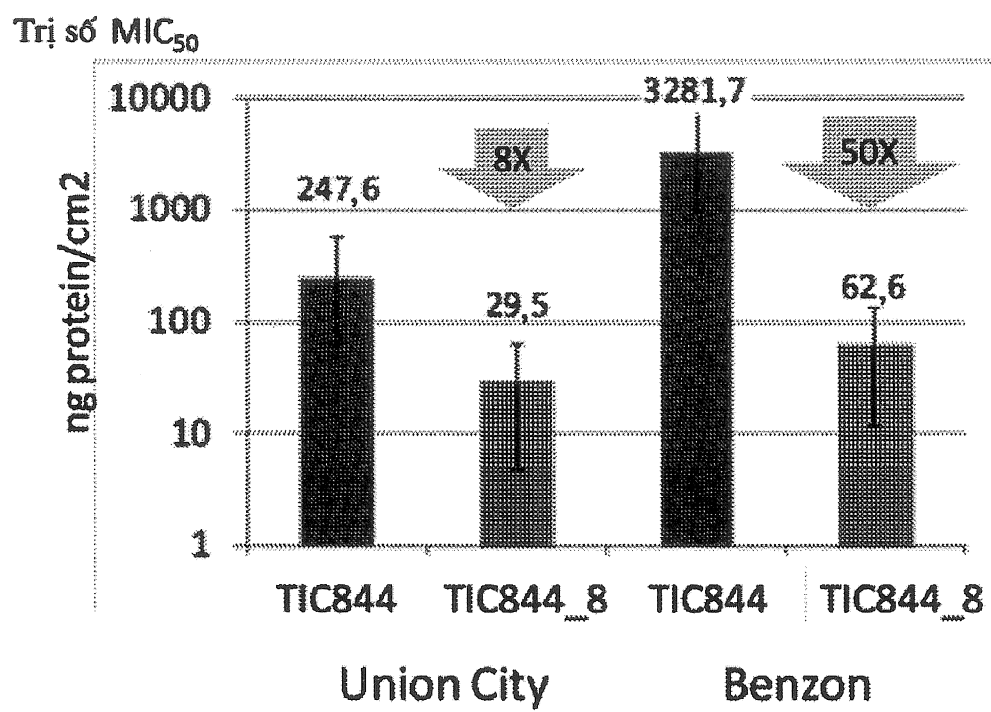
11. Phương pháp phòng trừ sinh vật gây hại cánh vảy, phương pháp này bao gồm cho sinh vật gây hại cánh vảy tiếp xúc với protein trừ côn trùng đã được xử lý theo điểm 1 với lượng có tác dụng ức chế.

12. Tế bào thực vật, thực vật, hoặc bộ phận của thực vật chuyển gen, chứa protein trừ côn trùng đã được xử lý theo điểm 1.

13. Phương pháp phòng trừ sinh vật gây hại cánh vảy, phương pháp này bao gồm cho sinh vật gây hại tiếp xúc với tế bào thực vật, thực vật, hoặc bộ phận của thực vật chuyển gen theo điểm 12, trong đó tế bào thực vật, thực vật, hoặc bộ phận của thực vật chuyển gen này biểu hiện protein trừ côn trùng đã được xử lý với lượng có tác dụng ức chế sinh vật gây hại cánh vảy.

14. Sản phẩm hàng hóa thu được từ tế bào thực vật, thực vật, hoặc bộ phận của thực vật chuyển gen theo điểm 12, trong đó sản phẩm này chứa protein trừ côn trùng đã được xử lý chứa trình tự axit amin bất kỳ trong số các trình tự nêu trong SEQ ID NO:44, SEQ ID NO:40, SEQ ID NO:12, hoặc SEQ ID NO:26 với lượng phát hiện được.

HÌNH 1



## DANH MỤC TRÌNH TỰ

<110> Monsanto Technology LLC  
 Baum, James A  
 Cerruti, Thomas  
 Flasiński, Stanislaw  
 Fu, Xiaoran  
 Howe, Arlene R  
 Salvador, Sara

<120> Protein trừ côn trùng đã được xử lý, chế phẩm ức chế côn trùng chứa protein này và phương pháp phòng trừ sinh vật gây hại cánh vảy

<130> P34223WO00/0022270.00078

<150> US 62/064994  
 <151> 2014-10-16

<150> US 62/065017  
 <151> 2014-10-17

<160> 44

<170> PatentIn phiên bản 3.5

<210> 1  
 <211> 3498  
 <212> ADN  
 <213> Bacillus thuringiensis

<220>  
 <221> đặc tính khác  
 <222> (1)..(3498)  
 <223> Trình tự nucleotit được dùng để biểu hiện ở tế bào vi khuẩn ghi mã Cry1D<sub>1</sub>.

<400> 1  
 atggaaataa ataatcaaaa ccaatgtgtg cttacaatt gtttaagtaa tcctaaggag 60  
 ataatattag gcgaggaaag gctagaaaca gggaatactg tagcagacat ttcattaggg 120  
 cttattaatt ttctatatc taattttgta ccaggaggag gatttatagt aggtttacta 180  
 gaattaatat ggggatttat agggccttcg caatgggata tttttttagc tcaaattgag 240  
 caattgatta gtcaaagaat agaagaattt gctaggaatc aggcaatttc aagattggag 300  
 gggctaagca atctttataa ggtctatgtt agagcgttta gcgactggga gaaagatcct 360  
 actaatcctg ctttaaggga agaatgcgt atacaattta atgacatgaa tagtgctctc 420  
 ataacggcta ttccactttt tagagttcaa aattatgaag ttgctctttt atctgtatat 480  
 gttcaagccg caaacttaca tttatctatt ttaagggatg tttcagtttt cggagaaaga 540  
 tggggatatg atacagcgac tatcaataat cgctatagtg atctgactag ccttattcat 600  
 gtttatacta accattgtgt ggatacgtat aatcagggat taaggcgttt ggaaggctgt 660  
 tttcttagcg attggattgt atataatcgt ttccggagac aattgacaat ttcagtatta 720  
 gatattgttg cgttttttcc aaattatgat attagaacat atccaattca aacagctact 780  
 cagctaacga ggggaagtcta tctggattta ctttttatta atgaaaatct ttctcctgca 840

|                                                                     |      |
|---------------------------------------------------------------------|------|
| gcaagctatc caaccttttc agctgctgaa agtgctataa ttagaagtcc tcatttagta   | 900  |
| gacttttttaa atagcttttac catttataca gatagtctgg cacgttatgc atattgggga | 960  |
| gggcacttgg taaattcttt ccgcacagga accactacta atttgataag atccccttta   | 1020 |
| tatggaaggg aaggaaatac agagcgcccc gtaactatta ccgcatacacc tagcgtacca  | 1080 |
| atatttagaa cactttcata tattacaggc cttgacaatt caaatcctgt agctggaatc   | 1140 |
| gagggagtgg aattccaaaa tactataagt agaagtatct atcgtaaaag cggccaata    | 1200 |
| gattctttta gtgaattacc acctcaagat gccagcgtat ctctgcaat tgggtatagt    | 1260 |
| caccgtttat gccatgcaac attttttagaa cggattagtg gaccaagaat agcaggcacc  | 1320 |
| gtattttctt ggacacaccg tagtgccagc cctactaacg aagtaagtcc atctagaatt   | 1380 |
| acacaaattc catgggtaaa ggcgcatact cttgcatctg gtgcctccgt cattaaggt    | 1440 |
| cctggattta caggtggaga tattctgact aggaatagta tgggcgagct ggggacctta   | 1500 |
| cgagtaacct tcacaggaag attaccacaa agttattata tacgtttccg ttatgcttcg   | 1560 |
| gtagcaaata ggagtgttac atttagatat tcacagccac ctctgtatgg aatttcattt   | 1620 |
| ccaaaaacta tggacgcagg tgaaccacta acatctcggt cgttcgctca tacaacactc   | 1680 |
| ttcactcaa taaccttttc acgagctcaa gaagaatttg atctatacat ccaatcgggt    | 1740 |
| gtttatatag atogaattga atttataccg gttactgcaa catttgaggc agaatatgat   | 1800 |
| ttagaaagag cgcaaaaggt ggtgaatgcc ctgtttacgt ctacaaacca actagggcta   | 1860 |
| aaaacagatg tgacggatta tcatattgat caggtatcca atctagttgc gtgtttatcg   | 1920 |
| gatgaatttt gtctggatga aaagagagaa ttgtccgaga aagttaaaca tgcaaagcga   | 1980 |
| ctcagtgatg agcggaattt acttcaagat ccaaacttca gagggatcaa taggcaacca   | 2040 |
| gaccgtggct ggagaggaag tacggatatt actatccaag gaggagatga cgtattcaaa   | 2100 |
| gagaattacg ttacgctacc gggtaacctt gatgagtgt atccaacgta tttatatcaa    | 2160 |
| aaaatagatg agtcgaaatt aaaagcctat acccgttatc aattaagagg gtatatcgaa   | 2220 |
| gatagtcaag acttagaaat ctatttaatt cgttacaatg caaaacacga aatagtaa     | 2280 |
| gtaccaggta caggaagttt atggcctctt tctgtagaaa atcaaattgg accttggtga   | 2340 |
| gaaccgaatc gatgcgcgcc acaccttgaa tggaatcctg atttacctg ttctgcaga     | 2400 |
| gacggggaaa aatgtgcaca tcattctcat catttctctt tggacattga tgttggtgt    | 2460 |
| acagacttaa atgaggactt aggtgtatgg gtgatattca agattaagac gcaagatggc   | 2520 |
| cacgcacgac tagggaatct agagtttctc gaagagaaac cattattagg agaagcacta   | 2580 |
| gctcgtgtga aaagagcgga gaaaaatgg agagacaaac gcgaaacatt acaattggaa    | 2640 |
| acaactatcg tttataaaga ggcaaaagaa tctgtagatg ctttatttgt aaactctcaa   | 2700 |
| tatgatagat tacaagcgga tacgaacatc gcgatgattc atgcggcaga taaacgcgtt   | 2760 |

```

catagaattc gagaagcgta tctgccggag ctgtctgtga ttccgggtgt caatgctggct 2820
atTTTTgaag aattagaaga gcgtatTTTt actgcatttt ccctatatga tgcgagaaat 2880
attattaaaa atggcgattt caataatggc ttattatgct ggaacgtgaa agggcatgta 2940
gaggtagaag aacaaaacaa tcaccgttca gtctgtggtta tcccagaatg ggaggcagaa 3000
gtgtcacaag aggttcgtgt ctgtccaggt cgtggctata tccttcgtgt tacagcgtac 3060
aaagagggat atggagaagg ttgcgtaacg atccatgaga tcgagaacaa tacagacgaa 3120
ctgaaattca acaactgtgt agaagaggaa gtatatccaa acaacacggg aacgtgtatt 3180
aattatactg cgactcaaga agaatatgag ggtacgtaca ctctctgtaa tcgaggatat 3240
gacgaagcct atggtaataa cccttccgta ccagctgatt atgcgtcagt ctatgaagaa 3300
aaatcgtata cagatagacg aagagagaat ccttgtgaat ctaacagagg atatggagat 3360
tacacaccac taccagctgg ttatgtaaca aaggaattag agtacttccc agagaccgat 3420
aaggatatga ttgagattgg agaaacagaa ggaacattca tcgtggacag cgtggaatta 3480
ctccttatgg aggaatag 3498

```

```

<210> 2
<211> 1165
<212> PRT
<213> Bacillus thuringiensis

<220>
<221> đặc tính_khác
<222> (1)..(1165)
<223> Trình tự axit amin của protein Cry1Dal.

```

```
<400> 2
```

```

Met Glu Ile Asn Asn Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Ser
1           5           10           15

Asn Pro Lys Glu Ile Ile Leu Gly Glu Glu Arg Leu Glu Thr Gly Asn
20           25           30

Thr Val Ala Asp Ile Ser Leu Gly Leu Ile Asn Phe Leu Tyr Ser Asn
35           40           45

Phe Val Pro Gly Gly Gly Phe Ile Val Gly Leu Leu Glu Leu Ile Trp
50           55           60

Gly Phe Ile Gly Pro Ser Gln Trp Asp Ile Phe Leu Ala Gln Ile Glu
65           70           75           80

Gln Leu Ile Ser Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala Ile
85           90           95

Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Lys Val Tyr Val Arg Ala
100          105          110

Phe Ser Asp Trp Glu Lys Asp Pro Thr Asn Pro Ala Leu Arg Glu Glu
115          120          125

Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Ile Thr Ala Ile

```

| 130                 | 135                     | 140                     |
|---------------------|-------------------------|-------------------------|
| Pro Leu Phe Arg Val | Gln Asn Tyr Glu Val     | Ala Leu Leu Ser Val Tyr |
| 145                 | 150                     | 155 160                 |
| Val Gln Ala Ala Asn | Leu His Leu Ser Ile     | Leu Arg Asp Val Ser Val |
|                     | 165                     | 170 175                 |
| Phe Gly Glu Arg Trp | Gly Tyr Asp Thr Ala Thr | Ile Asn Asn Arg Tyr     |
|                     | 180                     | 185 190                 |
| Ser Asp Leu Thr Ser | Leu Ile His Val Tyr Thr | Asn His Cys Val Asp     |
|                     | 195                     | 200 205                 |
| Thr Tyr Asn Gln Gly | Leu Arg Arg Leu Glu Gly | Arg Phe Leu Ser Asp     |
|                     | 210                     | 215 220                 |
| Trp Ile Val Tyr Asn | Arg Phe Arg Arg Gln Leu | Thr Ile Ser Val Leu     |
|                     | 225                     | 230 235 240             |
| Asp Ile Val Ala Phe | Phe Pro Asn Tyr Asp Ile | Arg Thr Tyr Pro Ile     |
|                     | 245                     | 250 255                 |
| Gln Thr Ala Thr Gln | Leu Thr Arg Glu Val Tyr | Leu Asp Leu Pro Phe     |
|                     | 260                     | 265 270                 |
| Ile Asn Glu Asn Leu | Ser Pro Ala Ala Ser Tyr | Pro Thr Phe Ser Ala     |
|                     | 275                     | 280 285                 |
| Ala Glu Ser Ala Ile | Ile Arg Ser Pro His Leu | Val Asp Phe Leu Asn     |
|                     | 290                     | 295 300                 |
| Ser Phe Thr Ile Tyr | Thr Asp Ser Leu Ala Arg | Tyr Ala Tyr Trp Gly     |
|                     | 305                     | 310 315 320             |
| Gly His Leu Val Asn | Ser Phe Arg Thr Gly Thr | Thr Thr Asn Leu Ile     |
|                     | 325                     | 330 335                 |
| Arg Ser Pro Leu Tyr | Gly Arg Glu Gly Asn Thr | Glu Arg Pro Val Thr     |
|                     | 340                     | 345 350                 |
| Ile Thr Ala Ser Pro | Ser Val Pro Ile Phe Arg | Thr Leu Ser Tyr Ile     |
|                     | 355                     | 360 365                 |
| Thr Gly Leu Asp Asn | Ser Asn Pro Val Ala Gly | Ile Glu Gly Val Glu     |
|                     | 370                     | 375 380                 |
| Phe Gln Asn Thr Ile | Ser Arg Ser Ile Tyr Arg | Lys Ser Gly Pro Ile     |
|                     | 385                     | 390 395 400             |
| Asp Ser Phe Ser Glu | Leu Pro Pro Gln Asp Ala | Ser Val Ser Pro Ala     |
|                     | 405                     | 410 415                 |
| Ile Gly Tyr Ser His | Arg Leu Cys His Ala Thr | Phe Leu Glu Arg Ile     |
|                     | 420                     | 425 430                 |
| Ser Gly Pro Arg Ile | Ala Gly Thr Val Phe Ser | Trp Thr His Arg Ser     |
|                     | 435                     | 440 445                 |
| Ala Ser Pro Thr Asn | Glu Val Ser Pro Ser Arg | Ile Thr Gln Ile Pro     |
|                     | 450                     | 455 460                 |
| Trp Val Lys Ala His | Thr Leu Ala Ser Gly Ala | Ser Val Ile Lys Gly     |
|                     | 465                     | 470 475 480             |

Pro Gly Phe Thr Gly Gly Asp Ile Leu Thr Arg Asn Ser Met Gly Glu  
 485 490 495  
 Leu Gly Thr Leu Arg Val Thr Phe Thr Gly Arg Leu Pro Gln Ser Tyr  
 500 505 510  
 Tyr Ile Arg Phe Arg Tyr Ala Ser Val Ala Asn Arg Ser Gly Thr Phe  
 515 520 525  
 Arg Tyr Ser Gln Pro Pro Ser Tyr Gly Ile Ser Phe Pro Lys Thr Met  
 530 535 540  
 Asp Ala Gly Glu Pro Leu Thr Ser Arg Ser Phe Ala His Thr Thr Leu  
 545 550 555 560  
 Phe Thr Pro Ile Thr Phe Ser Arg Ala Gln Glu Glu Phe Asp Leu Tyr  
 565 570 575  
 Ile Gln Ser Gly Val Tyr Ile Asp Arg Ile Glu Phe Ile Pro Val Thr  
 580 585 590  
 Ala Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala Gln Lys Val Val  
 595 600 605  
 Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Leu Lys Thr Asp Val  
 610 615 620  
 Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Ala Cys Leu Ser  
 625 630 635 640  
 Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys  
 645 650 655  
 His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn  
 660 665 670  
 Phe Arg Gly Ile Asn Arg Gln Pro Asp Arg Gly Trp Arg Gly Ser Thr  
 675 680 685  
 Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val  
 690 695 700  
 Thr Leu Pro Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln  
 705 710 715 720  
 Lys Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg  
 725 730 735  
 Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr  
 740 745 750  
 Asn Ala Lys His Glu Ile Val Asn Val Pro Gly Thr Gly Ser Leu Trp  
 755 760 765  
 Pro Leu Ser Val Glu Asn Gln Ile Gly Pro Cys Gly Glu Pro Asn Arg  
 770 775 780  
 Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu His Cys Ser Cys Arg  
 785 790 795 800  
 Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile  
 805 810 815

Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile  
                   820                  825                  830  
 Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu  
                   835                  840                  845  
 Phe Leu Glu Glu Lys Pro Leu Leu Gly Glu Ala Leu Ala Arg Val Lys  
                   850                  855                  860  
 Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Thr Leu Gln Leu Glu  
                   865                  870                  875                  880  
 Thr Thr Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe  
                   885                  890                  895  
 Val Asn Ser Gln Tyr Asp Arg Leu Gln Ala Asp Thr Asn Ile Ala Met  
                   900                  905                  910  
 Ile His Ala Ala Asp Lys Arg Val His Arg Ile Arg Glu Ala Tyr Leu  
                   915                  920                  925  
 Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu  
                   930                  935                  940  
 Leu Glu Glu Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn  
                   945                  950                  955                  960  
 Ile Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Leu Cys Trp Asn Val  
                   965                  970                  975  
 Lys Gly His Val Glu Val Glu Glu Gln Asn Asn His Arg Ser Val Leu  
                   980                  985                  990  
 Val Ile Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val Cys  
                   995                  1000                  1005  
 Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu Gly  
                   1010                  1015                  1020  
 Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn Thr  
                   1025                  1030                  1035  
 Asp Glu Leu Lys Phe Asn Asn Cys Val Glu Glu Glu Val Tyr Pro  
                   1040                  1045                  1050  
 Asn Asn Thr Val Thr Cys Ile Asn Tyr Thr Ala Thr Gln Glu Glu  
                   1055                  1060                  1065  
 Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly Tyr Asp Glu Ala  
                   1070                  1075                  1080  
 Tyr Gly Asn Asn Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr  
                   1085                  1090                  1095  
 Glu Glu Lys Ser Tyr Thr Asp Arg Arg Arg Glu Asn Pro Cys Glu  
                   1100                  1105                  1110  
 Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu Pro Ala Gly Tyr  
                   1115                  1120                  1125  
 Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp  
                   1130                  1135                  1140  
 Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val

|                             |      |      |  |
|-----------------------------|------|------|--|
| 1145                        | 1150 | 1155 |  |
| Glu Leu Leu Leu Met Glu Glu |      |      |  |
| 1160                        | 1165 |      |  |

<210> 3  
 <211> 3498  
 <212> ADN  
 <213> Trình tự nhân tạo

<220>  
 <223> Trình tự nucleotit được dùng để biểu hiện ở tế bào vi khuẩn ghi mã Cry1Da1\_3.

<400> 3

|                                                                     |      |
|---------------------------------------------------------------------|------|
| atggaaataa ataatcaaaa ccaatgtgtg ccttacaatt gtttaagtaa tcctaaggag   | 60   |
| ataatattag gcgaggaaaag gctagaaaaca gggaatactg tagcagacat ttcattaggg | 120  |
| cttattaatt ttctatatc taattttgta ccaggaggag gatttatagt aggtttacta    | 180  |
| gaattaatat ggggatttat agggccttcg caatgggata tttttttagc tcaaattgag   | 240  |
| caattgatta gtcaaagaat agaagaattt gctaggaatc aggcaatttc aagattggag   | 300  |
| gggctaagca atctttataa ggtctatgtt agagcggtta gcgactggga gaaagatcct   | 360  |
| actaatcctg ctttaaggga agaatgcgt atacaattta atgacatgaa tagtgctctc    | 420  |
| ataacggcta ttccactttt tagagttcaa aattatgaag ttgctctttt atctgtatat   | 480  |
| gttcaagccg caaacttaca tttatctatt ttaagggatg tttcagtttt cggagaaaga   | 540  |
| tggggatatg atacagcgac tatcaataat cgctatagtg atctgactag ccttattcat   | 600  |
| gtttatacta accatttgtt ggatacgtat aatcagggat taaggcgttt ggaaggctgt   | 660  |
| tttcttagcg attggattgt atataatcgt ttccggagac aattgacaat ttcagtatta   | 720  |
| gatattgttg cgttttttcc aaattatgat attagaacat atccaattca aacagctact   | 780  |
| cagctaacga gggaagtcta tctggattta ccttttatta atgaaaatct ttctcctgca   | 840  |
| gcaagctatc caaccttttc agctgctgaa agtgctataa ttagaagtcc tcatttagta   | 900  |
| gactttttta atagctttac catttataca gatagtctgg cacgtagtgc atattgggga   | 960  |
| gggcacttgg taaattcttt ccgcacagga accactacta atttgataag atccccctta   | 1020 |
| tatggaaggg aaggaaatac agagcgcccc gtaactatta ccgcatacacc tagcgtacca  | 1080 |
| atatttagaa cactttcata tattacaggc cttgacaatt caaatcctgt agctggaatc   | 1140 |
| gagggagtgg aattccaaaa tactataagt agaagtatct atcgtaaaag cggccaata    | 1200 |
| gattctttta gtgaattacc acctcaagat gccagcgtat ctctgcaat tgggtatagt    | 1260 |
| caccgtttat gccatgcaac attttttagaa cggattagtg gaccaagaat agcaggcacc  | 1320 |
| gtattttctt ggacacaccg tagtgccagc cctactaacg aagtaagtcc atctagaatt   | 1380 |
| acacaaattc catgggtaaa ggcgcatact cttgcatctg gtgcctccgt cattaaaggt   | 1440 |
| cctggattta caggtggaga tattctgact aggaatagta tgggcgagct ggggacctta   | 1500 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| cgagtaacct tcacaggaag attaccacaa agttattata tacgtttccg ttatgottcg  | 1560 |
| gtagcaaata ggagtgttac atttagatat tcacagccac ctctgtatgg aatttcattt  | 1620 |
| ccaaaaacta tggacgcagg tgaaccacta acatctcggt cgttcgctca tacaacactc  | 1680 |
| ttcactccaa taaccttttc acgagctcaa gaagaatttg atctatacat ccaatcgggt  | 1740 |
| gtttatatag atcgaattga atttataccg gttactgcaa catttgaggc agaatatgat  | 1800 |
| ttagaaagag cgcaaaagggt ggtgaatgcc ctgtttacgt ctacaaacca actagggcta | 1860 |
| aaaacagatg tgacggatta tcatattgat caggtatcca atctagttgc gtgtttatcg  | 1920 |
| gatgaatttt gtctggatga aaagagagaa ttgtccgaga aagttaaaca tgcaaagcga  | 1980 |
| ctcagtgatg agcggaattt acttcaagat ccaaacttca gagggatcaa taggcaacca  | 2040 |
| gaccgtggct ggagaggaag tacggatatt actatccaag gagggatga cgtattcaaa   | 2100 |
| gagaattacg ttacgctacc gggtaacctt gatgagtgtc atccaacgta tttatatcaa  | 2160 |
| aaaatagatg agtcgaaatt aaaagcctat acccgttatc aattaagagg gtatatcgaa  | 2220 |
| gatagtcaag acttagaaat ctatttaatt cgttacaatg caaaacacga aatagtaaat  | 2280 |
| gtaccaghta caggaagttt atggcctctt tctgtagaaa atcaaattgg acctgtgga   | 2340 |
| gaaccgaatc gatgcgcgcc acaccttgaa tggaatcctg atttactg ttccctgcaga   | 2400 |
| gacggggaaa aatgtgcaca tcattctcat catttctctt tggacattga tgttgatgt   | 2460 |
| acagacttaa atgaggactt aggtgtatgg gtgatattca agattaagac gcaagatggc  | 2520 |
| cacgcacgac tagggaatct agagtttctc gaagagaaac cattattagg agaagcacta  | 2580 |
| gctcgtgtga aaagagcgga gaaaaaatgg agagacaaac gcgaaacatt acaattggaa  | 2640 |
| acaactatcg tttataaaga ggcaaaagaa tctgtagatg ctttatttgt aaactctcaa  | 2700 |
| tatgatagat tacaagcgga tacgaacatc gcgatgattc atgcggcaga taaacgcgtt  | 2760 |
| catagaattc gagaagcgta tctgccggag ctgtctgtga ttccgggtgt caatgcggct  | 2820 |
| atttttgaag aattagaaga gcgtattttc actgcatttt ccctatatga tgcgagaaat  | 2880 |
| attattaaaa atggcgattt caataatggc ttattatgct ggaacgtgaa agggcatgta  | 2940 |
| gaggtagaag aacaaaacaa tcaccgttca gtccgtggtta tcccagaatg ggaggcagaa | 3000 |
| gtgtcacaag aggttcgtgt ctgtccaggt cgtggctata tccttcgtgt tacagcgtac  | 3060 |
| aaagagggat atggagaagg ttgcgtaacg atccatgaga tcgagaacaa tacagacgaa  | 3120 |
| ctgaaattca acaactgtgt agaagaggaa gtatatccaa acaacacggt aacgtgtatt  | 3180 |
| aattatactg cgactcaaga agaatatgag ggtacgtaca cttctcgtaa tcgaggatat  | 3240 |
| gacgaagcct atggtaataa cccttccgta ccagctgatt atgcgtcagt ctatgaagaa  | 3300 |
| aatcgtata cagatagacg aagagagaat ccttgtgaat ctaacagagg atatggagat   | 3360 |
| tacacaccac taccagctgg ttatgtaaca aaggaattag agtacttccc agagaccgat  | 3420 |

aaggtatgga ttgagattgg agaaacagaa ggaacattca tcgtggacag cgtggaatta 3480

ctccttatgg aggaatag 3498

<210> 4

<211> 1165

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Trình tự axit amin của protein đã được xử lý có hoạt tính trừ côn trùng Cry1Dal\_3.

<400> 4

Met Glu Ile Asn Asn Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Ser  
1 5 10 15

Asn Pro Lys Glu Ile Ile Leu Gly Glu Glu Arg Leu Glu Thr Gly Asn  
20 25 30

Thr Val Ala Asp Ile Ser Leu Gly Leu Ile Asn Phe Leu Tyr Ser Asn  
35 40 45

Phe Val Pro Gly Gly Gly Phe Ile Val Gly Leu Leu Glu Leu Ile Trp  
50 55 60

Gly Phe Ile Gly Pro Ser Gln Trp Asp Ile Phe Leu Ala Gln Ile Glu  
65 70 75 80

Gln Leu Ile Ser Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala Ile  
85 90 95

Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Lys Val Tyr Val Arg Ala  
100 105 110

Phe Ser Asp Trp Glu Lys Asp Pro Thr Asn Pro Ala Leu Arg Glu Glu  
115 120 125

Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Ile Thr Ala Ile  
130 135 140

Pro Leu Phe Arg Val Gln Asn Tyr Glu Val Ala Leu Leu Ser Val Tyr  
145 150 155 160

Val Gln Ala Ala Asn Leu His Leu Ser Ile Leu Arg Asp Val Ser Val  
165 170 175

Phe Gly Glu Arg Trp Gly Tyr Asp Thr Ala Thr Ile Asn Asn Arg Tyr  
180 185 190

Ser Asp Leu Thr Ser Leu Ile His Val Tyr Thr Asn His Cys Val Asp  
195 200 205

Thr Tyr Asn Gln Gly Leu Arg Arg Leu Glu Gly Arg Phe Leu Ser Asp  
210 215 220

Trp Ile Val Tyr Asn Arg Phe Arg Arg Gln Leu Thr Ile Ser Val Leu  
225 230 235 240

Asp Ile Val Ala Phe Phe Pro Asn Tyr Asp Ile Arg Thr Tyr Pro Ile  
245 250 255

Gln Thr Ala Thr Gln Leu Thr Arg Glu Val Tyr Leu Asp Leu Pro Phe  
 260 265 270  
 Ile Asn Glu Asn Leu Ser Pro Ala Ala Ser Tyr Pro Thr Phe Ser Ala  
 275 280 285  
 Ala Glu Ser Ala Ile Ile Arg Ser Pro His Leu Val Asp Phe Leu Asn  
 290 295 300  
 Ser Phe Thr Ile Tyr Thr Asp Ser Leu Ala Arg Ser Ala Tyr Trp Gly  
 305 310 315 320  
 Gly His Leu Val Asn Ser Phe Arg Thr Gly Thr Thr Thr Asn Leu Ile  
 325 330 335  
 Arg Ser Pro Leu Tyr Gly Arg Glu Gly Asn Thr Glu Arg Pro Val Thr  
 340 345 350  
 Ile Thr Ala Ser Pro Ser Val Pro Ile Phe Arg Thr Leu Ser Tyr Ile  
 355 360 365  
 Thr Gly Leu Asp Asn Ser Asn Pro Val Ala Gly Ile Glu Gly Val Glu  
 370 375 380  
 Phe Gln Asn Thr Ile Ser Arg Ser Ile Tyr Arg Lys Ser Gly Pro Ile  
 385 390 395 400  
 Asp Ser Phe Ser Glu Leu Pro Pro Gln Asp Ala Ser Val Ser Pro Ala  
 405 410 415  
 Ile Gly Tyr Ser His Arg Leu Cys His Ala Thr Phe Leu Glu Arg Ile  
 420 425 430  
 Ser Gly Pro Arg Ile Ala Gly Thr Val Phe Ser Trp Thr His Arg Ser  
 435 440 445  
 Ala Ser Pro Thr Asn Glu Val Ser Pro Ser Arg Ile Thr Gln Ile Pro  
 450 455 460  
 Trp Val Lys Ala His Thr Leu Ala Ser Gly Ala Ser Val Ile Lys Gly  
 465 470 475 480  
 Pro Gly Phe Thr Gly Gly Asp Ile Leu Thr Arg Asn Ser Met Gly Glu  
 485 490 495  
 Leu Gly Thr Leu Arg Val Thr Phe Thr Gly Arg Leu Pro Gln Ser Tyr  
 500 505 510  
 Tyr Ile Arg Phe Arg Tyr Ala Ser Val Ala Asn Arg Ser Gly Thr Phe  
 515 520 525  
 Arg Tyr Ser Gln Pro Pro Ser Tyr Gly Ile Ser Phe Pro Lys Thr Met  
 530 535 540  
 Asp Ala Gly Glu Pro Leu Thr Ser Arg Ser Phe Ala His Thr Thr Leu  
 545 550 555 560  
 Phe Thr Pro Ile Thr Phe Ser Arg Ala Gln Glu Glu Phe Asp Leu Tyr  
 565 570 575  
 Ile Gln Ser Gly Val Tyr Ile Asp Arg Ile Glu Phe Ile Pro Val Thr  
 580 585 590  
 Ala Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala Gln Lys Val Val

| 595 |     |     |     |     | 600 |     |     |     |     | 605 |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asn | Ala | Leu | Phe | Thr | Ser | Thr | Asn | Gln | Leu | Gly | Leu | Lys | Thr | Asp | Val |
| 610 |     |     |     |     |     | 615 |     |     |     |     | 620 |     |     |     |     |
| Thr | Asp | Tyr | His | Ile | Asp | Gln | Val | Ser | Asn | Leu | Val | Ala | Cys | Leu | Ser |
| 625 |     |     |     |     | 630 |     |     |     |     | 635 |     |     |     |     | 640 |
| Asp | Glu | Phe | Cys | Leu | Asp | Glu | Lys | Arg | Glu | Leu | Ser | Glu | Lys | Val | Lys |
|     |     |     | 645 |     |     |     |     |     | 650 |     |     |     |     | 655 |     |
| His | Ala | Lys | Arg | Leu | Ser | Asp | Glu | Arg | Asn | Leu | Leu | Gln | Asp | Pro | Asn |
|     |     |     | 660 |     |     |     |     | 665 |     |     |     |     | 670 |     |     |
| Phe | Arg | Gly | Ile | Asn | Arg | Gln | Pro | Asp | Arg | Gly | Trp | Arg | Gly | Ser | Thr |
|     | 675 |     |     |     |     | 680 |     |     |     |     |     | 685 |     |     |     |
| Asp | Ile | Thr | Ile | Gln | Gly | Gly | Asp | Asp | Val | Phe | Lys | Glu | Asn | Tyr | Val |
| 690 |     |     |     |     | 695 |     |     |     |     |     | 700 |     |     |     |     |
| Thr | Leu | Pro | Gly | Thr | Phe | Asp | Glu | Cys | Tyr | Pro | Thr | Tyr | Leu | Tyr | Gln |
| 705 |     |     |     |     | 710 |     |     |     |     | 715 |     |     |     |     | 720 |
| Lys | Ile | Asp | Glu | Ser | Lys | Leu | Lys | Ala | Tyr | Thr | Arg | Tyr | Gln | Leu | Arg |
|     |     |     |     | 725 |     |     |     |     | 730 |     |     |     |     | 735 |     |
| Gly | Tyr | Ile | Glu | Asp | Ser | Gln | Asp | Leu | Glu | Ile | Tyr | Leu | Ile | Arg | Tyr |
|     |     | 740 |     |     |     |     | 745 |     |     |     |     |     | 750 |     |     |
| Asn | Ala | Lys | His | Glu | Ile | Val | Asn | Val | Pro | Gly | Thr | Gly | Ser | Leu | Trp |
|     | 755 |     |     |     |     |     | 760 |     |     |     |     | 765 |     |     |     |
| Pro | Leu | Ser | Val | Glu | Asn | Gln | Ile | Gly | Pro | Cys | Gly | Glu | Pro | Asn | Arg |
|     | 770 |     |     |     | 775 |     |     |     |     |     | 780 |     |     |     |     |
| Cys | Ala | Pro | His | Leu | Glu | Trp | Asn | Pro | Asp | Leu | His | Cys | Ser | Cys | Arg |
| 785 |     |     |     |     | 790 |     |     |     |     | 795 |     |     |     |     | 800 |
| Asp | Gly | Glu | Lys | Cys | Ala | His | His | Ser | His | His | Phe | Ser | Leu | Asp | Ile |
|     |     |     |     | 805 |     |     |     |     | 810 |     |     |     |     | 815 |     |
| Asp | Val | Gly | Cys | Thr | Asp | Leu | Asn | Glu | Asp | Leu | Gly | Val | Trp | Val | Ile |
|     |     |     | 820 |     |     |     |     | 825 |     |     |     |     | 830 |     |     |
| Phe | Lys | Ile | Lys | Thr | Gln | Asp | Gly | His | Ala | Arg | Leu | Gly | Asn | Leu | Glu |
|     | 835 |     |     |     |     |     | 840 |     |     |     |     | 845 |     |     |     |
| Phe | Leu | Glu | Glu | Lys | Pro | Leu | Leu | Gly | Glu | Ala | Leu | Ala | Arg | Val | Lys |
|     | 850 |     |     |     |     | 855 |     |     |     |     | 860 |     |     |     |     |
| Arg | Ala | Glu | Lys | Lys | Trp | Arg | Asp | Lys | Arg | Glu | Thr | Leu | Gln | Leu | Glu |
| 865 |     |     |     |     | 870 |     |     |     |     | 875 |     |     |     |     | 880 |
| Thr | Thr | Ile | Val | Tyr | Lys | Glu | Ala | Lys | Glu | Ser | Val | Asp | Ala | Leu | Phe |
|     |     |     | 885 |     |     |     |     |     | 890 |     |     |     |     | 895 |     |
| Val | Asn | Ser | Gln | Tyr | Asp | Arg | Leu | Gln | Ala | Asp | Thr | Asn | Ile | Ala | Met |
|     |     |     | 900 |     |     |     |     | 905 |     |     |     |     | 910 |     |     |
| Ile | His | Ala | Ala | Asp | Lys | Arg | Val | His | Arg | Ile | Arg | Glu | Ala | Tyr | Leu |
|     | 915 |     |     |     |     |     | 920 |     |     |     |     | 925 |     |     |     |
| Pro | Glu | Leu | Ser | Val | Ile | Pro | Gly | Val | Asn | Ala | Ala | Ile | Phe | Glu | Glu |
|     | 930 |     |     |     |     | 935 |     |     |     |     | 940 |     |     |     |     |

Leu Glu Glu Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn  
945 950 955 960

Ile Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Leu Cys Trp Asn Val  
965 970 975

Lys Gly His Val Glu Val Glu Glu Gln Asn Asn His Arg Ser Val Leu  
980 985 990

Val Ile Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val Cys  
995 1000 1005

Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu Gly  
1010 1015 1020

Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn Thr  
1025 1030 1035

Asp Glu Leu Lys Phe Asn Asn Cys Val Glu Glu Glu Val Tyr Pro  
1040 1045 1050

Asn Asn Thr Val Thr Cys Ile Asn Tyr Thr Ala Thr Gln Glu Glu  
1055 1060 1065

Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly Tyr Asp Glu Ala  
1070 1075 1080

Tyr Gly Asn Asn Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr  
1085 1090 1095

Glu Glu Lys Ser Tyr Thr Asp Arg Arg Arg Glu Asn Pro Cys Glu  
1100 1105 1110

Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu Pro Ala Gly Tyr  
1115 1120 1125

Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp  
1130 1135 1140

Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val  
1145 1150 1155

Glu Leu Leu Leu Met Glu Glu  
1160 1165

<210> 5

<211> 3498

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Trình tự nucleotit được dùng để biểu hiện ở tế bào vi khuẩn  
ghi mã Cry1Da1\_4.

<400> 5

atggaaataa ataatcaaaa ccaatgtgtg ccttacaatt gtttaagtaa tcctaaggag 60

ataatattag gcgaggaaag gctagaaaaca gggaatactg tagcagacat ttcattaggg 120

cttattaatt ttctatatc taattttgta ccaggaggag gatttatagt aggtttacta 180

gaattaatat ggggatttat agggccttcg caatgggata ttttttttagc tcaaattgag 240

|                                   |                                      |      |
|-----------------------------------|--------------------------------------|------|
| caattgatta gtcaaagaat agaagaat    | ttt gctaggaatc aggcaatttc aagattggag | 300  |
| gggctaagca atctttataa ggtctatgtt  | agagcgttta gcgactggga gaaagatcct     | 360  |
| actaatcctg ctttaaggga agaaatgcgt  | atacaattta atgacatgaa tagtgctctc     | 420  |
| ataacggcta ttccactttt tagagttcaa  | aattatgaag ttgctctttt atctgtatat     | 480  |
| gttcaagccg caaacttaca tttatctatt  | ttaagggatg tttcagtttt cggagaaaga     | 540  |
| tggggatag atacagcgac tatcaataat   | cgctatagtg atctgactag ccttattcat     | 600  |
| gtttatacta accattgtgt ggatacgtat  | aatcagggat taaggcgttt ggaaggctgt     | 660  |
| tttcttagcg attggattgt atataatcgt  | ttccggagac aattgacaat ttcagtatta     | 720  |
| gatattgttg cgttttttcc aaattatgat  | attagaacat atccaattca aacagctact     | 780  |
| cagctaacga ggggaagtcta tctggattta | ccttttatta atgaaaatct ttctcctgca     | 840  |
| gcaagctatc caaccttttc agctgctgaa  | agtgcataaa ttagaagtcc tcatttagta     | 900  |
| gactttttta atagctttac catttataca  | gatagtctgg cacgttatgc atattgggga     | 960  |
| gggcacttgg taaattcttt ccgcacagga  | accactacta atttgataag atccccctta     | 1020 |
| tatggaaggg aaggaaatac agagcgcccc  | gtaactatta ccgcatacacc tagcgtacca    | 1080 |
| atatttagaa cactttcata tattacaggc  | cttgacaatc gtaatcctgt agctggaatc     | 1140 |
| gagggagtgg aattccaaaa tactataagt  | agaagtatct atcgtaaaag cggccaata      | 1200 |
| gattctttta gtgaattacc acctcaagat  | gccagcgatc ctctcgcaat tgggtatagt     | 1260 |
| caccgtttat gccatgcaac attttttagaa | cggattagtg gaccaagaat agcaggcacc     | 1320 |
| gtattttctt ggacacaccg tagtgccagc  | cctactaacg aagtaagtcc atctagaatt     | 1380 |
| acacaaatc catgggtaaa gggcatact    | cttgcatctg gtgcctccgt cattaaaggt     | 1440 |
| cctggattta caggtggaga tattctgact  | aggaatagta tgggcgagct ggggacctta     | 1500 |
| cgagtaacct tcacaggaag attaccacaa  | agttattata tacgtttccg ttatgcttcg     | 1560 |
| gtagcaaata ggagtggtag atttagatat  | tcacagccac ctctgtatgg aatttcattt     | 1620 |
| ccaaaaacta tggacgcagg tgaaccacta  | acatctcggt cgttcgctca tacaacactc     | 1680 |
| ttcactccaa taaccttttc acgagctcaa  | gaagaatttg atctatacat ccaatcgggt     | 1740 |
| gtttatatag atcgaattga atttataccg  | gttactgcaa catttgaggc agaatatgat     | 1800 |
| ttagaaagag cgcaaaaggt ggtgaatgcc  | ctgtttacgt ctacaaacca actagggcta     | 1860 |
| aaaacagatg tgacggatta tcatattgat  | caggtatcca atctagttgc gtgtttatcg     | 1920 |
| gatgaatttt gtctggatga aaagagagaa  | ttgtccgaga aagttaaaca tgcaaagcga     | 1980 |
| ctcagtgatg agcggaattt acttcaagat  | ccaaacttca gagggatcaa taggcaacca     | 2040 |
| gaccgtggct ggagaggaag tacggatatt  | actatccaag gaggagatga cgtattcaaa     | 2100 |
| gagaattacg ttacgctacc ggttaccttt  | gatgagtgct atccaacgta tttatatcaa     | 2160 |

```

aaaatagatg agtcgaaatt aaaagcctat acccgttatc aattaagagg gtatatcgaa      2220
gatagtcaag acttagaaat ctatttaatt cgttacaatg caaaacacga aatagtaa      2280
gtaccaggta caggaagttt atggcctctt tctgtagaaa atcaaattgg accttggtga      2340
gaaccgaatc gatgcgcgcc acaccttgaa tggaatcctg atttacctg ttctgcaga      2400
gacggggaaa aatgtgcaca tcattctcat catttctctt tggacattga tgttgatgt      2460
acagacttaa atgaggactt aggtgtatgg gtgatattca agattaagac gcaagatggc      2520
cacgcacgac tagggaatct agagtttctc gaagagaaac cattattagg agaagcacta      2580
gctcgtgtga aaagagcggg gaaaaaatgg agagacaaac gcgaaacatt acaattggaa      2640
acaactatcg ttataaaaga ggcaaaagaa tctgtagatg ctttatttgt aaactctcaa      2700
tatgatagat tacaagcggg tacgaacatc gcgatgattc atgcggcaga taaacgcgtt      2760
catagaattc gagaagcgtg tctgccggag ctgtctgtga ttccgggtgt caatgcggct      2820
atttttgaag aattagaaga gcgtattttc actgcatttt ccctatatga tgcgagaaat      2880
attattaaaa atggcgattt caataatggc ttattatgct ggaacgtgaa agggcatgta      2940
gaggtagaag aacaaaacaa tcaccgttca gtcttggtta tcccagaatg ggaggcagaa      3000
gtgtcacaag aggttcgtgt ctgtccaggt cgtggctata tccttcgtgt tacagcgtac      3060
aaagagggat atggagaagg ttgcgtaacg atocatgaga tcgagaacaa tacagacgaa      3120
ctgaaattca acaactgtgt agaagaggaa gtatatccaa acaacacggg aacgtgtatt      3180
aattatactg cgactcaaga agaatatgag ggtacgtaca cttctcgtaa tcgaggatat      3240
gacgaagcct atggtaataa cccttccgta ccagctgatt atgcgtcagt ctatgaagaa      3300
aaatcgtata cagatagacg aagagagaat ccttgtgaat ctaacagagg atatggagat      3360
tacacaccac taccagctgg ttatgtaaca aaggaattag agtacttccc agagaccgat      3420
aaggtatgga ttgagattgg agaaacagaa ggaacattca tcgtggacag cgtggaatta      3480
ctccttatgg aggaatag                                     3498

```

<210> 6

<211> 1165

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Trình tự axit amin của protein đã được xử lý có hoạt tính trừ côn trùng Cry1Da1\_4.

<400> 6

```

Met Glu Ile Asn Asn Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Ser
1           5           10           15

```

```

Asn Pro Lys Glu Ile Ile Leu Gly Glu Glu Arg Leu Glu Thr Gly Asn
20           25           30

```

```

Thr Val Ala Asp Ile Ser Leu Gly Leu Ile Asn Phe Leu Tyr Ser Asn

```

| 35  |     |     |     |     | 40  |     |     |     |     | 45  |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Phe | Val | Pro | Gly | Gly | Gly | Phe | Ile | Val | Gly | Leu | Leu | Glu | Leu | Ile | Trp |
| 50  |     |     |     |     |     | 55  |     |     |     |     | 60  |     |     |     |     |
| Gly | Phe | Ile | Gly | Pro | Ser | Gln | Trp | Asp | Ile | Phe | Leu | Ala | Gln | Ile | Glu |
| 65  |     |     |     |     | 70  |     |     |     |     | 75  |     |     |     |     | 80  |
| Gln | Leu | Ile | Ser | Gln | Arg | Ile | Glu | Glu | Phe | Ala | Arg | Asn | Gln | Ala | Ile |
|     |     |     |     | 85  |     |     |     |     | 90  |     |     |     |     | 95  |     |
| Ser | Arg | Leu | Glu | Gly | Leu | Ser | Asn | Leu | Tyr | Lys | Val | Tyr | Val | Arg | Ala |
|     |     |     | 100 |     |     |     |     | 105 |     |     |     |     | 110 |     |     |
| Phe | Ser | Asp | Trp | Glu | Lys | Asp | Pro | Thr | Asn | Pro | Ala | Leu | Arg | Glu | Glu |
|     |     | 115 |     |     |     |     | 120 |     |     |     |     | 125 |     |     |     |
| Met | Arg | Ile | Gln | Phe | Asn | Asp | Met | Asn | Ser | Ala | Leu | Ile | Thr | Ala | Ile |
| 130 |     |     |     |     |     | 135 |     |     |     |     | 140 |     |     |     |     |
| Pro | Leu | Phe | Arg | Val | Gln | Asn | Tyr | Glu | Val | Ala | Leu | Leu | Ser | Val | Tyr |
| 145 |     |     |     |     | 150 |     |     |     |     | 155 |     |     |     |     | 160 |
| Val | Gln | Ala | Ala | Asn | Leu | His | Leu | Ser | Ile | Leu | Arg | Asp | Val | Ser | Val |
|     |     |     |     | 165 |     |     |     |     | 170 |     |     |     |     | 175 |     |
| Phe | Gly | Glu | Arg | Trp | Gly | Tyr | Asp | Thr | Ala | Thr | Ile | Asn | Asn | Arg | Tyr |
|     |     |     | 180 |     |     |     |     | 185 |     |     |     |     |     | 190 |     |
| Ser | Asp | Leu | Thr | Ser | Leu | Ile | His | Val | Tyr | Thr | Asn | His | Cys | Val | Asp |
|     |     | 195 |     |     |     |     | 200 |     |     |     |     | 205 |     |     |     |
| Thr | Tyr | Asn | Gln | Gly | Leu | Arg | Arg | Leu | Glu | Gly | Arg | Phe | Leu | Ser | Asp |
|     | 210 |     |     |     |     | 215 |     |     |     |     | 220 |     |     |     |     |
| Trp | Ile | Val | Tyr | Asn | Arg | Phe | Arg | Arg | Gln | Leu | Thr | Ile | Ser | Val | Leu |
| 225 |     |     |     |     | 230 |     |     |     |     | 235 |     |     |     |     | 240 |
| Asp | Ile | Val | Ala | Phe | Phe | Pro | Asn | Tyr | Asp | Ile | Arg | Thr | Tyr | Pro | Ile |
|     |     |     |     | 245 |     |     |     |     | 250 |     |     |     |     | 255 |     |
| Gln | Thr | Ala | Thr | Gln | Leu | Thr | Arg | Glu | Val | Tyr | Leu | Asp | Leu | Pro | Phe |
|     |     |     | 260 |     |     |     |     | 265 |     |     |     |     | 270 |     |     |
| Ile | Asn | Glu | Asn | Leu | Ser | Pro | Ala | Ala | Ser | Tyr | Pro | Thr | Phe | Ser | Ala |
|     |     | 275 |     |     |     |     | 280 |     |     |     |     | 285 |     |     |     |
| Ala | Glu | Ser | Ala | Ile | Ile | Arg | Ser | Pro | His | Leu | Val | Asp | Phe | Leu | Asn |
|     | 290 |     |     |     |     | 295 |     |     |     |     | 300 |     |     |     |     |
| Ser | Phe | Thr | Ile | Tyr | Thr | Asp | Ser | Leu | Ala | Arg | Tyr | Ala | Tyr | Trp | Gly |
| 305 |     |     |     |     | 310 |     |     |     |     | 315 |     |     |     |     | 320 |
| Gly | His | Leu | Val | Asn | Ser | Phe | Arg | Thr | Gly | Thr | Thr | Thr | Asn | Leu | Ile |
|     |     |     |     | 325 |     |     |     |     | 330 |     |     |     |     | 335 |     |
| Arg | Ser | Pro | Leu | Tyr | Gly | Arg | Glu | Gly | Asn | Thr | Glu | Arg | Pro | Val | Thr |
|     |     |     | 340 |     |     |     |     | 345 |     |     |     |     | 350 |     |     |
| Ile | Thr | Ala | Ser | Pro | Ser | Val | Pro | Ile | Phe | Arg | Thr | Leu | Ser | Tyr | Ile |
|     |     | 355 |     |     |     |     | 360 |     |     |     |     | 365 |     |     |     |
| Thr | Gly | Leu | Asp | Asn | Arg | Asn | Pro | Val | Ala | Gly | Ile | Glu | Gly | Val | Glu |
|     | 370 |     |     |     |     | 375 |     |     |     |     | 380 |     |     |     |     |

Phe Gln Asn Thr Ile Ser Arg Ser Ile Tyr Arg Lys Ser Gly Pro Ile  
 385 390 395 400  
 Asp Ser Phe Ser Glu Leu Pro Pro Gln Asp Ala Ser Val Ser Pro Ala  
 405 410 415  
 Ile Gly Tyr Ser His Arg Leu Cys His Ala Thr Phe Leu Glu Arg Ile  
 420 425 430  
 Ser Gly Pro Arg Ile Ala Gly Thr Val Phe Ser Trp Thr His Arg Ser  
 435 440 445  
 Ala Ser Pro Thr Asn Glu Val Ser Pro Ser Arg Ile Thr Gln Ile Pro  
 450 455 460  
 Trp Val Lys Ala His Thr Leu Ala Ser Gly Ala Ser Val Ile Lys Gly  
 465 470 475 480  
 Pro Gly Phe Thr Gly Gly Asp Ile Leu Thr Arg Asn Ser Met Gly Glu  
 485 490 495  
 Leu Gly Thr Leu Arg Val Thr Phe Thr Gly Arg Leu Pro Gln Ser Tyr  
 500 505 510  
 Tyr Ile Arg Phe Arg Tyr Ala Ser Val Ala Asn Arg Ser Gly Thr Phe  
 515 520 525  
 Arg Tyr Ser Gln Pro Pro Ser Tyr Gly Ile Ser Phe Pro Lys Thr Met  
 530 535 540  
 Asp Ala Gly Glu Pro Leu Thr Ser Arg Ser Phe Ala His Thr Thr Leu  
 545 550 555 560  
 Phe Thr Pro Ile Thr Phe Ser Arg Ala Gln Glu Glu Phe Asp Leu Tyr  
 565 570 575  
 Ile Gln Ser Gly Val Tyr Ile Asp Arg Ile Glu Phe Ile Pro Val Thr  
 580 585 590  
 Ala Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala Gln Lys Val Val  
 595 600 605  
 Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Leu Lys Thr Asp Val  
 610 615 620  
 Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Ala Cys Leu Ser  
 625 630 635 640  
 Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys  
 645 650 655  
 His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn  
 660 665 670  
 Phe Arg Gly Ile Asn Arg Gln Pro Asp Arg Gly Trp Arg Gly Ser Thr  
 675 680 685  
 Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val  
 690 695 700  
 Thr Leu Pro Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln  
 705 710 715 720

Lys Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg  
 725 730 735  
 Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr  
 740 745 750  
 Asn Ala Lys His Glu Ile Val Asn Val Pro Gly Thr Gly Ser Leu Trp  
 755 760 765  
 Pro Leu Ser Val Glu Asn Gln Ile Gly Pro Cys Gly Glu Pro Asn Arg  
 770 775 780  
 Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu His Cys Ser Cys Arg  
 785 790 795 800  
 Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile  
 805 810 815  
 Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile  
 820 825 830  
 Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu  
 835 840 845  
 Phe Leu Glu Glu Lys Pro Leu Leu Gly Glu Ala Leu Ala Arg Val Lys  
 850 855 860  
 Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Thr Leu Gln Leu Glu  
 865 870 875 880  
 Thr Thr Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe  
 885 890 895  
 Val Asn Ser Gln Tyr Asp Arg Leu Gln Ala Asp Thr Asn Ile Ala Met  
 900 905 910  
 Ile His Ala Ala Asp Lys Arg Val His Arg Ile Arg Glu Ala Tyr Leu  
 915 920 925  
 Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu  
 930 935 940  
 Leu Glu Glu Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn  
 945 950 955 960  
 Ile Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Leu Cys Trp Asn Val  
 965 970 975  
 Lys Gly His Val Glu Val Glu Glu Gln Asn Asn His Arg Ser Val Leu  
 980 985 990  
 Val Ile Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val Cys  
 995 1000 1005  
 Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu Gly  
 1010 1015 1020  
 Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn Thr  
 1025 1030 1035  
 Asp Glu Leu Lys Phe Asn Asn Cys Val Glu Glu Glu Val Tyr Pro  
 1040 1045 1050  
 Asn Asn Thr Val Thr Cys Ile Asn Tyr Thr Ala Thr Gln Glu Glu

|                                                                                        |      |      |
|----------------------------------------------------------------------------------------|------|------|
| 1055                                                                                   | 1060 | 1065 |
| Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly Tyr Asp Glu Ala                            |      |      |
| 1070                                                                                   | 1075 | 1080 |
| Tyr Gly Asn Asn Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr                            |      |      |
| 1085                                                                                   | 1090 | 1095 |
| Glu Glu Lys Ser Tyr Thr Asp Arg Arg Arg Glu Asn Pro Cys Glu                            |      |      |
| 1100                                                                                   | 1105 | 1110 |
| Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu Pro Ala Gly Tyr                            |      |      |
| 1115                                                                                   | 1120 | 1125 |
| Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp                            |      |      |
| 1130                                                                                   | 1135 | 1140 |
| Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val                            |      |      |
| 1145                                                                                   | 1150 | 1155 |
| Glu Leu Leu Leu Met Glu Glu                                                            |      |      |
| 1160                                                                                   | 1165 |      |
| <210> 7                                                                                |      |      |
| <211> 3498                                                                             |      |      |
| <212> ADN                                                                              |      |      |
| <213> Trình tự nhân tạo                                                                |      |      |
| <220>                                                                                  |      |      |
| <223> Trình tự nucleotit được dùng để biểu hiện ở tế bào vi khuẩn<br>ghi mã CrylDa1_5. |      |      |
| <400> 7                                                                                |      |      |
| atggaaataa ataatcaaaa ccaatgtgtg ccttacaatt gtttaagtaa tcctaaggag                      |      | 60   |
| ataatattag gcgaggaaag gctagaaaca gggaatactg tagcagacat ttcattaggg                      |      | 120  |
| cttattaatt ttctatattc taattttgta ccaggaggag gatttatagt aggtttacta                      |      | 180  |
| gaattaatat ggggatttat agggccttcg caatgggata ttttttttagc tcaaattgag                     |      | 240  |
| caattgatta gtcaaagaat agaagaattt gctaggaatc aggcaatttc aagattggag                      |      | 300  |
| gggctaagca atctttataa ggtctatggt agagcgttta gcgactggga gaaagatcct                      |      | 360  |
| actaatcctg ctttaaggga agaaatgcgt atacaattta atgacatgaa tagtgctctc                      |      | 420  |
| ataacggcta ttccactttt tagagttcaa aattatgaag ttgctctttt atctgtatat                      |      | 480  |
| gttcaagccg caaacttaca tttatctatt ttaagggatg tttcagtttt cggagaaaga                      |      | 540  |
| tggggatatg atacagcgac tatcaataat cgctatagtg atctgactag ccttattcat                      |      | 600  |
| gtttatacta accattgtgt ggatacgtat aatcagggat taaggcgttt ggaaggctgt                      |      | 660  |
| tttcttagcg attggattgt atataatcgt ttccggagac aattgacaat ttcagtatta                      |      | 720  |
| gatattgttg cgtttttttc aaattatgat attagaacat atccaattca aacagctact                      |      | 780  |
| cagctaacga gggaagtcta tctggattta ctttttatta atgaaaatct ttctcctgca                      |      | 840  |
| gcaagctatc caaccttttc agctgctgaa agtgctataa ttagaagtcc tcatttagta                      |      | 900  |
| gactttttta atagctttac catttataca gatagtctgg cacgttctgc atattgggga                      |      | 960  |

```

gggcacttgg taaattcttt cgcacagga accactacta atttgataag atcccccttta 1020
tatggaaggg aaggaaatac agagcgcccc gtaactatta ccgcatcacc tagcgtacca 1080
atatttagaa cactttcata tagaacaggc cttgacaatt caaatcctgt agctggaatc 1140
gagggagtgg aattccaaaa tactataagt agaagtatct atcgtaaaag cgtccaata 1200
gattctttta gtgaattacc acctcaagat gccagcgtat ctctgcaat tgggtatagt 1260
caccgtttat gccatgcaac attttttagaa cggattagt gaccaagaat agcaggcacc 1320
gtattttctt ggacacaccg tagtgccagc cctactaacg aagtaagtcc atctagaatt 1380
acacaaattc catgggtaaa ggcgcatact cttgcatctg gtgcctccgt cattaaaggt 1440
cctggattta caggtggaga tattctgact aggaatagta tgggcgagct ggggacctta 1500
cgagtaacct tcacaggaag attaccacaa agttattata tacgtttccg ttatgcttcg 1560
gtagcaaata ggagtgttac atttagatat tcacagccac cttcgtatgg aatttcattt 1620
ccaaaaacta tggacgcag tgaaccacta acatctcggt cgttcgctca tacaacactc 1680
ttcactccaa taaccttttc acgagctcaa gaagaatttg atctatacat ccaatcgggt 1740
gtttatatag atcgaattga atttataccg gttactgcaa catttgaggc agaatatgat 1800
ttagaaagag cgcaaaaggt ggtgaatgcc ctgtttacgt ctacaaacca actagggcta 1860
aaaacagatg tgacggatta tcatattgat caggtatcca atctagttgc gtgtttatcg 1920
gatgaatttt gtctggatga aaagagagaa ttgtccgaga aagttaaaca tgcaaagcga 1980
ctcagtgatg agcggaattt acttcaagat ccaaacttca gagggatcaa taggcaacca 2040
gaccgtggct ggagaggaag tacggatatt actatccaag gaggagatga cgtattcaaa 2100
gagaattacg ttacgctacc gggtagcttt gatgagtgt atccaacgta tttatatcaa 2160
aaaatagatg agtcgaaatt aaaagcctat acccgttatc aattaagagg gtatatcgaa 2220
gatagtcaag acttagaaat ctatttaatt cgttacaatg caaaacacga aatagtaaat 2280
gtaccaggta caggaagttt atggcctctt totgtagaaa atcaaattgg acctgtgga 2340
gaaccgaatc gatgcgcgcc acaccttgaa tggaatcctg atttacactg ttctgcaga 2400
gacggggaaa aatgtgcaca tcattctcat catttctctt tggacattga tgttgatgt 2460
acagacttaa atgaggactt aggtgtatgg gtgatattca agattaagac gcaagatggc 2520
cacgcacgac tagggaatct agagtttctc gaagagaaac cattattagg agaagcacta 2580
gctcgtgtga aaagagcgga gaaaaaatgg agagacaaac gcgaaacatt acaattggaa 2640
acaactatcg ttataaaga ggcaaaagaa totgtagatg ctttatttgt aaactctcaa 2700
tatgatagat tacaagcgga tacgaacatc gcgatgattc atgcggcaga taaacgcgtt 2760
catagaattc gagaagcgta tctgccggag ctgtctgtga ttccgggtgt caatgcggct 2820
atttttgaag aattagaaga gcgtattttc actgcatttt ccctatatga tgcgagaaat 2880

```

```

attattaaaa atggcgattt caataatggc ttattatgct ggaacgtgaa agggcatgta      2940
gaggtagaag aacaaaacaa tcaccgttca gtcctgggta tcccagaatg ggaggcagaa      3000
gtgtcacaag aggttcgtgt ctgtccaggt cgtgggtata tccttcgtgt tacagcgtac      3060
aaagagggat atggagaagg ttgcgtaacg atccatgaga tcgagaacaa tacagacgaa      3120
ctgaaattca acaactgtgt agaagaggaa gtatatccaa acaacacggg aacgtgtatt      3180
aattatactg cgactcaaga agaatatgag ggtacgtaca cttctcgtaa tcgaggatat      3240
gacgaagcct atggtaataa cccttcgcta ccagctgatt atgcgtcagt ctatgaagaa      3300
aaatcgtata cagatagacg aagagagaat ccttgtgaat ctaacagagg atatggagat      3360
tacacaccac taccagctgg ttatgtaaca aaggaattag agtacttccc agagaccgat      3420
aaggtatgga ttgagattgg agaaacagaa ggaacattca tcgtggacag cgtggaatta      3480
ctccttatgg aggaatag                                                    3498

```

<210> 8

<211> 1165

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Trình tự axit amin của protein đã được xử lý có hoạt tính trừ côn trùng Cry1Da1\_5.

<400> 8

```

Met Glu Ile Asn Asn Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Ser
1              5              10              15

Asn Pro Lys Glu Ile Ile Leu Gly Glu Glu Arg Leu Glu Thr Gly Asn
20              25              30

Thr Val Ala Asp Ile Ser Leu Gly Leu Ile Asn Phe Leu Tyr Ser Asn
35              40              45

Phe Val Pro Gly Gly Gly Phe Ile Val Gly Leu Leu Glu Leu Ile Trp
50              55              60

Gly Phe Ile Gly Pro Ser Gln Trp Asp Ile Phe Leu Ala Gln Ile Glu
65              70              75              80

Gln Leu Ile Ser Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala Ile
85              90              95

Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Lys Val Tyr Val Arg Ala
100             105             110

Phe Ser Asp Trp Glu Lys Asp Pro Thr Asn Pro Ala Leu Arg Glu Glu
115             120             125

Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Ile Thr Ala Ile
130             135             140

Pro Leu Phe Arg Val Gln Asn Tyr Glu Val Ala Leu Leu Ser Val Tyr
145             150             155             160

```

Val Gln Ala Ala Asn Leu His Leu Ser Ile Leu Arg Asp Val Ser Val  
 165 170 175  
 Phe Gly Glu Arg Trp Gly Tyr Asp Thr Ala Thr Ile Asn Asn Arg Tyr  
 180 185 190  
 Ser Asp Leu Thr Ser Leu Ile His Val Tyr Thr Asn His Cys Val Asp  
 195 200 205  
 Thr Tyr Asn Gln Gly Leu Arg Arg Leu Glu Gly Arg Phe Leu Ser Asp  
 210 215 220  
 Trp Ile Val Tyr Asn Arg Phe Arg Arg Gln Leu Thr Ile Ser Val Leu  
 225 230 235 240  
 Asp Ile Val Ala Phe Phe Pro Asn Tyr Asp Ile Arg Thr Tyr Pro Ile  
 245 250 255  
 Gln Thr Ala Thr Gln Leu Thr Arg Glu Val Tyr Leu Asp Leu Pro Phe  
 260 265 270  
 Ile Asn Glu Asn Leu Ser Pro Ala Ala Ser Tyr Pro Thr Phe Ser Ala  
 275 280 285  
 Ala Glu Ser Ala Ile Ile Arg Ser Pro His Leu Val Asp Phe Leu Asn  
 290 295 300  
 Ser Phe Thr Ile Tyr Thr Asp Ser Leu Ala Arg Ser Ala Tyr Trp Gly  
 305 310 315 320  
 Gly His Leu Val Asn Ser Phe Arg Thr Gly Thr Thr Thr Asn Leu Ile  
 325 330 335  
 Arg Ser Pro Leu Tyr Gly Arg Glu Gly Asn Thr Glu Arg Pro Val Thr  
 340 345 350  
 Ile Thr Ala Ser Pro Ser Val Pro Ile Phe Arg Thr Leu Ser Tyr Arg  
 355 360 365  
 Thr Gly Leu Asp Asn Ser Asn Pro Val Ala Gly Ile Glu Gly Val Glu  
 370 375 380  
 Phe Gln Asn Thr Ile Ser Arg Ser Ile Tyr Arg Lys Ser Gly Pro Ile  
 385 390 395 400  
 Asp Ser Phe Ser Glu Leu Pro Pro Gln Asp Ala Ser Val Ser Pro Ala  
 405 410 415  
 Ile Gly Tyr Ser His Arg Leu Cys His Ala Thr Phe Leu Glu Arg Ile  
 420 425 430  
 Ser Gly Pro Arg Ile Ala Gly Thr Val Phe Ser Trp Thr His Arg Ser  
 435 440 445  
 Ala Ser Pro Thr Asn Glu Val Ser Pro Ser Arg Ile Thr Gln Ile Pro  
 450 455 460  
 Trp Val Lys Ala His Thr Leu Ala Ser Gly Ala Ser Val Ile Lys Gly  
 465 470 475 480  
 Pro Gly Phe Thr Gly Gly Asp Ile Leu Thr Arg Asn Ser Met Gly Glu  
 485 490 495  
 Leu Gly Thr Leu Arg Val Thr Phe Thr Gly Arg Leu Pro Gln Ser Tyr

| 500 |     |     |     |     | 505 |     |     |     |     | 510 |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Tyr | Ile | Arg | Phe | Arg | Tyr | Ala | Ser | Val | Ala | Asn | Arg | Ser | Gly | Thr | Phe |
|     |     | 515 |     |     |     |     | 520 |     |     |     |     | 525 |     |     |     |
| Arg | Tyr | Ser | Gln | Pro | Pro | Ser | Tyr | Gly | Ile | Ser | Phe | Pro | Lys | Thr | Met |
|     |     | 530 |     |     |     |     | 535 |     |     |     |     | 540 |     |     |     |
| Asp | Ala | Gly | Glu | Pro | Leu | Thr | Ser | Arg | Ser | Phe | Ala | His | Thr | Thr | Leu |
|     |     |     |     |     |     |     | 550 |     |     |     |     | 555 |     |     | 560 |
| Phe | Thr | Pro | Ile | Thr | Phe | Ser | Arg | Ala | Gln | Glu | Glu | Phe | Asp | Leu | Tyr |
|     |     |     |     | 565 |     |     |     |     | 570 |     |     |     |     | 575 |     |
| Ile | Gln | Ser | Gly | Val | Tyr | Ile | Asp | Arg | Ile | Glu | Phe | Ile | Pro | Val | Thr |
|     |     |     | 580 |     |     |     |     | 585 |     |     |     |     | 590 |     |     |
| Ala | Thr | Phe | Glu | Ala | Glu | Tyr | Asp | Leu | Glu | Arg | Ala | Gln | Lys | Val | Val |
|     |     | 595 |     |     |     |     | 600 |     |     |     |     | 605 |     |     |     |
| Asn | Ala | Leu | Phe | Thr | Ser | Thr | Asn | Gln | Leu | Gly | Leu | Lys | Thr | Asp | Val |
|     |     | 610 |     |     |     |     | 615 |     |     |     |     | 620 |     |     |     |
| Thr | Asp | Tyr | His | Ile | Asp | Gln | Val | Ser | Asn | Leu | Val | Ala | Cys | Leu | Ser |
|     |     | 625 |     |     |     |     | 630 |     |     |     |     | 635 |     |     | 640 |
| Asp | Glu | Phe | Cys | Leu | Asp | Glu | Lys | Arg | Glu | Leu | Ser | Glu | Lys | Val | Lys |
|     |     |     |     | 645 |     |     |     |     | 650 |     |     |     |     | 655 |     |
| His | Ala | Lys | Arg | Leu | Ser | Asp | Glu | Arg | Asn | Leu | Leu | Gln | Asp | Pro | Asn |
|     |     |     | 660 |     |     |     |     | 665 |     |     |     |     | 670 |     |     |
| Phe | Arg | Gly | Ile | Asn | Arg | Gln | Pro | Asp | Arg | Gly | Trp | Arg | Gly | Ser | Thr |
|     |     | 675 |     |     |     |     | 680 |     |     |     |     |     | 685 |     |     |
| Asp | Ile | Thr | Ile | Gln | Gly | Gly | Asp | Asp | Val | Phe | Lys | Glu | Asn | Tyr | Val |
|     |     | 690 |     |     |     |     | 695 |     |     |     |     | 700 |     |     |     |
| Thr | Leu | Pro | Gly | Thr | Phe | Asp | Glu | Cys | Tyr | Pro | Thr | Tyr | Leu | Tyr | Gln |
|     |     | 705 |     |     |     |     | 710 |     |     |     |     | 715 |     |     | 720 |
| Lys | Ile | Asp | Glu | Ser | Lys | Leu | Lys | Ala | Tyr | Thr | Arg | Tyr | Gln | Leu | Arg |
|     |     |     |     | 725 |     |     |     |     | 730 |     |     |     |     | 735 |     |
| Gly | Tyr | Ile | Glu | Asp | Ser | Gln | Asp | Leu | Glu | Ile | Tyr | Leu | Ile | Arg | Tyr |
|     |     |     | 740 |     |     |     |     | 745 |     |     |     |     | 750 |     |     |
| Asn | Ala | Lys | His | Glu | Ile | Val | Asn | Val | Pro | Gly | Thr | Gly | Ser | Leu | Trp |
|     |     | 755 |     |     |     |     | 760 |     |     |     |     | 765 |     |     |     |
| Pro | Leu | Ser | Val | Glu | Asn | Gln | Ile | Gly | Pro | Cys | Gly | Glu | Pro | Asn | Arg |
|     |     | 770 |     |     |     |     | 775 |     |     |     |     | 780 |     |     |     |
| Cys | Ala | Pro | His | Leu | Glu | Trp | Asn | Pro | Asp | Leu | His | Cys | Ser | Cys | Arg |
|     |     | 785 |     |     |     |     | 790 |     |     |     |     | 795 |     |     | 800 |
| Asp | Gly | Glu | Lys | Cys | Ala | His | His | Ser | His | His | Phe | Ser | Leu | Asp | Ile |
|     |     |     |     | 805 |     |     |     |     | 810 |     |     |     |     | 815 |     |
| Asp | Val | Gly | Cys | Thr | Asp | Leu | Asn | Glu | Asp | Leu | Gly | Val | Trp | Val | Ile |
|     |     |     | 820 |     |     |     |     | 825 |     |     |     |     | 830 |     |     |
| Phe | Lys | Ile | Lys | Thr | Gln | Asp | Gly | His | Ala | Arg | Leu | Gly | Asn | Leu | Glu |
|     |     | 835 |     |     |     |     | 840 |     |     |     |     | 845 |     |     |     |

Phe Leu Glu Glu Lys Pro Leu Leu Gly Glu Ala Leu Ala Arg Val Lys  
 850 855 860  
 Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Thr Leu Gln Leu Glu  
 865 870 875 880  
 Thr Thr Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe  
 885 890 895  
 Val Asn Ser Gln Tyr Asp Arg Leu Gln Ala Asp Thr Asn Ile Ala Met  
 900 905 910  
 Ile His Ala Ala Asp Lys Arg Val His Arg Ile Arg Glu Ala Tyr Leu  
 915 920 925  
 Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu  
 930 935 940  
 Leu Glu Glu Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn  
 945 950 955 960  
 Ile Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Leu Cys Trp Asn Val  
 965 970 975  
 Lys Gly His Val Glu Val Glu Glu Gln Asn Asn His Arg Ser Val Leu  
 980 985 990  
 Val Ile Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val Cys  
 995 1000 1005  
 Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu Gly  
 1010 1015 1020  
 Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn Thr  
 1025 1030 1035  
 Asp Glu Leu Lys Phe Asn Asn Cys Val Glu Glu Glu Val Tyr Pro  
 1040 1045 1050  
 Asn Asn Thr Val Thr Cys Ile Asn Tyr Thr Ala Thr Gln Glu Glu  
 1055 1060 1065  
 Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly Tyr Asp Glu Ala  
 1070 1075 1080  
 Tyr Gly Asn Asn Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr  
 1085 1090 1095  
 Glu Glu Lys Ser Tyr Thr Asp Arg Arg Arg Glu Asn Pro Cys Glu  
 1100 1105 1110  
 Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu Pro Ala Gly Tyr  
 1115 1120 1125  
 Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp  
 1130 1135 1140  
 Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val  
 1145 1150 1155  
 Glu Leu Leu Leu Met Glu Glu  
 1160 1165

<210> 9  
 <211> 3498  
 <212> ADN  
 <213> Trình tự nhân tạo

<220>  
 <223> Trình tự nucleotit được dùng để biểu hiện ở tế bào vi khuẩn ghi mã Cry1Da1\_6.

<400> 9  
 atggaaataa ataatcaaaa ccaatgtgtg cottacaatt gtttaagtaa tcctaaggag 60  
 ataatattag gcgaggaaag gctagaaaca gggaatactg tagcagacat ttcattaggg 120  
 cttattaatt ttctatattc taattttgta ccaggaggag gatttatagt aggtttacta 180  
 gaattaatat ggggatttat agggccttcg caatgggata tttttttagc tcaaattgag 240  
 caattgatta gtcaaagaat agaagaatgt gctaggaatc aggcaatttc aagattggag 300  
 gggctaagca atctttataa ggtctatgtt agagcgttta gcgactggga gaaagatcct 360  
 actaatcctg ctttaaggga agaaatgcgt atacaattta atgacatgaa tagtgctctc 420  
 ataacggcta ttccactttt tagagttcaa aattatgaag ttgctctttt atctgtatat 480  
 gttcaagccg caaacttaca tttatctatt ttaagggatg tttcagtttt cggagaaaga 540  
 tggggatatg atacagcgac tatcaataat cgtatagtg atctgactag ccttattcat 600  
 gtttatacta accattgtgt ggatacgtat aatcagggat taaggcgttt ggaaggctgt 660  
 tttcttagcg attggattgt atataatcgt ttccggagac aattgacaat ttcagtatta 720  
 gatattgttg cgttttttcc aaattatgat attagaacat atccaattca aacagctact 780  
 cagctaacga gggaagtcta tctggattta ctttttatta atgaaaatct ttctcctgca 840  
 gcaaaatatt caaccttttc agctgctgaa agtgctataa ttagaagtc tcathtagta 900  
 gactttttta atagctttac catttataca gatagtctgg cacgttctgc atattgggga 960  
 gggcacttgg taaattcttt ccgcacagga accactacta atttgataag atccccttta 1020  
 tatggaaggg aaggaaatac agagcgcccc gtaactatta ccgcatcacc tagcgtacca 1080  
 atatttagaa cactttcata tccaacaggc cttgacaatt caaatcctgt agctggaatc 1140  
 gagggagtgg aattccaaaa tactataagt agaagtatct atcgtaaaag cggccaata 1200  
 gattctttta gtgaattacc acctcaagat gccagcgtat ctctgcaat tgggtatagt 1260  
 caccgtttat gccatgcaac attttttagaa cggattagtg gaccaagaat agcaggcacc 1320  
 gtatttttct ggacacaccg tagtgccagc cctactaacg aagtaagtc atctagaatt 1380  
 acacaaattc catgggtaaa ggcgatact cttgcatctg gtgcctccgt cattaaagg 1440  
 cctggattta cagggtggaga tattctgact aggaatagta tgggcgagct ggggacctta 1500  
 cgagtaacct tcacaggaag attaccacaa agttattata tacgtttccg ttatgcttcg 1560  
 gtagcaaata ggagtgttac atttagatat tcacagccac cttcgtatgg aatttcattt 1620

|                                                                    |      |
|--------------------------------------------------------------------|------|
| ccaaaaacta tggacgcagg tgaaccacta acatctcggt cggtcgctca tacaacactc  | 1680 |
| ttcactccaa taaccttttc acgagctcaa gaagaatttg atctatacat ccaatcgggt  | 1740 |
| gtttatatag atcgaattga atttataccg gttactgcaa catttgaggc agaatatgat  | 1800 |
| ttagaaagag cgcaaaaggt ggtgaatgcc ctgtttacgt ctacaaacca actagggcta  | 1860 |
| aaaacagatg tgacggatta tcatattgat caggtatcca atctagttgc gtgtttatcg  | 1920 |
| gatgaatttt gtctggatga aaagagagaa ttgtccgaga aagttaaaca tgcaaagcga  | 1980 |
| ctcagtgatg agcggaattt acttcaagat ccaaacttca gagggatcaa taggcaacca  | 2040 |
| gaccgtggct ggagaggaag tacggatatt actatccaag gagggatga cgtattcaaa   | 2100 |
| gagaattacg ttacgctacc gggtagcttt gatgagtgc atccaacgta tttatatcaa   | 2160 |
| aaaatagatg agtcgaaatt aaaagcctat acccgttatc aattaagagg gtatatcgaa  | 2220 |
| gatagtcaag acttagaaat ctatttaatt cggtacaatg caaaacacga aatagtaaat  | 2280 |
| gtaccaggta caggaagttt atggcctctt tctgtagaaa atcaaattgg accttgtgga  | 2340 |
| gaaccgaatc gatgcgcgcc acaccttgaa tggaatcctg atttacactg ttacctgcaga | 2400 |
| gacggggaaa aatgtgcaca tcattctcat catttctctt tggacattga tgttggatgt  | 2460 |
| acagacttaa atgaggactt aggtgtatgg gtgatattca agattaagac gcaagatggc  | 2520 |
| cacgcacgac tagggaatct agagtttctc gaagagaaac cattattagg agaagcacta  | 2580 |
| gctcgtgtga aaagagcgga gaaaaaatgg agagacaaac gcgaaacatt acaattggaa  | 2640 |
| acaactatcg tttataaaga ggcaaaagaa tctgtagatg ctttatttgt aaactctcaa  | 2700 |
| tatgatagat tacaagcgga tacgaacatc gcgatgattc atgcggcaga taaacgcgtt  | 2760 |
| catagaattc gagaagcgta tctgccggag ctgtctgtga ttccgggtgt caatgcggct  | 2820 |
| atttttgaag aattagaaga gcgtattttc actgcatttt ccctatatga tgcgagaaat  | 2880 |
| attattaaaa atggcgattt caataatggc ttattatgct ggaacgtgaa agggcatgta  | 2940 |
| gaggtagaag aacaaaacaa tcaccgttca gtocgtggtta tcccagaatg ggaggcagaa | 3000 |
| gtgtcacaaag aggttcgtgt ctgtccaggt cgtggctata tccttcgtgt tacagcgtac | 3060 |
| aaagagggat atggagaagg ttgcgtaacg atccatgaga tcgagaacaa tacagacgaa  | 3120 |
| ctgaaattca acaactgtgt agaagaggaa gtatatccaa acaacacggt aacgtgtatt  | 3180 |
| aattatactg cgactcaaga agaatatgag ggtacgtaca cttctcgtaa tcgaggatat  | 3240 |
| gacgaagcct atggtaataa cccttcgta ccagctgatt atgcgtcagt ctatgaagaa   | 3300 |
| aatcgtata cagatagacg aagagagaat ccttgtgaat ctaacagagg atatggagat   | 3360 |
| tacacaccac taccagctgg ttatgtaaca aaggaattag agtacttccc agagaccgat  | 3420 |
| aaggtatgga ttgagattgg agaaacagaa ggaacattca tcgtggacag cgtggaatta  | 3480 |
| ctccttatgg aggaatag                                                | 3498 |

<210> 10  
 <211> 1165  
 <212> PRT  
 <213> Trình tự nhân tạo  
  
 <220>  
 <223> Trình tự axit amin của protein đã được xử lý có hoạt tính trừ côn trùng Cry1Dal\_6.  
  
 <400> 10  
  
 Met Glu Ile Asn Asn Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Ser  
 1 5 10 15  
 Asn Pro Lys Glu Ile Ile Leu Gly Glu Glu Arg Leu Glu Thr Gly Asn  
 20 25 30  
 Thr Val Ala Asp Ile Ser Leu Gly Leu Ile Asn Phe Leu Tyr Ser Asn  
 35 40 45  
 Phe Val Pro Gly Gly Gly Phe Ile Val Gly Leu Leu Glu Leu Ile Trp  
 50 55 60  
 Gly Phe Ile Gly Pro Ser Gln Trp Asp Ile Phe Leu Ala Gln Ile Glu  
 65 70 75 80  
 Gln Leu Ile Ser Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala Ile  
 85 90 95  
 Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Lys Val Tyr Val Arg Ala  
 100 105 110  
 Phe Ser Asp Trp Glu Lys Asp Pro Thr Asn Pro Ala Leu Arg Glu Glu  
 115 120 125  
 Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Ile Thr Ala Ile  
 130 135 140  
 Pro Leu Phe Arg Val Gln Asn Tyr Glu Val Ala Leu Leu Ser Val Tyr  
 145 150 155 160  
 Val Gln Ala Ala Asn Leu His Leu Ser Ile Leu Arg Asp Val Ser Val  
 165 170 175  
 Phe Gly Glu Arg Trp Gly Tyr Asp Thr Ala Thr Ile Asn Asn Arg Tyr  
 180 185 190  
 Ser Asp Leu Thr Ser Leu Ile His Val Tyr Thr Asn His Cys Val Asp  
 195 200 205  
 Thr Tyr Asn Gln Gly Leu Arg Arg Leu Glu Gly Arg Phe Leu Ser Asp  
 210 215 220  
 Trp Ile Val Tyr Asn Arg Phe Arg Arg Gln Leu Thr Ile Ser Val Leu  
 225 230 235 240  
 Asp Ile Val Ala Phe Phe Pro Asn Tyr Asp Ile Arg Thr Tyr Pro Ile  
 245 250 255  
 Gln Thr Ala Thr Gln Leu Thr Arg Glu Val Tyr Leu Asp Leu Pro Phe  
 260 265 270  
 Ile Asn Glu Asn Leu Ser Pro Ala Ala Lys Tyr Pro Thr Phe Ser Ala  
 275 280 285

Ala Glu Ser Ala Ile Ile Arg Ser Pro His Leu Val Asp Phe Leu Asn  
 290 295 300  
 Ser Phe Thr Ile Tyr Thr Asp Ser Leu Ala Arg Ser Ala Tyr Trp Gly  
 305 310 315 320  
 Gly His Leu Val Asn Ser Phe Arg Thr Gly Thr Thr Thr Asn Leu Ile  
 325 330 335  
 Arg Ser Pro Leu Tyr Gly Arg Glu Gly Asn Thr Glu Arg Pro Val Thr  
 340 345 350  
 Ile Thr Ala Ser Pro Ser Val Pro Ile Phe Arg Thr Leu Ser Tyr Pro  
 355 360 365  
 Thr Gly Leu Asp Asn Ser Asn Pro Val Ala Gly Ile Glu Gly Val Glu  
 370 375 380  
 Phe Gln Asn Thr Ile Ser Arg Ser Ile Tyr Arg Lys Ser Gly Pro Ile  
 385 390 395 400  
 Asp Ser Phe Ser Glu Leu Pro Pro Gln Asp Ala Ser Val Ser Pro Ala  
 405 410 415  
 Ile Gly Tyr Ser His Arg Leu Cys His Ala Thr Phe Leu Glu Arg Ile  
 420 425 430  
 Ser Gly Pro Arg Ile Ala Gly Thr Val Phe Ser Trp Thr His Arg Ser  
 435 440 445  
 Ala Ser Pro Thr Asn Glu Val Ser Pro Ser Arg Ile Thr Gln Ile Pro  
 450 455 460  
 Trp Val Lys Ala His Thr Leu Ala Ser Gly Ala Ser Val Ile Lys Gly  
 465 470 475 480  
 Pro Gly Phe Thr Gly Gly Asp Ile Leu Thr Arg Asn Ser Met Gly Glu  
 485 490 495  
 Leu Gly Thr Leu Arg Val Thr Phe Thr Gly Arg Leu Pro Gln Ser Tyr  
 500 505 510  
 Tyr Ile Arg Phe Arg Tyr Ala Ser Val Ala Asn Arg Ser Gly Thr Phe  
 515 520 525  
 Arg Tyr Ser Gln Pro Pro Ser Tyr Gly Ile Ser Phe Pro Lys Thr Met  
 530 535 540  
 Asp Ala Gly Glu Pro Leu Thr Ser Arg Ser Phe Ala His Thr Thr Leu  
 545 550 555 560  
 Phe Thr Pro Ile Thr Phe Ser Arg Ala Gln Glu Glu Phe Asp Leu Tyr  
 565 570 575  
 Ile Gln Ser Gly Val Tyr Ile Asp Arg Ile Glu Phe Ile Pro Val Thr  
 580 585 590  
 Ala Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala Gln Lys Val Val  
 595 600 605  
 Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Leu Lys Thr Asp Val  
 610 615 620

```

Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Ala Cys Leu Ser
625                               630                               635                               640

Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys
                               645                               650                               655

His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn
                               660                               665                               670

Phe Arg Gly Ile Asn Arg Gln Pro Asp Arg Gly Trp Arg Gly Ser Thr
                               675                               680                               685

Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val
690                               695                               700

Thr Leu Pro Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln
705                               710                               715                               720

Lys Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg
                               725                               730                               735

Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr
                               740                               745                               750

Asn Ala Lys His Glu Ile Val Asn Val Pro Gly Thr Gly Ser Leu Trp
                               755                               760                               765

Pro Leu Ser Val Glu Asn Gln Ile Gly Pro Cys Gly Glu Pro Asn Arg
770                               775                               780

Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu His Cys Ser Cys Arg
785                               790                               795                               800

Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile
                               805                               810                               815

Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile
                               820                               825                               830

Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu
835                               840                               845

Phe Leu Glu Glu Lys Pro Leu Leu Gly Glu Ala Leu Ala Arg Val Lys
850                               855                               860

Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Thr Leu Gln Leu Glu
865                               870                               875                               880

Thr Thr Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe
                               885                               890                               895

Val Asn Ser Gln Tyr Asp Arg Leu Gln Ala Asp Thr Asn Ile Ala Met
                               900                               905                               910

Ile His Ala Ala Asp Lys Arg Val His Arg Ile Arg Glu Ala Tyr Leu
915                               920                               925

Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu
930                               935                               940

Leu Glu Glu Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn
945                               950                               955                               960

Ile Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Leu Cys Trp Asn Val

```

965 970 975  
 Lys Gly His Val Glu Val Glu Glu Gln Asn Asn His Arg Ser Val Leu  
 980 985 990  
 Val Ile Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val Cys  
 995 1000 1005  
 Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu Gly  
 1010 1015 1020  
 Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn Thr  
 1025 1030 1035  
 Asp Glu Leu Lys Phe Asn Asn Cys Val Glu Glu Glu Val Tyr Pro  
 1040 1045 1050  
 Asn Asn Thr Val Thr Cys Ile Asn Tyr Thr Ala Thr Gln Glu Glu  
 1055 1060 1065  
 Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly Tyr Asp Glu Ala  
 1070 1075 1080  
 Tyr Gly Asn Asn Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr  
 1085 1090 1095  
 Glu Glu Lys Ser Tyr Thr Asp Arg Arg Arg Glu Asn Pro Cys Glu  
 1100 1105 1110  
 Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu Pro Ala Gly Tyr  
 1115 1120 1125  
 Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp  
 1130 1135 1140  
 Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val  
 1145 1150 1155  
 Glu Leu Leu Leu Met Glu Glu  
 1160 1165

<210> 11

<211> 3498

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Trình tự nucleotit được dùng để biểu hiện ở tế bào vi khuẩn ghi mã Cry1Da1\_7.

<400> 11

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| atggaaataa ataatcaaaa ccaatgtgtg ccttacaatt gtttaagtaa tcctaaggag  | 60  |
| ataatattag gcgaggaaag gctagaaaca gggaatactg tagcagacat ttcattaggg  | 120 |
| cttattaatt ttctatattc taattttgtg ccaggaggag gatttatagt aggtttacta  | 180 |
| gaattaatat ggggatttat agggccttcg caatgggata tttttttagc tcaaattgag  | 240 |
| caattgatta gtcaaagaat agaagaattt gctaggaatc aggcaatttc aagattggag  | 300 |
| gggctaagca atctttataa ggtctatgtt agagcgttta gcgactggga gaaagatcct  | 360 |
| actaatcctg ctttaaggga agaaatgcgt atacaattta atgacatgaa tagtgcctctc | 420 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| ataacggcta ttccactttt tagagttcaa aattatgaag ttgctctttt atctgtatat  | 480  |
| gttcaagccg caaacttaca tttatctatt ttaagggatg tttcagtttt cggagaaaga  | 540  |
| tggggatatg atacagcgac tatcaataat cgctatagtg atctgactag ccttattcat  | 600  |
| gtttatacta accattgtgt ggatacgtat aatcagggat taaggcgttt ggaaggctgt  | 660  |
| tttcttagcg attggattgt atataatcgt ttccggagac aattgacaat ttcagtatta  | 720  |
| gatattgttg cgttttttcc aaattatgat attagaacat atccaattca aacagctact  | 780  |
| cagctaacga gggaagtcta tctggattta ctttttatta atgaaaatct ttctcctgca  | 840  |
| gcagtatatc caaccttttc agctgctgaa agtgcataa ttagaagtcc tcatttagta   | 900  |
| gactttttta atagctttac cttttataca gatagtctgg cacgttctgc atattgggga  | 960  |
| gggcacttgg taaattcttt ccgcacagga accactacta atttgataag atccccctta  | 1020 |
| tatggaaggg aaggaaatac agagcgcccc gtaactatta ccgcatcacc tagcgtacca  | 1080 |
| atatttagaa cactttcata tccaacaggc cttgacaatt caaatcctgt agctggaatc  | 1140 |
| gagggagtgg aattccaaaa tactataagt agaagtatct atcgtaaaag cggccaata   | 1200 |
| gattctttta gtgaattacc acctcaagat gccagcgtat ctctgcaat tgggtatagt   | 1260 |
| caccgtttat gccatgcaac attttttagaa cggattagtg gaccaagaat agcaggcacc | 1320 |
| gtattttctt ggacacaccg tagtgccagc cctactaacg aagtaagtcc atctagaatt  | 1380 |
| acacaaattc catgggtaaa ggcgcatact cttgcatctg gtgcctccgt cattaaaggt  | 1440 |
| cctggattta caggtggaga tattctgact aggaatagta tgggcgagct ggggacctta  | 1500 |
| cgagtaacct tcacaggaag attaccacaa agttattata tacgtttccg ttatgcttcg  | 1560 |
| gtagcaaata ggagtgttac atttagatat tcacagccac ctctgtatgg aatttcattt  | 1620 |
| ccaaaaacta tggacgcagg tgaaccacta acatctcggt cgttcgctca tacaacactc  | 1680 |
| ttcactccaa taaccttttc acgagctcaa gaagaatttg atctatacat ccaatcgggt  | 1740 |
| gtttatatag atcgaattga atttataccg gttactgcaa catttgaggc agaatatgat  | 1800 |
| ttagaaagag cgcaaaaggt ggtgaatgcc ctgtttacgt ctacaaacca actagggcta  | 1860 |
| aaaacagatg tgacggatta tcatattgat caggtatcca atctagttgc gtgtttatcg  | 1920 |
| gatgaatttt gtctggatga aaagagagaa ttgtccgaga aagttaaaca tgcaaagcga  | 1980 |
| ctcagtgatg agcggaattt acttcaagat ccaaacttca gagggatcaa taggcaacca  | 2040 |
| gaccgtggct ggagaggaag tacggatatt actatccaag gaggagatga cgtattcaaa  | 2100 |
| gagaattacg ttacgctacc gggtaacctt gatgagtgt atccaacgta tttatatcaa   | 2160 |
| aaaatagatg agtcgaaatt aaaagcctat acccgttatc aattaagagg gtatatcgaa  | 2220 |
| gatagtcaag acttagaaat ctatttaatt cgttacaatg caaaacacga aatagtaa    | 2280 |
| gtaccaggta caggaagttt atggcctctt tctgtagaaa atcaaattgg accttgtgga  | 2340 |

```

gaaccgaatc gatgcgcgcc acaccttgaa tggaatcctg atttacactg ttcctgcaga      2400
gacggggaaa aatgtgcaca tcattctcat catttctctt tggacattga tgttggatgt      2460
acagacttaa atgaggactt aggtgtatgg gtgatattca agattaagac gcaagatggc      2520
cacgcacgac tagggaatct agagtttctc gaagagaaac cattattagg agaagcacta      2580
gctcgtgtga aaagagcggg gaaaaaatgg agagacaaac gcgaaacatt acaattggaa      2640
acaactatcg ttataaaga ggcaaaagaa tctgtagatg ctttatttgt aaactotcaa      2700
tatgatagat tacaagcggg tacgaacatc gcgatgattc atgcggcaga taaacgcgtt      2760
catagaattc gagaagcgtg tctgccggag ctgtctgtga ttccgggtgt caatgcggct      2820
atttttgaag aattagaaga gcgtattttc actgcatttt ccctatatga tgcgagaaat      2880
attattaaaa atggcgattt caataatggc ttattatgct ggaacgtgaa agggcatgta      2940
gaggtagaag aacaaaacaa tcaccgttca gtcctgggta tcccagaatg ggaggcagaa      3000
gtgtcacaag aggttcgtgt ctgtccaggt cgtggctata tccttcgtgt tacagcgtac      3060
aaagagggat atggagaagg ttgcgtaacg atccatgaga tcgagaacaa tacagaogaa      3120
ctgaaattca acaactgtgt agaagaggaa gtatatccaa acaacacggt aacgtgtatt      3180
aattatactg cgactcaaga agaatatgag ggtacgtaca cttctcgtaa tcgaggatat      3240
gacgaagcct atggtaataa cccttccgta ccagctgatt atgcgtcagt ctatgaagaa      3300
aaatcgtata cagatagacg aagagagaat ccttgtgaat ctaacagagg atatggagat      3360
tacacaccac taccagctgg ttatgtaaca aaggaattag agtacttccc agagaccgat      3420
aaggatatga ttgagattgg agaaacagaa ggaacattca tcgtggacag cgtggaatta      3480
ctccttatgg aggaatag                                          3498

```

<210> 12

<211> 1165

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Trình tự axit amin của protein đã được xử lý có hoạt tính trừ côn trùng Cry1Da1\_7.

<400> 12

```

Met Glu Ile Asn Asn Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Ser
1           5           10           15

```

```

Asn Pro Lys Glu Ile Ile Leu Gly Glu Glu Arg Leu Glu Thr Gly Asn
          20           25           30

```

```

Thr Val Ala Asp Ile Ser Leu Gly Leu Ile Asn Phe Leu Tyr Ser Asn
          35           40           45

```

```

Phe Val Pro Gly Gly Gly Phe Ile Val Gly Leu Leu Glu Leu Ile Trp
          50           55           60

```

Gly Phe Ile Gly Pro Ser Gln Trp Asp Ile Phe Leu Ala Gln Ile Glu  
 65 70 75 80  
 Gln Leu Ile Ser Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala Ile  
 85 90 95  
 Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Lys Val Tyr Val Arg Ala  
 100 105 110  
 Phe Ser Asp Trp Glu Lys Asp Pro Thr Asn Pro Ala Leu Arg Glu Glu  
 115 120 125  
 Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Ile Thr Ala Ile  
 130 135 140  
 Pro Leu Phe Arg Val Gln Asn Tyr Glu Val Ala Leu Leu Ser Val Tyr  
 145 150 155 160  
 Val Gln Ala Ala Asn Leu His Leu Ser Ile Leu Arg Asp Val Ser Val  
 165 170 175  
 Phe Gly Glu Arg Trp Gly Tyr Asp Thr Ala Thr Ile Asn Asn Arg Tyr  
 180 185 190  
 Ser Asp Leu Thr Ser Leu Ile His Val Tyr Thr Asn His Cys Val Asp  
 195 200 205  
 Thr Tyr Asn Gln Gly Leu Arg Arg Leu Glu Gly Arg Phe Leu Ser Asp  
 210 215 220  
 Trp Ile Val Tyr Asn Arg Phe Arg Arg Gln Leu Thr Ile Ser Val Leu  
 225 230 235 240  
 Asp Ile Val Ala Phe Phe Pro Asn Tyr Asp Ile Arg Thr Tyr Pro Ile  
 245 250 255  
 Gln Thr Ala Thr Gln Leu Thr Arg Glu Val Tyr Leu Asp Leu Pro Phe  
 260 265 270  
 Ile Asn Glu Asn Leu Ser Pro Ala Ala Val Tyr Pro Thr Phe Ser Ala  
 275 280 285  
 Ala Glu Ser Ala Ile Ile Arg Ser Pro His Leu Val Asp Phe Leu Asn  
 290 295 300  
 Ser Phe Thr Ile Tyr Thr Asp Ser Leu Ala Arg Ser Ala Tyr Trp Gly  
 305 310 315 320  
 Gly His Leu Val Asn Ser Phe Arg Thr Gly Thr Thr Thr Asn Leu Ile  
 325 330 335  
 Arg Ser Pro Leu Tyr Gly Arg Glu Gly Asn Thr Glu Arg Pro Val Thr  
 340 345 350  
 Ile Thr Ala Ser Pro Ser Val Pro Ile Phe Arg Thr Leu Ser Tyr Pro  
 355 360 365  
 Thr Gly Leu Asp Asn Ser Asn Pro Val Ala Gly Ile Glu Gly Val Glu  
 370 375 380  
 Phe Gln Asn Thr Ile Ser Arg Ser Ile Tyr Arg Lys Ser Gly Pro Ile  
 385 390 395 400  
 Asp Ser Phe Ser Glu Leu Pro Pro Gln Asp Ala Ser Val Ser Pro Ala

| 405 |     |     |     |     |     |     |     |     |     | 410 |     |     |     |     | 415 |  |  |  |  |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|--|--|--|
| Ile | Gly | Tyr | Ser | His | Arg | Leu | Cys | His | Ala | Thr | Phe | Leu | Glu | Arg | Ile |  |  |  |  |
|     |     |     | 420 |     |     |     |     | 425 |     |     |     |     | 430 |     |     |  |  |  |  |
| Ser | Gly | Pro | Arg | Ile | Ala | Gly | Thr | Val | Phe | Ser | Trp | Thr | His | Arg | Ser |  |  |  |  |
|     |     | 435 |     |     |     |     | 440 |     |     |     |     | 445 |     |     |     |  |  |  |  |
| Ala | Ser | Pro | Thr | Asn | Glu | Val | Ser | Pro | Ser | Arg | Ile | Thr | Gln | Ile | Pro |  |  |  |  |
|     | 450 |     |     |     |     | 455 |     |     |     |     | 460 |     |     |     |     |  |  |  |  |
| Trp | Val | Lys | Ala | His | Thr | Leu | Ala | Ser | Gly | Ala | Ser | Val | Ile | Lys | Gly |  |  |  |  |
| 465 |     |     |     |     | 470 |     |     |     |     | 475 |     |     |     |     | 480 |  |  |  |  |
| Pro | Gly | Phe | Thr | Gly | Gly | Asp | Ile | Leu | Thr | Arg | Asn | Ser | Met | Gly | Glu |  |  |  |  |
|     |     |     |     | 485 |     |     |     |     | 490 |     |     |     |     | 495 |     |  |  |  |  |
| Leu | Gly | Thr | Leu | Arg | Val | Thr | Phe | Thr | Gly | Arg | Leu | Pro | Gln | Ser | Tyr |  |  |  |  |
|     |     |     | 500 |     |     |     |     | 505 |     |     |     |     | 510 |     |     |  |  |  |  |
| Tyr | Ile | Arg | Phe | Arg | Tyr | Ala | Ser | Val | Ala | Asn | Arg | Ser | Gly | Thr | Phe |  |  |  |  |
|     |     | 515 |     |     |     |     | 520 |     |     |     |     | 525 |     |     |     |  |  |  |  |
| Arg | Tyr | Ser | Gln | Pro | Pro | Ser | Tyr | Gly | Ile | Ser | Phe | Pro | Lys | Thr | Met |  |  |  |  |
|     | 530 |     |     |     |     | 535 |     |     |     |     | 540 |     |     |     |     |  |  |  |  |
| Asp | Ala | Gly | Glu | Pro | Leu | Thr | Ser | Arg | Ser | Phe | Ala | His | Thr | Thr | Leu |  |  |  |  |
| 545 |     |     |     |     | 550 |     |     |     |     | 555 |     |     |     |     | 560 |  |  |  |  |
| Phe | Thr | Pro | Ile | Thr | Phe | Ser | Arg | Ala | Gln | Glu | Glu | Phe | Asp | Leu | Tyr |  |  |  |  |
|     |     |     |     | 565 |     |     |     |     | 570 |     |     |     |     | 575 |     |  |  |  |  |
| Ile | Gln | Ser | Gly | Val | Tyr | Ile | Asp | Arg | Ile | Glu | Phe | Ile | Pro | Val | Thr |  |  |  |  |
|     |     |     | 580 |     |     |     |     | 585 |     |     |     |     | 590 |     |     |  |  |  |  |
| Ala | Thr | Phe | Glu | Ala | Glu | Tyr | Asp | Leu | Glu | Arg | Ala | Gln | Lys | Val | Val |  |  |  |  |
|     |     | 595 |     |     |     |     | 600 |     |     |     |     | 605 |     |     |     |  |  |  |  |
| Asn | Ala | Leu | Phe | Thr | Ser | Thr | Asn | Gln | Leu | Gly | Leu | Lys | Thr | Asp | Val |  |  |  |  |
|     | 610 |     |     |     |     | 615 |     |     |     |     | 620 |     |     |     |     |  |  |  |  |
| Thr | Asp | Tyr | His | Ile | Asp | Gln | Val | Ser | Asn | Leu | Val | Ala | Cys | Leu | Ser |  |  |  |  |
| 625 |     |     |     |     | 630 |     |     |     |     | 635 |     |     |     |     | 640 |  |  |  |  |
| Asp | Glu | Phe | Cys | Leu | Asp | Glu | Lys | Arg | Glu | Leu | Ser | Glu | Lys | Val | Lys |  |  |  |  |
|     |     |     |     | 645 |     |     |     |     | 650 |     |     |     |     | 655 |     |  |  |  |  |
| His | Ala | Lys | Arg | Leu | Ser | Asp | Glu | Arg | Asn | Leu | Leu | Gln | Asp | Pro | Asn |  |  |  |  |
|     |     |     | 660 |     |     |     |     | 665 |     |     |     |     | 670 |     |     |  |  |  |  |
| Phe | Arg | Gly | Ile | Asn | Arg | Gln | Pro | Asp | Arg | Gly | Trp | Arg | Gly | Ser | Thr |  |  |  |  |
|     |     | 675 |     |     |     |     | 680 |     |     |     |     | 685 |     |     |     |  |  |  |  |
| Asp | Ile | Thr | Ile | Gln | Gly | Gly | Asp | Asp | Val | Phe | Lys | Glu | Asn | Tyr | Val |  |  |  |  |
|     | 690 |     |     |     |     | 695 |     |     |     |     | 700 |     |     |     |     |  |  |  |  |
| Thr | Leu | Pro | Gly | Thr | Phe | Asp | Glu | Cys | Tyr | Pro | Thr | Tyr | Leu | Tyr | Gln |  |  |  |  |
| 705 |     |     |     |     | 710 |     |     |     |     | 715 |     |     |     |     | 720 |  |  |  |  |
| Lys | Ile | Asp | Glu | Ser | Lys | Leu | Lys | Ala | Tyr | Thr | Arg | Tyr | Gln | Leu | Arg |  |  |  |  |
|     |     |     |     | 725 |     |     |     |     | 730 |     |     |     |     | 735 |     |  |  |  |  |
| Gly | Tyr | Ile | Glu | Asp | Ser | Gln | Asp | Leu | Glu | Ile | Tyr | Leu | Ile | Arg | Tyr |  |  |  |  |
|     |     |     | 740 |     |     |     |     | 745 |     |     |     |     | 750 |     |     |  |  |  |  |

```

Asn Ala Lys His Glu Ile Val Asn Val Pro Gly Thr Gly Ser Leu Trp
  755                      760                      765

Pro Leu Ser Val Glu Asn Gln Ile Gly Pro Cys Gly Glu Pro Asn Arg
  770                      775                      780

Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu His Cys Ser Cys Arg
  785                      790                      795                      800

Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile
                      805                      810                      815

Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile
                      820                      825                      830

Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu
  835                      840                      845

Phe Leu Glu Glu Lys Pro Leu Leu Gly Glu Ala Leu Ala Arg Val Lys
  850                      855                      860

Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Thr Leu Gln Leu Glu
  865                      870                      875                      880

Thr Thr Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe
                      885                      890                      895

Val Asn Ser Gln Tyr Asp Arg Leu Gln Ala Asp Thr Asn Ile Ala Met
                      900                      905                      910

Ile His Ala Ala Asp Lys Arg Val His Arg Ile Arg Glu Ala Tyr Leu
  915                      920                      925

Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu
  930                      935                      940

Leu Glu Glu Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn
  945                      950                      955                      960

Ile Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Leu Cys Trp Asn Val
                      965                      970                      975

Lys Gly His Val Glu Val Glu Glu Gln Asn Asn His Arg Ser Val Leu
                      980                      985                      990

Val Ile Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val Cys
  995                      1000                      1005

Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu Gly
  1010                      1015                      1020

Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn Thr
  1025                      1030                      1035

Asp Glu Leu Lys Phe Asn Asn Cys Val Glu Glu Glu Val Tyr Pro
  1040                      1045                      1050

Asn Asn Thr Val Thr Cys Ile Asn Tyr Thr Ala Thr Gln Glu Glu
  1055                      1060                      1065

Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly Tyr Asp Glu Ala
  1070                      1075                      1080

```

Tyr Gly Asn Asn Pro Ser Val Pro Ala Asp Tyr Ala Ser Val Tyr  
 1085 1090 1095  
 Glu Glu Lys Ser Tyr Thr Asp Arg Arg Arg Glu Asn Pro Cys Glu  
 1100 1105 1110  
 Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu Pro Ala Gly Tyr  
 1115 1120 1125  
 Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val Trp  
 1130 1135 1140  
 Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser Val  
 1145 1150 1155  
 Glu Leu Leu Leu Met Glu Glu  
 1160 1165

<210> 13

<211> 3420

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Trình tự nucleotit được dùng để biểu hiện ở tế bào vi khuẩn ghi mã TIC844.

<400> 13

|                                                                    |      |
|--------------------------------------------------------------------|------|
| atggagatca acaaccagaa ccagtgcgtc ccgtacaact gcctgagcaa ccctaaggag  | 60   |
| atcatcctgg gtgaggaacg cctggagacc ggcaacaccg tagccgacat tagcctgggc  | 120  |
| ctcatcaact tcctctacag caacttcgtg cccggcgggtg gcttcacgtg gggcctcctg | 180  |
| gagcttatct ggggcttcat cggcccgtcc cagtgggaca tcttcctcgc ccagatcgag  | 240  |
| caactgatca gccagcggat cgaggagttc gctaggaacc aggccatctc ccgcctggag  | 300  |
| ggactctcca acctctacaa ggtgtacgtg gcgcggttca gcgactggga gaaggaccgc  | 360  |
| accaaccgcg ccctccgcga ggaaatgcgt atccagttca acgatatgaa ctgcggccctc | 420  |
| atcaccgcca tcccgctctt ccgcgtgcag aactacgagg tggccctcct gtccgtgtac  | 480  |
| gttcaagccg ccaacctcca cctctccatc ctccgcgacg tgagcgtggt cggcgagcgc  | 540  |
| tggggctacg acaccgccac catcaacaac cgctactccg acctcacctc cctcatccac  | 600  |
| gtttacacca accactgcgt ggacacgtac aaccagggcc tccgccgcct ggagggccgc  | 660  |
| ttcctctccg actggatcgt gtacaaccgc ttccgccgcc agctcaccat ctccgtcctg  | 720  |
| gacatcgtcg ccttctttcc caactacgac atccgcacct accctatcca gaccgccacc  | 780  |
| cagctcaccg gcgaggtcta cctcgacctc ccgttcatca acgagaacct cagcccggcc  | 840  |
| gccagctacc cgaccttctc cgccgctgag tccgccatca ttgcagccc gcacctcgtg   | 900  |
| gacttcctca actccttcac catctacacc gactccctcg cccgctacgc ctactggggc  | 960  |
| ggtcacctcg tgaactcctt ccgcaccggc accactacca acctcatccg cagcccgcgc  | 1020 |
| tacggccgcg agggcaaacac cgagcgcccc gtgaccatca ccgccagccc gagcgtgccc | 1080 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| atcttccgca ccctcagcta catcacccgc ctggacaaca gcaaccctgt ggcgggcatc  | 1140 |
| gagggcggtg agttccagaa caccatctcc aggagcatct accgcaagag cggccctatc  | 1200 |
| gacagcttca gcgagctgcc tctcaggac gccagcgtga gccctgccat cggctacagc   | 1260 |
| cacaggctgt gccacgccac ctctctggag cgcacacagc gccctcgcac cgcgggcacc  | 1320 |
| gtgttctcgt ggaccacccg cagcgcctct cctacgaacg aggtgtctcc tagtcgcac   | 1380 |
| accagatcc cttgggtcaa ggcccacacc ctggctagtgt gcgctagtgt catcaagggc  | 1440 |
| cctggcttca ccggtggtga catcctgacc aggaactcta tgggcgagct gggcactctg  | 1500 |
| agggtcactt tcaactggcg cctgcctcag tcttactaca tccgcttccg ctacgctagt  | 1560 |
| gtcgctaacc gctctggtac tttccgctac tctcagcctc cgtcttacgg tatctctttc  | 1620 |
| cctaagacta tggacgctgg tgagcctctg accagtagga gcttcgctca cactactctg  | 1680 |
| ttcactccta tcactttctc tagggctcag gaggagttcg acctgtacat ccagtctggt  | 1740 |
| gtgtacatcg acaggatcga gttcatcccc gtgaccgccg cgttcgaggg cgagtacgac  | 1800 |
| cttgagcgcg ccagaaaggc tgtcaatgag ctcttcacgt ccagcaatca gatcggcctg  | 1860 |
| aagaccgacg tcaactgacta ccacatcgac caagtctcca acctcgtgga gtgcctctcc | 1920 |
| gatgagttct gcctcgacga gaagaaggag ctgtccgaga aggtgaagca tgccaagcgt  | 1980 |
| ctcagcgacg agaggaatct cctccaggac cccaatttcc gcggcatcaa caggcagctc  | 2040 |
| gaccgcggct ggcgcgggcag caccgacatc acgatccagg gcggcgacga tgtgttcaag | 2100 |
| gagaactacg tgactctcct gggcactttc gacgagtgtc accctaccta cttgtaccag  | 2160 |
| aagatcgatg agtccaagct caaggcttac actcgctacc agctccgagg ctacatcgaa  | 2220 |
| gacagccaag acctcgagat ttacctgatc cgctacaacg ccaagcacga gaccgtcaac  | 2280 |
| gtgcccggta ctgggtccct ctggccgctg agcgccccca gcccgatcgg caagtgtgcc  | 2340 |
| caccacagcc accacttctc cttggacatc gatgtgggct gcaccgacct gaacgaggac  | 2400 |
| ctcgaggtct gggtcactct caagatcaag acccaggacg gccacgcgcg cctgggcaac  | 2460 |
| ctggagttcc tcgaggagaa gcccttggtc ggtgaggctc tggccagggc caagagggct  | 2520 |
| gagaagaagt ggagggacaa gcgcgagaag ctcgagtggg agaccaacat cgtttacaag  | 2580 |
| gaggccaagg agagcgtcga cgccctgttc gtgaactccc agtacgaccg cctgcaggcc  | 2640 |
| gacaccaaca tcgccatgat ccacgctgcc gacaagaggg tgcacagcat tcgcgaggcc  | 2700 |
| tacctgcctg agctgtccgt gatccctggt gtgaacgctg ccatctttga ggagctggag  | 2760 |
| ggccgcatct ttaccgcatt ctccctgtac gacgcccgcg acgtgatcaa gaacggtgac  | 2820 |
| ttcaacaatg gcctcagctg ctggaacgct aagggccacg tggacgtcga ggaacagaac  | 2880 |
| aaccaccgct ccgtcctggt cgtcccagag tgggaggctg aggtctccca agaggtccgc  | 2940 |
| gtctgcccag gccgcggcta cattctcagg gtcaccgctt acaaggaggg ctacgggtgag | 3000 |

```

ggctgtgtga ccatccacga gatcgagaac aacaccgacg agcttaagtt ctccaactgc 3060
gtggaggagg aggtgtaccc aaacaacacc gttacttgca acgactacac cgccaccag 3120
gaggagtacg agggcaccta cacttccagg aacaggggct acgatggtgc ctacgagagc 3180
aacagcagcg ttctgtctga ctacgcttcc gcctacgagg agaaggccta cacggatggc 3240
cgcagggaca acccttgcca gagcaaccgc ggctacggcg actacactcc cctgcccgcc 3300
ggctacgtta ccaaggagct ggagtacttc ccggagactg acaaggtgtg gatcgagatc 3360
ggcgagaccg agggcacctt catcgtggac agcgtggagc tgctcctgat ggaggagtag 3420

```

<210> 14

<211> 1139

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Trình tự axit amin của protein khảm TIC844.

<400> 14

```

Met Glu Ile Asn Asn Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Ser
1           5           10           15

Asn Pro Lys Glu Ile Ile Leu Gly Glu Glu Arg Leu Glu Thr Gly Asn
20           25           30

Thr Val Ala Asp Ile Ser Leu Gly Leu Ile Asn Phe Leu Tyr Ser Asn
35           40           45

Phe Val Pro Gly Gly Gly Phe Ile Val Gly Leu Leu Glu Leu Ile Trp
50           55           60

Gly Phe Ile Gly Pro Ser Gln Trp Asp Ile Phe Leu Ala Gln Ile Glu
65           70           75           80

Gln Leu Ile Ser Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala Ile
85           90           95

Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Lys Val Tyr Val Arg Ala
100          105          110

Phe Ser Asp Trp Glu Lys Asp Pro Thr Asn Pro Ala Leu Arg Glu Glu
115          120          125

Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Ile Thr Ala Ile
130          135          140

Pro Leu Phe Arg Val Gln Asn Tyr Glu Val Ala Leu Leu Ser Val Tyr
145          150          155          160

Val Gln Ala Ala Asn Leu His Leu Ser Ile Leu Arg Asp Val Ser Val
165          170          175

Phe Gly Glu Arg Trp Gly Tyr Asp Thr Ala Thr Ile Asn Asn Arg Tyr
180          185          190

Ser Asp Leu Thr Ser Leu Ile His Val Tyr Thr Asn His Cys Val Asp
195          200          205

Thr Tyr Asn Gln Gly Leu Arg Arg Leu Glu Gly Arg Phe Leu Ser Asp

```

| 210                                                                                | 215 | 220 |
|------------------------------------------------------------------------------------|-----|-----|
| Trp Ile Val Tyr Asn Arg Phe Arg Arg Gln Leu Thr Ile Ser Val Leu<br>225 230 235 240 |     |     |
| Asp Ile Val Ala Phe Phe Pro Asn Tyr Asp Ile Arg Thr Tyr Pro Ile<br>245 250 255     |     |     |
| Gln Thr Ala Thr Gln Leu Thr Arg Glu Val Tyr Leu Asp Leu Pro Phe<br>260 265 270     |     |     |
| Ile Asn Glu Asn Leu Ser Pro Ala Ala Ser Tyr Pro Thr Phe Ser Ala<br>275 280 285     |     |     |
| Ala Glu Ser Ala Ile Ile Arg Ser Pro His Leu Val Asp Phe Leu Asn<br>290 295 300     |     |     |
| Ser Phe Thr Ile Tyr Thr Asp Ser Leu Ala Arg Tyr Ala Tyr Trp Gly<br>305 310 315 320 |     |     |
| Gly His Leu Val Asn Ser Phe Arg Thr Gly Thr Thr Thr Asn Leu Ile<br>325 330 335     |     |     |
| Arg Ser Pro Leu Tyr Gly Arg Glu Gly Asn Thr Glu Arg Pro Val Thr<br>340 345 350     |     |     |
| Ile Thr Ala Ser Pro Ser Val Pro Ile Phe Arg Thr Leu Ser Tyr Ile<br>355 360 365     |     |     |
| Thr Gly Leu Asp Asn Ser Asn Pro Val Ala Gly Ile Glu Gly Val Glu<br>370 375 380     |     |     |
| Phe Gln Asn Thr Ile Ser Arg Ser Ile Tyr Arg Lys Ser Gly Pro Ile<br>385 390 395 400 |     |     |
| Asp Ser Phe Ser Glu Leu Pro Pro Gln Asp Ala Ser Val Ser Pro Ala<br>405 410 415     |     |     |
| Ile Gly Tyr Ser His Arg Leu Cys His Ala Thr Phe Leu Glu Arg Ile<br>420 425 430     |     |     |
| Ser Gly Pro Arg Ile Ala Gly Thr Val Phe Ser Trp Thr His Arg Ser<br>435 440 445     |     |     |
| Ala Ser Pro Thr Asn Glu Val Ser Pro Ser Arg Ile Thr Gln Ile Pro<br>450 455 460     |     |     |
| Trp Val Lys Ala His Thr Leu Ala Ser Gly Ala Ser Val Ile Lys Gly<br>465 470 475 480 |     |     |
| Pro Gly Phe Thr Gly Gly Asp Ile Leu Thr Arg Asn Ser Met Gly Glu<br>485 490 495     |     |     |
| Leu Gly Thr Leu Arg Val Thr Phe Thr Gly Arg Leu Pro Gln Ser Tyr<br>500 505 510     |     |     |
| Tyr Ile Arg Phe Arg Tyr Ala Ser Val Ala Asn Arg Ser Gly Thr Phe<br>515 520 525     |     |     |
| Arg Tyr Ser Gln Pro Pro Ser Tyr Gly Ile Ser Phe Pro Lys Thr Met<br>530 535 540     |     |     |
| Asp Ala Gly Glu Pro Leu Thr Ser Arg Ser Phe Ala His Thr Thr Leu<br>545 550 555 560 |     |     |

Phe Thr Pro Ile Thr Phe Ser Arg Ala Gln Glu Glu Phe Asp Leu Tyr  
 565 570 575  
 Ile Gln Ser Gly Val Tyr Ile Asp Arg Ile Glu Phe Ile Pro Val Thr  
 580 585 590  
 Ala Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala Gln Lys Ala Val  
 595 600 605  
 Asn Glu Leu Phe Thr Ser Ser Asn Gln Ile Gly Leu Lys Thr Asp Val  
 610 615 620  
 Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu Ser  
 625 630 635 640  
 Asp Glu Phe Cys Leu Asp Glu Lys Lys Glu Leu Ser Glu Lys Val Lys  
 645 650 655  
 His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn  
 660 665 670  
 Phe Arg Gly Ile Asn Arg Gln Leu Asp Arg Gly Trp Arg Gly Ser Thr  
 675 680 685  
 Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val  
 690 695 700  
 Thr Leu Leu Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln  
 705 710 715 720  
 Lys Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg  
 725 730 735  
 Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr  
 740 745 750  
 Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp  
 755 760 765  
 Pro Leu Ser Ala Pro Ser Pro Ile Gly Lys Cys Ala His His Ser His  
 770 775 780  
 His Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp  
 785 790 795 800  
 Leu Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala  
 805 810 815  
 Arg Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val Gly Glu  
 820 825 830  
 Ala Leu Ala Arg Val Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg  
 835 840 845  
 Glu Lys Leu Glu Trp Glu Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu  
 850 855 860  
 Ser Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Arg Leu Gln Ala  
 865 870 875 880  
 Asp Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg Val His Ser  
 885 890 895

Ile Arg Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile Pro Gly Val Asn  
 900 905 910  
 Ala Ala Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser  
 915 920 925  
 Leu Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly  
 930 935 940  
 Leu Ser Cys Trp Asn Val Lys Gly His Val Asp Val Glu Glu Gln Asn  
 945 950 955 960  
 Asn His Arg Ser Val Leu Val Val Pro Glu Trp Glu Ala Glu Val Ser  
 965 970 975  
 Gln Glu Val Arg Val Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr  
 980 985 990  
 Ala Tyr Lys Glu Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile  
 995 1000 1005  
 Glu Asn Asn Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu  
 1010 1015 1020  
 Glu Val Tyr Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Ala  
 1025 1030 1035  
 Thr Gln Glu Glu Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly  
 1040 1045 1050  
 Tyr Asp Gly Ala Tyr Glu Ser Asn Ser Ser Val Pro Ala Asp Tyr  
 1055 1060 1065  
 Ala Ser Ala Tyr Glu Glu Lys Ala Tyr Thr Asp Gly Arg Arg Asp  
 1070 1075 1080  
 Asn Pro Cys Glu Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu  
 1085 1090 1095  
 Pro Ala Gly Tyr Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr  
 1100 1105 1110  
 Asp Lys Val Trp Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile  
 1115 1120 1125  
 Val Asp Ser Val Glu Leu Leu Leu Met Glu Glu  
 1130 1135

<210> 15

<211> 3420

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Trình tự nucleotit được dùng để biểu hiện ở tế bào vi khuẩn ghi mã TIC844\_2.

<400> 15

atggaaataa ataatcaaaa ccaatgtgtg cottacaatt gtttaagtaa tcctaaggag 60

ataatattag gcgaggaaag gctagaaaca gggaatactg tagcagacat ttcattaggg 120

ottattaatt ttctatattc taattttgta ccaggaggag gatttatagt aggtttacta 180

|            |            |             |            |             |             |      |
|------------|------------|-------------|------------|-------------|-------------|------|
| gaattaatat | ggggatttat | agggccttcg  | caatgggata | tttttttagc  | tcaaattgag  | 240  |
| caattgatta | gtcaaagaat | agaagaat    | gctaggaatc | aggcaatttc  | aagattggag  | 300  |
| gggctaagca | atctttataa | ggtctatggt  | agagcgttta | gcgactggga  | gaaagatcct  | 360  |
| actaatcctg | ctttaaggga | agaaatgcgt  | atacaattta | atgacatgaa  | tagtgctctc  | 420  |
| ataacggcta | ttccactttt | tagagttcaa  | aattatgaag | ttgctctttt  | atctgtatat  | 480  |
| gttcaagcgg | caaacttaca | tttatctatt  | ttaagggatg | tttcagtttt  | cggagaaaga  | 540  |
| tggggatatg | atacagcgac | tatcaataat  | cgctatagtg | atctgactag  | ccttattcat  | 600  |
| gtttatacta | accattgtgt | ggatacgtat  | aatcagggat | taaggcgttt  | ggaaggctgt  | 660  |
| tttcttagcg | attggattgt | atataatcgt  | ttccggagac | aattgacaat  | ttcagtatta  | 720  |
| gatattgttg | cgttttttcc | aaattatgat  | attagaacat | atccaattca  | aacagctact  | 780  |
| cagctaacga | gggaagtcta | tctggattta  | ccttttatta | atgaaaatct  | ttctcctgca  | 840  |
| gcaagctatc | caaccttttc | agctgctgaa  | agtgcataaa | ttagaagtcc  | tcatttagta  | 900  |
| gactttttta | atagctttac | catttataca  | gatagtctgg | cacgttctgc  | atattgggga  | 960  |
| gggcacttgg | taaattcttt | ccgcacagga  | accactacta | atttgataag  | atccccttta  | 1020 |
| tatggaaggg | aaggaaatac | agagcgcccc  | gtaactatta | ccgcatacacc | tagcgtacca  | 1080 |
| atatttagaa | cactttcata | tattacaggc  | cttgacaatt | cacatcctgt  | agctggaatc  | 1140 |
| gagggagtgg | aattccaaaa | tactataagt  | agaagtatct | atcgtaaaag  | cggccaataa  | 1200 |
| gattctttta | gtgaattacc | acctcaagat  | gccagcgtat | ctcctgcaat  | tgggtatagt  | 1260 |
| caccgtttat | gccatgcaac | attttttagaa | cggttaagtg | gaccaagaat  | agcaggcacc  | 1320 |
| gtattttctt | ggacacaccg | tagtgccagc  | cctactaacg | aagtaagtcc  | atctagaatt  | 1380 |
| acacaaattc | catgggtaaa | ggcgcatact  | cttgcatctg | gtgcctccgt  | cattaaaggt  | 1440 |
| cctggattta | caggtggaga | tattctgact  | aggaatagta | tgggcgagct  | ggggacctta  | 1500 |
| cgagtaacct | tcacaggaag | attaccacaa  | agttattata | tacgtttccg  | ttatgcttcg  | 1560 |
| gtagcaaata | ggagtgttac | atttagatat  | tcacagccac | cttcgtatgg  | aatttcattt  | 1620 |
| ccaaaaacta | tggacgcagg | tgaaccacta  | acatctcggt | cgttcgctca  | tacaacactc  | 1680 |
| ttcactccaa | taaccttttc | acgagctcaa  | gaagaatttg | atctatacat  | ccaatcgggt  | 1740 |
| gtttatatag | atcgaattga | atttataccg  | gttactgcaa | catttgaggc  | agaatatgat  | 1800 |
| ttagaaagag | cacaaaaggc | ggtgaatgag  | ctgtttactt | cttccaatca  | aatcgggtta  | 1860 |
| aaaacagatg | tgacggatta | tcatatgtat  | caagtatcca | atttagttga  | gtgtttatct  | 1920 |
| gatgaatttt | gtctggatga | aaaaaaagaa  | ttgtccgaga | aagtcaaaca  | tgcgaaagcga | 1980 |
| cttagtgatg | agcggaat   | acttcaagat  | ccaaacttta | gagggatcaa  | tagacaacta  | 2040 |
| gaccgtggct | ggagaggaag | tacggatatt  | accatccaag | gaggcgatga  | cgtattcaaa  | 2100 |

gagaattacg ttacgctatt gggtagcttt gatgagtgt atccaacgta tttatatcaa 2160  
 aaaatagatg agtcgaaatt aaaagcctat acccgttacc aattaagagg gtatatcgaa 2220  
 gatagtcaag acttagaaat ctatttaatt cgctacaatg ccaaacacga aacagtaaatt 2280  
 gtgccaggta cgggttcctt atggccgctt tcagcccaa gtccaatcg aaatgtgcc 2340  
 catcattccc atcatttctc cttggacatt gatgttggat gtacagactt aaatgaggac 2400  
 ttaggtgtat ggggtgatatt caagattaag acgcaagatg gccatgcaag actaggaaat 2460  
 ctagaatttc tcgaagagaa accattagta ggagaagcac tagctcgtgt gaaaagagcg 2520  
 gagaaaaaat ggagagacaa acgtgaaaaa ttggaatggg aaacaaatat tgtttataaa 2580  
 gaggcaaaag aatctgtaga tgctttattt gtaaaactctc aatatgatag attacaagcg 2640  
 gataccaaca tcgcgatgat tcatgcggca gataaacgcg ttcatagcat tcgagaagct 2700  
 tatctgcctg agctgtctgt gattccgggt gtcaatgcgg ctatTTTTga agaattagaa 2760  
 gggcgtatTT tctactgcatt ctccctatat gatgcgagaa atgtcattaa aaatgggtgat 2820  
 ttttaataatg gcttatcctg ctggaacgtg aaagggcatg tagatgtaga agaacaaaac 2880  
 aaccaccgtt cggtccttgt tggtccggaa tgggaagcag aagtgtcaca agaagttcgt 2940  
 gtctgtccgg gtcgtggcta tctcttctgt gtcacagcgt acaaggagg atatggagaa 3000  
 ggttgcgtaa ccattcatga gatcgagaac aatacagacg aactgaagtt tagcaactgt 3060  
 gtagaagagg aagtatatcc aaacaacacg gtaacgtgta atgattatac tgcgactcaa 3120  
 gaagaatatg agggtagcta cacttctcgt aatcgaggat atgacggagc ctatgaaagc 3180  
 aattcttctg taccagctga ttatgcatca gcctatgaag aaaaagcata tacagatgga 3240  
 cgaagagaca atccttgtga atctaacaga ggatatgggg attacacacc actaccagct 3300  
 ggctatgtga caaaagaatt agagtacttc ccagaaaccg ataaggtatg gattgagatc 3360  
 ggagaaacgg aaggaacatt catcgtggac agcgtggaat tacttcttat ggaggaatag 3420

<210> 16

<211> 1139

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Trình tự axit amin của protein khám đã xử lý có hoạt tính trừ côn trùng TIC844\_2.

<400> 16

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Glu | Ile | Asn | Asn | Gln | Asn | Gln | Cys | Val | Pro | Tyr | Asn | Cys | Leu | Ser |
| 1   |     |     |     | 5   |     |     |     | 10  |     |     |     |     |     | 15  |     |
| Asn | Pro | Lys | Glu | Ile | Ile | Leu | Gly | Glu | Glu | Arg | Leu | Glu | Thr | Gly | Asn |
|     |     |     | 20  |     |     |     |     | 25  |     |     |     |     |     | 30  |     |
| Thr | Val | Ala | Asp | Ile | Ser | Leu | Gly | Leu | Ile | Asn | Phe | Leu | Tyr | Ser | Asn |
|     |     | 35  |     |     |     |     | 40  |     |     |     |     |     | 45  |     |     |

Phe Val Pro Gly Gly Gly Phe Ile Val Gly Leu Leu Glu Leu Ile Trp  
 50 55 60  
 Gly Phe Ile Gly Pro Ser Gln Trp Asp Ile Phe Leu Ala Gln Ile Glu  
 65 70 75 80  
 Gln Leu Ile Ser Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala Ile  
 85 90 95  
 Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Lys Val Tyr Val Arg Ala  
 100 105 110  
 Phe Ser Asp Trp Glu Lys Asp Pro Thr Asn Pro Ala Leu Arg Glu Glu  
 115 120 125  
 Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Ile Thr Ala Ile  
 130 135 140  
 Pro Leu Phe Arg Val Gln Asn Tyr Glu Val Ala Leu Leu Ser Val Tyr  
 145 150 155 160  
 Val Gln Ala Ala Asn Leu His Leu Ser Ile Leu Arg Asp Val Ser Val  
 165 170 175  
 Phe Gly Glu Arg Trp Gly Tyr Asp Thr Ala Thr Ile Asn Asn Arg Tyr  
 180 185 190  
 Ser Asp Leu Thr Ser Leu Ile His Val Tyr Thr Asn His Cys Val Asp  
 195 200 205  
 Thr Tyr Asn Gln Gly Leu Arg Arg Leu Glu Gly Arg Phe Leu Ser Asp  
 210 215 220  
 Trp Ile Val Tyr Asn Arg Phe Arg Arg Gln Leu Thr Ile Ser Val Leu  
 225 230 235 240  
 Asp Ile Val Ala Phe Phe Pro Asn Tyr Asp Ile Arg Thr Tyr Pro Ile  
 245 250 255  
 Gln Thr Ala Thr Gln Leu Thr Arg Glu Val Tyr Leu Asp Leu Pro Phe  
 260 265 270  
 Ile Asn Glu Asn Leu Ser Pro Ala Ala Ser Tyr Pro Thr Phe Ser Ala  
 275 280 285  
 Ala Glu Ser Ala Ile Ile Arg Ser Pro His Leu Val Asp Phe Leu Asn  
 290 295 300  
 Ser Phe Thr Ile Tyr Thr Asp Ser Leu Ala Arg Ser Ala Tyr Trp Gly  
 305 310 315 320  
 Gly His Leu Val Asn Ser Phe Arg Thr Gly Thr Thr Thr Asn Leu Ile  
 325 330 335  
 Arg Ser Pro Leu Tyr Gly Arg Glu Gly Asn Thr Glu Arg Pro Val Thr  
 340 345 350  
 Ile Thr Ala Ser Pro Ser Val Pro Ile Phe Arg Thr Leu Ser Tyr Ile  
 355 360 365  
 Thr Gly Leu Asp Asn Ser His Pro Val Ala Gly Ile Glu Gly Val Glu  
 370 375 380  
 Phe Gln Asn Thr Ile Ser Arg Ser Ile Tyr Arg Lys Ser Gly Pro Ile

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 385 |     |     |     |     | 390 |     |     |     |     | 395 |     |     |     | 400 |     |
| Asp | Ser | Phe | Ser | Glu | Leu | Pro | Pro | Gln | Asp | Ala | Ser | Val | Ser | Pro | Ala |
|     |     |     |     | 405 |     |     |     |     | 410 |     |     |     |     | 415 |     |
| Ile | Gly | Tyr | Ser | His | Arg | Leu | Cys | His | Ala | Thr | Phe | Leu | Glu | Arg | Leu |
|     |     |     | 420 |     |     |     |     | 425 |     |     |     |     | 430 |     |     |
| Ser | Gly | Pro | Arg | Ile | Ala | Gly | Thr | Val | Phe | Ser | Trp | Thr | His | Arg | Ser |
|     |     | 435 |     |     |     |     | 440 |     |     |     |     | 445 |     |     |     |
| Ala | Ser | Pro | Thr | Asn | Glu | Val | Ser | Pro | Ser | Arg | Ile | Thr | Gln | Ile | Pro |
|     | 450 |     |     |     |     | 455 |     |     |     |     | 460 |     |     |     |     |
| Trp | Val | Lys | Ala | His | Thr | Leu | Ala | Ser | Gly | Ala | Ser | Val | Ile | Lys | Gly |
| 465 |     |     |     |     | 470 |     |     |     |     | 475 |     |     |     |     | 480 |
| Pro | Gly | Phe | Thr | Gly | Gly | Asp | Ile | Leu | Thr | Arg | Asn | Ser | Met | Gly | Glu |
|     |     |     |     | 485 |     |     |     |     | 490 |     |     |     |     | 495 |     |
| Leu | Gly | Thr | Leu | Arg | Val | Thr | Phe | Thr | Gly | Arg | Leu | Pro | Gln | Ser | Tyr |
|     |     |     | 500 |     |     |     |     | 505 |     |     |     |     | 510 |     |     |
| Tyr | Ile | Arg | Phe | Arg | Tyr | Ala | Ser | Val | Ala | Asn | Arg | Ser | Gly | Thr | Phe |
|     |     | 515 |     |     |     |     | 520 |     |     |     |     | 525 |     |     |     |
| Arg | Tyr | Ser | Gln | Pro | Pro | Ser | Tyr | Gly | Ile | Ser | Phe | Pro | Lys | Thr | Met |
|     | 530 |     |     |     |     | 535 |     |     |     |     | 540 |     |     |     |     |
| Asp | Ala | Gly | Glu | Pro | Leu | Thr | Ser | Arg | Ser | Phe | Ala | His | Thr | Thr | Leu |
| 545 |     |     |     |     | 550 |     |     |     |     | 555 |     |     |     |     | 560 |
| Phe | Thr | Pro | Ile | Thr | Phe | Ser | Arg | Ala | Gln | Glu | Glu | Phe | Asp | Leu | Tyr |
|     |     |     |     | 565 |     |     |     |     | 570 |     |     |     |     | 575 |     |
| Ile | Gln | Ser | Gly | Val | Tyr | Ile | Asp | Arg | Ile | Glu | Phe | Ile | Pro | Val | Thr |
|     |     |     | 580 |     |     |     |     | 585 |     |     |     |     | 590 |     |     |
| Ala | Thr | Phe | Glu | Ala | Glu | Tyr | Asp | Leu | Glu | Arg | Ala | Gln | Lys | Ala | Val |
|     |     | 595 |     |     |     |     | 600 |     |     |     |     | 605 |     |     |     |
| Asn | Glu | Leu | Phe | Thr | Ser | Ser | Asn | Gln | Ile | Gly | Leu | Lys | Thr | Asp | Val |
|     | 610 |     |     |     |     | 615 |     |     |     |     | 620 |     |     |     |     |
| Thr | Asp | Tyr | His | Ile | Asp | Gln | Val | Ser | Asn | Leu | Val | Glu | Cys | Leu | Ser |
| 625 |     |     |     |     | 630 |     |     |     |     | 635 |     |     |     |     | 640 |
| Asp | Glu | Phe | Cys | Leu | Asp | Glu | Lys | Lys | Glu | Leu | Ser | Glu | Lys | Val | Lys |
|     |     |     |     | 645 |     |     |     |     | 650 |     |     |     |     | 655 |     |
| His | Ala | Lys | Arg | Leu | Ser | Asp | Glu | Arg | Asn | Leu | Leu | Gln | Asp | Pro | Asn |
|     |     |     | 660 |     |     |     |     | 665 |     |     |     |     | 670 |     |     |
| Phe | Arg | Gly | Ile | Asn | Arg | Gln | Leu | Asp | Arg | Gly | Trp | Arg | Gly | Ser | Thr |
|     |     | 675 |     |     |     |     | 680 |     |     |     |     | 685 |     |     |     |
| Asp | Ile | Thr | Ile | Gln | Gly | Gly | Asp | Asp | Val | Phe | Lys | Glu | Asn | Tyr | Val |
|     | 690 |     |     |     |     | 695 |     |     |     |     | 700 |     |     |     |     |
| Thr | Leu | Leu | Gly | Thr | Phe | Asp | Glu | Cys | Tyr | Pro | Thr | Tyr | Leu | Tyr | Gln |
| 705 |     |     |     |     | 710 |     |     |     |     | 715 |     |     |     |     | 720 |
| Lys | Ile | Asp | Glu | Ser | Lys | Leu | Lys | Ala | Tyr | Thr | Arg | Tyr | Gln | Leu | Arg |
|     |     |     |     | 725 |     |     |     |     | 730 |     |     |     |     | 735 |     |

Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr  
 740 745 750  
 Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp  
 755 760 765  
 Pro Leu Ser Ala Pro Ser Pro Ile Gly Lys Cys Ala His His Ser His  
 770 775 780  
 His Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp  
 785 790 795 800  
 Leu Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala  
 805 810 815  
 Arg Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val Gly Glu  
 820 825 830  
 Ala Leu Ala Arg Val Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg  
 835 840 845  
 Glu Lys Leu Glu Trp Glu Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu  
 850 855 860  
 Ser Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Arg Leu Gln Ala  
 865 870 875 880  
 Asp Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg Val His Ser  
 885 890 895  
 Ile Arg Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile Pro Gly Val Asn  
 900 905 910  
 Ala Ala Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser  
 915 920 925  
 Leu Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly  
 930 935 940  
 Leu Ser Cys Trp Asn Val Lys Gly His Val Asp Val Glu Glu Gln Asn  
 945 950 955 960  
 Asn His Arg Ser Val Leu Val Val Pro Glu Trp Glu Ala Glu Val Ser  
 965 970 975  
 Gln Glu Val Arg Val Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr  
 980 985 990  
 Ala Tyr Lys Glu Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile  
 995 1000 1005  
 Glu Asn Asn Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu  
 1010 1015 1020  
 Glu Val Tyr Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Ala  
 1025 1030 1035  
 Thr Gln Glu Glu Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly  
 1040 1045 1050  
 Tyr Asp Gly Ala Tyr Glu Ser Asn Ser Ser Val Pro Ala Asp Tyr  
 1055 1060 1065

Ala Ser Ala Tyr Glu Glu Lys Ala Tyr Thr Asp Gly Arg Arg Asp  
 1070 1075 1080

Asn Pro Cys Glu Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu  
 1085 1090 1095

Pro Ala Gly Tyr Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr  
 1100 1105 1110

Asp Lys Val Trp Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile  
 1115 1120 1125

Val Asp Ser Val Glu Leu Leu Leu Met Glu Glu  
 1130 1135

<210> 17  
 <211> 3420  
 <212> ADN  
 <213> Trình tự nhân tạo

<220>  
 <223> Trình tự nucleotit được dùng để biểu hiện ở tế bào vi khuẩn  
 ghi mã TIC844\_4.

<400> 17  
 atggaaataa ataatcaaaa ccaatgtgtg ccttacaatt gtttaagtaa tcctaaggag 60  
 ataattattag gcgaggaaaag gctagaaaca gggaatactg tagcagacat ttcattaggg 120  
 cttattaatt ttctatattc taattttgta ccaggaggag gatttatagt aggtttacta 180  
 gaattaatat ggggatttat agggccttcg caatgggata tttttttagc tcaaattgag 240  
 caattgatta gtcaaagaat agaagaatgt gctaggaatc aggcaatttc aagattggag 300  
 gggctaagca atctttataa ggtctatgtt agagcgttta gcgactggga gaaagatcct 360  
 actaatcctg ctttaaggga agaaatgcgt atacaattta atgacatgaa tagtgctctc 420  
 ataacggcta ttccactttt tagagttcaa aattatgaag ttgctctttt atctgtatat 480  
 gttcaagccg caaacttaca tttatctatt ttaagggatg tttcagtttt cggagaaaga 540  
 tggggatatg atacagcgac tatcaataat cgctatagtg atctgactag ccttattcat 600  
 gtttatacta accattgtgt ggatacgtat aatcagggat taaggcgttt ggaaggctgt 660  
 tttcttagcg attggattgt atataatcgt ttccggagac aattgacaat ttcagtatta 720  
 gatattgttg cgttttttcc aaattatgat attagaacat atccaattca aacagctact 780  
 cagctaacga gggaagtcta tctggattta ccttttatta atgaaaatct ttctcctgca 840  
 gcaagctatc caaccttttc agctgctgaa agtgcataaa ttagaagtcc tcatttagta 900  
 gactttttta atagctttac catttataca gatagtctgg cacgtagtgc atattgggga 960  
 gggcacttgg taaattcttt ccgcacagga accactacta atttgataag atccccttta 1020  
 tatggaaggg aaggaaatac agagcgcccc gtaactatta ccgcatacacc tagcgtacca 1080  
 atatttagaa cactttcata tattacaggc cttgacaatt caaatcctgt agctggaatc 1140  
 gagggagtgg aattccaaaa tactataagt agaagtatct atcgtaaaag cggccaata 1200

|            |             |             |            |            |             |      |
|------------|-------------|-------------|------------|------------|-------------|------|
| gattctttta | gtgaattacc  | acctcaagat  | gccagcgtat | ctcctgcaat | tgggtatagt  | 1260 |
| caccgtttat | gccatgcaac  | atTTTTtagaa | cggattagt  | gaccaagaat | agcaggcacc  | 1320 |
| gtatTTTctt | ggacacaccg  | tagtgccagc  | cctactaacg | aagtaagtcc | atctagaatt  | 1380 |
| acacaaattc | catgggtaaa  | ggcgcatact  | cttgcatctg | gtgcctccgt | cattaaaggt  | 1440 |
| cctggattta | caggtggaga  | tattctgact  | aggaatagta | tgggcgagct | ggggacctta  | 1500 |
| cgagtaacct | tcacaggaag  | attaccacaa  | agttattata | tacgtttccg | ttatgcttcg  | 1560 |
| gtagcaaata | ggagtgttac  | atttagatat  | tcacagccac | cttcgtatgg | aatttcattt  | 1620 |
| ccaaaaacta | tggacgcagg  | tgaaccacta  | acatctcggt | cgttcgctca | tacaacactc  | 1680 |
| ttcactccaa | taaccttttc  | acgagctcaa  | gaagaatttg | atctatacat | ccaatcgggt  | 1740 |
| gtttatatag | atcgaattga  | atttataccg  | gttactgcaa | catttgaggc | agaatatgat  | 1800 |
| ttagaaagag | cacaaaaggc  | ggtgaatgag  | ctgtttactt | cttccaatca | aatcgggtta  | 1860 |
| aaaacagatg | tgacggatta  | tcatatgtat  | caagtatcca | atttagttga | gtgtttatct  | 1920 |
| gatgaatttt | gtctggatga  | aaaaaaagaa  | ttgtccgaga | aagtcaaaca | tgcggaagcga | 1980 |
| cttagtgatg | agcggaaattt | acttcaagat  | ccaaacttta | gagggatcaa | tagacaacta  | 2040 |
| gaccgtggct | ggagaggaag  | tacggatatt  | accatccaag | gaggcgatga | cgtattcaaa  | 2100 |
| gagaattacg | ttacgctatt  | gggtaccttt  | gatgagtgt  | atccaacgta | tttatatcaa  | 2160 |
| aaaatagatg | agtcgaaatt  | aaaagcctat  | acccgttacc | aattaagagg | gtatatcgaa  | 2220 |
| gatagtcaag | acttagaaat  | ctatttaatt  | cgctacaatg | ccaaacacga | aacagtaa    | 2280 |
| gtgccaggta | cgggttcctt  | atggccgctt  | tcagccccaa | gtccaatcgg | aaaatgtgcc  | 2340 |
| catcattccc | atcatttctc  | cttggacatt  | gatgttggat | gtacagactt | aaatgaggac  | 2400 |
| ttaggtgtat | gggtgatatt  | caagattaag  | acgcaagatg | gccatgcaag | actaggaaat  | 2460 |
| ctagaatttc | tcgaagagaa  | accattagta  | ggagaagcac | tagctcgtgt | gaaaagagcg  | 2520 |
| gagaaaaaat | ggagagacaa  | acgtgaaaaa  | ttggaatggg | aaacaaatat | tgtttataaa  | 2580 |
| gaggcaaaag | aatctgtaga  | tgctttat    | gtaaactctc | aatatgatag | attacaagcg  | 2640 |
| gataccaaca | tcgcgatgat  | tcatgcggca  | gataaacgcg | ttcatagcat | tcgagaagct  | 2700 |
| tatctgcctg | agctgtctgt  | gattccgggt  | gtcaatgcgg | ctatTTTTga | agaattagaa  | 2760 |
| gggcgtat   | ttactgcatt  | ctccctatat  | gatgcgagaa | atgtcattaa | aaatgggtgat | 2820 |
| tttaataatg | gcttatcctg  | ctggaacgtg  | aaagggcatg | tagatgtaga | agaacaaaac  | 2880 |
| aaccaccgtt | cggtccttgt  | tgttccggaa  | tgggaagcag | aagtgtcaca | agaagtctgt  | 2940 |
| gtctgtccgg | gtcgtggcta  | tatccttcgt  | gtcacagcgt | acaaggaggg | atatggagaa  | 3000 |
| ggttgcgtaa | ccattcatga  | gatcgagaac  | aatacagacg | aactgaagtt | tagcaactgt  | 3060 |
| gtagaagagg | aagtatatcc  | aaacaacacg  | gtaacgtgta | atgattatac | tgcgactcaa  | 3120 |

gaagaatatg aggggtacgta cacttctcgt aatcgaggat atgacggagc ctatgaaagc 3180  
aattcttctg taccagctga ttatgcatca gcctatgaag aaaaagcata tacagatgga 3240  
cgaagagaca atccttgtga atctaacaga ggatatgggg attacacacc actaccagct 3300  
ggctatgtga caaaagaatt agagtacttc ccagaaaccg ataaggtatg gattgagatc 3360  
ggagaaacgg aaggaacatt catcgtggac agcgtggaat tacttcttat ggaggaatag 3420

<210> 18

<211> 1139

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Trình tự axit amin của protein khám đã xử lý có hoạt tính trừ côn trùng TIC844\_4.

<400> 18

Met Glu Ile Asn Asn Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Ser  
1 5 10 15

Asn Pro Lys Glu Ile Ile Leu Gly Glu Glu Arg Leu Glu Thr Gly Asn  
20 25 30

Thr Val Ala Asp Ile Ser Leu Gly Leu Ile Asn Phe Leu Tyr Ser Asn  
35 40 45

Phe Val Pro Gly Gly Gly Phe Ile Val Gly Leu Leu Glu Leu Ile Trp  
50 55 60

Gly Phe Ile Gly Pro Ser Gln Trp Asp Ile Phe Leu Ala Gln Ile Glu  
65 70 75 80

Gln Leu Ile Ser Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala Ile  
85 90 95

Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Lys Val Tyr Val Arg Ala  
100 105 110

Phe Ser Asp Trp Glu Lys Asp Pro Thr Asn Pro Ala Leu Arg Glu Glu  
115 120 125

Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Ile Thr Ala Ile  
130 135 140

Pro Leu Phe Arg Val Gln Asn Tyr Glu Val Ala Leu Leu Ser Val Tyr  
145 150 155 160

Val Gln Ala Ala Asn Leu His Leu Ser Ile Leu Arg Asp Val Ser Val  
165 170 175

Phe Gly Glu Arg Trp Gly Tyr Asp Thr Ala Thr Ile Asn Asn Arg Tyr  
180 185 190

Ser Asp Leu Thr Ser Leu Ile His Val Tyr Thr Asn His Cys Val Asp  
195 200 205

Thr Tyr Asn Gln Gly Leu Arg Arg Leu Glu Gly Arg Phe Leu Ser Asp  
210 215 220

Trp Ile Val Tyr Asn Arg Phe Arg Arg Gln Leu Thr Ile Ser Val Leu  
 225 230 235 240  
 Asp Ile Val Ala Phe Phe Pro Asn Tyr Asp Ile Arg Thr Tyr Pro Ile  
 245 250 255  
 Gln Thr Ala Thr Gln Leu Thr Arg Glu Val Tyr Leu Asp Leu Pro Phe  
 260 265 270  
 Ile Asn Glu Asn Leu Ser Pro Ala Ala Ser Tyr Pro Thr Phe Ser Ala  
 275 280 285  
 Ala Glu Ser Ala Ile Ile Arg Ser Pro His Leu Val Asp Phe Leu Asn  
 290 295 300  
 Ser Phe Thr Ile Tyr Thr Asp Ser Leu Ala Arg Ser Ala Tyr Trp Gly  
 305 310 315 320  
 Gly His Leu Val Asn Ser Phe Arg Thr Gly Thr Thr Thr Asn Leu Ile  
 325 330 335  
 Arg Ser Pro Leu Tyr Gly Arg Glu Gly Asn Thr Glu Arg Pro Val Thr  
 340 345 350  
 Ile Thr Ala Ser Pro Ser Val Pro Ile Phe Arg Thr Leu Ser Tyr Ile  
 355 360 365  
 Thr Gly Leu Asp Asn Ser Asn Pro Val Ala Gly Ile Glu Gly Val Glu  
 370 375 380  
 Phe Gln Asn Thr Ile Ser Arg Ser Ile Tyr Arg Lys Ser Gly Pro Ile  
 385 390 395 400  
 Asp Ser Phe Ser Glu Leu Pro Pro Gln Asp Ala Ser Val Ser Pro Ala  
 405 410 415  
 Ile Gly Tyr Ser His Arg Leu Cys His Ala Thr Phe Leu Glu Arg Ile  
 420 425 430  
 Ser Gly Pro Arg Ile Ala Gly Thr Val Phe Ser Trp Thr His Arg Ser  
 435 440 445  
 Ala Ser Pro Thr Asn Glu Val Ser Pro Ser Arg Ile Thr Gln Ile Pro  
 450 455 460  
 Trp Val Lys Ala His Thr Leu Ala Ser Gly Ala Ser Val Ile Lys Gly  
 465 470 475 480  
 Pro Gly Phe Thr Gly Gly Asp Ile Leu Thr Arg Asn Ser Met Gly Glu  
 485 490 495  
 Leu Gly Thr Leu Arg Val Thr Phe Thr Gly Arg Leu Pro Gln Ser Tyr  
 500 505 510  
 Tyr Ile Arg Phe Arg Tyr Ala Ser Val Ala Asn Arg Ser Gly Thr Phe  
 515 520 525  
 Arg Tyr Ser Gln Pro Pro Ser Tyr Gly Ile Ser Phe Pro Lys Thr Met  
 530 535 540  
 Asp Ala Gly Glu Pro Leu Thr Ser Arg Ser Phe Ala His Thr Thr Leu  
 545 550 555 560  
 Phe Thr Pro Ile Thr Phe Ser Arg Ala Gln Glu Glu Phe Asp Leu Tyr

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|--|--|--|-----|--|--|--|--|--|--|--|--|--|--|-----|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|
| 565 |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     | 570 |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  | 575 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Ile | Gln | Ser | Gly | Val | Tyr | Ile | Asp | Arg | Ile | Glu | Phe | Ile | Pro | Val | Thr |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|     |     |     | 580 |     |     |     |     |     | 585 |     |     |     |     |     | 590 |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Ala | Thr | Phe | Glu | Ala | Glu | Tyr | Asp | Leu | Glu | Arg | Ala | Gln | Lys | Ala | Val |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|     |     |     | 595 |     |     |     |     |     | 600 |     |     |     |     |     | 605 |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Asn | Glu | Leu | Phe | Thr | Ser | Ser | Asn | Gln | Ile | Gly | Leu | Lys | Thr | Asp | Val |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|     |     |     | 610 |     |     |     |     |     | 615 |     |     |     |     |     | 620 |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Thr | Asp | Tyr | His | Ile | Asp | Gln | Val | Ser | Asn | Leu | Val | Glu | Cys | Leu | Ser |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|     |     |     | 625 |     |     |     |     |     | 630 |     |     |     |     |     | 635 |     |  |  |  |  | 640 |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Asp | Glu | Phe | Cys | Leu | Asp | Glu | Lys | Lys | Glu | Leu | Ser | Glu | Lys | Val | Lys |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|     |     |     | 645 |     |     |     |     |     | 650 |     |     |     |     |     | 655 |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| His | Ala | Lys | Arg | Leu | Ser | Asp | Glu | Arg | Asn | Leu | Leu | Gln | Asp | Pro | Asn |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|     |     |     | 660 |     |     |     |     |     | 665 |     |     |     |     |     | 670 |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Phe | Arg | Gly | Ile | Asn | Arg | Gln | Leu | Asp | Arg | Gly | Trp | Arg | Gly | Ser | Thr |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|     |     |     | 675 |     |     |     |     |     | 680 |     |     |     |     |     | 685 |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Asp | Ile | Thr | Ile | Gln | Gly | Gly | Asp | Asp | Val | Phe | Lys | Glu | Asn | Tyr | Val |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|     |     |     | 690 |     |     |     |     |     | 695 |     |     |     |     |     | 700 |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Thr | Leu | Leu | Gly | Thr | Phe | Asp | Glu | Cys | Tyr | Pro | Thr | Tyr | Leu | Tyr | Gln |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|     |     |     | 705 |     |     |     |     |     | 710 |     |     |     |     |     | 715 |     |  |  |  |  | 720 |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Lys | Ile | Asp | Glu | Ser | Lys | Leu | Lys | Ala | Tyr | Thr | Arg | Tyr | Gln | Leu | Arg |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|     |     |     | 725 |     |     |     |     |     | 730 |     |     |     |     |     | 735 |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Gly | Tyr | Ile | Glu | Asp | Ser | Gln | Asp | Leu | Glu | Ile | Tyr | Leu | Ile | Arg | Tyr |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|     |     |     | 740 |     |     |     |     |     | 745 |     |     |     |     |     | 750 |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Asn | Ala | Lys | His | Glu | Thr | Val | Asn | Val | Pro | Gly | Thr | Gly | Ser | Leu | Trp |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|     |     |     | 755 |     |     |     |     |     | 760 |     |     |     |     |     | 765 |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Pro | Leu | Ser | Ala | Pro | Ser | Pro | Ile | Gly | Lys | Cys | Ala | His | His | Ser | His |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|     |     |     | 770 |     |     |     |     |     | 775 |     |     |     |     |     | 780 |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| His | Phe | Ser | Leu | Asp | Ile | Asp | Val | Gly | Cys | Thr | Asp | Leu | Asn | Glu | Asp |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|     |     |     | 785 |     |     |     |     |     | 790 |     |     |     |     |     | 795 |     |  |  |  |  | 800 |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Leu | Gly | Val | Trp | Val | Ile | Phe | Lys | Ile | Lys | Thr | Gln | Asp | Gly | His | Ala |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|     |     |     | 805 |     |     |     |     |     | 810 |     |     |     |     |     | 815 |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Arg | Leu | Gly | Asn | Leu | Glu | Phe | Leu | Glu | Glu | Lys | Pro | Leu | Val | Gly | Glu |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|     |     |     | 820 |     |     |     |     |     | 825 |     |     |     |     |     | 830 |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Ala | Leu | Ala | Arg | Val | Lys | Arg | Ala | Glu | Lys | Lys | Trp | Arg | Asp | Lys | Arg |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|     |     |     | 835 |     |     |     |     |     | 840 |     |     |     |     |     | 845 |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Glu | Lys | Leu | Glu | Trp | Glu | Thr | Asn | Ile | Val | Tyr | Lys | Glu | Ala | Lys | Glu |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|     |     |     | 850 |     |     |     |     |     | 855 |     |     |     |     |     | 860 |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Ser | Val | Asp | Ala | Leu | Phe | Val | Asn | Ser | Gln | Tyr | Asp | Arg | Leu | Gln | Ala |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|     |     |     | 865 |     |     |     |     |     | 870 |     |     |     |     |     | 875 |     |  |  |  |  | 880 |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Asp | Thr | Asn | Ile | Ala | Met | Ile | His | Ala | Ala | Asp | Lys | Arg | Val | His | Ser |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|     |     |     | 885 |     |     |     |     |     | 890 |     |     |     |     |     | 895 |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Ile | Arg | Glu | Ala | Tyr | Leu | Pro | Glu | Leu | Ser | Val | Ile | Pro | Gly | Val | Asn |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|     |     |     | 900 |     |     |     |     |     | 905 |     |     |     |     |     | 910 |     |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |     |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

Ala Ala Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser  
915 920 925

Leu Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly  
930 935 940

Leu Ser Cys Trp Asn Val Lys Gly His Val Asp Val Glu Glu Gln Asn  
945 950 955 960

Asn His Arg Ser Val Leu Val Val Pro Glu Trp Glu Ala Glu Val Ser  
965 970 975

Gln Glu Val Arg Val Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr  
980 985 990

Ala Tyr Lys Glu Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile  
995 1000 1005

Glu Asn Asn Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu  
1010 1015 1020

Glu Val Tyr Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Ala  
1025 1030 1035

Thr Gln Glu Glu Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly  
1040 1045 1050

Tyr Asp Gly Ala Tyr Glu Ser Asn Ser Ser Val Pro Ala Asp Tyr  
1055 1060 1065

Ala Ser Ala Tyr Glu Glu Lys Ala Tyr Thr Asp Gly Arg Arg Asp  
1070 1075 1080

Asn Pro Cys Glu Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu  
1085 1090 1095

Pro Ala Gly Tyr Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr  
1100 1105 1110

Asp Lys Val Trp Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile  
1115 1120 1125

Val Asp Ser Val Glu Leu Leu Leu Met Glu Glu  
1130 1135

<210> 19

<211> 3420

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Trình tự nucleotit được dùng để biểu hiện ở tế bào vi khuẩn  
ghi mã TIC844\_5.

<400> 19

atggaaataa ataatcaaaa ccaatgtgtg cottacaatt gttaaagtaa tcctaaggag 60

ataatattag gcgaggaaag gctagaaaca gggaatactg tagcagacat ttcattaggg 120

cttattaatt ttctatattc taattttgta ccaggaggag gatttatagt aggtttacta 180

gaattaatat ggggatttat agggccttcg caatgggata ttttttttagc tcaaattgag 240

|                                                                    |      |
|--------------------------------------------------------------------|------|
| caattgatta gtcaaagaat agaagaatth gctaggaatc aggcaatthc aagattggag  | 300  |
| gggctaagca atctttataa ggtctatgth agagcgttht gcgactggga gaaagatcct  | 360  |
| actaatcctg ctttaaggga agaaatgcgt atacaattht atgacatgaa tagtgctctc  | 420  |
| ataacggcta ttccacttht tagagttcaa aattatgaag ttgctcttht atctgtatat  | 480  |
| gttcaagccg caaacttaca tttatctatt ttaagggatg tttcagtht cggagaaaga   | 540  |
| tggggatatg atacagcgac tatcaataat cgctatagtg atctgactag ccttattcat  | 600  |
| gtttatacta accattgtgt ggatacgtat aatcagggat taaggcgtht ggaaggctgt  | 660  |
| tttcttagcg attggattgt atataatcgt ttccggagac aattgacaat ttcagtatta  | 720  |
| gatattgttg cgtthtthtcc aaattatgat attagaacat atccaattca aacagctact | 780  |
| cagctaacga gggaagtcta tctggattth cctthtatta atgaaaatct ttctcctgca  | 840  |
| gcaaaaatc caaccttht agctgctgaa agtgctataa ttagaagtcc tcatttagta    | 900  |
| gactthttaa atagcttht catttataca gatagtctgg cacgttctgc atattgggga   | 960  |
| gggcacttgg taaattctth ccgcacagga accactacta atttgataag atcccctht   | 1020 |
| tatggaaggg aaggaaatac agagcgcccc gtaactatta ccgcatcacc tagcgtagca  | 1080 |
| atatttagaa cactthcata tccaacaggc cttgacaatt caaatcctgt agctggaatc  | 1140 |
| gagggagtgg aattccaaaa tactataagt agaagtatct atcgtaaaag cggccaata   | 1200 |
| gattcttht gtgaattacc acctcaagat gccagcgtat ctctgcaat tgggtatagt    | 1260 |
| caccgttht gccatgcaac atthttagaa cggattagtg gaccaagaat agcaggcacc   | 1320 |
| gtatthtctt ggacacaccg tagtgccagc cttactaacg aagtaagtcc atctagaatt  | 1380 |
| acacaaatc catgggtaaa gggcatact cttgcatctg gtgcctccgt cattaaaggt    | 1440 |
| cctggattth caggtggaga tattctgact aggaatagta tgggcgagct ggggacctta  | 1500 |
| cgagtaacct tcacaggaag attaccacaa agttattata tacgtthtccg ttatgcttcg | 1560 |
| gtagcaaata ggagtgttac atthtagatat tcacagccac cttcgtatgg aatttcattt | 1620 |
| ccaaaaacta tggacgcagg tgaaccacta acatctcgtt cgttcgctca tacaacactc  | 1680 |
| ttcactccaa taaccttht acgagctcaa gaagaattht atctatacat ccaatcgggt   | 1740 |
| gtttatatag atcgaattga atthtaccg gttactgcaa catttgaggc agaatatgat   | 1800 |
| ttagaaagag caaaaaggc ggtgaatgag ctgtthtactt cttccaatca aatcgggtta  | 1860 |
| aaaacagatg tgacggatta tcatattgat caagtatcca atthagttha gtgtttatct  | 1920 |
| gatgaattht gtctggatga aaaaaagaa ttgtccgaga aagtcaaaca tgcgaagcga   | 1980 |
| cttagtgatg agcggaatth acttcaagat ccaaaactth gagggatcaa tagacaacta  | 2040 |
| gaccgtggct ggagaggaag tacggatatt accatccaag gaggcgatga cgtattcaaa  | 2100 |
| gagaattacg ttacgctatt gggtaacctt gatgagtgt atccaacgta tttatatcaa   | 2160 |

```

aaaatagatg agtcgaaatt aaaagcctat acccgttacc aattaagagg gtatatcgaa      2220
gatagtcaag acttagaaat ctatttaatt cgctacaatg ccaaacacga aacagtaa      2280
gtgccaggta cgggttcctt atggccgctt tcagcccca gtccaatcgg aaaatgtgcc      2340
catcattccc atcatttctc cttggacatt gatgttggat gtacagactt aaatgaggac      2400
ttaggtgtat gggtgatatt caagattaag acgcaagatg gccatgcaag actaggaa      2460
ctagaatttc tcgaagagaa accattagta ggagaagcac tagctcgtgt gaaaagagcg      2520
gagaaaaaat ggagagacaa acgtgaaaaa ttggaatggg aaacaaatat tgtttataaa      2580
gaggcaaaag aatctgtaga tgctttatct gtaaactctc aatatgatag attacaagcg      2640
gataccaaca tcgcatgat tcatgctgga gataaacgct ttcatagcat tcgagaagct      2700
tatctgcctg agctgtctgt gattccgggt gtcaatgcgg ctatttttga agaattagaa      2760
gggcgtatct tcaactgcatt ctccctatat gatgagagaa atgtcattaa aaatgggtgat      2820
tttaataatg gcttatcctg ctggaacgtg aaagggcatg tagatgtaga agaacaaaac      2880
aaccaccgtt cggtccttgt tggtccggaa tgggaagcag aagtgtcaca agaagtctgt      2940
gtctgtccgg gtcgtggcta tatccttcgt gtcacagcgt acaaggaggg atatggagaa      3000
ggttgcgtaa ccattcatga gatcgagaac aatacagacg aactgaagtt tagcaactgt      3060
gtagaagagg aagtatatcc aaacaacacg gtaacgtgta atgattatac tgcgactcaa      3120
gaagaatatg aggggtacga cacttctcgt aatcgaggat atgacggagc ctatgaaagc      3180
aattcttctg taccagctga ttatgcatca gcctatgaag aaaaagcata tacagatgga      3240
cgaagagaca atccttgtga atctaacaga ggatatgggg attacacacc actaccagct      3300
ggctatgtga caaaagaatt agagtacttc ccagaaaccg ataaggatat gattgagatc      3360
ggagaaacgg aaggaacatt catcgtggac agcgtggaat tacttcttat ggaggaatag      3420

```

<210> 20

<211> 1139

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Trình tự axit amin của protein khám đã xử lý có hoạt tính trừ côn trùng TIC844\_5.

<400> 20

```

Met Glu Ile Asn Asn Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Ser
1              5              10              15

```

```

Asn Pro Lys Glu Ile Ile Leu Gly Glu Glu Arg Leu Glu Thr Gly Asn
              20              25              30

```

```

Thr Val Ala Asp Ile Ser Leu Gly Leu Ile Asn Phe Leu Tyr Ser Asn
              35              40              45

```

```

Phe Val Pro Gly Gly Gly Phe Ile Val Gly Leu Leu Glu Leu Ile Trp
              50              55              60

```

Gly Phe Ile Gly Pro Ser Gln Trp Asp Ile Phe Leu Ala Gln Ile Glu  
 65 70 75 80  
 Gln Leu Ile Ser Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala Ile  
 85 90 95  
 Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Lys Val Tyr Val Arg Ala  
 100 105 110  
 Phe Ser Asp Trp Glu Lys Asp Pro Thr Asn Pro Ala Leu Arg Glu Glu  
 115 120 125  
 Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Ile Thr Ala Ile  
 130 135 140  
 Pro Leu Phe Arg Val Gln Asn Tyr Glu Val Ala Leu Leu Ser Val Tyr  
 145 150 155 160  
 Val Gln Ala Ala Asn Leu His Leu Ser Ile Leu Arg Asp Val Ser Val  
 165 170 175  
 Phe Gly Glu Arg Trp Gly Tyr Asp Thr Ala Thr Ile Asn Asn Arg Tyr  
 180 185 190  
 Ser Asp Leu Thr Ser Leu Ile His Val Tyr Thr Asn His Cys Val Asp  
 195 200 205  
 Thr Tyr Asn Gln Gly Leu Arg Arg Leu Glu Gly Arg Phe Leu Ser Asp  
 210 215 220  
 Trp Ile Val Tyr Asn Arg Phe Arg Arg Gln Leu Thr Ile Ser Val Leu  
 225 230 235 240  
 Asp Ile Val Ala Phe Phe Pro Asn Tyr Asp Ile Arg Thr Tyr Pro Ile  
 245 250 255  
 Gln Thr Ala Thr Gln Leu Thr Arg Glu Val Tyr Leu Asp Leu Pro Phe  
 260 265 270  
 Ile Asn Glu Asn Leu Ser Pro Ala Ala Lys Tyr Pro Thr Phe Ser Ala  
 275 280 285  
 Ala Glu Ser Ala Ile Ile Arg Ser Pro His Leu Val Asp Phe Leu Asn  
 290 295 300  
 Ser Phe Thr Ile Tyr Thr Asp Ser Leu Ala Arg Ser Ala Tyr Trp Gly  
 305 310 315 320  
 Gly His Leu Val Asn Ser Phe Arg Thr Gly Thr Thr Thr Asn Leu Ile  
 325 330 335  
 Arg Ser Pro Leu Tyr Gly Arg Glu Gly Asn Thr Glu Arg Pro Val Thr  
 340 345 350  
 Ile Thr Ala Ser Pro Ser Val Pro Ile Phe Arg Thr Leu Ser Tyr Pro  
 355 360 365  
 Thr Gly Leu Asp Asn Ser Asn Pro Val Ala Gly Ile Glu Gly Val Glu  
 370 375 380  
 Phe Gln Asn Thr Ile Ser Arg Ser Ile Tyr Arg Lys Ser Gly Pro Ile  
 385 390 395 400

Asp Ser Phe Ser Glu Leu Pro Pro Gln Asp Ala Ser Val Ser Pro Ala  
 405 410 415  
 Ile Gly Tyr Ser His Arg Leu Cys His Ala Thr Phe Leu Glu Arg Ile  
 420 425 430  
 Ser Gly Pro Arg Ile Ala Gly Thr Val Phe Ser Trp Thr His Arg Ser  
 435 440 445  
 Ala Ser Pro Thr Asn Glu Val Ser Pro Ser Arg Ile Thr Gln Ile Pro  
 450 455 460  
 Trp Val Lys Ala His Thr Leu Ala Ser Gly Ala Ser Val Ile Lys Gly  
 465 470 475 480  
 Pro Gly Phe Thr Gly Gly Asp Ile Leu Thr Arg Asn Ser Met Gly Glu  
 485 490 495  
 Leu Gly Thr Leu Arg Val Thr Phe Thr Gly Arg Leu Pro Gln Ser Tyr  
 500 505 510  
 Tyr Ile Arg Phe Arg Tyr Ala Ser Val Ala Asn Arg Ser Gly Thr Phe  
 515 520 525  
 Arg Tyr Ser Gln Pro Pro Ser Tyr Gly Ile Ser Phe Pro Lys Thr Met  
 530 535 540  
 Asp Ala Gly Glu Pro Leu Thr Ser Arg Ser Phe Ala His Thr Thr Leu  
 545 550 555 560  
 Phe Thr Pro Ile Thr Phe Ser Arg Ala Gln Glu Glu Phe Asp Leu Tyr  
 565 570 575  
 Ile Gln Ser Gly Val Tyr Ile Asp Arg Ile Glu Phe Ile Pro Val Thr  
 580 585 590  
 Ala Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala Gln Lys Ala Val  
 595 600 605  
 Asn Glu Leu Phe Thr Ser Ser Asn Gln Ile Gly Leu Lys Thr Asp Val  
 610 615 620  
 Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu Ser  
 625 630 635 640  
 Asp Glu Phe Cys Leu Asp Glu Lys Lys Glu Leu Ser Glu Lys Val Lys  
 645 650 655  
 His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn  
 660 665 670  
 Phe Arg Gly Ile Asn Arg Gln Leu Asp Arg Gly Trp Arg Gly Ser Thr  
 675 680 685  
 Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val  
 690 695 700  
 Thr Leu Leu Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln  
 705 710 715 720  
 Lys Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg  
 725 730 735  
 Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr

|     |      |     |     |     |     |     |      |     |     |     |      |     |     |     |     |  |  |  |
|-----|------|-----|-----|-----|-----|-----|------|-----|-----|-----|------|-----|-----|-----|-----|--|--|--|
|     | 740  |     |     |     |     |     |      | 745 |     |     |      | 750 |     |     |     |  |  |  |
| Asn | Ala  | Lys | His | Glu | Thr | Val | Asn  | Val | Pro | Gly | Thr  | Gly | Ser | Leu | Trp |  |  |  |
|     | 755  |     |     |     |     |     | 760  |     |     |     | 765  |     |     |     |     |  |  |  |
| Pro | Leu  | Ser | Ala | Pro | Ser | Pro | Ile  | Gly | Lys | Cys | Ala  | His | His | Ser | His |  |  |  |
|     | 770  |     |     |     |     |     | 775  |     |     |     | 780  |     |     |     |     |  |  |  |
| His | Phe  | Ser | Leu | Asp | Ile | Asp | Val  | Gly | Cys | Thr | Asp  | Leu | Asn | Glu | Asp |  |  |  |
|     | 785  |     |     |     |     |     | 790  |     |     |     | 795  |     |     |     |     |  |  |  |
| Leu | Gly  | Val | Trp | Val | Ile | Phe | Lys  | Ile | Lys | Thr | Gln  | Asp | Gly | His | Ala |  |  |  |
|     | 805  |     |     |     |     |     | 810  |     |     |     | 815  |     |     |     |     |  |  |  |
| Arg | Leu  | Gly | Asn | Leu | Glu | Phe | Leu  | Glu | Glu | Lys | Pro  | Leu | Val | Gly | Glu |  |  |  |
|     | 820  |     |     |     |     |     | 825  |     |     |     | 830  |     |     |     |     |  |  |  |
| Ala | Leu  | Ala | Arg | Val | Lys | Arg | Ala  | Glu | Lys | Lys | Trp  | Arg | Asp | Lys | Arg |  |  |  |
|     | 835  |     |     |     |     |     | 840  |     |     |     | 845  |     |     |     |     |  |  |  |
| Glu | Lys  | Leu | Glu | Trp | Glu | Thr | Asn  | Ile | Val | Tyr | Lys  | Glu | Ala | Lys | Glu |  |  |  |
|     | 850  |     |     |     |     |     | 855  |     |     |     | 860  |     |     |     |     |  |  |  |
| Ser | Val  | Asp | Ala | Leu | Phe | Val | Asn  | Ser | Gln | Tyr | Asp  | Arg | Leu | Gln | Ala |  |  |  |
|     | 865  |     |     |     |     |     | 870  |     |     |     | 875  |     |     |     |     |  |  |  |
| Asp | Thr  | Asn | Ile | Ala | Met | Ile | His  | Ala | Ala | Asp | Lys  | Arg | Val | His | Ser |  |  |  |
|     | 885  |     |     |     |     |     | 890  |     |     |     | 895  |     |     |     |     |  |  |  |
| Ile | Arg  | Glu | Ala | Tyr | Leu | Pro | Glu  | Leu | Ser | Val | Ile  | Pro | Gly | Val | Asn |  |  |  |
|     | 900  |     |     |     |     |     | 905  |     |     |     | 910  |     |     |     |     |  |  |  |
| Ala | Ala  | Ile | Phe | Glu | Glu | Leu | Glu  | Gly | Arg | Ile | Phe  | Thr | Ala | Phe | Ser |  |  |  |
|     | 915  |     |     |     |     |     | 920  |     |     |     | 925  |     |     |     |     |  |  |  |
| Leu | Tyr  | Asp | Ala | Arg | Asn | Val | Ile  | Lys | Asn | Gly | Asp  | Phe | Asn | Asn | Gly |  |  |  |
|     | 930  |     |     |     |     |     | 935  |     |     |     | 940  |     |     |     |     |  |  |  |
| Leu | Ser  | Cys | Trp | Asn | Val | Lys | Gly  | His | Val | Asp | Val  | Glu | Glu | Gln | Asn |  |  |  |
|     | 945  |     |     |     |     |     | 950  |     |     |     | 955  |     |     |     |     |  |  |  |
| Asn | His  | Arg | Ser | Val | Leu | Val | Val  | Pro | Glu | Trp | Glu  | Ala | Glu | Val | Ser |  |  |  |
|     | 965  |     |     |     |     |     | 970  |     |     |     | 975  |     |     |     |     |  |  |  |
| Gln | Glu  | Val | Arg | Val | Cys | Pro | Gly  | Arg | Gly | Tyr | Ile  | Leu | Arg | Val | Thr |  |  |  |
|     | 980  |     |     |     |     |     | 985  |     |     |     | 990  |     |     |     |     |  |  |  |
| Ala | Tyr  | Lys | Glu | Gly | Tyr | Gly | Glu  | Gly | Cys | Val | Thr  | Ile | His | Glu | Ile |  |  |  |
|     | 995  |     |     |     |     |     | 1000 |     |     |     | 1005 |     |     |     |     |  |  |  |
| Glu | Asn  | Asn | Thr | Asp | Glu | Leu | Lys  | Phe | Ser | Asn | Cys  | Val | Glu | Glu |     |  |  |  |
|     | 1010 |     |     |     |     |     | 1015 |     |     |     | 1020 |     |     |     |     |  |  |  |
| Glu | Val  | Tyr | Pro | Asn | Asn | Thr | Val  | Thr | Cys | Asn | Asp  | Tyr | Thr | Ala |     |  |  |  |
|     | 1025 |     |     |     |     |     | 1030 |     |     |     | 1035 |     |     |     |     |  |  |  |
| Thr | Gln  | Glu | Glu | Tyr | Glu | Gly | Thr  | Tyr | Thr | Ser | Arg  | Asn | Arg | Gly |     |  |  |  |
|     | 1040 |     |     |     |     |     | 1045 |     |     |     | 1050 |     |     |     |     |  |  |  |
| Tyr | Asp  | Gly | Ala | Tyr | Glu | Ser | Asn  | Ser | Ser | Val | Pro  | Ala | Asp | Tyr |     |  |  |  |
|     | 1055 |     |     |     |     |     | 1060 |     |     |     | 1065 |     |     |     |     |  |  |  |
| Ala | Ser  | Ala | Tyr | Glu | Glu | Lys | Ala  | Tyr | Thr | Asp | Gly  | Arg | Arg | Asp |     |  |  |  |
|     | 1070 |     |     |     |     |     | 1075 |     |     |     | 1080 |     |     |     |     |  |  |  |

Asn Pro Cys Glu Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu  
1085 1090 1095

Pro Ala Gly Tyr Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr  
1100 1105 1110

Asp Lys Val Trp Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile  
1115 1120 1125

Val Asp Ser Val Glu Leu Leu Leu Met Glu Glu  
1130 1135

<210> 21

<211> 3420

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Trình tự nucleotit được dùng để biểu hiện ở tế bào vi khuẩn  
ghi mã TIC844\_6.

<400> 21

|                                                                    |      |
|--------------------------------------------------------------------|------|
| atggaaataa ataatcaaaa ccaatgtgtg ccttacaatt gtttaagtaa tcctaaggag  | 60   |
| ataatattag gcgaggaaag gctagaaaca gggaatactg tagcagacat ttcattaggg  | 120  |
| cttattaatt ttctatattc taattttgta ccaggaggag gatttatagt aggtttacta  | 180  |
| gaattaatat ggggatttat agggccttcg caatgggata tttttttagc tcaaattgag  | 240  |
| caattgatta gtcaaagaat agaagaatgt gctaggaatc aggcaatttc aagattggag  | 300  |
| gggctaagca atctttataa ggtctatgtt agagcgttta gcgactggga gaaagatcct  | 360  |
| actaatcctg ctttaaggga agaaatgctg atacaattta atgacatgaa tagtgctctc  | 420  |
| ataacggcta ttccactttt tagagttcaa aattatgaag ttgctctttt atctgtatat  | 480  |
| gttcaagccg caaacttaca tttatctatt ttaagggatg tttcagtttt cggagaaaga  | 540  |
| tggggatatg atacagcgac tatcaataat cgctatagtg atctgactag cottattcat  | 600  |
| gtttatacta accattgtgt ggatacgtat aatcagggat taaggcgttt ggaaggctgt  | 660  |
| tttcttagcg attggattgt atataatcgt ttccggagac aattgacaat ttcagtatta  | 720  |
| gatattgttg cgttttttcc aaattatgat attagaacat atccaattca aacagctact  | 780  |
| cagctaacga gggaagtcta tctggattta ctttttatta atgaaaatct ttctcctgca  | 840  |
| gcaagctatc caaccttttc agctgctgaa agtgctataa ttagaagtcc tcatttagta  | 900  |
| gactttttta atagctttac catttataca gatagtctgg caggttatgc atattgggga  | 960  |
| gggcacttgg taaattcttt ccgcacagga accactacta atttgataag atccccctta  | 1020 |
| tatggaaggg aaggaaatac agagcgcccc gtaactatta ccgcattcacc tagcgtacca | 1080 |
| atatttagaa cactttcata tattacaggc cttgacaatc gtaatcctgt agctggaatc  | 1140 |
| gagggagtgg aattccaaaa tactataagt agaagtatct atcgtaaaag cgggtccaata | 1200 |
| gattctttta gtgaattacc acctcaagat gccagcgtat ctctgcaat tgggtatagt   | 1260 |

|                                                                    |      |
|--------------------------------------------------------------------|------|
| caccgtttat gccatgcaac attttttagaa cggattagtg gaccaagaat agcaggcacc | 1320 |
| gtatttttctt ggacacaccg tagtgccagc cctactaacg aagtaagtcc atctagaatt | 1380 |
| acacaaaattc catgggtaaa ggcgcatact cttgcatctg gtgcctccgt cattaaaggt | 1440 |
| cctggatttta caggtggaga tattctgact aggaatagta tgggcgagct ggggacctta | 1500 |
| cgagtaacct tcacaggaag attaccacaa agttattata tacgtttccg ttatgcttcg  | 1560 |
| gtagcaaata ggagtggtag atttagatat tcacagccac cttcgtatgg aatttcattt  | 1620 |
| ccaaaaacta tggacgcagg tgaaccacta acatctcggt cgttcgctca tacaacactc  | 1680 |
| ttcactccaa taaccttttc acgagctcaa gaagaatttg atctatacat ccaatcgggt  | 1740 |
| gtttatatag atcgaattga atttataaccg gttactgcaa catttgaggc agaatatgat | 1800 |
| ttagaaagag cacaaaaggc ggtgaatgag ctgtttactt cttccaatca aatcgggtta  | 1860 |
| aaaacagatg tgacggatta tcatattgat caagtatcca atttagttga gtgtttatct  | 1920 |
| gatgaatttt gtctggatga aaaaaagaa ttgtccgaga aagtcaaaca tgcgaagcga   | 1980 |
| cttagtgatg agcggaattt acttcaagat ccaaacttta gagggatcaa tagacaacta  | 2040 |
| gaccgtggct ggagaggaag tacggatatt accatccaag gaggcgatga cgtattcaaa  | 2100 |
| gagaattacg ttacgctatt gggtagcttt gatgagtgt atccaacgta tttatatcaa   | 2160 |
| aaaatagatg agtcgaaatt aaaagcctat acccgttacc aattaagagg gtatatcgaa  | 2220 |
| gatagtcaag acttagaaat ctatttaatt cgctacaatg ccaaacacga aacagtaaat  | 2280 |
| gtgccaggta cgggttcctt atggccgctt tcagcccaa gtccaatcg aaatgtgcc     | 2340 |
| catcattccc atcatttctc cttggacatt gatgttggat gtacagactt aaatgaggac  | 2400 |
| ttagggtgat ggggtgatatt caagattaag acgcaagatg gccatgcaag actaggaaat | 2460 |
| ctagaatttc tcgaagagaa accattagta ggagaagcac tagctcgtgt gaaaagagcg  | 2520 |
| gagaaaaaat ggagagacaa acgtgaaaaa ttggaatggg aaacaaatat tgtttataaa  | 2580 |
| gaggcaaaag aatctgtaga tgctttatct gtaaaactctc aatatgatag attacaagcg | 2640 |
| gataccaaca tcgcatgat tcatgcggca gataaacgct ttcatagcat tcgagaagct   | 2700 |
| tatctgcctg agctgtctgt gattccgggt gtcaatgcgg ctatttttga agaattagaa  | 2760 |
| gggcgtatct tcaactgatt ctccctatat gatgcgagaa atgtcattaa aaatgggtgat | 2820 |
| tttaataaatg gcttatcctg ctggaacgtg aaagggcatg tagatgtaga agaacaaaac | 2880 |
| aaccaccgtt cggtccttgt tgttccggaa tgggaagcag aagtgtcaca agaagtccgt  | 2940 |
| gtctgtccgg gtcgtggcta tatccttcgt gtcacagcgt acaaggaggg atatggagaa  | 3000 |
| ggttgcgtaa ccattcatga gatcgagaac aatacagacg aactgaagtt tagcaactgt  | 3060 |
| gtagaagagg aagtatatcc aaacaacacg gtaacgtgta atgattatac tgcgactcaa  | 3120 |
| gaagaatatg agggtagcta cacttctcgt aatcgaggat atgacggagc ctatgaaagc  | 3180 |

```

aattcttctg taccagctga ttatgcatca gcctatgaag aaaaagcata tacagatgga      3240
cgaagagaca atccttgtga atctaacaga ggatatggggg attacacacc actaccagct      3300
ggctatgtga caaaagaatt agagtacttc ccagaaaccg ataaggatatg gattgagatc      3360
ggagaaacgg aaggaacatt catcgtggac agcgtggaat tacttcttat ggaggaatag      3420

```

<210> 22

<211> 1139

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Trình tự axit amin của protein khám đã xử lý có hoạt tính trừ côn trùng TIC844\_6.

<400> 22

```

Met Glu Ile Asn Asn Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Ser
1          5          10          15

```

```

Asn Pro Lys Glu Ile Ile Leu Gly Glu Glu Arg Leu Glu Thr Gly Asn
          20          25          30

```

```

Thr Val Ala Asp Ile Ser Leu Gly Leu Ile Asn Phe Leu Tyr Ser Asn
          35          40          45

```

```

Phe Val Pro Gly Gly Gly Phe Ile Val Gly Leu Leu Glu Leu Ile Trp
          50          55          60

```

```

Gly Phe Ile Gly Pro Ser Gln Trp Asp Ile Phe Leu Ala Gln Ile Glu
65          70          75          80

```

```

Gln Leu Ile Ser Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala Ile
          85          90          95

```

```

Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Lys Val Tyr Val Arg Ala
          100          105          110

```

```

Phe Ser Asp Trp Glu Lys Asp Pro Thr Asn Pro Ala Leu Arg Glu Glu
          115          120          125

```

```

Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Ile Thr Ala Ile
          130          135          140

```

```

Pro Leu Phe Arg Val Gln Asn Tyr Glu Val Ala Leu Leu Ser Val Tyr
145          150          155          160

```

```

Val Gln Ala Ala Asn Leu His Leu Ser Ile Leu Arg Asp Val Ser Val
          165          170          175

```

```

Phe Gly Glu Arg Trp Gly Tyr Asp Thr Ala Thr Ile Asn Asn Arg Tyr
          180          185          190

```

```

Ser Asp Leu Thr Ser Leu Ile His Val Tyr Thr Asn His Cys Val Asp
          195          200          205

```

```

Thr Tyr Asn Gln Gly Leu Arg Arg Leu Glu Gly Arg Phe Leu Ser Asp
          210          215          220

```

```

Trp Ile Val Tyr Asn Arg Phe Arg Arg Gln Leu Thr Ile Ser Val Leu
225          230          235          240

```

Asp Ile Val Ala Phe Phe Pro Asn Tyr Asp Ile Arg Thr Tyr Pro Ile  
 245 250 255  
 Gln Thr Ala Thr Gln Leu Thr Arg Glu Val Tyr Leu Asp Leu Pro Phe  
 260 265 270  
 Ile Asn Glu Asn Leu Ser Pro Ala Ala Ser Tyr Pro Thr Phe Ser Ala  
 275 280 285  
 Ala Glu Ser Ala Ile Ile Arg Ser Pro His Leu Val Asp Phe Leu Asn  
 290 295 300  
 Ser Phe Thr Ile Tyr Thr Asp Ser Leu Ala Arg Tyr Ala Tyr Trp Gly  
 305 310 315 320  
 Gly His Leu Val Asn Ser Phe Arg Thr Gly Thr Thr Thr Asn Leu Ile  
 325 330 335  
 Arg Ser Pro Leu Tyr Gly Arg Glu Gly Asn Thr Glu Arg Pro Val Thr  
 340 345 350  
 Ile Thr Ala Ser Pro Ser Val Pro Ile Phe Arg Thr Leu Ser Tyr Ile  
 355 360 365  
 Thr Gly Leu Asp Asn Arg Asn Pro Val Ala Gly Ile Glu Gly Val Glu  
 370 375 380  
 Phe Gln Asn Thr Ile Ser Arg Ser Ile Tyr Arg Lys Ser Gly Pro Ile  
 385 390 395 400  
 Asp Ser Phe Ser Glu Leu Pro Pro Gln Asp Ala Ser Val Ser Pro Ala  
 405 410 415  
 Ile Gly Tyr Ser His Arg Leu Cys His Ala Thr Phe Leu Glu Arg Ile  
 420 425 430  
 Ser Gly Pro Arg Ile Ala Gly Thr Val Phe Ser Trp Thr His Arg Ser  
 435 440 445  
 Ala Ser Pro Thr Asn Glu Val Ser Pro Ser Arg Ile Thr Gln Ile Pro  
 450 455 460  
 Trp Val Lys Ala His Thr Leu Ala Ser Gly Ala Ser Val Ile Lys Gly  
 465 470 475 480  
 Pro Gly Phe Thr Gly Gly Asp Ile Leu Thr Arg Asn Ser Met Gly Glu  
 485 490 495  
 Leu Gly Thr Leu Arg Val Thr Phe Thr Gly Arg Leu Pro Gln Ser Tyr  
 500 505 510  
 Tyr Ile Arg Phe Arg Tyr Ala Ser Val Ala Asn Arg Ser Gly Thr Phe  
 515 520 525  
 Arg Tyr Ser Gln Pro Pro Ser Tyr Gly Ile Ser Phe Pro Lys Thr Met  
 530 535 540  
 Asp Ala Gly Glu Pro Leu Thr Ser Arg Ser Phe Ala His Thr Thr Leu  
 545 550 555 560  
 Phe Thr Pro Ile Thr Phe Ser Arg Ala Gln Glu Glu Phe Asp Leu Tyr  
 565 570 575

```

Ile Gln Ser Gly Val Tyr Ile Asp Arg Ile Glu Phe Ile Pro Val Thr
      580                      585                      590

Ala Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala Gln Lys Ala Val
      595                      600                      605

Asn Glu Leu Phe Thr Ser Ser Asn Gln Ile Gly Leu Lys Thr Asp Val
      610                      615                      620

Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu Ser
      625                      630                      635                      640

Asp Glu Phe Cys Leu Asp Glu Lys Lys Glu Leu Ser Glu Lys Val Lys
      645                      650                      655

His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn
      660                      665                      670

Phe Arg Gly Ile Asn Arg Gln Leu Asp Arg Gly Trp Arg Gly Ser Thr
      675                      680                      685

Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val
      690                      695                      700

Thr Leu Leu Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln
      705                      710                      715                      720

Lys Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg
      725                      730                      735

Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr
      740                      745                      750

Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp
      755                      760                      765

Pro Leu Ser Ala Pro Ser Pro Ile Gly Lys Cys Ala His His Ser His
      770                      775                      780

His Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp
      785                      790                      795                      800

Leu Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala
      805                      810                      815

Arg Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val Gly Glu
      820                      825                      830

Ala Leu Ala Arg Val Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg
      835                      840                      845

Glu Lys Leu Glu Trp Glu Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu
      850                      855                      860

Ser Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Arg Leu Gln Ala
      865                      870                      875                      880

Asp Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg Val His Ser
      885                      890                      895

Ile Arg Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile Pro Gly Val Asn
      900                      905                      910

Ala Ala Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser

```

915                                      920                                      925  
 Leu Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly  
     930                                      935                                      940  
 Leu Ser Cys Trp Asn Val Lys Gly His Val Asp Val Glu Glu Gln Asn  
     945                                      950                                      955                                      960  
 Asn His Arg Ser Val Leu Val Val Pro Glu Trp Glu Ala Glu Val Ser  
                                     965                                      970                                      975  
 Gln Glu Val Arg Val Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr  
                                     980                                      985                                      990  
 Ala Tyr Lys Glu Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile  
                                     995                                      1000                                      1005  
 Glu Asn Asn Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu  
     1010                                      1015                                      1020  
 Glu Val Tyr Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Ala  
     1025                                      1030                                      1035  
 Thr Gln Glu Glu Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly  
     1040                                      1045                                      1050  
 Tyr Asp Gly Ala Tyr Glu Ser Asn Ser Ser Val Pro Ala Asp Tyr  
     1055                                      1060                                      1065  
 Ala Ser Ala Tyr Glu Glu Lys Ala Tyr Thr Asp Gly Arg Arg Asp  
     1070                                      1075                                      1080  
 Asn Pro Cys Glu Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu  
     1085                                      1090                                      1095  
 Pro Ala Gly Tyr Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr  
     1100                                      1105                                      1110  
 Asp Lys Val Trp Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile  
     1115                                      1120                                      1125  
 Val Asp Ser Val Glu Leu Leu Leu Met Glu Glu  
     1130                                      1135

<210> 23

<211> 3420

<212> ADN

<213> Trình tự nhân tạo

<220>

<223> Trình tự nucleotit được dùng để biểu hiện ở tế bào vi khuẩn  
ghi mã TIC844\_7.

<400> 23

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| atggaaataa ataatcaaaa ccaatgtgtg ccttacaatt gtttaagtaa tcctaaggag | 60  |
| ataatattag gcgaggaaag gctagaaaca gggaatactg tagcagacat ttcattaggg | 120 |
| cttattaatt ttctatatcc taattttgta ccaggaggag gatttatagt aggtttacta | 180 |
| gaattaatat ggggatttat agggccttcg caatgggata tttttttagc tcaaattgag | 240 |
| caattgatta gtcaaagaat agaagaattt gctaggaatc aggcaatttc aagattggag | 300 |

```

gggctaagca atctttataa ggtctatgtt agagcgttta gcgactggga gaaagatcct 360
actaatcctg ctttaaggga agaaatgcgt atacaattta atgacatgaa tagtgctctc 420
ataacggcta ttccactttt tagagttcaa aattatgaag ttgctctttt atctgtatat 480
gttcaagccg caaacttaca tttatctatt ttaagggatg tttcagtttt cggagaaaga 540
tggggatatg atacagcgac tatcaataat cgctatagtg atctgactag ccttattcat 600
gtttatacta accattgtgt ggatacgtat aatcagggat taaggcgttt ggaaggctgt 660
tttcttagcg attggattgt atataatcgt ttccggagac aattgacaat ttcagtatta 720
gatattgttg cgttttttcc aaattatgat attagaacat atccaattca aacagctact 780
cagctaacga gggaagtcta tctggattta ctttttatta atgaaaatct ttctcctgca 840
gcaagctatc caaccttttc agctgctgaa agtgctataa ttagaagtcc tcatttagta 900
gactttttta atagctttac catttataca gatagtctgg cacgttctgc atattgggga 960
gggcacttgg taaattcttt ccgcacagga accactacta atttgataag atccccttta 1020
tatggaaggg aaggaaatac agagcgcccc gtaactatta ccgcatcacc tagcgtagca 1080
atatttagaa cactttcata tagaacaggc cttgacaatt caaatcctgt agctggaatc 1140
gagggagtgg aattccaaaa tactataagt agaagtatct atcgtaaaag cggccaata 1200
gattctttta gtgaattacc acctcaagat gccagcgtat ctctgcaat tgggtatagt 1260
caccgtttat gccatgcaac attttttagaa cggattagtg gaccaagaat agcaggcacc 1320
gtattttctt ggacacaccg tagtgccagc cctactaacg aagtaagtcc atctagaatt 1380
acacaaattc catgggtaaa ggcgcatact cttgcatctg gtgcctccgt cattaaaggt 1440
cctggattta caggtggaga tattctgact aggaatagta tgggcgagct ggggacctta 1500
cgagtaacct tcacaggaag attaccacaa agttattata tacgtttccg ttatgcttcg 1560
gtagcaaata ggagtgttac atttagatat tcacagccac cttcgtagtg aatttcattt 1620
ccaaaaacta tggacgcagg tgaaccacta acatctcggt cgttcgctca tacaacactc 1680
ttcactccaa taaccttttc acgagctcaa gaagaatttg atctatacat ccaatcgggt 1740
gtttatatag atogaattga atttataccg gttactgcaa catttgaggc agaatatgat 1800
ttagaaagag cacaaaaggc ggtgaatgag ctgtttactt cttccaatca aatcgggtta 1860
aaaacagatg tgacggatta tcatattgat caagtatcca atttagttga gtgtttatct 1920
gatgaatttt gtctggatga aaaaaaagaa ttgtccgaga aagtcaaaca tgcgaagcga 1980
cttagtgatg agcggaattt acttcaagat ccaaacttta gagggatcaa tagacaacta 2040
gaccgtggct ggagaggaag tacggatatt accatccaag gaggcgatga cgtattcaaa 2100
gagaattacg ttacgctatt gggtagcttt gatgagtgtc atccaacgta tttatatcaa 2160
aaaatagatg agtcgaaatt aaaagcctat acccgttacc aattaagagg gtatatcgaa 2220

```

gatagtcaag acttagaaat ctattttaatt cgctacaatg ccaaacacga aacagtaaatt 2280  
gtgccaggta cgggttcctt atggccgctt tcagcccca gtccaatcgg aaaatgtgcc 2340  
catcattccc atcattttctc cttggacatt gatgttggat gtacagactt aaatgaggac 2400  
ttaggtgtat ggggtgatatt caagattaag acgcaagatg gccatgcaag actaggaaat 2460  
ctagaatttc tcgaagagaa accattagta ggagaagcac tagctcgtgt gaaaagagcg 2520  
gagaaaaaat ggagagacaa acgtgaaaaa ttggaatggg aaacaaatat tgtttataaa 2580  
gaggcaaaag aatctgtaga tgctttatit gtaactctc aatatgatag attacaagcg 2640  
gataccaaca tcgcgatgat tcatgcggca gataaacgcg ttcatagcat tcgagaagct 2700  
tatctgcctg agctgtctgt gattccgggt gtcaatgcgg ctatttttga agaattagaa 2760  
gggcgtatit tcaactgcatt ctccctatat gatgcgagaa atgtcattaa aaatgggtgat 2820  
tttaataatg gcttatcctg ctggaacgtg aaagggcatg tagatgtaga agaacaaaac 2880  
aaccaccgtt cggtccttgt tggtccggaa tgggaagcag aagtgtcaca agaagttcgt 2940  
gtctgtccgg gtcgtggcta tacccttctg gtcacagcgt acaaggaggg atatggagaa 3000  
ggttgcgtaa ccattcatga gatcgagaac aatacagacg aactgaagtt tagcaactgt 3060  
gtagaagagg aagtatatcc aaacaacacg gtaacgtgta atgattatac tgcgactcaa 3120  
gaagaatatg aggggtacgta cacttctcgt aatcgaggat atgacggagc ctatgaaagc 3180  
aattcttctg taccagctga ttatgcatca gcctatgaag aaaaagcata tacagatgga 3240  
cgaagagaca atccttgtga atctaacaga ggatatgggg attacacacc actaccagct 3300  
ggctatgtga caaaagaatt agagtacttc ccagaaaccg ataaggtatg gattgagatc 3360  
ggagaaacgg aaggaacatt catcgtggac agcgtggaat tacttcttat ggaggaatag 3420

<210> 24

<211> 1139

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Trình tự axit amin của protein khám đã xử lý có hoạt tính trừ côn trùng TIC844\_7.

<400> 24

Met Glu Ile Asn Asn Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Ser  
1 5 10 15

Asn Pro Lys Glu Ile Ile Leu Gly Glu Glu Arg Leu Glu Thr Gly Asn  
20 25 30

Thr Val Ala Asp Ile Ser Leu Gly Leu Ile Asn Phe Leu Tyr Ser Asn  
35 40 45

Phe Val Pro Gly Gly Gly Phe Ile Val Gly Leu Leu Glu Leu Ile Trp  
50 55 60

Gly Phe Ile Gly Pro Ser Gln Trp Asp Ile Phe Leu Ala Gln Ile Glu

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 65  |     |     |     |     | 70  |     |     |     |     |     | 75  |     |     |     | 80  |     |
| Gln | Leu | Ile | Ser |     | Gln | Arg | Ile | Glu | Glu | Phe | Ala | Arg | Asn | Gln | Ala | Ile |
|     |     |     |     |     | 85  |     |     |     |     | 90  |     |     |     | 95  |     |     |
| Ser | Arg | Leu | Glu | Gly | Leu | Ser | Asn | Leu | Tyr | Lys | Val | Tyr | Val | Arg | Ala |     |
|     |     |     | 100 |     |     |     |     | 105 |     |     |     |     | 110 |     |     |     |
| Phe | Ser | Asp | Trp | Glu | Lys | Asp | Pro | Thr | Asn | Pro | Ala | Leu | Arg | Glu | Glu |     |
|     |     | 115 |     |     |     |     | 120 |     |     |     |     | 125 |     |     |     |     |
| Met | Arg | Ile | Gln | Phe | Asn | Asp | Met | Asn | Ser | Ala | Leu | Ile | Thr | Ala | Ile |     |
|     | 130 |     |     |     |     | 135 |     |     |     |     | 140 |     |     |     |     |     |
| Pro | Leu | Phe | Arg | Val | Gln | Asn | Tyr | Glu | Val | Ala | Leu | Leu | Ser | Val | Tyr |     |
| 145 |     |     |     |     | 150 |     |     |     |     | 155 |     |     |     |     | 160 |     |
| Val | Gln | Ala | Ala | Asn | Leu | His | Leu | Ser | Ile | Leu | Arg | Asp | Val | Ser | Val |     |
|     |     |     |     | 165 |     |     |     |     | 170 |     |     |     |     | 175 |     |     |
| Phe | Gly | Glu | Arg | Trp | Gly | Tyr | Asp | Thr | Ala | Thr | Ile | Asn | Asn | Arg | Tyr |     |
|     |     |     | 180 |     |     |     |     | 185 |     |     |     |     | 190 |     |     |     |
| Ser | Asp | Leu | Thr | Ser | Leu | Ile | His | Val | Tyr | Thr | Asn | His | Cys | Val | Asp |     |
|     |     | 195 |     |     |     |     | 200 |     |     |     |     | 205 |     |     |     |     |
| Thr | Tyr | Asn | Gln | Gly | Leu | Arg | Arg | Leu | Glu | Gly | Arg | Phe | Leu | Ser | Asp |     |
|     | 210 |     |     |     |     | 215 |     |     |     |     | 220 |     |     |     |     |     |
| Trp | Ile | Val | Tyr | Asn | Arg | Phe | Arg | Arg | Gln | Leu | Thr | Ile | Ser | Val | Leu |     |
| 225 |     |     |     |     | 230 |     |     |     |     | 235 |     |     |     |     | 240 |     |
| Asp | Ile | Val | Ala | Phe | Phe | Pro | Asn | Tyr | Asp | Ile | Arg | Thr | Tyr | Pro | Ile |     |
|     |     |     |     | 245 |     |     |     |     | 250 |     |     |     |     | 255 |     |     |
| Gln | Thr | Ala | Thr | Gln | Leu | Thr | Arg | Glu | Val | Tyr | Leu | Asp | Leu | Pro | Phe |     |
|     |     |     | 260 |     |     |     |     | 265 |     |     |     |     | 270 |     |     |     |
| Ile | Asn | Glu | Asn | Leu | Ser | Pro | Ala | Ala | Ser | Tyr | Pro | Thr | Phe | Ser | Ala |     |
|     |     | 275 |     |     |     |     | 280 |     |     |     |     | 285 |     |     |     |     |
| Ala | Glu | Ser | Ala | Ile | Ile | Arg | Ser | Pro | His | Leu | Val | Asp | Phe | Leu | Asn |     |
|     | 290 |     |     |     |     | 295 |     |     |     |     | 300 |     |     |     |     |     |
| Ser | Phe | Thr | Ile | Tyr | Thr | Asp | Ser | Leu | Ala | Arg | Ser | Ala | Tyr | Trp | Gly |     |
| 305 |     |     |     |     | 310 |     |     |     |     | 315 |     |     |     |     | 320 |     |
| Gly | His | Leu | Val | Asn | Ser | Phe | Arg | Thr | Gly | Thr | Thr | Thr | Asn | Leu | Ile |     |
|     |     |     |     | 325 |     |     |     |     | 330 |     |     |     |     | 335 |     |     |
| Arg | Ser | Pro | Leu | Tyr | Gly | Arg | Glu | Gly | Asn | Thr | Glu | Arg | Pro | Val | Thr |     |
|     |     |     | 340 |     |     |     |     | 345 |     |     |     |     | 350 |     |     |     |
| Ile | Thr | Ala | Ser | Pro | Ser | Val | Pro | Ile | Phe | Arg | Thr | Leu | Ser | Tyr | Arg |     |
|     |     | 355 |     |     |     |     | 360 |     |     |     |     | 365 |     |     |     |     |
| Thr | Gly | Leu | Asp | Asn | Ser | Asn | Pro | Val | Ala | Gly | Ile | Glu | Gly | Val | Glu |     |
|     | 370 |     |     |     |     | 375 |     |     |     |     | 380 |     |     |     |     |     |
| Phe | Gln | Asn | Thr | Ile | Ser | Arg | Ser | Ile | Tyr | Arg | Lys | Ser | Gly | Pro | Ile |     |
| 385 |     |     |     |     | 390 |     |     |     |     | 395 |     |     |     |     | 400 |     |
| Asp | Ser | Phe | Ser | Glu | Leu | Pro | Pro | Gln | Asp | Ala | Ser | Val | Ser | Pro | Ala |     |
|     |     |     |     | 405 |     |     |     |     | 410 |     |     |     |     | 415 |     |     |

Ile Gly Tyr Ser His Arg Leu Cys His Ala Thr Phe Leu Glu Arg Ile  
 420 425 430  
 Ser Gly Pro Arg Ile Ala Gly Thr Val Phe Ser Trp Thr His Arg Ser  
 435 440 445  
 Ala Ser Pro Thr Asn Glu Val Ser Pro Ser Arg Ile Thr Gln Ile Pro  
 450 455 460  
 Trp Val Lys Ala His Thr Leu Ala Ser Gly Ala Ser Val Ile Lys Gly  
 465 470 475 480  
 Pro Gly Phe Thr Gly Gly Asp Ile Leu Thr Arg Asn Ser Met Gly Glu  
 485 490 495  
 Leu Gly Thr Leu Arg Val Thr Phe Thr Gly Arg Leu Pro Gln Ser Tyr  
 500 505 510  
 Tyr Ile Arg Phe Arg Tyr Ala Ser Val Ala Asn Arg Ser Gly Thr Phe  
 515 520 525  
 Arg Tyr Ser Gln Pro Pro Ser Tyr Gly Ile Ser Phe Pro Lys Thr Met  
 530 535 540  
 Asp Ala Gly Glu Pro Leu Thr Ser Arg Ser Phe Ala His Thr Thr Leu  
 545 550 555 560  
 Phe Thr Pro Ile Thr Phe Ser Arg Ala Gln Glu Glu Phe Asp Leu Tyr  
 565 570 575  
 Ile Gln Ser Gly Val Tyr Ile Asp Arg Ile Glu Phe Ile Pro Val Thr  
 580 585 590  
 Ala Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala Gln Lys Ala Val  
 595 600 605  
 Asn Glu Leu Phe Thr Ser Ser Asn Gln Ile Gly Leu Lys Thr Asp Val  
 610 615 620  
 Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu Ser  
 625 630 635 640  
 Asp Glu Phe Cys Leu Asp Glu Lys Lys Glu Leu Ser Glu Lys Val Lys  
 645 650 655  
 His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn  
 660 665 670  
 Phe Arg Gly Ile Asn Arg Gln Leu Asp Arg Gly Trp Arg Gly Ser Thr  
 675 680 685  
 Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val  
 690 695 700  
 Thr Leu Leu Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln  
 705 710 715 720  
 Lys Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg  
 725 730 735  
 Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr  
 740 745 750

Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp  
 755 760 765  
 Pro Leu Ser Ala Pro Ser Pro Ile Gly Lys Cys Ala His His Ser His  
 770 775 780  
 His Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp  
 785 790 795 800  
 Leu Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala  
 805 810 815  
 Arg Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val Gly Glu  
 820 825 830  
 Ala Leu Ala Arg Val Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg  
 835 840 845  
 Glu Lys Leu Glu Trp Glu Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu  
 850 855 860  
 Ser Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Arg Leu Gln Ala  
 865 870 875 880  
 Asp Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg Val His Ser  
 885 890 895  
 Ile Arg Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile Pro Gly Val Asn  
 900 905 910  
 Ala Ala Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser  
 915 920 925  
 Leu Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly  
 930 935 940  
 Leu Ser Cys Trp Asn Val Lys Gly His Val Asp Val Glu Glu Gln Asn  
 945 950 955 960  
 Asn His Arg Ser Val Leu Val Val Pro Glu Trp Glu Ala Glu Val Ser  
 965 970 975  
 Gln Glu Val Arg Val Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr  
 980 985 990  
 Ala Tyr Lys Glu Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile  
 995 1000 1005  
 Glu Asn Asn Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu  
 1010 1015 1020  
 Glu Val Tyr Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Ala  
 1025 1030 1035  
 Thr Gln Glu Glu Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly  
 1040 1045 1050  
 Tyr Asp Gly Ala Tyr Glu Ser Asn Ser Ser Val Pro Ala Asp Tyr  
 1055 1060 1065  
 Ala Ser Ala Tyr Glu Glu Lys Ala Tyr Thr Asp Gly Arg Arg Asp  
 1070 1075 1080  
 Asn Pro Cys Glu Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu

|                                                             |             |      |  |
|-------------------------------------------------------------|-------------|------|--|
| 1085                                                        | 1090        | 1095 |  |
| Pro Ala Gly Tyr Val Thr Lys Glu Leu Glu Tyr Phe             | Pro Glu Thr |      |  |
| 1100                                                        | 1105        | 1110 |  |
| Asp Lys Val Trp Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile |             |      |  |
| 1115                                                        | 1120        | 1125 |  |
| Val Asp Ser Val Glu Leu Leu Leu Met Glu Glu                 |             |      |  |
| 1130                                                        | 1135        |      |  |

<210> 25  
 <211> 3420  
 <212> ADN  
 <213> Trình tự nhân tạo

<220>  
 <223> Trình tự nucleotit được dùng để biểu hiện ở tế bào vi khuẩn ghi mã TIC844\_8.

<400> 25  
 atggaaataa ataatcaaaa ccaatgtgtg ccttacaatt gtttaagtaa tcctaaggag 60  
 ataatattag gcgaggaaag gctagaaaca gggaatactg tagcagacat ttcattaggg 120  
 cttattaatt ttctatattc taatttttga ccaggaggag gatttatagt aggtttacta 180  
 gaattaatat ggggatttat agggccttcg caatgggata ttttttttagc tcaaattgag 240  
 caattgatta gtcaaagaat agaagaattt gctaggaatc aggcaatttc aagattggag 300  
 gggctaagca atctttataa ggtctatgtt agagcgttta gcgactggga gaaagatcct 360  
 actaatcctg ctttaaggga agaaatgcgt atacaattta atgacatgaa tagtgctctc 420  
 ataacggcta ttccactttt tagagttcaa aattatgaag ttgctctttt atctgtatat 480  
 gttcaagccg caaacttaca tttatctatt ttaagggatg tttcagtttt cgagaaaaga 540  
 tggggatag atacagcgac tatcaataat cgctatagtg atctgactag ccttattcat 600  
 gtttatacta accattgtgt ggatacgtat aatcagggat taaggcgttt ggaaggctgt 660  
 tttcttagcg attggattgt atataatcgt ttccggagac aattgacaat ttcagtatta 720  
 gatattgttg cgttttttcc aaattatgat attagaacat atccaattca aacagctact 780  
 cagctaacga gggaagtcta tctggattta ctttttatta atgaaaatct ttctcctgca 840  
 gcagtatatc caaccttttc agctgctgaa agtgctataa ttagaagtcc tcatttagta 900  
 gactttttta atagctttac catttataca gatagtctgg cacgttctgc atattgggga 960  
 gggcacttg taaattcttt ccgcacagga accactacta atttgataag atccccctta 1020  
 tatggaaggg aaggaaatac agagcgcccc gtaactatta ccgcatacacc tagcgtacca 1080  
 atatttagaa cactttcata tccaacaggc cttgacaatt caaatcctgt agctggaatc 1140  
 gagggagtgg aattccaaaa tactataagt agaagtatct atcgtaaaag cggccaata 1200  
 gattctttta gtgaattacc acctcaagat gccagcgtat ctctgcaat tgggtatagt 1260  
 caccgtttat gccatgcaac attttttagaa cggatttagtg gaccaagaat agcaggcacc 1320

|                                                                    |      |
|--------------------------------------------------------------------|------|
| gtattttctt ggacacaccg tagtgccagc cctactaacg aagtaagtcc atctagaatt  | 1380 |
| acacaaaattc catgggtaaa ggcgcatact cttgcatctg gtgcctccgt cattaaaggt | 1440 |
| cctggattta caggtggaga tattctgact aggaatagta tgggcgagct ggggacctta  | 1500 |
| cgagtaacct tcacaggaag attaccacaa agttattata tacgtttccg ttatgcttcg  | 1560 |
| gtagcaaata ggagtgttac atttagatat tcacagccac cttcgtatgg aatttcattt  | 1620 |
| ccaaaaacta tggacgcagg tgaaccacta acatctcgtt cgttcgtca tacaacactc   | 1680 |
| ttcactccaa taaccttttc acgagctcaa gaagaatttg atctatacat ccaatcgggt  | 1740 |
| gtttatatag atcgaattga atttataccg gttactgcaa catttgaggc agaatatgat  | 1800 |
| ttagaaagag caaaaaaggc ggtgaatgag ctgtttactt cttccaatca aatcgggtta  | 1860 |
| aaaacagatg tgacggatta tcatattgat caagtatcca atttagttga gtgtttatct  | 1920 |
| gatgaatttt gtctggatga aaaaaagaa ttgtccgaga aagtcaaaca tgcgaagcga   | 1980 |
| cttagtgatg agcggaattt acttcaagat ccaaacttta gagggatcaa tagacaacta  | 2040 |
| gaccgtggct ggagaggaag tacggatatt accatccaag gaggcgatga cgtattcaaa  | 2100 |
| gagaattacg ttacgctatt gggtagcttt gatgagtgtc atccaacgta tttatatcaa  | 2160 |
| aaaatagatg agtcgaaatt aaaagcctat acccgttacc aattaagagg gtatatcgaa  | 2220 |
| gatagtcaag acttagaaat ctatttaatt cgctacaatg ccaaacacga aacagtaaat  | 2280 |
| gtgccaggta cgggttcctt atggccgctt tcagccccaa gtccaatcgg aaaatgtgcc  | 2340 |
| catcattccc atcatttctc cttggacatt gatgttgat gtacagactt aaatgaggac   | 2400 |
| ttaggtgtat gggtgatatt caagattaag acgcaagatg gccatgcaag actaggaaat  | 2460 |
| ctagaatttc tcgaagagaa accattagta ggagaagcac tagctcgtgt gaaagagcg   | 2520 |
| gagaaaaaat ggagagacaa acgtgaaaaa ttggaatggg aaacaaatat tgtttataaa  | 2580 |
| gaggcaaaag aatctgtaga tgctttattt gtaaaactctc aatatgatag attacaagcg | 2640 |
| gataccaaca tcgcgatgat tcatgcggca gataaacgcg ttcatagcat tcgagaagct  | 2700 |
| tatctgcctg agctgtctgt gattccgggt gtcaatgcgg ctatttttga agaattagaa  | 2760 |
| gggcgtatth tcactgcatt ctccctatat gatgcgagaa atgtcattaa aaatggtgat  | 2820 |
| tttaataatg gcttatcctg ctggaacgtg aaagggcatg tagatgtaga agaacaaaac  | 2880 |
| aaccaccgtt cggtccttgt tgttccggaa tgggaagcag aagtgtcaca agaagttcgt  | 2940 |
| gtctgtccgg gtcgtggcta tatccttcgt gtcacagcgt acaaggagg atatggagaa   | 3000 |
| ggttgcgtaa ccattcatga gatcgagaac aatacagacg aactgaagtt tagcaactgt  | 3060 |
| gtagaagagg aagtatatcc aaacaacacg gtaacgtgta atgattatac tgcgactcaa  | 3120 |
| gaagaatatg agggtagcta cacttctcgt aatcgaggat atgacggagc ctatgaaagc  | 3180 |
| aattcttctg taccagctga ttatgcatca gcctatgaag aaaaagcata tacagatgga  | 3240 |

cgaagagaca atccttgtga atctaacaga ggatatgggg attacacacc actaccagct 3300  
 ggctatgtga caaaagaatt agagtacttc ccagaaaccg ataaggatatg gattgagatc 3360  
 ggagaaacgg aaggaacatt catcgtggac agcgtggaat tacttcttat ggaggaatag 3420

<210> 26  
 <211> 1139  
 <212> PRT  
 <213> Trình tự nhân tạo

<220>  
 <223> Trình tự axit amin của protein khám đã xử lý có hoạt tính trừ côn  
 trùng TIC844\_8.

<400> 26

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Glu | Ile | Asn | Asn | Gln | Asn | Gln | Cys | Val | Pro | Tyr | Asn | Cys | Leu | Ser | 1   | 5   | 10  | 15  |
| Asn | Pro | Lys | Glu | Ile | Ile | Leu | Gly | Glu | Glu | Arg | Leu | Glu | Thr | Gly | Asn | 20  | 25  | 30  |     |
| Thr | Val | Ala | Asp | Ile | Ser | Leu | Gly | Leu | Ile | Asn | Phe | Leu | Tyr | Ser | Asn | 35  | 40  | 45  |     |
| Phe | Val | Pro | Gly | Gly | Gly | Phe | Ile | Val | Gly | Leu | Leu | Glu | Leu | Ile | Trp | 50  | 55  | 60  |     |
| Gly | Phe | Ile | Gly | Pro | Ser | Gln | Trp | Asp | Ile | Phe | Leu | Ala | Gln | Ile | Glu | 65  | 70  | 75  | 80  |
| Gln | Leu | Ile | Ser | Gln | Arg | Ile | Glu | Glu | Phe | Ala | Arg | Asn | Gln | Ala | Ile | 85  | 90  | 95  |     |
| Ser | Arg | Leu | Glu | Gly | Leu | Ser | Asn | Leu | Tyr | Lys | Val | Tyr | Val | Arg | Ala | 100 | 105 | 110 |     |
| Phe | Ser | Asp | Trp | Glu | Lys | Asp | Pro | Thr | Asn | Pro | Ala | Leu | Arg | Glu | Glu | 115 | 120 | 125 |     |
| Met | Arg | Ile | Gln | Phe | Asn | Asp | Met | Asn | Ser | Ala | Leu | Ile | Thr | Ala | Ile | 130 | 135 | 140 |     |
| Pro | Leu | Phe | Arg | Val | Gln | Asn | Tyr | Glu | Val | Ala | Leu | Leu | Ser | Val | Tyr | 145 | 150 | 155 | 160 |
| Val | Gln | Ala | Ala | Asn | Leu | His | Leu | Ser | Ile | Leu | Arg | Asp | Val | Ser | Val | 165 | 170 | 175 |     |
| Phe | Gly | Glu | Arg | Trp | Gly | Tyr | Asp | Thr | Ala | Thr | Ile | Asn | Asn | Arg | Tyr | 180 | 185 | 190 |     |
| Ser | Asp | Leu | Thr | Ser | Leu | Ile | His | Val | Tyr | Thr | Asn | His | Cys | Val | Asp | 195 | 200 | 205 |     |
| Thr | Tyr | Asn | Gln | Gly | Leu | Arg | Arg | Leu | Glu | Gly | Arg | Phe | Leu | Ser | Asp | 210 | 215 | 220 |     |
| Trp | Ile | Val | Tyr | Asn | Arg | Phe | Arg | Arg | Gln | Leu | Thr | Ile | Ser | Val | Leu | 225 | 230 | 235 | 240 |
| Asp | Ile | Val | Ala | Phe | Phe | Pro | Asn | Tyr | Asp | Ile | Arg | Thr | Tyr | Pro | Ile |     |     |     |     |

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| 245 |     |     |     |     |     |     |     | 250 |     |     |     |     | 255 |     |     |  |
| Gln | Thr | Ala | Thr | Gln | Leu | Thr | Arg | Glu | Val | Tyr | Leu | Asp | Leu | Pro | Phe |  |
|     |     |     | 260 |     |     |     |     | 265 |     |     |     | 270 |     |     |     |  |
| Ile | Asn | Glu | Asn | Leu | Ser | Pro | Ala | Ala | Val | Tyr | Pro | Thr | Phe | Ser | Ala |  |
|     |     | 275 |     |     |     |     |     | 280 |     |     |     | 285 |     |     |     |  |
| Ala | Glu | Ser | Ala | Ile | Ile | Arg | Ser | Pro | His | Leu | Val | Asp | Phe | Leu | Asn |  |
|     |     | 290 |     |     |     |     | 295 |     |     |     | 300 |     |     |     |     |  |
| Ser | Phe | Thr | Ile | Tyr | Thr | Asp | Ser | Leu | Ala | Arg | Ser | Ala | Tyr | Trp | Gly |  |
| 305 |     |     |     | 310 |     |     |     |     | 315 |     |     |     | 320 |     |     |  |
| Gly | His | Leu | Val | Asn | Ser | Phe | Arg | Thr | Gly | Thr | Thr | Thr | Asn | Leu | Ile |  |
|     |     |     | 325 |     |     |     |     | 330 |     |     |     | 335 |     |     |     |  |
| Arg | Ser | Pro | Leu | Tyr | Gly | Arg | Glu | Gly | Asn | Thr | Glu | Arg | Pro | Val | Thr |  |
|     |     |     | 340 |     |     |     |     | 345 |     |     |     | 350 |     |     |     |  |
| Ile | Thr | Ala | Ser | Pro | Ser | Val | Pro | Ile | Phe | Arg | Thr | Leu | Ser | Tyr | Pro |  |
|     |     | 355 |     |     |     |     |     | 360 |     |     |     | 365 |     |     |     |  |
| Thr | Gly | Leu | Asp | Asn | Ser | Asn | Pro | Val | Ala | Gly | Ile | Glu | Gly | Val | Glu |  |
|     |     | 370 |     |     |     |     | 375 |     |     |     | 380 |     |     |     |     |  |
| Phe | Gln | Asn | Thr | Ile | Ser | Arg | Ser | Ile | Tyr | Arg | Lys | Ser | Gly | Pro | Ile |  |
| 385 |     |     |     | 390 |     |     |     |     | 395 |     |     |     | 400 |     |     |  |
| Asp | Ser | Phe | Ser | Glu | Leu | Pro | Pro | Gln | Asp | Ala | Ser | Val | Ser | Pro | Ala |  |
|     |     |     | 405 |     |     |     |     | 410 |     |     |     | 415 |     |     |     |  |
| Ile | Gly | Tyr | Ser | His | Arg | Leu | Cys | His | Ala | Thr | Phe | Leu | Glu | Arg | Ile |  |
|     |     |     | 420 |     |     |     |     | 425 |     |     |     | 430 |     |     |     |  |
| Ser | Gly | Pro | Arg | Ile | Ala | Gly | Thr | Val | Phe | Ser | Trp | Thr | His | Arg | Ser |  |
|     |     | 435 |     |     |     |     | 440 |     |     |     | 445 |     |     |     |     |  |
| Ala | Ser | Pro | Thr | Asn | Glu | Val | Ser | Pro | Ser | Arg | Ile | Thr | Gln | Ile | Pro |  |
|     |     | 450 |     |     |     |     | 455 |     |     |     | 460 |     |     |     |     |  |
| Trp | Val | Lys | Ala | His | Thr | Leu | Ala | Ser | Gly | Ala | Ser | Val | Ile | Lys | Gly |  |
| 465 |     |     |     | 470 |     |     |     |     | 475 |     |     |     | 480 |     |     |  |
| Pro | Gly | Phe | Thr | Gly | Gly | Asp | Ile | Leu | Thr | Arg | Asn | Ser | Met | Gly | Glu |  |
|     |     |     | 485 |     |     |     |     | 490 |     |     |     | 495 |     |     |     |  |
| Leu | Gly | Thr | Leu | Arg | Val | Thr | Phe | Thr | Gly | Arg | Leu | Pro | Gln | Ser | Tyr |  |
|     |     |     | 500 |     |     |     |     | 505 |     |     |     | 510 |     |     |     |  |
| Tyr | Ile | Arg | Phe | Arg | Tyr | Ala | Ser | Val | Ala | Asn | Arg | Ser | Gly | Thr | Phe |  |
|     |     | 515 |     |     |     |     | 520 |     |     |     | 525 |     |     |     |     |  |
| Arg | Tyr | Ser | Gln | Pro | Pro | Ser | Tyr | Gly | Ile | Ser | Phe | Pro | Lys | Thr | Met |  |
|     |     | 530 |     |     |     |     | 535 |     |     |     | 540 |     |     |     |     |  |
| Asp | Ala | Gly | Glu | Pro | Leu | Thr | Ser | Arg | Ser | Phe | Ala | His | Thr | Thr | Leu |  |
| 545 |     |     |     | 550 |     |     |     |     | 555 |     |     |     | 560 |     |     |  |
| Phe | Thr | Pro | Ile | Thr | Phe | Ser | Arg | Ala | Gln | Glu | Glu | Phe | Asp | Leu | Tyr |  |
|     |     |     | 565 |     |     |     |     | 570 |     |     |     | 575 |     |     |     |  |
| Ile | Gln | Ser | Gly | Val | Tyr | Ile | Asp | Arg | Ile | Glu | Phe | Ile | Pro | Val | Thr |  |
|     |     |     | 580 |     |     |     |     | 585 |     |     |     | 590 |     |     |     |  |

Ala Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala Gln Lys Ala Val  
 595 600 605  
 Asn Glu Leu Phe Thr Ser Ser Asn Gln Ile Gly Leu Lys Thr Asp Val  
 610 615 620  
 Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu Ser  
 625 630 635 640  
 Asp Glu Phe Cys Leu Asp Glu Lys Lys Glu Leu Ser Glu Lys Val Lys  
 645 650 655  
 His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn  
 660 665 670  
 Phe Arg Gly Ile Asn Arg Gln Leu Asp Arg Gly Trp Arg Gly Ser Thr  
 675 680 685  
 Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val  
 690 695 700  
 Thr Leu Leu Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln  
 705 710 715 720  
 Lys Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg  
 725 730 735  
 Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr  
 740 745 750  
 Asn Ala Lys His Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu Trp  
 755 760 765  
 Pro Leu Ser Ala Pro Ser Pro Ile Gly Lys Cys Ala His His Ser His  
 770 775 780  
 His Phe Ser Leu Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp  
 785 790 795 800  
 Leu Gly Val Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala  
 805 810 815  
 Arg Leu Gly Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val Gly Glu  
 820 825 830  
 Ala Leu Ala Arg Val Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg  
 835 840 845  
 Glu Lys Leu Glu Trp Glu Thr Asn Ile Val Tyr Lys Glu Ala Lys Glu  
 850 855 860  
 Ser Val Asp Ala Leu Phe Val Asn Ser Gln Tyr Asp Arg Leu Gln Ala  
 865 870 875 880  
 Asp Thr Asn Ile Ala Met Ile His Ala Ala Asp Lys Arg Val His Ser  
 885 890 895  
 Ile Arg Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile Pro Gly Val Asn  
 900 905 910  
 Ala Ala Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe Thr Ala Phe Ser  
 915 920 925

Leu Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn Gly  
 930 935 940  
 Leu Ser Cys Trp Asn Val Lys Gly His Val Asp Val Glu Glu Gln Asn  
 945 950 955 960  
 Asn His Arg Ser Val Leu Val Val Pro Glu Trp Glu Ala Glu Val Ser  
 965 970 975  
 Gln Glu Val Arg Val Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr  
 980 985 990  
 Ala Tyr Lys Glu Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile  
 995 1000 1005  
 Glu Asn Asn Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu Glu  
 1010 1015 1020  
 Glu Val Tyr Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr Ala  
 1025 1030 1035  
 Thr Gln Glu Glu Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly  
 1040 1045 1050  
 Tyr Asp Gly Ala Tyr Glu Ser Asn Ser Ser Val Pro Ala Asp Tyr  
 1055 1060 1065  
 Ala Ser Ala Tyr Glu Glu Lys Ala Tyr Thr Asp Gly Arg Arg Asp  
 1070 1075 1080  
 Asn Pro Cys Glu Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu  
 1085 1090 1095  
 Pro Ala Gly Tyr Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr  
 1100 1105 1110  
 Asp Lys Val Trp Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile  
 1115 1120 1125  
 Val Asp Ser Val Glu Leu Leu Leu Met Glu Glu  
 1130 1135

<210> 27  
 <211> 3498  
 <212> ADN  
 <213> Trình tự nhân tạo

<220>  
 <223> Trình tự ADN tổng hợp được thiết kế để biểu hiện ở thực vật ghi mã  
 Cry1Da1.

<400> 27  
 atggagatca acaaccagaa ccagtgcgtc ccgtacaact gcctgagcaa ccctaaggag 60  
 atcatcctgg gtgaggaacg cctggagacc ggcaacaccg tagccgacat tagcctgggc 120  
 ctcatcaact tcctctacag caacttcgtg cccggcggtg gcttcatcgt gggcctcctg 180  
 gagcttatct ggggcttcat cggcccgtcc cagtgggaca tcttcctcgc ccagatcgag 240  
 caactgatca gccagcggat cgaggagttc gctaggaacc aggccatctc ccgcctggag 300  
 ggactctcca acctctacaa ggtgtacgtg cgcgcgttca gcgactggga gaaggaccgc 360

|                                                                    |      |
|--------------------------------------------------------------------|------|
| accaacccgg cctccgcga ggaaatgCGT atccagttca acgatatgaa ctCGGCCctc   | 420  |
| atcaccgcca tcccgctctt ccgCGtgCag aactacgagG tggccctcct gtccgtgtac  | 480  |
| gttcaagccg ccaacctcca cctctccatc ctccgcgacg tgagcgtgtt cggcgagcgc  | 540  |
| tggggctacg acaccgccac catcaacaac cGctactccg acctcacctc cctcatccac  | 600  |
| gtttacacca accactgCGt ggacacgtac aaccagggcc tccgCCgcct ggagggccgc  | 660  |
| ttcctctccg actggatCGt gtacaaccgc ttccgCCgcC agctcaccat ctccgtcctg  | 720  |
| gacatcgtcg ctttctttcc caactacgac atccgcacct accctatcca gaccgccacc  | 780  |
| cagctcaccC gcgaggtcta cctcgacctc ccgttcatca acgagaacct cagcccgGCC  | 840  |
| gccagctacc cGaccttctc cGccgctgag tccgccatca ttcgCagccc gcaacctCGtG | 900  |
| gacttcctca actccttcac catctacacc gactccctcg cccgctacgc ctactggggc  | 960  |
| ggtcacctcg tgaactcctt ccgcaccggc accactacca acctcatccg cagcccgctc  | 1020 |
| tacggccgcg agggcaacac cGagcgcCCg gtgaccatca ccgccagccc gagcgtgccc  | 1080 |
| atcttccgca ccctcagcta catcaccggc ctggacaaca gcaacctgtg ggcgggcatc  | 1140 |
| gagggcgtgg agttccagaa caccatctcc aggagcatct accgcaagag cggccctatc  | 1200 |
| gacagcttca gcgagctgcc tctcaggac gccagcgtga gccctgccat cggtacagc    | 1260 |
| cacaggctgt gccacgccac cttcctggag cGcatcagcg gccctcgcat cGcgggcacc  | 1320 |
| gtgttctcgt ggaccaccg cagcgcctct cctacgaacg aggtgtctcc tagtgcctc    | 1380 |
| accagatcc cttgggtcaa ggccacacc ctggctagtG gcgctagtgt catcaagggc    | 1440 |
| cctggcttca ccggtggtga catcctgacc aggaactcta tgggcgagct gggcactctg  | 1500 |
| agggctcactt tcactggccg cctgcctcag tottactaca tccgcttccg ctacgctagt | 1560 |
| gtcgctaacc gctctggtac ttccgctac tctcagcctc cgtcttacgG tatctctttc   | 1620 |
| cctaagacta tggacgctgg tgagcctctg accagtagga gcttcgctca cactactctg  | 1680 |
| ttcactccta tcactttctc tagggctcag gaggagtctg acctgtacat ccagtctggt  | 1740 |
| gtgtacatcg acaggatcga gttcatcccc gtgaccgcca cgttcgaggc cGagtacgac  | 1800 |
| cttgagcgcg ccagaaggt ggtgaacgcc ctcttacta gcaactaacca gctaggcctg   | 1860 |
| aagactgacg tgaccgacta ccacatcgac caagtgaGca acctagtggc ctgcctctcc  | 1920 |
| gacgagttct gcctcgacga gaagcgcgag ctgtccgaga aggtgaagca cgccaagcgc  | 1980 |
| ctctccgacg agcGcaacct gctccaggac cccaacttca ggggcatcaa caggcagccc  | 2040 |
| gaccgcggct ggcgcggtc caccgacatc accatccagg gcggtgacga cgtattcaag   | 2100 |
| gagaactacg ttaccctccc cggcaccttc gacgagtgtt accccaccta cctctaccag  | 2160 |
| aagatcgacg agtccaagct gaaggcctac acccgctacc agctccgcgG ctacatcgag  | 2220 |
| gactcccagG acctggaaat ctacctatc cGctacaacg ccaagcacga gatcgtgaac   | 2280 |

```

gtgcctggca ccggcagcct ctggcctctc agcgtggaga accagatcgg cccttgcggc 2340
gagcctaacc gctgcgcccc tcacctcgag tggaaccctg acctccactg ctcgtcgagg 2400
gacggcgaga agtgcgccca ccatagccac cacttctctc tggacatcga cgtgggctgc 2460
accgacctga acgaggacct gggcgtgtgg gttatcttca agatcaagac ccaggacggt 2520
cacgccaggc tgggtaacct ggagttcctt gaggaaaagc ctctgctggg tgaggccctg 2580
gccagggctca agagggctga gaagaaatgg agggataaga gggagaccct gcagctggag 2640
accactatcg tctacaagga ggctaaggag tctgtcgatg ctctgttcgt caactctcag 2700
tacgatagac tgcaagctga taccaacatc gctatgatcc acgctgcgga taagcgggtc 2760
caccggatcc gggaggctta ccttcgggag ctttctgtca tcccgggtgt caacgctgcg 2820
atcttcgagg aacttgagga acggatcttc actgcgttta gtctttacga tgcgcggaac 2880
atcatcaaga acggggactt caacaatggt ctgctgtgct ggaacgtcaa gggcatgtc 2940
gaggtcgagg aacaaaacaa tcatcgtagt gtccctgtca ttcttgagtg ggaggcggag 3000
gtctctcaag aggtccgtgt ttgccggggg cgtgggtaca ttcttcgtgt tactgcgtac 3060
aaggaggggt acggggaggg gtgcgttact attcatgaga ttgagaacaa tactgatgag 3120
cttaagttca acaattgtgt tgaggaggag gtttaccoga acaatactgt tacgtgcac 3180
aactacacgg caacgcaaga ggaatacgag gggacgtaca cctcgcgtaa tagaggggat 3240
gatgaggcgt acggaaacaa cccgtcgggt ccagcagatt atgcctcggg ttatgaggag 3300
aagtcgtaca cggatagacg acgcgagaat ccatgtgagt caaatcgagg atacggagat 3360
tacacaccat taccagcagg atacgttaca aaggagttgg aatacttccc ggaaacagat 3420
aaagtttga ttgaaatcgg agaaacagaa ggaacattca tcgtcgactc agtagaattg 3480
ttgttgatgg aagaatga 3498

```

<210> 28  
 <211> 1165  
 <212> PRT  
 <213> Trình tự nhân tạo

<220>  
 <223> Trình tự axit amin của CrylDa1 được ghi mã bởi trình tự ADN tổng hợp.

<400> 28

```

Met Glu Ile Asn Asn Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu Ser
1           5           10           15

```

```

Asn Pro Lys Glu Ile Ile Leu Gly Glu Glu Arg Leu Glu Thr Gly Asn
20           25           30

```

```

Thr Val Ala Asp Ile Ser Leu Gly Leu Ile Asn Phe Leu Tyr Ser Asn
35           40           45

```

```

Phe Val Pro Gly Gly Gly Phe Ile Val Gly Leu Leu Glu Leu Ile Trp

```

| 50      |     |         |         |         | 55     |         |         |         |         | 60      |     |     |         |     |         |
|---------|-----|---------|---------|---------|--------|---------|---------|---------|---------|---------|-----|-----|---------|-----|---------|
| Gly 65  | Phe | Ile     | Gly     | Pro     | Ser 70 | Gln     | Trp     | Asp     | Ile     | Phe 75  | Leu | Ala | Gln     | Ile | Glu 80  |
| Gln     | Leu | Ile     | Ser     | Gln 85  | Arg    | Ile     | Glu     | Glu     | Phe 90  | Ala     | Arg | Asn | Gln     | Ala | Ile 95  |
| Ser     | Arg | Leu     | Glu 100 | Gly     | Leu    | Ser     | Asn     | Leu 105 | Tyr     | Lys     | Val | Tyr | Val 110 | Arg | Ala     |
| Phe     | Ser | Asp 115 | Trp     | Glu     | Lys    | Asp     | Pro 120 | Thr     | Asn     | Pro     | Ala | Leu | Arg     | Glu | Glu 125 |
| Met 130 | Arg | Ile     | Gln     | Phe     | Asn    | Asp 135 | Met     | Asn     | Ser     | Ala     | Leu | Ile | Thr     | Ala | Ile 140 |
| Pro 145 | Leu | Phe     | Arg     | Val     | Gln    | Asn     | Tyr     | Glu     | Val     | Ala 155 | Leu | Leu | Ser     | Val | Tyr 160 |
| Val     | Gln | Ala     | Ala     | Asn 165 | Leu    | His     | Leu     | Ser     | Ile     | Leu     | Arg | Asp | Val     | Ser | Val 170 |
| Phe     | Gly | Glu     | Arg     | Trp 180 | Gly    | Tyr     | Asp     | Thr     | Ala     | Thr     | Ile | Asn | Asn     | Arg | Tyr 185 |
| Ser     | Asp | Leu 195 | Thr     | Ser     | Leu    | Ile     | His 200 | Val     | Tyr     | Thr     | Asn | His | Cys     | Val | Asp 205 |
| Thr 210 | Tyr | Asn     | Gln     | Gly     | Leu    | Arg     | Arg 215 | Leu     | Glu     | Gly     | Arg | Phe | Leu     | Ser | Asp 220 |
| Trp 225 | Ile | Val     | Tyr     | Asn     | Arg    | Phe     | Arg 230 | Arg     | Gln     | Leu     | Thr | Ile | Ser     | Val | Leu 235 |
| Asp     | Ile | Val     | Ala     | Phe 245 | Phe    | Pro     | Asn     | Tyr     | Asp 250 | Ile     | Arg | Thr | Tyr     | Pro | Ile 255 |
| Gln     | Thr | Ala     | Thr 260 | Gln     | Leu    | Thr     | Arg     | Glu     | Val     | Tyr     | Leu | Asp | Leu     | Pro | Phe 265 |
| Ile     | Asn | Glu     | Asn 275 | Leu     | Ser    | Pro     | Ala     | Ala     | Ser     | Tyr     | Pro | Thr | Phe     | Ser | Ala 280 |
| Ala     | Glu | Ser     | Ala 290 | Ile     | Ile    | Arg     | Ser     | Pro     | His     | Leu     | Val | Asp | Phe     | Leu | Asn 300 |
| Ser 305 | Phe | Thr     | Ile     | Tyr     | Thr    | Asp     | Ser     | Leu     | Ala     | Arg     | Tyr | Ala | Tyr     | Trp | Gly 310 |
| Gly     | His | Leu     | Val     | Asn 325 | Ser    | Phe     | Arg     | Thr     | Gly     | Thr     | Thr | Thr | Asn     | Leu | Ile 330 |
| Arg     | Ser | Pro     | Leu 340 | Tyr     | Gly    | Arg     | Glu     | Gly     | Asn     | Thr     | Glu | Arg | Pro     | Val | Thr 345 |
| Ile     | Thr | Ala     | Ser 355 | Pro     | Ser    | Val     | Pro     | Ile     | Phe     | Arg     | Thr | Leu | Ser     | Tyr | Ile 360 |
| Thr     | Gly | Leu     | Asp 370 | Asn     | Ser    | Asn     | Pro     | Val     | Ala     | Gly     | Ile | Glu | Gly     | Val | Glu 375 |
| Phe     | Gln | Asn     | Thr     | Ile     | Ser    | Arg     | Ser     | Ile     | Tyr     | Arg     | Lys | Ser | Gly     | Pro | Ile 380 |
| 385     |     |         |         |         | 390    |         |         |         |         | 395     |     |     |         |     | 400     |

Asp Ser Phe Ser Glu Leu Pro Pro Gln Asp Ala Ser Val Ser Pro Ala  
 405 410 415  
 Ile Gly Tyr Ser His Arg Leu Cys His Ala Thr Phe Leu Glu Arg Ile  
 420 425 430  
 Ser Gly Pro Arg Ile Ala Gly Thr Val Phe Ser Trp Thr His Arg Ser  
 435 440 445  
 Ala Ser Pro Thr Asn Glu Val Ser Pro Ser Arg Ile Thr Gln Ile Pro  
 450 455 460  
 Trp Val Lys Ala His Thr Leu Ala Ser Gly Ala Ser Val Ile Lys Gly  
 465 470 475 480  
 Pro Gly Phe Thr Gly Gly Asp Ile Leu Thr Arg Asn Ser Met Gly Glu  
 485 490 495  
 Leu Gly Thr Leu Arg Val Thr Phe Thr Gly Arg Leu Pro Gln Ser Tyr  
 500 505 510  
 Tyr Ile Arg Phe Arg Tyr Ala Ser Val Ala Asn Arg Ser Gly Thr Phe  
 515 520 525  
 Arg Tyr Ser Gln Pro Pro Ser Tyr Gly Ile Ser Phe Pro Lys Thr Met  
 530 535 540  
 Asp Ala Gly Glu Pro Leu Thr Ser Arg Ser Phe Ala His Thr Thr Leu  
 545 550 555 560  
 Phe Thr Pro Ile Thr Phe Ser Arg Ala Gln Glu Glu Phe Asp Leu Tyr  
 565 570 575  
 Ile Gln Ser Gly Val Tyr Ile Asp Arg Ile Glu Phe Ile Pro Val Thr  
 580 585 590  
 Ala Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala Gln Lys Val Val  
 595 600 605  
 Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Leu Lys Thr Asp Val  
 610 615 620  
 Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Ala Cys Leu Ser  
 625 630 635 640  
 Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val Lys  
 645 650 655  
 His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro Asn  
 660 665 670  
 Phe Arg Gly Ile Asn Arg Gln Pro Asp Arg Gly Trp Arg Gly Ser Thr  
 675 680 685  
 Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr Val  
 690 695 700  
 Thr Leu Pro Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr Gln  
 705 710 715 720  
 Lys Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu Arg  
 725 730 735

Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg Tyr  
 740 745 750  
 Asn Ala Lys His Glu Ile Val Asn Val Pro Gly Thr Gly Ser Leu Trp  
 755 760 765  
 Pro Leu Ser Val Glu Asn Gln Ile Gly Pro Cys Gly Glu Pro Asn Arg  
 770 775 780  
 Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu His Cys Ser Cys Arg  
 785 790 795 800  
 Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp Ile  
 805 810 815  
 Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val Ile  
 820 825 830  
 Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu Glu  
 835 840 845  
 Phe Leu Glu Glu Lys Pro Leu Leu Gly Glu Ala Leu Ala Arg Val Lys  
 850 855 860  
 Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Thr Leu Gln Leu Glu  
 865 870 875 880  
 Thr Thr Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu Phe  
 885 890 895  
 Val Asn Ser Gln Tyr Asp Arg Leu Gln Ala Asp Thr Asn Ile Ala Met  
 900 905 910  
 Ile His Ala Ala Asp Lys Arg Val His Arg Ile Arg Glu Ala Tyr Leu  
 915 920 925  
 Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu Glu  
 930 935 940  
 Leu Glu Glu Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg Asn  
 945 950 955 960  
 Ile Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Leu Cys Trp Asn Val  
 965 970 975  
 Lys Gly His Val Glu Val Glu Glu Gln Asn Asn His Arg Ser Val Leu  
 980 985 990  
 Val Ile Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val Cys  
 995 1000 1005  
 Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu Gly  
 1010 1015 1020  
 Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn Thr  
 1025 1030 1035  
 Asp Glu Leu Lys Phe Asn Asn Cys Val Glu Glu Glu Val Tyr Pro  
 1040 1045 1050  
 Asn Asn Thr Val Thr Cys Ile Asn Tyr Thr Ala Thr Gln Glu Glu  
 1055 1060 1065  
 Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly Tyr Asp Glu Ala

|                                                                                                                                       |                     |             |
|---------------------------------------------------------------------------------------------------------------------------------------|---------------------|-------------|
| 1070                                                                                                                                  | 1075                | 1080        |
| Tyr Gly Asn Asn Pro Ser Val                                                                                                           | Pro Ala Asp Tyr Ala | Ser Val Tyr |
| 1085                                                                                                                                  | 1090                | 1095        |
| Glu Glu Lys Ser Tyr Thr Asp                                                                                                           | Arg Arg Arg Glu Asn | Pro Cys Glu |
| 1100                                                                                                                                  | 1105                | 1110        |
| Ser Asn Arg Gly Tyr Gly Asp                                                                                                           | Tyr Thr Pro Leu Pro | Ala Gly Tyr |
| 1115                                                                                                                                  | 1120                | 1125        |
| Val Thr Lys Glu Leu Glu Tyr                                                                                                           | Phe Pro Glu Thr Asp | Lys Val Trp |
| 1130                                                                                                                                  | 1135                | 1140        |
| Ile Glu Ile Gly Glu Thr Glu                                                                                                           | Gly Thr Phe Ile Val | Asp Ser Val |
| 1145                                                                                                                                  | 1150                | 1155        |
| Glu Leu Leu Leu Met Glu Glu                                                                                                           |                     |             |
| 1160                                                                                                                                  | 1165                |             |
| <210> 29                                                                                                                              |                     |             |
| <211> 3501                                                                                                                            |                     |             |
| <212> ADN                                                                                                                             |                     |             |
| <213> Trình tự nhân tạo                                                                                                               |                     |             |
| <220>                                                                                                                                 |                     |             |
| <223> Trình tự ADN tổng hợp được thiết kế để biểu hiện ở thực vật ghi mã Cry1Da1 có gốc Alanin bổ sung được cài xen vào tại vị trí 2. |                     |             |
| <400> 29                                                                                                                              |                     |             |
| atggctgaga tcaacaacca gaaccagtgc gtcccgtaca actgcctgag caaccctaag                                                                     |                     | 60          |
| gagatcatcc tgggtgagga acgcctggag accggcaaca ccgtagccga cattagcctg                                                                     |                     | 120         |
| ggcctcatca acttcctcta cagcaacttc gtgcccggcg gtggcttcat cgtgggcctc                                                                     |                     | 180         |
| ctggagctta tctggggcct catcggcccg tccagtgagg acatcttcct cgcccagatc                                                                     |                     | 240         |
| gagcaactga tcagccagcg gatcgaggag ttcgctagga accaggccat ctccgcctg                                                                      |                     | 300         |
| gagggactct ccaacctcta caaggtgtac gtgcgcgcgt tcagcgactg ggagaaggac                                                                     |                     | 360         |
| ccgaccaacc cggccctccg cgaggaaatg cgtatccagt tcaacgatat gaactcggcc                                                                     |                     | 420         |
| ctcatcaccg ccatcccgtt cttccgcgtg cagaactacg aggtggccct cctgtccgtg                                                                     |                     | 480         |
| tacgttcaag cgcceaacct ccacctctcc atcctccgcg acgtgagcgt gttcggcgag                                                                     |                     | 540         |
| cgctggggct acgacaccgc caccatcaac aaccgctact ccgacctcac ctccctcatc                                                                     |                     | 600         |
| cacgtttaca ccaaccactg cgtggacacg tacaaccagg gcctccgccg cctggagggc                                                                     |                     | 660         |
| cgcttcctct ccgactggat cgtgtacaac cgcttcgcc gccagctcac catctccgtc                                                                      |                     | 720         |
| ctggacatcg tcgccttctt tcccaactac gacatccgca cctaccctat ccagaccgcc                                                                     |                     | 780         |
| accagctca cccgcgaggt ctacctcgac ctccggttca tcaacgagaa cctcagcccg                                                                      |                     | 840         |
| gccgccagct acccgacctt ctccgcgcgt gagtccgcca tcattcgag cccgcacctc                                                                      |                     | 900         |
| gtggacttcc tcaactcctt caccatctac accgactccc tcgcccgcta cgcctactgg                                                                     |                     | 960         |
| ggcggtcacc tcgtgaactc cttccgcacc ggcaccacta ccaacctcat ccgcagcccg                                                                     |                     | 1020        |

|                                                                     |      |
|---------------------------------------------------------------------|------|
| ctctacggcc gcgagggcaa caccgagcgc ccggtgacca tcaccgccag ccgagcgtg    | 1080 |
| cccatcttcc gcacctcag ctacatcacc ggcctggaca acagcaacc tgtggcgggc     | 1140 |
| atcgagggcg tggagttcca gaacaccatc tcaggagca tctaccgcaa gagcggccct    | 1200 |
| atcgacagct tcagcgagct gcctcctcag gacgccagcg tgagccctgc catcggtac    | 1260 |
| agccacaggc tgtgccacgc caccttcctg gagcgcatca gcggccctcg catcggggc    | 1320 |
| accgtgttct cgtggacca ccgcagcgc tctcctacga acgaggtgtc tcctagtcgc     | 1380 |
| atcaccaga tcccttggt caaggccac accctggcta gtggcgctag tgtcatcaag      | 1440 |
| ggccctggct tcaccggtg tgacatcctg accaggaact ctatgggca gctgggact      | 1500 |
| ctgagggta ctttactgg ccgcctgcct cagtcttact acatccgctt ccgctacgt      | 1560 |
| agtgtcgcta accgctctg tactttccgc tactctcagc ctccgtctta cggtatctct    | 1620 |
| ttccctaaga ctatggacgc tggtagacct ctgaccagta ggagcttcgc tcacactact   | 1680 |
| ctgttcactc ctatcacttt ctctagggct caggaggagt tcgacctgta catccagtct   | 1740 |
| ggtgtgtaca tcgacaggat cgagttcatc ccggtgaccg ccacgttcga ggccgagtac   | 1800 |
| gaccttgagc gcgccagaa ggtggtgaac gccctcttca ctagcactaa ccagctaggc    | 1860 |
| ctgaagactg acgtgaccga ctaccacatc gaccaagtga gcaacctagt ggctgcctc    | 1920 |
| tccgacgagt tctgcctcga cgagaagcgc gagctgtccg agaagggtgaa gcacgccaa   | 1980 |
| cgctctccg acgagcgcaa cctgctccag gaccccaact tcaggggcat caacaggcag    | 2040 |
| cccgaccgcg gctggcgcg ctccaccgac atcaccatcc agggcggtga cgacgtattc    | 2100 |
| aaggagaact acgttaccct ccccggcacc ttcgacgagt gttacccac ctacctctac    | 2160 |
| cagaagatcg acgagtccaa gctgaaggcc tacaccgct accagctccg cggctacatc    | 2220 |
| gaggactccc aggacctgga aatctacctc atccgctaca acgccaaagca cgagatcgctg | 2280 |
| aacgtgcctg gcaccggcag cctctggcct ctacagctgg agaaccagat cggcccttgc   | 2340 |
| ggcgagccta accgctgcgc ccctcacctc gagtggaacc ctgacctca ctgctcgtgc    | 2400 |
| agggacggcg agaagtgcgc ccaccatagc caccattct ctctggacat cgacgtgggc    | 2460 |
| tgcaccgacc tgaacgagga cctgggcgtg tgggttatct tcaagatcaa gaccaggac    | 2520 |
| ggtcacgcca ggctgggtaa cctggagttc cttgaggaaa agcctctgct gggtgaggcc   | 2580 |
| ctggccaggg tcaagagggc tgagaagaaa tggagggata agaggagac cctgcagctg    | 2640 |
| gagaccacta tcgtctacaa ggaggctaag gagtctgtcg atgctctgtt cgtcaactct   | 2700 |
| cagtacgata gactgcaagc tgataccaac atcgctatga tccacgctgc ggataagcgg   | 2760 |
| gtccaccgga tccgggaggc ttaccttcg gagctttctg tcatccggg tgtcaacgct     | 2820 |
| gcgatcttcg aggaacttga ggaacggatc ttoactgcgt ttagtcttta cgatgocgg    | 2880 |
| aacatcatca agaacgggga cttcaacaat ggtctgctgt gctggaacgt caagggtcat   | 2940 |

```

gtcgagggtcg aggaacaaaa caatcatcgt agtgtccttg tcattcctga gtgggaggcg      3000
gaggtctctc aagaggtccg tgtttgcccg gggcgtgggt acattcttcg tgttactgcg      3060
tacaaggagg ggtacgggga ggggtgcgtt actattcatg agattgagaa caatactgat      3120
gagcttaagt tcaacaattg tgttgaggag gaggtttacc cgaacaatac tgttacgtgc      3180
atcaactaca cggcaacgca agaggaatac gaggggacgt acacctcgcg taatagaggg      3240
tatgatgagg cgtacggaaa caaccgctcg gttccagcag attatgcctc ggtttatgag      3300
gagaagtcgt acacggatag acgacgcgag aatccatgtg agtcaaatcg aggatacgga      3360
gattacacac cattaccagc aggatacgtt acaaaggagt tggaatactt cccggaaaca      3420
gataaagttt ggattgaaat cggagaaaca gaaggaacat tcatcgtcga ctcagtagaa      3480
tgtttgttga tggaagaatg a                                     3501

```

<210> 30

<211> 1166

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Trình tự axit amin của CrylDal được ghi mã bởi trình tự ADN tổng hợp trong đó gốc Alanin bổ sung được cài xen vào tại vị trí 2.

<400> 30

```

Met Ala Glu Ile Asn Asn Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu
1           5           10           15

```

```

Ser Asn Pro Lys Glu Ile Ile Leu Gly Glu Glu Arg Leu Glu Thr Gly
          20           25           30

```

```

Asn Thr Val Ala Asp Ile Ser Leu Gly Leu Ile Asn Phe Leu Tyr Ser
          35           40           45

```

```

Asn Phe Val Pro Gly Gly Gly Phe Ile Val Gly Leu Leu Glu Leu Ile
          50           55           60

```

```

Trp Gly Phe Ile Gly Pro Ser Gln Trp Asp Ile Phe Leu Ala Gln Ile
          65           70           75           80

```

```

Glu Gln Leu Ile Ser Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala
          85           90           95

```

```

Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Lys Val Tyr Val Arg
          100          105          110

```

```

Ala Phe Ser Asp Trp Glu Lys Asp Pro Thr Asn Pro Ala Leu Arg Glu
          115          120          125

```

```

Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Ile Thr Ala
          130          135          140

```

```

Ile Pro Leu Phe Arg Val Gln Asn Tyr Glu Val Ala Leu Leu Ser Val
          145          150          155          160

```

```

Tyr Val Gln Ala Ala Asn Leu His Leu Ser Ile Leu Arg Asp Val Ser
          165          170          175

```

Val Phe Gly Glu Arg Trp Gly Tyr Asp Thr Ala Thr Ile Asn Asn Arg  
 180 185 190  
 Tyr Ser Asp Leu Thr Ser Leu Ile His Val Tyr Thr Asn His Cys Val  
 195 200 205  
 Asp Thr Tyr Asn Gln Gly Leu Arg Arg Leu Glu Gly Arg Phe Leu Ser  
 210 215 220  
 Asp Trp Ile Val Tyr Asn Arg Phe Arg Arg Gln Leu Thr Ile Ser Val  
 225 230 235 240  
 Leu Asp Ile Val Ala Phe Phe Pro Asn Tyr Asp Ile Arg Thr Tyr Pro  
 245 250 255  
 Ile Gln Thr Ala Thr Gln Leu Thr Arg Glu Val Tyr Leu Asp Leu Pro  
 260 265 270  
 Phe Ile Asn Glu Asn Leu Ser Pro Ala Ala Ser Tyr Pro Thr Phe Ser  
 275 280 285  
 Ala Ala Glu Ser Ala Ile Ile Arg Ser Pro His Leu Val Asp Phe Leu  
 290 295 300  
 Asn Ser Phe Thr Ile Tyr Thr Asp Ser Leu Ala Arg Tyr Ala Tyr Trp  
 305 310 315 320  
 Gly Gly His Leu Val Asn Ser Phe Arg Thr Gly Thr Thr Thr Asn Leu  
 325 330 335  
 Ile Arg Ser Pro Leu Tyr Gly Arg Glu Gly Asn Thr Glu Arg Pro Val  
 340 345 350  
 Thr Ile Thr Ala Ser Pro Ser Val Pro Ile Phe Arg Thr Leu Ser Tyr  
 355 360 365  
 Ile Thr Gly Leu Asp Asn Ser Asn Pro Val Ala Gly Ile Glu Gly Val  
 370 375 380  
 Glu Phe Gln Asn Thr Ile Ser Arg Ser Ile Tyr Arg Lys Ser Gly Pro  
 385 390 395 400  
 Ile Asp Ser Phe Ser Glu Leu Pro Pro Gln Asp Ala Ser Val Ser Pro  
 405 410 415  
 Ala Ile Gly Tyr Ser His Arg Leu Cys His Ala Thr Phe Leu Glu Arg  
 420 425 430  
 Ile Ser Gly Pro Arg Ile Ala Gly Thr Val Phe Ser Trp Thr His Arg  
 435 440 445  
 Ser Ala Ser Pro Thr Asn Glu Val Ser Pro Ser Arg Ile Thr Gln Ile  
 450 455 460  
 Pro Trp Val Lys Ala His Thr Leu Ala Ser Gly Ala Ser Val Ile Lys  
 465 470 475 480  
 Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu Thr Arg Asn Ser Met Gly  
 485 490 495  
 Glu Leu Gly Thr Leu Arg Val Thr Phe Thr Gly Arg Leu Pro Gln Ser  
 500 505 510  
 Tyr Tyr Ile Arg Phe Arg Tyr Ala Ser Val Ala Asn Arg Ser Gly Thr

| 515 |     |     |     |     | 520 |     |     |     |     | 525 |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Phe | Arg | Tyr | Ser | Gln | Pro | Pro | Ser | Tyr | Gly | Ile | Ser | Phe | Pro | Lys | Thr |
| 530 |     |     |     |     |     | 535 |     |     |     |     | 540 |     |     |     |     |
| Met | Asp | Ala | Gly | Glu | Pro | Leu | Thr | Ser | Arg | Ser | Phe | Ala | His | Thr | Thr |
| 545 |     |     |     |     | 550 |     |     |     |     | 555 |     |     |     |     | 560 |
| Leu | Phe | Thr | Pro | Ile | Thr | Phe | Ser | Arg | Ala | Gln | Glu | Glu | Phe | Asp | Leu |
|     |     |     |     | 565 |     |     |     |     | 570 |     |     |     |     | 575 |     |
| Tyr | Ile | Gln | Ser | Gly | Val | Tyr | Ile | Asp | Arg | Ile | Glu | Phe | Ile | Pro | Val |
|     |     |     | 580 |     |     |     |     | 585 |     |     |     |     | 590 |     |     |
| Thr | Ala | Thr | Phe | Glu | Ala | Glu | Tyr | Asp | Leu | Glu | Arg | Ala | Gln | Lys | Val |
|     |     | 595 |     |     |     |     | 600 |     |     |     |     | 605 |     |     |     |
| Val | Asn | Ala | Leu | Phe | Thr | Ser | Thr | Asn | Gln | Leu | Gly | Leu | Lys | Thr | Asp |
| 610 |     |     |     |     |     | 615 |     |     |     |     | 620 |     |     |     |     |
| Val | Thr | Asp | Tyr | His | Ile | Asp | Gln | Val | Ser | Asn | Leu | Val | Ala | Cys | Leu |
| 625 |     |     |     |     | 630 |     |     |     |     | 635 |     |     |     |     | 640 |
| Ser | Asp | Glu | Phe | Cys | Leu | Asp | Glu | Lys | Arg | Glu | Leu | Ser | Glu | Lys | Val |
|     |     |     |     | 645 |     |     |     |     | 650 |     |     |     |     | 655 |     |
| Lys | His | Ala | Lys | Arg | Leu | Ser | Asp | Glu | Arg | Asn | Leu | Leu | Gln | Asp | Pro |
|     |     |     | 660 |     |     |     |     | 665 |     |     |     |     | 670 |     |     |
| Asn | Phe | Arg | Gly | Ile | Asn | Arg | Gln | Pro | Asp | Arg | Gly | Trp | Arg | Gly | Ser |
|     |     | 675 |     |     |     |     | 680 |     |     |     |     | 685 |     |     |     |
| Thr | Asp | Ile | Thr | Ile | Gln | Gly | Gly | Asp | Asp | Val | Phe | Lys | Glu | Asn | Tyr |
|     | 690 |     |     |     |     | 695 |     |     |     |     | 700 |     |     |     |     |
| Val | Thr | Leu | Pro | Gly | Thr | Phe | Asp | Glu | Cys | Tyr | Pro | Thr | Tyr | Leu | Tyr |
| 705 |     |     |     |     | 710 |     |     |     |     | 715 |     |     |     |     | 720 |
| Gln | Lys | Ile | Asp | Glu | Ser | Lys | Leu | Lys | Ala | Tyr | Thr | Arg | Tyr | Gln | Leu |
|     |     |     |     | 725 |     |     |     |     | 730 |     |     |     |     | 735 |     |
| Arg | Gly | Tyr | Ile | Glu | Asp | Ser | Gln | Asp | Leu | Glu | Ile | Tyr | Leu | Ile | Arg |
|     |     |     | 740 |     |     |     |     | 745 |     |     |     |     | 750 |     |     |
| Tyr | Asn | Ala | Lys | His | Glu | Ile | Val | Asn | Val | Pro | Gly | Thr | Gly | Ser | Leu |
|     |     | 755 |     |     |     |     | 760 |     |     |     |     | 765 |     |     |     |
| Trp | Pro | Leu | Ser | Val | Glu | Asn | Gln | Ile | Gly | Pro | Cys | Gly | Glu | Pro | Asn |
|     | 770 |     |     |     |     | 775 |     |     |     |     | 780 |     |     |     |     |
| Arg | Cys | Ala | Pro | His | Leu | Glu | Trp | Asn | Pro | Asp | Leu | His | Cys | Ser | Cys |
| 785 |     |     |     |     | 790 |     |     |     |     | 795 |     |     |     |     | 800 |
| Arg | Asp | Gly | Glu | Lys | Cys | Ala | His | His | Ser | His | His | Phe | Ser | Leu | Asp |
|     |     |     |     | 805 |     |     |     |     | 810 |     |     |     |     | 815 |     |
| Ile | Asp | Val | Gly | Cys | Thr | Asp | Leu | Asn | Glu | Asp | Leu | Gly | Val | Trp | Val |
|     |     |     | 820 |     |     |     |     | 825 |     |     |     |     | 830 |     |     |
| Ile | Phe | Lys | Ile | Lys | Thr | Gln | Asp | Gly | His | Ala | Arg | Leu | Gly | Asn | Leu |
|     |     | 835 |     |     |     |     | 840 |     |     |     |     | 845 |     |     |     |
| Glu | Phe | Leu | Glu | Glu | Lys | Pro | Leu | Leu | Gly | Glu | Ala | Leu | Ala | Arg | Val |
|     | 850 |     |     |     |     | 855 |     |     |     |     | 860 |     |     |     |     |

Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Thr Leu Gln Leu  
 865 870 875 880  
 Glu Thr Thr Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu  
 885 890 895  
 Phe Val Asn Ser Gln Tyr Asp Arg Leu Gln Ala Asp Thr Asn Ile Ala  
 900 905 910  
 Met Ile His Ala Ala Asp Lys Arg Val His Arg Ile Arg Glu Ala Tyr  
 915 920 925  
 Leu Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu  
 930 935 940  
 Glu Leu Glu Glu Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg  
 945 950 955 960  
 Asn Ile Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Leu Cys Trp Asn  
 965 970 975  
 Val Lys Gly His Val Glu Val Glu Glu Gln Asn Asn His Arg Ser Val  
 980 985 990  
 Leu Val Ile Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val  
 995 1000 1005  
 Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu  
 1010 1015 1020  
 Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn  
 1025 1030 1035  
 Thr Asp Glu Leu Lys Phe Asn Asn Cys Val Glu Glu Glu Val Tyr  
 1040 1045 1050  
 Pro Asn Asn Thr Val Thr Cys Ile Asn Tyr Thr Ala Thr Gln Glu  
 1055 1060 1065  
 Glu Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly Tyr Asp Glu  
 1070 1075 1080  
 Ala Tyr Gly Asn Asn Pro Ser Val Pro Ala Asp Tyr Ala Ser Val  
 1085 1090 1095  
 Tyr Glu Glu Lys Ser Tyr Thr Asp Arg Arg Arg Glu Asn Pro Cys  
 1100 1105 1110  
 Glu Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu Pro Ala Gly  
 1115 1120 1125  
 Tyr Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val  
 1130 1135 1140  
 Trp Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser  
 1145 1150 1155  
 Val Glu Leu Leu Leu Met Glu Glu  
 1160 1165  
 <210> 31  
 <211> 3501  
 <212> ADN

<213> Trình tự nhân tạo

<220>

<223> Trình tự ADN tổng hợp được thiết kế để biểu hiện ở thực vật ghi mã Cry1Da1\_3 có gốc Alanin bổ sung được cài xen vào tại vị trí 2.

<400> 31

```

atggctgaga tcaacaacca gaaccagtgc gtcccgtaca actgcctgag caaccctaag      60
gagatcatcc tgggtgagga acgcctggag accggcaaca ccgtagccga cattagcctg      120
ggcctcatca acttcctcta cagcaacttc gtgcccggcg gtggcttcat cgtgggcctc      180
ctggagctta tctggggcct catcggcccg tccagtgagg acatcttcct cgcccagatc      240
gagcaactga tcagccagcg gatcgaggag ttcgctagga accaggccat ctccgcctg      300
gagggactct ccaacctcta caaggtgtac gtgcgcgcgt tcagcgactg ggagaaggac      360
ccgaccaacc cggccctccg cgaggaaatg cgtatccagt tcaacgatat gaactcggcc      420
ctcatcaccg ccatcccgtc cttccgcgtg cagaactacg aggtggccct cctgtccgtg      480
tacgttcaag cgcacaacct ccacctctcc atcctccgcg acgtgagcgt gttcggcgag      540
cgctggggct acgacaccgc caccatcaac aaccgctact ccgacctcac ctccctcatc      600
cacgtttaca ccaaccactg cgtggacacg tacaaccagg gcctccgccg cctggagggc      660
cgcttcctct ccgactggat cgtgtacaac cgcttcgcc gccagctcac catctccgtc      720
ctggacatcg tcgccttctt tcccaactac gacatccgca cctaccctat ccagaccgcc      780
accagctca cccgcgaggt ctacctcgac ctcccgttca tcaacgagaa cctcagcccg      840
gccgccagct acccgacctt ctccgcgcgt gagtccgcc tcatctgcag cccgcacctc      900
gtggacttcc tcaactcctt caccatctac accgactccc tcgcccgcag cgcctactgg      960
ggcgggtcacc tcgtgaactc cttccgcacc ggcaccacta ccaacctcat ccgcagcccg      1020
ctctacggcc gcgagggcaa caccgagcgc ccggtgacca tcaccgccag cccgagcgtg      1080
cccatcttcc gcaccctcag ctacatcacc ggccctggaca acagcaaccc tgtggcgggc      1140
atcgagggcg tggagttcca gaacaccatc tccaggagca tctaccgcaa gagcggccct      1200
atcgacagct tcagcgagct gcctcctcag gacgccagcg tgagccctgc catcggtac      1260
agccacaggc tgtgccacgc caccttcctg gagcgcatca gcggccctcg catcgcgggc      1320
accgtgttct cgtggacca cgcagcgcc tctctacga acgaggtgtc tcctagtcgc      1380
atcaccaga toccttgggt caaggccac accctggcta gtggcgctag tgtcatcaag      1440
ggccctggct tcaccggtg tgacatcctg accaggaact ctatgggcga gctgggcact      1500
ctgaggttca ctttactg cgcctgcct cagtcttact acatccgctt ccgctacgct      1560
agtgtcgcta accgctctg tactttccgc tactctcagc ctccgtctta cgttatctct      1620
ttccctaaga ctatggacgc tggtagcct ctgaccagta ggagcttcgc tcacactact      1680
ctgttcactc ctatcacttt ctctagggct caggaggagt tcgacctgta catccagtct      1740

```

|                                                                     |      |
|---------------------------------------------------------------------|------|
| ggtgtgtaca tcgacaggat cgagttcatc cccgtgaccg ccacgttcga ggccgagtag   | 1800 |
| gaccttgagc gcgccagaa ggtggtgaac gccctcttca ctagcactaa ccagctaggc    | 1860 |
| ctgaagactg acgtgaccga ctaccacatc gaccaagtga gcaacctagt ggctgcctc    | 1920 |
| tccgacgagt tctgcctcga cgagaagcgc gagctgtccg agaagggtgaa gcacgccaaag | 1980 |
| cgcctctccg acgagcgcaa cctgctccag gaccccaact tcaggggcat caacaggcag   | 2040 |
| cccgaccgcg gctggcgcg ctcaccgac atcaccatcc agggcggtga cgacgtattc     | 2100 |
| aaggagaact acgtttaccct ccccggcacc ttcgacgagt gttacccac ctacctctac   | 2160 |
| cagaagatcg acgagtccaa gctgaaggcc tacaccgct accagctccg cggtacatc     | 2220 |
| gaggactccc aggacctgga aatctacctc atccgctaca acgccaagca cgagatcgtg   | 2280 |
| aacgtgcctg gcaccggcag cctctggcct ctacgctgg agaaccagat cggcccttgc    | 2340 |
| ggcgagccta accgctgcgc cctcacctc gagtggaacc ctgacctcca ctgctcgtgc    | 2400 |
| agggacggcg agaagtgcgc ccaccatagc caccatttct ctctggacat cgacgtgggc   | 2460 |
| tgcaccgacc tgaacgagga cctgggcgtg tgggttatct tcaagatcaa gaccaggac    | 2520 |
| ggtcacgcca ggctgggtaa cctggagttc cttgaggaaa agcctctgct gggtagggcc   | 2580 |
| ctggccaggg tcaagagggc tgagaagaaa tggagggata agagggagac cctgcagctg   | 2640 |
| gagaccacta tcgtctacaa ggaggctaag gagtctgtcg atgctctgtt cgtcaactct   | 2700 |
| cagtacgata gactgcaagc tgataccaac atcgctatga tccacgctgc ggataagcgg   | 2760 |
| gtccaccgga tccgggaggc ttaccttccg gagctttctg tcatcccggg tgtcaacgct   | 2820 |
| gcgatcttcg aggaacttga ggaacgcatc ttcactgcgt ttagtcttta cgatgcgcgg   | 2880 |
| aacatcatca agaacgggga cttcaacaat ggtctgctgt gctggaacgt caagggtcat   | 2940 |
| gtcgaggtcg aggaacaaaa caatcatcgt agtgtccttg tcatctctga gtgggaggcg   | 3000 |
| gaggtctctc aagaggtccg tgtttgcccg gggcgtgggt acattcttcg tgttactgcg   | 3060 |
| tacaaggagg ggtacgggga ggggtgcgtt actattcatg agattgagaa caatactgat   | 3120 |
| gagcttaagt tcaacaattg tgttgaggag gaggtttacc cgaacaatac tgttacgtgc   | 3180 |
| atcaactaca cggcaacgca agaggaatac gaggggacgt acacctcgcg taatagaggg   | 3240 |
| tatgatgagg cgtacggaaa caaccgctcg gttccagcag attatgcctc ggtttatgag   | 3300 |
| gagaagtcgt acacggatag acgacgcgag aatccatgtg agtcaaatac aggatacggg   | 3360 |
| gattacacac cattaccagc aggatacgtt acaaggagt tggaatactt cccggaaaca    | 3420 |
| gataaagttt ggattgaaat cggagaaaca gaaggaacat tcatcgtcga ctcagtagaa   | 3480 |
| ttgttggtga tggaagaatg a                                             | 3501 |
| <210> 32                                                            |      |
| <211> 1166                                                          |      |
| <212> PRT                                                           |      |

<213> Trình tự nhân tạo

<220>

<223> Trình tự axit amin của protein đã được xử lý có hoạt tính trừ côn trùng Cry1Da1\_3 được ghi mã bởi trình tự ADN tổng hợp trong đó gốc Alanin bổ sung được cài xen vào tại vị trí 2.

<400> 32

```

Met Ala Glu Ile Asn Asn Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu
1           5           10           15

Ser Asn Pro Lys Glu Ile Ile Leu Gly Glu Glu Arg Leu Glu Thr Gly
          20           25           30

Asn Thr Val Ala Asp Ile Ser Leu Gly Leu Ile Asn Phe Leu Tyr Ser
          35           40           45

Asn Phe Val Pro Gly Gly Gly Phe Ile Val Gly Leu Leu Glu Leu Ile
          50           55           60

Trp Gly Phe Ile Gly Pro Ser Gln Trp Asp Ile Phe Leu Ala Gln Ile
65           70           75           80

Glu Gln Leu Ile Ser Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala
          85           90           95

Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Lys Val Tyr Val Arg
          100          105          110

Ala Phe Ser Asp Trp Glu Lys Asp Pro Thr Asn Pro Ala Leu Arg Glu
          115          120          125

Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Ile Thr Ala
          130          135          140

Ile Pro Leu Phe Arg Val Gln Asn Tyr Glu Val Ala Leu Leu Ser Val
          145          150          155          160

Tyr Val Gln Ala Ala Asn Leu His Leu Ser Ile Leu Arg Asp Val Ser
          165          170          175

Val Phe Gly Glu Arg Trp Gly Tyr Asp Thr Ala Thr Ile Asn Asn Arg
          180          185          190

Tyr Ser Asp Leu Thr Ser Leu Ile His Val Tyr Thr Asn His Cys Val
          195          200          205

Asp Thr Tyr Asn Gln Gly Leu Arg Arg Leu Glu Gly Arg Phe Leu Ser
          210          215          220

Asp Trp Ile Val Tyr Asn Arg Phe Arg Arg Gln Leu Thr Ile Ser Val
          225          230          235          240

Leu Asp Ile Val Ala Phe Phe Pro Asn Tyr Asp Ile Arg Thr Tyr Pro
          245          250          255

Ile Gln Thr Ala Thr Gln Leu Thr Arg Glu Val Tyr Leu Asp Leu Pro
          260          265          270

Phe Ile Asn Glu Asn Leu Ser Pro Ala Ala Ser Tyr Pro Thr Phe Ser
          275          280          285

Ala Ala Glu Ser Ala Ile Ile Arg Ser Pro His Leu Val Asp Phe Leu

```

| 290 |     |     |     |     | 295 |     |     |     |     | 300 |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asn | Ser | Phe | Thr | Ile | Tyr | Thr | Asp | Ser | Leu | Ala | Arg | Ser | Ala | Tyr | Trp |
| 305 |     |     |     |     | 310 |     |     |     |     | 315 |     |     |     |     | 320 |
| Gly | Gly | His | Leu | Val | Asn | Ser | Phe | Arg | Thr | Gly | Thr | Thr | Thr | Asn | Leu |
|     |     |     |     | 325 |     |     |     |     | 330 |     |     |     |     | 335 |     |
| Ile | Arg | Ser | Pro | Leu | Tyr | Gly | Arg | Glu | Gly | Asn | Thr | Glu | Arg | Pro | Val |
|     |     |     | 340 |     |     |     |     | 345 |     |     |     |     | 350 |     |     |
| Thr | Ile | Thr | Ala | Ser | Pro | Ser | Val | Pro | Ile | Phe | Arg | Thr | Leu | Ser | Tyr |
|     |     | 355 |     |     |     |     | 360 |     |     |     |     | 365 |     |     |     |
| Ile | Thr | Gly | Leu | Asp | Asn | Ser | Asn | Pro | Val | Ala | Gly | Ile | Glu | Gly | Val |
|     | 370 |     |     |     |     | 375 |     |     |     |     | 380 |     |     |     |     |
| Glu | Phe | Gln | Asn | Thr | Ile | Ser | Arg | Ser | Ile | Tyr | Arg | Lys | Ser | Gly | Pro |
| 385 |     |     |     |     | 390 |     |     |     |     | 395 |     |     |     |     | 400 |
| Ile | Asp | Ser | Phe | Ser | Glu | Leu | Pro | Pro | Gln | Asp | Ala | Ser | Val | Ser | Pro |
|     |     |     |     | 405 |     |     |     |     | 410 |     |     |     |     | 415 |     |
| Ala | Ile | Gly | Tyr | Ser | His | Arg | Leu | Cys | His | Ala | Thr | Phe | Leu | Glu | Arg |
|     |     |     | 420 |     |     |     | 425 |     |     |     |     |     | 430 |     |     |
| Ile | Ser | Gly | Pro | Arg | Ile | Ala | Gly | Thr | Val | Phe | Ser | Trp | Thr | His | Arg |
|     |     | 435 |     |     |     |     | 440 |     |     |     |     | 445 |     |     |     |
| Ser | Ala | Ser | Pro | Thr | Asn | Glu | Val | Ser | Pro | Ser | Arg | Ile | Thr | Gln | Ile |
|     | 450 |     |     |     |     | 455 |     |     |     |     | 460 |     |     |     |     |
| Pro | Trp | Val | Lys | Ala | His | Thr | Leu | Ala | Ser | Gly | Ala | Ser | Val | Ile | Lys |
| 465 |     |     |     |     | 470 |     |     |     |     | 475 |     |     |     |     | 480 |
| Gly | Pro | Gly | Phe | Thr | Gly | Gly | Asp | Ile | Leu | Thr | Arg | Asn | Ser | Met | Gly |
|     |     |     |     | 485 |     |     |     |     | 490 |     |     |     |     | 495 |     |
| Glu | Leu | Gly | Thr | Leu | Arg | Val | Thr | Phe | Thr | Gly | Arg | Leu | Pro | Gln | Ser |
|     |     |     | 500 |     |     |     |     | 505 |     |     |     |     | 510 |     |     |
| Tyr | Tyr | Ile | Arg | Phe | Arg | Tyr | Ala | Ser | Val | Ala | Asn | Arg | Ser | Gly | Thr |
|     |     | 515 |     |     |     |     | 520 |     |     |     |     | 525 |     |     |     |
| Phe | Arg | Tyr | Ser | Gln | Pro | Pro | Ser | Tyr | Gly | Ile | Ser | Phe | Pro | Lys | Thr |
|     |     | 530 |     |     |     | 535 |     |     |     |     | 540 |     |     |     |     |
| Met | Asp | Ala | Gly | Glu | Pro | Leu | Thr | Ser | Arg | Ser | Phe | Ala | His | Thr | Thr |
| 545 |     |     |     |     | 550 |     |     |     |     | 555 |     |     |     |     | 560 |
| Leu | Phe | Thr | Pro | Ile | Thr | Phe | Ser | Arg | Ala | Gln | Glu | Glu | Phe | Asp | Leu |
|     |     |     |     | 565 |     |     |     |     | 570 |     |     |     |     | 575 |     |
| Tyr | Ile | Gln | Ser | Gly | Val | Tyr | Ile | Asp | Arg | Ile | Glu | Phe | Ile | Pro | Val |
|     |     |     | 580 |     |     |     |     | 585 |     |     |     |     | 590 |     |     |
| Thr | Ala | Thr | Phe | Glu | Ala | Glu | Tyr | Asp | Leu | Glu | Arg | Ala | Gln | Lys | Val |
|     |     | 595 |     |     |     |     | 600 |     |     |     |     | 605 |     |     |     |
| Val | Asn | Ala | Leu | Phe | Thr | Ser | Thr | Asn | Gln | Leu | Gly | Leu | Lys | Thr | Asp |
|     | 610 |     |     |     |     | 615 |     |     |     |     | 620 |     |     |     |     |
| Val | Thr | Asp | Tyr | His | Ile | Asp | Gln | Val | Ser | Asn | Leu | Val | Ala | Cys | Leu |
| 625 |     |     |     |     | 630 |     |     |     |     | 635 |     |     |     |     | 640 |

```

Ser Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val
      645                      650                      655

Lys His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro
      660                      665                      670

Asn Phe Arg Gly Ile Asn Arg Gln Pro Asp Arg Gly Trp Arg Gly Ser
      675                      680                      685

Thr Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr
      690                      695                      700

Val Thr Leu Pro Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr
      705                      710                      715                      720

Gln Lys Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu
      725                      730                      735

Arg Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg
      740                      745                      750

Tyr Asn Ala Lys His Glu Ile Val Asn Val Pro Gly Thr Gly Ser Leu
      755                      760                      765

Trp Pro Leu Ser Val Glu Asn Gln Ile Gly Pro Cys Gly Glu Pro Asn
      770                      775                      780

Arg Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu His Cys Ser Cys
      785                      790                      795                      800

Arg Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp
      805                      810                      815

Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val
      820                      825                      830

Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu
      835                      840                      845

Glu Phe Leu Glu Glu Lys Pro Leu Leu Gly Glu Ala Leu Ala Arg Val
      850                      855                      860

Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Thr Leu Gln Leu
      865                      870                      875                      880

Glu Thr Thr Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu
      885                      890                      895

Phe Val Asn Ser Gln Tyr Asp Arg Leu Gln Ala Asp Thr Asn Ile Ala
      900                      905                      910

Met Ile His Ala Ala Asp Lys Arg Val His Arg Ile Arg Glu Ala Tyr
      915                      920                      925

Leu Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu
      930                      935                      940

Glu Leu Glu Glu Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg
      945                      950                      955                      960

Asn Ile Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Leu Cys Trp Asn
      965                      970                      975

```

Val Lys Gly His Val Glu Val Glu Glu Gln Asn Asn His Arg Ser Val  
980 985 990

Leu Val Ile Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val  
995 1000 1005

Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu  
1010 1015 1020

Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn  
1025 1030 1035

Thr Asp Glu Leu Lys Phe Asn Asn Cys Val Glu Glu Glu Val Tyr  
1040 1045 1050

Pro Asn Asn Thr Val Thr Cys Ile Asn Tyr Thr Ala Thr Gln Glu  
1055 1060 1065

Glu Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly Tyr Asp Glu  
1070 1075 1080

Ala Tyr Gly Asn Asn Pro Ser Val Pro Ala Asp Tyr Ala Ser Val  
1085 1090 1095

Tyr Glu Glu Lys Ser Tyr Thr Asp Arg Arg Arg Glu Asn Pro Cys  
1100 1105 1110

Glu Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu Pro Ala Gly  
1115 1120 1125

Tyr Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val  
1130 1135 1140

Trp Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser  
1145 1150 1155

Val Glu Leu Leu Leu Met Glu Glu  
1160 1165

<210> 33  
<211> 3501  
<212> ADN  
<213> Trình tự nhân tạo

<220>  
<223> Trình tự ADN tổng hợp được thiết kế để biểu hiện ở thực vật ghi mã  
Cry1Da1\_4 có gốc Alanin bổ sung được cài xen vào tại vị trí 2.

<400> 33  
atggctgaga tcaacaacca gaaccagtgc gtcccgtaga actgcctgag caaccctaag 60  
gagatcatcc tgggtgagga acgcctggag accggcaaca ccgtagccga cattagcctg 120  
ggcctcatca acttcctcta cagcaacttc gtgccggcg gtggcttcat cgtgggcctc 180  
ctggagctta tctggggctt catcggccg tccagtgagg acatcttctt cgccagatc 240  
gagcaactga tcagccagcg gatcgaggag ttcgctagga accaggccat ctcccgcctg 300  
gagggactct ccaacctcta caaggtgtac gtgcgcgcgt tcagcgactg ggagaaggac 360  
ccgaccaacc cggccctccg cgaggaaatg cgtatccagt tcaacgatat gaactcggcc 420  
ctcatcaccg ccatcccgct cttccgcgtg cagaactacg aggtggccct cctgtccgtg 480

|                                                                    |      |
|--------------------------------------------------------------------|------|
| tacgttcaag cgcgaacct ccacctctcc atcctccgcg acgtgagcgt gttcggcgag   | 540  |
| cgctggggct acgacaccgc caccatcaac aaccgctact ccgacctcac ctccctcatc  | 600  |
| cacgtttaca ccaaccactg cgtggacacg tacaaccagg gcctccgccg cctggagggc  | 660  |
| cgcttcctct ccgactggat cgtgtacaac cgcttccgcc gccagctcac catctccgtc  | 720  |
| ctggacatcg tcgccttctt tcccaactac gacatccgca cctaccctat ccagaccgcc  | 780  |
| accagctca cccgcgaggt ctacctcgac ctcccgttca tcaacgagaa cctcagcccc   | 840  |
| gccgccagct acccgacctt ctccgccgct gagtccgcca tcattcgag cccgcacctc   | 900  |
| gtggacttcc tcaactcctt caccatctac accgactccc tcgcccgcta cgctactgg   | 960  |
| ggcggtcacc tcgtgaactc cttccgcacc ggcaccacta ccaacctcat ccgcagcccc  | 1020 |
| ctctacggcc gcgagggcaa caccgagcgc ccggtgacca tcaccgccag cccgagcgtg  | 1080 |
| cccatcttcc gcacctcag ctacatcacc ggcttgaca acaggaacct tgtggcgggc    | 1140 |
| atcgagggcg tggagttcca gaacaccatc tccaggagca tctaccgcaa gagcggcct   | 1200 |
| atcgacagct tcagcgagct gcctcctcag gacgccagcg tgagccctgc catcggtac   | 1260 |
| agccacaggc tgtgccacgc caccttcctg gacgcgatca gcggccctcg catcgcgggc  | 1320 |
| accgtgttct cgtggacca cgcgagcgc tctctacga acgaggtgtc tcctagtgc      | 1380 |
| atcaccaga tcccttgggt caaggccac accctggcta gtggcgctag tgtcatcaag    | 1440 |
| ggccctggct tcaccggtg tgacatcctg accaggaact ctatgggcga gctgggcact   | 1500 |
| ctgagggta ctttactgg ccgcctgcct cagtcttact acatccgctt ccgctacgct    | 1560 |
| agtgtcgcta accgctctgg tactttccgc tactctcagc ctccgtctta cggatatctt  | 1620 |
| ttccctaaga ctatggacgc tggtagcct ctgaccagta ggagcttcgc tcacactact   | 1680 |
| ctgttcactc ctatcacttt ctctagggct caggaggagt tcgacctgta catccagtct  | 1740 |
| ggtgtgtaca tcgacaggat cgagttcatc cccgtgaccg ccacgttcga ggccgagtac  | 1800 |
| gaccttgagc gcgccagaa ggtggtgaac gccctcttca ctagcactaa ccagctaggc   | 1860 |
| ctgaagactg acgtgaccga ctaccacatc gaccaagtga gcaacctagt ggctgcctc   | 1920 |
| tccgacgagt tctgcctcga cgagaagcgc gagctgtccg agaagggtgaa gcacgccaag | 1980 |
| cgcctctccg acgagcgcaa cctgctccag gaccccaact tcaggggcat caacaggcag  | 2040 |
| ccgaccgcg gctggcgcg ctccaccgac atcaccatcc agggcggtga cgacgtattc    | 2100 |
| aaggagaact acgttaccct ccccggcacc ttcgacgagt gttacccac ctacctctac   | 2160 |
| cagaagatcg acgagtccaa gctgaaggcc tacaccgct accagctccg cggctacatc   | 2220 |
| gaggactccc aggacctgga aatctacctc atccgctaca acgccaagca cgagatcgtg  | 2280 |
| aacgtgcctg gcaccggcag cctctggcct ctcagcgtgg agaaccagat cggcccttgc  | 2340 |
| ggcgagccta accgctgcgc cctcacctc gagtggaacc ctgacctcca ctgctcgtgc   | 2400 |

```

agggacggcg agaagtgcgc ccaccatagc caccacttct ctctggacat cgacgtgggc      2460
tgcaccgacc tgaacgagga cctgggctgt tgggttatct tcaagatcaa gaccaggac      2520
ggtcacgcca ggctgggtaa cctggagttc cttgaggaaa agcctctgct gggtgaggcc      2580
ctggccaggg tcaagagggc tgagaagaaa tggagggata agagggagac cctgcagctg      2640
gagaccacta tcgtctacaa ggaggctaag gagtctgtcg atgctctgtt cgtcaactct      2700
cagtacgata gactgcaagc tgataccaac atcgctatga tccacgctgc ggataagcgg      2760
gtccaccgga tccgggaggc ttaccttcgc gagctttctg tcatcccggt tgtcaacgct      2820
gcgatcttcg aggaacttga ggaacggatc ttcactgcgt ttagtcttta cgatgcgcgg      2880
aacatcatca agaacgggga cttcaacaat ggtctgctgt gctggaacgt caagggtcat      2940
gtcgaggtcg aggaacaaaa caatcatcgt agtgtccttg tcattcctga gtgggaggcg      3000
gaggtctctc aagaggtccg tgtttgcccg gggcgtgggt acattcttcg tgttactgcg      3060
tacaaggagg ggtacgggga ggggtgcgtt actattcatg agattgagaa caatactgat      3120
gagcttaagt tcaacaattg tgttgaggag gaggtttacc cgaacaatac tgttacgtgc      3180
atcaactaca cggcaacgca agaggaatac gaggggacgt acacctcgcg taatagaggg      3240
tatgatgagg cgtacggaaa caaccgctcg gttccagcag attatgcctc ggtttatgag      3300
gagaagtcgt acacggatag acgacgcgag aatccatgtg agtcaaatcg aggatacgg      3360
gattacacac cattaccagc aggatacgtt acaaaggagt tggaatactt cccggaaaca      3420
gataaagttt ggattgaaat cggagaaaca gaaggaacat tcatcgctga ctcagtagaa      3480
ttgttggtga tggaagaatg a                                          3501

```

<210> 34

<211> 1166

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Trình tự axit amin của protein đã được xử lý có hoạt tính trừ côn trùng Cry1Da1\_4 được ghi mã bởi trình tự ADN tổng hợp trong đó gốc Alanin bổ sung được cài xen vào tại vị trí 2.

<400> 34

```

Met Ala Glu Ile Asn Asn Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu
1           5           10           15

```

```

Ser Asn Pro Lys Glu Ile Ile Leu Gly Glu Glu Arg Leu Glu Thr Gly
          20           25           30

```

```

Asn Thr Val Ala Asp Ile Ser Leu Gly Leu Ile Asn Phe Leu Tyr Ser
          35           40           45

```

```

Asn Phe Val Pro Gly Gly Gly Phe Ile Val Gly Leu Leu Glu Leu Ile
          50           55           60

```

```

Trp Gly Phe Ile Gly Pro Ser Gln Trp Asp Ile Phe Leu Ala Gln Ile

```

| 65                  | 70                  | 75                      | 80  |
|---------------------|---------------------|-------------------------|-----|
| Glu Gln Leu Ile Ser | Gln Arg Ile Glu Glu | Phe Ala Arg Asn Gln Ala |     |
|                     | 85                  | 90                      | 95  |
| Ile Ser Arg Leu Glu | Gly Leu Ser Asn Leu | Tyr Lys Val Tyr Val Arg |     |
|                     | 100                 | 105                     | 110 |
| Ala Phe Ser Asp Trp | Glu Lys Asp Pro Thr | Asn Pro Ala Leu Arg Glu |     |
|                     | 115                 | 120                     | 125 |
| Glu Met Arg Ile Gln | Phe Asn Asp Met Asn | Ser Ala Leu Ile Thr Ala |     |
|                     | 130                 | 135                     | 140 |
| Ile Pro Leu Phe Arg | Val Gln Asn Tyr Glu | Val Ala Leu Leu Ser Val |     |
| 145                 | 150                 | 155                     | 160 |
| Tyr Val Gln Ala Ala | Asn Leu His Leu Ser | Ile Leu Arg Asp Val Ser |     |
|                     | 165                 | 170                     | 175 |
| Val Phe Gly Glu Arg | Trp Gly Tyr Asp Thr | Ala Thr Ile Asn Asn Arg |     |
|                     | 180                 | 185                     | 190 |
| Tyr Ser Asp Leu Thr | Ser Leu Ile His Val | Tyr Thr Asn His Cys Val |     |
|                     | 195                 | 200                     | 205 |
| Asp Thr Tyr Asn Gln | Gly Leu Arg Arg Leu | Glu Gly Arg Phe Leu Ser |     |
|                     | 210                 | 215                     | 220 |
| Asp Trp Ile Val Tyr | Asn Arg Phe Arg Arg | Gln Leu Thr Ile Ser Val |     |
| 225                 | 230                 | 235                     | 240 |
| Leu Asp Ile Val Ala | Phe Phe Pro Asn Tyr | Asp Ile Arg Thr Tyr Pro |     |
|                     | 245                 | 250                     | 255 |
| Ile Gln Thr Ala Thr | Gln Leu Thr Arg Glu | Val Tyr Leu Asp Leu Pro |     |
|                     | 260                 | 265                     | 270 |
| Phe Ile Asn Glu Asn | Leu Ser Pro Ala Ala | Ser Tyr Pro Thr Phe Ser |     |
|                     | 275                 | 280                     | 285 |
| Ala Ala Glu Ser Ala | Ile Ile Arg Ser Pro | His Leu Val Asp Phe Leu |     |
|                     | 290                 | 295                     | 300 |
| Asn Ser Phe Thr Ile | Tyr Thr Asp Ser Leu | Ala Arg Tyr Ala Tyr Trp |     |
| 305                 | 310                 | 315                     | 320 |
| Gly Gly His Leu Val | Asn Ser Phe Arg Thr | Gly Thr Thr Thr Asn Leu |     |
|                     | 325                 | 330                     | 335 |
| Ile Arg Ser Pro Leu | Tyr Gly Arg Glu Gly | Asn Thr Glu Arg Pro Val |     |
|                     | 340                 | 345                     | 350 |
| Thr Ile Thr Ala Ser | Pro Ser Val Pro Ile | Phe Arg Thr Leu Ser Tyr |     |
|                     | 355                 | 360                     | 365 |
| Ile Thr Gly Leu Asp | Asn Arg Asn Pro Val | Ala Gly Ile Glu Gly Val |     |
|                     | 370                 | 375                     | 380 |
| Glu Phe Gln Asn Thr | Ile Ser Arg Ser Ile | Tyr Arg Lys Ser Gly Pro |     |
| 385                 | 390                 | 395                     | 400 |
| Ile Asp Ser Phe Ser | Glu Leu Pro Pro Gln | Asp Ala Ser Val Ser Pro |     |
|                     | 405                 | 410                     | 415 |

Ala Ile Gly Tyr Ser His Arg Leu Cys His Ala Thr Phe Leu Glu Arg  
 420 425 430  
 Ile Ser Gly Pro Arg Ile Ala Gly Thr Val Phe Ser Trp Thr His Arg  
 435 440 445  
 Ser Ala Ser Pro Thr Asn Glu Val Ser Pro Ser Arg Ile Thr Gln Ile  
 450 455 460  
 Pro Trp Val Lys Ala His Thr Leu Ala Ser Gly Ala Ser Val Ile Lys  
 465 470 475 480  
 Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu Thr Arg Asn Ser Met Gly  
 485 490 495  
 Glu Leu Gly Thr Leu Arg Val Thr Phe Thr Gly Arg Leu Pro Gln Ser  
 500 505 510  
 Tyr Tyr Ile Arg Phe Arg Tyr Ala Ser Val Ala Asn Arg Ser Gly Thr  
 515 520 525  
 Phe Arg Tyr Ser Gln Pro Pro Ser Tyr Gly Ile Ser Phe Pro Lys Thr  
 530 535 540  
 Met Asp Ala Gly Glu Pro Leu Thr Ser Arg Ser Phe Ala His Thr Thr  
 545 550 555 560  
 Leu Phe Thr Pro Ile Thr Phe Ser Arg Ala Gln Glu Glu Phe Asp Leu  
 565 570 575  
 Tyr Ile Gln Ser Gly Val Tyr Ile Asp Arg Ile Glu Phe Ile Pro Val  
 580 585 590  
 Thr Ala Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala Gln Lys Val  
 595 600 605  
 Val Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Leu Lys Thr Asp  
 610 615 620  
 Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Ala Cys Leu  
 625 630 635 640  
 Ser Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val  
 645 650 655  
 Lys His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro  
 660 665 670  
 Asn Phe Arg Gly Ile Asn Arg Gln Pro Asp Arg Gly Trp Arg Gly Ser  
 675 680 685  
 Thr Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr  
 690 695 700  
 Val Thr Leu Pro Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr  
 705 710 715 720  
 Gln Lys Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu  
 725 730 735  
 Arg Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg  
 740 745 750

Tyr Asn Ala Lys His Glu Ile Val Asn Val Pro Gly Thr Gly Ser Leu  
 755 760 765  
 Trp Pro Leu Ser Val Glu Asn Gln Ile Gly Pro Cys Gly Glu Pro Asn  
 770 775 780  
 Arg Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu His Cys Ser Cys  
 785 790 795 800  
 Arg Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp  
 805 810 815  
 Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val  
 820 825 830  
 Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu  
 835 840 845  
 Glu Phe Leu Glu Glu Lys Pro Leu Leu Gly Glu Ala Leu Ala Arg Val  
 850 855 860  
 Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Thr Leu Gln Leu  
 865 870 875 880  
 Glu Thr Thr Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu  
 885 890 895  
 Phe Val Asn Ser Gln Tyr Asp Arg Leu Gln Ala Asp Thr Asn Ile Ala  
 900 905 910  
 Met Ile His Ala Ala Asp Lys Arg Val His Arg Ile Arg Glu Ala Tyr  
 915 920 925  
 Leu Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu  
 930 935 940  
 Glu Leu Glu Glu Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg  
 945 950 955 960  
 Asn Ile Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Leu Cys Trp Asn  
 965 970 975  
 Val Lys Gly His Val Glu Val Glu Glu Gln Asn Asn His Arg Ser Val  
 980 985 990  
 Leu Val Ile Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val  
 995 1000 1005  
 Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu  
 1010 1015 1020  
 Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn  
 1025 1030 1035  
 Thr Asp Glu Leu Lys Phe Asn Asn Cys Val Glu Glu Glu Val Tyr  
 1040 1045 1050  
 Pro Asn Asn Thr Val Thr Cys Ile Asn Tyr Thr Ala Thr Gln Glu  
 1055 1060 1065  
 Glu Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly Tyr Asp Glu  
 1070 1075 1080  
 Ala Tyr Gly Asn Asn Pro Ser Val Pro Ala Asp Tyr Ala Ser Val

|                                                             |      |      |  |
|-------------------------------------------------------------|------|------|--|
| 1085                                                        | 1090 | 1095 |  |
| Tyr Glu Glu Lys Ser Tyr Thr Asp Arg Arg Arg Glu Asn Pro Cys |      |      |  |
| 1100                                                        | 1105 | 1110 |  |
| Glu Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu Pro Ala Gly |      |      |  |
| 1115                                                        | 1120 | 1125 |  |
| Tyr Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val |      |      |  |
| 1130                                                        | 1135 | 1140 |  |
| Trp Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser |      |      |  |
| 1145                                                        | 1150 | 1155 |  |
| Val Glu Leu Leu Leu Met Glu Glu                             |      |      |  |
| 1160                                                        | 1165 |      |  |

<210> 35  
 <211> 3501  
 <212> ADN  
 <213> Trình tự nhân tạo

<220>  
 <223> Trình tự ADN tổng hợp được thiết kế để biểu hiện ở thực vật ghi mã Cry1Da1\_5 có gốc Alanin bổ sung được cài xen vào tại vị trí 2.

<400> 35  
 atggctgaga tcaacaacca gaaccagtgc gtcccgtaga actgcctgag caaccctaag 60  
 gagatcatcc tgggtgagga acgcctggag accggcaaca ccgtagccga cattagcctg 120  
 ggccatcatca acttcctcta cagcaacttc gtgcccggcg gtggcttcat cgtgggcctc 180  
 ctggagctta tctggggctt catcggcccg tccagtgagg acatcttctc cgcccagatc 240  
 gagcaactga tcagccagcg gatcgaggag ttogctagga accaggccat ctccgcctg 300  
 gagggactct ccaacctcta caaggtgtac gtgcgcgcgt tcagcgactg ggagaaggac 360  
 ccgaccaacc cggccctccg cgaggaaatg cgtatccagt tcaacgatat gaactoggcc 420  
 ctcatcaccg ccatcccgtc cttccgcgtg cagaactacg aggtggccct cctgtccgtg 480  
 tacgttcaag ccgccaacct ccacctctcc atcctccgcg acgtgagcgt gttcggcgag 540  
 cgctgggggt acgacaccgc caccatcaac aaccgctact ccgacctcac ctccctcatc 600  
 cacgtttaca ccaaccactg cgtggacacg tacaaccagg gcctccgccg cctggagggc 660  
 cgcttctctc ccgactggat cgtgtacaac cgcttccgcc gccagctcac catctccgtc 720  
 ctggacatcg tcgccttctt tcccaactac gacatccgca cctaccctat ccagaccgcc 780  
 acccagctca cccgcgaggt ctacctcgac ctcccgttca tcaacgagaa cctcagcccc 840  
 gccgccagct acccgacctt ctccgccgct gagtccgcca tcattogcag cccgcacctc 900  
 gtggacttcc tcaactcctt caccatctac accgactccc tcgcccgcag cgcctactgg 960  
 ggcggtcacc tcgtgaactc cttccgcacc ggcaccacta ccaacctcat ccgcagcccc 1020  
 ctctacggcc gcgagggcaa caccgagcgc ccggtgacca tcaccgccag cccgagcgtg 1080  
 cccatcttcc gcaccctcag ctaccgcacc ggccctggaca acagcaaccc tgtggcgggc 1140

|                                                                    |      |
|--------------------------------------------------------------------|------|
| atcgagggcg tggagttcca gaacaccatc tccaggagca tctaccgcaa gagcggccct  | 1200 |
| atcgacagct tcagcgagct gcctcctcag gacgccagcg tgagccctgc catcggctac  | 1260 |
| agccacaggc tgtgccacgc caccttctcg gagcgcatca ggggccctcg catcgcgggc  | 1320 |
| accgtgttct cgtggacca cgcagcgcc tctcctacga acgaggtgtc tctagtcgc     | 1380 |
| atcaccaga tcccttgggt caaggccac accctggcta gtggcgctag tgtcatcaag    | 1440 |
| ggccctggct tcaccggtgg tgacatcctg accaggaact ctatgggcga gctgggcact  | 1500 |
| ctgaggttca ctttacttgg ccgcctgcct cagtcttact acatccgctt ccgctacgt   | 1560 |
| agtgtcgcta accgctctgg tactttccgc tactctcagc ctccgtctta cggtatctct  | 1620 |
| ttccctaaga ctatggacgc tggtagacct ctgaccagta ggagcttcgc tcacactact  | 1680 |
| ctgttctact ctatcacttt ctctagggct caggaggagt tcgacctgta catccagtct  | 1740 |
| ggtgtgtaca tcgacaggat cgagttcatc cccgtgaccg ccacgttcga ggccgagtac  | 1800 |
| gaccttgagc gcgccagaa ggtggtgaac gccctcttca ctagcactaa ccagctaggc   | 1860 |
| ctgaagactg acgtgaccga ctaccacatc gaccaagtga gcaacctagt ggccctgcctc | 1920 |
| tccgacgagt tctgcctcga cgagaagcg gagctgtccg agaaggtgaa gcacgccaa    | 1980 |
| cgctctccg acgagcgcaa cctgctccag gaccccaact tcaggggcat caacaggcag   | 2040 |
| cccgaccgcg gctggcgcg ctccaccgac atcaccatcc agggcggtga cgacgtattc   | 2100 |
| aaggagaact acgttaccct ccccggcacc ttgcacgagt gttacccac ctacctctac   | 2160 |
| cagaagatcg acgagtccaa gctgaaggcc tacaccgct accagctccg cggctacatc   | 2220 |
| gaggactccc aggacctgga aatctacctc atccgctaca acgccaagca cgagatcgtg  | 2280 |
| aacgtgcctg gcaccggcag cctctggcct ctcagcgtag agaaccagat cggcccttgc  | 2340 |
| ggcgagccta accgctgcgc cctcacctc gagtggaaac ctgacctca ctgctcgtgc    | 2400 |
| agggacggcg agaagtgcgc ccaccatagc caccacttct ctctggacat cgacgtgggc  | 2460 |
| tgcaccgacc tgaacgagga cctgggcgtg tgggttatct tcaagatcaa gaccaggac   | 2520 |
| ggtcacgcca ggctgggtaa cctggagttc cttgaggaaa agcctctgct gggtagggcc  | 2580 |
| ctggccaggg tcaagagggc tgagaagaaa tggagggata agaggagac cctgcagctg   | 2640 |
| gagaccacta tcgtctacaa ggaggctaag gagtctgtcg atgctctgtt cgtcaactct  | 2700 |
| cagtacgata gactgcaagc tgataccaac atcgctatga tccacgctgc ggataagcg   | 2760 |
| gtccaccgga tccgggaggc ttaccttccg gagctttctg tcatcccggt tgtcaacgct  | 2820 |
| gcgatcttgc aggaacttga ggaacgatc ttactgcgt ttagtcttta cgatgcgcgg    | 2880 |
| aacatcatca agaacgggga cttcaacaat ggtctgctgt gctggaacgt caagggtcat  | 2940 |
| gtcgaggtcg aggaacaaaa caatcatcgt agtgtccttgc tcatcctga gtgggaggcg  | 3000 |
| gaggtctctc aagaggtccg tgtttgcccg gggcgtaggt acattcttgc tgttactgcg  | 3060 |

```

tacaaggagg ggtacgggga ggggtgcggtt actattcatg agattgagaa caatactgat      3120
gagcttaagt tcaacaattg tggtgaggag gaggtttacc cgaacaatac tgttacgtgc      3180
atcaactaca cggcaacgca agaggaatac gaggggacgt acacctcgcg taatagaggg      3240
tatgatgagg cgtacggaaa caaccgctcg gttccagcag attatgcctc ggtttatgag      3300
gagaagtcgt acacggatag acgacgcgag aatccatgtg agtcaaatcg aggatacggg      3360
gattacacac cattaccagc aggatacgtt acaaaggagt tggaatactt cccggaaaca      3420
gataaagttt ggattgaaat cggagaaaca gaaggaacat tcatcgtcga ctcagtagaa      3480
ttgttggtga tggaagaatg a                                              3501

```

<210> 36

<211> 1166

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Trình tự axit amin của protein đã được xử lý có hoạt tính trừ côn trùng Cry1Dal\_5 được ghi mã bởi trình tự ADN tổng hợp trong đó gốc Alanin bổ sung được cài xen vào tại vị trí 2.

<400> 36

```

Met Ala Glu Ile Asn Asn Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu
1              5              10              15

Ser Asn Pro Lys Glu Ile Ile Leu Gly Glu Glu Arg Leu Glu Thr Gly
              20              25              30

Asn Thr Val Ala Asp Ile Ser Leu Gly Leu Ile Asn Phe Leu Tyr Ser
              35              40              45

Asn Phe Val Pro Gly Gly Gly Phe Ile Val Gly Leu Leu Glu Leu Ile
              50              55              60

Trp Gly Phe Ile Gly Pro Ser Gln Trp Asp Ile Phe Leu Ala Gln Ile
65              70              75              80

Glu Gln Leu Ile Ser Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala
              85              90              95

Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Lys Val Tyr Val Arg
              100             105             110

Ala Phe Ser Asp Trp Glu Lys Asp Pro Thr Asn Pro Ala Leu Arg Glu
              115             120             125

Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Ile Thr Ala
              130             135             140

Ile Pro Leu Phe Arg Val Gln Asn Tyr Glu Val Ala Leu Leu Ser Val
145             150             155             160

Tyr Val Gln Ala Ala Asn Leu His Leu Ser Ile Leu Arg Asp Val Ser
              165             170             175

Val Phe Gly Glu Arg Trp Gly Tyr Asp Thr Ala Thr Ile Asn Asn Arg
              180             185             190

```

Tyr Ser Asp Leu Thr Ser Leu Ile His Val Tyr Thr Asn His Cys Val  
 195 200 205  
 Asp Thr Tyr Asn Gln Gly Leu Arg Arg Leu Glu Gly Arg Phe Leu Ser  
 210 215 220  
 Asp Trp Ile Val Tyr Asn Arg Phe Arg Arg Gln Leu Thr Ile Ser Val  
 225 230 235 240  
 Leu Asp Ile Val Ala Phe Phe Pro Asn Tyr Asp Ile Arg Thr Tyr Pro  
 245 250 255  
 Ile Gln Thr Ala Thr Gln Leu Thr Arg Glu Val Tyr Leu Asp Leu Pro  
 260 265 270  
 Phe Ile Asn Glu Asn Leu Ser Pro Ala Ala Ser Tyr Pro Thr Phe Ser  
 275 280 285  
 Ala Ala Glu Ser Ala Ile Ile Arg Ser Pro His Leu Val Asp Phe Leu  
 290 295 300  
 Asn Ser Phe Thr Ile Tyr Thr Asp Ser Leu Ala Arg Ser Ala Tyr Trp  
 305 310 315 320  
 Gly Gly His Leu Val Asn Ser Phe Arg Thr Gly Thr Thr Thr Asn Leu  
 325 330 335  
 Ile Arg Ser Pro Leu Tyr Gly Arg Glu Gly Asn Thr Glu Arg Pro Val  
 340 345 350  
 Thr Ile Thr Ala Ser Pro Ser Val Pro Ile Phe Arg Thr Leu Ser Tyr  
 355 360 365  
 Arg Thr Gly Leu Asp Asn Ser Asn Pro Val Ala Gly Ile Glu Gly Val  
 370 375 380  
 Glu Phe Gln Asn Thr Ile Ser Arg Ser Ile Tyr Arg Lys Ser Gly Pro  
 385 390 395 400  
 Ile Asp Ser Phe Ser Glu Leu Pro Pro Gln Asp Ala Ser Val Ser Pro  
 405 410 415  
 Ala Ile Gly Tyr Ser His Arg Leu Cys His Ala Thr Phe Leu Glu Arg  
 420 425 430  
 Ile Ser Gly Pro Arg Ile Ala Gly Thr Val Phe Ser Trp Thr His Arg  
 435 440 445  
 Ser Ala Ser Pro Thr Asn Glu Val Ser Pro Ser Arg Ile Thr Gln Ile  
 450 455 460  
 Pro Trp Val Lys Ala His Thr Leu Ala Ser Gly Ala Ser Val Ile Lys  
 465 470 475 480  
 Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu Thr Arg Asn Ser Met Gly  
 485 490 495  
 Glu Leu Gly Thr Leu Arg Val Thr Phe Thr Gly Arg Leu Pro Gln Ser  
 500 505 510  
 Tyr Tyr Ile Arg Phe Arg Tyr Ala Ser Val Ala Asn Arg Ser Gly Thr  
 515 520 525

Phe Arg Tyr Ser Gln Pro Pro Ser Tyr Gly Ile Ser Phe Pro Lys Thr  
 530 535 540  
 Met Asp Ala Gly Glu Pro Leu Thr Ser Arg Ser Phe Ala His Thr Thr  
 545 550 555 560  
 Leu Phe Thr Pro Ile Thr Phe Ser Arg Ala Gln Glu Glu Phe Asp Leu  
 565 570 575  
 Tyr Ile Gln Ser Gly Val Tyr Ile Asp Arg Ile Glu Phe Ile Pro Val  
 580 585 590  
 Thr Ala Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala Gln Lys Val  
 595 600 605  
 Val Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Leu Lys Thr Asp  
 610 615 620  
 Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Ala Cys Leu  
 625 630 635 640  
 Ser Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val  
 645 650 655  
 Lys His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro  
 660 665 670  
 Asn Phe Arg Gly Ile Asn Arg Gln Pro Asp Arg Gly Trp Arg Gly Ser  
 675 680 685  
 Thr Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr  
 690 695 700  
 Val Thr Leu Pro Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr  
 705 710 715 720  
 Gln Lys Ile Asp Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu  
 725 730 735  
 Arg Gly Tyr Ile Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg  
 740 745 750  
 Tyr Asn Ala Lys His Glu Ile Val Asn Val Pro Gly Thr Gly Ser Leu  
 755 760 765  
 Trp Pro Leu Ser Val Glu Asn Gln Ile Gly Pro Cys Gly Glu Pro Asn  
 770 775 780  
 Arg Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu His Cys Ser Cys  
 785 790 795 800  
 Arg Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp  
 805 810 815  
 Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val  
 820 825 830  
 Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu  
 835 840 845  
 Glu Phe Leu Glu Glu Lys Pro Leu Leu Gly Glu Ala Leu Ala Arg Val  
 850 855 860  
 Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Thr Leu Gln Leu

|                                                                 |  |      |  |      |  |      |
|-----------------------------------------------------------------|--|------|--|------|--|------|
| 865                                                             |  | 870  |  | 875  |  | 880  |
| Glu Thr Thr Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu |  |      |  |      |  |      |
|                                                                 |  | 885  |  | 890  |  | 895  |
| Phe Val Asn Ser Gln Tyr Asp Arg Leu Gln Ala Asp Thr Asn Ile Ala |  |      |  |      |  |      |
|                                                                 |  | 900  |  | 905  |  | 910  |
| Met Ile His Ala Ala Asp Lys Arg Val His Arg Ile Arg Glu Ala Tyr |  |      |  |      |  |      |
|                                                                 |  | 915  |  | 920  |  | 925  |
| Leu Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu |  |      |  |      |  |      |
|                                                                 |  | 930  |  | 935  |  | 940  |
| Glu Leu Glu Glu Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg |  |      |  |      |  |      |
| 945                                                             |  | 950  |  | 955  |  | 960  |
| Asn Ile Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Leu Cys Trp Asn |  |      |  |      |  |      |
|                                                                 |  | 965  |  | 970  |  | 975  |
| Val Lys Gly His Val Glu Val Glu Glu Gln Asn Asn His Arg Ser Val |  |      |  |      |  |      |
|                                                                 |  | 980  |  | 985  |  | 990  |
| Leu Val Ile Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val |  |      |  |      |  |      |
|                                                                 |  | 995  |  | 1000 |  | 1005 |
| Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu     |  |      |  |      |  |      |
|                                                                 |  | 1010 |  | 1015 |  | 1020 |
| Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn     |  |      |  |      |  |      |
|                                                                 |  | 1025 |  | 1030 |  | 1035 |
| Thr Asp Glu Leu Lys Phe Asn Asn Cys Val Glu Glu Glu Val Tyr     |  |      |  |      |  |      |
|                                                                 |  | 1040 |  | 1045 |  | 1050 |
| Pro Asn Asn Thr Val Thr Cys Ile Asn Tyr Thr Ala Thr Gln Glu     |  |      |  |      |  |      |
|                                                                 |  | 1055 |  | 1060 |  | 1065 |
| Glu Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly Tyr Asp Glu     |  |      |  |      |  |      |
|                                                                 |  | 1070 |  | 1075 |  | 1080 |
| Ala Tyr Gly Asn Asn Pro Ser Val Pro Ala Asp Tyr Ala Ser Val     |  |      |  |      |  |      |
|                                                                 |  | 1085 |  | 1090 |  | 1095 |
| Tyr Glu Glu Lys Ser Tyr Thr Asp Arg Arg Arg Glu Asn Pro Cys     |  |      |  |      |  |      |
|                                                                 |  | 1100 |  | 1105 |  | 1110 |
| Glu Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu Pro Ala Gly     |  |      |  |      |  |      |
|                                                                 |  | 1115 |  | 1120 |  | 1125 |
| Tyr Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val     |  |      |  |      |  |      |
|                                                                 |  | 1130 |  | 1135 |  | 1140 |
| Trp Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser     |  |      |  |      |  |      |
|                                                                 |  | 1145 |  | 1150 |  | 1155 |
| Val Glu Leu Leu Leu Met Glu Glu                                 |  |      |  |      |  |      |
|                                                                 |  | 1160 |  | 1165 |  |      |

<210> 37  
 <211> 3501  
 <212> ADN  
 <213> Trình tự nhân tạo

&lt;220&gt;

<223> Trình tự ADN tổng hợp được thiết kế để biểu hiện ở thực vật ghi mã Cry1Da1\_6 có gốc Alanin bổ sung được cài xen vào tại vị trí 2.

&lt;400&gt; 37

```

atggctgaga tcaacaacca gaaccagtgc gtcccgtaga actgcctgag caaccctaag      60
gagatcatcc tgggtgagga acgcctggag accggcaaca ccgtagccga cattagcctg      120
ggcctcatca acttcctcta cagcaacttc gtgcccgggc gtggcttcat cgtgggcctc      180
ctggagctta tctggggcctt catcgccccg tccagtgagg acatcttcct cggccagatc      240
gagcaactga tcagccagcg gatcgaggag ttcgctagga accaggccat ctccgcctg      300
gagggactct ccaacctcta caaggtgtac gtgcgcgcgt tcagcgactg ggagaaggac      360
ccgaccaacc cggccctccg cgaggaaatg cgtatccagt tcaacgatat gaactcggcc      420
ctcatcaccc ccatcccgtt cttccgcgtg cagaactacg aggtggccct cctgtccgtg      480
tacgttcaag ccgccaacct ccacctctcc atcctccgcg acgtgagcgt gttcggcgag      540
cgctggggct acgacaccgc caccatcaac aaccgctact ccgacctcac ctccctcatc      600
cacgtttaca ccaaccactg cgtggacacg tacaaccagg gcctccgccg cctggagggc      660
cgcttcctct ccgactggat cgtgtacaac cgcttcgccg gccagctcac catctccgtc      720
ctggacatcg tcgccttctt tcccaactac gacatccgca cctaccctat ccagaccgcc      780
accagctca cccgcgaggt ctacctcgac ctcccgttca tcaacgagaa cctcagcccg      840
gccgccaagt acccgacctt ctccgccgtt gagtccgccg tcattcgag cccgaacctc      900
gtggacttcc tcaactcctt caccatctac accgactccc tcgcccgag cgctacttgg      960
ggcggtcacc tcgtgaactc cttccgcacc ggcaccacta ccaacctcat ccgcagcccg     1020
ctctacggcc gcgagggcaa caccgagcgc ccggtgacca tcaaccgag cccgagcgtg     1080
cccatcttcc gcaccctcag ctaccccacc ggcttggaac acagcaacct tgtggcgggc     1140
atcgagggcg tggagttcca gaacaccatc tccaggagca tctaccgcaa gagcggccct     1200
atcgacagct tcagcgagct gcctcctcag gacgccagcg tgagccctgc catcggtac     1260
agccacaggc tgtgccacgc caccttcctg gagcgcatca gcggccctcg catcgcgggc     1320
accgtgttct cgtggacca cgcagcgcc tctcctacga acgaggtgtc tcctagtcgc     1380
atcaccaga tcccttgggt caaggccac accctggcta gtggcgctag tgtcatcaag     1440
ggccctggct tcaccggtgg tgacatcctg accaggaact ctatgggcga gctgggcact     1500
ctgaggttca ctttacttgg ccgcctgcct cagtcttact acatccgctt ccgctacgct     1560
agtgtcgcta accgctctgg tactttccgc tactctcagc ctccgtctta cgttatctct     1620
ttccctaaga ctatggacgc tggtagacct ctgaccagta ggagcttcgc tcacactact     1680
ctgttcactc ctatcacttt ctctagggct caggaggagt tcgacctgta catccagtct     1740
ggtgtgtaca tcgacaggat cgagttcatc ccggtgaccg ccacgttcga ggccgagtac     1800

```

|                                                                    |      |
|--------------------------------------------------------------------|------|
| gaccttgagc ggcgccagaa ggtggtgaac gccctcttca ctagcactaa ccagctaggc  | 1860 |
| ctgaagactg acgtgaccga ctaccacatc gaccaagtga gcaacctagt ggctgcctc   | 1920 |
| tccgacgagt tctgcctcga cgagaagcgc gagctgtccg agaaggtgaa gcacgccaa   | 1980 |
| cgcctctccg acgagcgcaa cctgctccag gaccccaact tcaggggcat caacaggcag  | 2040 |
| cccgaccgcg gctggcgcgg ctccaccgac atcaccatcc agggcggtga cgacgtattc  | 2100 |
| aaggagaact acgttaccct ccccggcacc ttogacgagt gttacccac ctacctctac   | 2160 |
| cagaagatcg acgagtccaa gctgaaggcc tacaccgct accagctccg cggtacatc    | 2220 |
| gaggactccc aggacctgga aatctacctc atccgctaca acgccaagca cgagatcgtg  | 2280 |
| aacgtgcctg gcaccggcag cctctggcct ctacgctgg agaaccagat cggcccttgc   | 2340 |
| ggcgagccta accgctgcgc cctcacctc gagtggaacc ctgacctca ctgctcgtgc    | 2400 |
| agggacggcg agaagtgcgc ccaccatagc caccacttct ctctggacat cgacgtgggc  | 2460 |
| tgcaccgacc tgaacgagga cctgggctg tgggttatct tcaagatcaa gaccaggac    | 2520 |
| ggtcacgcca ggctgggtaa cctggagtgc cttgaggaaa agcctctgct gggtgaggcc  | 2580 |
| ctggccaggg tcaagagggc tgagaagaaa tggagggata agagggagac cctgcagctg  | 2640 |
| gagaccacta tcgtctacaa ggaggctaag gagtctgtcg atgctctgtt cgtcaactct  | 2700 |
| cagtacgata gactgcaagc tgataccaac atcgctatga tccacgctgc ggataagcgg  | 2760 |
| gtccaccgga tccgggaggc ttaccttccg gagctttctg tcatccggg tgtcaacgct   | 2820 |
| gcgatcttcg aggaacttga ggaacggatc ttcactgcgt ttagtcttta cgatgcgcgg  | 2880 |
| aacatcatca agaacgggga cttcaacaat ggtctgctgt gctggaacgt caagggtcat  | 2940 |
| gtcgaggctg aggaacaaaa caatcatcgt agtgtccttg tcatcctga gtgggaggcg   | 3000 |
| gaggtctctc aagaggtccg tgtttgcccg gggcggtgggt acattcttcg tgttactgcg | 3060 |
| tacaaggagg ggtacgggga ggggtgctgt actattcatg agattgagaa caatactgat  | 3120 |
| gagcttaagt tcaacaattg tgttgaggag gaggtttacc cgaacaatac tgttacgtgc  | 3180 |
| atcaactaca cggcaacgca agaggaatac gaggggacgt acacctcgcg taatagaggg  | 3240 |
| tatgatgagg cgtacggaaa caaccgctcg gttocagcag attatgcctc ggtttatgag  | 3300 |
| gagaagtcgt acacggatag acgacgcgag aatccatgtg agtcaaatac aggatacggg  | 3360 |
| gattacacac cattaccagc aggatacgtt acaaaggagt tggaatactt cccggaaaca  | 3420 |
| gataaagttt ggattgaaat cggagaaaca gaaggaacat tcatcgtcga ctacgtagaa  | 3480 |
| ttgttgttga tggaagaatg a                                            | 3501 |
| <210> 38                                                           |      |
| <211> 1166                                                         |      |
| <212> PRT                                                          |      |
| <213> Trình tự nhân tạo                                            |      |

&lt;220&gt;

<223> Trình tự axit amin của protein đã được xử lý có hoạt tính trừ côn trùng Cry1Da1\_6 được ghi mã bởi trình tự ADN tổng hợp trong đó gốc Alanin bổ sung được cài xen vào tại vị trí 2.

&lt;400&gt; 38

```

Met Ala Glu Ile Asn Asn Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu
1           5           10           15

Ser Asn Pro Lys Glu Ile Ile Leu Gly Glu Glu Arg Leu Glu Thr Gly
          20           25           30

Asn Thr Val Ala Asp Ile Ser Leu Gly Leu Ile Asn Phe Leu Tyr Ser
          35           40           45

Asn Phe Val Pro Gly Gly Gly Phe Ile Val Gly Leu Leu Glu Leu Ile
          50           55           60

Trp Gly Phe Ile Gly Pro Ser Gln Trp Asp Ile Phe Leu Ala Gln Ile
65           70           75           80

Glu Gln Leu Ile Ser Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala
          85           90           95

Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Lys Val Tyr Val Arg
          100          105          110

Ala Phe Ser Asp Trp Glu Lys Asp Pro Thr Asn Pro Ala Leu Arg Glu
          115          120          125

Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Ile Thr Ala
          130          135          140

Ile Pro Leu Phe Arg Val Gln Asn Tyr Glu Val Ala Leu Leu Ser Val
          145          150          155          160

Tyr Val Gln Ala Ala Asn Leu His Leu Ser Ile Leu Arg Asp Val Ser
          165          170          175

Val Phe Gly Glu Arg Trp Gly Tyr Asp Thr Ala Thr Ile Asn Asn Arg
          180          185          190

Tyr Ser Asp Leu Thr Ser Leu Ile His Val Tyr Thr Asn His Cys Val
          195          200          205

Asp Thr Tyr Asn Gln Gly Leu Arg Arg Leu Glu Gly Arg Phe Leu Ser
          210          215          220

Asp Trp Ile Val Tyr Asn Arg Phe Arg Arg Gln Leu Thr Ile Ser Val
          225          230          235          240

Leu Asp Ile Val Ala Phe Phe Pro Asn Tyr Asp Ile Arg Thr Tyr Pro
          245          250          255

Ile Gln Thr Ala Thr Gln Leu Thr Arg Glu Val Tyr Leu Asp Leu Pro
          260          265          270

Phe Ile Asn Glu Asn Leu Ser Pro Ala Ala Lys Tyr Pro Thr Phe Ser
          275          280          285

Ala Ala Glu Ser Ala Ile Ile Arg Ser Pro His Leu Val Asp Phe Leu
          290          295          300

```

```

Asn Ser Phe Thr Ile Tyr Thr Asp Ser Leu Ala Arg Ser Ala Tyr Trp
305                      310                      315                      320

Gly Gly His Leu Val Asn Ser Phe Arg Thr Gly Thr Thr Thr Asn Leu
                      325                      330                      335

Ile Arg Ser Pro Leu Tyr Gly Arg Glu Gly Asn Thr Glu Arg Pro Val
                      340                      345                      350

Thr Ile Thr Ala Ser Pro Ser Val Pro Ile Phe Arg Thr Leu Ser Tyr
                      355                      360                      365

Pro Thr Gly Leu Asp Asn Ser Asn Pro Val Ala Gly Ile Glu Gly Val
                      370                      375                      380

Glu Phe Gln Asn Thr Ile Ser Arg Ser Ile Tyr Arg Lys Ser Gly Pro
385                      390                      395                      400

Ile Asp Ser Phe Ser Glu Leu Pro Pro Gln Asp Ala Ser Val Ser Pro
                      405                      410                      415

Ala Ile Gly Tyr Ser His Arg Leu Cys His Ala Thr Phe Leu Glu Arg
                      420                      425                      430

Ile Ser Gly Pro Arg Ile Ala Gly Thr Val Phe Ser Trp Thr His Arg
                      435                      440                      445

Ser Ala Ser Pro Thr Asn Glu Val Ser Pro Ser Arg Ile Thr Gln Ile
                      450                      455                      460

Pro Trp Val Lys Ala His Thr Leu Ala Ser Gly Ala Ser Val Ile Lys
465                      470                      475                      480

Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu Thr Arg Asn Ser Met Gly
                      485                      490                      495

Glu Leu Gly Thr Leu Arg Val Thr Phe Thr Gly Arg Leu Pro Gln Ser
                      500                      505                      510

Tyr Tyr Ile Arg Phe Arg Tyr Ala Ser Val Ala Asn Arg Ser Gly Thr
                      515                      520                      525

Phe Arg Tyr Ser Gln Pro Pro Ser Tyr Gly Ile Ser Phe Pro Lys Thr
                      530                      535                      540

Met Asp Ala Gly Glu Pro Leu Thr Ser Arg Ser Phe Ala His Thr Thr
545                      550                      555                      560

Leu Phe Thr Pro Ile Thr Phe Ser Arg Ala Gln Glu Glu Phe Asp Leu
                      565                      570                      575

Tyr Ile Gln Ser Gly Val Tyr Ile Asp Arg Ile Glu Phe Ile Pro Val
                      580                      585                      590

Thr Ala Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala Gln Lys Val
                      595                      600                      605

Val Asn Ala Leu Phe Thr Ser Thr Asn Gln Leu Gly Leu Lys Thr Asp
                      610                      615                      620

Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Ala Cys Leu
625                      630                      635                      640

Ser Asp Glu Phe Cys Leu Asp Glu Lys Arg Glu Leu Ser Glu Lys Val

```

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|     |     |     |     | 645 |     |     |     | 650 |     |     |     | 655 |     |     |     |
| Lys | His | Ala | Lys | Arg | Leu | Ser | Asp | Glu | Arg | Asn | Leu | Leu | Gln | Asp | Pro |
|     |     |     | 660 |     |     |     | 665 |     |     |     |     |     | 670 |     |     |
| Asn | Phe | Arg | Gly | Ile | Asn | Arg | Gln | Pro | Asp | Arg | Gly | Trp | Arg | Gly | Ser |
|     |     |     | 675 |     |     |     | 680 |     |     |     |     |     | 685 |     |     |
| Thr | Asp | Ile | Thr | Ile | Gln | Gly | Gly | Asp | Asp | Val | Phe | Lys | Glu | Asn | Tyr |
|     |     |     | 690 |     |     |     | 695 |     |     |     |     |     | 700 |     |     |
| Val | Thr | Leu | Pro | Gly | Thr | Phe | Asp | Glu | Cys | Tyr | Pro | Thr | Tyr | Leu | Tyr |
|     |     |     | 705 |     |     |     | 710 |     |     |     |     |     | 715 |     |     |
| Gln | Lys | Ile | Asp | Glu | Ser | Lys | Leu | Lys | Ala | Tyr | Thr | Arg | Tyr | Gln | Leu |
|     |     |     | 725 |     |     |     | 730 |     |     |     |     |     | 735 |     |     |
| Arg | Gly | Tyr | Ile | Glu | Asp | Ser | Gln | Asp | Leu | Glu | Ile | Tyr | Leu | Ile | Arg |
|     |     |     | 740 |     |     |     | 745 |     |     |     |     |     | 750 |     |     |
| Tyr | Asn | Ala | Lys | His | Glu | Ile | Val | Asn | Val | Pro | Gly | Thr | Gly | Ser | Leu |
|     |     |     | 755 |     |     |     | 760 |     |     |     |     |     | 765 |     |     |
| Trp | Pro | Leu | Ser | Val | Glu | Asn | Gln | Ile | Gly | Pro | Cys | Gly | Glu | Pro | Asn |
|     |     |     | 770 |     |     |     | 775 |     |     |     |     |     | 780 |     |     |
| Arg | Cys | Ala | Pro | His | Leu | Glu | Trp | Asn | Pro | Asp | Leu | His | Cys | Ser | Cys |
|     |     |     | 785 |     |     |     | 790 |     |     |     |     |     | 795 |     |     |
| Arg | Asp | Gly | Glu | Lys | Cys | Ala | His | His | Ser | His | His | Phe | Ser | Leu | Asp |
|     |     |     | 805 |     |     |     | 810 |     |     |     |     |     | 815 |     |     |
| Ile | Asp | Val | Gly | Cys | Thr | Asp | Leu | Asn | Glu | Asp | Leu | Gly | Val | Trp | Val |
|     |     |     | 820 |     |     |     | 825 |     |     |     |     |     | 830 |     |     |
| Ile | Phe | Lys | Ile | Lys | Thr | Gln | Asp | Gly | His | Ala | Arg | Leu | Gly | Asn | Leu |
|     |     |     | 835 |     |     |     | 840 |     |     |     |     |     | 845 |     |     |
| Glu | Phe | Leu | Glu | Glu | Lys | Pro | Leu | Leu | Gly | Glu | Ala | Leu | Ala | Arg | Val |
|     |     |     | 850 |     |     |     | 855 |     |     |     |     |     | 860 |     |     |
| Lys | Arg | Ala | Glu | Lys | Lys | Trp | Arg | Asp | Lys | Arg | Glu | Thr | Leu | Gln | Leu |
|     |     |     | 865 |     |     |     | 870 |     |     |     |     |     | 875 |     |     |
| Glu | Thr | Thr | Ile | Val | Tyr | Lys | Glu | Ala | Lys | Glu | Ser | Val | Asp | Ala | Leu |
|     |     |     | 885 |     |     |     | 890 |     |     |     |     |     | 895 |     |     |
| Phe | Val | Asn | Ser | Gln | Tyr | Asp | Arg | Leu | Gln | Ala | Asp | Thr | Asn | Ile | Ala |
|     |     |     | 900 |     |     |     | 905 |     |     |     |     |     | 910 |     |     |
| Met | Ile | His | Ala | Ala | Asp | Lys | Arg | Val | His | Arg | Ile | Arg | Glu | Ala | Tyr |
|     |     |     | 915 |     |     |     | 920 |     |     |     |     |     | 925 |     |     |
| Leu | Pro | Glu | Leu | Ser | Val | Ile | Pro | Gly | Val | Asn | Ala | Ala | Ile | Phe | Glu |
|     |     |     | 930 |     |     |     | 935 |     |     |     |     |     | 940 |     |     |
| Glu | Leu | Glu | Glu | Arg | Ile | Phe | Thr | Ala | Phe | Ser | Leu | Tyr | Asp | Ala | Arg |
|     |     |     | 945 |     |     |     | 950 |     |     |     |     |     | 955 |     |     |
| Asn | Ile | Ile | Lys | Asn | Gly | Asp | Phe | Asn | Asn | Gly | Leu | Leu | Cys | Trp | Asn |
|     |     |     | 965 |     |     |     | 970 |     |     |     |     |     | 975 |     |     |
| Val | Lys | Gly | His | Val | Glu | Val | Glu | Glu | Gln | Asn | Asn | His | Arg | Ser | Val |
|     |     |     | 980 |     |     |     | 985 |     |     |     |     |     | 990 |     |     |

Leu Val Ile Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val  
 995 1000 1005  
 Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu  
 1010 1015 1020  
 Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn  
 1025 1030 1035  
 Thr Asp Glu Leu Lys Phe Asn Asn Cys Val Glu Glu Glu Val Tyr  
 1040 1045 1050  
 Pro Asn Asn Thr Val Thr Cys Ile Asn Tyr Thr Ala Thr Gln Glu  
 1055 1060 1065  
 Glu Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly Tyr Asp Glu  
 1070 1075 1080  
 Ala Tyr Gly Asn Asn Pro Ser Val Pro Ala Asp Tyr Ala Ser Val  
 1085 1090 1095  
 Tyr Glu Glu Lys Ser Tyr Thr Asp Arg Arg Arg Glu Asn Pro Cys  
 1100 1105 1110  
 Glu Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu Pro Ala Gly  
 1115 1120 1125  
 Tyr Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val  
 1130 1135 1140  
 Trp Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser  
 1145 1150 1155  
 Val Glu Leu Leu Leu Met Glu Glu  
 1160 1165

<210> 39  
 <211> 3501  
 <212> ADN  
 <213> Trình tự nhân tạo

<220>  
 <223> Trình tự ADN tổng hợp được thiết kế để biểu hiện ở thực vật ghi mã  
 Cry1Da1\_7 có gốc Alanin bổ sung được cài xen vào tại vị trí 2.

<400> 39  
 atggctgaga tcaacaacca gaaccagtgc gtcccgtaca actgcctgag caaccctaag 60  
 gagatcatcc tgggtgagga acgcctggag accggcaaca ccgtagccga cattagcctg 120  
 ggcctcatca acttcctcta cagcaacttc gtgcccggcg gtggcttcat cgtgggcctc 180  
 ctggagctta tctggggctt catcgccccg tccagtgagg acatcttcct cgtccagatc 240  
 gagcaactga tcagccagcg gatcgaggag ttogctagga accaggccat ctcccgcctg 300  
 gagggactct ccaacctcta caaggtgtac gtgcgcgcgt tcagcgactg ggagaaggac 360  
 ccgaccaacc cggccctccg cgaggaaatg cgtatccagt tcaacgatat gaactcggcc 420  
 ctcattaccg ccatcccgtc cttccgcgtg cagaactacg aggtggccct cctgtccgtg 480  
 tacgttcaag ccgccaacct ccacctctcc atoctccgcg acgtgagcgt gttcggcgag 540

|                                                                    |      |
|--------------------------------------------------------------------|------|
| cgctggggct acgacaccgc caccatcaac aaccgctact ccgacctcac ctccctcatc  | 600  |
| cacgtttaca ccaaccactg cgtggacacg tacaaccagg gcctccgccg cctggagggc  | 660  |
| cgcttcctct ccgactggat cgtgtacaac cgcttcgcc gccagctcac catctccgtc   | 720  |
| ctggacatcg tcgccttctt tcccaactac gacatccgca cctaccctat ccagaccgcc  | 780  |
| accagctca cccgcgaggt ctacctcgac ctcccgttca tcaacgagaa cctcagccc    | 840  |
| gccgccgtct acccgacctt ctccgccgct gagtccgcc tcatcgcag cccgcacctc    | 900  |
| gtggacttcc tcaactcctt caccatctac accgactccc tcgccgcag cgctactgg    | 960  |
| ggcggtcacc tcgtgaactc cttccgcacc ggcaccacta ccaacctcat ccgcagccc   | 1020 |
| ctctacggcc gcgagggcaa caccgagcgc ccggtgacca tcaccgccag cccgagcgtg  | 1080 |
| cccatcttcc gcacctcag ctaccccacc ggccctggaca acagcaacct tgtggggggc  | 1140 |
| atcgagggcg tggagttcca gaacaccatc tccaggagca tctaccgcaa gagcgccct   | 1200 |
| atcgacagct tcagcgagct gcctcctcag gacgccagcg tgagccctgc catcggtac   | 1260 |
| agccacaggc tgtgccacgc caccttcctg gagcgcatca gcggccctcg catcgcgggc  | 1320 |
| accgtgttct cgtggacca cgcgagcgc tctcctacga acgaggtgtc tcctagtgc     | 1380 |
| atcaccaga tcccttgggt caaggccac accctggcta gtggcgctag tgtcatcaag    | 1440 |
| ggccctggct tcaccggtgg tgacatcctg accaggaact ctatgggcga gctgggcact  | 1500 |
| ctgaggttca ctttactgg ccgcctgcct cagtottact acatccgctt ccgctaogct   | 1560 |
| agtgtcgcta accgctctgg tactttccgc tactctcagc ctccgtctta cgttatctct  | 1620 |
| ttccctaaga ctatggacgc tgggtgagcct ctgaccagta ggagcttcgc tcacactact | 1680 |
| ctgttcactc ctatcacttt ctctagggct caggaggagt tcgacctgta catccagtct  | 1740 |
| ggtgtgtaca tcgacaggat cgagttcatc cccgtgaccg ccacgttcga ggccgagtac  | 1800 |
| gaccttgagc gcgccagaa ggtggtgaac gccctcttca ctagcactaa ccagctaggc   | 1860 |
| ctgaagactg acgtgaccga ctaccacatc gaccaagtga gcaacctagt ggccctgctc  | 1920 |
| tccgacgagt tctgcctcga cgagaagcgc gagctgtccg agaaggtgaa gcacgccaa   | 1980 |
| cgcctctccg acgagcgcaa cctgctccag gaccccaact tcaggggcat caacaggcag  | 2040 |
| cccgaccgcg gctggcgcg ctccaccgac atcaccatcc agggcggtga cgacgtattc   | 2100 |
| aaggagaact acgttaccct ccccggcacc ttcgacgagt gttacccac ctacctctac   | 2160 |
| cagaagatcg acgagtccaa gctgaaggcc tacaccgct accagctccg cggctacatc   | 2220 |
| gaggactccc aggacctgga aatctacctc atccgctaca acgccaagca cgagatcgtg  | 2280 |
| aacgtgcctg gcaccggcag cctctggcct ctcagcgtgg agaaccagat cggcccttgc  | 2340 |
| ggcgagccta accgctgcgc cctcacctc gagtgggaacc ctgacctcca ctgctcgtgc  | 2400 |
| agggacggcg agaagtgcgc ccacatagc caccattct ctctggacat cgacgtgggc    | 2460 |

```

tgcaccgacc tgaacgagga cctgggctgt tgggttatct tcaagatcaa gaccaggac      2520
ggtcacgcca ggctgggtaa cctggagttc cttgaggaaa agcctctgct gggtagggcc      2580
ctggccaggg tcaagagggc tgagaagaaa tggagggata agagggagac cctgcagctg      2640
gagaccacta tcgtctacaa ggaggctaag gagtctgtcg atgctctgtt cgtcaactct      2700
cagtacgata gactgcaagc tgataccaac atcgctatga tccacgctgc ggataagcgg      2760
gtccaccgga tccgggaggc ttaccttcgc gagctttctg tcatcccggg tgtcaacgct      2820
gcatcttcgc aggaacttga ggaacggatc ttactgctg ttagtcttta cgatgcgcgg      2880
aacatcatca agaacgggga cttcaacaat ggtctgctgt gctggaacgt caagggctcat      2940
gtcgaggctg aggaacaaaa caatcatcgt agtgtccttg tcattcctga gtgggaggcg      3000
gaggtctctc aagagggtccg tgtttgcccg gggcggtggg acattcttcg tgttactgcg      3060
tacaaggagg ggtacgggga ggggtgctt actattcatg agattgagaa caatactgat      3120
gagcttaagt tcaacaattg tgttgaggag gaggtttacc cgaacaatac tgttactgct      3180
atcaactaca cggcaacgca agaggaatac gaggggacgt acacctcgcg taatagaggg      3240
tatgatgagg cgtacggaaa caaccgctcg gttccagcag attatgcctc ggtttatgag      3300
gagaagtcgt acacggatag acgacgcgag aatccatgtg agtcaaactg aggatacggg      3360
gattacacac cattaccagc aggatacgtt acaaaggagt tggaatactt cccggaaaca      3420
gataaagttt ggattgaaat cggagaaaaca gaaggaacat tcatcgctga ctcagtagaa      3480
ttgttgttga tggaagaatg a                                              3501

```

<210> 40

<211> 1166

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Trình tự axit amin của protein đã được xử lý có hoạt tính trừ côn trùng Cry1Da1\_7 được ghi mã bởi trình tự ADN tổng hợp trong đó gốc Alanin bổ sung được cài xen vào tại vị trí 2.

<400> 40

```

Met Ala Glu Ile Asn Asn Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu
1           5           10           15

```

```

Ser Asn Pro Lys Glu Ile Ile Leu Gly Glu Glu Arg Leu Glu Thr Gly
          20           25           30

```

```

Asn Thr Val Ala Asp Ile Ser Leu Gly Leu Ile Asn Phe Leu Tyr Ser
          35           40           45

```

```

Asn Phe Val Pro Gly Gly Gly Phe Ile Val Gly Leu Leu Glu Leu Ile
          50           55           60

```

```

Trp Gly Phe Ile Gly Pro Ser Gln Trp Asp Ile Phe Leu Ala Gln Ile
65           70           75           80

```

Glu Gln Leu Ile Ser Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala  
 85 90 95  
 Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Lys Val Tyr Val Arg  
 100 105 110  
 Ala Phe Ser Asp Trp Glu Lys Asp Pro Thr Asn Pro Ala Leu Arg Glu  
 115 120 125  
 Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Ile Thr Ala  
 130 135 140  
 Ile Pro Leu Phe Arg Val Gln Asn Tyr Glu Val Ala Leu Leu Ser Val  
 145 150 155 160  
 Tyr Val Gln Ala Ala Asn Leu His Leu Ser Ile Leu Arg Asp Val Ser  
 165 170 175  
 Val Phe Gly Glu Arg Trp Gly Tyr Asp Thr Ala Thr Ile Asn Asn Arg  
 180 185 190  
 Tyr Ser Asp Leu Thr Ser Leu Ile His Val Tyr Thr Asn His Cys Val  
 195 200 205  
 Asp Thr Tyr Asn Gln Gly Leu Arg Arg Leu Glu Gly Arg Phe Leu Ser  
 210 215 220  
 Asp Trp Ile Val Tyr Asn Arg Phe Arg Arg Gln Leu Thr Ile Ser Val  
 225 230 235 240  
 Leu Asp Ile Val Ala Phe Phe Pro Asn Tyr Asp Ile Arg Thr Tyr Pro  
 245 250 255  
 Ile Gln Thr Ala Thr Gln Leu Thr Arg Glu Val Tyr Leu Asp Leu Pro  
 260 265 270  
 Phe Ile Asn Glu Asn Leu Ser Pro Ala Ala Val Tyr Pro Thr Phe Ser  
 275 280 285  
 Ala Ala Glu Ser Ala Ile Ile Arg Ser Pro His Leu Val Asp Phe Leu  
 290 295 300  
 Asn Ser Phe Thr Ile Tyr Thr Asp Ser Leu Ala Arg Ser Ala Tyr Trp  
 305 310 315 320  
 Gly Gly His Leu Val Asn Ser Phe Arg Thr Gly Thr Thr Thr Asn Leu  
 325 330 335  
 Ile Arg Ser Pro Leu Tyr Gly Arg Glu Gly Asn Thr Glu Arg Pro Val  
 340 345 350  
 Thr Ile Thr Ala Ser Pro Ser Val Pro Ile Phe Arg Thr Leu Ser Tyr  
 355 360 365  
 Pro Thr Gly Leu Asp Asn Ser Asn Pro Val Ala Gly Ile Glu Gly Val  
 370 375 380  
 Glu Phe Gln Asn Thr Ile Ser Arg Ser Ile Tyr Arg Lys Ser Gly Pro  
 385 390 395 400  
 Ile Asp Ser Phe Ser Glu Leu Pro Pro Gln Asp Ala Ser Val Ser Pro  
 405 410 415  
 Ala Ile Gly Tyr Ser His Arg Leu Cys His Ala Thr Phe Leu Glu Arg

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 420 |     |     |     |     | 425 |     |     |     |     | 430 |     |     |     |     |     |
| Ile | Ser | Gly | Pro | Arg | Ile | Ala | Gly | Thr | Val | Phe | Ser | Trp | Thr | His | Arg |
| 435 |     |     |     |     |     | 440 |     |     | 445 |     |     |     |     |     |     |
| Ser | Ala | Ser | Pro | Thr | Asn | Glu | Val | Ser | Pro | Ser | Arg | Ile | Thr | Gln | Ile |
| 450 |     |     |     |     |     | 455 |     |     | 460 |     |     |     |     |     |     |
| Pro | Trp | Val | Lys | Ala | His | Thr | Leu | Ala | Ser | Gly | Ala | Ser | Val | Ile | Lys |
| 465 |     |     | 470 |     |     |     |     |     | 475 |     |     | 480 |     |     |     |
| Gly | Pro | Gly | Phe | Thr | Gly | Gly | Asp | Ile | Leu | Thr | Arg | Asn | Ser | Met | Gly |
|     |     |     | 485 |     |     |     |     |     | 490 |     |     | 495 |     |     |     |
| Glu | Leu | Gly | Thr | Leu | Arg | Val | Thr | Phe | Thr | Gly | Arg | Leu | Pro | Gln | Ser |
|     |     |     | 500 |     |     | 505 |     |     |     |     |     | 510 |     |     |     |
| Tyr | Tyr | Ile | Arg | Phe | Arg | Tyr | Ala | Ser | Val | Ala | Asn | Arg | Ser | Gly | Thr |
| 515 |     |     |     |     |     | 520 |     |     | 525 |     |     |     |     |     |     |
| Phe | Arg | Tyr | Ser | Gln | Pro | Pro | Ser | Tyr | Gly | Ile | Ser | Phe | Pro | Lys | Thr |
| 530 |     |     | 535 |     |     | 540 |     |     |     |     |     |     |     |     |     |
| Met | Asp | Ala | Gly | Glu | Pro | Leu | Thr | Ser | Arg | Ser | Phe | Ala | His | Thr | Thr |
| 545 |     |     | 550 |     |     | 555 |     |     | 560 |     |     |     |     |     |     |
| Leu | Phe | Thr | Pro | Ile | Thr | Phe | Ser | Arg | Ala | Gln | Glu | Glu | Phe | Asp | Leu |
|     |     |     | 565 |     |     | 570 |     |     | 575 |     |     |     |     |     |     |
| Tyr | Ile | Gln | Ser | Gly | Val | Tyr | Ile | Asp | Arg | Ile | Glu | Phe | Ile | Pro | Val |
| 580 |     |     | 585 |     |     | 590 |     |     |     |     |     |     |     |     |     |
| Thr | Ala | Thr | Phe | Glu | Ala | Glu | Tyr | Asp | Leu | Glu | Arg | Ala | Gln | Lys | Val |
| 595 |     |     | 600 |     |     | 605 |     |     |     |     |     |     |     |     |     |
| Val | Asn | Ala | Leu | Phe | Thr | Ser | Thr | Asn | Gln | Leu | Gly | Leu | Lys | Thr | Asp |
| 610 |     |     | 615 |     |     | 620 |     |     |     |     |     |     |     |     |     |
| Val | Thr | Asp | Tyr | His | Ile | Asp | Gln | Val | Ser | Asn | Leu | Val | Ala | Cys | Leu |
| 625 |     |     | 630 |     |     | 635 |     |     | 640 |     |     |     |     |     |     |
| Ser | Asp | Glu | Phe | Cys | Leu | Asp | Glu | Lys | Arg | Glu | Leu | Ser | Glu | Lys | Val |
|     |     |     | 645 |     |     | 650 |     |     | 655 |     |     |     |     |     |     |
| Lys | His | Ala | Lys | Arg | Leu | Ser | Asp | Glu | Arg | Asn | Leu | Leu | Gln | Asp | Pro |
| 660 |     |     | 665 |     |     | 670 |     |     |     |     |     |     |     |     |     |
| Asn | Phe | Arg | Gly | Ile | Asn | Arg | Gln | Pro | Asp | Arg | Gly | Trp | Arg | Gly | Ser |
| 675 |     |     | 680 |     |     | 685 |     |     |     |     |     |     |     |     |     |
| Thr | Asp | Ile | Thr | Ile | Gln | Gly | Gly | Asp | Asp | Val | Phe | Lys | Glu | Asn | Tyr |
| 690 |     |     | 695 |     |     | 700 |     |     |     |     |     |     |     |     |     |
| Val | Thr | Leu | Pro | Gly | Thr | Phe | Asp | Glu | Cys | Tyr | Pro | Thr | Tyr | Leu | Tyr |
| 705 |     |     | 710 |     |     | 715 |     |     | 720 |     |     |     |     |     |     |
| Gln | Lys | Ile | Asp | Glu | Ser | Lys | Leu | Lys | Ala | Tyr | Thr | Arg | Tyr | Gln | Leu |
|     |     |     | 725 |     |     | 730 |     |     | 735 |     |     |     |     |     |     |
| Arg | Gly | Tyr | Ile | Glu | Asp | Ser | Gln | Asp | Leu | Glu | Ile | Tyr | Leu | Ile | Arg |
| 740 |     |     | 745 |     |     | 750 |     |     |     |     |     |     |     |     |     |
| Tyr | Asn | Ala | Lys | His | Glu | Ile | Val | Asn | Val | Pro | Gly | Thr | Gly | Ser | Leu |
| 755 |     |     | 760 |     |     | 765 |     |     |     |     |     |     |     |     |     |

Trp Pro Leu Ser Val Glu Asn Gln Ile Gly Pro Cys Gly Glu Pro Asn  
 770 775 780  
 Arg Cys Ala Pro His Leu Glu Trp Asn Pro Asp Leu His Cys Ser Cys  
 785 790 795 800  
 Arg Asp Gly Glu Lys Cys Ala His His Ser His His Phe Ser Leu Asp  
 805 810 815  
 Ile Asp Val Gly Cys Thr Asp Leu Asn Glu Asp Leu Gly Val Trp Val  
 820 825 830  
 Ile Phe Lys Ile Lys Thr Gln Asp Gly His Ala Arg Leu Gly Asn Leu  
 835 840 845  
 Glu Phe Leu Glu Glu Lys Pro Leu Leu Gly Glu Ala Leu Ala Arg Val  
 850 855 860  
 Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys Arg Glu Thr Leu Gln Leu  
 865 870 875 880  
 Glu Thr Thr Ile Val Tyr Lys Glu Ala Lys Glu Ser Val Asp Ala Leu  
 885 890 895  
 Phe Val Asn Ser Gln Tyr Asp Arg Leu Gln Ala Asp Thr Asn Ile Ala  
 900 905 910  
 Met Ile His Ala Ala Asp Lys Arg Val His Arg Ile Arg Glu Ala Tyr  
 915 920 925  
 Leu Pro Glu Leu Ser Val Ile Pro Gly Val Asn Ala Ala Ile Phe Glu  
 930 935 940  
 Glu Leu Glu Glu Arg Ile Phe Thr Ala Phe Ser Leu Tyr Asp Ala Arg  
 945 950 955 960  
 Asn Ile Ile Lys Asn Gly Asp Phe Asn Asn Gly Leu Leu Cys Trp Asn  
 965 970 975  
 Val Lys Gly His Val Glu Val Glu Glu Gln Asn Asn His Arg Ser Val  
 980 985 990  
 Leu Val Ile Pro Glu Trp Glu Ala Glu Val Ser Gln Glu Val Arg Val  
 995 1000 1005  
 Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val Thr Ala Tyr Lys Glu  
 1010 1015 1020  
 Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu Ile Glu Asn Asn  
 1025 1030 1035  
 Thr Asp Glu Leu Lys Phe Asn Asn Cys Val Glu Glu Glu Val Tyr  
 1040 1045 1050  
 Pro Asn Asn Thr Val Thr Cys Ile Asn Tyr Thr Ala Thr Gln Glu  
 1055 1060 1065  
 Glu Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg Gly Tyr Asp Glu  
 1070 1075 1080  
 Ala Tyr Gly Asn Asn Pro Ser Val Pro Ala Asp Tyr Ala Ser Val  
 1085 1090 1095

Tyr Glu Glu Lys Ser Tyr Thr Asp Arg Arg Arg Glu Asn Pro Cys  
 1100 1105 1110  
 Glu Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro Leu Pro Ala Gly  
 1115 1120 1125  
 Tyr Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu Thr Asp Lys Val  
 1130 1135 1140  
 Trp Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe Ile Val Asp Ser  
 1145 1150 1155  
 Val Glu Leu Leu Leu Met Glu Glu  
 1160 1165

<210> 41  
 <211> 3423  
 <212> ADN  
 <213> Trình tự nhân tạo

<220>  
 <223> Trình tự ADN tổng hợp được thiết kế để biểu hiện ở thực vật ghi mã  
 TIC844 có gốc Alanin bổ sung được cài xen vào tại vị trí 2.

<400> 41  
 atggctgaga tcaacaacca gaaccagtgc gtcccgtaga actgcctgag caaccctaag 60  
 gagatcatcc tgggtgagga acgcctggag accggcaaca ccgtagccga cattagcctg 120  
 ggcctcatca acttcctcta cagcaacttc gtgccggcg gtggcttcat cgtgggcctc 180  
 ctggagctta tctggggctt catcgccccg tccagtgagg acatcttctt cgcccagatc 240  
 gagcaactga tcagccagcg gatcgaggag ttgcctagga accaggccat ctcccgcctg 300  
 gagggactct ccaacctcta caaggtgtac gtgcgcgcgt tcagcgactg ggagaaggac 360  
 ccgaccaacc cggccctccg cgaggaaatg cgtatccagt tcaacgatat gaactcggcc 420  
 ctcatcaccc ccatcccgtt cttccgcgtg cagaactacg aggtggccct cctgtccgtg 480  
 tacgttcaag ccgccaacct ccacctctcc atcctccgcg acgtgagcgt gttcggcgag 540  
 cgctggggct acgacaccgc caccatcaac aaccgctact ccgacctcac ctccctcatc 600  
 cacgtttaca ccaaccactg cgtggacacg tacaaccagg gcctccgccg cctggagggc 660  
 cgcttctctt ccgactggat cgtgtacaac cgcttccgcc gccagctcac catctccgtc 720  
 ctggacatcg tcgccttctt tcccaactac gacatccgca cctaccctat ccagacggcc 780  
 acccagctca cccgcgaggt ctacctcgac ctcccgttca tcaacgagaa cctcagcccg 840  
 gccgccagct acccgacctt ctccgcgcgt gagtccgcca tcattcgag cccgcacctc 900  
 gtggacttcc tcaactcctt caccatctac accgactccc tcgcccgtta cgccacttgg 960  
 ggcggtcacc tcgtgaactc cttccgcacc ggcaccacta ccaacctcat ccgcagcccg 1020  
 ctctacggcc gcgagggcaa caccgagcgc ccggtgacca tcaccgccag cccgagcgtg 1080  
 cccatcttcc gcacctcag ctacatcacc ggccctggaca acagcaacct tgtggcgggc 1140  
 atcgagggcg tggagttcca gaacaccatc tccaggagca tctaccgcaa gagcggccct 1200

|                                                                    |      |
|--------------------------------------------------------------------|------|
| atcgacagct tcagcgagct gcctcctcag gacgccagcg tgagccctgc catcggtac   | 1260 |
| agccacaggc tgtgccacgc caccttcctg gagcgcatca gcggccctcg catcgcgggc  | 1320 |
| accgtgttct cgtggacca ccgcagcgcc tctcctacga acgaggtgtc tcctagtgcg   | 1380 |
| atcaccaga tcccttgggt caaggccac accctggcta gtggcgctag tgtcatcaag    | 1440 |
| ggccctggct tcaccggtgg tgacatcctg accaggaact ctatgggcga gctgggcact  | 1500 |
| ctgagggtea ctttactgg ccgcctgcct cagtcttact acatccgctt ccgctacgct   | 1560 |
| agtgtcgta accgctctgg tactttccgc tactctcagc ctccgtctta cggtatctct   | 1620 |
| ttccctaaga ctatggacgc tggtgagcct ctgaccagta ggagcttcgc tcacactact  | 1680 |
| ctgttcactc ctatcacttt ctctagggct caggaggagt tcgacctgta catccagtct  | 1740 |
| ggtgtgtaca tcgacaggat cgagttcatc cccgtgaccg ccacgttcga ggccgagtac  | 1800 |
| gaccttgagc gcgcccagaa ggctgtcaat gagctcttca cgtccagcaa tcagatcggc  | 1860 |
| ctgaagaccg acgtcactga ctaccacatc gaccaagtct ccaacctcgt ggagtgcctc  | 1920 |
| tccgatgagt tctgcctcga cgagaagaag gagctgtccg agaagggtgaa gcatgccaa  | 1980 |
| cgtctcagcg acgagaggaa tctcctccag gaccccaatt tccgcggcat caacaggcag  | 2040 |
| ctcgaccgcg gctggcgcgg cagcaccgac atcacgatcc agggcggcga cgatgtgttc  | 2100 |
| aaggagaact acgtgactct cctgggcact ttcgacgagt gctaccctac ctacttgtac  | 2160 |
| cagaagatcg atgagtccaa gctcaaggct tacactcgct accagctccg cggtacatc   | 2220 |
| gaagacagcc aagacctcga gatttacctg atccgctaca acgccaagca cgagaccgtc  | 2280 |
| aacgtgcccg gtactggttc cctctggccg ctgagcgccc ccagcccgat cggcaagtgt  | 2340 |
| gccaccaca gccaccactt ctcttggac atcgatgtgg gctgcaccga cctgaacgag    | 2400 |
| gacctcgag tctgggtcat cttcaagatc aagaccagc acggccacgc gcgcctgggc    | 2460 |
| aacctggagt tcctcgagga gaagcccctg gtcggtgagg ctctggccag ggtcaagagg  | 2520 |
| gctgagaaga agtggaggga caagcgcgag aagctcgagt gggagaccaa catcgtttac  | 2580 |
| aaggaggcca aggagagcgt cgacgccctg ttctgtaact ccagtagca ccgcctgcag   | 2640 |
| gccgacacca acatcgccat gatccacgct gccgacaaga ggggtgcacag cattogcgag | 2700 |
| gcctacctgc ctgagctgtc cgtgatccct ggtgtgaacg ctgccatctt tgaggagctg  | 2760 |
| gagggccgca tctttaccgc attctccctg tacgacgccc gcaacgtgat caagaacggt  | 2820 |
| gacttcaaca atggcctcag ctgctggaac gtcaagggcc acgtggacgt cgaggaacag  | 2880 |
| aacaaccacc gctccgtcct ggtcgtccca gagtgggagg ctgaggtctc ccaagaggtc  | 2940 |
| cgcgctctgc caggccgcgg ctacattctc agggtcaccg cttacaagga gggctacggt  | 3000 |
| gagggctgtg tgaccatcca cgagatcgag aacaacaccg acgagcttaa gttctccaac  | 3060 |
| tgcggtggagg aggaggtgta cccaaacaac accgttactt gcaacgacta caccgccacc | 3120 |

```

caggaggagt acgagggcac ctacacttcc aggaacaggg gctacgatgg tgcctacgag      3180
agcaacagca gcgttcctgc tgactacgct tccgcctacg aggagaaggc ctacacggat      3240
ggccgcaggg acaacccttg cgagagcaac cgcggtctacg gcgactacac tccccgtccc      3300
gccggctacg ttaccaagga gctggagtac ttcccggaga ctgacaaggt gtggatcgag      3360
atcggcgaga ccgagggcac cttcatcgtg gacagcgtgg agctgctcct gatggaggag      3420
tag                                                                    3423

```

<210> 42

<211> 1140

<212> PRT

<213> Trình tự nhân tạo

<220>

<223> Trình tự axit amin của protein khảm TIC844 được ghi mã bởi trình tự ADN tổng hợp trong đó gốc Alanin bổ sung được cài xen vào tại vị trí 2.

<400> 42

```

Met Ala Glu Ile Asn Asn Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu
1              5              10              15

```

```

Ser Asn Pro Lys Glu Ile Ile Leu Gly Glu Glu Arg Leu Glu Thr Gly
              20              25              30

```

```

Asn Thr Val Ala Asp Ile Ser Leu Gly Leu Ile Asn Phe Leu Tyr Ser
              35              40              45

```

```

Asn Phe Val Pro Gly Gly Gly Phe Ile Val Gly Leu Leu Glu Leu Ile
              50              55              60

```

```

Trp Gly Phe Ile Gly Pro Ser Gln Trp Asp Ile Phe Leu Ala Gln Ile
65              70              75              80

```

```

Glu Gln Leu Ile Ser Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala
              85              90              95

```

```

Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Lys Val Tyr Val Arg
              100              105              110

```

```

Ala Phe Ser Asp Trp Glu Lys Asp Pro Thr Asn Pro Ala Leu Arg Glu
              115              120              125

```

```

Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Ile Thr Ala
              130              135              140

```

```

Ile Pro Leu Phe Arg Val Gln Asn Tyr Glu Val Ala Leu Leu Ser Val
145              150              155              160

```

```

Tyr Val Gln Ala Ala Asn Leu His Leu Ser Ile Leu Arg Asp Val Ser
              165              170              175

```

```

Val Phe Gly Glu Arg Trp Gly Tyr Asp Thr Ala Thr Ile Asn Asn Arg
              180              185              190

```

```

Tyr Ser Asp Leu Thr Ser Leu Ile His Val Tyr Thr Asn His Cys Val
              195              200              205

```

```

Asp Thr Tyr Asn Gln Gly Leu Arg Arg Leu Glu Gly Arg Phe Leu Ser
210                               215                               220

Asp Trp Ile Val Tyr Asn Arg Phe Arg Arg Gln Leu Thr Ile Ser Val
225                               230                               235                               240

Leu Asp Ile Val Ala Phe Phe Pro Asn Tyr Asp Ile Arg Thr Tyr Pro
245                               245                               250                               255

Ile Gln Thr Ala Thr Gln Leu Thr Arg Glu Val Tyr Leu Asp Leu Pro
260                               265                               270

Phe Ile Asn Glu Asn Leu Ser Pro Ala Ala Ser Tyr Pro Thr Phe Ser
275                               280                               285

Ala Ala Glu Ser Ala Ile Ile Arg Ser Pro His Leu Val Asp Phe Leu
290                               295                               300

Asn Ser Phe Thr Ile Tyr Thr Asp Ser Leu Ala Arg Tyr Ala Tyr Trp
305                               310                               315                               320

Gly Gly His Leu Val Asn Ser Phe Arg Thr Gly Thr Thr Thr Asn Leu
325                               330                               335

Ile Arg Ser Pro Leu Tyr Gly Arg Glu Gly Asn Thr Glu Arg Pro Val
340                               345                               350

Thr Ile Thr Ala Ser Pro Ser Val Pro Ile Phe Arg Thr Leu Ser Tyr
355                               360                               365

Ile Thr Gly Leu Asp Asn Ser Asn Pro Val Ala Gly Ile Glu Gly Val
370                               375                               380

Glu Phe Gln Asn Thr Ile Ser Arg Ser Ile Tyr Arg Lys Ser Gly Pro
385                               390                               395                               400

Ile Asp Ser Phe Ser Glu Leu Pro Pro Gln Asp Ala Ser Val Ser Pro
405                               410                               415

Ala Ile Gly Tyr Ser His Arg Leu Cys His Ala Thr Phe Leu Glu Arg
420                               425                               430

Ile Ser Gly Pro Arg Ile Ala Gly Thr Val Phe Ser Trp Thr His Arg
435                               440                               445

Ser Ala Ser Pro Thr Asn Glu Val Ser Pro Ser Arg Ile Thr Gln Ile
450                               455                               460

Pro Trp Val Lys Ala His Thr Leu Ala Ser Gly Ala Ser Val Ile Lys
465                               470                               475                               480

Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu Thr Arg Asn Ser Met Gly
485                               490                               495

Glu Leu Gly Thr Leu Arg Val Thr Phe Thr Gly Arg Leu Pro Gln Ser
500                               505                               510

Tyr Tyr Ile Arg Phe Arg Tyr Ala Ser Val Ala Asn Arg Ser Gly Thr
515                               520                               525

Phe Arg Tyr Ser Gln Pro Pro Ser Tyr Gly Ile Ser Phe Pro Lys Thr
530                               535                               540

Met Asp Ala Gly Glu Pro Leu Thr Ser Arg Ser Phe Ala His Thr Thr

```

|                 |                         |                         |  |     |     |     |
|-----------------|-------------------------|-------------------------|--|-----|-----|-----|
| 545             |                         | 550                     |  | 555 |     | 560 |
| Leu Phe Thr Pro | Ile Thr Phe Ser Arg Ala | Gln Glu Glu Phe Asp Leu |  |     |     |     |
|                 | 565                     | 570                     |  |     | 575 |     |
| Tyr Ile Gln Ser | Gly Val Tyr Ile Asp Arg | Ile Glu Phe Ile Pro Val |  |     |     |     |
|                 | 580                     | 585                     |  |     | 590 |     |
| Thr Ala Thr Phe | Glu Ala Glu Tyr Asp Leu | Glu Arg Ala Gln Lys Ala |  |     |     |     |
|                 | 595                     | 600                     |  |     | 605 |     |
| Val Asn Glu Leu | Phe Thr Ser Ser Asn Gln | Ile Gly Leu Lys Thr Asp |  |     |     |     |
|                 | 610                     | 615                     |  |     | 620 |     |
| Val Thr Asp Tyr | His Ile Asp Gln Val Ser | Asn Leu Val Glu Cys Leu |  |     |     |     |
|                 | 625                     | 630                     |  |     | 635 | 640 |
| Ser Asp Glu Phe | Cys Leu Asp Glu Lys Lys | Glu Leu Ser Glu Lys Val |  |     |     |     |
|                 | 645                     | 650                     |  |     | 655 |     |
| Lys His Ala Lys | Arg Leu Ser Asp Glu Arg | Asn Leu Leu Gln Asp Pro |  |     |     |     |
|                 | 660                     | 665                     |  |     | 670 |     |
| Asn Phe Arg Gly | Ile Asn Arg Gln Leu Asp | Arg Gly Trp Arg Gly Ser |  |     |     |     |
|                 | 675                     | 680                     |  |     | 685 |     |
| Thr Asp Ile Thr | Ile Gln Gly Gly Asp Asp | Val Phe Lys Glu Asn Tyr |  |     |     |     |
|                 | 690                     | 695                     |  |     | 700 |     |
| Val Thr Leu Leu | Gly Thr Phe Asp Glu Cys | Tyr Pro Thr Tyr Leu Tyr |  |     |     |     |
|                 | 705                     | 710                     |  |     | 715 | 720 |
| Gln Lys Ile Asp | Glu Ser Lys Leu Lys Ala | Tyr Thr Arg Tyr Gln Leu |  |     |     |     |
|                 | 725                     | 730                     |  |     | 735 |     |
| Arg Gly Tyr Ile | Glu Asp Ser Gln Asp Leu | Glu Ile Tyr Leu Ile Arg |  |     |     |     |
|                 | 740                     | 745                     |  |     | 750 |     |
| Tyr Asn Ala Lys | His Glu Thr Val Asn Val | Pro Gly Thr Gly Ser Leu |  |     |     |     |
|                 | 755                     | 760                     |  |     | 765 |     |
| Trp Pro Leu Ser | Ala Pro Ser Pro Ile Gly | Lys Cys Ala His His Ser |  |     |     |     |
|                 | 770                     | 775                     |  |     | 780 |     |
| His His Phe Ser | Leu Asp Ile Asp Val Gly | Cys Thr Asp Leu Asn Glu |  |     |     |     |
|                 | 785                     | 790                     |  |     | 795 | 800 |
| Asp Leu Gly Val | Trp Val Ile Phe Lys Ile | Lys Thr Gln Asp Gly His |  |     |     |     |
|                 | 805                     | 810                     |  |     | 815 |     |
| Ala Arg Leu Gly | Asn Leu Glu Phe Leu Glu | Glu Lys Pro Leu Val Gly |  |     |     |     |
|                 | 820                     | 825                     |  |     | 830 |     |
| Glu Ala Leu Ala | Arg Val Lys Arg Ala Glu | Lys Lys Trp Arg Asp Lys |  |     |     |     |
|                 | 835                     | 840                     |  |     | 845 |     |
| Arg Glu Lys Leu | Glu Trp Glu Thr Asn Ile | Val Tyr Lys Glu Ala Lys |  |     |     |     |
|                 | 850                     | 855                     |  |     | 860 |     |
| Glu Ser Val Asp | Ala Leu Phe Val Asn Ser | Gln Tyr Asp Arg Leu Gln |  |     |     |     |
|                 | 865                     | 870                     |  |     | 875 | 880 |
| Ala Asp Thr Asn | Ile Ala Met Ile His Ala | Ala Asp Lys Arg Val His |  |     |     |     |
|                 | 885                     | 890                     |  |     | 895 |     |

Ser Ile Arg Glu Ala Tyr Leu Pro Glu Leu Ser Val Ile Pro Gly Val  
 900 905 910  
 Asn Ala Ala Ile Phe Glu Glu Leu Glu Gly Arg Ile Phe Thr Ala Phe  
 915 920 925  
 Ser Leu Tyr Asp Ala Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn  
 930 935 940  
 Gly Leu Ser Cys Trp Asn Val Lys Gly His Val Asp Val Glu Glu Gln  
 945 950 955 960  
 Asn Asn His Arg Ser Val Leu Val Val Pro Glu Trp Glu Ala Glu Val  
 965 970 975  
 Ser Gln Glu Val Arg Val Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val  
 980 985 990  
 Thr Ala Tyr Lys Glu Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu  
 995 1000 1005  
 Ile Glu Asn Asn Thr Asp Glu Leu Lys Phe Ser Asn Cys Val Glu  
 1010 1015 1020  
 Glu Glu Val Tyr Pro Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr  
 1025 1030 1035  
 Ala Thr Gln Glu Glu Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg  
 1040 1045 1050  
 Gly Tyr Asp Gly Ala Tyr Glu Ser Asn Ser Ser Val Pro Ala Asp  
 1055 1060 1065  
 Tyr Ala Ser Ala Tyr Glu Glu Lys Ala Tyr Thr Asp Gly Arg Arg  
 1070 1075 1080  
 Asp Asn Pro Cys Glu Ser Asn Arg Gly Tyr Gly Asp Tyr Thr Pro  
 1085 1090 1095  
 Leu Pro Ala Gly Tyr Val Thr Lys Glu Leu Glu Tyr Phe Pro Glu  
 1100 1105 1110  
 Thr Asp Lys Val Trp Ile Glu Ile Gly Glu Thr Glu Gly Thr Phe  
 1115 1120 1125  
 Ile Val Asp Ser Val Glu Leu Leu Leu Met Glu Glu  
 1130 1135 1140

<210> 43  
 <211> 3423  
 <212> ADN  
 <213> Nhân tạo

<220>

<223> Trình tự ADN tổng hợp được thiết kế để biểu hiện ở thực vật ghi mã  
 TIC844\_8 có gốc Alanin bổ sung được cài xen vào tại vị trí 2.

<400> 43

atggctgaga tcaacaacca gaaccagtgc gtcccgtaga actgcctgag caaccctaag 60  
 gagatcatcc tgggtgagga acgcctggag accggcaaca ccgtagccga cattagcctg 120  
 ggccctcatca acttcctcta cagcaacttc gtgcccggcg gtggcttcat cgtgggcctc 180

|                                                                    |      |
|--------------------------------------------------------------------|------|
| ctggagctta tctggggctt catcggcccg tcccagtggg acatcttctt cgcccagatc  | 240  |
| gagcaactga tcagccagcg gatcgaggag ttcgctagga accaggccat ctccgcctg   | 300  |
| gagggactct ccaacctcta caaggtgtac gtgcgcgcgt tcagcgactg ggagaaggac  | 360  |
| ccgaccaacc cggccctccg cgaggaaatg cgtatccagt tcaacgatat gaactcggcc  | 420  |
| ctcatcaccg ccatcccgtt cttccgcgtg cagaactacg aggtggccct cctgtccgtg  | 480  |
| tacgttcaag cgcgcaacct ccacctctcc atcctccgcy acgtgagcgt gttcggcgag  | 540  |
| cgctggggct acgacaccgc caccatcaac aaccgctact ccgacctcac ctccctcatc  | 600  |
| cacgtttaca ccaaccactg cgtggacacg tacaaccagg gcctccgccg cctggagggc  | 660  |
| cgcttcctct ccgactggat cgtgtacaac cgcttccgcc gccagctcac catctccgtc  | 720  |
| ctggacatcg tcgccttctt tcccaactac gacatccgca cctaccctat ccagaccgcc  | 780  |
| accagctca cccgcgaggt ctacctcgac ctcccgttca tcaacgagaa cctcagcccg   | 840  |
| gccgccgtct acccgacctt ctccgccgtt gagtccgcca tcattcgag cccgcacctc   | 900  |
| gtggacttcc tcaactcctt caccatctac accgactccc tcgccgcag cgcctactgg   | 960  |
| ggcggtcacc tcgtgaactc cttccgcacc ggcaccacta ccaacctcat ccgcagcccc  | 1020 |
| ctctacggcc gcgagggcaa caccgagcgc ccggtgacca tcaccgccag cccgagcgtg  | 1080 |
| cccatcttcc gcacctcag ctaccccacc ggcttggaaca acagcaacct tgtggcgggc  | 1140 |
| atcgagggcg tggagttcca gaacaccatc tcaggagca tctaccgcaa gagcggccct   | 1200 |
| atcgacagct tcagcgagct gcctcctcag gacgccagcg tgagccctgc catcggttac  | 1260 |
| agccacaggc tgtgccacgc caccttctct gagcgcatca gcggccctcg catcgcgggc  | 1320 |
| accgtgttct cgtggaccca ccgcagcgcc tctctacga acgaggtgtc tcctagtctc   | 1380 |
| atcaccacaga tcccttgggt caaggccac accctggcta gtggcgctag tgtcatcaag  | 1440 |
| ggccctggct tcaccggtgg tgacatctg accaggaact ctatgggcga gctgggcact   | 1500 |
| ctgagggtea ctttacttgg ccgcctgcct cagtcttact acatccgctt ccgctacgtt  | 1560 |
| agtgtcgcta accgctctgg tactttccgc tactctcagc ctccgtctta cggtatctct  | 1620 |
| ttccctaaga ctatggacgc tggtagacct ctgaccagta ggagcttcgc tcacactact  | 1680 |
| ctgttcactc ctatcacttt ctctagggct caggaggagt tcgacctgta catocagtct  | 1740 |
| ggtgtgtaca tcgacaggat cgagttcatc cccgtgaccg ccacgttcga ggccgagtac  | 1800 |
| gaccttgagc gcgcccagaa ggctgtcaat gagctcttca cgtocagcaa tcagatcggc  | 1860 |
| ctgaagaccg acgtcactga ctaccacatc gaccaagtct ccaacctcgt ggagtgcctc  | 1920 |
| tccgatgagt tctgcctcga cgagaagaag gagctgtccg agaagggtgaa gcatgccaa  | 1980 |
| cgtctcagcg acgagaggaa tctcctccag gaccccaatt tccgcggcat caacaggcag  | 2040 |
| ctcgaccgcg gctggcgcgg cagcaccgac atcacgatcc agggcgggcga cgatgtgttc | 2100 |

```

aaggagaact acgtgactct cctgggcaact ttcgacgagt gctaccctac ctacttgtac      2160
cagaagatcg atgagtccaa gctcaaggct tacactcgct accagctccg cggctacatc      2220
gaagacagcc aagacctoga gatttacctg atccgctaca acgccaagca cgagaccgtc      2280
aacgtgcccg gtactgggtc cctctggccg ctgagcgccc ccagcccgat cggcaagtgt      2340
gccaccaca gccaccactt ctcttggac atcgatgtgg gctgcaccga cctgaacgag      2400
gacctcgag tctgggtcat cttcaagatc aagaccagc acggccacgc gcgcctgggc      2460
aacctggagt tcctcgagga gaagcccctg gtcggtgagg ctctggccag ggtcaagagg      2520
gctgagaaga agtggaggga caagcgcgag aagctcgagt gggagaccaa catcgtttac      2580
aaggaggcca aggagagcgt cgacgccctg ttcgtgaact ccagtagca ccgcctgcag      2640
gccgacacca acatcgccat gatccacgt gccgacaaga gggcgcacag cattcgcgag      2700
gcctacctgc ctgagctgtc cgtgatccct ggtgtgaacg ctgccatctt tgaggagctg      2760
gagggccgca tctttaccgc attctccctg tacgacgccc gcaacgtgat caagaacggt      2820
gacttcaaca atggcctcag ctgctggaac gtcaagggcc acgtggacgt cgaggaacag      2880
aacaaccacc gctccgtcct ggtcgtccca gagtgggagg ctgaggtctc ccaagaggtc      2940
cgcgctctgc caggccgcgg ctacattctc agggtcaccg cttacaagga gggctacggt      3000
gagggctgtg tgaccatcca cgagatcgag aacaacaccg acgagcttaa gttctccaac      3060
tgcggtggagg aggaggtgta cccaaacaac accgttactt gcaacgacta caccgccacc      3120
caggaggagt acgagggcac ctacacttcc aggaacaggg gctacgatgg tgcctacgag      3180
agcaacagca gcgttcctgc tgactacgct tccgcctacg aggagaaggc ctacacggat      3240
ggccgcaggg acaacccttg cgagagcaac cgcggtctacg gcgaactacac tcccctgccc      3300
gccggctacg ttaccaagga gctggagtac ttcccggaga ctgacaaggt gtggatcgag      3360
atcggcgaga ccgagggcac cttcatcgtg gacagcgtgg agctgctcct gatggaggag      3420
tag                                                                           3423

```

<210> 44

<211> 1140

<212> PRT

<213> Nhân tạo

<220>

<223> Trình tự axit amin của protein khám đã xử lý có hoạt tính trừ côn trùng TIC844\_8 được ghi mã bởi trình tự ADN tổng hợp trong đó gốc Alanin bổ sung được cài xen vào tại vị trí 2.

<400> 44

```

Met Ala Glu Ile Asn Asn Gln Asn Gln Cys Val Pro Tyr Asn Cys Leu
1           5           10           15

```

```

Ser Asn Pro Lys Glu Ile Ile Leu Gly Glu Glu Arg Leu Glu Thr Gly
          20           25           30

```

```

Asn Thr Val Ala Asp Ile Ser Leu Gly Leu Ile Asn Phe Leu Tyr Ser
   35                               40                       45

Asn Phe Val Pro Gly Gly Gly Phe Ile Val Gly Leu Leu Glu Leu Ile
   50                               55                       60

Trp Gly Phe Ile Gly Pro Ser Gln Trp Asp Ile Phe Leu Ala Gln Ile
   65                               70                       75                       80

Glu Gln Leu Ile Ser Gln Arg Ile Glu Glu Phe Ala Arg Asn Gln Ala
                        85                               90                       95

Ile Ser Arg Leu Glu Gly Leu Ser Asn Leu Tyr Lys Val Tyr Val Arg
                        100                               105                       110

Ala Phe Ser Asp Trp Glu Lys Asp Pro Thr Asn Pro Ala Leu Arg Glu
      115                               120                       125

Glu Met Arg Ile Gln Phe Asn Asp Met Asn Ser Ala Leu Ile Thr Ala
      130                               135                       140

Ile Pro Leu Phe Arg Val Gln Asn Tyr Glu Val Ala Leu Leu Ser Val
      145                               150                       155                       160

Tyr Val Gln Ala Ala Asn Leu His Leu Ser Ile Leu Arg Asp Val Ser
                        165                               170                       175

Val Phe Gly Glu Arg Trp Gly Tyr Asp Thr Ala Thr Ile Asn Asn Arg
      180                               185                       190

Tyr Ser Asp Leu Thr Ser Leu Ile His Val Tyr Thr Asn His Cys Val
      195                               200                       205

Asp Thr Tyr Asn Gln Gly Leu Arg Arg Leu Glu Gly Arg Phe Leu Ser
      210                               215                       220

Asp Trp Ile Val Tyr Asn Arg Phe Arg Arg Gln Leu Thr Ile Ser Val
      225                               230                       235                       240

Leu Asp Ile Val Ala Phe Phe Pro Asn Tyr Asp Ile Arg Thr Tyr Pro
                        245                               250                       255

Ile Gln Thr Ala Thr Gln Leu Thr Arg Glu Val Tyr Leu Asp Leu Pro
      260                               265                       270

Phe Ile Asn Glu Asn Leu Ser Pro Ala Ala Val Tyr Pro Thr Phe Ser
      275                               280                       285

Ala Ala Glu Ser Ala Ile Ile Arg Ser Pro His Leu Val Asp Phe Leu
      290                               295                       300

Asn Ser Phe Thr Ile Tyr Thr Asp Ser Leu Ala Arg Ser Ala Tyr Trp
      305                               310                       315                       320

Gly Gly His Leu Val Asn Ser Phe Arg Thr Gly Thr Thr Thr Asn Leu
                        325                               330                       335

Ile Arg Ser Pro Leu Tyr Gly Arg Glu Gly Asn Thr Glu Arg Pro Val
      340                               345                       350

Thr Ile Thr Ala Ser Pro Ser Val Pro Ile Phe Arg Thr Leu Ser Tyr
      355                               360                       365

```

```

Pro Thr Gly Leu Asp Asn Ser Asn Pro Val Ala Gly Ile Glu Gly Val
 370                               380

Glu Phe Gln Asn Thr Ile Ser Arg Ser Ile Tyr Arg Lys Ser Gly Pro
385                               395                               400

Ile Asp Ser Phe Ser Glu Leu Pro Pro Gln Asp Ala Ser Val Ser Pro
                               405                               410                               415

Ala Ile Gly Tyr Ser His Arg Leu Cys His Ala Thr Phe Leu Glu Arg
                               420                               425                               430

Ile Ser Gly Pro Arg Ile Ala Gly Thr Val Phe Ser Trp Thr His Arg
                               435                               440                               445

Ser Ala Ser Pro Thr Asn Glu Val Ser Pro Ser Arg Ile Thr Gln Ile
                               450                               455                               460

Pro Trp Val Lys Ala His Thr Leu Ala Ser Gly Ala Ser Val Ile Lys
465                               470                               475                               480

Gly Pro Gly Phe Thr Gly Gly Asp Ile Leu Thr Arg Asn Ser Met Gly
                               485                               490                               495

Glu Leu Gly Thr Leu Arg Val Thr Phe Thr Gly Arg Leu Pro Gln Ser
                               500                               505                               510

Tyr Tyr Ile Arg Phe Arg Tyr Ala Ser Val Ala Asn Arg Ser Gly Thr
                               515                               520                               525

Phe Arg Tyr Ser Gln Pro Pro Ser Tyr Gly Ile Ser Phe Pro Lys Thr
                               530                               535                               540

Met Asp Ala Gly Glu Pro Leu Thr Ser Arg Ser Phe Ala His Thr Thr
545                               550                               555                               560

Leu Phe Thr Pro Ile Thr Phe Ser Arg Ala Gln Glu Glu Phe Asp Leu
                               565                               570                               575

Tyr Ile Gln Ser Gly Val Tyr Ile Asp Arg Ile Glu Phe Ile Pro Val
                               580                               585                               590

Thr Ala Thr Phe Glu Ala Glu Tyr Asp Leu Glu Arg Ala Gln Lys Ala
                               595                               600                               605

Val Asn Glu Leu Phe Thr Ser Ser Asn Gln Ile Gly Leu Lys Thr Asp
                               610                               615                               620

Val Thr Asp Tyr His Ile Asp Gln Val Ser Asn Leu Val Glu Cys Leu
625                               630                               635                               640

Ser Asp Glu Phe Cys Leu Asp Glu Lys Lys Glu Leu Ser Glu Lys Val
                               645                               650                               655

Lys His Ala Lys Arg Leu Ser Asp Glu Arg Asn Leu Leu Gln Asp Pro
                               660                               665                               670

Asn Phe Arg Gly Ile Asn Arg Gln Leu Asp Arg Gly Trp Arg Gly Ser
                               675                               680                               685

Thr Asp Ile Thr Ile Gln Gly Gly Asp Asp Val Phe Lys Glu Asn Tyr
690                               695                               700

Val Thr Leu Leu Gly Thr Phe Asp Glu Cys Tyr Pro Thr Tyr Leu Tyr

```

|                     |                                                 |     |  |      |  |      |
|---------------------|-------------------------------------------------|-----|--|------|--|------|
| 705                 |                                                 | 710 |  | 715  |  | 720  |
| Gln Lys Ile Asp     | Glu Ser Lys Leu Lys Ala Tyr Thr Arg Tyr Gln Leu |     |  |      |  |      |
|                     | 725                                             |     |  | 730  |  | 735  |
| Arg Gly Tyr Ile     | Glu Asp Ser Gln Asp Leu Glu Ile Tyr Leu Ile Arg |     |  |      |  |      |
|                     | 740                                             |     |  | 745  |  | 750  |
| Tyr Asn Ala Lys His | Glu Thr Val Asn Val Pro Gly Thr Gly Ser Leu     |     |  |      |  |      |
|                     | 755                                             |     |  | 760  |  | 765  |
| Trp Pro Leu Ser Ala | Pro Ser Pro Ile Gly Lys Cys Ala His His Ser     |     |  |      |  |      |
|                     | 770                                             |     |  | 775  |  | 780  |
| His His Phe Ser Leu | Asp Ile Asp Val Gly Cys Thr Asp Leu Asn Glu     |     |  |      |  |      |
|                     | 785                                             |     |  | 790  |  | 795  |
| Asp Leu Gly Val     | Trp Val Ile Phe Lys Ile Lys Thr Gln Asp Gly His |     |  |      |  |      |
|                     | 805                                             |     |  | 810  |  | 815  |
| Ala Arg Leu Gly     | Asn Leu Glu Phe Leu Glu Glu Lys Pro Leu Val Gly |     |  |      |  |      |
|                     | 820                                             |     |  | 825  |  | 830  |
| Glu Ala Leu Ala Arg | Val Lys Arg Ala Glu Lys Lys Trp Arg Asp Lys     |     |  |      |  |      |
|                     | 835                                             |     |  | 840  |  | 845  |
| Arg Glu Lys Leu Glu | Trp Glu Thr Asn Ile Val Tyr Lys Glu Ala Lys     |     |  |      |  |      |
|                     | 850                                             |     |  | 855  |  | 860  |
| Glu Ser Val Asp Ala | Leu Phe Val Asn Ser Gln Tyr Asp Arg Leu Gln     |     |  |      |  |      |
|                     | 865                                             |     |  | 870  |  | 875  |
| Ala Asp Thr Asn Ile | Ala Met Ile His Ala Ala Asp Lys Arg Val His     |     |  |      |  |      |
|                     | 885                                             |     |  | 890  |  | 895  |
| Ser Ile Arg Glu Ala | Tyr Leu Pro Glu Leu Ser Val Ile Pro Gly Val     |     |  |      |  |      |
|                     | 900                                             |     |  | 905  |  | 910  |
| Asn Ala Ala Ile Phe | Glu Glu Leu Glu Gly Arg Ile Phe Thr Ala Phe     |     |  |      |  |      |
|                     | 915                                             |     |  | 920  |  | 925  |
| Ser Leu Tyr Asp Ala | Arg Asn Val Ile Lys Asn Gly Asp Phe Asn Asn     |     |  |      |  |      |
|                     | 930                                             |     |  | 935  |  | 940  |
| Gly Leu Ser Cys Trp | Asn Val Lys Gly His Val Asp Val Glu Glu Gln     |     |  |      |  |      |
|                     | 945                                             |     |  | 950  |  | 955  |
| Asn Asn His Arg Ser | Val Leu Val Val Pro Glu Trp Glu Ala Glu Val     |     |  |      |  |      |
|                     | 965                                             |     |  | 970  |  | 975  |
| Ser Gln Glu Val Arg | Val Cys Pro Gly Arg Gly Tyr Ile Leu Arg Val     |     |  |      |  |      |
|                     | 980                                             |     |  | 985  |  | 990  |
| Thr Ala Tyr Lys Glu | Gly Tyr Gly Glu Gly Cys Val Thr Ile His Glu     |     |  |      |  |      |
|                     | 995                                             |     |  | 1000 |  | 1005 |
| Ile Glu Asn Asn Thr | Asp Glu Leu Lys Phe Ser Asn Cys Val Glu         |     |  |      |  |      |
|                     | 1010                                            |     |  | 1015 |  | 1020 |
| Glu Glu Val Tyr Pro | Asn Asn Thr Val Thr Cys Asn Asp Tyr Thr         |     |  |      |  |      |
|                     | 1025                                            |     |  | 1030 |  | 1035 |
| Ala Thr Gln Glu Glu | Tyr Glu Gly Thr Tyr Thr Ser Arg Asn Arg         |     |  |      |  |      |
|                     | 1040                                            |     |  | 1045 |  | 1050 |

|      |     |     |     |     |     |      |     |     |     |     |      |     |     |     |
|------|-----|-----|-----|-----|-----|------|-----|-----|-----|-----|------|-----|-----|-----|
| Gly  | Tyr | Asp | Gly | Ala | Tyr | Glu  | Ser | Asn | Ser | Ser | Val  | Pro | Ala | Asp |
| 1055 |     |     |     |     |     | 1060 |     |     |     |     | 1065 |     |     |     |
| Tyr  | Ala | Ser | Ala | Tyr | Glu | Glu  | Lys | Ala | Tyr | Thr | Asp  | Gly | Arg | Arg |
| 1070 |     |     |     |     |     | 1075 |     |     |     |     | 1080 |     |     |     |
| Asp  | Asn | Pro | Cys | Glu | Ser | Asn  | Arg | Gly | Tyr | Gly | Asp  | Tyr | Thr | Pro |
| 1085 |     |     |     |     |     | 1090 |     |     |     |     | 1095 |     |     |     |
| Leu  | Pro | Ala | Gly | Tyr | Val | Thr  | Lys | Glu | Leu | Glu | Tyr  | Phe | Pro | Glu |
| 1100 |     |     |     |     |     | 1105 |     |     |     |     | 1110 |     |     |     |
| Thr  | Asp | Lys | Val | Trp | Ile | Glu  | Ile | Gly | Glu | Thr | Glu  | Gly | Thr | Phe |
| 1115 |     |     |     |     |     | 1120 |     |     |     |     | 1125 |     |     |     |
| Ile  | Val | Asp | Ser | Val | Glu | Leu  | Leu | Leu | Met | Glu | Glu  |     |     |     |
| 1130 |     |     |     |     |     | 1135 |     |     |     |     | 1140 |     |     |     |